RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

RF/MICROWAVE CIRCUIT DESIGN FOR WIRELESS APPLICATIONS

RF/MICROWAVE CIRCUIT DESIGN FOR WIRELESS APPLICATIONS

Ulrich L. Rohde
Synergy Microwave Corporation

David P. Newkirk
Ansoft Corporation

A WILEY-INTERSCIENCE PUBLICATION
JOHN WILEY & SONS, INC.
New York / Chichester / Weinheim / Brisbane / Singapore / Toronto

Designations used by companies to distinguish their products are often claimed as trademarks. In all instances where John Wiley & Sons, Inc., is aware of a claim, the product names appear in initial capital or ALL CAPITAL LETTERS. Readers, however, should contact the appropriate companies for more complete information regarding trademarks and registration.

Copyright © 2000 by John Wiley & Sons, Inc. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic or mechanical, including uploading, downloading, printing, decompiling, recording or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the Publisher. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, (212) 850-6011, fax (212) 850-6008, E-Mail: PERMREQ @ WILEY.COM.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold with the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional person should be sought.

ISBN 0-471-22413-8

This title is also available in print as ISBN 0-471-29818-2.

For more information about Wiley products, visit our web site at www.Wiley.com.

To Professor Vittorio Rizzoli

who has been instrumental in the development of the powerful harmonic-balance analysis tool, specifically Microwave Harmonica, which is part of Ansoft's Serenade Design Environment. Most of the success enjoyed by Compact Software, now part of Ansoft, continues to be based on his far-reaching contributions.

CONTENTS

Foreword xiii

Preface xv

1 Introduction to Wireless Circuit Design 1

1-1 Overview / 1
1-2 System Functions / 3
1-3 The Radio Channel and Modulation Requirements / 5
1-3-1 Introduction / 5
1-3-2 Channel Impulse Response / 7
1-3-3 Doppler Effect / 13
1-3-4 Transfer Function / 14
1-3-5 Time Response of Channel Impulse Response and Transfer Function / 14
1-3-6 Lessons Learned / 17
1-3-7 Wireless Signal Example: The TDMA System in GSM / 18
1-4 About Bits, Symbols, and Waveforms / 29
1-4-1 Introduction / 29
1-4-2 Some Fundamentals of Digital Modulation Techniques / 38
1-5 Analysis of Wireless Systems / 47
1-5-1 Analog and Digital Receiver Designs / 47
1-5-2 Transmitters / 58
1-6 Building Blocks / 81
1-7 System Specifications and Their Relationship to Circuit Design / 83
1-7-1 System Noise and Noise Floor / 83
1-7-2 System Amplitude and Phase Behavior / 88
1-8 Testing / 114
1-8-1 Introduction / 114
1-8-2 Transmission and Reception Quality / 114
1-8-3 Base-Station Simulation / 118
1-8-4 GSM / 118

1-8-5 DECT / 118
1-9 Converting C/N or SNR to E_b/N_0 / 120

2 Models for Active Devices 123

2-1 Diodes / 124
2-1-1 Large-Signal Diode Model / 124
2-1-2 Mixer and Detector Diodes / 128
2-1-3 PIN Diodes / 135
2-1-4 Tuning Diodes / 153
2-2 Bipolar Transistors / 198
2-2-1 Transistor Structure Types / 198
2-2-2 Large-Signal Behavior of Bipolar Transistors / 199
2-2-3 Large-Signal Transistors in the Forward-Active Region / 209
2-2-4 Effects of Collector Voltage on Large-Signal Characteristics in the Forward-Active Region / 225
2-2-5 Saturation and Inverse Active Regions / 227
2-2-6 Small-Signal Models of Bipolar Transistors / 232
2-3 Field-Effect Transistors / 237
2-3-1 Large-Signal Behavior of JFETs / 246
2-3-2 Small-Signal Behavior of JFETs / 249
2-3-3 Large-Signal Behavior of MOSFETs / 254
2-3-4 Small-Signal Model of the MOS Transistor in Saturation / 262
2-3-5 Short-Channel Effects in FETs / 266
2-3-6 Small-Signal Models of MOSFETs / 271
2-3-7 GaAs MESFETs / 301
2-3-8 Small-Signal GaAs MESFET Model / 310
2-4 Parameter Extraction of Active Devices / 322
2-4-1 Introduction / 322
2-4-2 Typical SPICE Parameters / 322
2-4-3 Noise Modeling / 323
2-4-4 Scalable Device Models / 333
2-4-5 Conclusions / 348
2-4-6 Device Libraries / 359
2-4-7 A Novel Approach for Simulation at Low Voltage and Near Pinchoff Voltage / 359
2-4-8 Example: Improving the BFR193W Model / 370

3 Amplifier Design with BJTs and FETs 375

3-1 Properties of Amplifiers / 375
3-1-1 Introduction / 375
3-1-2 Gain / 380
3-1-3 Noise Figure (NF) / 385
3-1-4 Linearity / 415
3-1-5 AGC / 431
3-1-6 Bias and Power Voltage and Current (Power Consumption) / 436

3-2 Amplifier Gain, Stability, and Matching / 441
3-2-1 Scattering Parameter Relationships / 442
3-2-2 Low-Noise Amplifiers / 448
3-2-3 High-Gain Amplifiers / 466
3-2-4 Low-Voltage Open-Collector Design / 477
3-3 Single-Stage FeedBack Amplifiers / 490
3-3-1 Lossless or Noiseless Feedback / 495
3-3-2 Broadband Matching / 496
3-4 Two-Stage Amplifiers / 497
3-5 Amplifiers with Three or More Stages / 507
3-5-1 Stability of Multistage Amplifiers / 512
3-6 A Novel Approach to Voltage-Controlled Tuned Filters Including CAD Validation / 513
3-6-1 Diode Performance / 513
3-6-2 A VHF Example / 516
3-6-3 An HF/VHF Voltage-Controlled Filter / 518
3-6-4 Improving the VHF Filter / 521
3-6-5 Conclusion / 521
3-7 Differential Amplifiers / 522
3-8 Frequency Doublers / 526
3-9 Multistage Amplifiers with Automatic Gain Control (AGC) / 532
3-10 Biasing / 534
3-10-1 RF Biasing / 543
3-10-2 dc Biasing / 543
3-10-3 dc Biasing of IC-Type Amplifiers / 547
3-11 Push–Pull/Parallel Amplifiers / 547
3-12 Power Amplifiers / 550
3-12-1 Example 1: 7-W Class C BJT Amplifier for 1.6 GHz / 550
3-12-2 Impedance Matching Networks Applied to RF Power Transistors / 565
3-12-3 Example 2: Low-Noise Amplifier Using Distributed Elements / 585
3-12-4 Example 3: 1-W Amplifier Using the CLY15 / 589
3-12-5 Example 4: 90-W Push–Pull BJT Amplifier at 430 MHz / 598
3-12-6 Quasiparallel Transistors for Improved Linearity / 600
3-12-7 Distribution Amplifiers / 602
3-12-8 Stability Analysis of a Power Amplifier / 602
3-13 Power Amplifier Datasheets and Manufacturer-Recommended Applications / 611

4 Mixer Design 636

4-1 Introduction / 636
4-2 Properties of Mixers / 639
4-2-1 Conversion Gain/Loss / 639
4-2-2 Noise Figure / 641
4-2-3 Linearity / 645
4-2-4 LO Drive Level / 647

4-2-5 Interport Isolation / 647
4-2-6 Port VSWR / 647
4-2-7 dc Offset / 647
4-2-8 dc Polarity / 649
4-2-9 Power Consumption / 649
4-3 Diode Mixers / 649
4-3-1 Single-Diode Mixer / 650
4-3-2 Single-Balanced Mixer / 652
4-3-3 Diode-Ring Mixer / 659
4-4 Transistor Mixers / 678
4-4-1 BJT Gilbert Cell / 679
4-4-2 BJT Gilbert Cell with Feedback / 682
4-4-3 FET Mixers / 684
4-4-4 MOSFET Gilbert Cell / 693
4-4-5 GaAsFET Single-Gate Switch / 694

5 RF/Wireless Oscillators 716

5-1 Introduction to Frequency Control / 716
5-2 Background / 716
5-3 Oscillator Design / 719
5-3-1 Basics of Oscillators / 719
5-4 Oscillator Circuits / 735
5-4-1 Hartley / 735
5-4-2 Colpitts / 735
5-4-3 Clapp–Gouriet / 736
5-5 Design of RF Oscillators / 736
5-5-1 General Thoughts on Transistor Oscillators / 736
5-5-2 Two-Port Microwave/RF Oscillator Design / 741
5-5-3 Ceramic-Resonator Oscillators / 745
5-5-4 Using a Microstrip Inductor as the Oscillator Resonator / 748
5-5-5 Hartley Microstrip Resonator Oscillator / 756
5-5-6 Crystal Oscillators / 756
5-5-7 Voltage-Controlled Oscillators / 758
5-5-8 Diode-Tuned Resonant Circuits / 765
5-5-9 Practical Circuits / 771
5-6 Noise in Oscillators / 778
5-6-1 Linear Approach to the Calculation of Oscillator Phase Noise / 778
5-6-2 AM-to-PM Conversion / 788
5-6-3 Nonlinear Approach to the Calculation of Oscillator Phase Noise / 798
5-7 Oscillators in Practice / 813
5-7-1 Oscillator Specifications / 813
5-7-2 More Practical Circuits / 814
5-8 Design of RF Oscillators Using CAD / 825
5-8-1 Harmonic-Balance Simulation / 825
5-8-2 Time-Domain Simulation / 831

5-9 Phase-Noise Improvements of Integrated RF and Millimeter-Wave Oscillators / 831
5-9-1 Introduction / 831
5-9-2 Review of Noise Analysis / 831
5-9-3 Workarounds / 833
5-9-4 Reduction of Flicker Noise / 834
5-9-5 Applications to Integrated Oscillators / 835
5-9-6 Summary / 842

6 Wireless Synthesizers 848

6-1 Introduction / 848
6-2 Phase-Locked Loops / 848
6-2-1 PLL Basics / 848
6-2-2 Phase/Frequency Comparators / 851
6-2-3 Filters for Phase Detectors Providing Voltage Output / 863
6-2-4 Charge-Pump-Based Phase-Locked Loops / 867
6-2-5 How to Do a Practical PLL Design Using CAD / 876
6-3 Fractional-*N*-Division PLL Synthesis / 880
6-3-1 The Fractional-*N* Principle / 880
6-3-2 Spur-Suppression Techniques / 882
6-4 Direct Digital Synthesis / 889

APPENDIXES

A HBT High-Frequency Modeling and Integrated Parameter Extraction 900

A-1 Introduction / 900
A-2 High-Frequency HBT Modeling / 901
A-2-1 dc and Small-Signal Model / 902
A-2-2 Linearized T Model / 904
A-2-3 Linearized Hybrid-π Model / 906
A-3 Integrated Parameter Extraction / 907
A-3-1 Formulation of Integrated Parameter Extraction / 908
A-3-2 Model Optimization / 908
A-4 Noise Model Validation / 909
A-5 Parameter Extraction of an HBT Model / 913
A-6 Conclusions / 921

B Nonlinear Microwave Circuit Design Using Multiharmonic Load-Pull Simulation Technique 923

B-1 Introduction / 923
B-2 Multiharmonic Load-Pull Simulation Using Harmonic Balance / 924
B-2-1 Formulation of Multiharmonic Load-Pull Simulation / 924
B-2-2 Systematic Design Procedure / 925

B-3 Application of Multiharmonic Load-Pull Simulation / 927
B-3-1 Narrowband Power Amplifier Design / 927
B-3-2 Frequency Doubler Design / 933
B-4 Conclusions / 937
B-5 Note on the Practicality of Load-Pull-Based Design / 937

INDEX **939**

FOREWORD

One of the wonderful things about living in these times is the chance to witness, and occasionally be part of, major technological trends with often profound impacts on society and people's lives. At the risk of stating the obvious, one of the greatest technological trends has been the growth of wireless personal communication—the development and success of a variety of cellular and personal communication system technologies, such as GSM, CDMA, and Wireless Data and Messaging, and the spreading of the systems enabled by these technologies worldwide. The impact on people's lives has been significant, not only in their ability to stay in touch with their business associates and with their families, but often in the ability to save lives and prevent crime. On some occasions, people who have never before used a plain old telephone have made their first long distance communication using the most advanced satellite or digital cellular technology. This growth of wireless communication has encompassed new frequencies, driven efforts to standardize communication protocols and frequencies to enable people to communicate better as part of a global network, and has encompassed new wireless applications. The wireless web is with us, and advances in wireless global positioning technology are likely to provide more examples of lifesaving experiences due to the ability to send help precisely and rapidly to where help is urgently needed.

RF and microwave circuit design has been the key enabler for this growth and success in wireless communication. To a very large extent, the ability to mass produce high quality, dependable wireless products has been achieved through the advances of some incredible RF design engineers, sometimes working alone, oftentimes working and sharing ideas as part of a virtual community of RF engineers. During these past few years, these advances have generated a gradual demystification of RF and microwave circuitry, moving RF techniques ever so reluctantly from "black art" to science. Dr. Ulrich Rohde has long impressed many of us as one of the principal leaders in these advances.

In this book, *RF/Microwave Circuit Design for Wireless Applications*, Dr. Rohde helps clarify RF theory and its reduction to practical applications in developing RF circuits. The book provides insights into the semiconductor technologies, and how appropriate technology decisions can be made. Then, the book discusses—first in overview, then in detail—each of the RF circuit blocks involved in wireless applications: the amplifiers, mixers, oscillators, and frequency synthesizers that work together to amplify and extract the signal from an often hostile environment of noise and reflected signals. Dr. Rohde's unique expertise in VCO and PLL design is particularly valuable in these unusually difficult designs.

It is a personal pleasure to write this foreword—Dr. Rohde has provided guest lectures to engineers at Motorola, and provided suggestions on paths to take and paths to avoid to several design engineers. The value his insights have provided are impossible to measure, but are so substantial that we owe him a "thanks" that can never be expressed strongly enough. I believe that his impact on the larger RF community is even more substantial. This book helps share his expertise in a widely available form.

ERIC MAASS
Director of Operations, Wireless Transceiver Products
Motorola, SPS

PREFACE

When I started two years ago to write a book on wireless technology—specifically, circuit design—I had hoped that the explosion of the technology had stabilized. To my surprise, however, the technology is far from settled, and I found myself in a constant chase to catch up with the latest developments. Such a chase requires a fast engine like the Concorde.

In the case of this somewhat older technology, its speed still has not been surpassed by any other commercial approach. This tells us there is a lot of design technology that needs to be understood or modified to handle today's needs. Because of the very demanding calculation effort required in circuit design, this book makes heavy use of the most modern CAD tools. Hewlett-Packard was kind enough to provide us with a copy of their Advanced Design System (ADS), which also comes with matching synthesis and a wideband CDMA library. Unfortunately, some of the mechanics of getting us started on the software collided with the already delayed publication schedule of this book, and we were only in a position to reference their advanced capability and not really demonstrate it. The use of this software,

including the one from Eagleware, which was also provided to us, needed to be deferred to the next edition of this book. To give a consistent presentation, we decided to stay with the Ansoft tools. One of the most time-consuming efforts was the actual modeling job, since we wanted to make sure all circuits would work properly. There are too many publications showing incomplete or nonworking designs.

On the positive side, trade journals give valuable insight into state-of-the-art designs, and it is recommended that all engineers subscribe to them. Some of the major publications include:

Applied Microwave & Wireless
Electronic Design
Electronic Engineering Europe
Microwave Journal
Microwaves & RF
Microwave Product Digest (MPD)
RF Design
Wireless Systems Design

There are also several conferences that have excellent proceedings, which can be obtained either in book form or on CD:

GaAs IC Symposium (annual; sponsored by IEEE-EDS, IEEE-MTT)
IEEE International Solid-State Circuits Conference (annual)
IEEE MTT-S International Microwave Symposium (annual)

There may be other useful conferences along these lines that are announced in the trade journals mentioned above. There are also workshops associated with conferences, such as the recent "Designing RF Receivers for Wireless Systems," associated with the IEEE MTT-S.

Other useful tools include courses, such as *Introduction to RF/MW Design*, a four-day short course offered by Besser Associates.

Wireless design can be split into a digital part, which has to do with the various modulation and demodulation capabilities (advantages and disadvantages), and an analog part, the description of which comprises most of this book.

The analog part is complicated by the fact that we have three competing technologies. Given the fact that cost, space, and power consumption are issues for handheld and battery-operated applications, CMOS has been a strong contender in the area of cordless telephones because of its relaxed signal-to-noise-ratio specifications compared with cellular telephones. CMOS is much noisier than bipolar and GaAs technologies. One of the problems then is the input/output stage at UHF/SHF frequencies. Here we find a fierce battle between silicon-germanium (SiGe) transistors and GaAs technology. Most prescalers are bipolar, and most power amplifiers are based on GaAs FETs or LDMOS transistors for base stations. The most competitive technologies are the SiGe transistors and, of course, GaAs, the latter being the most expensive of the three mentioned. In the silicon-germanium area, IBM and Maxim seem to be the leaders, with many others trying to catch up.

Another important issue is differentiation between handheld or battery-operated applications and base stations. Most designers, who are tasked to look into battery-operated devices, ultimately resort to using available integrated circuits, which seem to change every six to nine months, with new offerings. Given the multiple choices, we have not yet seen a

systematic approach to selecting the proper IC families and their members. We have therefore decided to give some guidelines for the designer applications of ICs, focusing mainly on high-performance applications. In the case of high-performance applications, low power consumption is not that big an issue; dynamic range in its various forms tends to be more important. Most of these circuits are designed in discrete portions or use discrete parts. Anyone who has a reasonable antenna and has a line of sight to New York City, with the antenna connected to a spectrum analyzer, will immediately understand this. Between telephones, both cordless and cellular, high-powered pagers, and other services, the spectrum analyzer will be overwhelmed by these signals. IC applications for handsets and other applications already value their parts as "good." Their third-order intercept points are better than −10 dBm, while the real professional having to design a fixed station is looking for at least +10 dBm, if not more. This applies not only to amplifiers but also to mixer and oscillator performance. We therefore decided to give examples of this dynamic range. The brief surveys of current ICs included in Chapter 1 were assembled for the purpose of showing typical specifications and practical needs. It is useful that large companies make both cellular telephones and integrated circuits or their discrete implementation for base stations. We strongly believe that the circuits selected by us will be useful for all applications.

Chapter 1 is an introduction to digital modulation, which forms the foundation of wireless radiocommunication and its performance evaluation. We decided to leave the discussion of actual implementation to more qualified individuals. Since the standards for these modulations are still in a state of flux, we felt it would not be possible to attack all angles. Chapter 1 contains some very nice material from various sources including tutorial material from my German company, Rohde & Schwarz in Munich—specifically, from the digital modulation portion of their 1998 *Introductory Training for Sales Engineers* CD. *Note*: On a few rare occasions, we have used either a picture or an equation more than once so the reader need not refer to a previous chapter for full understanding of a discussion.

Chapter 2 is a comprehensive introduction to the various semiconductor technologies to enable the designer to make an educated decision. Relevant material such as PIN diodes have also been covered. In many applications, the transistors are being used close to their electrical limits, such as a combination of low voltage and low current. The f_T dependence, noise figure, and large-signal performance have to be evaluated. Another important application for diodes is their use as switches, as well as variable capacitances frequently referred to as tuning diodes. In order for the reader to better understand the meaning of the various semiconductor parameters, we have included a variety of datasheets and some small applications showing which technology is best for a particular application. In linear applications, noise figure is extremely important; in nonlinear applications, the distortion products need to be known. Therefore, this chapter includes not only the linear performance of semiconductors, but also their nonlinear behavior, including even some details on parameter extraction. Given the number of choices the designer has today and the frequent lack of complete data from manufacturers, these are important issues.

Chapter 3, the longest chapter, has the most detailed analysis and guidelines for discrete and integrated amplifiers, providing deep insight into semiconductor performance and circuitry necessary to get the best results from the devices. We deal with the properties of the amplifiers, gain stability, and matching, and we evaluate one-, two-, and three-stage amplifiers with internal dc coupling and feedback, as are frequently found in integrated circuits. In doing so, we also provide examples of ICs currently on the market, knowing that every six months more sophisticated devices will appear. Another important topic in this chapter is the choice of bias point and matching for digital signal handling, and we provide

insight into such complex issues as the adjacent channel power ratio, which is related to a form of distortion caused by the amplifier in its particular operating mode. To connect these amplifiers, impedance matching is a big issue, and we evaluate some couplers and broadband matching circuits useful at these high frequencies, as well as providing a tracking filter as preselector, using tuning diodes. Discussion of differential amplifiers, frequency doublers, AGC, biasing and push-pull/parallel amplifiers comes next, followed by an in-depth section on power amplifiers, including several practical examples and an investigation of amplifier stability analysis. A selection of power-amplifier datasheets and manufacturer-recommended applications rounds out this chapter.

Chapter 4 is a detailed analysis of the available mixer circuits that are applicable to the wireless frequency range. The design and the necessary mathematics to calculate the difference between insertion loss and noise figure are both presented. The reader is given insight into the differences between passive and active mixers, additive and multiplicative mixers, and other useful hints. We have also added some very clever circuits from companies such as Motorola and Siemens, as they are available as ICs.

Chapter 5, on oscillators, is a logical next step, as many amplifiers turn out to oscillate. After a brief introduction explaining why voltage-controlled oscillators (VCOs) are needed, we cover the necessary conditions for oscillation and its resulting phase noise for various configurations, including microwave oscillators and the very important ceramic-resonator-based oscillator. This chapter walks the reader through the various noise-contributing factors and the performance differences between discrete and integrated oscillators and their performance. Here too, a large number of novel circuits are covered.

Chapter 6 deals with the frequency synthesizer, which depends heavily on the oscillators shown in Chapter 5 and different system configurations to obtain the best performance. All components of a synthesizer, such as loop filters and phase/frequency discriminators, are evaluated along with their actual performance. Included are further applications for commercial synthesizer chips. Of course, the principles of the direct digital frequency synthesizer, as well as the fractional-*N*-division synthesizer, are covered. The fractional-*N*-division synthesizer is probably one of the most exciting implementations of synthesizers, and we have added patent information for those interested in coming up with their own designs.

The book then ends with two appendixes. Appendix A is an exciting approach to high-frequency modeling and integrated parameter extraction for HBTs. An enhanced noise model has been developed that gives significant improvement in the accuracy of determining the performance of these devices.

Appendix B is another CAD-based application for determining circuit performance—specifically, how to implement load-pulling simulation.

Appendix C is an electronic reproduction of a manual for a GSM handset application board that can be downloaded via web browser or ftp program from Wiley's public ftp area at ftp://ftp.wiley.com/public/sci-tech-med/microwave. It is probably the most exciting portion for the reader who would like to know how everything is put together for a mobile wireless application. Again, since every few months more clever ICs are available, some of the power consumption parameters and applications may vary relative to the system discussed, but all new designs will certainly be based on its general principles.

We would like to thank the many engineers from Ansoft, Alpha Industries, Motorola, National Semiconductor, Philips, Rohde & Schwarz, and Siemens Semiconductor (now Infineon Technologies) for supplying current information and giving permission to reproduce some excellent material.

In the area of permissions, National Semiconductor has specifically asked us to include the following passage, which applies to all their permissions:

LIFE SUPPORT POLICY

NATIONAL'S PRODUCTS ARE NOT AUTHORIZED FOR USE AS CRITICAL COMPONENTS IN LIFE SUPPORT DEVICES OR SYSTEMS WITHOUT THE EXPRESS WRITTEN APPROVAL OF THE PRESIDENT OF NATIONAL SEMICONDUCTOR CORPORATION.

As used herein:

1. Life support devices or systems are devices or systems which (a) are intended for surgical implant into the body or (b) support or sustain life and whose failure to perform, when properly used in accordance with instructions for use provided in the labeling, can be reasonably expected to result in a significant injury to the user.
2. A critical component is any component of a life support device or system whose failure to perform can be reasonably expected to cause the failure of the life support device or system, or to affect its safety or effectiveness.

I am also grateful to John Wiley & Sons, specifically George Telecki, for tolerating the several slips in schedule, which were the result of the complexity of this effort.

ULRICH L. ROHDE

Upper Saddle River, New Jersey
March, 2000

RF/MICROWAVE CIRCUIT DESIGN FOR WIRELESS APPLICATIONS

INDEX

Abrupt junction, 155–157
Abrupt-junction diode, capacitance versus total junction, 155–156
Acceptor, 140
Access burst, 28, 29
Acoustic measurements, 115
AD7008 DDS modulator, 892–893
Additive JFET mixer, 691, 693
Additive mixing:
  BJT, 637
  MOSFET, 638, 691
Adjacent-channel power ratio, 103–104, 114
  high-gain amplifiers, 470
AGC, 431, 433–436
AlGaAs/InGaAs HEMT, 313–317
Alloyed diodes, distortion product reduction, 170
Alternating voltage, modulating diode capacitance by, 186
Amplifiers
  adjacent-channel power ratio as function of RF source power, 429
  AGC, 431, 433–436
  biasing, *see* Biasing
  BJT, 439, 441
  class A, B, and C operation, 375–376
  compression, 415, 417
  constant-gain circles, 446
  differential, 522–525
  distributed, 378–379
  dynamic range, 415, 417
  emitter–ground connection, 436
  figure of merit, 446
  frequency doublers, 526–532
  gain, 380–385
  intermodulation distortion, 415, 417
  linearity
    analysis, 420–429
    requirements for digital modulation, 417
low-voltage open-collector design, 477–490
  collector–emitter voltage, 480, 482
  dc load line, 480, 482
  flexible matching circuit, 488–490
  open collector with inductor, 483–486
  open collector with inductor and R_{LOAD}, 487–489
  open collector with R_{LOAD}, 481–482, 484
  R_C as source resistor, 477–478
  transistor analysis, 477, 479
multistage, 507–512
  with automatic gain control, 532–534
noise factor, 386
noise figure, 377–378, 385–415
  bias-dependent noise parameters, 403–405
  cascaded networks, 396
  determining noise parameters, 414–415
  influence of external parasitic elements, 399–405
  measurements, 389–391
  noise circles, 405–408
  noise correlation in linear two-parts using correlation matrices, 408–412
  noisy two-port, 391–396
  signal-to-noise ratio, 387–389
  test equipment, 412–414
output, modulation signal, 423
π/4-DQPSK, circuit analysis, 429–432
potentially unstable, design, 451
power consumption, 436–442
properties, 375–380
push–pull/parallel, 547–550
single-stage feedback, 490–497
S parameter relationships, 442, 444–447
stability factor, 381–382

Amplifiers (*continued*)
time-domain magnitude of complex modulation signal, 429–429
transducer power, 445–446
two-stage, 497–507
voltage gain, 445
see also High-gain amplifiers; Low-noise amplifiers; Power amplifiers
Amplitude-imbalance errors, 672
Amplitude linearity, issues, 89, 91
Amplitude nonlinearity, 88–89
Amplitude shift keying, *see* ASK
Amplitude stability, oscillators, 731
AM-to-PM conversion, 101–102, 788–797
Analog FM, 62
Analog modulation:
single-sideband, 62–63
spectral considerations, 89–90
Analog receiver:
C/N, 47–48
design, 47–49
selectivity measurement, 109
Angelov FET model, dc *I–V* curves, 365
Ansoft physics-based MESFET model, 335
AP-to-PM distortion, 101
ASK:
bit error rate, 40–41
in frequency domain, 38–39
in I/Q plane, 38–39
in time domain, 38
AT21400 chip, 784–785
AT-41435 silicon tripolar transistor, noise parameters versus feedback, 402
Attenuation, versus angular frequency, 581–582
Automatic gain control, 148

BA243/244, specifications, 194
BA110 diode, capacitance/voltage characteristic, 173
Baluns, 713
Bandpass filter:
conversion of low-pass filter into, 582–583
networks, broadband matching using, 578, 580–585
Band spreading, 17–18
Bandwidth, effect on fading, 16
Barkhausen criteria, 720
Barrier height, Schottky diode, 133–134
Barrier potential, 127
Baseband modulation inputs, SA900, 64
Baseband waveforms, mapping data onto, 34–35
Base current, 222–223
Base-station
identification code, 28
simulation, 118
Base transport factor, 224
BAT 14-099, 654–657
BB141, capacitance/voltage characteristic, 174–175
BB142, capacitance/voltage characteristic, 174–175
BCR400 bias controller, 440–441, 546
BF995, 281–290
BF999, 276–280
BFG235, 472, 474
BFP420, 442–443
transistors in parallel, 492–493
BFP420 matched amplifier, 460–461
narrowband, 462–466
frequency-dependent gain, matching, and noise performance, 462, 468
frequency response, 464, 466
inductance for resonance, 462
input filter, 464–465
schematic, 463
BFP420 transistor, noise parameters, 403–405
BFP450 amplifier, 586–589
with distributed-element matching, 587–588
BFR193W, 370–371
Biasing, amplifiers, 436–439, 534–547
correction elements, 541–542
dc, 543–547
IC-type, 546–547
Lange coupler, 539
multiple coupled lines element, 539–540
OPEN element, 541–542
radial stubs, 540–541
RF, 543
STEP element, 541–542
T junction, cross, and Y junction, 536–538
transmission line, 534, 536
via holes, 540–541
Binary phase shift keying, *see* BPSK
Bipolar devices, scaling, 333
Bipolar junction transistor, *see* BJT
Bipolar transistors, 198–236
base current, 222–223
efficiency, 201–202
electrical characteristics, 202–218
ac characteristics, 203–218
collector–base capacitance, 208
collector–base time constant, 208
dc characteristics, 202–203
maximum frequency of oscillation, 208–209
reverse *I–V* characteristics, 202–203
S parameter, 203–206
transition frequency, 206–208
emitter current, 223
inverse current gain, 230
large-signal, forward-active region, 209, 219–224
collector voltage effects, 225–227
large-signal behavior, 199–209
leakage current effect, 229, 231–232
noise factor, 200–201, 341
npn planar structure, 219–220
output characteristics, 226
performance characteristics, 200–202

power gain, 200
power output, 201
saturation and inverse active regions, 227–232
sign convention, 199
small-signal models, 232–236
Bit error rate, 114
after channel equalizer, 12
noise and, 85–86
Rayleigh channel, 7–8
Bit synchronization, 24
BJT:
additive mixing, 637
amplifiers, 439, 441
Colpitts oscillator, input impedance, 721–722
high-frequency, noise factor, 396–397
noise model, 326–328
90-W push–pull amplifier, 598–600
BJT amplifier, 7-W class, 550–564
conducting angle, 551
dc *I–V* curves, 556, 559
efficiency, 552–553
frequency response, 556, 558
gain, 556, 558
as function of drive, 556, 563
heat sink, thermal resistance, 553
input matching network, 554
large-signal *S* parameters, 563
load line, 556, 559
output, 556, 560–562
matching network, 555
schematic, 557
BJT-based oscillators:
microwave, phase noise, 828
with noise feedback, 837–838
BJT DRO, 828–831
BJT Gilbert cell:
advantages, 679
with feedback, 682–690
validation circuit, 680
BJT microwave oscillator, 827–828
BJT model, 232–236
BJT oscillator, phase noise, 814, 819, 817, 824
as function of supply voltage, 812
BJT RF amplifier:
with distributed elements, 535, 543
with lumped elements, 535
Blocking, 92
dynamic range, 92
Bode equation, 581
Bode plot, phase-locked loops, 878–879
Body effect, 262
Boltzmann approximation, Fermi–Dirac distribution function, 220
BPF450 amplifier:
frequency-dependent responses, 591–592
schematic, 590–591
BPSK, 669
bandwidth requirements, 40, 42
bit error rate, 40–41, 43
constellation diagram, 40, 42
in frequency domain, 38–39
maximum interference voltages, 40, 42
Breakdown voltage:
versus capacitance ratio, testing, 162
PIN diodes, 142–143
testing, 180–181
Broadband matching:
single-stage feedback amplifiers, 496–497
using bandpass filter networks, 578, 580–585
Broadband modulation, 17
Burst:
structures, 23–29
bit synchronization, 24
compensation of multipath reception, 25–26
delay correction, 26–28
guard period, 26–27
information bits, 23–25
training sequence, 24–26
types, 28–29
Burst noise, JFET, 254

Capacitance:
adding across tuning diode, 794
connected in parallel or series with tuner diode, 183–186, 767–768
gate–source, MOS, 264
microstrip, 752
minimum, determining, 184–185
PIN diodes, 143–145
RF power transistors, 566–567
temperature coefficient, 162–164
testing, 174–177
as function of junction temperature, 175–176
modulating by applied ac voltage, 186
Capacitance diodes, 513–514
equivalent circuits, 174
Capacitance equations, MESFETs, 341–342
Capacitance ratio, 764, 767
determining, 184–185
testing, 167
Capacitors, interdigital, 539–540
Carrier concentrations, saturated *npn* transistor, 227
Carrier rejection, 672–674
Carrier-to-noise ratio, converting to energy per bit/normalized noise power, 119
Cascade amplifier, 497, 500–502
Cascaded networks, noise figure, 88, 396–399
Cascaded sigma-delta modulator, power spectral response, 884
CDMA, advantages and disadvantages, 20–21
CDMA signal, 17
CD4046 phase/frequency comparator, 858–860
Cellular telephone:
growth, 1

Cellular telephone (*continued*)
parameters, 56
standard, 55
system functions, 3–5
Ceramic-resonator oscillators, equivalent circuit calculation, 747–750
CFY77, 313–317
CGY94 GaAs MMIC power amplifier, 419–420
simulated signal, 423–428
CGY96 GaAs MMIC power amplifier, 417–418
CGY121A, 435–439
application circuit and parts list, 437–438
block diagram, 436
gain versus $V_{control}$, 439
Channel impulse response, 7–13, 26
delay spread, 9
echoes, 8–10
equalization, 9, 11–12
estimation, 11
time response, 14–16
Charge pump, 848, 853
external, 868, 870–872
Charge-pump-based phase-locked loops, 867–868, 870–876
Clapp–Gouriet oscillator, 730, 736–737
Clock recovery circuitry, 51, 53
CLY10, 927
CLY15, 317–321
output and power characteristics, 592–593
1-W amplifier, 589, 591–598
CLY15 amplifier:
frequency-dependent responses, 595, 598
schematic, 597
CMOS, 255
CMY91, 705, 708
CMY210, circuit, 698, 708
Code-division multiple access, *see* CDMA
Coherence bandwidth, 14
Coherent demodulation, 37–38
Collector–base capacitance, 208
Collector–base time constant, 208
Collector current, saturation region, 229–230
Collector efficiency, 202
Collector–emitter voltage, amplifiers, low-voltage open-collector design, 480, 482
Collector voltage, effects on large-signal bipolar transistors, 225–227
Colpitts oscillator, 725–727, 735–736, 773–775, 778
using RF negative feedback, 804, 806
Compression, amplifiers, 415, 417
Compression point, 1-dB, mixers, 645
Conduction angle, low-noise amplifiers, 448–449
Congruence transformation, 411
Constant-gain circles, 446
Contact potential, 132–133
Conversion gain/loss, mixers, 639–640
Cordless telephone:
parameters, 56
standards, 55
Correlation admittance, 393–394
Correlation matrix:
from *ABCD* matrix, 411–412
noise correlation in linear two-ports, 408–412
Correlation receiver, 36–37
Cross, 537
Cross-modulation, 99–100
PIN diodes, 149
testing, 168–170, 188–190
Crystal oscillators, 66, 716–717, 756–763
abbreviated circuit, 803–804
Colpitts, 758
electrical equivalent, 757
input impedance, 759
noise-sideband performance, 797
output, 761
parameters, 757
phase noise, 760, 763
phase noise versus reference frequency, 877
ultra-low-phase-noise applications, 762
Curtice cubic model, NE71000, 352
Cutoff frequency, 164
testing, 179–180

Damping factor, 864–865
Databank, generating for parameter extraction, 334
dc biasing, 543–547
IC-type amplifiers, 546–547
dc-coupled oscillator, 771–772, 775
dc models, comparison, 348–350
dc offset, mixers, 647
dc polarity, mixers, 649
dc-stabilized oscillator, 776–778
DECT, testing, 118–119
Delay correction, 26–28
Delay line, principles, 834–835
Delay spread, 9
Demodulation, digitally modulated carriers, 36–38
Depletion FETs, 309–310
Depletion zone, 143–144
Desensitization, 92
Desensitization point, 1-dB, mixers, 645
Detector diodes, 128–135
Device libraries, FETs, 359–361
Differential amplifiers, 522–525
Differential gain, 385
Differential group delay, 103–104
Differential phase, 385
Differential phase modulation, 38
Diffusion charge, 127
Diffusion current density, 220
Digital FM, 62
Digital I/Q modulator, 33

Digital modulation:
linearity requirements, 417
spectral considerations, 89–90
techniques, 38–46
Digital modulator, 30
Digital radiocommunication tester, 116–117
Digital receivers, selectivity measurement, 109
Digital recursion relation, 891
Digital tristate comparators, 855–863
Diode attenuator/switch, 670–671
Diode diffusion capacitance, 640
Diode loss, testing, 163–168
Diode mixers, 649–678
BAT 14-099, 654–657
diode-ring mixer, *see* Diode-ring mixer
single-balanced, 652–653, 658–660
single-diode, 650–653
subharmonically pumped single-balanced mixer, 659, 661
20 GHz, 706–708
Diode noise model, 323, 325–326
Diode-ring mixer, 659–660, 662–678
abode-cathode voltage, 666, 668
binary phase shift keying modulator, 669
conversion gain and noise figure, 662–663
diode attenuator/switch, 670–671
IF-output voltage, 667
image-reject mixer, 670–671
in-phase/quadrature modulator, 671–677
output, 664–665
phase detector, 669
quadrature IF mixer, 670
quadrature phase sift keying modulator, 669–670
responses for LO levels, 666
Rohde & Schwarz subharmonically pumped DBM, 677–678
schematic, 662
single-sideband modulator, 671–677
termination-insensitive mixer, 668–669
triple-balanced mixer, 676–677
two-tone testing, 666–667
Diode rings, phase/frequency comparators, 851–852
Diodes, 124–197
capacitance, 513–514
modeling, 124–125, 127
capacitance–voltage characteristic, 764
detector, 128–135
diffusion charge, 127
double-balanced mixer, noise figure and conversion gain versus LO power, 644
equivalent noise circuit, 325
hyperabrupt-junction, 516–518
I–V curves, 128
junction capacitance, 132–133
versus frequency, 134–136
large-signal model, 124–128
linear model, 135,137
mixer, 128–135, 137
noise figure versus LO power, 134
performance, 513–516
Schottky barriers, electrical characteristics and physics, 128–130
silicon versus GaAs, 134
small-signal parameters, 131–132
SPICE parameters, 126
see also PIN diodes; Testing
Diode switch, 191–197
as bandswitch, 193–196
data, 193–194
resonant circuits incorporating, 193–196
technology, 191–193
use in television receiver, 197
Diode-tuned resonant circuits, 765–769, 771
Direct digital synthesis, 889, 891–896
block diagram, 892–894
design guidelines, 891
digital recursion relation, 891
low-power, drawback, 892
Distortion, effects, power amplifiers, 416–420
Distortion ratio, 94–95
Distribution amplifiers, 602
DMOS, cross section, 269–270
Donor, 140
Dopants, 140
Doppler effect, 13–14
phase uncertainty, 16
Double-balanced mixers:
interport isolation, 660, 662–663
Rohde & Schwarz subharmonically pumped, 677–678
Doubly balanced “star” mixer, 708
Drain current, KGF1608, 357
Drain–source voltage, FET, 420–421, 423
Dual-conversion receiver, block diagram, 108
Dual-downconversion receiver, schematic, 47
Dual-gate MOS/GaAs mixers, 692, 694
DUALTX output matching network, 67–68
Dummy burst, 28–29
Dynamic measure, 96–99
Dynamic range, 96, 111
mixers, 645

Early voltage effect, 484–485
Ebers–Moll equations, 230–231
Echo profiles, 8–9, 13
Edge-triggered JK master–slave flip-flops, phase/frequency comparators, 852–855
Efficiency, bipolar transistors, 201–202
EG8021 monolithic amplifier, 376–378
Electrical properties, testing, 178–181
Emitter current, 223
saturation region, 229–230
Enhancement FETs, 309–310
Envelope delay, 103–104

Epitaxial-collector, 199
Equivalent noise conductance, 394–395
ESH2/ESH3 test receiver, 769, 771
Excess noise, 398
Excess noise ratio, 413
Exponential transmission lines, 578
Eye diagrams, 422–423
  π/4-DQPSK, 429–430

Fading, 5–6
  effect of bandwidth, 16
  simulator, 12
FDMA, advantages and disadvantages, 18–19
Feedback amplifier, elements, 494
Feedback oscillator, 733
Fermi–Dirac distribution function, Boltzmann approximation, 220
FET amplifier, 381–383
  circuit diagram, 381
  single-tone RF power sweep analysis, 420–421
FETs, 237–321
  device libraries, 359–361
  drain current, 556, 564
  drain–source voltage, 420–421, 423
  equivalent noise circuit, 251, 253
  forward-based gate model, 342
  linear model, 251
  models
    ac errors, 359
    dc errors, 348
  modified Materka model, dc *I–V* curves, 367–368
  MOSFETs, 254–262
  noise modeling, 323, 325–333
  operating parameters, 237, 240
  parameter extraction, 338–339, 341
    generating databank, 334–337
  scalable device models, 333–334
  short-channel effects, 266–271
  simulation at low voltage and near pinchoff voltage, 359, 365–370
  SPICE parameters, 322–325
  types, 237–239
Field-effect transistors, *see* FETs
Figure of merit:
  amplifiers, 446
  amplitude linearity, 89, 91
  dynamic measure, 96–99
  error vector magnitude, 111–113
  I-dB compression point, 92
  intermodulation intercept point, 93–95
  maximum frequency of oscillation, 208–209
  noise figure, *see* Noise figure
  noise power ratio, 100–101
  transition frequency, 206–208
  triple-beat distortion, 99–100
Film resistor, equivalent model, 79
Filter attenuator, π-mode, 150–151
Filters:
  frequency response/phase-noise analysis graph, 883
  phase detectors providing voltage output, 863–870
  phase-locked loops, passive, 872–876
  voltage-controlled tuned, 513–522
Flicker corner frequency, 326–327, 329, 332
Flicker noise, 782, 784
  cleaning up, 834, 836
  effect on noise-sideband performance, 789–790
  integrated RF and millimeter-wave oscillators, 834–835, 837–838
Flicker noise coefficient, 326–327, 329, 332
Forward current, as function of diode voltage, 134–135
Forward error correction, 114
Forward transconductance curve, 246–247
Four-reactance networks, 573–578
Fractional-*N*-division PLL synthesis, 880–890
  spur-suppression techniques, 882–890
Fractional-*N*-division synthesizer, phase noise, 886–887
Fractional-*N* principle, 880–882
Fractional-*N* synthesizer, block diagram, 884
Frequency shift keying, 35
Frequency correction burst, 28
Frequency-division duplex transceiver, 63
Frequency-division multiple access, *see* FDMA
Frequency doubler:
  circuit topology, 934
  conversion purity, 935–936
  dc *I–V* curves, 531–532
  design, using multiharmonic load-pull simulation, 933–937
  frequency-dependent gain, 529–530
  input and output voltage waveforms, 935, 937
  output spectrum, 529, 531
  schematic, 526–527
  spectral purity, 934–936
Frequency doublers, 526–532
Frequency pushing, 813
Frequency ratio, output voltage as function of, 857–858
Frequency shift, testing, 188
Frequency synthesizer, block diagram, 717
Fukui's expression, 408
Fundamental angle-modulation theory, 46

GaAs, testing, 158–159
GaAsFET amplifier, dc-coupled, 502–503, 506–507
GaAsFET feedback amplifier, 466–468
GaAsFET single-gate switch, 694–713
  circuit, 695
  physical layout of, 696
GaAsFET wideband amplifiers, 382–385
GaAs MESFETs, 325
  datasheet, 317–321
  disadvantages, 303

extrinsic model, 305
large-signal behavior, 301, 303–310
large-signal equations, 304, 306–307
linear equivalent circuit, 310–311
modified Materka-Kacprzak model, 304, 307–309
noise model, 328–330
package model, 305
small-signal model, 310–321
structure, 302
types, 309–310
GaAs MMIC, 699–704
Gain:
amplifiers, 380–385
circles, 406
compression, 92–93
multiple-signal, 100
definitions, 383
differential, 385
as function of drive, 556, 563
saturation, 92
Gaussian minimum shift keying, 35, 62
GMSK, 35, 62
Graded junction, 513–514
Group delay, 103–104
Groupe Special Mobile:
pulsed signal, 432
testing, 118
see also TDMA, in GSM
Guard period, 26–27
Gummel–Poon BJT model, 209, 219, 326

Handheld transceiver, block diagram, 3–4
Harmonic-balance simulation, 923–924
multiharmonic load-pull simulation using, 924–927
RF oscillators, 825–282
Harmonic distortion, testing, 170–171
Harmonic generation, 188
Harmonic intermodulation products, mixers, 645–646
Harmonic mixing, 674
Hartley microstrip resonator oscillator, 756
Hartley oscillator, 725–726, 735–736
Health effects, potential, 1–2
Heat sink, thermal resistance, 553
Heterojunction bipolar transistors, 900–921
integrated parameter extraction, 907–909
intrinsic noise parameters, 907
model
dc and small-signal, 902–904
dc *I–V* curves, 914
equivalent circuit, 901–902
linearized hybrid-π, 906–907
linearized T, 904–906
noise figures, 918–920
optimization, 908–909
parameter extraction, 913–920
S parameters, 915–918
modeling, 901–907
noise figure, 904–905
noise model, validation, 909–913
package parasitics, 902
HF/VHF voltage-controlled filter, 518–521
High-frequency field, PIN diodes applications, 147–148
High-frequency signals, amplitude control, PIN diodes, 148, 150–151
High-gain amplifiers, 466, 468–477
adjacent-channel power ratio, 470
BFG235, 472, 474
class A, B, and C operation, 466, 468–469
dc *I–V* curves, 469–470
noise figure, 469
third-order intercept point, 470–471
three-tone analysis, 470–471, 473
tuned circuits, 468
Hopf bifurcation, 608
Hybrid synthesizer, 893, 896
Hyperabrupt-junction diode, 158–159
Hyperabrupt-junction tuning diodes, 516–518

ICOM IC-736 HF/6-meter transceiver, 893–894
IC-type amplifiers, dc biasing, 546–547
IF image, 636–637
Image-reject mixer, 670–671
Impact ionization, 273–274
Impedance:
input
Colpitts oscillator, 721–722
crystal oscillator, 759
negative-resistance oscillator, 728–729
RF power transistors, 565–566
junction, 191–192
output
matching, SA900, 67–68
RF power transistors, 565–567
transformation equation, 380
Impedance inverters, 582, 584
Impedance matching networks, applied to RF power transistors, 565–585
broadband matching using bandpass filter networks, 578, 580–585
exponential lines, 578
four-reactance networks, 573–578
matching networks using quarter-wave transformers, 578–580
three-reactance matching networks, 570–574
two-resistance networks, 567–570
use of transmission lines and inductors, 570–571
Inductors, printed, 536, 538
Information channel, 31
In-phase/quadrature modulator, 671–677
Input matching network, CLY15, 592–593, 595–596
Input selectivity, 108

Integrated parameter extraction, HBT:
formulation, 908
model optimization, 908–909
Integrated RF and millimeter-wave oscillators:
flicker noise reduction, 834–835, 837–838
phase-noise improvements, 831–842
applications, 835, 838–842
workarounds, 833–836
Interdigital capacitors, 539–540
Intermodulation:
large-signal effects, 100, 102
PIN diodes, 149
testing, 170
Intermodulation distortion, 92–96
amplifiers, 415, 417
mixers, 646–647
quasiparallel transistors, 600–602
Intermodulation intercept point, 93–95
Intersymbol interference, 26
Inverse current gain, 230
I/Q generator, digital FM baseband, 62
I/Q modulation, 34
transmitters, 58–63
I/Q modulator:
equations, 76–77
mathematical representation, 58–59
IS-54 front-end chipset, 63–65
IS-54 handsets, configurations, 66
ISM band application, SA900, 73, 76
Isolation, mixers, 647–648

JFETs:
burst noise, 254
datasheet, 241–245
large-signal behavior, 246–250
lowest-noise, 784
modified Materka model, 246
noise characteristics, 253
noise model, 328–330
nonlinear model, 250
small-signal behavior, 249, 251–254
static characteristics, 246–247
structure, 302
Johnson noise, resistor, 387–388
Jones cell, 710
Junction capacitance:
versus frequency, 134–136
range versus voltage, 134–136
Schottky barrier chip, 132–133
Junction field-effect transistor, *see* JFETs
Junction impedance, 191–192

Ka-band MMIC voltage-controlled oscillator, 838, 840–841
KGF1608, 348, 354–358
dc *I–V* curves, 356
drain current, 357
output power, 358
Kirchhoff's equations, 468

Lange coupler, 539
four-strip version, 548–549
Large-signal diode model, 124–128
Large-signal effects, 100, 102
LDMOS FETs, 270–271, 325, 612–616, 693, 819
Leakage current, testing, 180–181
Leeson equation, 736–737
Lifetime, 141
Linear digital modulation, 60–62
Linear diode model, 135, 137
Linear distortion, 88
Linearized hybrid-π model, 906–907
Linearized T model, 904–906
Linearly graded junction, testing, 156–158
Linear modulations, 34–35
LMX2350-based synthesizer, 888–890
LO drive level, mixers, 647
Load-pull technique, 923–938
Logical symbols, 30
LO harmonics, 48–49
Loop-filter design,improper, 106
LO outputs, 64, 66
LO power, versus noise figure, diodes, 134
Lossless feedback, single-stage feedback amplifiers, 495–496
Low-noise amplifiers, 448–468
BFP420 amplifier
matched, 460–461
narrowband, 462–466
conduction angle, 448–449
design guidelines, 451–452
effective FR voltage, 451
fundamental and harmonic currents, 450–451
GaAsFET feedback amplifier, 466–468
NE68133 matched amplifier, 452–459
power gain, 448
saturation voltage, 448
using distributed elements, 585–592
push–pull BJT amplifier, 598–600
1-W amplifier using CLY15, 589, 591–598
Low-pass filter, conversion into bandpass filter, 582–583
Lumped-resonator oscillator, 744–745

Maas mixer, 707
Mapping equation, 925
M-ary phase shift keying modulation, *see* MPSK
Materka FET, scaling, 334
Materka FET model, modified, dc *I–V* curves, 367–368
Materka-Kacprzak model, modified, GaAs MESFETs, 304, 307–309
Materka model:
modified, 246
NE71000, 351

Maxim Integrated Products, 77, 79
Maximum available gain, 200
Maximum frequency of oscillation, 208–209
MBE MESFET, 362–364
MC1350/1490, 532–534
MC 12040 phase/frequency comparator, 858, 860
MC12148 ECL oscillator, 815, 817, 822–823
MC13109FB, test circuit, 78–79
MC13143, frequency responses, 683
MC13144, 501–504
Mesa processing, 159–160
MESFET doubler:
  gain comparison, 335–336
  layout, 337
MESFETs, 927–929
  capacitance equations, 341–342
  circuit partitioned into linear and nonlinear subcircuits, 826
  GaAs, *see* GaAs MESFETs
  intrinsic model and complete chip/package model, 340
  parameter extraction, 340–348
  physics-based modeling, 359, 362–364
  R_{DS}, 369
MEXTRAM, 556, 563
MGA64135 MMIC amplifier, 472, 475–477
Microstrip inductor:
  high-Q, 751–753
  as oscillator resonator, 748–756
Microwave diode, scaling, 333
Miller effect, 586
Minimum detectable signal, 83
Minimum shift keying, 35
Minority-carrier charge, 221–222
Minority-carrier concentration, saturated transistor, 227–229
Mixed-mode MFSK communication system, 50–57
  baseband circuitry, 50, 52
  BER versus SNR, 51, 54
  block diagram, 50
  clock recovery circuitry, 51, 53
  PLL CAD simulation, 51, 53–57
  received signal, 51, 53
  RF section, 51–52
Mixer diodes, 128–135, 135, 137
  equivalent circuit, 640
Mixers, 636–713
  conversion gain/loss, 639–640
  dc offset, 647
  dc polarity, 649
  dynamic range, 645
  harmonic intermodulation products, 645–646
  intermodulation distortion, 646–647
  interport isolation, 647–648
  linearity, 645–647
  LO drive level, 647
  noise figure, 641–645
  port VSWR, 647, 649
  power consumption, 649
  SSB versus DSB noise figure, 645
  see also Diode mixers; Transistor mixers
Mobile station, synchronization, 27
Modulation noise analysis, 803
Modulator, cascaded sigma-delta modulator, power spectral response, 884
MOS:
  devices, transfer characteristics, 255–261
  electron drift velocity versus tangent electric field, 266–267
  gate–source capacitance, 264
  I–V characteristics, 259, 261
  small-signal model in saturation, 262–265
  threshold voltage, 258
  voltage limitations, 261–262
MOSFET Gilbert cell, 693–694
MOSFET oscillator, phase noise, 814, 819
MOSFETs:
  additive mixing, 638, 691
  equivalent noise circuit, 331
  f_T, 265
  large-signal behavior, 254–262
  model of velocity saturation, 268
  multiplicative mixing, 638
  noise model, 331–333
  structure, 302
  substrate flow, 273–274
  subthreshold conduction, 271–273
MPSK, 15–16
MRF186, 617–623
MRF899, 625–630
MRF5003, 291–300
MSA-0375 MMIC amplifier, 501, 505
Multiharmonic load-pull simulation, 923–937
  circuit topology, 924–925, 927
  design procedure using, 926
  formulation, 924–925
  frequency doubler design, 933–937
  narrowband power amplifier design, 927–934
  output power spectrum, 931, 933
  practicality, 937
  second-harmonic, 931–932
  systematic design procedure, 925–927
Multipath reception, compensation, 25–26
Multiplicative mixing, MOSFET, 638
Multistage amplifiers, 507–512
  with automatic gain control, 532–534
  stability, 512

Narrowband modulation, 17
Narrowband power amplifier, design, 927–934
NE67300, nonlinear device library datasheet, 360–361
NE71000:
  Curtice cubic model, 352
  dc *I–V* curves, 343

NE71000 (*continued*)
Materka model, 351
S parameters, 344–348
TOM model, 353
NE42484A, 786–787
NE5204A IC, 512
NEC UPC2749, 507–509
Negative-resistance oscillator, input impedance, 728–729
NE68133 matched amplifier, 452–459
circles for gain, noise figure, and source and load-plane stability, 453, 455
input matching-network extraction, 455–456
intermodulation distortion outputs, 458
optimized performance, 456–457
output constellation, 458–459
output matching-network extraction, 455–456
NE/SA5204A amplifier, 508–509
N-JFET mixer, 691, 693
NMOS:
with bias voltages applied, 258–259
depletion region, 255, 257
enhancement-mode structure, 255–258
transfer characteristic, 271–273
Noise:
conversion analysis, 801, 803
excess, 398
mechanisms, 800–801
modeling, 323, 325–333
in oscillators, 778–812
AM-to-PM conversion, 788–797
causes, 782
generation, 798
sideband, 789–790
sources, 393–394
see also Phase noise; Signal-to-noise ratio; System noise
Noise analysis, review, 831–833
Noise bandwidth, 388–389
Noise circles, 405–408
Noise correlation, linear two-ports, using correlation matrices, 408–412
Noise correlation matrix, 906
Noise equivalent resistance, 394
Noise factor, 86–88
amplifiers, 386
bipolar transistors, 200–201, 341
mixer, exact mathematical nonlinear approach, 642–644
in terms of correlation matrix, 412
Noise figure, 86–88, 201
amplifiers, 377–378
for antennas and antenna systems, 87
cascaded networks, 88, 396–399
as function of external feedback, 402–403
HBT, 904–905, 918–920
high-gain amplifiers, 469
versus LO power, diodes, 134
lowest, 585–586
mixers, 641–645
SSB versus DSB, 645
temperature dependency, 913
test equipment, 412–414
see also Amplifiers, noise figure
Noise floor, 83
Noiseless feedback, single-stage feedback amplifiers, 495–496
Noise matrix, transformation, 410–411
Noise model:
bijunction transistor, 326–328
GaAs MESFETs, 328–330
JFET, 328–330
MOSFET, 331–333
validation, HBT, 909–913
Noise parameters:
bias-dependent, 403–405, 911–913
determining, 414–415
transformation matrix, 400–401
Noise performance, RF oscillators, 736
Noise power, thermal, 386
Noise power ratio, 100–101
Noise-sideband:
crystal oscillator, 797
as function of flicker frequency, 789–790
influence of tuning diodes, 791–792
power, 112
Noise temperature, 88
Noisy nonlinear circuit, equivalent representation, 798–799
Noisy two-port, 391–396
ABCD- matrix description, 392
cascaded, 396–399
noise correlation using correlation matrices, 408–412
S-parameter form, 392–393
Nonlinear distortion, 88
npn, 198
NPN silicon RF power transistor, 625–630
Nyquist criterion, 720
Nyquist's equation, 394, 788
Nyquist stability analysis, power amplifiers, 603, 606–607
NZA, datasheet, 241–245

Offset QPSK, 45–46
On-chip clocks, 68, 70
Oscillating amplifier, phase noise, 608–610
Oscillation:
approximate frequency, 606, 608–610
where it begins, 608, 610–611
Oscillators:
ac load line, 810
amplitude stability, 731
background, 716, 718

Barkhausen criteria, 720
block diagram, 719
with capacitive voltage divider, 720–722
Clapp–Gouriet, 730, 736–737
coarse and fine tuning, 769–771, 775
Colpitts, 725–727, 735–736
conversion noise analysis, 801, 803
dc *I–V* curves, 810
design, 719–735
equivalent representation, noisy nonlinear circuit, 798–799
experimental variations, 803–805
feedback, 733
frequency conversion approach, 798–802
Hartley, 725–726, 735–736
input impedance, 721–722
lumped-resonator, 744–745
modulation noise analysis, 803
NE42484A, 786–787
noise, 105
Nyquist criterion, 720
output, 808–809
phase noise
causes, 782–783
comparison between predicted and measured, 807
equivalent feedback models, 780–782
linear approach to calculating, 778–788
nonlinear approach to calculating, 798–812
optimization, 805, 811–812
phase stability, 731–735
practical circuits, 814–824
push–pull, 814, 817
short-term frequency stability, 732
silicon/GaAs-based integrated VCOs, 817–822, 825
specifications, 813–814
three-reactance oscillators, 723–728
two-port oscillator, 728–731
types, 716
see also Integrated RF and millimeter-wave oscillators; Noise, in oscillators; RF oscillators
Output impedance matching, SA900, 67–68
Output load, RF power transistors, 566–567
Output matching network, CLY15, 592, 594–596
Output power, KGF1608, 358

Parallel-resonant circuit:
testing, 181–183
tuning range, 769, 771
Parameter extraction:
generating databank, 334–337
MESFETs, 340–348
test setup, 338
Parasitic effects, amplifiers, 399–405
Phase-cancellation network, 672
Phase constellation, 112–113
Phase detectors:
diode-ring mixer, 669
providing voltage output, filters, 863–870
Phase errors, 111–112
Phase feedback loop, closed-loop response, 781
Phase/frequency comparators, 851–863
with antibacklash circuit, 862–863
digital tristate, 855–863
diode rings, 851–852
edge-triggered JK master–slave flip-flops, 852–855
Phase-imbalance errors, 672
Phase-locked loops, 848–880
basics, 848–851
Bode plot, 878–879
charge-pump-based, 867–868, 870–876
damping factor, 864–865
design using CAD, 876–880
external charge pump, 868, 870–872
filter
passive, 872–876
for phase detectors providing voltage output, 863–870
fractional-*N*-division synthesis, 880–890
linearized model, 850
nonlinear, 850
phase/frequency comparators, 851–863
second-order, 864
third-order, 866
reference-energy suppression, 873–874
transient response, 867–870
VCO operation, 850
Phase-locked-loop synthesizer, 748, 750
block diagram, 848–849
Phase-locked loop system, CAD-based, 51, 53–57
block diagram, 51, 54
phase noise, 51, 53–55
Phase noise, 111–112
added to carrier, 778–779
BJT oscillator, 817, 824
ceramic-resonator-based oscillator, 749–750
comparison of BJT and MOSFET oscillators, 814, 819
crystal oscillator, 760, 763
effects, 103, 105–107
fractional-*N*-division synthesizer, 886–887
as function of supply voltage, 812
microwave BJT oscillator, 828
modeled by noise-free amplifier and phase modulator, 780
oscillating amplifier, 608–610
with oscillator output, 734–735
oscillators
causes, 782–783
comparison between predicted and measured, 807
equivalent feedback models, 780–782

Phase noise (*continued*)
oscillators (*continued*)
linear approach to calculating, 778–788
nonlinear approach to calculating, 798–812
optimization, 805, 811–812
versus reference frequency, 877
RF oscillators, 738–739
Siemens IC oscillator, 816
spectral density, 780–781
two-differential-amplifier oscillator circuit, 821
VCO, 774, 777, 831–832
optimization and, 805, 811
Phase nonlinearity, 88–89
Phase perturbation, 782, 784
Phase response, issues, 103
Phase-shift analysis, parallel tuned circuit, 732
Phase shift keying, 38–39
Phase stability, oscillators, 731–735
Phase uncertainty, Doppler effect, 16
π/4-DQPSK:
baseband generator, 60
circuit analysis, 429–432
eye diagram, 429–430
signal constellation, 60–61
Pinched-FET model, 342
Pinchoff voltage, 246
PIN diodes, 135–153
amplitude control of high-frequency signals, 148, 150–151
applications, 146–153
breakdown voltage, 142–143
capacitance, 143–145
cross-modulation, 149
current versus voltage, 143
dopants, 140
equivalent series circuit, 145
figure of merit, 145
insertion loss versus frequency, 145, 147
intermodulation, 149
large-signal model, 136–138
lifetime, 141
model keywords, 139
π network for TV tuners, 151–153
resistance, 141
forward, versus forward current, 148
as function of dc, 137, 139
reverse series, 145–146
reverse shunt, 147
series, as function of bias, 142
variable, 138, 140–142
ring, 670–671
scaling, 333
π network, TV tuners, PIN diodes, 151–153
Planar, 198
Planar process, 159–160
Plessey SL610 wideband amplifier with AGC, 433–435
PMOS, 255
Post-tuning drift, 167–168
Power amplifiers, 416–420, 550–611
BJT amplifier, 7-W class, 550–564
classes, 550
distribution amplifiers, 602
impedance matching networks, *see* Impedance matching networks, applied to RF power transistors
low-noise amplifier, using distributed elements, 585–592
MRF186, 617–623
MRF899, 625–630
Nyquist stability analysis, 603, 606–607
oscillation
approximate frequency, 606, 608–610
where it begins, 608, 610–611
output current, 550–551
PTF 10009, 612–616
quasiparallel transistors, improved linearity, 600–602
small-signal ac analysis, 603–605
stability analysis, 601–611
unstable, 606–608
Power consumption:
mixers, 649
Power gain, bipolar transistors, 200
Power ON time, SA900, 73, 75
Power output, bipolar transistors, 201
Power ratios–voltage ratios, 380
Printed inductors, 536, 538
PSK, 38–39
PTF 10009, 612–616
Punchthrough, 157
voltage, 144
Push–pull BJT amplifier, 598–600
Push–pull oscillator, 814, 817
using LDMOS FETs, 819
Push–pull/parallel amplifiers, 547–550

QAM, 43, 46
Q factor, 142–146
versus bias, 166
definitions, 163–165
testing, 163–168, 177–178
QPSK:
band-limited signal, 44, 46
bandwidth requirements, 40, 42
baseband generator, 60
bit error rate, 40–41, 43
constellation diagram, 40, 42
maximum interference voltages, 40, 42
modulation in time and frequency domains, 40–41
modulator, 40
serial-to-parallel conversion, 60–61
signal constellation, 60–61
spectrum, 40–41

pseudorandom binary sequence data, 44, 46
Quad-D circuit, 858
Quadrature amplitude modulation, 43, 46
Quadrature IF mixer, 670
Quadrature phase sift keying modulator, 669–670
Quadrature phase shift keying, *see* QPSK
Quality factor, "wide" microstrip, 752
Quarter-wave transformers, matching networks using, 578–580
Quasiparallel transistors, 600–602

Radial bend, 537
Radial stubs, 540–541
Radiation, "harmful," 2
Radio channel, characteristics, 5–7
Rayleigh channel, bit error rate, 7–8
Rayleigh distribution, 6–7
RC filter, schematic, 863–864
Receive signal, as a function of time or position, 6–7
Reception quality, 114–117
Reciprocal mixing, 105, 107–111
Reflection coefficient, 396
  between transformed load and generator, 580
  input, 445
  output, 408, 445
Resistor, Johnson noise, 387–388
Resonant circuits:
  diode-tuned, *see* Testing
  incorporating diode switches, 193–196
RF amplifier, with active biasing, 544–545
RF biasing, 543
RF carrier:
  digitally modulated
    demodulation, 36–38
    spectrum, 36
  modulated
    generation, 33–34
    waveform, 31–33
RF harmonics, 48–49
RFICs, selector guide, 79–82
RF image, 636–637
RF oscillators, 736–778
  buffered, 741–743
  ceramic-resonator oscillators, 745–750
  coarse and fine tuning, 769–771, 775
  Colpitts, 773–775, 778
  crystal oscillators, 756–763
  dc-coupled, 771–772, 775
  dc-stabilized, 776–778
  design flowchart, 744
  design using CAD, 825–831
    harmonic-balance simulation, 825–828
    time-domain simulation, 828–831
  diode-tuned resonant circuits, 765–769, 771
  Hartley microstrip resonator oscillator, 756
  increasing loaded Q, 749–751
  microstrip inductor, as oscillator resonator, 748–756
  noise performance, 736
  phase noise, 738–739
  two-port microwave/RF oscillator, 741–745
  UHF VCO using the tapped-inductor differential oscillator, 753–756
  voltage-controlled oscillators, 758, 764–766
  see also Integrated RF and millimeter-wave oscillators
RF parameters, versus local-oscillator drive level, 135–136
RF power FETs, 291–300, 617–623
RF power transistors:
  frequency response, 567–568
  impedance matching networks, *see* Impedance matching networks, applied to RF power transistors
  input impedance, 565–566
  output impedance, 565–567
  output load, 566–567
  termination reactance compensation, 569–570
RF source power, adjacent-channel power ratio as function of, 429
Rice distribution, 6–7
Richardson equation, 129–130
Rohde & Schwarz radiocommunication tester, 115–116
Rohde & Schwarz SMDU signal generator, 739–741
Rohde & Schwarz subharmonically pumped DBM, 677–678
Roll-off compensation network, 583, 585

SA620, 749, 751–752
  schematic, 755
SA900, 58–59
  amplitude and phase imbalance, 72
  architecture, 63–64
  baseband I/Q inputs, 64
  crystal oscillator, 66
  designing with, 64, 66–69
  ISM band application, 73, 76
  modes of operation, 68
  on-chip clocks, 68, 70
  output impedance matching, 67–68
  output matching using *S* parameters, 68–69
  performance, 70–71
  power ON time, 73, 75
  spectral mask, 73–75
  transmit local oscillator, 64, 66
  transmit modulator, 58–59
  VCO, 66–67
Saturation voltage, low-noise amplifiers, 448
Scaling, FETs, 333–334
Schottky barrier chip, junction capacitance, 132–133
Schottky barriers, electrical characteristics and physics, 128–130

Schottky diodes, 640
band diagrams, 133
barrier height, 133–134
chip cross section, 129
diode mixers, 652–653
as noise generator, 641
silicon versus GaAs, 134
Scout program, 338, 526
user interface, 338–339
Selectivity curves, four-reactance networks, 575–577
Sensitivity, 84–85
Series inductance, testing, 178–180
Series resistance, testing, 177–178
Series resonant frequency, testing, 178–180
Shockley equation, 124
Short-channel effects, FETs, 266–271
Siemens IC oscillator, 814–815
Siemens NPN silicon RF transistor, 210–218
Sigma-delta modulator, cascaded, power spectral response, 884
Signal generator, phase noise, 107
Signal representation, different forms, 33
Signal-to-noise ratio, 84–85
amplifiers, 387–389
converting to energy per bit/normalized noise power, 119
measurement, 389
Silicon, 198
testing, 158–159
Silicon-based BiCMOS, 835, 838–839
Silicon dual gate mixer, 710
Silicon dual Schottky diode, 654–657
Silicon/GaAs-based integrated VCOs, 817–822, 825
Silicon inductor, 526, 528–529
Silicon N channel MOSFET tetrode, 281–290
Silicon N channel MOSFET triode, 276–280
SINAD ratio, 85
Single-balanced mixer, 652–653, 658–660
subharmonically pumped, 659, 661
Single-BJT mixer, 678–679
Single-diffused diodes, distortion product reduction, 171
Single-diode mixer, 650–653
conversion gain and noise figure, 651
output spectrum, 651
Single-loop synthesizer, block diagram, 848–849
Single-sideband:
noise figure, measurements, 413–414
signal-to-noise ratio, 788
suppression contours, 73
Single-sideband AM, 62–63
Single-sideband modulator, 671–677
return loss, 678
Single-sideband phase noise, 105
Single-stage feedback amplifiers, 490–497
broadband matching, 496–497
lossless or noiseless feedback, 495–496
transconductance, 493–494
voltage gain, 490
Single-tone gain-compression factor, 92
Small-signal ac analysis, power amplifiers, 603–605
Smith charts, 444
Smith diagram, 585–586
S parameters, 203–206
amplifiers, relationships, 442, 444–447
BFP420, 443
HBT, 915–918
KGF1608, 355
linear noisy two-port, 392–393
NE71000, 344–348
two-port oscillators, 743
Spectral mask, SA900, 73–75
Spectral regrowth, 90, 103
SPICE noise model, enhanced, 328–329, 332
SPICE parameters, 322–325
BFR193W, 370
diodes, 126
SPICE shot noise model, 910
Splatter, 114
Spur-suppression techniques, 882–890
Stability analysis, power amplifier, 601–611
Stability factors, 381–382
two-port oscillators, 743
Stanford Microdevices, 77
Subharmonically pumped single-balanced mixer, 659, 661
Subharmonic mixing, 674
Substrate flow, MOSFETs, 273–274
Subthreshold conduction, MOSFETs, 271–273
Super low noise pseudomorphic HJ FET, 786–787
Switching FET mixer, simplified, 696–697
Synchronization burst, 28
System noise, 83–88
bit error rate and, 85–86
sensitivity, 84–85
SINAD ratio, 85

Tapped-microstrip resonator, differential oscillator, 753–756
TDA1053, internal circuitry, 151
TDMA:
advantages and disadvantages, 19–20
in GSM, 21–29
burst structures, 23–29
frame and multiframe, 21–23
RF data, 21–22
timers, 22–24
Television receiver, diode switch use, 197
Television tuners, π network, PIN diodes, 151–153
Temperature coefficient of capacitance, testing, as function of reverse voltage, 175, 177
Temperature-compensation circuit, 186–187
Termination-insensitive mixer, 668–669

Testing, 114
abrupt junction, 155–157
acoustic measurements, 115
base-station simulation, 118
breakdown voltage, 180–181
capacitance, 174–177
as function of junction temperature, 175–176
modulating by applied ac voltage, 186
temperature coefficient, 162–164
capacitance ratio, 160–162, 167
capacitances connected in parallel or series, 183–186
comparative, 167
compensating temperature dependence, 186–187
cross-modulation, 168–170, 188–190
current/voltage and capacitance/voltage characteristics, 173–174
cutoff frequency, 179–180
DECT, 118–119
differential forward resistance as function of forward current, 192–193
diode switch, 191–197
distortion products, 168–174
reduction, 170–174
dynamic stability, 187–190
electrical properties, 178–181
equivalent circuit, 174
equivalent shunt resistance, 182
frequency shift, 188
generating tuning voltage, 190–191
GSM, 118
harmonic distortion, 170–171
harmonic generation, 188
hyperabrupt junction, 158–159
intermodulation, 170
IS-95 parameters, 115
leakage current, 180–181
linearly graded junction, 156–158
matching, 181
parallel-resonant circuit, 181–183
physics, 155–160
planar versus mesa construction, 159–160
post-tuning drift, 167–168
Q factor, 163–168, 177–178
series inductance, 178–180
series resistance, 177–178
series resonant frequency, 178–180
silicon versus GaAs, 158–159
slope as function of the reverse voltage, 175–176
temperature coefficient of capacitance, as function of reverse voltage, 175, 177
tracking, 185–186
tuning range, 185
Thermal noise power, 386
Three-reactance matching networks, 570–574
Three-reactance oscillators, 723–728
Three-tone analysis, high-gain amplifiers, 470–471, 473
Time-division duplex transceiver, 63–64
Time-division multiple access, *see* TDMA
Time-domain simulation, RF oscillators, 828–831
Timing advance, 28
T junction, 537
TMOS, 269–270
TOM model, NE71000, 353
Tracking, 185–186, 771
Transceiver:
handheld, block diagram, 3–4
single-chip direct-conversion, 5
Transconductance:
differential amplifier, 522
single-stage feedback amplifiers, 493–494
Transfer characteristic, filter, 863–864
Transfer function, 14–15
time response, 15–16
Transformation equation, 380
Transformation matrix parameters, 400–401
Transformation paths, four-reactance networks, 575–576
Transient response, phase-locked loops, 867–870
Transistor mixers, 678–713
BJT Gilbert cell, 679–682
with feedback, 682–690
CMY210, 699–704
FET mixers, 684, 691–694
GaAsFET single-gate switch, 694–713
MC13143, 685–690
MOSFET Gilbert cell, 693–694
Transistor oscillators, 736–741
Transistors:
equivalent circuit, 399
with lowest noise figure, 783–784
structure types, 198–199
see also specific types of transistors
Transition frequency, 206–208
Transmission line, 534, 536
RF power transistors, 570–571
Transmission quality, 114–117
Transmit local oscillator, 64, 66
Transmitters, 58–77
I/Q modulation, 58–63
I/Q modulator equations, 76–77
system architecture, 63–66
see also SA900
Triple-balanced mixer, 676–677
Triple-beat distortion, 99–100
Tristate comparators, 855–863
Tristate detector, with antibacklash circuit, 862
Tuned filters, voltage-controlled, 513–522
diode performance, 513–516
HF/VHF, 518–521
third-order intercept point, 519–521
VHF

Tuned filters, voltage-controlled (*continued*)
example, 516–518
improving, 521–522
Tuning diodes, 153–197, 765
ac load line, 791, 793–796
capacitance
adding, 794
connected in parallel or series with, 767–768
influence on noise-sideband performance, 791–792
noise influence, 784
in parallel-resonant circuit, 765–767
Tuning range, 185
parallel-resonant circuit, 769, 771
Tuning voltage, generating, 190–191
Turbocharged, 601
Two-differential-amplifier oscillator, phase noise, 821
Two-port microwave/RF oscillator, 741–745
Two-port nonlinear circuit, schematic, 925–926
Two-port oscillator, 728–731
Two-ports:
parallel combination, 409
unconditionally stable, 447
Two-stage amplifiers, 497–507
TXLO inputs, 64, 66

UHF VCO using the tapped-inductor differential oscillator, 753–756
Ultra-high-frequency/super-high-frequency, *see* UHF/SHF
Ultra low power DC-2.4 GHz linear mixer, 685–690
UMA1018M dual-synthesizer chip, 867–870
UPC2710 electrical specifications, 508, 510–511
UPC2749 MMIC, 508

Varactors, 153
Varactor tuning diodes, 513–515
Varicaps, 153
VCO, 66–67
phase-locked loops, 850
phase noise, 774, 777, 831–832
optimization and, 805, 811
schematic, 791, 793
silicon/GaAs-based integrated, 817–822, 825
very-low-phase-noise, 776
VHF filter, 516–518
improving, 521–522
Via holes, 540–541
Viterbi algorithm, 11
VMOS:
cross section, 269–270
Voltage-controlled oscillator, 716, 719
Voltage gain, amplifiers, 445
VSWR, Lo-port, 647, 649

Wilkinson divider/combiners, 549
Wilkinson power dividers, 602
Wireless synthesizers, 848–896
direct digital synthesis, 889, 891–896
hybrid, 893, 896
see also Phase-locked loops

Y junction, 538

Zener diode, 190–191

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

1

INTRODUCTION TO WIRELESS CIRCUIT DESIGN

1-1 OVERVIEW

Wireless circuits are not that different from commonly known two-way radio, television, and broadcast arrangements. Some of them require high linearity in modulation (TV picture); some work via relay stations (two-way radio). The real differences lie in the fact that the cell sizes are much smaller, and that in most cases we attempt multiple channel use (reuse) using time-division multiplex, spread spectrum, or some other efficient means of reducing the bandwidth required for communication. One can argue that the wireless circuits include simple devices, such as garage-door openers and wireless keys for automobiles (we have seen many cases where strong interfering signals prevented the car owners from reclaiming their cars until the interfering signal disappeared). Another longtime favorite is cordless telephones: initially, 50-MHz models with essentially no privacy protection; later, more sophisticated models that operate at 900 MHz; and now, dual-band designs that use 900 MHz and 2.4 GHz.

The largest wireless growth area is probably the cellular telephones. The two major applications are the handsets, commonly referred to as cell phones or occasionally as "handies," and the base stations. The base stations have many more problems with large-signal-handling linearity at high power, although handset users may run into similar problems. An example of this is the waiting area of an airport, where many travelers are trying to conduct last-minute business: In one instance, we concluded that about 30% of all the people present were on the air! It would have been fun to evaluate this receiver-hostile environment with a spectrum analyzer.

From such use comes anxiety factors, the lesser of which is "When will my battery die?"—a spare battery tends to help—and the greater of which the ongoing question, "Will this cell-phone transmitter harm my body?" [22]. A brief comment for the self-proclaimed experts in this area: A 50–100-kW TV transmitter, specifically its video or picture portion, connected to a high-gain antenna, emits levels of energy in line-of-sight paths that by far exceed the pulsed energy from a cell phone. Specifically, the duration of energy is signifi-

cantly smaller, and the absolute energy is more than a thousandfold higher, than the radio frequency (RF) supposedly harming us from the cellular phone. Handheld two-way radios have been used for the last 30 years or so by police and other security interests, operating in the frequency range from 50 to 900 MHz with antennas close to the users' heads, and there are no known cases of cancer or any other illnesses caused by these handheld radios. Recent studies in England, debatably or not, showed that the reaction-time level of people using cell phones actually *increased*—but then there are always the skeptics and politically motivated who ignore the facts, try to influence the media, and have their 15 minutes of fame (as Andy Warhol used to say).

As to the "harmful" radiation, Figure 1-1 shows the simulated radiation of a Motorola flip phone. While there are no absolute values attached to the pattern colors, it is interesting to see that the antenna extension inside the plastic casing also radiates, but most of the energy definitely is emitted by the top of the antenna. It seems to be a good idea to hold the telephone in such a way that the antenna points away from the head, "just in case." The user will find a "warm" sensation that will have more to do with the efficiency of the RF power amplifier heating up the case than the effect of radiation.

With this introduction in place, we will first take a look at a typical ultra-high-frequency/super-high-frequency (UHF/SHF) transceiver and explain the path from the microphone to the antenna and back. After this, we will inspect the radio channel and its effect on various methods of digital modulation. Analysis of wireless receivers and transmitters will be next, followed by a look at available building blocks and how they affect the overall system. To validate proper system operation, a fairly large number of measurements and tests must be performed, and conveying their purpose and importance will necessitate the definition of a number of system characteristics and concepts, such as dynamic range. Finally, after this is done, we will look at the issue of wireless system testing. Again, we intend to

Figure 1-1 Simulated antenna radiation of a Motorola flip phone.

give guidance applicable to battery-operated, handheld operation as well as high-powered base stations.

1-2 SYSTEM FUNCTIONS

A cellular telephone is a hybrid between a double-sideband and frequency-modulated [FM; or phase-modulated (PM)] transceiver. The actual transmission is not continuous but is pulsed, and because of the pulse spectrum there is a signal bandwidth concern due to keying transients, not unlike intermodulation products of a single-sideband (SSB) transceiver cluttering up adjacent channels. The cellular telephone is also a linear transceiver in the sense that its signal-handling circuitry must be sufficiently amplitude- and phase-linear to preserve the modulation characteristics of the AM/PM hybrid emissions it transmits and receives. Containing such an emission's spectral regrowth, which affects operation on adjacent channels, is not unlike the linearity requirements we encounter in SSB transceivers—requirements so stringent that amplifiers must be run nearly in Class A to meet them. The time-division multiple access (TDMA) operating mode, which allows many stations to use the same frequency through the use of short, precisely timed transmissions, requires a system that transmits with a small duty cycle, putting much less thermal stress on a power amplifier than continuous operation. Power management, including a sleep mode, is another important issue in handset design.

Figure 1-2 shows the block diagram of a handheld transceiver. This is applicable for cellular telephones and other systems that allow full duplex. For those not too familiar with transceivers, here is a "walk" through the block diagram. The RF signal intercepted by the antenna is fed through a duplex filter into a front end consisting of a preamplifier, an additional filter, and a mixer. The duplexer is optimized more for separating transmit and receive frequencies than extreme selectivity, but because of the typical low field strengths of incoming signals, it provides enough selectivity to guard the receiver path against overload and intermodulation products. The preamplifier is either a single transistor or a cascode arrangement with a filter following it. These high-band filters, mostly supplied by Murata, are typically surface acoustic wave (SAW) filters with very small dimensions. We would already like to point out in this part of the block diagram that these filters typically have high-impedance inputs and outputs (somewhere between 200 Ω and 1 kΩ), therefore eliminating the nice test-setup possibilities typically provided in a 50-Ω system. Generally, integrated circuit (IC)-type mixers also operate at high impedances, which makes matching easier. The filter following the mixer is responsible for reducing the image, and then we go to the intermediate frequency (IF) and demodulation. The particular chip or chips mentioned here, supplied by Philips, are set out for a double-conversion receiver, and the demodulation is accomplished with a quadrature detector for FM analog modulation. The rest of the circuitry on the horizontal path does digital signal processing (DSP) and overall control functions. The four blocks at the far right refer to the central processor, which handles such things as display, power management, and information storage (such as frequently used telephone numbers). A nice overview about DSP in "readable" form is given by Kostic [1].

The transmit portion consists of an independent synthesizer that is modulated. There are dual synthesizer chips available to accommodate this. Both receive and transmit frequencies are controlled by a miniature temperature-compensated crystal oscillator (TCXO). One of its outputs is also used as the system master clock for all the digital activities. The output of the voltage-controlled oscillator (VCO) is then amplified and fed to the antenna through the

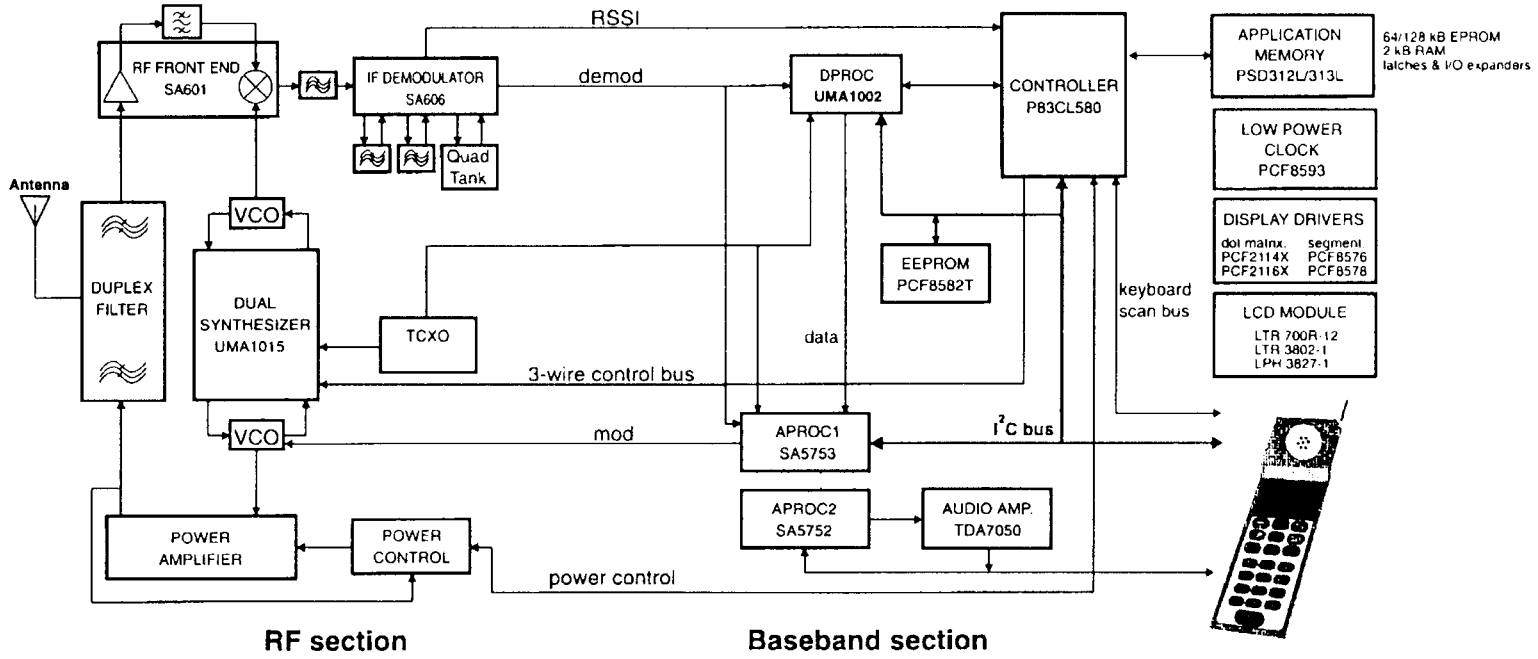


Figure 1-2 Block diagram of a handheld cellular telephone transceiver.

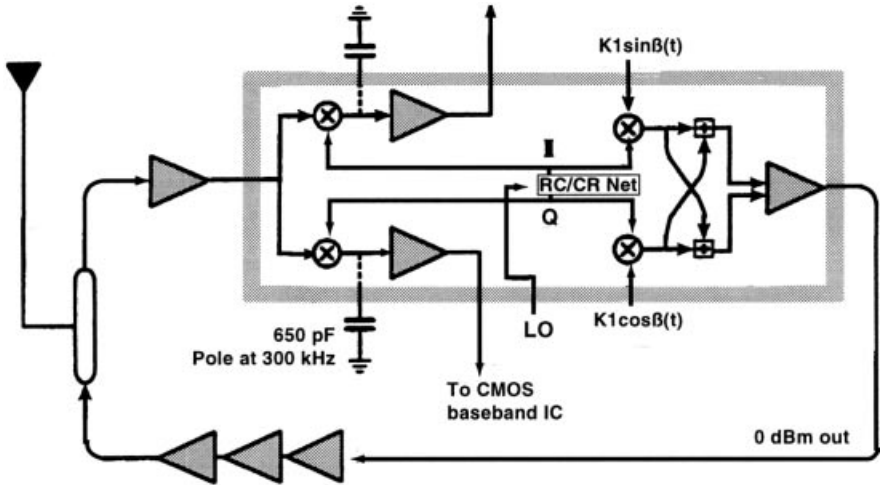


Figure 1-3 Single-chip direct-conversion transceiver by Alcatel. Channel selection is accomplished at baseband by low-pass switched-capacitor filters in a companion mixed-signal complementary metal-oxide semiconductor (CMOS) IC. A trimmed resistance-capacitance/capacitance-resistance (RC/CR) network generates the necessary quadrature signals for the chip's mixers.

same duplex filter as the receive portion. There are also schemes available for advanced modulation methods, specifically, code- , frequency-, and time-division multiple access (CDMA, FDMA, and TDMA, respectively). In these cases, the transmitter is not active all the time, and the duplexer can be replaced with a diode switch using a quarter-wavelength transmission line together with a PIN diode for the required switching.

Many modern devices use "zero IF" or direct conversion, which simplifies the IF or modulation portion of the unit significantly. Figure 1-3 shows an Alcatel single-chip direct-conversion transceiver. The signal is fed to an image-reject mixer with the local oscillator (LO) in quadrature, and the selectivity is obtained by manipulating the "audio bandwidth." Today we have a large number of implementations using different schemes that are beyond the scope of this book; therefore, we have decided to limit ourselves to a basic introduction because most of the relevant demodulation and coding are done in DSP, for which we will give appropriate references. A nice overview of different architectures is found in Razavi [2].

1-3 THE RADIO CHANNEL AND MODULATION REQUIREMENTS

1-3-1 Introduction

The transmission of information from a fixed station to a mobile is considerably influenced by the characteristics of the radio channel. The RF signal not only arrives at the receiving antenna on the direct path but is normally reflected by natural and artificial obstacles in its way. Consequently the signal arrives at the receiver several times in the form of echoes, which are superimposed on the direct signal (Figure 1-4). This superposition may be an advantage as the energy received in this case is greater than in single-path reception. This feature is made use of in the digital audio broadcasting (DAB) single-frequency network. However,

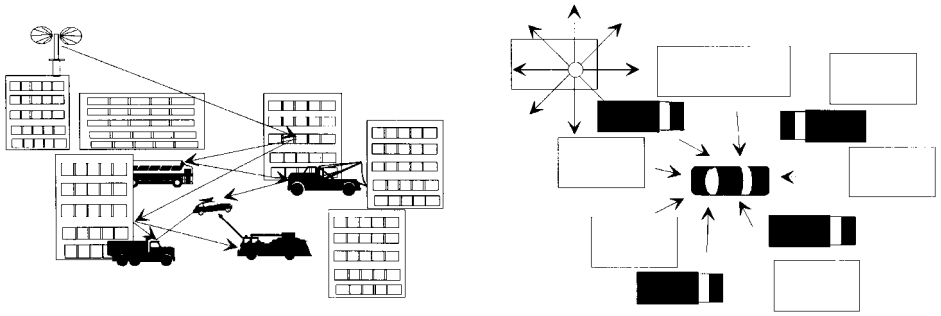


Figure 1-4 Mobile receiver affected by fading.

this characteristic may be a disadvantage when the different waves cancel each other under unfavorable phase conditions. In conventional car radio reception this effect is known as fading. It is particularly annoying when the vehicle stops in an area where the field strength is reduced because of fading (e.g., at traffic lights). Additional difficulties arise when digital signals are transmitted. If strong echo signals (compared to the directly received signal) arrive at the receiver with a delay on the order of a symbol period or more, time-adjacent symbols interfere with each other. In addition, the receive frequency may be falsified at high vehicle speeds because of the Doppler effect so that the receiver may have problems in estimating the instantaneous phase in the case of angle-modulated carriers. Both effects lead to a high symbol error rate even if the field strength is sufficiently high. Radio broadcasting systems using conventional frequency modulation are hardly affected by these interfering effects. If an analog system is replaced by a digital one that is expected to offer advantages over the previous system, it has to be ensured that these advantages—for example, better audiofrequency signal/noise (AF S/N) and the possibility of supplementary services for the subscriber—are not at the expense of reception in hilly terrain or at high vehicle speeds because of extreme fading.

For this reason a modulation method combined with suitable error protection has to be found for mobile reception in a typical radio channel, which is immune to fading, echo, and Doppler effects.

With a view to this, more detailed information on the radio channel is required. The channel can be described by means of a model. In the worst case, which may be the case for reception in built-up areas, it can be assumed that the mobile receives the signal on several indirect paths but not on a direct one. The signals are reflected, for example, by large buildings; the resulting signal delays are relatively long. In the vicinity of the receiver these paths are split up into a great number of subpaths; the delays of these signals are relatively short. These signals may again be reflected by buildings but also by other vehicles or natural obstacles like trees. Assuming the subpaths are statistically independent of each other, the superimposed signals at the antenna input cause considerable time- and position-dependent field-strength variations with an amplitude obeying the Rayleigh distribution (Figures 1-5 and 1-6).

If a direct path is received in addition, the distribution changes to the Rice distribution, and finally, when the direct path becomes dominant, the distribution follows the Gaussian distribution with the field strength of the direct path being used as the center value.

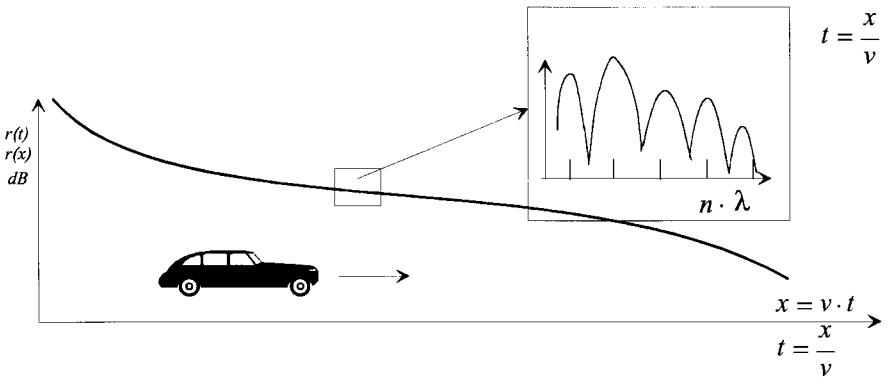


Figure 1-5 Receive signal as a function of time or position.

In a Rayleigh channel the bit error rate (BER) increases dramatically compared to the BER in an additive white Gaussian noise (AWGN) channel (Figure 1-7).

1-3-2 Channel Impulse Response

This scenario can be demonstrated by means of the channel impulse response. Let's assume that a very short pulse of extremely high amplitude [in the ideal case, a Dirac pulse $\delta(t)$] is sent by the transmitting antenna at a time $t_0 = 0$. This pulse arrives at the receiving antenna direct and in the form of reflections with different delays τ_i and different amplitudes because of path losses. The impulse response of the radio channel is the sum of all received pulses (Figure 1-8). Since the mobile receiver and also some of the reflecting objects are moving, the channel impulse response is a function of time and of delays τ_i; that is, it corresponds to

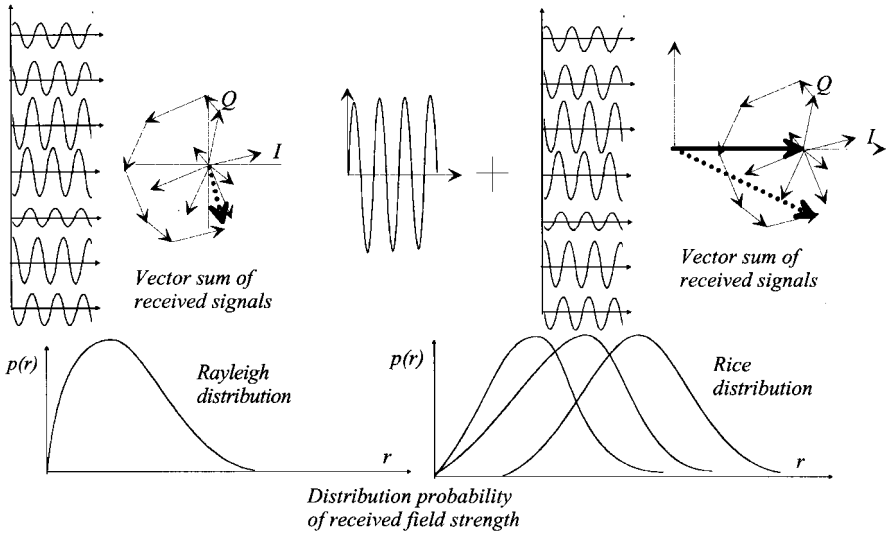


Figure 1-6 Rayleigh and Rice distributions.

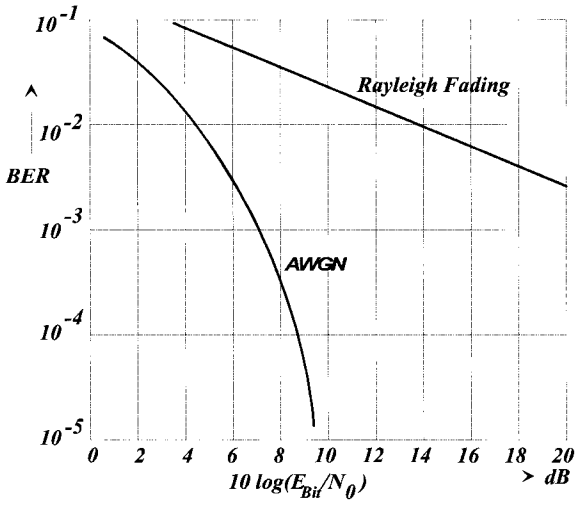


Figure 1-7 BER in a Rayleigh channel.

$$h(t,\tau) = \sum_{N} a_i \delta(t - \tau_i) \tag{1-1}$$

This shows that delta functions sent at different times t cause different reactions in the radio channel.

In many experimental investigations different landscape models with typical echo profiles were created. The most important are:

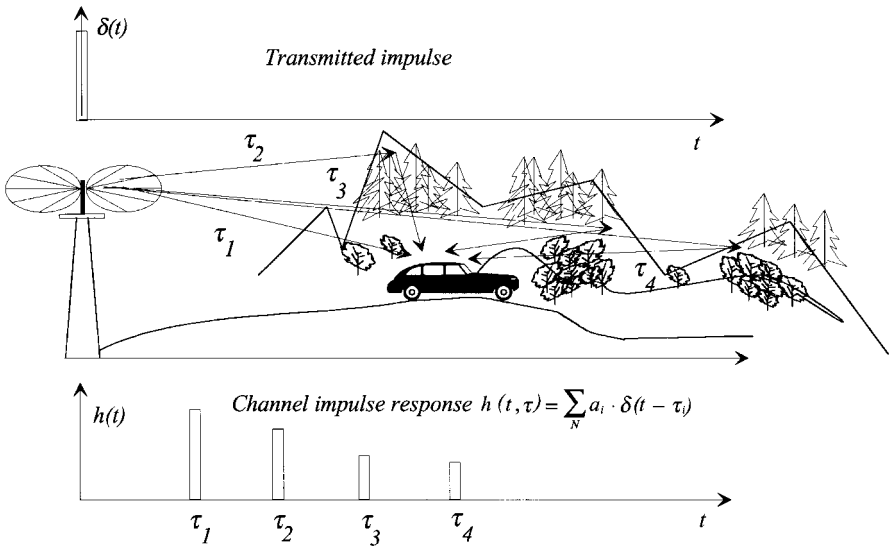


Figure 1-8 Channel impulse response.

- Rural area (RA)
- Typical urban area (TU)
- Bad urban area (BA)
- Hilly terrain (HT)

The channel impulse response informs on how the received power is distributed to the individual echoes. A parameter, the "delay spread," can be calculated from the channel impulse response, permitting an approximate description of typical landscape models (Figure 1-9).

The delay spread also roughly informs on the modulation parameters, carrier frequency, symbol period, and duration of guard interval, which have to be selected in relation to each other. If the receiver is located in an area with a high delay spread (e.g., in hilly terrain), echoes of the symbols sent at different times are superimposed when broadband modulation methods with a short symbol period are used. In the case of DAB, this problem is aggravated by the use of single-frequency networks. An adjacent transmitter emitting the same information on the same frequency has the effect of an artificial echo (Figure 1-10).

A constructive superposition of echoes is only possible if the symbol period is much greater than the delay spread. The following holds:

$$T_s > 10T_d \tag{1-2}$$

This has the consequence that relatively narrowband modulation methods have to be used. If this is not possible, channel equalizing is required.

For channel equalizing, a continuous estimation of the radio channel is necessary. The estimation is performed with the aid of a periodic transmission of data known to the receiver. In networks according to the Groupe Speciale Mobile (GSM) standards a midamble consisting of 26 bits—the training sequence—is transmitted with every burst. The training

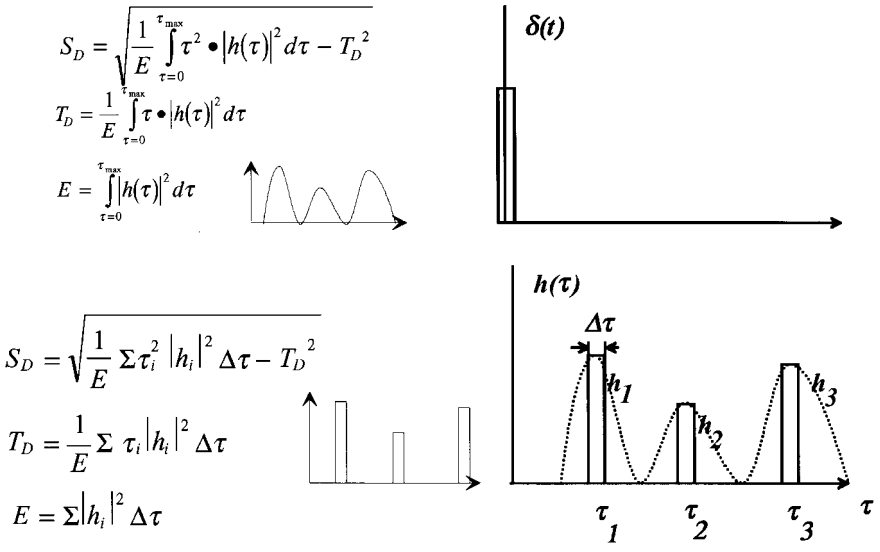


Figure 1-9 Calculation of delay spread.

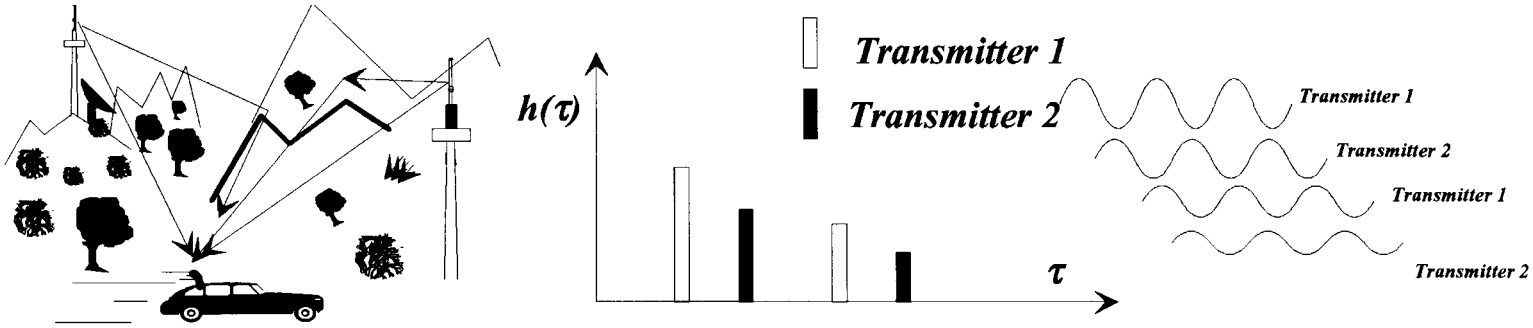


Figure 1-10 Artificial and natural echoes in the single-frequency network.

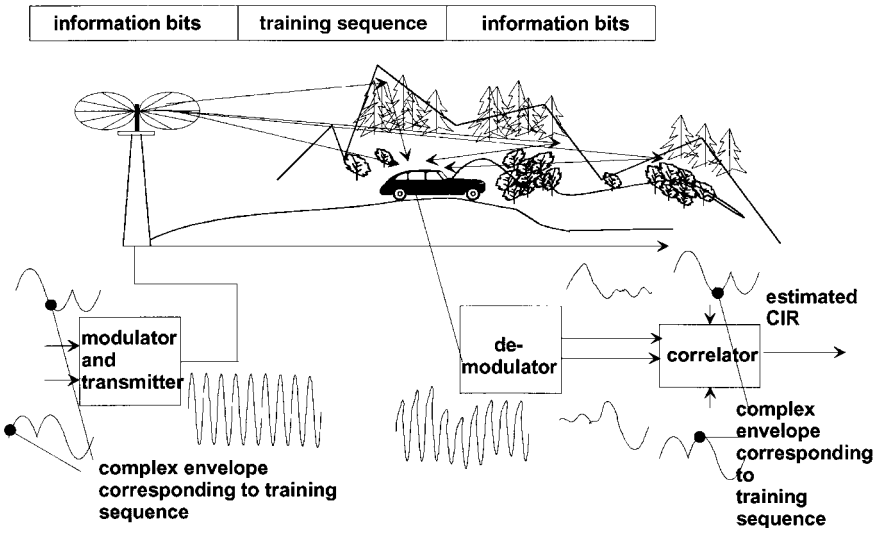


Figure 1-11 Channel estimation.

sequence corresponds to a characteristic pattern of I/Q signals that is kept in a memory in the receiver. The baseband signals of every received training sequence are correlated with the stored ones. From this correlation the channel can be estimated, and the properties of the estimated channel will then be fed to the equalizer (Figure 1-11).

The equalizer uses the Viterbi algorithm (maximum sequence likelihood estimation) for the estimation of the phases, which most likely have been sent at the sampling times. From these phases the information bits are calculated (Figure 1-12). A well-designed equalizer then will superimpose the energies of the single echoes constructively, so that the result in an area where the echoes are not too much delayed—delay times up to 16 μs have to be tolerated by a receiver—is better than in an area with no significant echoes (Figure 1-13).

Remaining bit errors are eliminated using another Viterbi decoder at the transmitter for the convolutionally encoded data sequences.

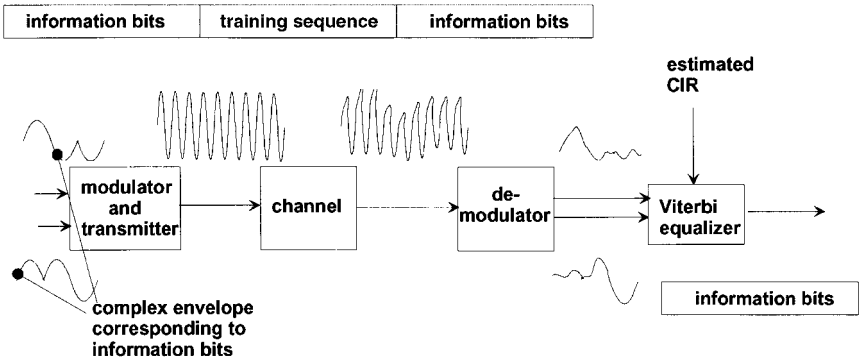


Figure 1-12 Channel equalization.

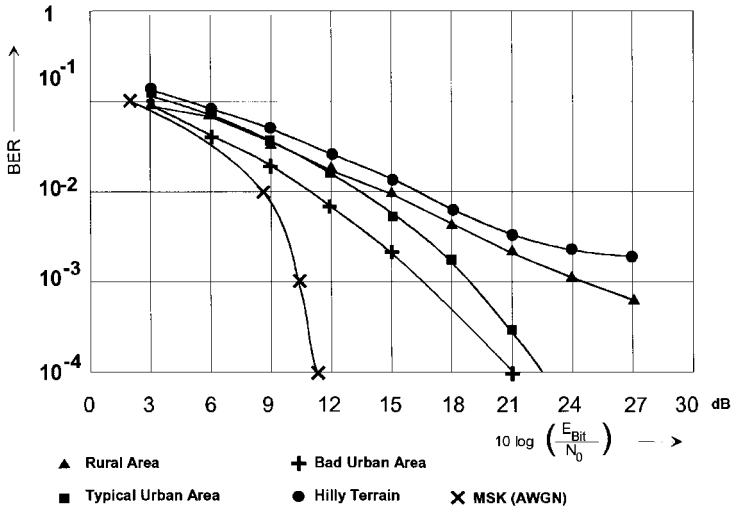


Figure 1-13 BERs after the channel equalizer in different areas.

The ability of a mobile receiver to work in a hostile environment, such as the radio channel with echoes, must be proved. The test is performed with the aid of a fading simulator. The fading simulator simulates different scenarios with different delay times and different Doppler profiles. A signal generator generates undistorted I/Q modulated RF signals, which are downconverted into the baseband. Here the I/Q signals are digitized and split into different channels, where they are delayed and attenuated, and where Doppler effects are superimposed. After combination of these distorted signals at the output of the baseband section of the simulator these signals modulate the RF carrier, which is the test signal for the receiver under test (Figure 1-14).

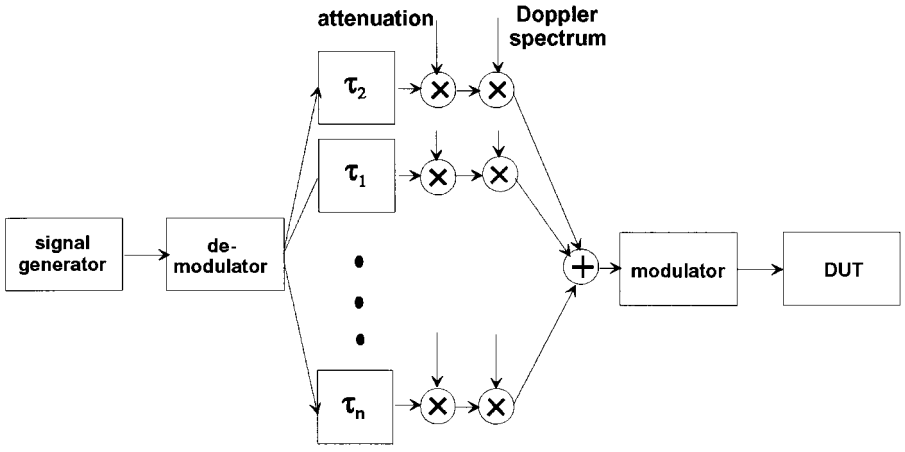


Figure 1-14 Fading simulator.

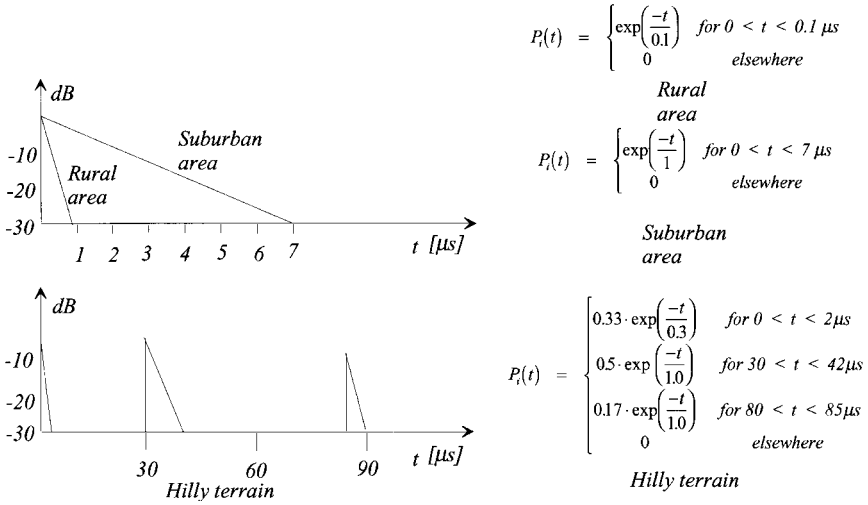


Figure 1-15 Typical landscape profiles.

To make the tests comparable, the Groupe Speciale Mobile (GSM) recommends typical profiles, for example,

- Rural area (RAx)
- Typical urban (TUx)
- Hilly terrain (HTx)

where number and strengths of the echoes and the Doppler spectra are prescribed (Figure 1-15).

1-3-3 Doppler Effect

Since the mobile receiver and some of the reflecting objects are in motion, the receive frequency is shifted because of the Doppler effect. In the case of single-path reception, this shift is calculated as follows:

$$f_d = \frac{v}{c} f_c \cos \alpha \tag{1-3}$$

where v = speed of vehicle
c = speed of light
f = carrier frequency
α = angle between v and the line connecting transmitter and receiver

In the case of multipath reception the signals on the individual paths arrive at the receiving antenna with different Doppler shifts because of the different angles α_i, and the receive spectrum is spread. Assuming an equal distribution of the angles of incidence, the power density spectrum can be calculated as follows:

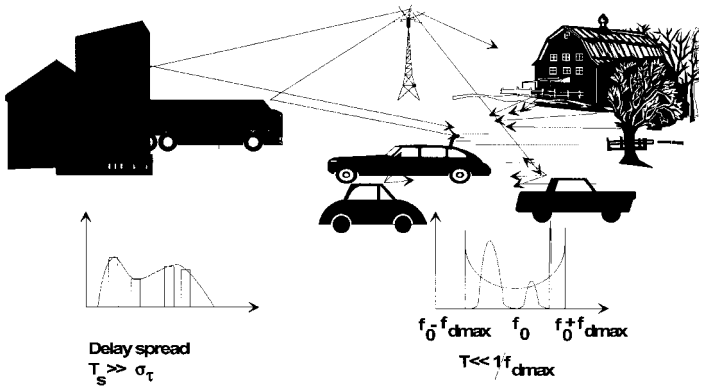


Figure 1-16 Doppler spread.

$$P(f) = \frac{1}{\pi} \frac{1}{\sqrt{f_d^2 - f^2}} \quad \text{for } |f| < |f_d| \tag{1-4}$$

where f_d = maximum Doppler frequency

Of course, other Doppler spectra are possible in addition to the pure Doppler shift described above: for example, spectra with a Gaussian distribution using one or several maxima. A Doppler spread can be calculated from the Doppler spectrum analogously to the delay spread (Figure 1-16).

1-3-4 Transfer Function

The fast Fourier transform (FFT) value of the channel impulse response is the transfer function $H(f, t)$ of the radio channel, which is also time dependent. The transfer function describes the attenuation of frequencies in the transmission channel. When examining the frequency dependence it will be evident that the influence of the transmission channel on two sine-wave signals of different frequencies becomes greater with increasing frequency difference. This behavior can adequately be described by the coherence bandwidth, which is approximately equal to the reciprocal delay spread; that is,

$$(\Delta f)_c = \frac{1}{T_d} \tag{1-5}$$

If the coherence bandwidth is sufficiently wide and, consequently, the associated delay spread is small, the channel is not frequency selective. This means that all frequencies are subject to the same fading. If the coherence bandwidth is narrow and the associated delay spread wide, even very close adjacent frequencies are attenuated differently by the channel. The effect on a broadband-modulated carrier with respect to the coherence bandwidth is obvious. The sidebands important for the transmitted information are attenuated to a different degree. The result is a considerable distortion of the receive signal combined with a high bit error rate even if the received field strength is high. This characteristic of the radio channel again speaks for the use of narrowband modulation methods (Figure 1-17).

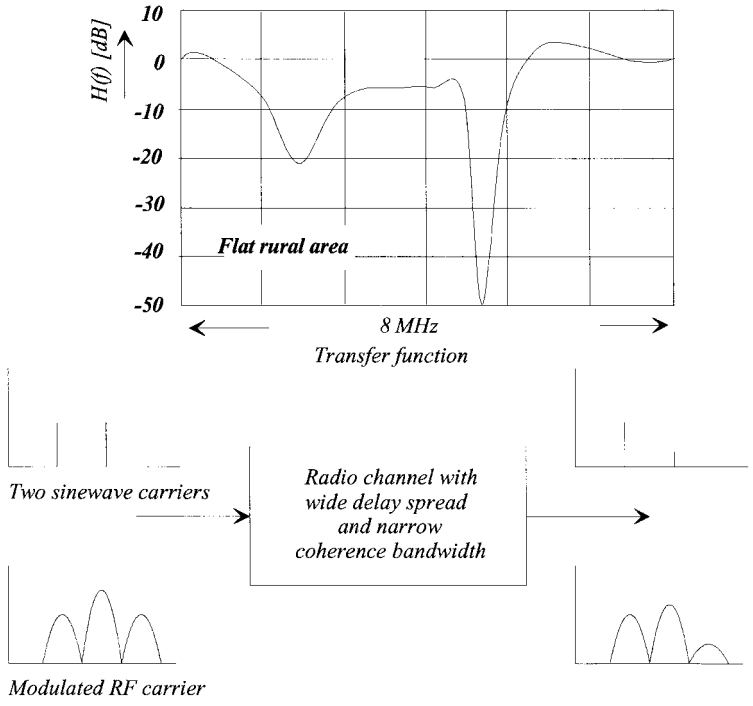


Figure 1-17 Effect of transfer function on modulated RF signals.

1-3-5 Time Response of Channel Impulse Response and Transfer Function

The time response of the radio channel can be derived from the Doppler spread. It is assumed that the channel rapidly varies at high vehicle speeds. The time variation of the radio channel can be described by a figure, the coherence time, which is analogous to the coherence bandwidth. This calculated value is the reciprocal bandwidth of the Doppler spectrum. A wide Doppler spectrum therefore indicates that the channel impulse response and the transfer function vary rapidly with time (Figure 1-18). If the Doppler spread is reduced to a single

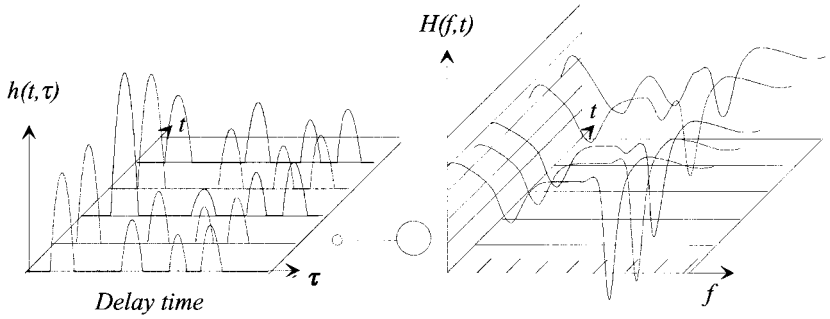


Figure 1-18 Channel impulse response and transfer function as a function of time.

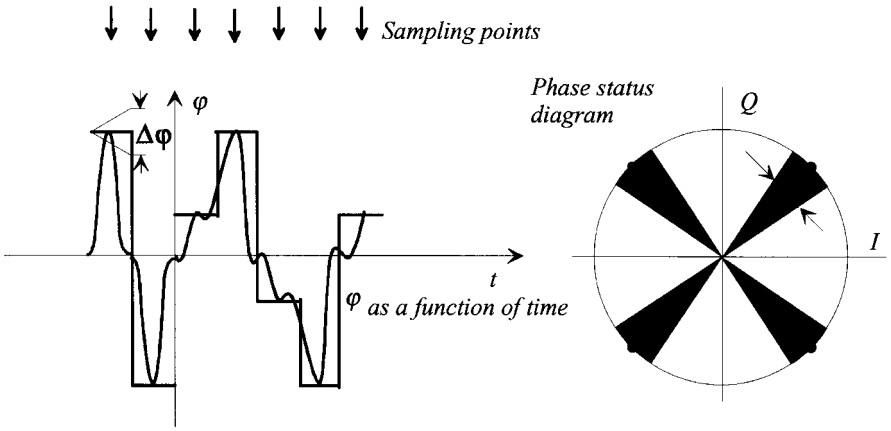


Figure 1-19 Phase uncertainty caused by Doppler effect.

line, the channel is time invariant. In other words, if the vehicle has stopped or moves at a constant speed in a terrain without reflecting objects, the channel impulse response and the transfer function measured at different times are the same.

The effect on information transmission will be illustrated in an example. In the case of *M*-ary phase shift keying (MPSK) modulation using hard keying, the transmitter holds the carrier phase for a certain period of time; that is, for the symbol period *T*. In the case of soft keying with low-pass-filtered baseband signals for limiting the modulated RF carrier, the nominal phase is reached at a specific time, the sampling time. In both cases the phase error $\varphi_f = f_d T_S$ is superimposed onto the nominal phase angle, which yields a phase uncertainty of $\Delta\varphi = 2\varphi_f$ at the receiver. The longer the symbol period the greater the angle deviation (Figure 1-19). Considering this characteristic of the transmission channel, a short symbol period of $T_S << (\Delta t)_c$ should be used. However, this requires broadband modulation methods.

Figure 1-20 shows the field strength or power arriving at the mobile receiver if the vehicle moves in a Rayleigh distribution channel. Since the phase depends on the vehicle position,

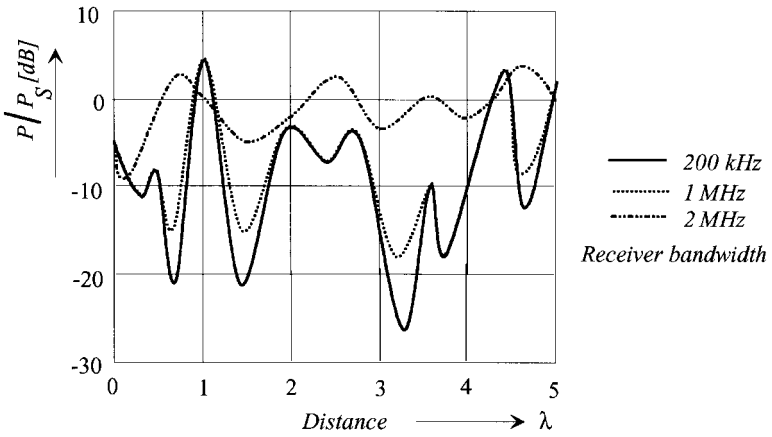


Figure 1-20 Effect of bandwidth on fading.

the receiver moves through positions of considerably differing field strength at different times (time dependence of radio channel). In the case of frequency-selective channels this applies to one frequency only; that is, to a receiver using a narrowband IF filter for narrowband emissions. As Figure 1-20 shows, this effect can be reduced by increasing the bandwidth of the emitted signal and consequently the receiver bandwidth.

1-3-6 Lessons Learned

The strongly frequency-selective channel causing inadmissible distortion of the broadband-modulated carrier and channel impulse responses like those expected in a hilly terrain speak in favor of narrowband modulation methods with long symbol periods. In hilly terrain extensively delayed echoes cause intersymbol interference when broadband modulation with short symbol periods is used. On the other hand, narrowband modulation has the disadvantage that the signals arrive at the receiver considerably attenuated and reception may be interrupted for an indefinite period of time. The burst errors occurring in digital information transmission cannot be corrected even with the most elaborate error protection methods.

These transmission interruptions can be avoided by using broadband modulation methods, but these are sensitive to strongly frequency-selective channels. Since broadband modulation is obtained through the use of short symbol periods, broadband modulation is unsuitable when greatly delayed echoes are expected.

It remains to be defined when a signal is considered a narrowband and when a broadband signal. This question shall be answered with the aid of an example. Apart from extremely narrowband analog modulation methods—as are used for sound broadcasting in the long-wave, mediumwave, and shortwave bands—FM sound broadcasting transmissions in the VHF bands are narrowband. In the case of digital modulation this means that transmissions with a rate of 400 kbps modulated onto a carrier with a bandwidth efficiency of 1.5 (bps)Hz over a bandwidth of approximately 300 kHz can be regarded as narrowband transmissions according to the definition above. Consequently, a DAB signal with this gross bit rate would be a narrowband signal and suitable for transmission on the radio channel with restrictions only. Based on the experience with conventional FM broadcasting systems in large cities and hilly terrain, this was obvious from the very beginning.

Consequently, it is necessary to find ways for spreading the band artificially without reducing the bandwidth efficiency. This means that a large band must be available for the transmission of several programs, the full bandwidth being used by all the programs without mutual interference.

Several approaches can be adopted to tackle the problem (Figure 1-21). One way would be a continuous change of the transmit and receive frequency according to a defined pattern (frequency hopping). This method is used in mobile radio, for instance, but only marginal investigations have been made in this respect for DAB.

Another possibility is to multiply the symbols of the individual programs with digital signals (pseudo-noise function) using a much higher bit rate so that a higher symbol rate is obtained. In this case the different programs are assigned different functions, which must be orthogonal to each other (code-division multiple access, CDMA). The “chopped” bit streams of the individual programs are modulated onto carriers of identical frequency and the modulated carriers are added. A correlation receiver knowing the pseudo-noise function divides the incoming CDMA signal into the individual programs. The disadvantage is obvious. Symbol periods are very short and elaborate means will be required for compensating the intersymbol interference.

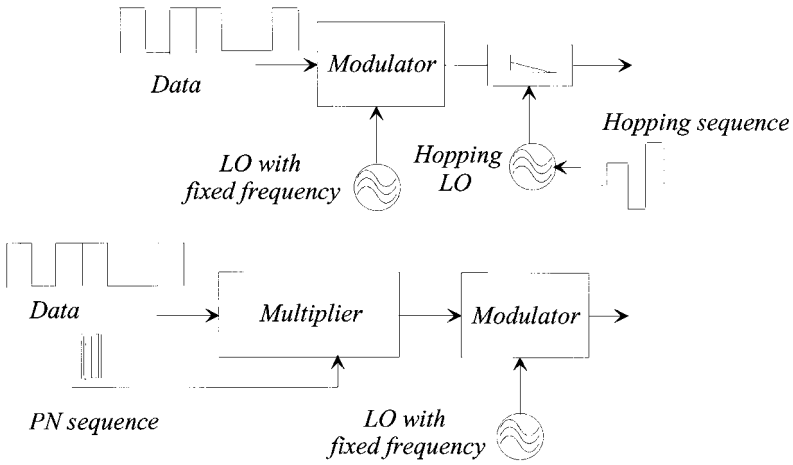


Figure 1-21 Band spreading by frequency hopping and CDMA.

A different approach has been chosen for DAB, which does not involve continuous and elaborate channel measurements, so that it would be possible to use favorably priced receivers as are demanded in the field of consumer electronics. The method used for DAB is a multicarrier method, where the information to be transmitted is spread onto many carriers using time and frequency interleaving. The terms time and frequency interleaving will be explained in the course of this discussion. The result is a broadband transmission method with long symbol periods. However, certain limitations caused by the Doppler effect will have to be accepted particularly at high carrier frequencies.

1-3-7 Wireless Signal Example: The TDMA System in GSM

Frequency-Division Multiple Access (FDMA). In analog radio systems the trend has always been toward a more efficient utilization of the available frequency spectrum by reducing the channel spacing. The number of radio channels obtained at a channel spacing of 12.5 kHz is of course twice that obtained at 25 kHz. However, any improvement brings about its disadvantage: the narrower the channel spacing, the higher the required frequency accuracy and the lower the possible maximum deviation of the frequency modulation. The latter leads to a poorer transmission quality due to the lower S/N ratio. Furthermore, the gaps between the channels, which must be a number of kilohertz wide for safety reasons, also reduce the available system bandwidth (see Figures 1-22 and 1-23).

The use of an available system spectrum divided into individual frequency channels enables the user to simultaneously access a multitude of different frequencies. This multiple access is called *frequency-division multiple access* (FDMA). Consequently, all radio systems with a spectrum divided into channels are FDMA systems. At present, the technically useful limit is reached with a channel spacing of 10–12.5 kHz.

Advantages of FDMA

- Simultaneous access to a given bandwidth by many subscribers
- Increase in the number of channels through reduction of channel spacing

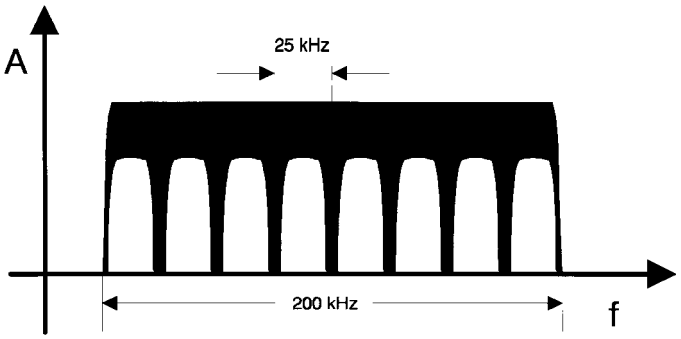


Figure 1-22 Channel spacing in broadband/narrowband systems.

Disadvantages of FDMA

- Higher frequency accuracy required
- Transmission quality decreasing with reduction of channel bandwidth
- Better rejection filters required
- One transmitter/receiver required per channel

Time-Division Multiple Access (TDMA). With TDMA systems, the available bandwidth is divided into considerably fewer, and therefore wider, channels than in FDMA systems. Each of these channels is available to several subscribers quasi-simultaneously. (See Figure 1-24.) However, a given subscriber can use the whole channel for a very short period (timeslot) only; for the rest of the time, no access is available. This serial access of several users is repeated within a fixed time frame.

Advantages of TDMA

- Simultaneous use of a specific bandwidth by a great number of subscribers
- Depending on the number of available timeslots, several subscribers can be served by one transmitter/receiver

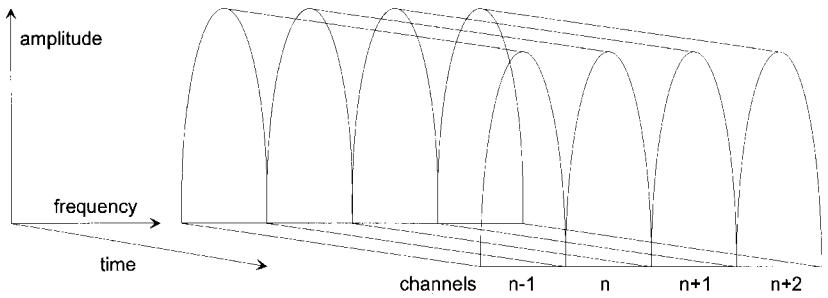


Figure 1-23 Frequency-division multiple access (FDMA).

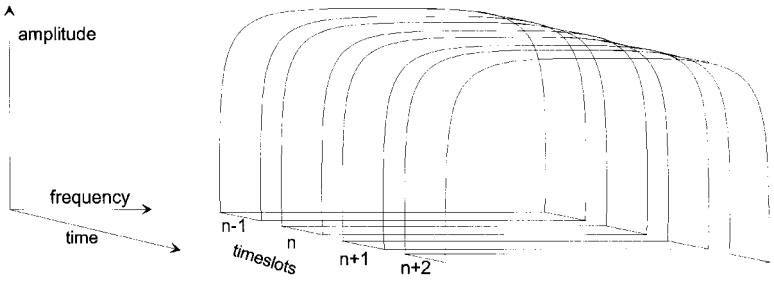


Figure 1-24 Time-division multiple access (TDMA).

- Transmitter and receiver are not permanently switched on (saves battery power)
- The RF section may carry out other tasks in the intervals between transmission and reception
- Reduced susceptibility to frequency-selective fading in the case of larger channel bandwidths

Disadvantages of TDMA

- Accurate time synchronization of subscribers required
- Higher processor capacity required
- Broadband modulators required

Code-Division Multiple Access (CDMA). The increasing use of low-priced and powerful signal processors allows a less common technique of multiple access to be employed in mass communication systems. In the case of *code-division multiple access* (CDMA), the whole system bandwidth is available to all subscribers at any time; that is, all send and receive simultaneously, with each using a specific code (Figure 1-25).

Logic "1" represents a certain bit sequence; logic "0" is the inversion of this sequence. The different signals are distinguished in the receiver by means of a cross-correlation of the received signal, which comprises a great number of codes, with the bit sequence expected so that the desired transmission signal can be detected.

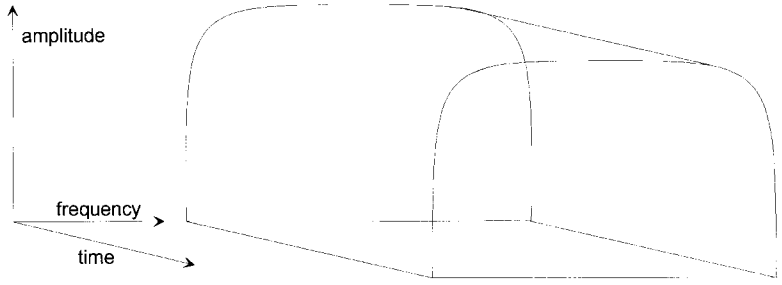


Figure 1-25 Code-division multiple access (CDMA).

Advantages of CDMA

- Simultaneous use of a specific channel or subband by many subscribers
- Several signals can be received simultaneously by one receiver
- Reduced susceptibility to frequency-selective fading in the case of large channel bandwidths
- More subscribers can be served
- Reduced costs for radio network planning

Disadvantages of CDMA

- Accurate time synchronization of subscribers required
- Fast transmitter power control over a wide dynamic range
- No experience with use in mass communication

TDMA in GSM

RF Data. In spite of the competition with other mobile radio systems, a common frequency band could be defined for GSM worldwide. All operators who signed the GSM memorandum of understanding committed themselves to install their GSM systems within the standardized frequency range. The competition for frequencies mainly affects countries using NMT900, the frequency range of which corresponds to the GSM P band. TACS also partly overlaps the GSM P band; the G1 band is completely within the TACS range. Cordless telephones operating in accordance with the CT1 standard also use the upper end of the GSM P band. CT1+ telephones, which had been assigned a frequency range below the P band years ago to protect them against GSM, have now been ousted by the G1 band. See Table 1-1.

Table 1-1 RF data for GSM900 and GSM1800

Variable	GSM900		GSM1800
Frequency range	P band	G1 band	
Uplink (MHz)			
(MS transmitting)	890–915	880–890	1710–1785
Downlink (MHz)			
(BTS transmitting)	935–960	925–935	1805–1880
Duplex spacing (MHz)	45	45	95
Spectrum (MHz)	2×25	2×10	2×75
Frequency channels	124	49	374
Channel numbers (ARFCN)	1–124	975–1023	512–885
Channel spacing	200 kHz		
Modulation	GMSK with BxT = 0.3		
Data transmission rate	270.833 kbps		
Bit duration	3.69 µs		

Since each frequency channel is divided into eight timeslots, transmitter and receiver operate in an intermittent mode, with the receive and transmit time in the upper and lower channels shifted by three timeslots (Figure 1-26). Although this alternative sending and receiving scheme operates in half-duplex rather than full-duplex operation, the received signal sounds continuous to the user.

TDMA Structure

FRAME AND MULTIFRAME. All GSM radio channels are organized in frames of approximately 4.62-ms duration. The frames are continuously repeated. Each frame is divided into 8 timeslots of approximately 577 μs each. A timeslot contains an information packet, the burst. The 26-type multiframes are used on all timeslots containing a traffic channel (voice and/or data); the 51-type multiframes on all timeslots are reserved for control channels. See Figure 1-27.

The TDMA structure uses other frame types above the multiframe level, as shown in Table 1-2.

TDMA TIMERS. The frame number within the hyperframe is counted continually so that counting of the TDMA clock restarts after approximately 3.5 hours. The frame number therefore represents a time unit in the GSM system. Similar to time counting, where the seconds are combined into minutes, hours, and days, GSM does not count the absolute frame numbers but uses timers instead. These timers are structured as shown in Table 1-3.

The absolute frame number is obtained by a multiplication of the three timers. However, on certain occasions a short version of the timers is used.

Burst Structures. Information between base station and mobile is sent in the timeslots. In each slot a certain amount of information—that is, a burst—can be transmitted. Normally the timeslot is occupied by the normal burst (Figure 1-28), which is used for signaling as well as for voice and data transmission.

Each part of the burst serves a specific purpose as described below.

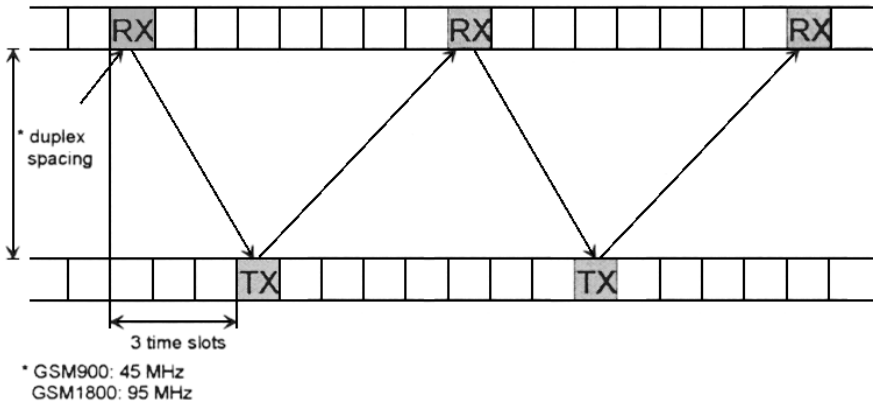


Figure 1-26 Duplex spacing of transmission and reception.

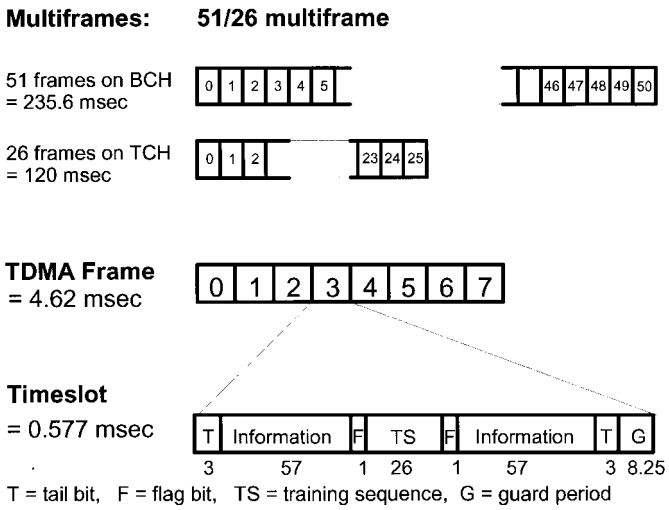


Figure 1-27 Timeslot, frame, and multiframe.

Information Bits. The normal burst is able to transmit 2×57 information bits. Since this information can be replaced every 4.62 ms during an ongoing call, the average theoretical transmission rate is

$$114 \text{ bits} \div 4.62 \text{ ms} \approx 24.7 \text{ kbps} \tag{1-6}$$

At the same time, this rate also represents the maximum transmission rate that can be obtained in the GSM system using one timeslot per transmission per time frame. Consequently, the transmission rate could be increased only if more than one timeslot is used for the transmission.

The bit rate is much lower in the control channels; that is, the above transmission rate is only attained by the mobile station if a traffic channel has been set up. In this case, the base station and the mobile station use a signaling channel, which also uses up capacity, in addition to the voice or data channel. Table 1-4 shows the assignment of the theoretically available capacity with a traffic channel setup.

Table 1-2 Superframes and hyperframes

Superframe	= 51 × 26 frames
	= 1326 frames
	= 6.12 s
Hyperframe	= 2048 frames
	= 2.715.648 frames
	= 3 h, 29 min, 3.5 s

Table 1-3 TDMA timers

T1:	= FN div. (26 × 51)	Range:	0–2047
T2:	= FN mod 26	Range:	0–25
T3:	= FN mod 51	Range:	0–50
FN (Frame Number)		Range:	0 2715647
FN_{max}	$= 51 \times 26 \times 2048 - 1$		

Training Sequence. In the middle of the normal burst, a 26-bit training sequence, the bit sequence of which is known to the receiver, is sent. The *training-sequence code* (TSC) can be one of eight different sequences. These sequences are stored in all receivers, and at the beginning of a transmission, the base transceiver station (BTS) decides on the TSC to be used. The training sequence serves two main purposes: bit synchronization and estimation of channel impulse response.

BIT SYNCHRONIZATION. Data transmitted via the air interface are in the asynchronous mode; that is, the receiver has to regenerate the bit clock from the data stream. To enable synchronization in the receiver, the transmitter adds synchronization bits to the information stream. Therefore, in normal data transmission, data telegrams start with a "...10101010..." sequence so that the receiver can regenerate the bit clock. A predefined bit word informs the receiver when the actual information (block synchronization) starts. A receiver synchronized in this way is able to decode the data stream online. The training sequence of the burst has to assume both synchronization tasks. Since it is in the middle of the burst, direct decoding is not possible.

Each burst must first be stored in the receiver and then decoded by postprocessing. The reason for using this method is the second task of the training sequence. Synchronization itself is carried out by means of cross-correlation; that is, the expected training sequence is compared (correlated) to the center of the received burst and to the beginning and end of the training sequence so that the bit clock is also known. A burst containing other than the expected training sequence cannot be synchronized and decoded.

COMPENSATION OF MULTIPATH RECEPTION. The signal from the transmitter (in Figure 1-29, BTS → MS; the same applies also in the opposite direction) arrives at the receiver not only along the direct path but also via various other paths as a result of reflection and diffraction caused by obstacles in the signal path.

Propagation conditions on these additional paths differ from those on the direct path. For instance, we can expect signals traveling via additional paths to exhibit:

- Longer travel times because of increased path length

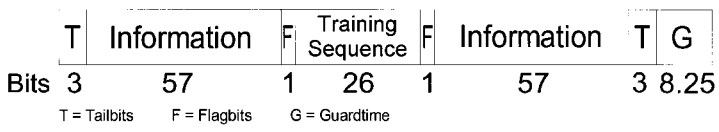


Figure 1-28 Normal burst.

Table 1-4 Transmission bit rates

	Information	Error Protection	Total
Traffic channel			22.8 kbps
Voice (full rate)	13.0 kbps	9.8 kbps	
Data	2.4 kbps	20.4 kbps	
	4.8 kbps	18.0 kbps	
	9.6 kbps	13.2 kbps	
Control channels			0.95 kbps
Idle frame			0.95 kbps
Total			**24.7 kbps**

- Various strengths
- Different Doppler shifts

Because of the different travel times, the signals arrive with a different phase at the receiving antenna. Depending on this phase, components may be canceled—that is, they may totally disappear—or added so that a high-quality signal is received for only a short period of time. RF level variations are statistically distributed; level shifts due to fading may be as great as 40 dB.

In addition to RF level fading, another annoying effect is encountered that, uncompensated, would make correct signal decoding rather difficult. Because of the additional distance the signal travels via the indirect path, the signal arriving at the receiving antenna exhibits time delay of its modulation in addition to variable phase shifts. The total of all channel responses to a single transmitted pulse is called *channel impulse response* (CIR, Figure 1-30). If the indirect path is only 1 kilometer longer, the GSM echo bit reaches the receiver later than the directly received bit and thus interferes with the next bit received. This *intersymbol interference* (ISI) may occur over several bits in succession. With delays of up to 15 µs,

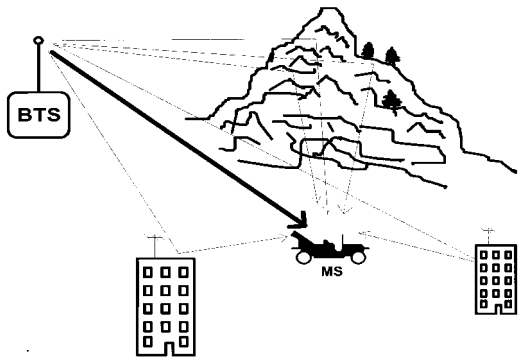


Figure 1-29 Multipath reception due to reflection and diffraction. The *base transceiver station* (BTS) is transmitting to the *mobile station* (MS).

differentiating the desired-signal components from echoes becomes more and more difficult. This problem can also be solved with the aid of the training sequence. The echoes on the delayed paths also contain the training sequence. The correlation used for detecting the original training sequence may also be used for detecting the training-sequence echoes as well as their delay and loss. With the aid of this information, the received signal can be corrected by a channel equalizer.

Guard Period. Transmission in each timeslot is terminated with a guard period (Figure 1-31) of 8.25 bit periods (≈ 30 μs). During this time the level of the burst must be reduced from nominal to a minimum value (by up to 70 dB) and the burst is modulated with so-called dummy bits (logic 1); that is, no information is transmitted. The user of the next timeslot should start sending during this guard period so that his burst has reached nominal power when the actual transmission in the timeslot starts. This means that the switching time that cannot be used for transmitting information is used twice.

Delay Correction. The integrity of a timeslot depends on whether the subscribers send only during the period assigned to them and otherwise keep quiet. This is only possible when all subscribers are accurately synchronized. For practical reasons, the clock signal is generated by the BTS and the mobile stations synchronize to it. Conflicts with adjacent timeslots may occur in the uplink where several subscribers have to share the same channel. Although several subscribers are addressed in the downlink, signals can only be transmitted by the BTS.

We will now examine the effect of a distance of 10 km between an MS and a BTS (Figure 1-32). Synchronization of the MS is as follows:

Delay over 10 km = 33.3 μs (distance ÷ velocity of light).
The sync signals from the BTS require this time for transmission.
→ MS is synchronized by 33.3 μs too late.

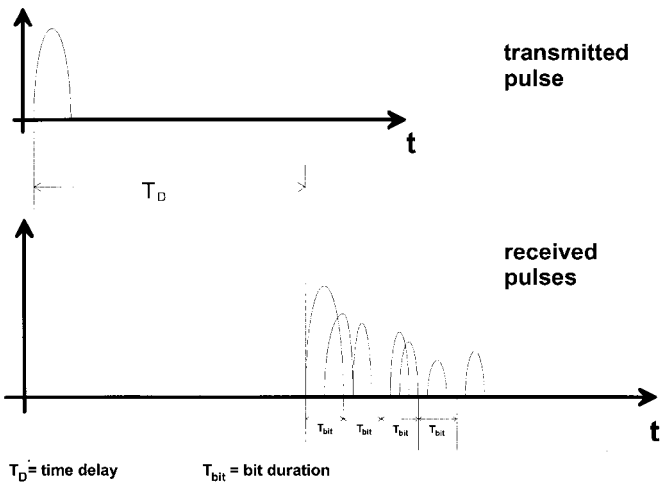


Figure 1-30 Channel impulse response.

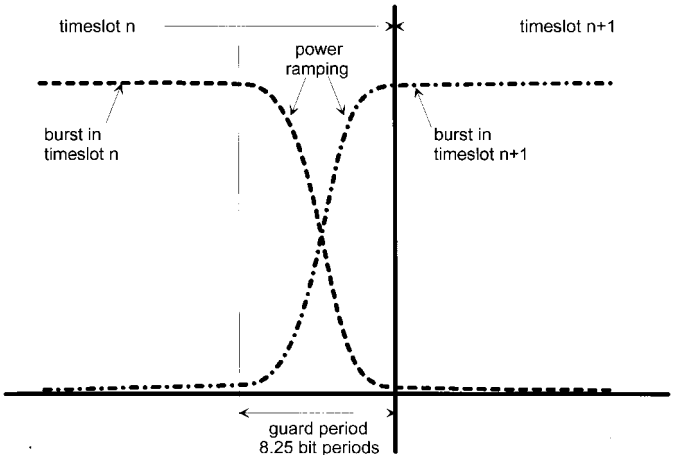


Figure 1-31 Guard period at the end of each timeslot.

MS sends a burst at the correct time from its point of view.
→ The burst is sent 33.3 µs too late.
Signal delay over a distance of 10 km → the burst requires another 33.3 µs. From the point of view of the BTS, the burst arrives with a delay of twice the delay time. Transmission cannot be made in the assigned timeslot and interferes with the next one.

The guard period at the end of each burst is only approximately 30 µs long and fully used in the example above. The greater the distance between the MS and BTS, the greater the effect of the signal delay. The only way to solve this problem is to make the MS send the burst at an earlier time. To do so, the distance between the MS and BTS must be known. The BTS determines the distance by means of a delay measurement and informs the MS of the actual delay. As a result, the MS corrects its transmission time so that the signals again arrive time-synchronized with the other mobiles at the BTS antenna.

In GSM, this procedure is called *timing advance* (TA) and is carried out continuously for all active mobile stations. Every 480 ms, a new TA value is sent to all active mobiles. A few limiting values of the GSM system can be deduced from the TA:

1. The TA is transmitted as a 6-bit word. Numerals from 0 to 63 can be represented with 6 bits. For instance, a TA of 10 means that for an MS from which bursts arrive with

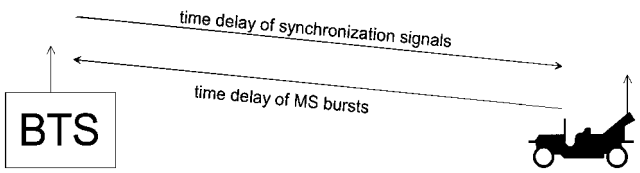


Figure 1-32 Signal delay and its effect.

a delay of 10 bit periods, transmission must be advanced by this amount. Since the TA is transmitted as an absolute value, a maximum delay of 232.5 μs (63 × 3.69 μs) can be signaled and corrected. This maximum double delay corresponds to a distance of approximately 34.9 km.

2. With the aid of the TA, the distance between the MS and the BTS can be determined. Since the smallest delay increment is a bit period, the resolution for a distance of one half of the delay corresponds to one bit period and therefore to approximately 550 m.

This synchronization scheme cannot solve an inherent problem: The delay of *the very first burst* sent by an MS cannot be corrected because, with contact between the MS and BTS yet to be established, the MS has not yet received an appropriate TA value from the BTS. Yet, if the mobile sends a normally timed burst, it may spill over into the next timeslot and interfere with transmission from another MS. To avoid causing interference in such cases, the mobile uses a special burst at the beginning of a transmission or when no valid TA has been received. Called the *access burst*, it is considerably shorter than a normal burst. The next section describes the access burst and other additional burst types in greater detail.

Burst Types. In addition to the normal burst described before, other bursts, including the access burst, are available for special purposes (Figure 1-33).

FREQUENCY CORRECTION BURST. The 142 "fixed bits" of the *frequency correction burst* (FCB) are all set to logic 0. With Gaussian minimum shift keying (GMSK), the type of modulation used in GSM, a stationary carrier frequency deviation, in this case approximately 67.7 kHz, is generated with this burst. The FCB is sent by the BTS only and used by the mobile for synchronization to the carrier frequency and for compensating a possible Doppler shift. It is sent by the BTS every 10 frames (approximately every 46 ms), but only in the timeslot 0 and only on a single (the C0) carrier.

SYNCHRONIZATION BURST. One frame after the frequency correction burst, but also in timeslot 0, the synchronization burst (SB) is transmitted. This burst is sent only by the BTS, and only on the C0 carrier. A particular difference between it and the normal burst is its considerably longer training sequence. Like the 26-bit training sequence of a normal burst, the SB's training sequence is also used for bit synchronization. Due to the great length of the sequence, synchronization can be more exact.

The twice 39 encrypted bits comprise timers T1, T2, and T3 in a coded form, and also the *base-station identification code* (BSIC). When this "GSM time" is received, the MS is synchronized to the BTS.

DUMMY BURST. A BTS must send continuously—that is, in all timeslots—on its C0 carrier because this carrier is used by the MS to find the nearest BTS and for evaluating reception quality. If no normal burst is available for transmission in a timeslot, the BTS sends dummy bursts instead, as the carrier cannot be transmitted without a modulation signal. Only the BTS sends dummy bursts, and only on the C0 carrier.

ACCESS BURST. As already pointed out, the access burst is sent when the MS first calls the BTS as a means of minimizing interference to other mobile transmitters (MTs) while initiating a delay measurement and the determination of the TA. In most cases, the access

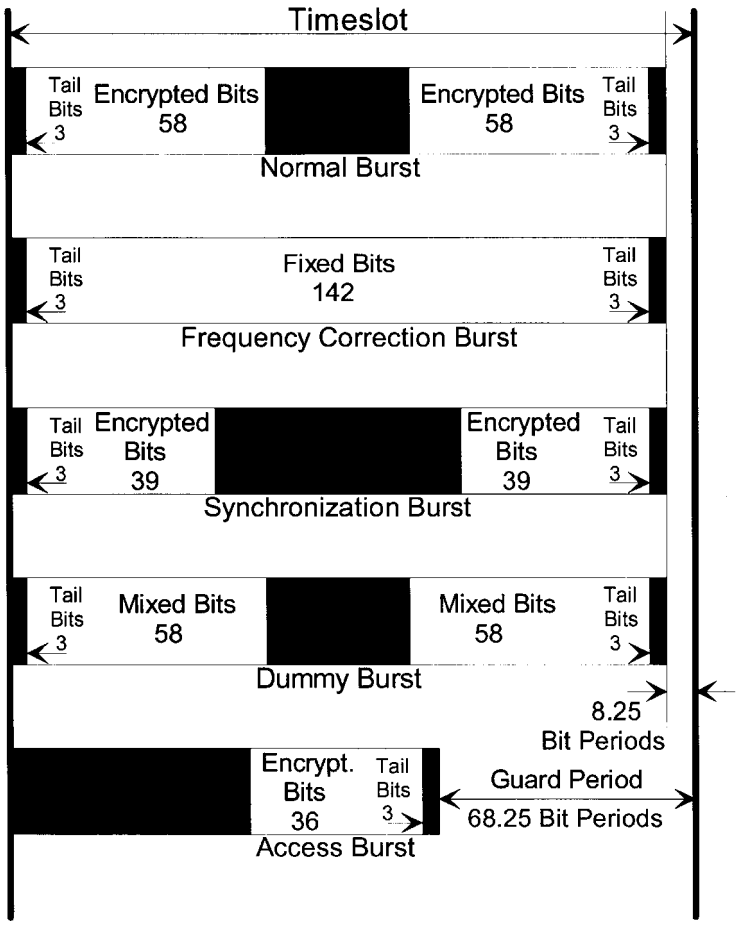


Figure 1-33 Burst types.

burst is also sent on the C0 carrier in the uplink direction, but in the case of a handover it may be transmitted on any carrier.

1-4 ABOUT BITS, SYMBOLS, AND WAVEFORMS

1-4-1 Introduction

Digital modulation of an RF carrier is the allocation of physically existing RF waveforms to the single elements of an alphabet of logical symbols where the number of allowed waveforms is equal to the number of logical elements of the alphabet (Figure 1-34). The most common alphabet is the binary one with the two logical symbols "0" and "1," but we will also deal with quaternary, octernary, and hexadecimal alphabets or more generally with *M*-ary alphabets comprising many more elements when discussing the signal generation with

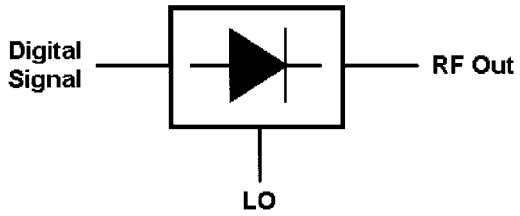


Figure 1-34 At base, digital modulation involves frequency shifting a baseband digital signal to RF. In practice, the process is more complicated than this because of bandwidth constraints on the resulting RF signal.

signal generators and dedicated software packages. The waveforms representing these symbols differ from each other by their parameter's amplitude $a(t)$, their frequency $f(t)$, and their phase $\varphi(t)$.

A modulator therefore is nothing more than a device by which this allocation is performed (Figure 1-35). From a coder it receives the logical symbols and emits at its output the corresponding waveforms $s_i(t)$. The waveform generation may be done by using a set of distinct generators (e.g., two oscillators to generate two signals with different frequencies in the case of binary frequency shift keying), by classical amplitude or frequency modulators, or by more sophisticated equipment such as I/Q modulators for M-ary modulations.

On their way across the RF channel from the transmitter to the receiver, these waveforms are distorted by noise and other disturbing properties of the RF channel.

The task of the receiver is to interpret the received waveforms $r_i(t)$ and to reallocate the proper logical symbols to them. For this purpose it is not necessary to reconstruct the original waveforms from the distorted ones (Figure 1-36). The important thing is to find out which

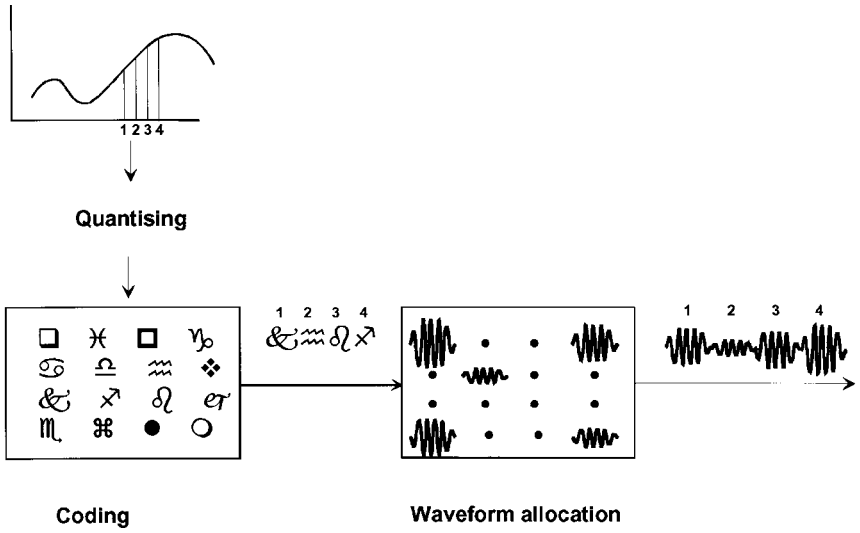


Figure 1-35 Digital modulator.

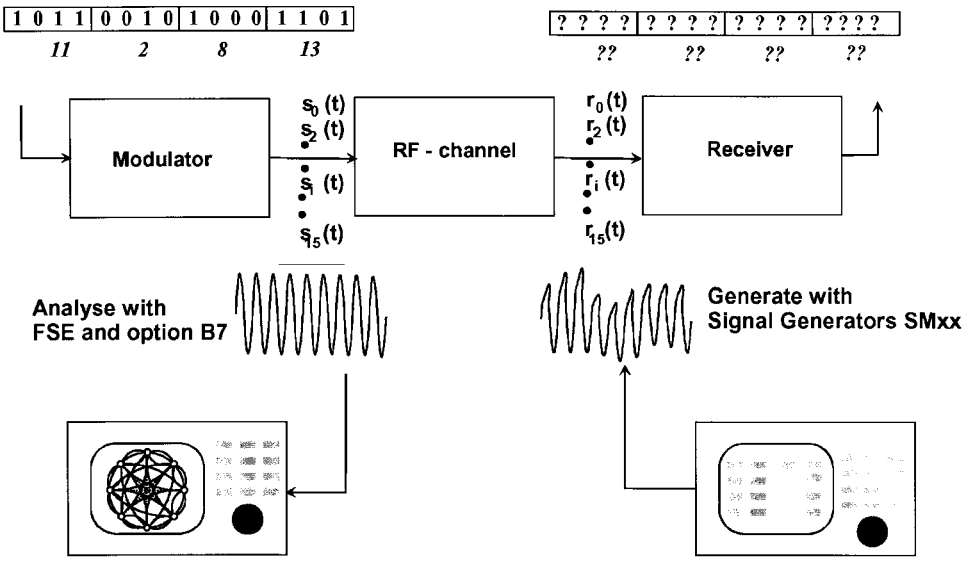


Figure 1-36 The information channel.

symbol has most probably been sent when a certain signal $r_i(t)$ has been received, a process that is known as maximum likelihood estimation.

For meaningful receiver tests, therefore, waveforms have to be generated that mimic real, distorted signals to prove the ability of a receiver to tolerate waveform distortions to a certain extent.

Representation of a Modulated RF Carrier. The waveform of a modulated RF carrier can be expressed as

$$s(t) = a(t)\cos[2\pi f_c(t)t + \varphi(t)] \tag{1-7}$$

and is defined by its amplitude $a(t)$, its carrier frequency $f_c(t)$, and its phase $\varphi(t)$. All three parameters are time variant and may be altered to generate different waveforms to represent logical symbols. If the occupied bandwidth of this modulated carrier is narrow compared to the carrier frequency f_c, we call this signal the *RF-bandpass signal* (Figure 1-37).

As any frequency variation causes a phase variation and vice versa a phase variation always causes a frequency variation, we can replace any frequency modulation by a corresponding phase modulation. Therefore we simplify the above equation to

$$s(t) = a(t)\cos[2\pi f_c t + \varphi(t)] \tag{1-8}$$

that is, we consider the carrier frequency as a constant and concentrate all frequency and phase variations into the parameter $\varphi(t)$.

For our purposes, a more suitable representation is

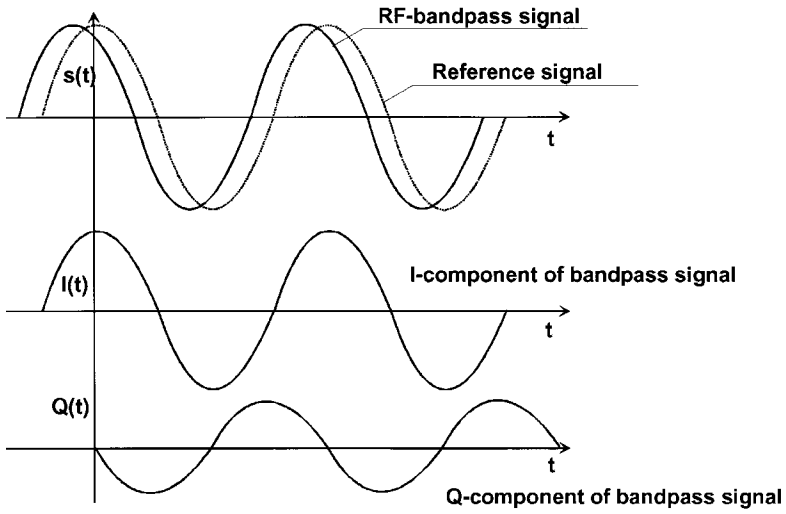


Figure 1-37 The bandpass signal and the I/Q representation of a carrier.

$$a(t)\cos[2\pi f_c t + \varphi(t)] = \cos[\varphi(t)]a(t)\cos(2\pi f_c t)$$

$$-\sin[\varphi(t)]a(t)\sin(2\pi f_c t) \tag{1-9}$$

which we call the *I/Q* representation of the RF signal. I/Q means that we have an I (in-phase) signal, namely, $\cos[\varphi(t)]a(t)\cos(2\pi f_c t)$, and a Q (quadrature) signal, namely, $-\sin[\varphi(t)]a(t)\sin(2\pi f_c t)$. These equations help us a lot in understanding an I/Q modulator. Because of the phase difference of 90° between the two carrier components, these are said to be *orthogonal* to each other.

All the information about the (modulated) carrier with the carrier frequency f_c is contained in the terms

$$c_I(t) = a(t)\cos[\varphi(t)] \tag{1-10}$$

$$c_Q(t) = a(t)\sin[\varphi(t)] \tag{1-11}$$

and, lazy as we are, we therefore disregard the terms $\cos(2\pi f_c t)$ and $-\sin(2\pi f_c t)$ from further consideration and denote the above signals $c_I(t)$ and $c_Q(t)$ as the components of the *complex baseband waveform* or *baseband signal*.

This leads us immediately to the *vector representation* of the signal, where we consider the two components $c_I(t)$ and $c_Q(t)$ of the complex baseband signal as the time-variant components of a time-variant vector with the vector length $a(t)$ and the angle to the I-axis $\varphi(t)$.We also get

$$a(t) = \sqrt{c_I^2(t) + c_Q^2(t)} \tag{1-12}$$

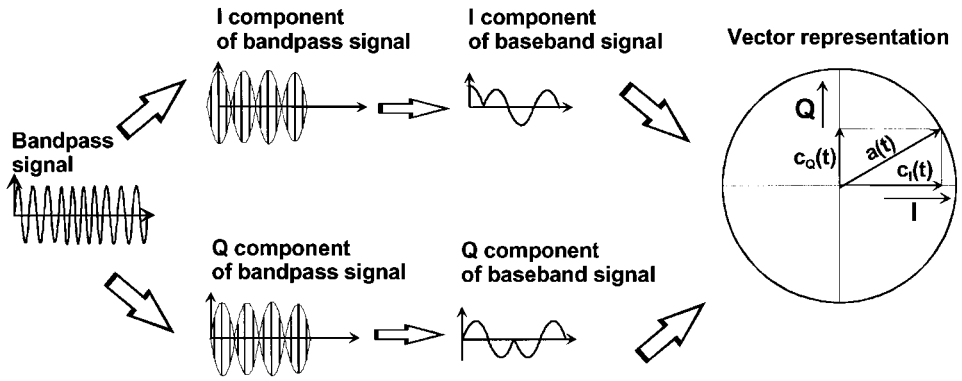


Figure 1-38 Different forms of signal representation.

$$\varphi(t) = \arctan\left(\frac{c_Q}{c_I}\right) \tag{1-13}$$

The vector can be depicted in the I/Q area (Figure 1-38).

Generation of the Modulated Carrier. Once we have realized that the modulated carrier can be represented as the sum of its I and Q components, which are the product of the two baseband components with two orthogonal RF carriers of the same frequency, it is easy to understand the hardware of the modulator (Figure 1-39). An unmodulated RF carrier is split up into two equal oscillations $\cos(2\pi ft)$, one of the two is then shifted by 0.5π and therefore is described by $-\sin(2\pi ft)$. The component $\cos(2\pi ft)$ is multiplied with the I component of the baseband signal $c_I(t)$; the other one, $-\sin(2\pi ft)$, is multiplied with the Q component $c_Q(t)$ of the baseband signal. Each multiplication may be performed using a double-balanced mixer. Afterward the two RF components are added in a simple power combiner. As it is difficult to shift the carrier by 90° over a broad frequency range, the modulated carrier is generated at an intermediate frequency and then upconverted to the wanted output frequency in a second mixer stage.

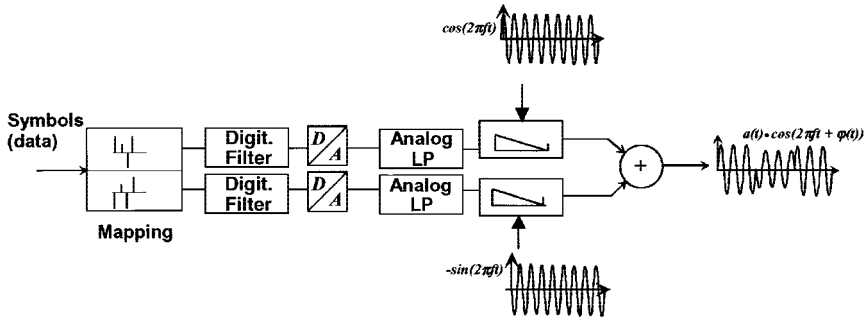


Figure 1-39 Principle of a digital I/Q modulator.

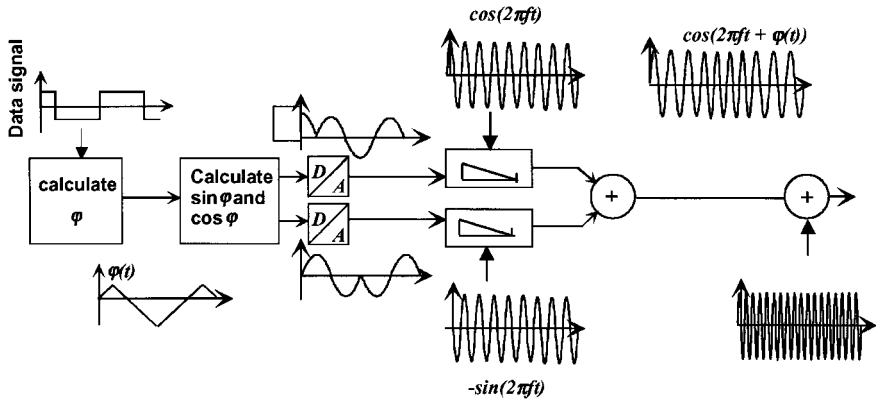


Figure 1-40 I/Q modulation (MSK and GMSK).

The baseband signals are generated by mapping every digital symbol onto a pair of digital pulses, which are fed to digital baseband filters. The output signal of these filters is D/A converted and smoothed by analog low-pass filters.

Figure 1-40 shows another example where, for a given modulation (MSK or GMSK), the instantaneous phase and then the corresponding cosinusoid and sinusoid, which modulate the two carrier components, are calculated from the data signal.

Digital designs of the modulator also exist, in which the IF-carrier generation, the time-variant phase shift, the multiplication with the baseband signals, and the sum of the components are calculated in a digital signal processor, the output of which is D/A converted and upconverted to the output frequency in the classical way. A further possibility is the generation of the modulated carrier with direct digital synthesis (DDS), as used in the Rohde & Schwarz SME signal generator.

Mapping the Data onto the Baseband Waveforms. The next question is: "How do we generate the baseband waveforms $c_I(t)$ and $c_Q(t)$?" There is no general answer to this question, as the generation of the baseband waveforms depends on the type of modulation. The following short descriptions will suffice for the moment.

Linear Modulations (All Kinds of Amplitude and Phase Shift Keying, and *M*-ary QAM)

- For binary amplitude and phase shift keying (ASK and BPSK), the data signal itself represented as a unipolar (ASK) or bipolar (BPSK) nonreturn-to-zero (NRZ) signal is the baseband waveform $c_I(t)$; the component $c_Q(t)$ does not exist.
- For *M*-ary phase shift keying and *M*-ary quadrature amplitude modulation, *N* bits are combined to form new symbols that are elements of an alphabet with $M = 2N$ elements. In the simplest case, every symbol is allocated an I and a Q amplitude during the symbol duration, which is *N* times the bit duration. The modulating signals $c_I(t)$ and $c_Q(t)$ are then staircase functions, and the modulated carrier has a time-varying envelope with the instantaneous amplitude $a(t)$ (Figure 1-41). Because the steps of the envelope cause unwanted side lobes of the RF spectrum, the baseband signals are filtered to smooth the shape of the RF envelope and reduce the occupied bandwidth of the modulated RF signal.

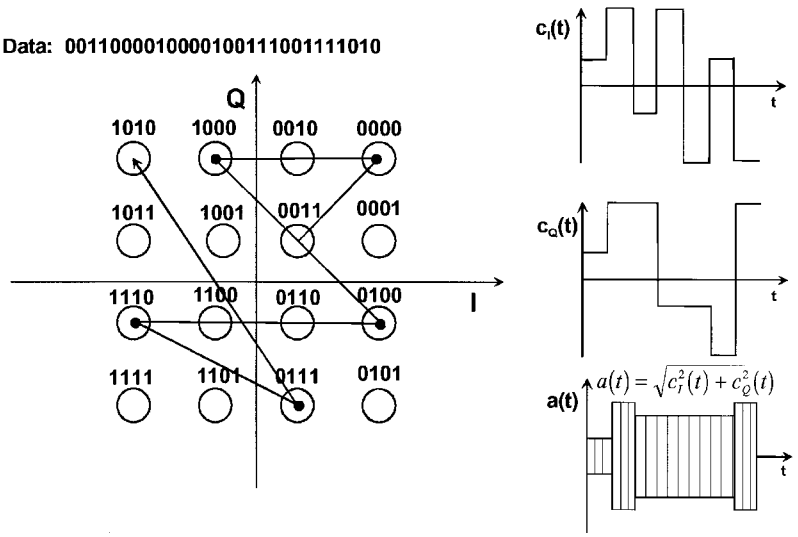


Figure 1-41 Constellation diagram, and baseband and RF signals of 16-QAM.

Nonlinear Modulations (Frequency Shift Keying, Minimum Shift Keying, and Gaussian Minimum Shift Keying)

- Despite the fact that *M*-ary frequency shift keying (FSK) could be performed using an I/Q modulator, for this type of modulation much simpler equipment such as a voltage-controlled oscillator is used as a frequency modulator. Figure 1-42 shows an example of quaternary frequency shift keying (4FSK), which also is known as 4 pulse-amplitude modulation/frequency modulation (4PAM/FM). This term indicates that every two bits are combined to a dibit that is mapped onto a baseband pulse with an amplitude taking on one of four possible levels. The pulse then is shaped by a baseband filter before being fed to the frequency modulator.
- If more precise modulations are required (e.g., MSK and GMSK, which also turn out to be frequency modulations), first the instantaneous phase of the modulated RF carrier is calculated from the data. The corresponding sine and cosine values that form the modulating baseband signals $c_I(t)$ and $c_Q(t)$ are determined from a lookup table. This operation is the reason for the fact that frequency modulation is called a nonlinear modulation.

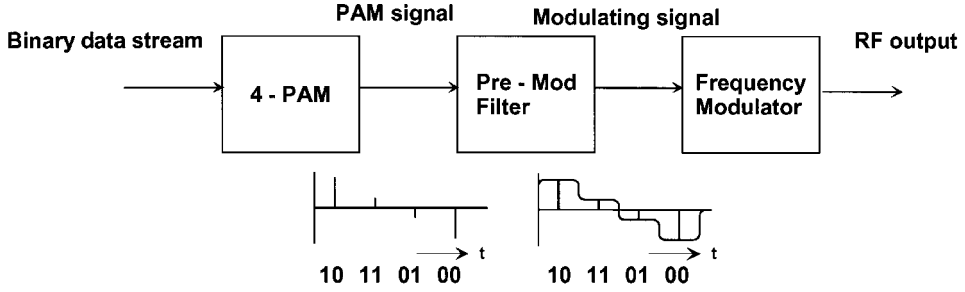


Figure 1-42 4PAM/FM.

The Spectrum of a Digitally Modulated Carrier. It is the task of any transmission process to occupy as little bandwidth as possible. The absolute lower limit in the baseband is half the symbol rate of the baseband signal, where for M-ary modulation the symbol rate r_{Symbol} is equal to the bit rate divided by $ld\,(M)$. This lower limit is only theoretical as ideal rectangular filters, which cannot be realized, were necessary. Therefore, in practice, a minimum baseband bandwidth of about $0.75r_{\text{Symbol}}$ has to be taken into account.

With linear modulation, the occupied bandwidth in the RF range is twice the occupied baseband bandwidth. This follows from the lag theorem, according to which the double-sided spectrum of a time function is shifted from $f = 0$ to the frequency $f = f_c$ when the time function is multiplied with $\cos(2\pi f_c t)$ (Figure 1-43).

Expressing this with formulas, we find

$$c(t) \quad \circ\!\!-\!\!\bullet \quad C(f) \tag{1-14}$$

$$e^{j2\pi f_0 t}c(t) \quad \circ\!\!-\!\!\bullet \quad C(f - f_0) \tag{1-15}$$

Therefore, if the baseband spectrum is limited by a low-pass filter, the RF spectrum is limited as if it were filtered by an RF bandpass filter with twice the bandwidth of the baseband filter.

Demodulation of Digitally Modulated Carriers. The demodulation process is an estimation process. The receiver compares the received waveform with all the possible waveforms and decides which symbol has been received. The waveform allocated is the one that is most similar to the received signal. This process is also called *maximum likelihood estimation*.

Figure 1-44 shows a possible demodulator. The received signals are cross-correlated with the stored ones in M correlators. High output from a given correlator indicates a match between its stored symbol and a symbol in the incoming signal.

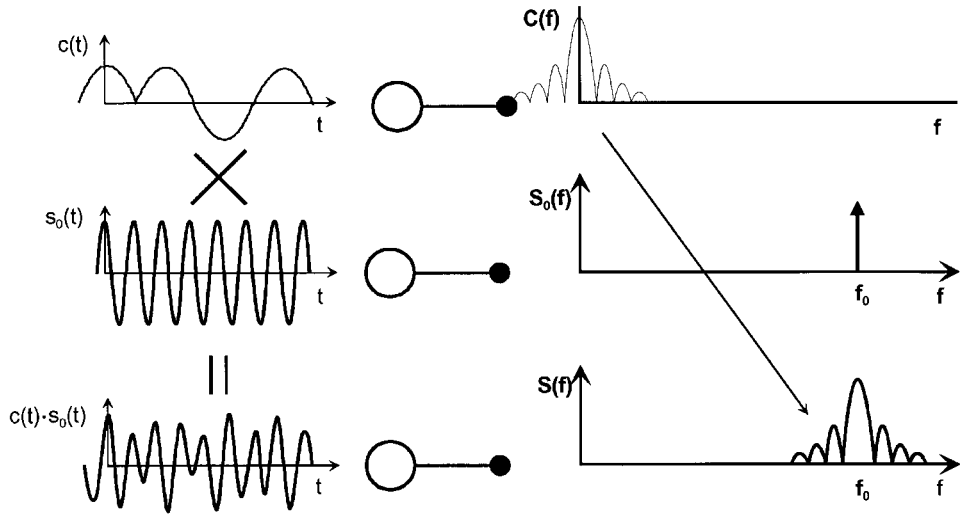


Figure 1-43 Occupied bandwidth in the baseband and the RF range.

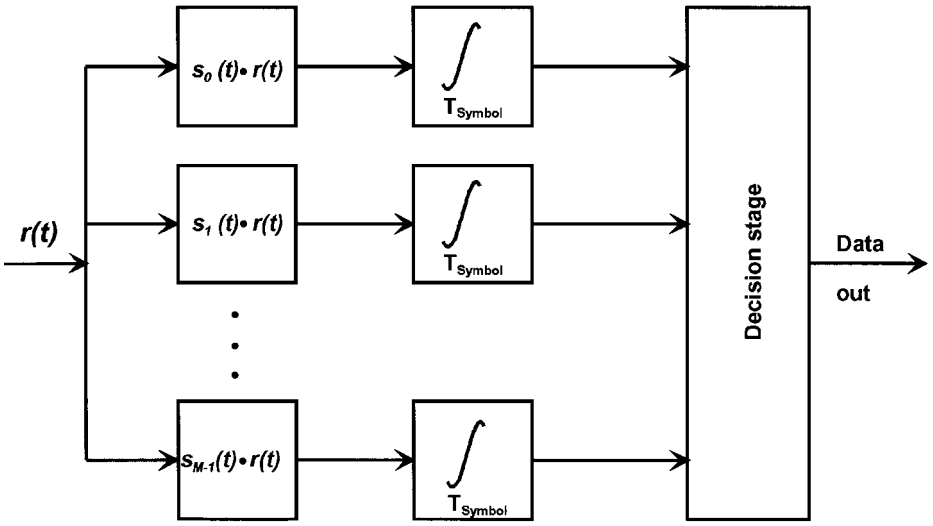


Figure 1-44 Correlation receiver.

As can easily be seen, such a receiver is quite complex. Therefore, other principles of receivers have been studied. The most often used receiver is based on an I/Q demodulator, the principle of which is shown in Figure 1-45.

The received signal is split into two equal components. One component is multiplied with the reconstructed carrier signal, the other with its orthogonal counterpart; that is, the $\pi/2$-shifted reconstructed carrier. After low-pass filtering, this multiplication delivers the I and the Q components of the baseband signal, which is more or less distorted by interference, multipath, and noise in the radio channel. These demodulated baseband signals then can be A/D converted and treated using simple or sophisticated algorithms to decide which symbols were originally sent. This type of demodulation requires an exact reconstruction of the unmodulated carrier, with respect not only to frequency but also to phase, and is called *coherent demodulation*. While the frequency recovery is in principle quite simple, the

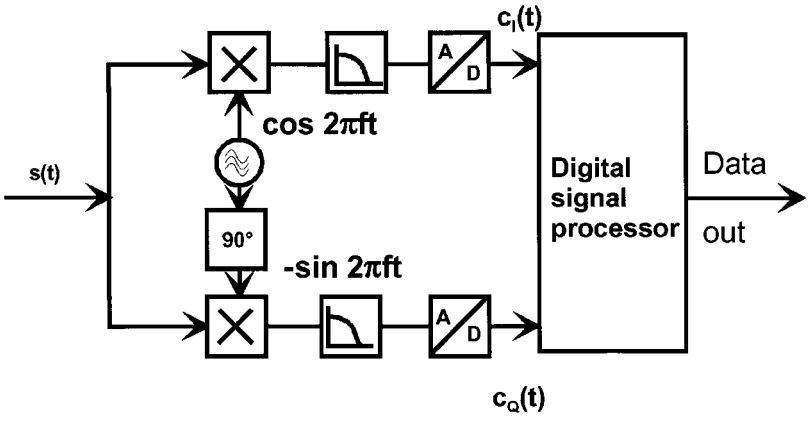


Figure 1-45 Coherent demodulation.

acquisition and tracking of the signal phase require complex signal processing of the baseband signal.

Coherent demodulation and its accompanying costly signal acquisition and tracking are necessary for all types of phase and quadrature amplitude modulation in which any part of the transmitted information is coded in the absolute phase of the carrier. It only can be avoided when the information is coded in the phase difference of two sequential waveforms. Such modulation schemes are called *differential phase modulation*. Examples of such modulation include differential binary phase shift keying (DBPSK) and differential quadriphase shift keying (DQPSK) and their derivatives.

The following list shows common digital modulation types and the telecommunication systems that use them:

BPSK—Satellite radio links, GPS, Inmarsat
QPSK—Satellite radio links, GPS, IS-95
OQPSK—IS-95 (reverse link)
π/4-QPSK—NADC, PHS, PDC. TETRA
GMSK—GSM, DCS 1800
GFSK (Gaussian FSK)—DECT
COFDM (coded orthogonal frequency-division multiplexing)—DAB, DVB

1-4-2 Some Fundamentals of Digital Modulation Techniques

The first waveforms that were used for "digital modulation" were amplitude shift keying (ASK), similar to on/off-keyed Morse code, and phase shift keying (similar to FM). Figure 1-46 shows the resulting waveforms in the time domain. It is also useful to look at them in the phase domain—the in-phase (I) and quadrature (Q) planes. From Figure 1-47 we see that ASK has only two states: no signal and signal at I (in-phase), while phase shift keying can have two states: +I and –I. In the frequency domain, ASK and BPSK have the output spectra

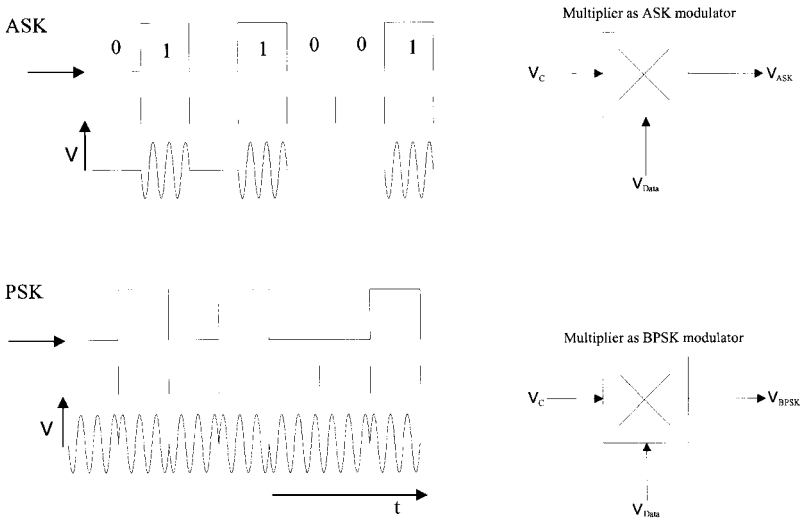


Figure 1-46 Amplitude shift keying (ASK) and phase shift keying (PSK) in the time domain.

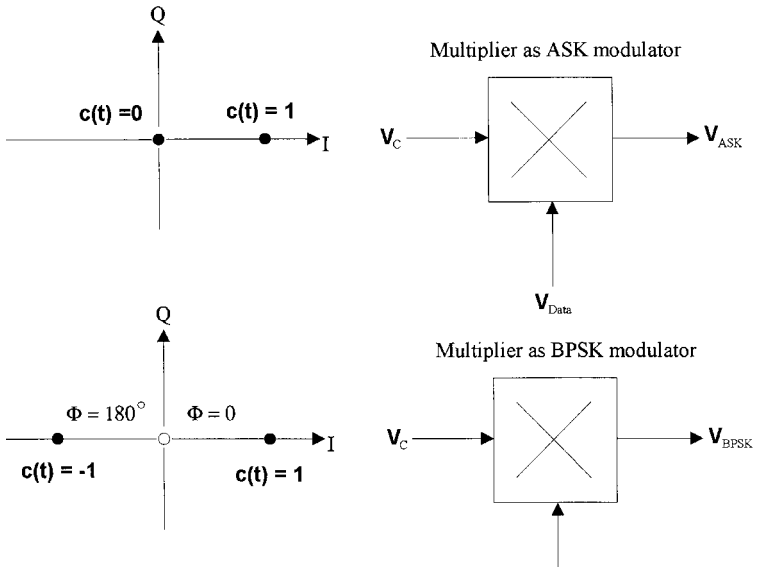


Figure 1-47 ASK and PSK in the I/Q plane.

shown in Figure 1-48. These output spectra, with their (sin *X*)/*X* appearance, depend on the rise and decay times and duty cycle of the modulating signal.

BPSK sends one data bit per signal state. Another way to put this is that each of BPSK's two signal states is a *symbol* that stands for just one bit—0 or 1. Transmitting more than one bit per symbol would allow us to improve our data throughput per unit of time, and this is

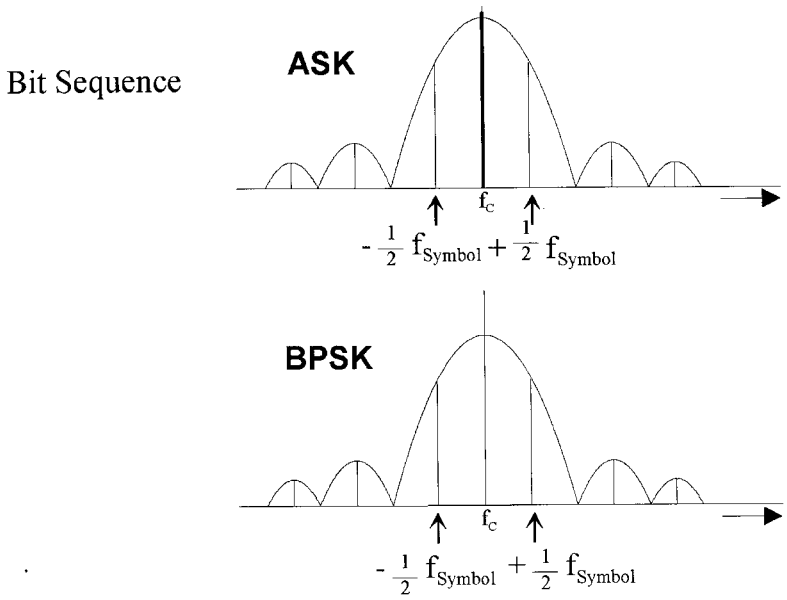
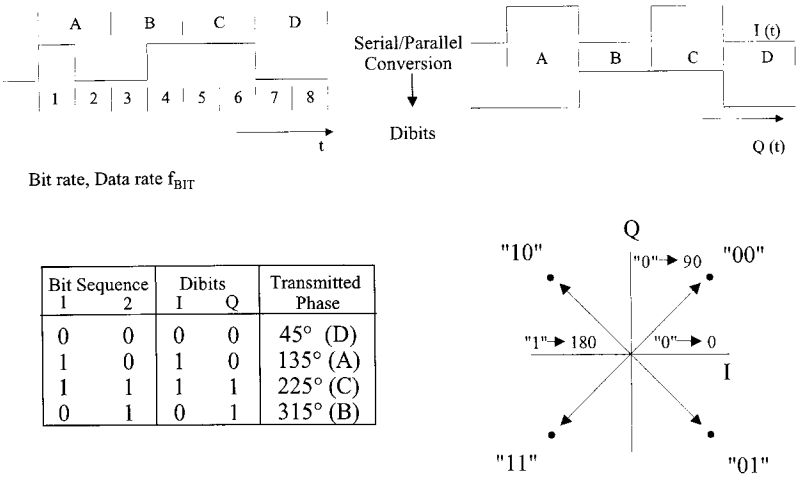


Figure 1-48 ASK and BPSK in the frequency domain.



Bit Sequence 1	2	Dibits I	Q	Transmitted Phase
0	0	0	0	45° (D)
1	0	1	0	135° (A)
1	1	1	1	225° (C)
0	1	0	1	315° (B)

Figure 1-49 Quadrature PSK (QPSK) modulator (I/Q modulator).

exactly what's done in quadrature PSK (QPSK), which uses *four* possible signal states as symbols for 2-bit sequences called *dibits* (Figure 1-49). Figure 1-50 shows a QPSK modulator and the QPSK constellation diagram; Figure 1-51 shows the result of QPSK in the time and frequency domains; Figure 1-52 shows a spectrogram of an actual QPSK emission.

Moving from ASK to BPSK or QPSK results in increased resistance to noise. Figure 1-53 illustrates this by graphing bit error rate versus signal-to-noise ratio (here expressed as 10 log (E_{bit}/N_0), where E_{bit} is energy per bit and N_0 is the noise [23]). Figure 1-54 compares the maximum interference voltages for BPSK and QPSK.

Figure 1-55 shows bandwidth requirements and constellation diagrams for BPSK and QPSK. Unfiltered BPSK and QPSK are purely angle-modulated emissions; for each, all symbols are transmitted at the same amplitude. As we'll see, however, bandwidth-limiting a BPSK or QPSK signal to minimize adjacent-channel interference results in envelope variations that must be preserved through careful circuit design if bit errors are to be minimized.

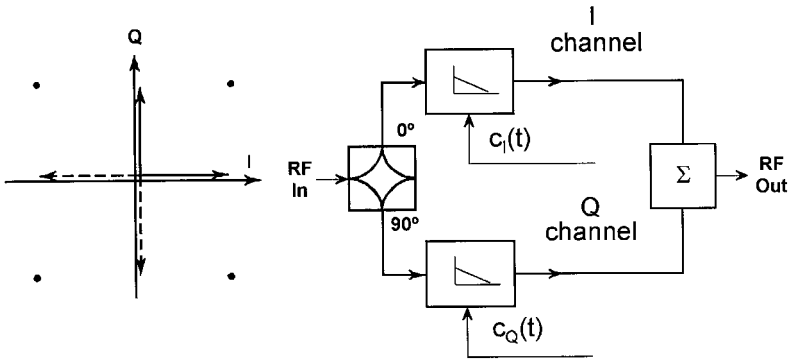


Figure 1-50 QPSK constellation diagram (*left*) and modulator (*right*).

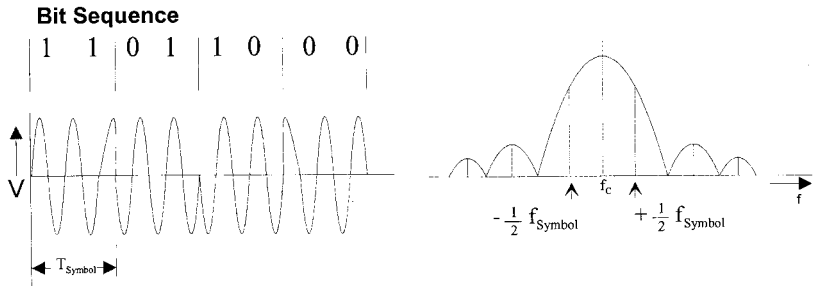


Figure 1-51 Result of QPSK modulation in the time and frequency domains.

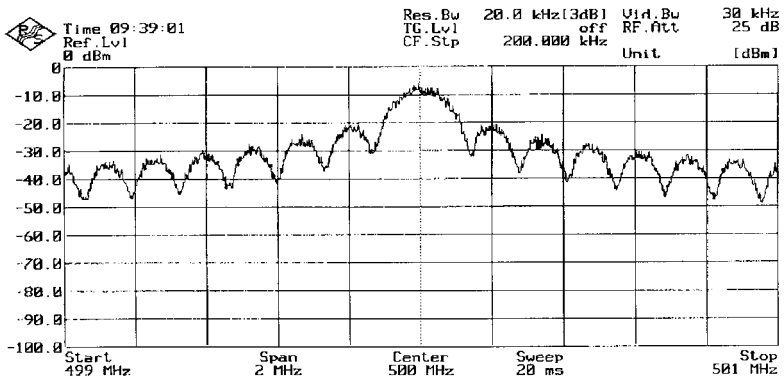


Figure 1-52 QPSK spectrum.

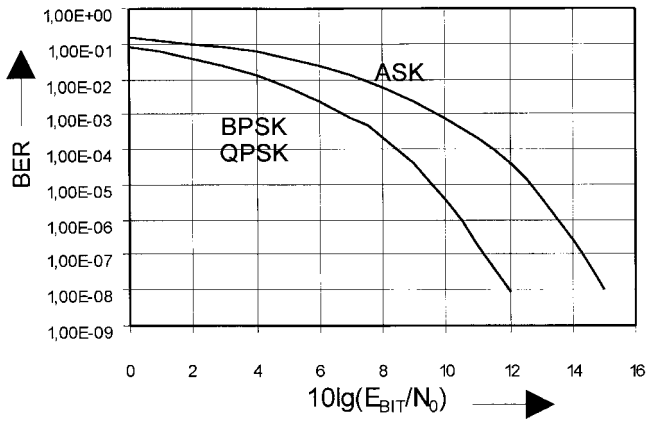


Figure 1-53 Bit error rate (BER) in terms of E_{bit}/N_0 for BPSK, QPSK, and ASK.

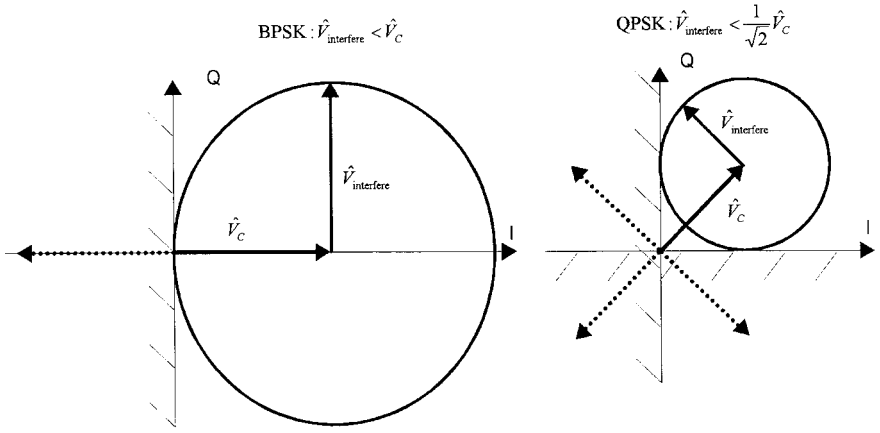


Figure 1-54 Maximum interference voltages for BPSK and QPSK, where $\hat{V}_C$ is the carrier voltage of the desired signal and $\hat{V}_{interfere}$ is the carrier voltage of the interfering signal.

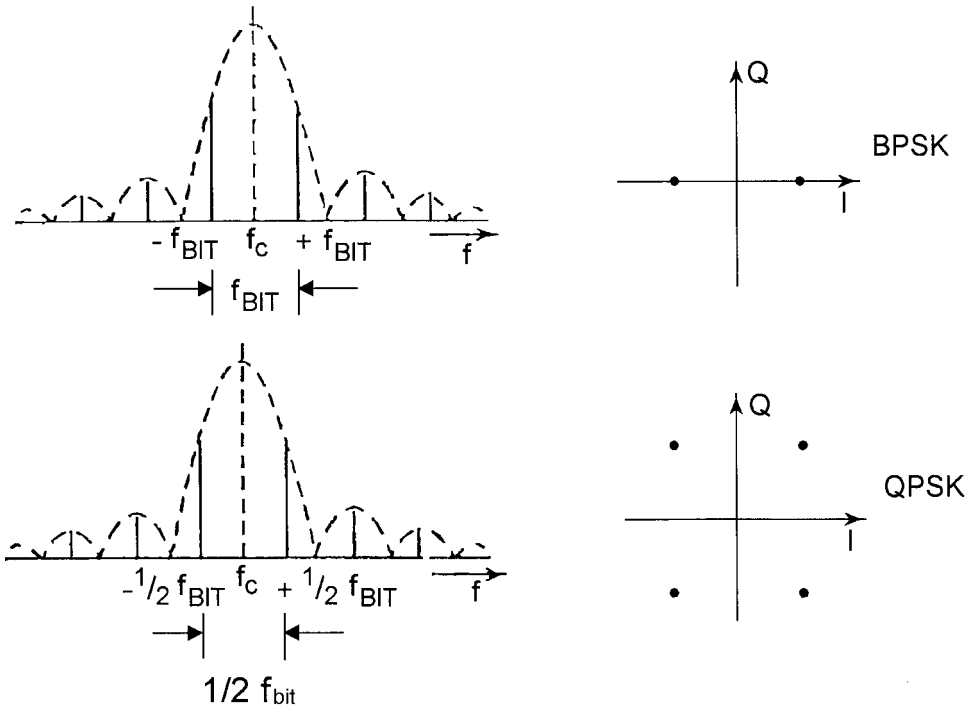


Figure 1-55 Bandwidth requirements (*left*) and constellation diagrams (*right*) for BPSK and QPSK. For each emission, all symbols are transmitted at the same amplitude, as is indicated by their equidistance from the constellation origin.

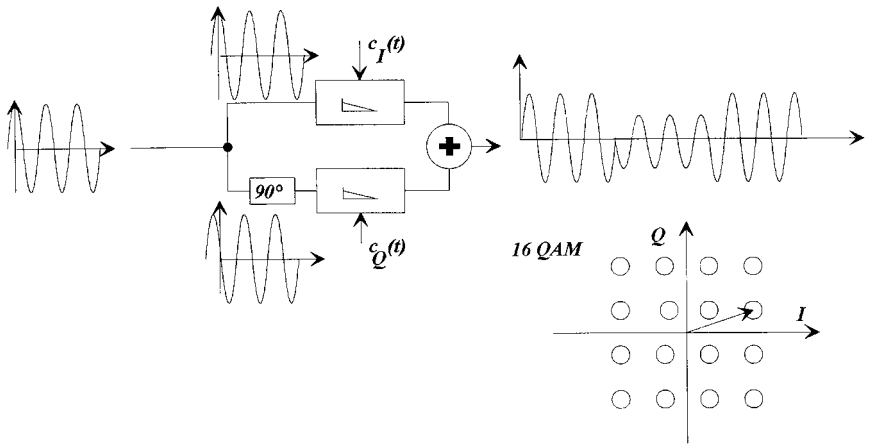


Figure 1-56 Quadrature amplitude modulation (QAM), with constellation diagram for 16-state QAM (16-QAM). Because transitions between QAM symbols involve shifts in phase *and amplitude*; even unfiltered QAM is a combination of amplitude and angle modulation.

Further increases in bit rate per symbol are possible with phase-modulation schemes that divide the 360° of each carrier cycle into increasingly smaller segments. Transmitting 3 bits per symbol, with each symbol spaced from its neighbors by 45°, gives 8-PSK; transmitting four bits per symbol, with each symbol spaced from its neighbors by 22.5°, results in 16-PSK. Increasing the data rate by increasing the number of bits per symbol doesn't give us something for nothing, however: The smaller the phase difference between adjacent symbols, the greater the likelihood that phase shifts resulting from phase noise, multipath reception, and other sources of phase perturbation will cause demodulation errors.

Modulation schemes that differentiate their data symbols through a combination of phase *and amplitude* shifts are susceptible to disturbances in amplitude in addition to phase. Figure 1-56 shows the modulator and constellation diagram for 16-state quadrature amplitude

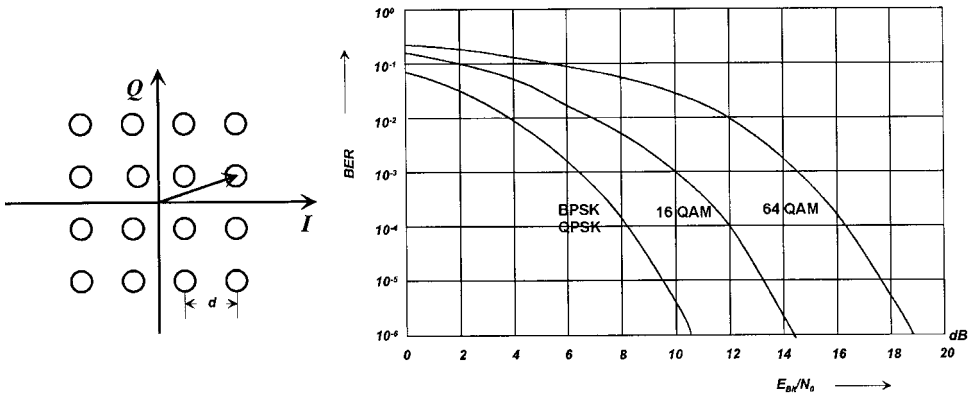


Figure 1-57 Bit error rate versus E_{bit}/N_0 for BPSK/QPSK, 16-QAM, and 64-QAM, showing the significantly greater SNR necessary for a given BER as the number of signal states is increased. In 64-QAM, each symbol represents 6 bits.

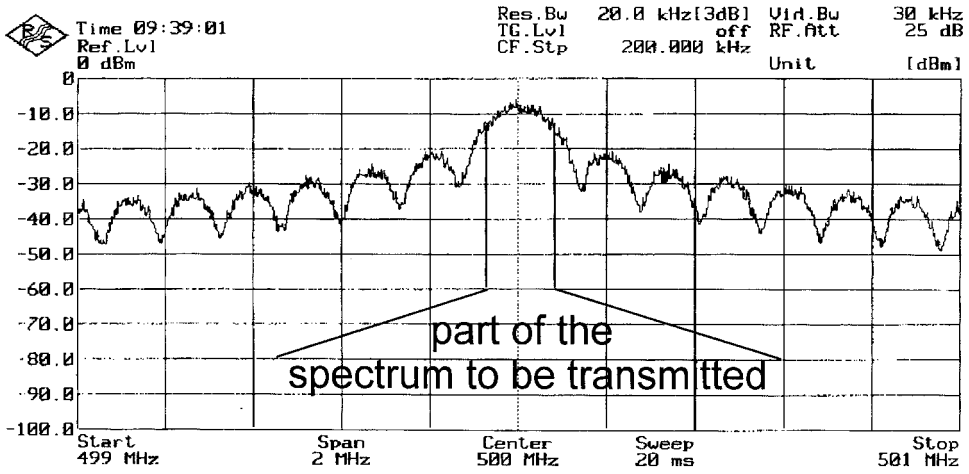


Figure 1-58 QPSK spectrum resulting from pseudorandom binary sequence (PRBS) data. Most of an angle-modulated emission's infinite sidebands are unnecessary for communication and can be removed by filtering—at the cost of introducing amplitude variations that must be sufficiently preserved to keep the bit error rate low.

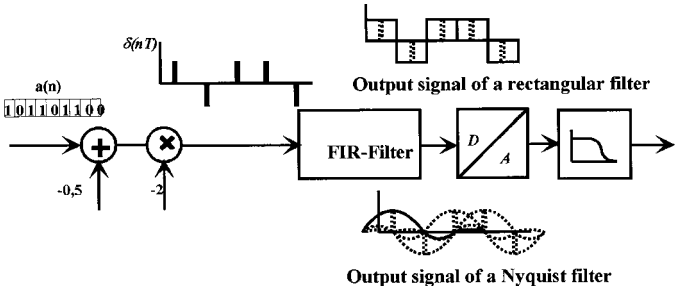


Figure 1-59 Baseband filtering. Figure 1-60 shows the result for QPSK.

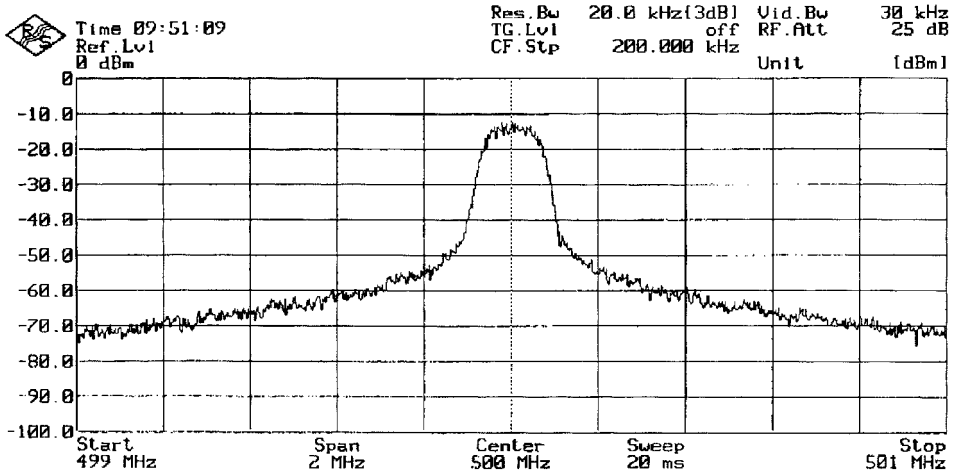


Figure 1-60 Spectrum of a band-limited QPSK signal.

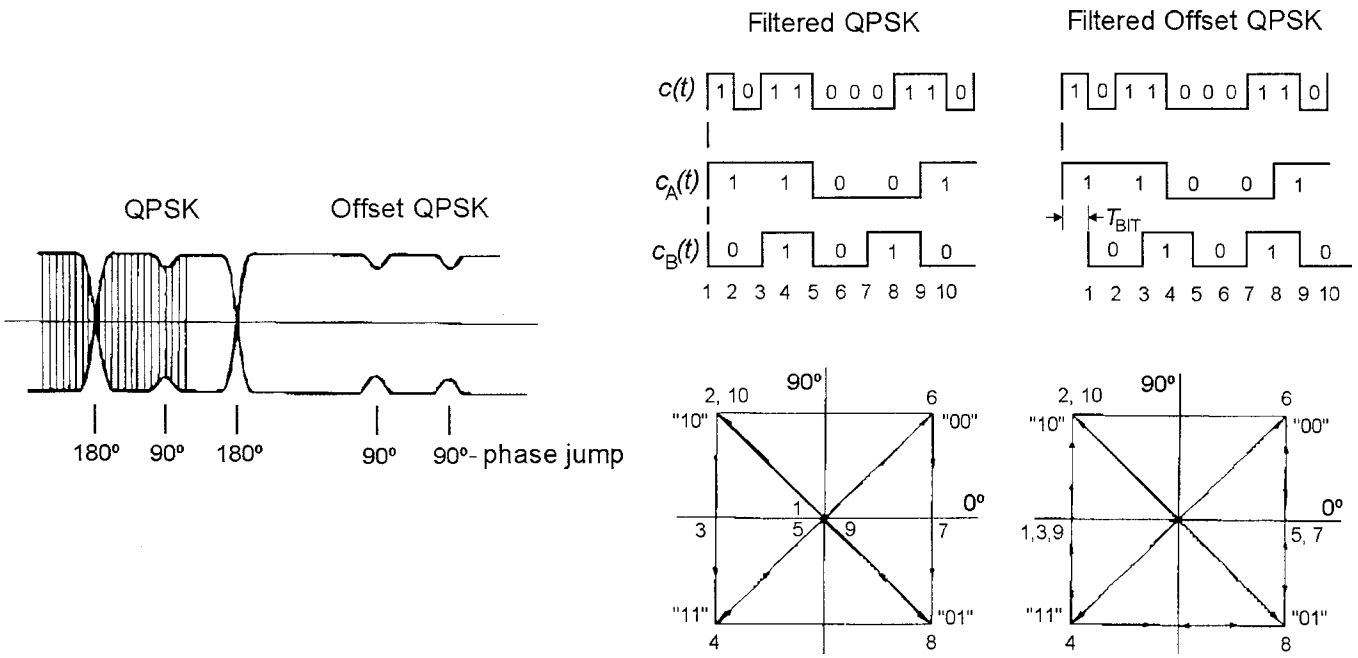


Figure 1-61 Offset QPSK (OQPSK) exhibits reduced envelope variations compared to standard QPSK by time-offsetting the I and Q channels.

modulation (16-QAM), in which each symbol represents 4 bits. Figure 1-57 shows bit error rate versus SNR for BPSK/QPSK, and 16- and 64-QAM. QAM systems can support very high bit rates—*if* sufficient SNR can be maintained and disturbances in signal amplitude and phase can be kept under control [3].

The infinite number of sidebands produced by angle modulation must be reduced by filtering to minimize the spectrum occupied by an emission to just that necessary for communication (Figure 1-58). Considerations of filter realizability and transmitter frequency agility dictate that this filtering must be done at baseband (Figure 1-59). The filter characteristics are optimized in accordance with the modulation scheme used (Figure 1-60). The filtering is done in using DSP, minimizing component count and resulting in characteristics that are essentially temperature independent and identical from unit to unit.

Fundamental angle-modulation theory tells us that the envelope of an angle-modulated signal does not vary in amplitude. Such an emission could be amplified in highly nonlinear stages without distortion. But because angle-modulated signals must be band-limited for practical use, and because band-limiting an angle-modulated signal strips away sideband energy that would contribute to its envelope constancy, the emissions used in modern wireless systems, even those that do not involve intentional amplitude modulation, vary in amplitude with modulation. How much a wireless emission's amplitude varies depends on the degree of filtering and the particular modulation scheme used. Various schemes, many of which rely on particular synergies of coding, modulation, and filtering, have been devised for minimizing amplitude variation in digitally modulated signals. Figure 1-61 shows one [offset QPSK (OQPSK)]; Figure 1-62 compares the spectra of three [QPSK, minimum shift keying (MSK), and Gaussian MSK (GMSK)]. Detailed discussion of such techniques is more concerned with coding, logic circuitry, and software than with radio hardware and is therefore beyond the scope of this book. What is important to the circuit designer is that stages handling the emission(s) on which a system depends must be sufficiently amplitude-linear to maintain the modulation integrity of the signal and, in the case of transmitters, to keep adjacent-channel interference and other spurious emissions within acceptable levels.

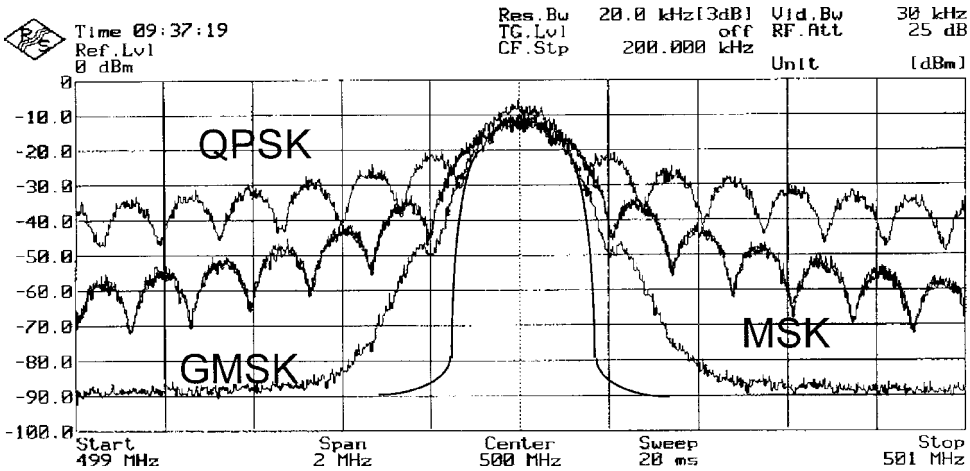


Figure 1-62 Comparison of QPSK, MSK, and GMSK.

1-5 ANALYSIS OF WIRELESS SYSTEMS

1-5-1 Analog and Digital Receiver Designs

For the purpose of showing the capability of modern computer-aided design (CAD), we now show first an analog receiver and following this an analog/digital receiver with its functionality and performance.

Receiver Design Examples

Analog Receiver Design. A transmitter modulates information signals onto an RF carrier for the purpose of efficient transmission over a noise-filled radio channel. The RF receiver's job is to demodulate that information signal while maintaining a sufficient signal-to-noise ratio (SNR). This must be done for widely varying input RF power level and with the presence of noise and interferers.

Modern communication standards place requirements on key system specifications, such as RF sensitivity and spurious response rejection. These system specifications must then be separated into individual circuit specifications via an accurate overall system model. CAD can play a major role in modeling systems and in determining the individual component requirements. A 2.4-GHz dual-down conversion system (see Figure 1-63) will be used as an example. The first IF is 200 MHz and the second IF is 45 MHz.

As a signal propagates from the transmitter to the receiver it is subject to path loss and multipath, resulting in extremely low signal levels at the receive antenna. RF sensitivity is a measure of how well a receiver can respond to these weak signals. It is specified differently for analog and digital receivers. For analog receivers there are several sensitivity measures, including minimum discernible signal (MDS), signal-to-noise plus and distortion ratio (SINAD), and noise figure. For digital receivers the typical sensitivity measure is maximum bit error rate (BER) at a given RF level. Typically, a required SINAD at the baseband demodulator output is specified over a given RF input power range. For example, audio measurements may require 12-dB SINAD at the audio output over RF input powers ranging from −110 to −35 dBm. This can then be translated to a minimum carrier-to-noise (C/N) ratio at the demodulator input to achieve a 12-dB SINAD at the demodulator output (see Figure 1-64).

Determining receiver sensitivity requires accurate determination of each component's noise contribution. Modern CAD software, such as Symphony by Ansoft, can model noise from each stage in the system including oscillator phase noise. The C/N ratio can then be plotted as a function of input power (see Figure 1-65).

This plot enables straightforward determination of the minimum input power level to achieve a certain C/N ratio. Once the sensitivity has been specified the necessary gain, or

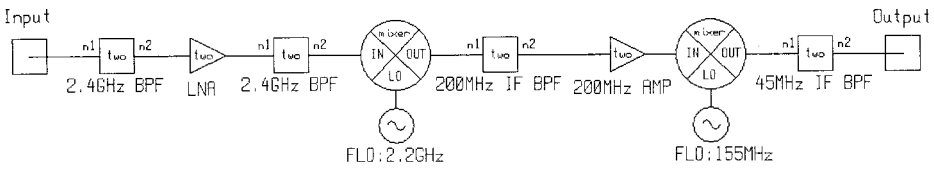


Figure 1-63 Dual-downconversion receiver schematic.

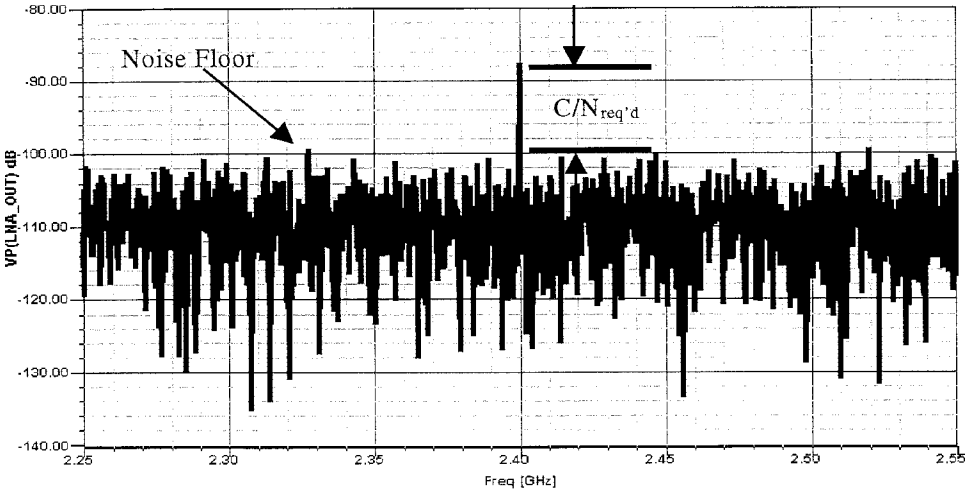


Figure 1-64 Receiver sensitivity measure showing required C/N.

loss, of each component can be determined. A budget calculation (see Table 1-5) can be used to examine the effects of each component on a particular system response.

Another key system parameter is receiver spurious response. The receiver's mixers typically cause the spurious responses. The RF and LO harmonics mix and create spurious responses at the desired IF frequency. These spurious responses can be characterized [4] by the equation

$$\pm mf_{\mathrm{RF}} \pm nf_{\mathrm{LO}} = \pm f_{\mathrm{IF}} \tag{1-16}$$

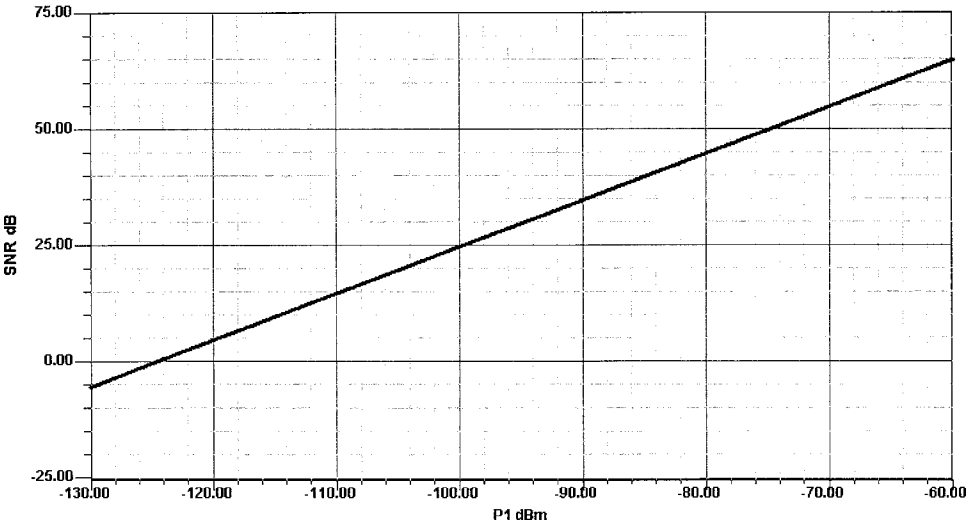


Figure 1-65 C/N ratio versus input RF power level.

Table 1-5 Budget calculation for the ΔS_{21} across each system component accounting for impedance mismatches

Names	ΔS_{21} (dB)
2.4 GHz BPF	–0.82
LNA	22.04
2.4 GHz BPF	–0.93
MIXER_1	–5.34
200 MHz IF BPF	–3.85
200 MHz AMP	22.21
MIXER_2	–8.43
45 MHz IF BPF	–0.42

Some spurious responses, or spurs, can be especially problematic because they may be too close to the intended IF to filter, thus masking the actual information-bearing signal. These spurious responses and their prediction can be especially troublesome if the receiver is operated near saturation. The analysis can then be refined by including a spur table that predicts the spur level relative to the IF signal. The spur's power at the output of a mixer is calculated using the spur table provided for this mixer. Once generated, each spur is carried through the remainder of the system with all mismatches accounted for. The output power level of each spur is shown in Table 1-6. Another useful output is the RF and LO indices, which indicate the origin of each spur.

In addition to sensitivity and spurious response calculations, an analog system can be analyzed in a CAD tool such as Symphony for gain, output power, noise figure, third-order intercept point, intermodulation distortion (IMD; due to multiple carriers), system budget, and dynamic range.

Table 1-6 Spur output powers and the corresponding RF and LO harmonic indices

Frequency (MHz)	P_{out} (dBm)
45.00	–26
20.00	–106
65.00	–86
90.00	–76
110.00	–66
135.00	–96
155.00	–44
175.00	–96
180.00	–106
200.00	–56
220.00	–96
245.00	–76
245.00	–46
265.00	–76

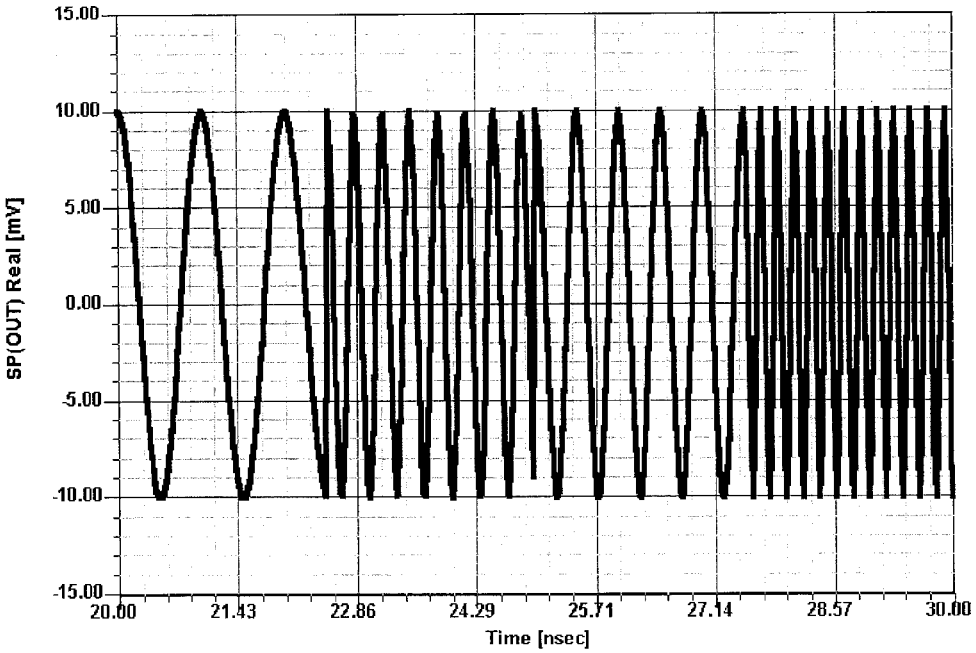


Figure 1-66 4FSK signal generated in Symphony.

Mixed-Mode MFSK Communication System. Next, a mixed-mode (digital and RF) communication system will be described and simulated. In this example digital symbols will be used to modulate an 8-GHz carrier. The system uses *M*-ary frequency shift keying (MFSK) bandpass modulation with a data rate of 40 Mbps. Convolutional coding is employed as a means of forward error correction. The system includes digital signal processing sections as well as RF sections and channel modeling. Several critical system parameters will be examined, including BER. FSK modulation can be described [5] by the equation

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos(\omega_i t + \phi) \tag{1-17}$$

where $i = 1, \ldots, M$. So the frequency term will have M discrete values with almost instantaneous jumps between each frequency value (see Figure 1-66). These rapid jumps between frequencies in an FSK system lead to increased spectral content.

Figure 1-67 shows a block diagram of the complete system. The system can be split into several major functional subsystems—the baseband modulator, the RF transmitter section, channel model, RF receiver, clock recovery circuitry, and baseband demodulator. Looking specifically at the baseband modulator circuitry (Figure 1-68), a pseudorandom bit source is used and the bit rate is set to 40 MHz.

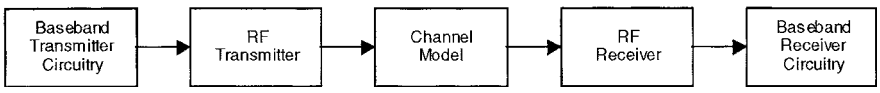


Figure 1-67 Mixed RF/DSP MFSK communication system simulated in Symphony.

A convolutional encoder then produces two coded bits per data bit and increases the bit rate to 80 MHz. The purpose of the convolutional encoder is to add redundancy to improve the received BER. A binary-to-*M*-ary encoder then assigns one symbol to every 2 bits, creating the four levels for the 4FSK, effectively halving the bit rate down to 40 MHz again. The signal is then scaled and upsampled. To decrease the bandwidth, a root-raised cosine filter is used to shape the pulses. The filtered signal then serves as the input to the frequency modulator. The RF section (see Figure 1-69) includes the transmitter that modulates the baseband signal onto an 8-GHz carrier.

That signal is amplified and then filtered to remove any harmonics. The signal is then passed through an additive white Gaussian noise model. The received signal is filtered, amplified, and downconverted twice to baseband. After carrier demodulation the signal is then sent through the clock recovery circuitry. Clock recovery is employed to ensure that sampling of the received signal is executed at the correct instances. This recovered timing information is then used as a clock signal for sample and hold circuitry. Clock recovery in this system was achieved using a phase-locked loop (PLL) configuration. The schematic for the clock recovery circuit is shown in Figure 1-70.

At the heart of the clock recovery circuit is the phase comparator element and the frequency modulator. The inputs to the phase comparator consist of a sample of the received data signal and the output of the feedback path that contains the frequency modulator. The frequency modulator reacts to phase differences in its own carrier and the received data signal. The output of the frequency modulator is fed back into the phase comparator, whose output is dependent on the phase differential at its inputs. The phase of the frequency modulator continually reacts to the output of the phase comparator and eventually lock is achieved. Once the timing of the received signal is locked onto, the clock that feeds the sample and hold will be properly aligned and correct signal sampling will be assured.

Figure 1-71 shows the received signal (filtered and unfiltered) as well as the recovered clock and the final data. Equalizer circuitry is then used to compensate for channel effects, which degrade the transmitted signal. The equalizer acts to undo or adapt the receiver to the effects of the channel. The equalizer employed in this MFSK system is the recursive least squares equalizer. The equalizer consists of a filter of *N* taps that undergoes an optimization in order to compensate for the channel effects. The equalizer depends on a known training sequence in order to adapt itself to the channel. The equalizer model updates the filter coefficients based on the input signal and the error signal (i.e., the difference between the output of the equalizer and the actual desired output). The update (optimization) is based on the recursive least squares algorithm. Several equalizers are available in Symphony, including complex least mean square equalizer, complex recursive least squares equalizer, least mean square equalizer, recursive least squares equalizer, and the Viterbi equalizer. After equalization, the BER of the system is analyzed versus SNR (see Figure 1-72).

PLL CAD SIMULATION. The Symphony system simulator also allows us to evaluate the performance of phase-locked loops. Figure 1-73 shows the block diagram of a PLL in which the VCO is synchronized against a reference. For the purpose of demonstrating the capability, we have selected a 1:1 loop with a crossover point of about 100 kHz, meaning that the loop gain is 1 at this frequency.

Figure 1-74 shows the system's simulated open- and closed-loop phase-noise performance. At the crossover point, the loop is running out of gain. The reason why the closed-loop noise increases above 2 kHz has to do with the noise contribution of various components of the loop system. The highest improvement occurs at 100 Hz, and as the loop gain decreases,

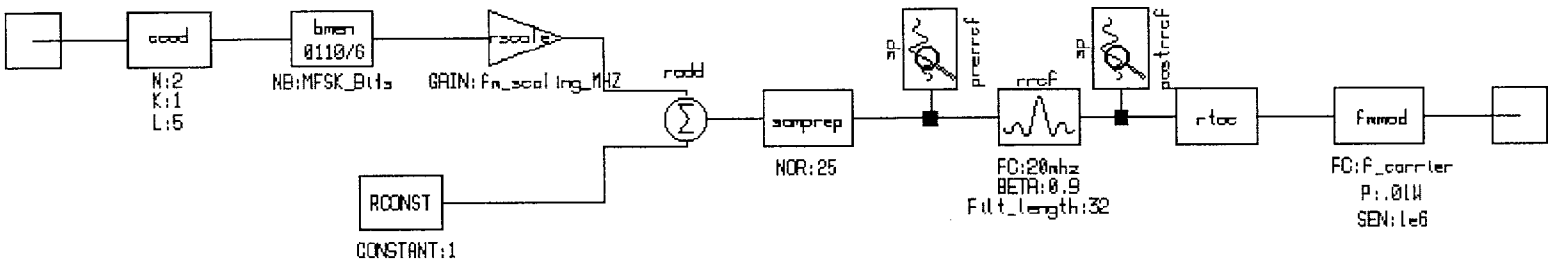


Figure 1-68 MFSK baseband circuitry including frequency modulator.

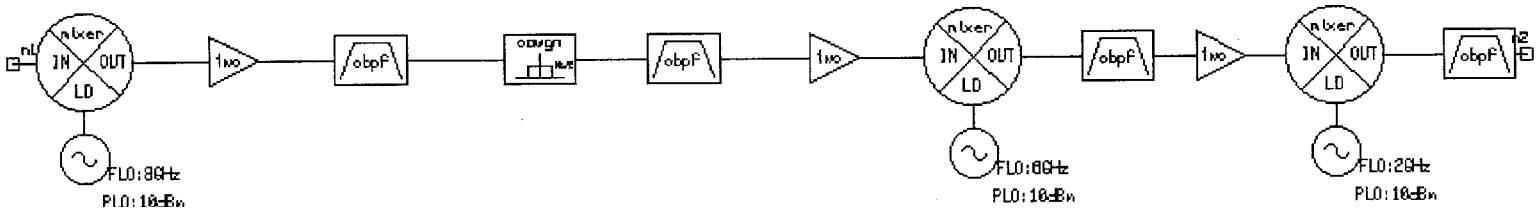


Figure 1-69 RF section including Gaussian noise channel model.

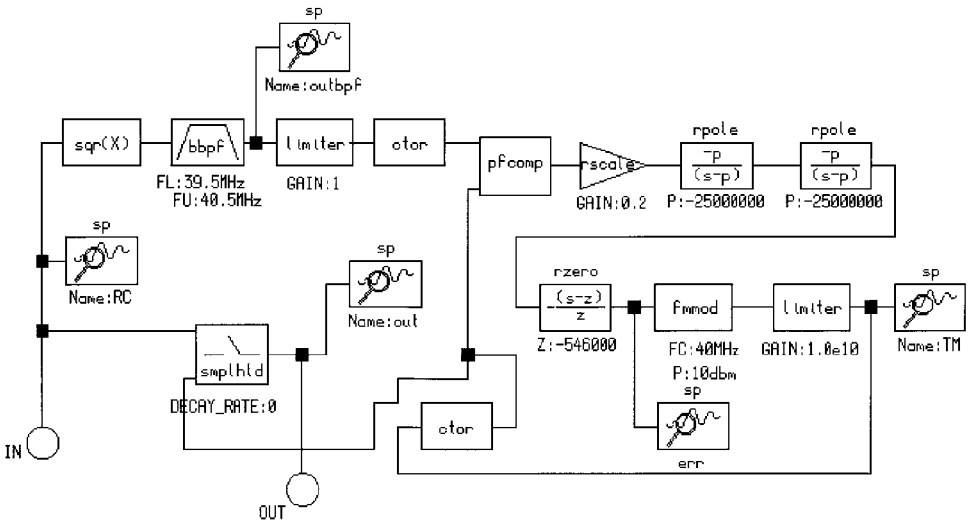


Figure 1-70 Clock recovery circuitry.

the improvement vanishes. Therefore it is desirable to make the loop bandwidth as wide as possible, as this also improves the loop's switching speed. On the other hand, as the phase noise of the free-running oscillator crosses over the phase noise of the reference noise, divider noise, and other noise contributors, one can make the noise actually worse than that of the free-running state. In a single-loop system, a compromise must be made between phase noise, switching speed, and bandwidth. A first-order approximation for switching speed is $2/f_L$, where f_L is the loop bandwidth. Assuming a loop bandwidth of 100 kHz, the switching speed

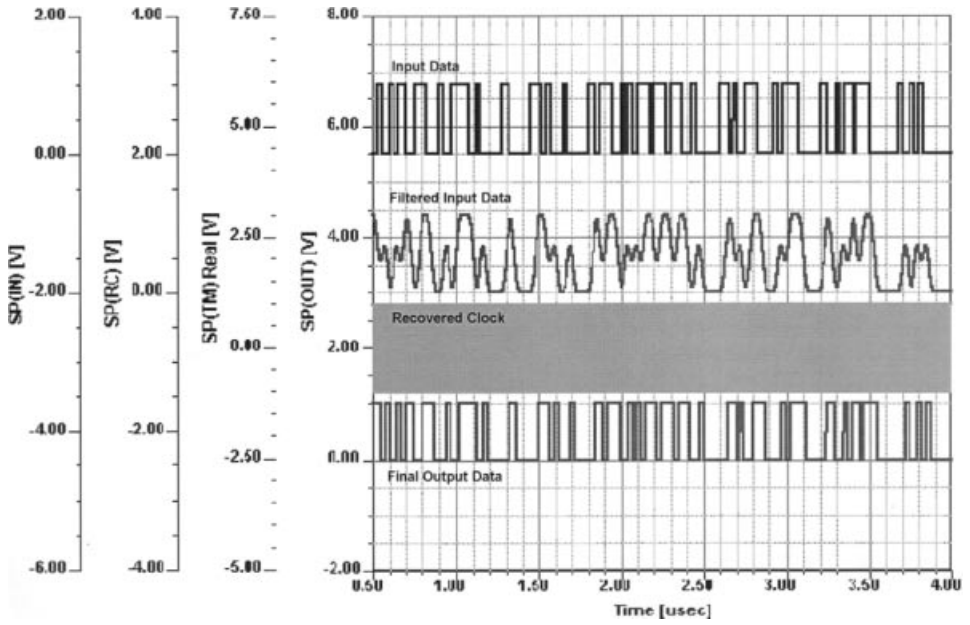


Figure 1-71 Input data, filtered input data, recovered clock, and final output data for the MFSK communication system.

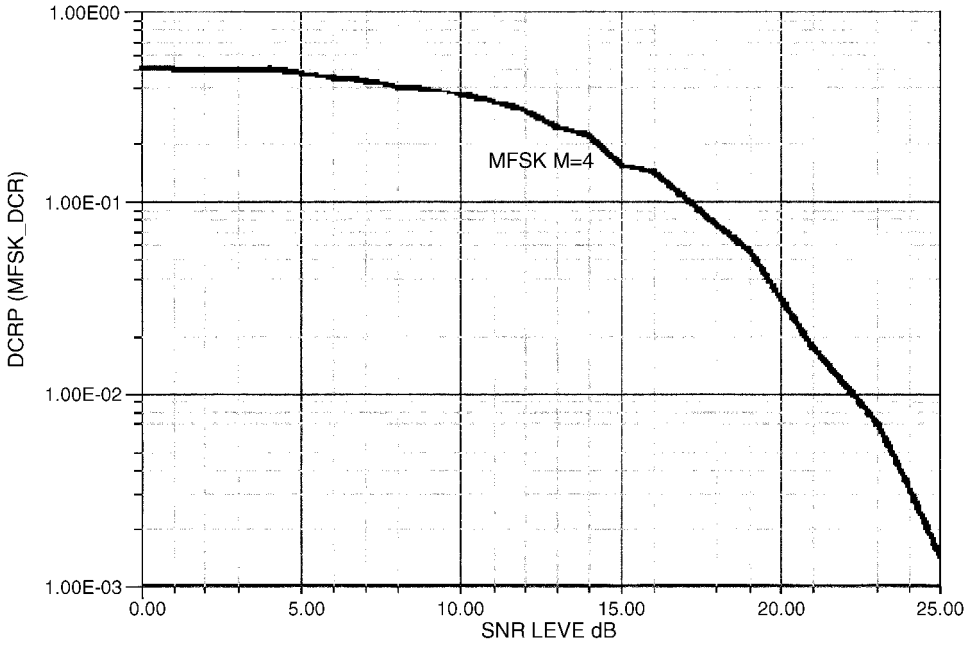


Figure 1-72 BER versus SNR for the MFSK system.

will be 20 μs. Looking at Figure 1-74, we can clearly see the trend from 200 Hz to 100 kHz. Because of the resolution of the sampling time (computation time), the open-loop phase noise below 150 Hz is too low and could be corrected by a straight-line extrapolation from 500 Hz toward 100 Hz. We have left the drawing uncorrected to show the effect of not-quite-adequate resolution.

Since we have already referenced analog and digital waveforms, we need to mention that the current system in place worldwide is referenced as the second-generation (2G) system.

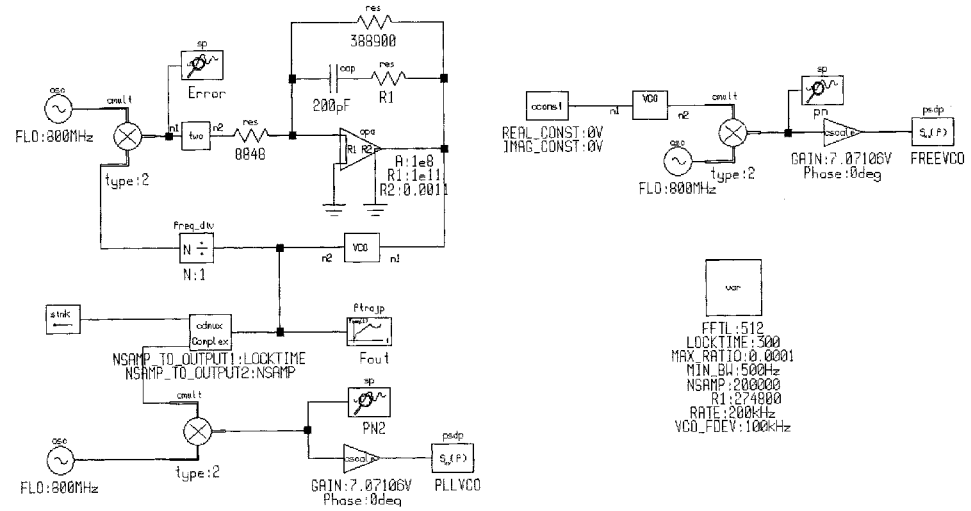


Figure 1-73 Block diagram of a CAD-based PLL system.

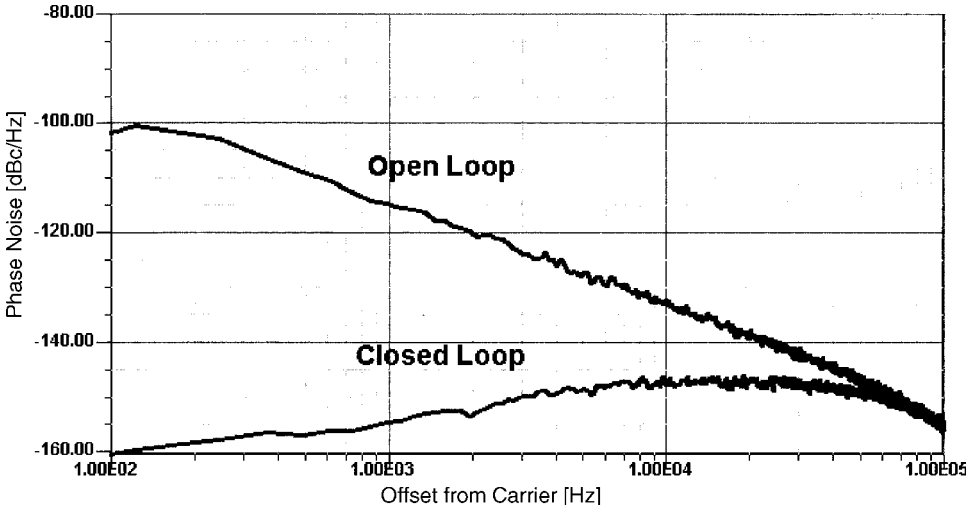


Figure 1-74 Open- and closed-loop phase noise for a CAD-based test phase-locked loop.

Table 1-7 Cellular and cordless standards

Cellular Telephone Networks	
Analog	Digital
AMPS	**IS-54**
Advanced Mobile Phone System	North American Digital Cellular
TACS	**IS-95**
Total Access Communication System	North American Digital Cellular
NMT	**GSM**
Nordic Mobile Telephone	Global System for Mobile Communications
	PDC
	Personal Digital Cellular
Cordless Telephone Networks	
Analog	Digital/PCN
CTO	**CT2/CT2+**
Cordless Telecom 0	Cordless Telecom 2
JCT	**DECT**
Japanese Cordless Telecom	Digital European Cordless Telecom
CT1/CT1+	**PHS**
Cordless Telecom 1	Personal Handy Phone System
	DCS-1800

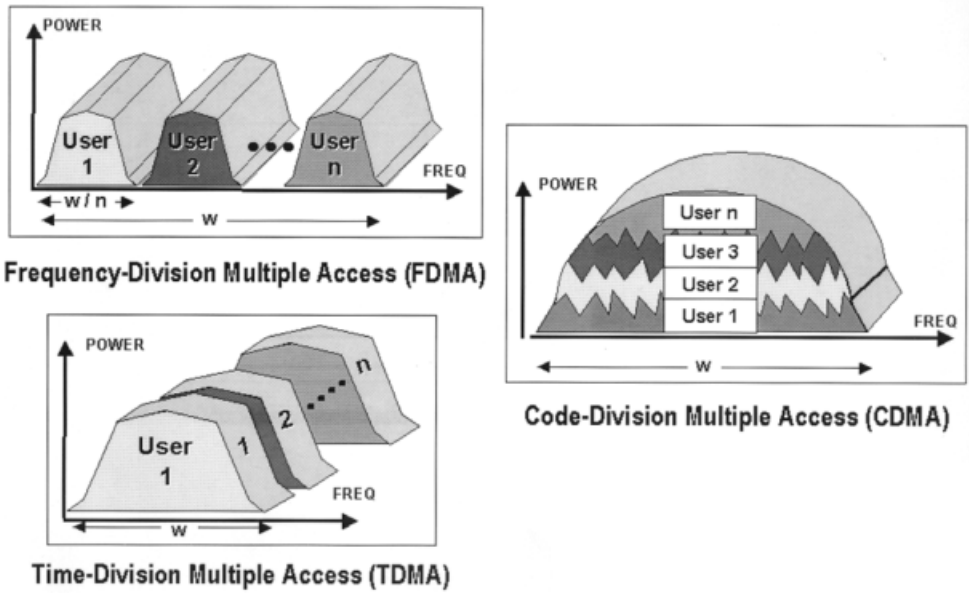


Figure 1-75 Multiplexing schemes.

Table 1-8 Parameters of cellular and cordless systems

	Cellular Telephones			Cordless Telephones		
Characteristics	IS-54	IS-95	GSM	CT2	DCS1800	DECT
Band (downlink)	869-894	869-894	935-960	864-868	1805-1880	1880-1900
Band (uplink)	824-849	824-849	890-915	864-868	1710-1785	1880-1900
Bandwidth	50 MHz	50MHz	50MHz	2 MHz	150 MHz	20 MHz
Channelization	TDMA/ FDMA	CDMA/ FDMA	TDMA/ FDMA	FDMA	TDMA/ FDM	TDMA
Channel spacing	30 kHz	1250 kHz	200 kHz	100 kHz	200 kHz	1728 kHz
Channels/carrier	3	55–62?	8	1	16	12
Number of channels	832 (3 users/ch.)	20 (798 users/ch.)	124 (8 users/ch.)	40	750 (16 users/ch.)	10 (12 users/ch.)
Duplex method	FDD	FDD	FDD	TDD	FDD	TDD
Channel bit rate	48.6 kbps	1.2288 Mbps	271 kbps	72 kbps	271 kbps	1152 kbps
Speech codec	VSELP	CELP	RPE-LTP	ADPCM	RPE-LTP	ADPCM
Bit rate (voice)	8 kbps	1.2–9.6 kbps	13 kbps	32 kbps	13 kbps	32 kbps
Modulation	π/4-DQPSK	BPSK/ OQPSK	GMSK	FSK	GMSK	GMSK
Mobile peak power	0.6–3 W	0.2–2 W	2–20 W	10 mW	0.25–2 W	250 mW
Mobile average power	0.6–3 W	0.2–2 W	0.25–2.5 W	5 mW	0.03–0.25 W	10 mW
Cell radius	30 miles	30 miles	1–5 miles	?	?	40–140 m
Cluster size (min)	7	1	3	N/A	3	N/A

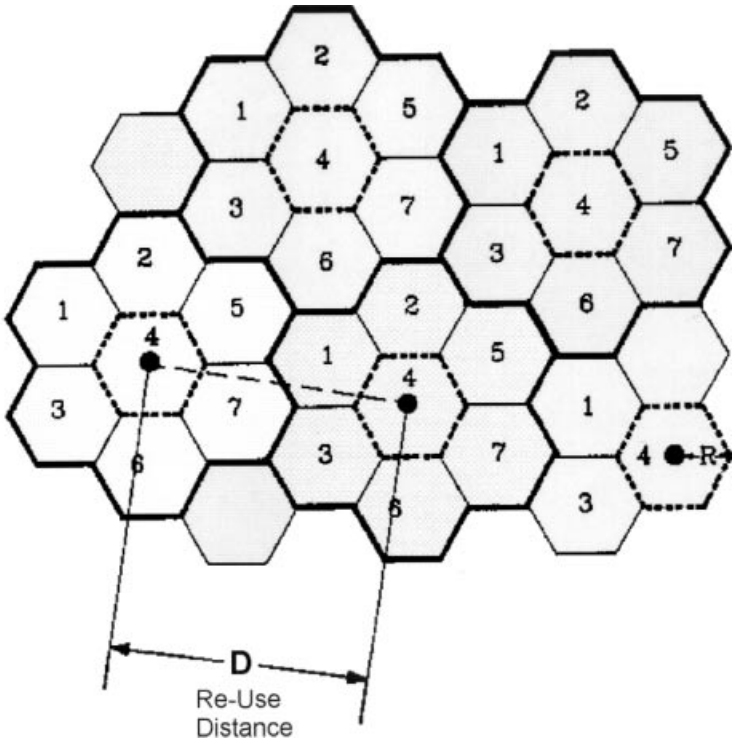


Figure 1-76 Cellular structure. $W = N \times B_C$, where W is the full bandwidth and B_C is the bandwidth per cell. The number of channels/cell = B_C/B_U, where B_U is bandwidth per user (FDMA) or N users (TDMA). $D = R\sqrt{3N}$.

The all-digital, adaptive-bandwidth third-generation (3G) system is being discussed in Europe, Japan, and the United States [6]. This is not without headaches. The Groupe Speciale Mobile (GSM) system in Europe still has potential to grow and is not under any urgent pressure to be upgraded. The opposite is true in Japan, because Japan is running out of capacity and needs to go to wideband CDMA. The important perceived advantage of this is that the number of simultaneous users can be extremely high: there is much less interference because of orthogonal modulation types with extremely low correlation and adaptive bandwidth. The actual definitions are still in flux, with the discussions dominated by the United States. This domination stems from the National Security Agency being determined to be able to eavesdrop on conversations that are none of their business (drug smuggling is always used as an excuse) but invites invasion of monitoring without proper court authorization. This is a very hot issue. The other similar problem has to do with patents. This type of technology is somewhat dominated by Qualcomm, a U.S. company, which likes to send a bill for patent infringement to everybody, possibly even to users. A similar speculation took place with the introduction of digital audio tapes and their duplication and is occurring nowadays with CDs carrying either music recordings or movies. This issue is driven by money and not by technology.

That said, Table 1-7 shows cellular and cordless standards, everything referring to 2G. Figure 1-75 shows the multiplexing schemes used. Table 1-8 shows the parameters of cellular and cordless systems. Figure 1-76 shows the cellular structure.

1-5-2 Transmitters

Since it is possible to generate any type of modulation using DSP, including those described in this chapter, a DSP-based transmitter can also be built. A good example of this is the Philips SA900, which is truly universal.

Introduction to the SA900.* The SA900 (Figure 1-77) is a truly universal in-phase and quadrature (I/Q) radio transmitter that can perform many types of analog and digital modulation including AM, FM, SSB, QAM, BPSK, QPSK, and FSK. It is a highly integrated system that saves space and cost for the manufacturers producing cellular and wireless products. The device allows baseband signals to directly modulate the I/Q carriers, which are generated by an internal phase shift network, in the 1-GHz range, and to maintain good linearity required for a linear modulation scheme (e.g., $\pi/4$-DQPSK). It contains an on-chip frequency divider, phase detector, and VCO, which can be built into a phase-locked loop (PLL) frequency synthesizer to create a transmit offset frequency. Its unique internal design allows frequency conversion without having an external image rejection filter for eliminating the sum term after mixing. The SA900 meets the specifications required by the IS-54, the industry standard for North America Digital Cellular (NADC) system. Here we review the basic concept of I/Q modulation and discuss the key points when designing the SA900 for an RF transmitter.

I/Q Modulation. Any bandpass RF signals can be represented in polar form by

$$s(t) = A(t)\cos[\omega_c t + \phi(t)] \tag{1-18}$$

where $A(t)$ is the signal envelope and $\phi(t)$ is the phase. By using the trigonometric identities, we can represent Eq. (1-18) in rectangular form by

$$s(t) = \mathrm{I}(t)\cos(\omega_c t) - \mathrm{Q}(t)\sin(\omega_c t) \tag{1-19}$$

where $\mathrm{I}(t) = A(t)\cos[\phi(t)]$
$\mathrm{Q}(t) = A(t)\sin[\phi(t)]$

Since the baseband signals $\mathrm{I}(t)$ and $\mathrm{Q}(t)$ modulate two exactly 90° out-of-phase carriers $\cos(\omega_c t)$ and $-\sin(\omega_c t)$, respectively, we call the system implementing Eq. (1-19) an in-phase and quadrature (I/Q) modulator. Figure 1-78 shows the mathematics and hardware implementation of an I/Q modulator.

The local oscillator, usually a VCO within a PLL, generates the carrier and is split into two equal signals. One goes directly into a double-balanced mixer to form the I-channel and the other one goes into the other mixer via a 90° phase shifter (realized by passive elements) to provide the Q-channel. The baseband signals $\mathrm{I}(t)$ and $\mathrm{Q}(t)$, either analog or digital in nature, modulate the carrier to produce the I and Q components, which are finally combined to form the desired RF transmitting signal. Since any RF signal can be represented in the I/Q form, any modulation scheme can be implemented by an I/Q modulator.

*Based on portions of the Philips Semiconductors/Signetics RF Communications Products Application Note AN1892, "SA900 I/Q Transmit Modulator for 1 GHz Applications," August 20, 1997. Used with permission.

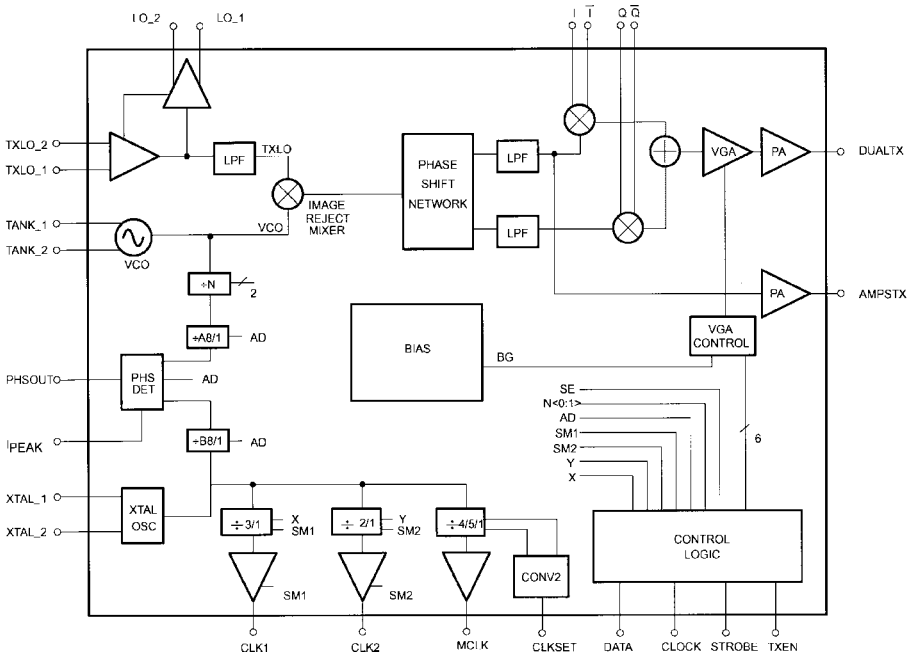


Figure 1-77 SA900 transmit modulator.

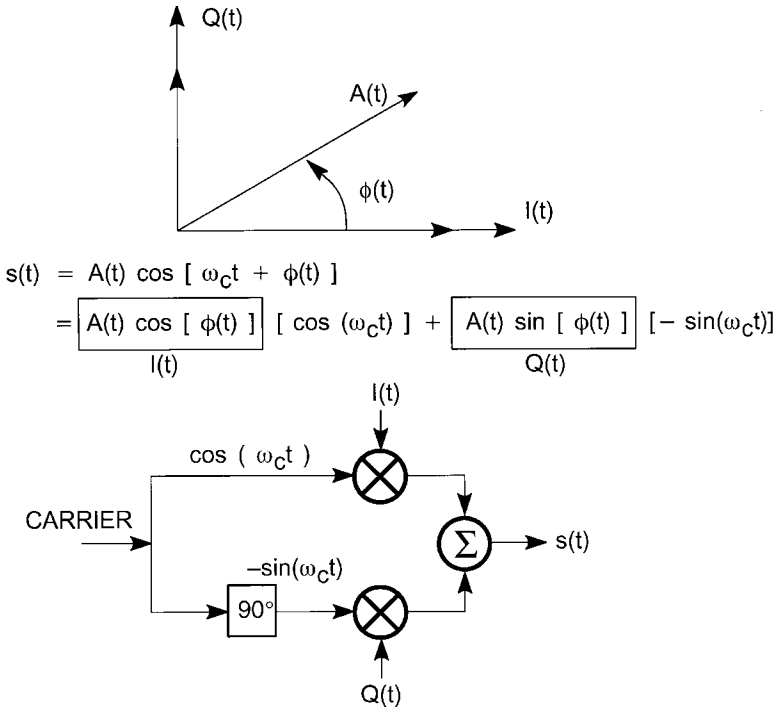


Figure 1-78 Mathematical representation and hardware implementation of I/Q modulator.

Linear Digital Modulation. Linear digital modulation techniques depend on varying the phase and/or magnitude of an analog carrier according to digital information: ones and zeros. This digital information can be the output of an analog-to-digital converter (e.g., voice codec), or it can be digital data in some standard formats (e.g., ASCII). The most popular digital signaling format is nonreturn-to-zero (NRZ), where 1's and 0's are converted into signals with amplitudes of 1 and –1, respectively, in a symbol duration. Since an NRZ signal has infinite bandwidth, transmit filters have to be used to limit the spectral spreading. To ensure each NRZ symbol does not smear into its neighbors due to low-pass filtering and channel distortion, causing intersymbol interference (ISI), the frequency response of the low-pass filter has to satisfy Nyquist criteria. One example of this type of filter is the linear phase square-root-raised cosine filter. Together with the same type of filter for receive low-pass filtering, the signal is guaranteed ISI-free in a Gaussian environment. One straightforward technique of transmitting such a band-limited signals through communication channels would be to apply it directly to the mixer of the I channel to generate an RF signal. This is known as binary phase shift keying (BPSK), where the phase of the carrier is shifted 180(to transmit a data change from 0 to 1 or 1 to 0.

Quadrature or quaternary phase shift keying (QPSK) is a much more common type of modulation scheme used in mobile and satellite communications. It has four possible states (90° apart) and each of them represents 2 bits of data. Figure 1-79 shows the baseband generator for QPSK (without the differential phase encoder). NRZ data bits go through the serial-to-parallel converter (see Figure 1-80) and are mapped in accordance to some rules to generate I and Q values. The generic rule states that the values of I and Q components are 1 and 1 for the data bits "11" (45°) and –1 and –1 for the data bits "00" (–135°). These discrete signals have to be band-limited by Nyquist low-pass filters to be ISI-free.

A more sophisticated way of mapping results in π/4-DQPSK (D for differential encoding), which is chosen for North America Digital Cellular (IS-54), Personal Digital Cellular (PDC) in Japan, and Personal Handy Phone System (PHS) in Japan. In this scheme, consecutive pairs of bits are encoded into one of the four possible phases: π/4 for "11," 3π/4 for "01," –3π/4 for "00," and –π/4 for "10." However, unlike the previous case that "11" is always π/4 and "00" is always –3π/4, the encoded phases are the degrees that the carrier has to shift at each sampling instances. Thus the information is contained in the phase difference (differential) instead of absolute phase for π/4-DQPSK.

A better way to tell the difference between QPSK and π/4-DQPSK is by looking at the signal constellation diagram, shown in Figure 1-81, which displays the possible values of I and Q vectors and changes of state. A constellation diagram is also known as a phase diagram because it shows the phase of the carrier at the sampling point. Note that the phases of QPSK are assigned for every 2 bits of data; therefore it can transmit twice as much information as BPSK in a given bandwidth; that is, it is more bandwidth-efficient. The 8-PSK is another type of modulation used for high-efficiency requirements. It maps 3 bits into eight phases, 45° apart, in the constellation. More spectrally efficient modulation can be created by

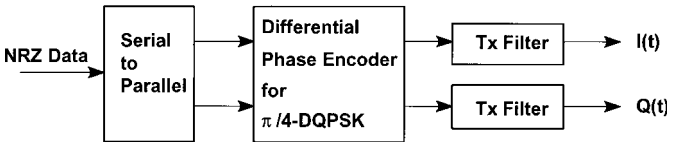


Figure 1-79 QPSK and π/4-DQPSK baseband generator.

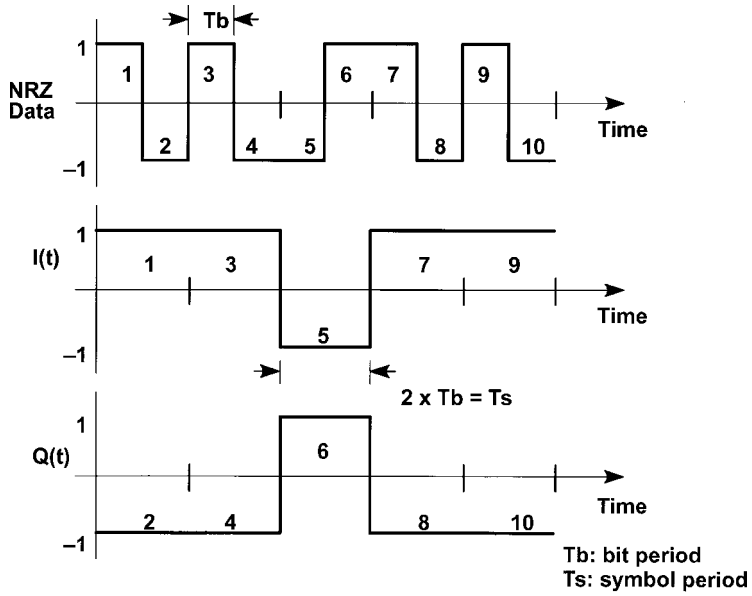


Figure 1-80 Serial-to-parallel conversion.

mapping more bits into one phase at each sampling point. However, as we add states to the signal constellation, the signal susceptibility to noise increases because the difference between states decreases. Then, a higher carrier-to-noise (C/N) ratio is required to maintain the same bit error rate.

One common misconception is that since π/4-DQPSK has eight states, it is just another type of 8-PSK. Note, however, that at every sampling instant, the carrier of π/4-DQPSK is allowed to switch to only one of the four possible states (see Figure 1-81). So we still have two data bits that get encoded into four phases. Thus π/4-DQPSK has the same spectral efficiency as QPSK for the same carrier power. The reason for using this modulation scheme is twofold. First, the envelope fluctuation, which causes spectral spreading due to nonlinearity of transmitter and amplifier, is reduced because the maximum phase shift is 135° instead

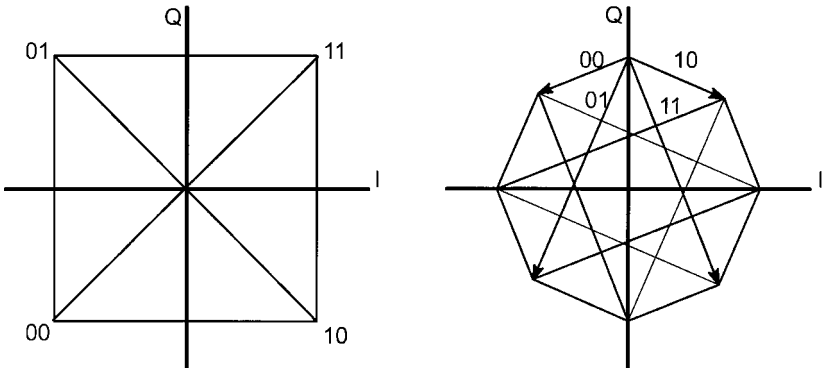


Figure 1-81 Signal constellation of QPSK and π/4-DQPSK.

of 180°. Second, the signal can be demodulated noncoherently, which simplifies the receiver circuitry by eliminating the need for carrier recovery.

Digital and Analog FM. Another family of digital modulation is categorized by frequency change of the carrier instead of phase and/or amplitude change. One member of this family is frequency shift keying (FSK), where the carrier switches between two frequencies. FSK is also known as digital FM because it can be generated by feeding the NRZ data stream into an analog VCO. FSK appears as a unit circle in the signal constellation because the RF signal envelope is constant and the phase is continuous. Baseband filtering is usually applied for FSK to limit the RF bandwidth of the signal so that more channels can fit into a given frequency band.

One common modulation of this type is known as Gaussian minimum shift keying (GMSK), which is used for GSM and some other wireless applications. GMSK can be generated by following its definition: band-limit the NRZ data stream by a Gaussian low-pass filter, then modulate a VCO with modulation index (2 × frequency deviation/bit rate) set to 0.5. In other words, the single-sided frequency deviation is one-fourth of the bit rate ($\Delta f = R/4$).

Another way of generating GMSK is by I/Q modulator. Referring back to Eq. (1-19), any RF signal can be split into I and Q components. Unlike the QPSK mentioned before, baseband I(t) and Q(t) are not discrete points for FM signals; rather, they are continuous functions of time. The way to produce FM is shown in Figure 1-82. We first store all the possible values of $\cos[\phi(t)]$ and $\sin[\phi(t)]$ in a ROM lookup table, which will be addressed by the incoming data to generate the I and Q samples. The output data from the ROM is then applied to D/A converters, after low-pass filtering for signal smoothing, to produce the analog baseband I and Q signals. This method guarantees the modulation index to be exactly 0.5, which is required for coherent detection of GMSK (e.g., the GSM system). The same I/Q principle can also be applied to generating analog FM signals.

Single-Sideband AM (SSB AM). AM signals can be divided into three types: the conventional AM, double sideband suppressed carrier AM (DSBSC AM), and SSBSC AM. The first type is not attractive because for 100% modulation, two-thirds of the transmit signal power appears in the carrier, which itself conveys none of the information added by modulation. By using a balanced mixer (e.g., a Gilbert cell), one can generate DSBSC AM, where the carrier is totally suppressed and only the upper and lower sidebands are present. However, this is still not the best because the information is transmitted twice, once in each sideband. To further increase the efficiency of transmission, only one sideband is needed to deliver the information. The SSBSC AM can be generated by an I/Q modulator with the baseband information feeding the modulator (by quadrature), as shown in Figure 1-83. This modulation technique can greatly reduce the bandwidth of the signal and allows more signals to be transmitted in a given frequency band. This topic is discussed in detail in [7].

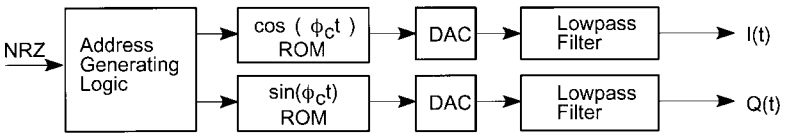


Figure 1-82 Digital FM (e.g., GMSK) baseband I/Q generator.

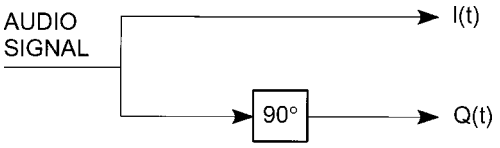


Figure 1-83 Baseband processing for SSBSC AM.

System Architecture. There are usually two schemes, the dual conversion and direct conversion, used for implementing transmit modulators. Dual conversion is simpler to implement by modulating an oscillator at lower frequency and then upconverting to the carrier frequency. This scheme, however, is more expensive due to the need for additional filtering and more printed circuit (PC) board space. By using only one mixer, direct conversion requires fewer components but is harder to implement.

The problems that direct conversion suffers are carrier leakage and modulated signal coupling. Poor RF isolation of the surface mount packages will allow the carrier to be present at the transmitter output, thus making it difficult to have –40 dBc carrier suppression. In addition, modulated RF signal would couple back to the oscillator (usually a VCO in a PLL synthesizer loop) and cause modulation distortion.

Based on the concept of dual conversion, the SA900 uses an image-reject mixer to eliminate the need for IF filtering and allow monolithic integration. The transmit carrier (LO) is downconverted by the frequency synthesized by the on-chip VCO, which operates from 90 to 140 MHz. This LO is then modulated by the baseband I/Q signals to obtain a complex modulation scheme. The image (sum term) after mixing and LO are sufficiently suppressed by the image-reject mixer. Any residual amounts can be further suppressed by an external duplex filter.

Figures 1-84 and 1-85, respectively, show how the SA900 can be used in frequency-division duplex (FDD) and time-division duplex (TDD) transceivers. Note that the LO for both systems is running at a frequency that is higher than the transmit frequency, thus minimizing carrier leakage. In the FDD system only one external VCO is required for generating both transmit and receive LO signals when using the SA900.

Figure 1-86 shows the IS-54 front-end chipset that consists of the SA601, SA7025, SA900, and SA637. This receiver architecture (SA637) supports a digital magnitude/phase baseband

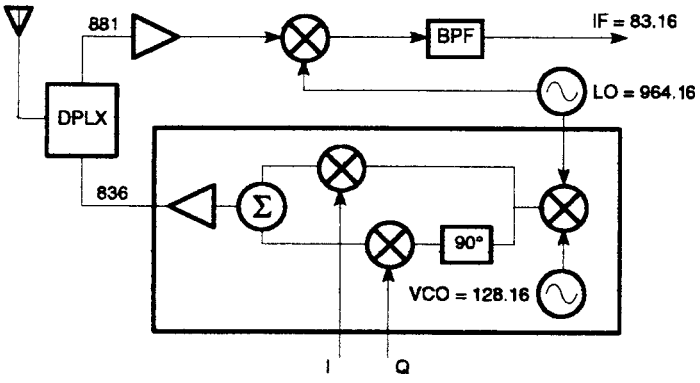


Figure 1-84 FDD system using SA900.

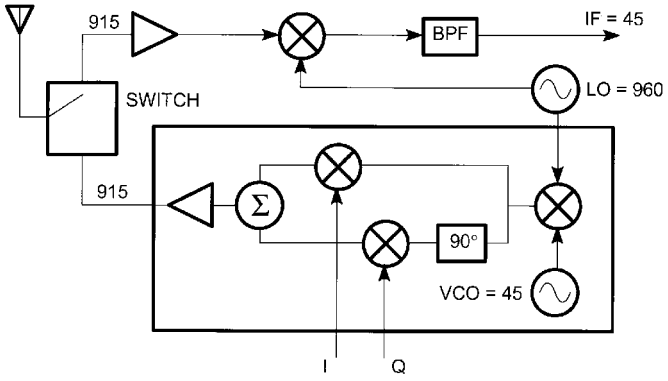


Figure 1-85 TDD system using SA900.

demodulator. An alternate configuration will be using the SA606 FM/IF receiver in conjunction with an external I/Q demodulator IC. Table 1-9 shows the possible configurations for the IS-54 handsets using the SA900 as transmitters.

Designing with the SA900

Baseband I/Q Inputs. The baseband modulation inputs are designed to be driven differentially for the SA900 to operate at its best. The I and Q inputs should have a dc offset of $V_{CC}/2$, which is externally provided by common DSP chips. If all four inputs are biased from the same source, the device can tolerate ±0.5-V dc error; however, inaccuracy of dc bias between I1/I2 or Q1/Q2 causes reduced suppression of the carrier. Thus, it is important to have a well-regulated dc supply for I and Q signal biasing. The bandwidth of the inputs is much higher than the specified 2 MHz. Approximately 2 dB of power loss will be experienced if the I and Q inputs are 50 MHz.

The SA900 generates a minimum of 0 dBm of power to a 50-Ω load when the amplitude of the I and Q signals are 400 mV P-P. The output power will decrease by 6 dB for every 50% decrease in I/Q amplitude. Single-ended I and Q sources can be used but are not recommended due to the degradation in carrier suppression (more than 10 dB compared to differential). In addition, the entire noise performance of the device will suffer. $V_{CC}/2$ should be applied to I2 and Q2 pins if the part is driven single-endedly.

Transmit Local Oscillator. The transmit local oscillator path consists of a TXLO input buffer, LO output buffer, VCO, image-reject mixer, and phase shift network. Together with a few external components, this section provides the I and Q carriers for modulation.

The TXLO inputs and LO outputs are designed to be used in an external PLL that synthesizes different frequencies for channel selection. The RF signal being generated is fed into TXLO inputs and then comes out of LO outputs to complete the system synthesizer loop. The TXLO inputs are differential in nature and have a voltage standing wave ratio (VSWR) of 2:1 with input impedance of 50 Ω. Single-ended sources can be used by ac grounding the TXLO_2 input, as done on the demo board. This signal should also be ac coupled into the TXLO_1 input. The frequency range for these inputs is from 900 to 1040 MHz while the input power should be between −10 and −13 dBm. The output level will be changed significantly if the input level is below −25 dBm.

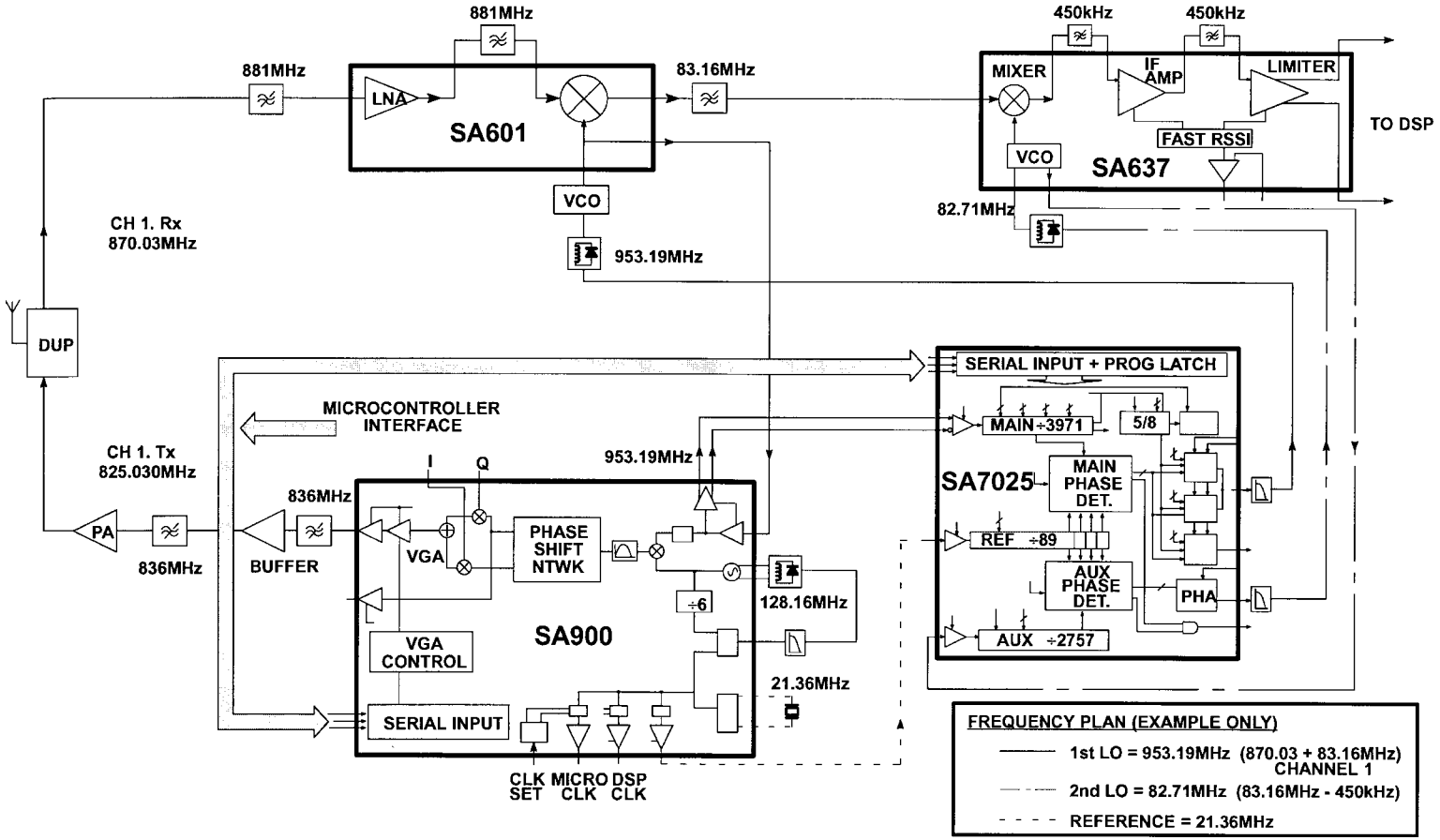


Figure 1-86 IS-54 front-end chipset from Philips.

Table 1-9 Possible configurations for IS-54 handsets

Rx First IF	On-Chip VCO Frequency	On-Chip ÷ *N* Value	Crystal Frequency
83.16 MHz	128.16 MHz	6	21.36 MHz
71.64 MHz	116.64 MHz	6	19.44 MHz
45 MHz	90 MHz	6	15 MHz
84.6 MHz	129.6 MHz	9	14.4 MHz

The output power of the LO buffered signal changes by about 2 dB when the SA900 is in a different mode of operation. Typical values are –13.5 dBm and –15.5 dBm for DUAL mode and STANDBY mode, respectively.

The 90° phase shift network, realized by resistor–capacitor (RC) networks, is capable of operating over a wide frequency range. Even though their frequency characteristics are optimized for cellular band, the part can also be used in other applications in a different band. In such cases, designers have to test the part experimentally to find out the performance, such as sideband suppression, carrier suppression, and image rejection.

Crystal Oscillator. The crystal oscillator (XTAL_1 and XTAL_2 pins) is used to provide a reference frequency between 10 and 45 MHz for the phase detector and the three on-chip clocks. It can be configured as a crystal oscillator using an external crystal and capacitors, or it can be driven by an external source. In the latter case, pin XTAL_2 can be left floating. Information regarding crystal oscillator design can be found in [8].

VCO. The VCO, together with the phase detector, the divider, and an external low-pass filter, can form a PLL for the transmit offset frequency. The image-reject mixer downconverts the TXLO signal to the RF carrier by the amount of VCO frequency. Thus the TXLO frequency should be the desired channel frequency plus the IF offset generated by the VCO. Note that the part will not function if the VCO section is not used.

The VCO is designed for generating IF frequency between 90 and 140 MHz. Together with an external varactor diode and resonator, it can be configured as an oscillator as shown in Figure 1-87.

The resonant frequency of such a circuit is

$$f_{\text{VCO}} = \frac{1}{2\pi\sqrt{LC_T}} \tag{1-20}$$

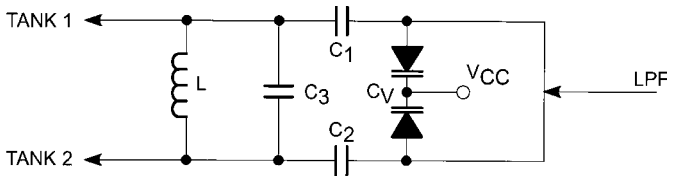


Figure 1-87 VCO tank circuit.

where $C_T = (C_1 \parallel C_2 \parallel C_V) + C_3$. C_V is a varactor diode, which changes capacitance as the voltage across it varies.

Calculation

$$C_1 = C_2 = 33 \text{ pF}$$

$$C_3 = 5.6 \text{ pF}$$

$$C_V = 33.5 \text{ pF @ } 2.5 \text{ V}$$

$$L = 100 \text{ nH}$$

$$C_T = 5.6 + (1/33 + 1/33 + 1/33.5)^{-1} = 16.7 \text{ pF}$$

$$f_{\text{VCO}} = \frac{1}{2\pi(100\text{E} - 9 \times 16.7\text{E} - 12)^{0.5}} = 123\text{MHz}$$

On the demo board, a 1:1 ratio RF transformer is also included to allow single-ended external source driving differential inputs when the VCO is not used.

When designing the VCO, careful printed circuit board (PCB) layout has to be made. Traces have to be short to avoid the parasitic capacitance and inductance that may cause unwanted oscillation. Referring to Eq. (1-20), there are a large number of combinations of L and C_T values that will give the same resonant frequency. If undesired spurs are found in the design due to PCB layout, experimenting with a different set of LC values may sometimes solve the problem.

Output Impedance Matching. The equivalent output impedance at the DUALTX pin is approximately equal to 600 Ω in parallel with 2 pF at 830 MHz. It has to be matched properly to generate maximum power into a 50-Ω load (e.g., a SAW filter). Figure 1-88 shows the recommended matching network. The shunt inductor (L_1) is used to provide maximum swing at the output (short at dc) and also provide reactance to make the real impedance 50 Ω looking

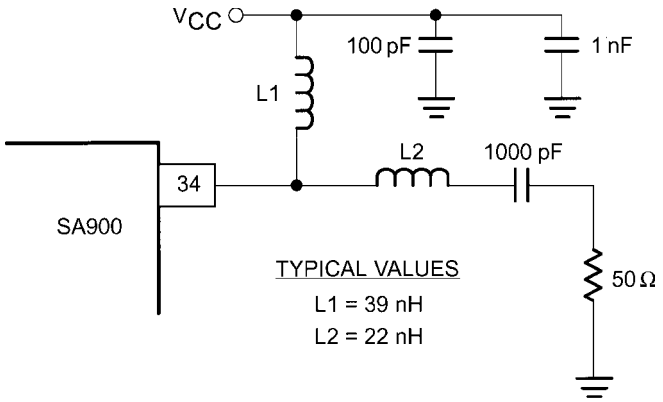


Figure 1-88 DUALTX output matching network.

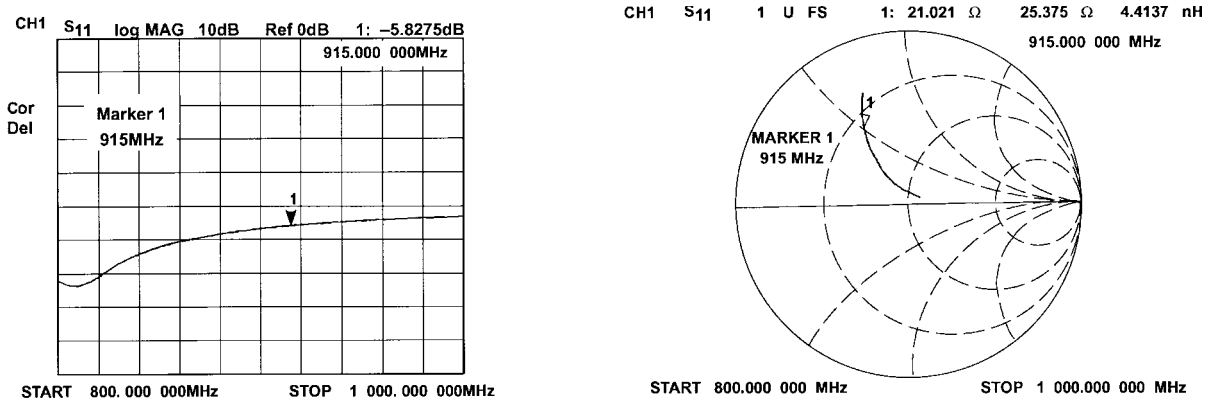
into the matching network. The remaining negative reactance is canceled by the series inductor (L_2). The values used on the demo board can be used as a reference but may not be suitable if a different layout is implemented. The two shunt capacitors are included to bypass the high-frequency RF signal, avoiding direct coupling into V_{CC}. The series ac coupling capacitor is used to maintain the proper bias for the output stage. Their values are big enough to be left out in the impedance matching calculation.

Using a network analyzer to measure the SA900's output *S* characteristic is necessary for obtaining optimum matching, which generates maximum output power. Figure 1-89 shows how to match the output impedance to a 50-Ω load at 915 MHz. First, calibrate the network analyzer to the DUALTX SMA connector on the demo board. Then, short the point where the series inductor is located and use the DELAY feature of the network analyzer to move the point of reference in the Smith Chart to the leftmost point. Now the network analyzer is calibrated to the beginning of the matching network, not just the SMA connector. The frequency response (Figure 1-89a) shows that the "dip" is around 830 MHz, the frequency where the board was originally matched. The Smith Chart shows that it requires less inductance to bring the marker to the center of the chart (50 Ω). By using a 15-nH series inductor, the "dip" was moved closer to 915 MHz (–15 dB) and a better match is achieved (Figure 1-89b).

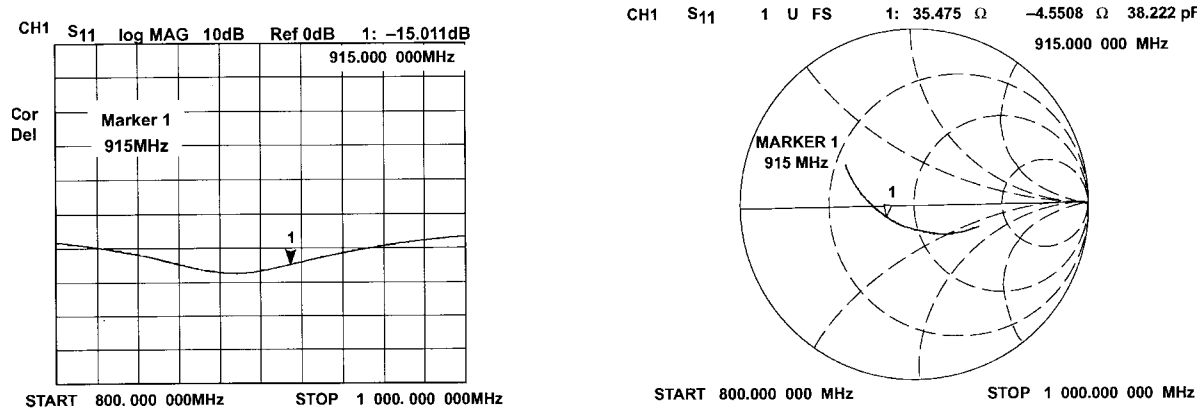
On-Chip Clocks. The crystal oscillator is buffered to provide three external clock signals: CLK1, CLK2, and MCLK. Table 1-10 shows the divide ratio and the controlling mechanism. CLK1 is usually used for the system synthesizer (e.g., SA7025) reference. Since MCLK is active all the time, it is ideal for providing the master clock for the microcontroller. When the device is in STANDBY mode, CLK1 and MCLK provide the clock signals necessary for receiving RF signals. CLK2 can also be used as a clock for the digital signal processing (DSP) chip.

Modes of Operation. The SA900 is intended for either AMPS mode (analog cellular) or DUAL mode (digital cellular, IS-54) operation. When the device is running in AMPS mode, the I/Q modulator, variable gain amplifier (VGA), and phase shifter are disabled. The fixed-gain amplifier is powered up during AMPS mode operation. However, since the divide ratio is too low (6, 7, or 8), the comparison frequency of the on-board PLL is too high, making it very difficult for the loop bandwidth to be less than 300 Hz for analog frequency modulation. The device includes two power-saving modes of operation that disable partial circuitry to reduce the power consumption of the overall chip. The SLEEP mode disables all the circuitry except the master clock (MCLK pin) of the SA900. The STANDBY mode shuts down everything except the TXLO buffer, MCLK, and CLK1, which allows the system synthesizer (e.g., SA7025) to continue running. These two power-saving modes are common to both AMPS and DUAL mode operation. The SA900 draws 60 mA in the DUAL mode, reduced to 3 mA and 8 mA, respectively, in the SLEEP and STANDBY modes.

The TXEN pin is for hardware powering down the modulator and synthesizer. The falling edge of the signal disables the modulator and synthesizer while the rising edge enables the modulator. To power down the synthesizer using software, send a data word with the SE bit set to "0" ("1" for enable). The synthesizer will be disabled right after the strobe signal is transmitted. Either SE or TXEN going low will turn off the synthesizer. This operation is common to both the AMPS and DUAL modes.



a. S_{11} Response and Smith Chart when L2 = 22nH



b. S_{11} Response and Smith Chart when L2 = 15nH

Figure 1-89 SA900 output matching using *S* parameters.

Table 1-10 Possible SA900 clock outputs

CLK1	divide by 3	X (bit 18) = 1
	divide by 1	X (bit 18) = 0
CLK2	divide by 2	Y (bit 19) = 1
	divide by 1	Y (bit 19) = 0
MCLK	divide by 4	CLKSET pin = V_{CC}
	divide by 5	CLKSET pin = $V_{CC}/2$
	divide by 1	CLKSET pin grounded

Performance of the SA900

Performance Criteria. Since the I/Q modulator is a universal transmitter, measuring only the frequency stability and modulation index of a generated FM signal would not be useful for other modulation schemes. Measurement parameters should be general enough so that they can represent the performance of modulators when applying different types of modulation and allow fair comparisons among different I/Q modulators. Based on this idea, two measurement techniques—in-phase modulation and quadrature modulation—are used for evaluating I/Q modulators.

The in-phase modulation relies on injecting two equal frequencies and phase signals at f_{mod} into the I and Q inputs. The result of this modulation is two sidebands appearing at f_{mod} offset from the carrier, with the carrier totally suppressed. This is also known as double-sideband (DSB) conversion. The quadrature modulation requires two equal frequencies (but 90° out-of-phase signals) being injected into the I and Q inputs. The result is a single-sideband suppressed carrier (SSBSC) signal with either the upper or lower sideband at f_{mod} carrier offset being suppressed. This is also known as single-sideband (SSB) upconversion. Figure 1-90 summarizes these two tests.

In a practical system, imperfection of an I/Q modulator is directly related to these two measurements. Sideband and carrier suppression from the quadrature modulation test will show the amount of gain imbalance, phase imbalance, and dc offset. On the other hand, intermodulation product suppression from the in-phase modulation test will show the linearity of an I/Q modulator. When making measurements, it is important to have well-balanced I and Q baseband modulating signals for measurement since the signal imperfection will translate into degraded sideband and carrier suppression.

Performance Graphs. In making those measurements for the demo board, the following parameters were used:

In-Phase Modulation

PIN 43 I1 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 0°
PIN 42 I2 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 180°
PIN 41 Q1 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 0°
PIN 40 Q2 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 180°

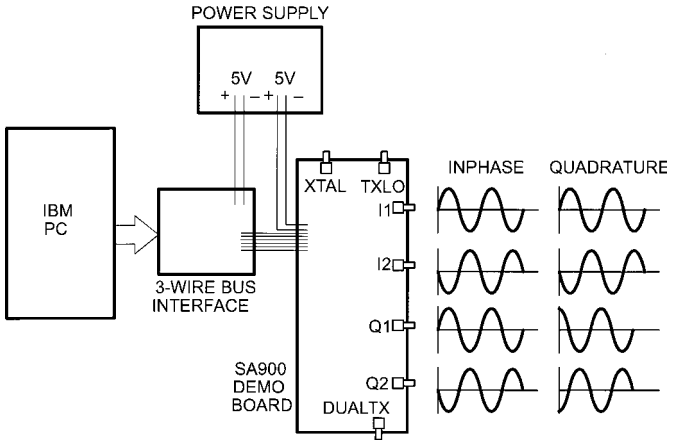


Figure 1-90 In-phase and quadrature modulation test.

Quadrature Modulation

PIN 43 I1 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 0°
PIN 42 I2 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 180°
PIN 41 Q1 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 90°
PIN 40 Q2 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 270°

Figure 1-91 illustrates what the typical output spectra would be if in-phase and quadrature modulation were applied to an I/Q modulator. Quadrature modulation will produce lower sideband (LSB) or upper sideband (USB) signal, depending on the phase angle between the I and Q signals. The SA900 was designed to have USB suppressed when the I signal leads the Q signal. The undesired signals are carrier breakthrough and the harmonic products of the baseband modulating signals sitting at $f_c \pm nf_{\text{mod}}$, where n is an integer ≥ 2.

Referring to Figure 1-91a, the output power is 1.3 dBm (cable loss = 0.7 dB) for the LSB while better than 38 dBc of carrier, sideband, and harmonics suppression is measured. The USB suppression of better than 26 dBc implies the residual AM of the transmit signal is better than 5%, a requirement of the IS-54 specification.

The in-phase modulation test will generate both LSB and USB. Beside these two tones, the carrier breakthrough and the harmonics, intermodulation (IM) products will all appear at the output. The odd IM products are dominant, and they satisfy the following rules: Let $f_1 = f_c - f_{\text{mod}}$, and $f_2 = f_c + f_{\text{mod}}$.

$$\text{Third-order IM:} \quad 2f_1 - f_2 = f_c - 3f_{\text{mod}}, \quad 2f_2 - f_1 = f_c + 3f_{\text{mod}}$$

$$\text{Fifth-order IM:} \quad 3f_1 - 2f_2 = f_c - 5f_{\text{mod}}, \quad 3f_2 - 2f_1 = f_c + 5f_{\text{mod}}$$

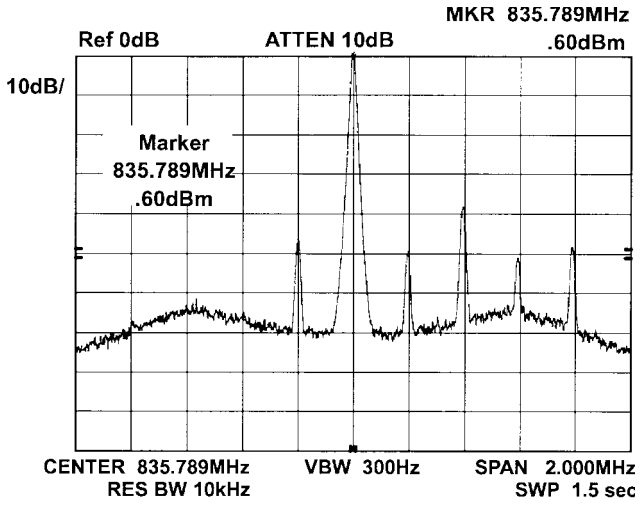
$$\text{Seventh order IM:} \quad 4f_1 - 3f_2 = f_c - 7f_{\text{mod}}, \quad 4f_2 - 3f_1 = f_c + 7f_{\text{mod}}$$

Referring to Figure 1-91b, the LSB and USB are –1.6 dBm (cable loss = 0.7 dB), which is 3 dB less than the measured power for the quadrature modulation test. The IM3 is better than –35 dBc. IM products of much higher orders are totally suppressed.

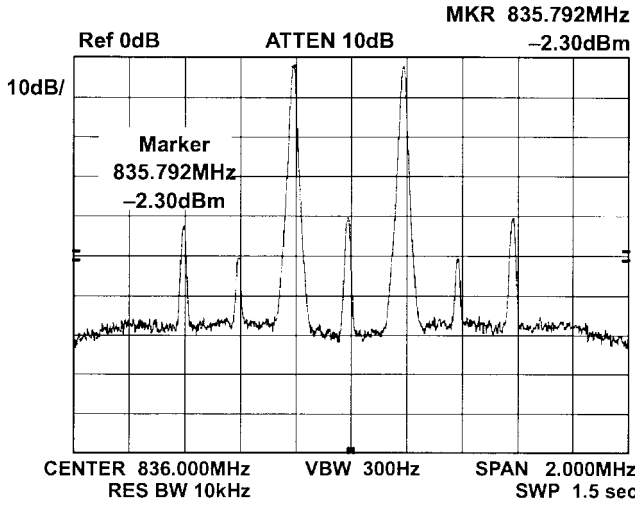
Amplitude and Phase Imbalance. Both amplitude and phase imbalance (error) of an I/Q modulator can be calculated directly from the SSB performance plots. Assuming that phase error equals ϕ radians and amplitude error equals K. The sideband suppression, X, in dBc can be expressed as (see the section entitled "I/Q Modulator Equations" for the derivation):

$$\text{SSB suppression,} \quad X(\text{dBc}) = 10 \log\left(\frac{K^2 + 2 \times K \times \cos(\phi) + 1}{K^2 - 2 \times K \times \cos(\phi) + 1}\right) \tag{1-21}$$

Collecting the like terms and expressing ϕ in terms of K and X, we find



a. Quadrature Test (SSB Up-conversion)



b. In-phase Test (DSB Up-conversion)

Figure 1-91 Typical SA900 output spectra.

$$\phi = \cos^{-1}\left(\frac{10^{X/10} \times K^2 + 10^{X/10} - 1 - K^2}{2 \times K + 2 \times K \times 10^{X/10}}\right) \tag{1-22}$$

For a given X, there will be a set of ϕ and K that satisfies Eq. (1-22). We can represent this relationship graphically, as shown in Figure 1-92. The contours show the phase and amplitude errors for SSB suppression, X, from 44 to 26 dBc. When X equals 40 dBc, phase error is less than 1.2° with an amplitude error of 0 dB. By the same token, the amplitude error is less than 0.2 dB with a 0° phase error.

Spectral Mask. To fully characterize the performance of an I/Q modulator, measurements of the power spectral density of various digital modulation schemes have to be made. Figures 1-93a and 1-93b show the measured spectral masks of IS-54 and PDC standards, which designate π/4-DQPSK as the modulation format.

GMSK is a digital modulation scheme widely used for wireless and mobile communications. Figure 1-93c shows the spectral mask of the modulation format required by GSM, the digital cellular standard in Europe. At carrier offsets of 200 and 300 kHz, the power of the signal is suppressed by 46 and 58 dB, respectively, which is well within the GSM specification.

Power ON Time. The power ON time for the SA900 is mainly determined by the loop bandwidth of the on-board PLL frequency synthesizer. It can be measured by using the HP 53310A modulation domain analyzer set to the EXTERNAL TRIGGERED mode. The STROBE signal from a three-wire bus is used to trigger the equipment. Figure 1-94 shows that the part can be powered up and locked in about 62 μs.

ISM Band Application. In the United States the industrial, scientific, and medical (ISM) band from 902 to 928 MHz is attractive because users are allowed, without having a license, to use transmit powers up to 1 watt in frequency hopping or direct sequence CDMA operation. The wide bandwidth nature of the SA900 fits well into this application. Figure 1-95 shows the output spectra of the SA900, illustrating how well the image-reject mixer works. A common IF (45 MHz) was chosen to be the offset frequency, and then injected externally into the VCO pins. The closest images (± 45 MHz), are down more than 36 dBc.

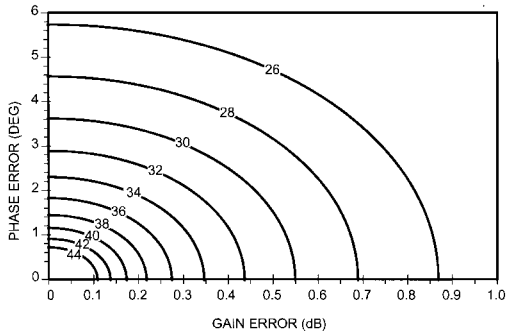
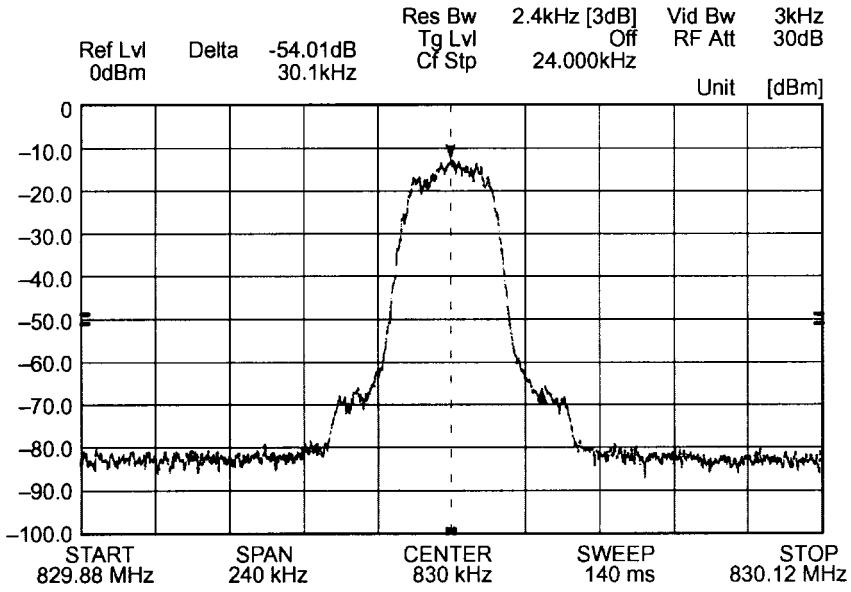
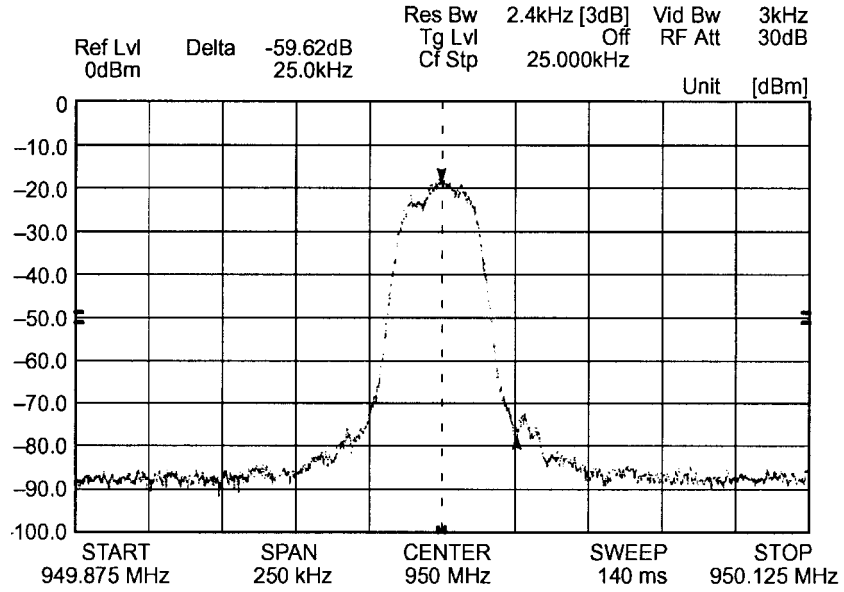


Figure 1-92 SSB suppression contours.

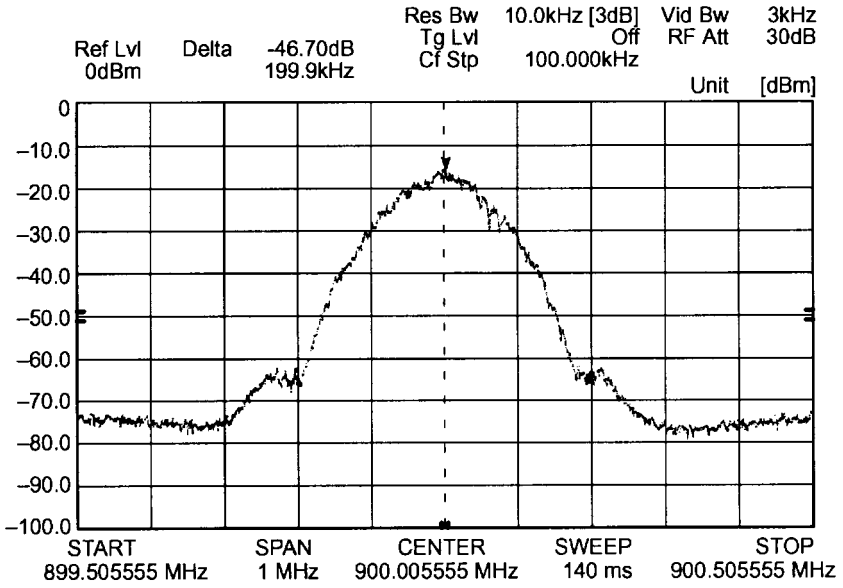


a. Power Spectrum of ADC Modulation
Bit rate = 48.6 kb/s; $\alpha = 0.35$



b. Power Spectrum of PDC Modulation
Bit rate = 42 kb/s; $\alpha = 0.5$

Figure 1-93 SA900 ADC, PDC, and GSM outputs.



c. Power Spectrum of GSM Modulation
Bit rate = 270.833 kb/s; $BT_b = 0.3$

Figure 1-93 Continued

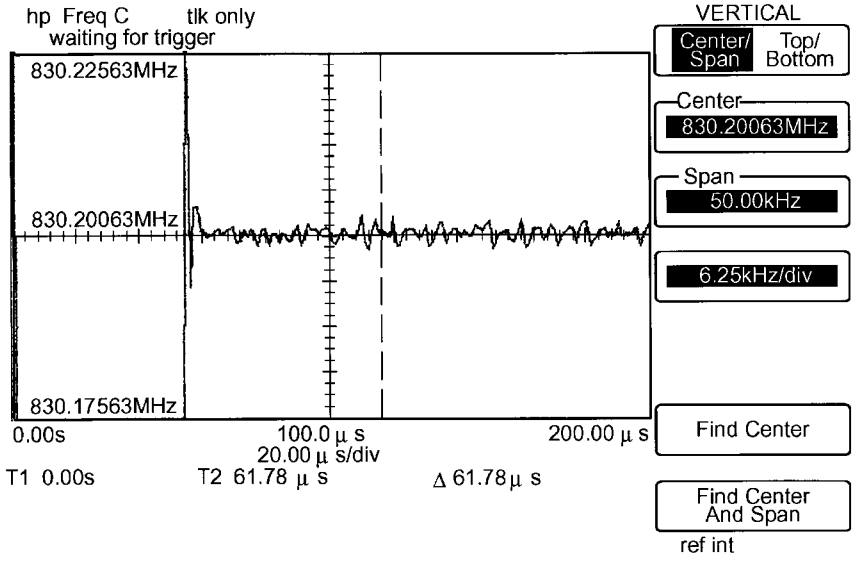


Figure 1-94 Power ON time measurement.

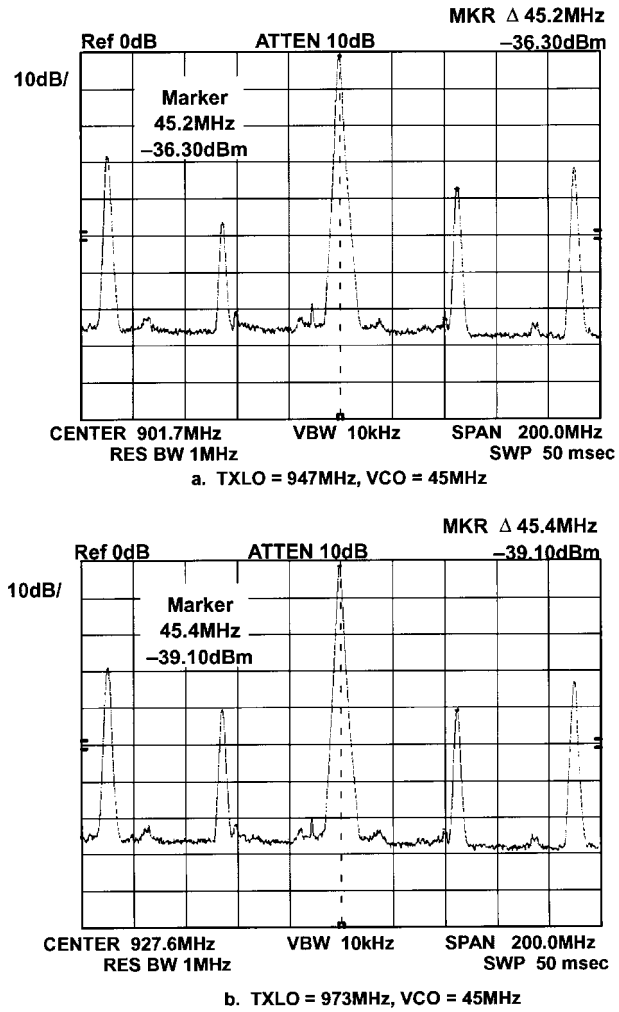


Figure 1-95 Output spectra of the SA900 in the 902-MHz ISM band.

I/Q Modulator Equations. Assume an imperfect I/Q modulator with gain error K and phase error ϕ, modulated by quadrature I/Q signals (SSB upconversion) ω_m. Then the signal, $s(t)$, at the output of the I/Q modulator becomes

$$s(t) = K\cos(\omega_c t + \phi)\cos(\omega_m t) - \sin(\omega_c t)\cos(\omega_m t + 90°) \tag{1-23}$$

Using trigonometric identity and letting $\omega_c - \omega_m = A$ and $\omega_c + \omega_m = B$, we obtain

$$s(t) = \frac{K}{2}\cos(At + \phi) + \frac{K}{2}\cos(Bt + \phi) + \frac{1}{2}\cos(At) - \frac{1}{2}\cos(Bt) \tag{1-24}$$

Assume the information is in the LSB—that is, A—and the spur is the USB—that is, B. We have

$$\text{Signal} = \frac{K}{2}\cos A \cos\phi + \frac{1}{2}\cos A - \frac{K}{2}\sin A \sin\phi \tag{1-25}$$

$$\text{Noise} = \frac{K}{2}\cos B \cos\phi + \frac{1}{2}\cos B - \frac{K}{2}\sin B \sin\phi \tag{1-26}$$

To find the power, we have to evaluate the envelope (amplitude) of these two signals. Recall that for any given bandpass signal in rectangular form, that is, bandpass signal = $X \cos \omega_t - Y \sin \omega_t$, the envelope is

$$\text{Envelope} = (X^2 + Y^2)^{0.5}$$

Therefore, from Eqs. (1-25) and (1-26)

$$\text{Signal} = \left[\left(\frac{K}{2}\cos\phi + \frac{1}{2}\right)^2 + \left(\frac{K}{2}\sin\phi\right)^2\right]^{0.5} \tag{1-27}$$

$$\text{Noise} = \left[\left(\frac{K}{2}\cos\phi - \frac{1}{2}\right)^2 + \left(\frac{K}{2}\sin\phi\right)^2\right]^{0.5} \tag{1-28}$$

Finally, the S/N ratio can be found by taking 10 log the ratio of Eqs. (1-27) and (1-28):

$$S/N = 10\log\left(\frac{K^2 + 2K\cos\phi + 1}{K^2 - 2K\cos\phi + 1}\right) \tag{1-29}$$

1-6 BUILDING BLOCKS

Our earlier block diagram already referred to a variety of integrated circuits that can be used to build a wireless system—in particular, a portable telephone. If we open any of the earlier-mentioned magazines and go through its advertising section, we will find many suppliers of subsystems/components that fit this requirement.

They can be split into various technologies. The following list is a product overview of next-generation advances in wireless technology taken from an ad of Stanford Microdevices, one of the very aggressive component/subsystem companies:

- New SX amplifiers featuring 2.4-GHz MMICs
- GaAs heterojunction bipolar transistor (HBT) discrete transistors (100 mW to 10 W)
- InP/GaAs (indium phosphate/gallium arsenide) high-linearity gain blocks
- Millimeter-wave product line featuring ICs for LMDS
- Silicon–germanium product line (gain blocks and low-noise amplifiers)
- SXQ power modules for cellular and PCS applications
- Power modules featuring 4–25-W power amplifiers for land mobile amplifiers
- Thirty- to 120-W silicon LDMOS power transistors for PCS, CDMA, and W-CDMA
- SAW filters featuring integrated transceiver modules
- Electro-optic and data communications products

More information about Stanford products is available at **http://www.stanfordmicro.com/**. Another interesting company along these lines is Maxim Integrated Products

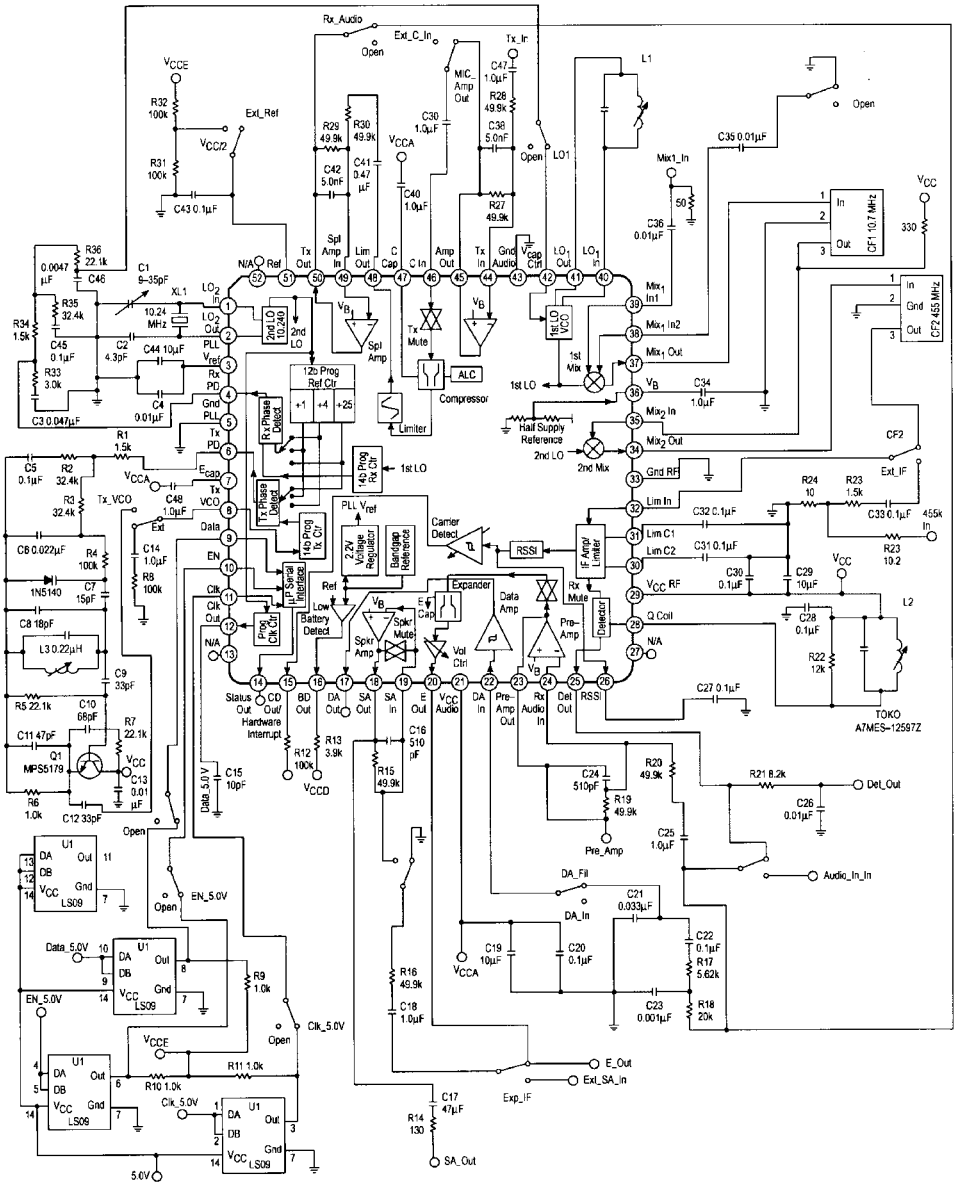


Figure 1-96 MC13109FB universal cordless telephone subsystem IC.

(**http://www.maxim-ic-com/**). They distribute an Analog Design Solutions brochure describing, among other things, the industry's first SiGe RFICs for improved radio front-end performance (1999). On the base-station side, companies such as Hewlett-Packard and Synergy Microwave produce high-performance components/modules as they are needed in the more hostile environment of base-station operation. The offerings in the field are changing very rapidly, and the best way to deal with this is to obtain subscriptions to the magazines mentioned earlier or contact manufacturers directly. Among those not mentioned are Alpha, Ericsson, Siemens (now Infineon Technologies), Motorola, and NEC.

In the beginning of this chapter we looked at block diagrams; here we focus on building blocks. Most manufacturers will try to move to the highest possible integration while avoiding external components. Figure 1-96, a test circuit for Motorola's MC13109FB cordless telephone subsystem IC, shows the level of complexity needed to test such an IC with all its functionality.

Another key issue at those frequencies is to really know a component's frequency dependencies. Figure 1-97 shows a good example. Typical components are capacitors and inductors as reactive elements, and resistors as passive elements. Depending on the type of integration, one may also have to look into the noise performance of integrated resistors. Examination of this issue is best achieved through discussions with the foundries. Along these lines, the dielectric material on which these surface-mounted components are mounted plays a big role and must be addressed individually.

Besser Associates of Los Altos, California, among other firms, offers a very nice one-day short course titled "Wireless Circuit Components: Measurements, Models and Data Extraction."

Table 1-11 presents a sampling of RFICs for wireless applications. These types of ICs, which were taken from Motorola, are made in similar form by many companies. The selection guide is useful in starting off with a design that is less integrated, therefore allowing the designer to have access to portions of the circuit that later "disappear" inside the IC.

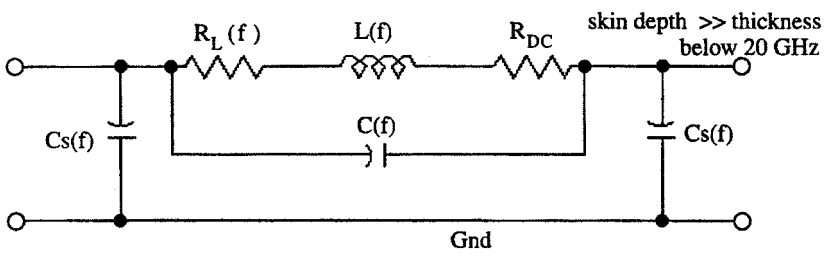


Figure 1-97 Equivalent model of a film resistor above a ground plane.

Table 1-11 Sample RF component/IC selector guide

RFICs

DOWNCONVERTERS

Device	RF Frequency Range (MHz)	Supply Voltage Range (V/dc)	Supply Current (mA) (Typ)	LNA Gain (dB) (Typ)	LNA NF (dB) (Typ)	Mixer Converter Gain (dB) (Typ)	Mixer NF (dB) (Typ)	Package	System Applicability
MC13142	dc to 1800	2.7 to 6.5	13.5	17	1.8	–3.0	12	SO-8	ISM, Cellular, PCS
MRFIC1502	1575	4.5 to 5.5	52	20	—	45	9.5	LQFP-48	GPS
MRFIC1814	1800 to 2000	2.7 to 4.5	10	17	2.5	8.0	10	TSSOP-16	CDS1800, PCS PHS

UPCONVERTERS/EXCITERS

Device	RF Frequency Range (MHz)	Supply Voltage Range (V/dc)	Supply Current (mA) (Typ)	Standby Current (mA) (Typ)	Converter Gain (dB) (Typ)	Output IP3 (dB) (Typ)	Package	System Applicability
MRFIC0954	800 to 1000	2.7 to 5.0	65	5.0	31	28	TSSOP-20EP	CDMA, TDMA, ISM
MRFIC1813	1700 to 2000	2.7 to 4.5	25	0.1	15	11	TSSOP-16	DCS1800, PCS
MRFIC1854	1700 to 2000	2.7 to 5.0	70	5.0	31	23	TSSOP-20EP	CDMA, TDMA, PCS

POWER AMPLIFIERS

Device	Frequency Range (MHz)	Supply Voltage Range (V/dc)	Saturated P_{out} (dBm) (Typ)	PAE (%) (Typ)	Gain P_{out}/P_{in} (dB) (Typ)	Package	System Applicability
MRFIC0913	800 to 1000	2.7 to 7.5	35	50	25	PFP-16	GSM
MRFIC0917	800 to 1000	2.7 to 5.5	34.5	45	22.5	PFP-16	GSM
MRFIC0919	800 to 1000	3.0 to 5.5	35.3	48	32.3	TSSOP-16EP	GSM
MRFIC1805	1500 to 2200	2.7 to 5.0	25	28	20	TSSOP-16	PHS, DECT, PCS
MRFIC1807	1500 to 2200	3.0 to 5.0	26.8	35	8.0	SO-16	DECT, PCS

MRFIC1817	1700 to 2000	2.7 to 5.0	33.5	42	30.5	PFP-16	DCS1800, PCS
MRFIC1818	1700 to 2000	2.7 to 6.0	34.5	42	31.5	PFP-16	DCS1800, PCS
MRFIC1819	1700 to 2000	3.0 to 5.0	33	40	27	TSSOP-16EP	DCS1800, PCS
MRFIC1856	800 to 1000	3.0 to 5.6	32	50	32	TSSOP-20EP	TDMA, CDMA, AMPS
	1700 to 2000		30	35	30		TDMA, CDMA, PCS
MRFIC2006	500 to 1000	1.8 to 4.0	15.5	25	23	SO-8	Cellular, ISM, CT

RF Building Blocks

AMPLIFIERS

Device	RF Frequency Range (MHz)	Supply Voltage Range (V/dc)	Supply Current (mA) (Typ)	Standby Current (μA) (Typ)	Small Signal Gain (dB) (Typ)	Output IP3 (dBm) (Typ)	NF (dB) (Typ)	Package	System Applicability
MC13144	100 to 2000	1.8 to 6.0	8.5	1	17	–5.0	1.4	SO-8	ISM, PCS, Cellular
MRFIC0915	100 to 2500	2.7 to 5.0	2.0	—	16.2	4.0	1.9	SOT-143	ISM, PCS, Cellular
MRFIC0916	100 to 2500	2.7 to 5.0	4.7	—	18.5	11	1.9	SOT-143	ISM, PCS, Cellular
MRFIC0930	800 to 1000	2.7 to 4.5	8.5	20	19	10	1.7	SO-8	GSM, AMPS, ISM
MRFIC1808DM	1700 to 2100	2.7 to 4.5	5.0	8.0	18	13	1.6	Micro-8	DCS1800, PCS
MRFIC1830	1700 to 2100	2.7 to 4.5	9.0	20	18.5	8.5	2.1	Micro-8	DCS1800, PCS
MRFIC1501	1000 to 2000	3.0 to 5.0	5.9	—	18	10	1.1	SO-8	GSP

MIXERS

Device	RF Frequency Range (MHz)	Supply Voltage Range (V/dc)	Supply Current (mA) (Typ)	Standby Current (μA) (Typ)	Converter Gain (dB) (Typ)	Input IP3 (dBm) (Typ)	Package	System Applicability
MC13143	DC to 2400	1.8 to 6.0	4.1	—	–2.6	16	SO-8	ISM, PCS, Cellular

(*continued*)

Table 1-11 Sample RF component/IC selector guide (Continued)

Frequency Synthesis

PLL SYNTHESIZERS

Frequency (MHz)	Supply Voltage (V/dc)	Nominal Supply Current (mA)	Phase Detector	Standby	Interface	Device	Suffix/Case
20 @ 5.0 V	3.0 to 9.0	7.5 @ 5 V	Single-ended three-state, double-ended	No	Parallel	MC145151-2	DW/751F
			Double-ended			MC145152-2	DW/751F
			Single-ended three-state, double-ended		Serial	MC145157-2	DW/751G
						MC145158-2	DW/751G
60 @ 3.0 V	2.5 to 5.5	3 @ 3 V	Two single-ended three-state	Yes		MC145162[a]	P/648, D/751B
85 @ 3.0 V	2.5 to 5.5	3 @ 3 V			Serial	MC145162-1[a]	D/751B
100 @ 3.0 V 185 @ 5.0 V	2.7 to 5.5	2 @ 3 V 6 @ 5 V	Single-ended three-state, double-ended	No		MC145170-2	P/648, D/751B, DT/948C
550, 60	1.8 to 3.6	3	Loop 1 = current source/sink/float Loop 2 = three-state	Yes		MC145181[a]	FTA/873C
1000	2.7 to 5.5	4.25	Current source/sink/float	No	Parallel	MC12181	D/751B
1100	4.5 to 5.5	7 @ 5 V	Current source/sink/float, double-ended	Yes	Serial	MC145191	F/751J, DT/948D
1100	2.7 to 5	6 @ 2.7 V				MC145192	F/751J, DT/948D
1100	2.7 to 5.5	12	Two current source/sink/float, double-ended			MC145220[a]	F/803C, DT/948D
1200, 550	1.8 to 3.6	4	Loop 1 = current source/sink/float Loop 2 = three-state			MC145225[a]	FTA/873C
2000	4.5 to 5.5	12 @ 5 V	Current source/sink, double-ended			MC145201	F/751J, DT/948D
2000	2.7 to 5.5	4 @ 3 V				MC145202	F/751J, DT/948D
2200, 550	1.8 to 3.6	5	Loop 1 = current source/sink/float Loop 2 = three-state			MC145230[a]	FTA/873C
2500	2.7 to 5.5	9.5	Current source/sink/float with dual outputs	No		MC12210	D/751B, DT/948E
2800	4.5 to 5.5	3.5	Current source/sink/float		None	MC12179	D/751

[a]Dual PLL.

1-7 SYSTEM SPECIFICATIONS AND THEIR RELATIONSHIP TO CIRCUIT DESIGN

Wireless communication involves a large range of signal powers—from levels on the order of 10^{-18} watt (at the receiver input) to 10^{2} watts (at a base-station transmitter output). A receiver must be able to demodulate signals that have been attenuated billions of times through propagation; a transmitter must be able to produce a properly modulated signal at a frequency suitable for propagation, at a level high enough to overcome worst-case propagation losses and provide a useful signal at the receiver. Gain is therefore an essential attribute of wireless systems. Because no single active device can provide all the gain required for transmission or reception, we must distribute the gain among multiple stages, designing each for optimum performance across the power span it bridges.

Two inescapable realities impose limits on the gain and absolute power output we may achieve with a given circuit: All real electrical and electronic networks generate noise to some degree, and all real electrical and electronic networks distort the signals applied to them to some degree. A signal weaker than a circuit's inherent noise cannot be amplified by that circuit because it remains indistinguishable from the noise. A signal that exceeds the power-handling capability of the circuit to which it is applied may be degraded, even rendered useless, by the resulting distortion. The following sections examine issues particular to system noise and linearity performance.

1-7-1 System Noise and Noise Floor

Assuming that a system's gain is sufficient, the weakest signal that may be processed satisfactorily, a figure of merit referred to as *noise floor* or (in receivers) *minimum detectable signal* (*MDS*), is limited by thermal noise, assumed to be equal to the noise power available from a resistor at 290 K (about 17 °C or 62 °F), an arbitrary reference value near standard room temperature. The noise power is equal to

$$P_n = kTB \tag{1-30}$$

where P_n is the noise power, k is Boltzmann's constant (1.38×10^{-23} watts per kelvin), T is the temperature in kelvins, and B is the bandwidth (in hertz) in which the noise appears. For $T = 290$ K, P_n is therefore 4.00×10^{-18} watts, or –174 dBm in a 1-Hz bandwidth. Increasing B to a value suitable for digital communications, such as 160 kHz for a GSM system, admits more noise to the network, raising the minimum noise power against which an incoming signal must compete to –122 dBm.

If the noise figure and bandwidth are known, the noise floor can be calculated using the equation

$$\text{Noise floor} = -174 \text{ dBm} + NF + 10 \log B \tag{1-31}$$

The trouble with this equation is that the "integrated" bandwidth depends so much on the selectivity shape factor, which is not always known. Figure 1-98 shows the translation of the bandwidth of a single tuned circuit with its Gaussian shape into its rectangular equivalent. The transformation is done by sizing the rectangle such that A′ = A and B′ = B; when this is true, the area of the rectangle equals the area under the curve.

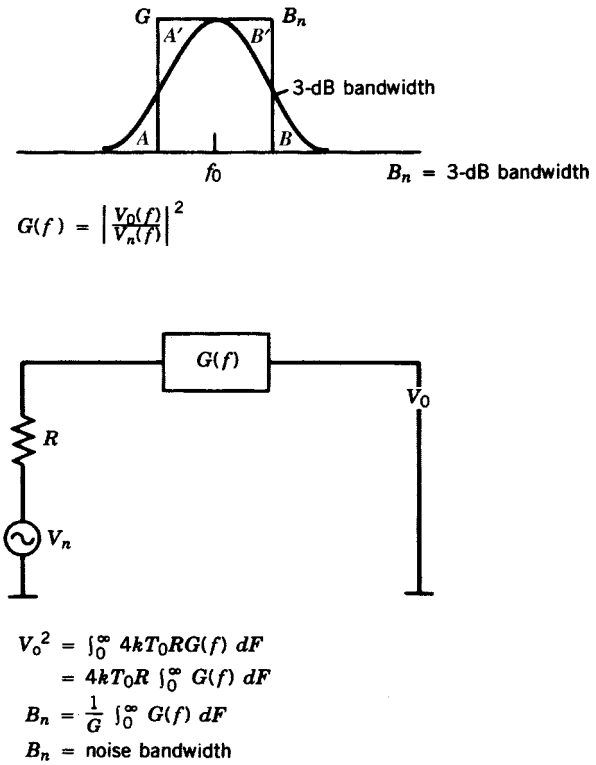


Figure 1-98 Graphical and mathematical explanation of the noise bandwidth from a comparison of the Gaussian-shaped bandwidth to the rectangular filter response.

Signal-to-Noise Ratio (S/N, SNR) and Sensitivity. Successful radiocommunication depends on the achievement of a specified minimum ratio of signal power to noise power, expressed in decibels, at the output of the receiver. The input voltage, expressed in absolute units or decibels relative to a microvolt (dBμV), necessary to achieve a particular signal-to-noise ratio in a particular bandwidth may be specified as a figure of merit called *sensitivity*. Because techniques used to measure S/N actually measure the ratio of signal-*plus-noise* to noise, specifications may refer to (or imply) S+N/N or (S+N)/N rather than S/N. The difference between (S+N)/N and S/N becomes negligible at high ratios of signal to noise; even at an SNR of 10 dB—a common value—the difference is only 0.46 dB [9].

Most receivers are designed to operate optimally when connected to an antenna system of a specified impedance (commonly 50 Ω), but relatively few receivers exhibit this design load impedance at their input terminals; that is, they are not designed for a conjugate input match. It is therefore customary to specify sensitivity in terms of "open circuit" voltage—the signal voltage that, with the receiver's antenna input terminated in its design antenna impedance, results in the desired ratio of signal to noise. If, as is commonly the case, the input voltage for a given SNR is determined using instrumentation calibrated in terms of *closed*-circuit voltage—that is, in terms of voltage across a load resistance equal to the instrument's source resistance—the voltage indicated will be one-half the open-circuit value for the SNR specified [10]. By convention, the open-circuit measurement condition is indicated by a sensitivity specification in volts of electromotive force (EMF). Specifying sensitivity in

terms of available signal power (usually decibels relative to 1 mW, or dBm) eliminates this open/closed-circuit confusion.

SINAD Ratio. Extending the measurement of signal-plus-noise to noise to include distortion results in a figure of merit called *SINAD* (signal-plus-noise-and-distortion), commonly applied to FM receivers:

$$SINAD = 10 \log_{10}\left(\frac{S+N+D}{N+D}\right) \tag{1-32}$$

where *SINAD* is in decibels, S is signal power, N is noise power, and D is distortion power. At a SINAD ratio of 12 dB—a common specification—the noise-and-distortion power is 25% that of the desired signal. As is true of (S+N)/N, SINAD closely approximates S/N at high ratios of signal to noise.

Bit Error Rate and Noise. For digital systems, signal-to-noise ratio and bit error rate are related. As introduced earlier, depending on the waveform, coding, and filtering, different BERs are related to particular SNRs (Figure 1-99). The SNR also depends on the interference from the synthesizer, as will be seen in the section on adjacent-channel power ratio (ACPR). Since a digital transmitter is always involved, in the on-the-air testing the laboratory measurement of SINAD cannot quite be correlated. At the IF level, we find the handover point between the transceiver's analog front end and its digital portion, and a standard noise-figure test setup, such as the one by Hewlett-Packard, is still a valuable instrument for evaluating noise figure and signal-to-noise ratio.

Here is an example for a digital measurement: Assume that we use a test generator like the SMIQ to evaluate a front end having an 8.8-dB noise figure. A 384-kbps π/4-DQPSK digital modulation signal (–5 dBm) is applied to the transceiver front end. Testing at 2.47 GHz, we find that the error vector magnitude (EVM; defined later in this section) is about 3.89% in transmit and about 1.37% in receive. To determine the receive sensitivity of the RF front end, the signal from the digitally modulated generator (at –95 dBm) is applied to the

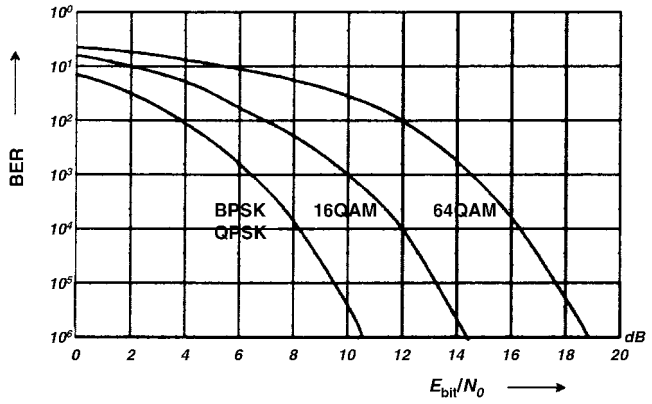


Figure 1-99 Bit error rate versus E_{bit}/N_0 for BPSK/QPSK, 16-QAM, and 64-QAM, showing the significantly greater SNR necessary for a given BER as the number of signal states is increased.

RF front end in receive mode. The measured EVM at 2.47 GHz is now 20%. By using the equation

$$S/N = -20 \log (EVM) \tag{1-33}$$

we see that the ratio that corresponds to an EVM of 20% is 14 dB, since

$$S/N = \frac{(E_S R_{S)}}{N_0 B} \tag{1-34}$$

where E_S is the energy of a symbol, R_S is the symbol rate, N_0 is the noise power density, and B is the bandwidth. Assuming a π/4-DQPSK signal and ratio of 14 dB, it can be determined that

$$E_S/N_0 = (S/N)\ (B/R_S) = 17 \text{ dB} \tag{1-35}$$

Assuming that the BER is 1E – 5, the required E_S/N_0 is 10 dB. Hence, from the definition of the sensitivity,

$$\begin{aligned} \text{Sensitivity} &= -144 + NF + R_S + E_S/N_0 \\ &= -144 + 8.8 + 10 \log(384/2) + 10 \\ &= -102 \text{ dBm} \end{aligned} \tag{1-36}$$

The value –144 is derived from kT_0 by taking the bandwidth into consideration.

The result obtained in Eq. (1-36) is consistent with the obtained E_S/N_0 of 17 dB from the measurement at –95 dBm input power [since –102 + (17 – 10) = –95]. Hence, it was determined that the receive sensitivity of the RF front end is approximately –102 dBm. We can then determine the dynamic range (DR) from

$$DR = \text{P}_{-1\text{dB}} - \text{Sensitivity} = -21 - 102 = 81 \text{ dB} \tag{1-37}$$

Noise Factor and Noise Figure. The noise floor predicted by Eq. (1-30) cannot be achieved and maintained in any real network or system of networks because all real networks generate noise. Determining how closely the SNR achieved at a given input level approaches the SNR achievable at that input level in a noiseless network is therefore of high interest to the circuit and system designer. The degree to which a network's noise contribution degrades the noise floor predicted by Eq. (1-30) is evaluated by its *noise factor* (*F*), which is expressed as the ratio

$$F = \frac{N_{\text{in}} + N_{\text{added}}}{N_{\text{in}}} \tag{1-38}$$

where F is the noise factor, N_{in} is the noise power available from the source, and N_{added} is the noise power added by the network, with both powers determined in the same bandwidth. Expressing this ratio in decibels ($10 \log_{10} F$) returns the *noise figure* (*NF*), a bandwidth-

independent figure of merit of great value in evaluating the noise performance of networks and communication systems. We can also express NF as the ratio of the network's input SNR to its output SNR:

$$NF = 10 \log_{10} \left[\frac{(S_{in}/N_{in})}{(S_{out}/N_{out})} \right] \tag{1-39}$$

where NF is the noise figure in decibels, S is signal power, and N is noise power, with the input and output values of these quantities signified by the subscripts and all powers determined in the same bandwidth. The noise figure of an ideal noiseless network is 0 dB; for all real, noisy networks, NF is positive. The NF of a lossy passive device is equal to its insertion loss.

Noise figure can also be defined for antennas and antenna systems:

$$NF_{ant} = 10 \log_{10} \frac{N_t + N_{ant}}{N_t} \tag{1-40}$$

where NF_{ant} is the antenna system noise figure in decibels, N_t is the antenna system's thermal noise power, as defined in Eq. (1-30), and N_{ant} is the total noise power picked up by the antenna system. From the lower end of the radio spectrum, and decreasingly up to approximately 400 MHz, noise intercepted by an antenna system from atmospheric, human-generated, and galactic sources will dominate NF_{ant} (Figure 1-100), and N_t can be considered as equivalent to the noise power of a resistor at 290 K. Atmospheric noise subsides above 40 MHz; from this region to perhaps 700 MHz, N_t is still largely negligible, with noise from human-generated and/or sky sources largely determining an antenna's noise figure. At these frequencies, an antenna's directivity and orientation relative to noise sources play a critical role in determining its noise figure; a strongly directive antenna located in a city suburb, for instance, exhibits a significantly higher noise figure when pointed toward the city center than

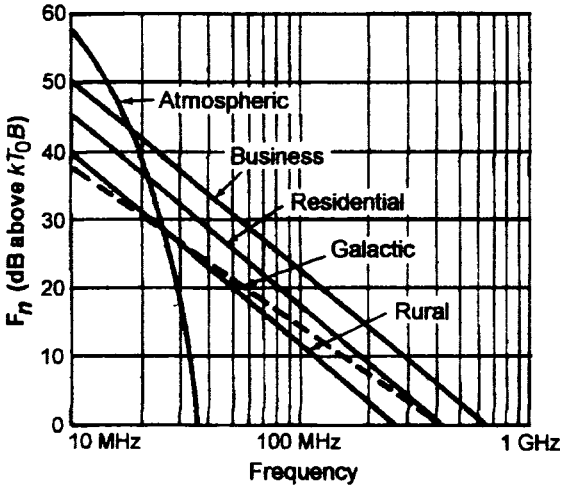


Figure 1-100 The higher the frequency, the more quiet the RF environment becomes, although the noise profile of specific sources may contradict this general rule.

it exhibits when pointed toward a more sparsely populated area. At frequencies above about 700 MHz, the RF environment is generally quieter, allowing receiver NF, not antenna NF, to more routinely limit a wireless receiver's sensitivity. At these frequencies, the concept of *noise temperature* is commonly used to evaluate the quietness of a receiver, its antenna system and its RF environment. Noise temperature is particularly useful in designing systems for space communication an antenna for which may be cooled, through radiation of a portion its own noise into the RF-cold sky, to a noise temperature an order of magnitude below 290 K.

Noise Figure of Cascaded Networks. The noise figure of two networks in cascade may be determined from

$$NF_{\text{total}} = 10 \log_{10}\left(F_1 + \frac{F_2 - 1}{G_1}\right) \tag{1-41}$$

where NF is the noise figure in dB, F_1 is the noise factor of the first network, F_2 is the noise factor of the second network, and G_1 is the gain (as a numerical ratio, *not* in dB) [11]. The noise figure of a system with more than two stages can be evaluated through repeated iterations of Eq. (1-41). Note that Eq. (1-41) assumes two conditions: (1) that F_1 and F_2 are determined in the same bandwidth, and (2) that the networks' input and output terminations are resistive—a condition that is commonly *not* true of RF amplifiers optimized for lowest noise. Equation (1-41) can be expanded to account for bandwidth, but accounting for complex terminations requires the use of *noise correlation matrix* techniques as described by Rohde et al. [9].

1-7-2 System Amplitude and Phase Behavior

If we could build electronic systems that were absolutely amplitude- and phase-linear, radiocommunication system design would be greatly simplified. An amplifier designed for a power gain of 10 dB, for example, would merely increase the power of all signals at its input by a factor of 10, regardless of their frequencies, while perfectly maintaining their relative phases. But all real electrical and electronic networks, even those designed (or supposed) to be amplitude- and phase-linear, exhibit amplitude and phase nonlinearity to some degree, just as they all generate noise to some degree.

The effects of amplitude nonlinearity, generically referred to as *nonlinear distortion*, include the generation, through *harmonic distortion* and/or *intermodulation distortion* (IMD), of output signals at frequencies not present at a system's input. Nonlinear distortion also results in *gain compression*—changes in system gain with changes in input-signal level. By convention, when workers in electronics refer to or consider a network's "linearity," they usually mean its *amplitude* linearity; likewise, by "distortion" they usually mean *nonlinear* distortion.

The effects of phase or frequency nonlinearity are generically referred to as *linear distortion* because they occur independently of signal amplitude and polarity. We often intentionally apply linear distortion through *filtering*, which modifies the amplitude relationships among existing spectral components of a signal without creating any new frequencies. Another linear-distortion effect, *phase* or *delay distortion*, results in the delay of signals of differing frequencies by differing amounts of time. In a system where signal phase conveys

information, as is true of most wireless links, phase distortion can seriously degrade communication.

The fact that all real networks are amplitude- and angle-nonlinear to some degree means that all real networks modify the amplitude and angle characteristics of the signals they handle. What is perhaps less obvious is that subjecting a signal to amplitude and angle nonlinearities causes "crosstalk" between its amplitude and angle characteristics. For example, through a nonlinear distortion effect called AM-to-PM conversion, changes in a signal's amplitude result in changes in its phase. Filtering an angle-modulated signal produces the reverse effect: Deprived by the filter's selectivity of spectral components necessary to maintain its envelope constancy, the signal emerges from the filter as a combination of angle and amplitude modulation.

Spectral Considerations of Analog and Digitally Modulated Signals. Many wireless modulation schemes exist and even more are proposed; all use angle (usually phase) modulation or a combination of phase and amplitude modulation using an emission format engineered to achieve multiple goals of information throughput, robustness, spectral and hardware efficiency, and reproducibility. Digital modulation is standard in mainstream wireless applications because it allows increased channel capacity, and immunity to noise and distortion, compared to analog systems. At the circuit-design level, the particulars of whether, or to what degree, a modulation scheme is AM, PM, analog, or digital matters less than the actual angle and amplitude characteristics of the signal(s) involved, and the tolerances within which these characteristics can be expected and allowed to vary.

A system's amplitude linearity is of major concern because of its relation to energy efficiency, receiver dynamic range, and transmitter spectral purity. Energy efficiency, important because many wireless applications use battery power, generally decreases inversely with amplitude linearity. Yet, sufficient amplitude linearity must be guaranteed in the front-end circuitry of wireless receivers subjected to multiple strong signals, and in all wireless receiver and transmitter stages handling variable-envelope signals.

Some digital modulation schemes, designed to be distortion-tolerant, produce constant-envelope signals; Gaussian minimum shift keying (GMSK) and Feher's quadrature phase shift keying (FQPSK) are examples of these. Such signals can be processed in highly nonlinear circuitry—for example, an RF power amplifier operated in saturation to maximize efficiency—without degrading their spectral composition. Other digital modulation schemes result in the emission—ideally—of a single PM sideband, with the carrier and all other sideband components suppressed. Such a signal exhibits, of necessity, phase and amplitude variations; the IS-54 cellular standard's $\pi/4$ differential quadrature phase shift keying ($\pi/4$-DQPSK) scheme, which results in envelope fluctuations of 3–6 dB, exemplifies this [12]. Intermodulation distortion among such a signal's spectral components can generate products that fall outside the bandwidth painstakingly controlled in the modulation process. This *spectral regrowth* must be minimized to prevent adjacent-channel interference (Figure 1-101).

For a better understanding, Figure 1-102 shows the various channels and the energy levels associated with them.

Amplitude Linearity Issues and Figures of Merit. A network's amplitude nonlinearity can be characterized by the expansion

$$y = k_1 f(x) + k_2[f(x)]^2 + k_3[f(x)]^3 + \text{higher-order terms} \tag{1-42}$$

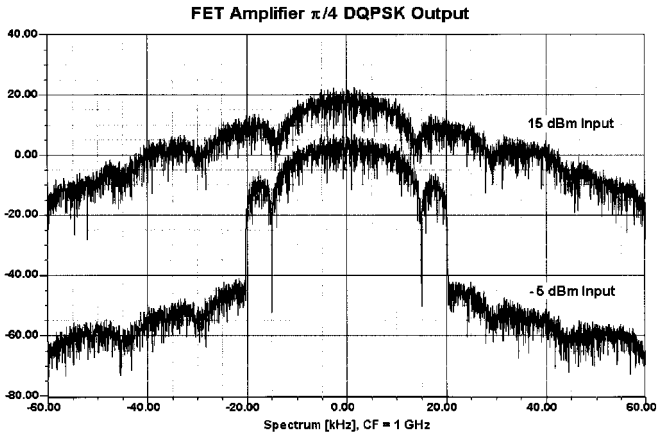


Figure 1-101 *Spectral regrowth* results when a variable-envelope emission, π/4-DQPSK in this case, is subjected to significant nonlinear distortion. This graph shows the simulated performance of a MESFET power amplifier operating at 1 GHz; the amplifier is 6 dB into compression when driven at 15 dBm.

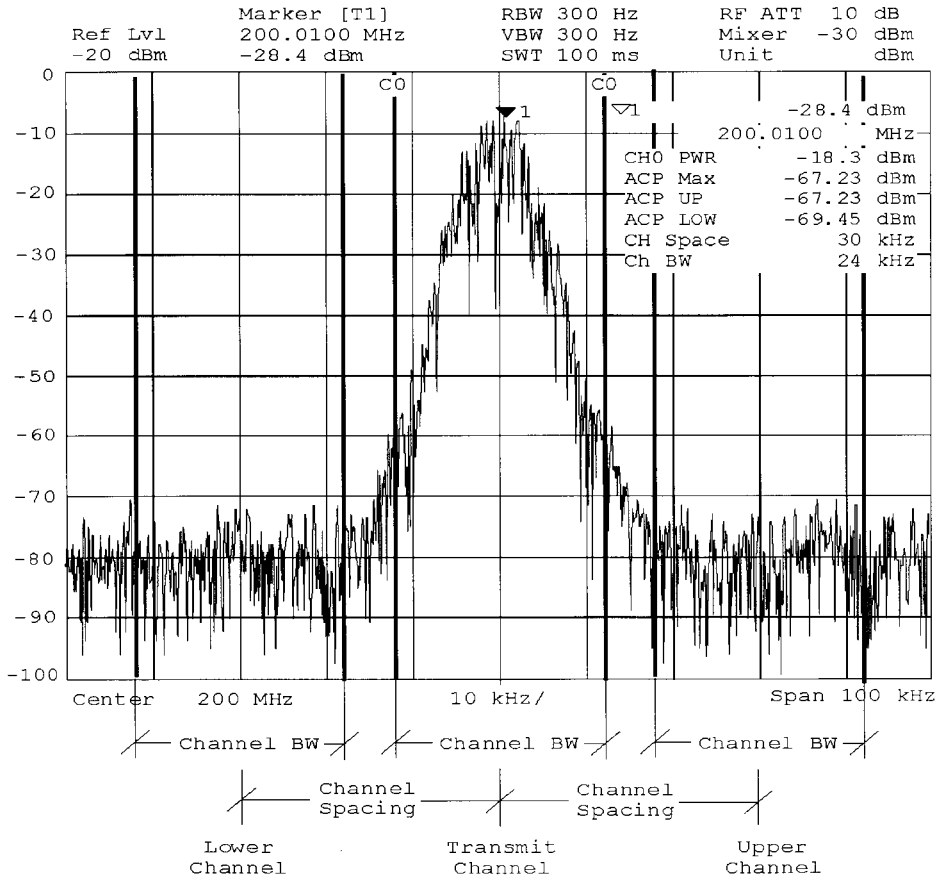


Figure 1-102 NADC signal and parameters, including channel spacing and channel bandwidth.

where y represents the output, the coefficients k_n represent complex quantities whose values can be determined by an analysis of the output waveforms, and $f(x)$ represents the input. Even though all practical networks exhibit amplitude nonlinearity, we can (and often do) refer to many networks as "linear." We say this of networks that are *sufficiently amplitude-linear for our purposes*—for example, weakly nonlinear networks in which small-signal operation is assumed even though the signal levels involved are sufficient to cause slight distortion. For many practical purposes, the first three terms of Eq. (1-42) adequately describe such a network's nonlinearity:

$$y = k_1 f(x) + k_2[f(x)]^2 + k_3[f(x)]^3 \tag{1-43}$$

In adopting this simplification, we assume also that the nonlinearity is frequency independent—that is, the network has sufficient bandwidth to allow all of the products predicted by Eq. (1-43) to appear at its output terminals unperturbed [13].

When multiple signals are present in a network, even weak nonlinearity can result in profound consequences. To illustrate this, we'll let $f(x)$ consist of two sinusoidal signals:

$$f(x) = A_1 \cos\omega_1 t + A_2 \cos\omega_2 t \tag{1-44}$$

We'll assume that ω_1 and ω_2 are close enough so that the coefficients k_i can be considered equal for both signals. We'll also assume for simplicity that all of the k_i are real. If Eq. (1-43) describes the network's response to an input $f(x)$, the response will be

$$\begin{aligned} y &= k_1(A_1 \cos\omega_1 t + A_2 \cos\omega_2 t) + k_2(A_1 \cos\omega_1 t + A_2 \cos\omega_2 t)^2 + k_3(A_1 \cos\omega_1 t + A_2 \cos\omega_2 t)^3 \\ &= k_1(A_2 \cos\omega_1 t + A_2 \cos\omega_2 t) \\ &\quad + k_2\left[A_1^2 \frac{1+\cos 2\omega_1 t}{2} + A_2^2 \frac{1+\cos 2\omega_2 t}{2} + A_1 A_2 \frac{\cos(\omega_1+\omega_2)t + \cos(\omega_1-\omega_2)t}{2}\right] \\ &\quad + k_3 \left\{ \begin{array}{l} \left[\left[A_1^3\left(\dfrac{\cos\omega_1 t}{2} + \dfrac{\cos\omega_2 t}{4} + \dfrac{\cos 3\omega_1 t}{4}\right) + A_2^3\left(\dfrac{3\cos\omega_2 t}{4} + \dfrac{\cos 3\omega_2 t}{4}\right)\right]\right] \\ + A_1^2 A_2[\tfrac{3}{2}\cos\omega_2 t + \tfrac{3}{4}\cos(2\omega_1 - \omega_2)t + \tfrac{3}{4}\cos(2\omega_1 - \omega_2)t] \\ + A_2^2 A_1[\tfrac{3}{2}\cos\omega_1 t + \tfrac{3}{4}\cos(2\omega_2 + \omega_1)t + \tfrac{3}{4}\cos(2\omega_2 - \omega_1)t] \end{array} \right\} \end{aligned} \tag{1-45}$$

The k_1 term of Eq. (1-45) represents the results of amplitude-linear behavior. No new frequency components have appeared; the two sine waves have merely been "rescaled" by k_1.

The second- and third-order terms of Eq. (1-45) represent the effects of harmonic distortion and intermodulation distortion. Second-order effects include second-harmonic distortion (the production of new signals at $2\omega_1$ and $2\omega_2$) and IMD (the production of new signals at $\omega_1 + \omega_2$ and $\omega_1 - \omega_2$). Third-order effects include gain compression, third-harmonic distortion (the production of new signals at $3\omega_1$ and $3\omega_2$), and IMD (the production of new signals at $2\omega_1 \pm \omega_2$ and $2\omega_2 \pm \omega_1$).

Gain Compression. Gain compression occurs when a network cannot increase its output amplitude in linear proportion to an amplitude increase at its input; gain *saturation* occurs when a network's output amplitude stops increasing (in practice, it may actually decrease) with increases in input amplitude. We can deduce from Eq. (1-45) that the amplitude of the $\cos\omega_1 t$ signal has become

$$A_1' = k_1 A_1 + k_3\left(\tfrac{3}{4}A_1^3 + \frac{3}{2}A_1 A_2^2\right) \tag{1-46}$$

Because k_3 will normally be negative, a large signal $A_2 \cos\omega_2 t$ can effectively mask a smaller signal $A_1 \cos\omega_1 t$ by reducing the network's gain. This third-order effect, known as *blocking* or *desensitization* when it occurs in a receiver, is a special case of gain compression. The presence of additional signals results in a greater reduction in gain; the gain reduction for each signal is a function of the relative levels of all signals present. A receiver's blocking behavior may be characterized in terms of the level of off-channel signal necessary to reduce the strength of an in-passband signal by a specified value, typically 1 dB; alternatively, the decibel ratio of the off-channel signal's power to the receiver's noise-floor power may be cited as *blocking dynamic range*. Desensitization may also be characterized in terms of the off-channel signal power necessary to degrade a system's SNR by a specified value.

Multiple signals need not be present for gain compression to occur. If only one signal is present, the ratio of gain with distortion to the network's idealized (linear) gain is

$$A_1' = \frac{k_1 + k_3\left(\tfrac{3}{4}A_1^2\right)}{k_1} \tag{1-47}$$

This is referred to as the *single-tone gain-compression factor*. Figure 1-103 shows how the k_3 term causes a network's gain to deviate from the ideal. The point at which a network's power gain is down 1 dB from the ideal for a single signal is a figure of merit known as the *1-dB compression point* (P_{-1dB}). Many networks (including many receiving and low-level transmitting circuits, such as low-noise amplifiers, mixers, and IF amplifiers) are usually operated under small-signal conditions—at levels sufficiently below P_{-1dB} to maintain high linearity. As we'll see, however, some networks (including power amplifiers for wireless systems) may be operated under large-signal conditions—near or in compression—to achieve optimum efficiency at some specified level of linearity. Figure 1-104 shows what happens when a digital emission that uses amplitude to convey information is subjected to amplitude compression.

Intermodulation. The new signals produced through intermodulation distortion (IMD) can profoundly affect the performance even of systems operated far below gain compression (Figure 1-105). IMD products of significant power can appear at frequencies remote from, in, and/or near the system passband, resulting in demodulation errors (in reception) and interference to other communications (in transmission). Where an IMD product appears relative to the passband depends on the passband width and center frequency, the frequencies of the signals present at the system input, and the order of the nonlinearity involved. These factors also determine the strength of an IMD product relative to the desired signal.

Second-order IMD (IM_2) results, for an input consisting of two signals ω_1 and ω_2, in the production of new signals at $\omega_1 + \omega_2$ and $\omega_1 - \omega_2$; third-order IMD (IM_3) results, for an input

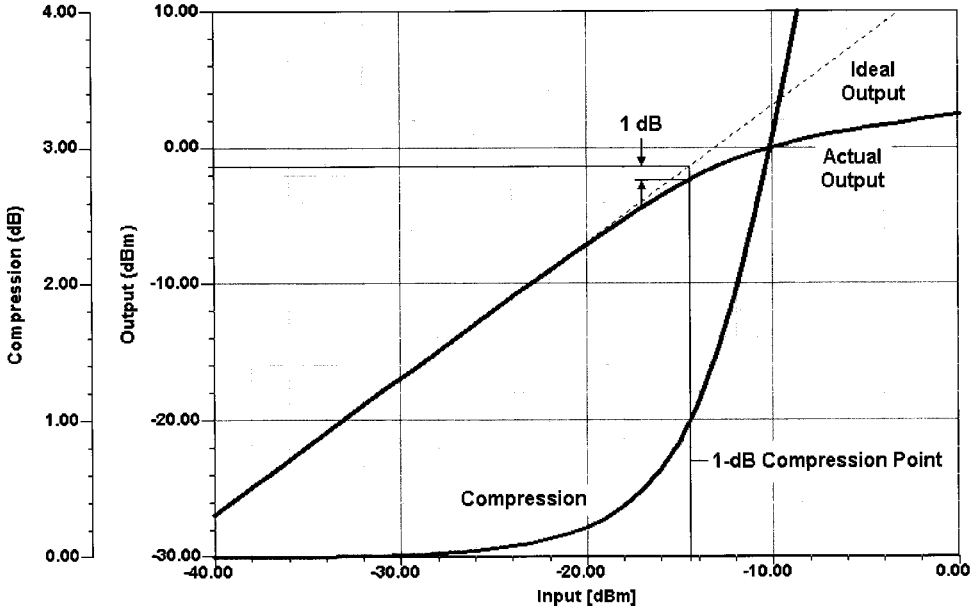


Figure 1-103 The power level at which a network's power output is down 1 dB relative to that of its ideally linear equivalent is a figure of merit known as the *1-dB compression point* (P_{-1dB}). The 1-dB compression point can be expressed relative to input power ($P_{-1dB,in}$) or output power ($P_{-1dB,out}$). For the amplifier simulated here, $P_{-1dB,in} \approx -14.5$ dBm and $P_{-1dB,out} \approx -1.3$ dBm.

consisting of two signals ω_1 and ω_2, in the production of new signals at $2\omega_1 \pm \omega_2$ and $2\omega_2 \pm \omega_1$.

Under small-signal conditions—that is, at levels well below compression—the power of an IM_2 product varies by 2 dB, and the power of an IM_3 product varies by 3 dB, per decibel change in input power level. This allows us to derive a network figure of merit, the *intermodulation intercept point* (*IP*), for a given IM order by extrapolating a network's linear and IM responses to their point of intersection (Figure 1-106)—the point at which their powers would be equal if compression did not occur. The intercept point for a given IM order

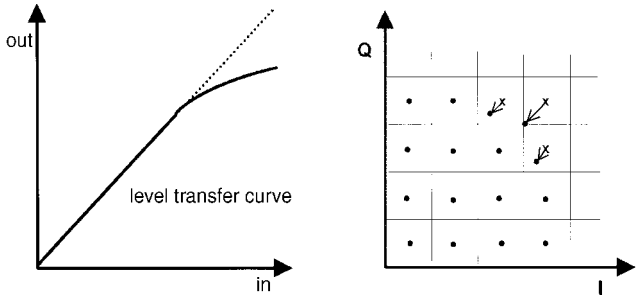


Figure 1-104 Influence of differential amplitude error (compression) on a QAM constellation.

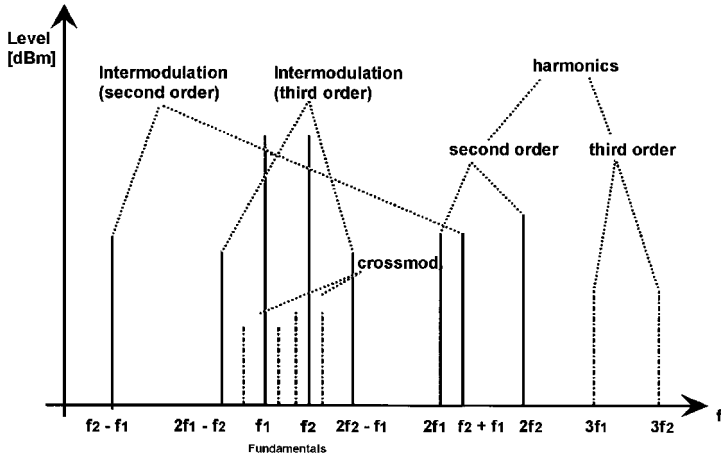


Figure 1-105 Relationships between fundamental and spurious signals, including harmonics and products of intermodulation.

n can be expressed, and should always be characterized, relative to input power ($IP_{n,\text{in}}$) or output power ($IP_{n,\text{out}}$); the IP_{in} and IP_{out} values differ by the network's linear gain. For equal-level test tones, $IP_{n,\text{in}}$ can be determined by

$$IP_{n,\text{in}} = \frac{nP_A - P_{\text{IM}_n}}{n-1} \tag{1-48}$$

where n is the order, P_A is the input power (of one tone), P_{IM_n} is the power of the IM product, and IP is the intercept point. The intercept point for cascaded networks can be determined from

$$IP_{2,\text{in}} = \frac{1}{\left(\frac{1}{\sqrt{IP1}} + \frac{G}{\sqrt{IP2}}\right)^2} \tag{1-49}$$

for IP_2 and from

$$IP_{3,\text{in}} = \frac{1}{\frac{1}{IP1} + \frac{G}{IP2}} \tag{1-50}$$

for IP_3. In both equations, $IP1$ is the input intercept of Stage 1 in watts, $IP2$ is the input intercept of Stage 2 in watts, and G is the gain of Stage 1 (as a numerical ratio, *not* in decibels). Both equations assume the worst-case condition, in which the distortion products of both stages add in-phase.

The ratio of the signal power to the IM-product power, the *distortion ratio*, can be expressed as

$$R_{dn} = (n-1)\left[IP_{n,\text{in}} - P_{\text{in}}\right] \tag{1-51}$$

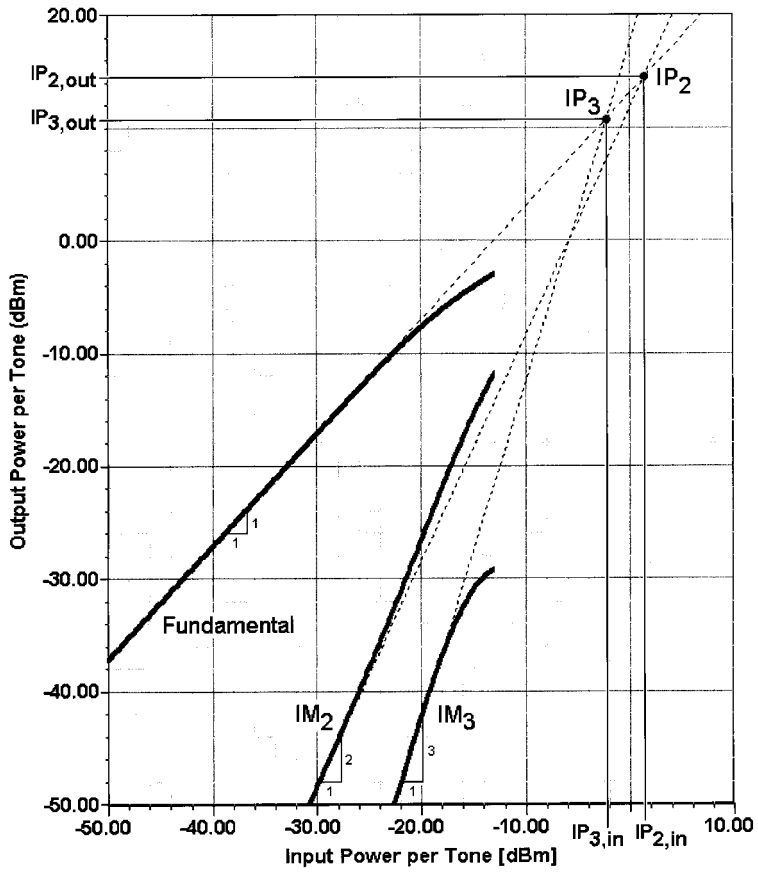


Figure 1-106 The level at which the power of one of a network's IM products equals that of the network's linear output is a figure of merit known as the *intermodulation intercept point* (*IP*). The intercept point for a given IM order *n* can be expressed, and should always be characterized, relative to input power ($IP_{n,in}$) or output power ($IP_{n,out}$); the IP_{in} and IP_{out} values differ by the network's linear gain. For the amplifier simulated here, $IP_{2,in} \approx 1.5$ dBm, $IP_{2,out} \approx 14.5$ dBm, $IP_{3,in} \approx -2.3$ dBm, and $IP_{3,out} \approx 10.7$ dBm. Each curve depicts the power in one tone of the response evaluated.

where n is the order, R_{dn} is the distortion ratio, $IP_{n,in}$ is the input intercept point, and P_{in} is the input power of one tone.

Discussions of IMD have traditionally downplayed the importance of IM_2 because the incidental distributed filtering contributed by the tuned circuitry once common in radiocommunication systems was usually enough to render out-of-passband IM_2 products caused by in-passband signals, and in-passband IM_2 products caused by out-of-passband signals, vanishingly weak compared to fundamental and IM_3 signals. In broadband systems that operate at bandwidths of an octave or more, however, in-passband signals may produce significantly strong in-passband IM_2 and second-harmonic products. In such applications, balanced circuit structures (such as push–pull amplifiers and balanced mixers) can be used to minimize IM_2 and other even-order nonlinear products.

As with IM_2, which IM_3 products are important depends on the spacing of the signals involved and the relative width of the system passband. If ω_1 and ω_2 are of approximately

the same frequency, the additive products $2\omega_1 + \omega_2$ and $2\omega_2 + \omega_1$ will be outside the passband of a narrowband system. The subtractive products $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$, however, will likely appear near or within the system passband. The IM_3 performance of any network subjected to multiple signals is therefore of critical importance, and an array of IM_3-related, sometimes application-specific, figures of merit has evolved as a result.

Dynamic Range. As we have seen, thermal noise sets the lower limit of the power span over which a network can operate. Distortion—that is, degradation by distortion of the signal's ability to convey information—sets the upper limit of a network's power span. Because the power level at which distortion becomes intolerable varies with signal type and application, a generic definition has evolved: The upper limit of a network's power span is the level at which the power of one IM product of a specified order is equal in power to the network's noise floor. The ratio of the noise-floor power to the upper-limit signal power is referred to as the network's *dynamic range* (*DR*), often more carefully characterized as *two-tone IMD dynamic range*, which, when evaluated with equal-power test tones, is a figure of merit commonly used to characterize receivers. When $IP_{n,\text{in}}$ and MDS are known, IMD DR can be determined from

$$DR_n = \frac{(n-1)[IP_{n,\text{in}} - MDS]}{n} \tag{1-52}$$

where *DR* is the dynamic range in decibels, *n* is the order, IP_{in} is the input intercept power in dBm, and *MDS* is the minimum detectable signal power in dBm.

Dynamic Measure: A New Receiver Figure of Merit. Taken by themselves, the IP_n and DR figures of merit can be misleading in the sense that broadband, resistive attenuation directly improves them by the amount of attenuation added while directly degrading NF by the same amount. The NF and MDS figures of merit can mislead us because they convey no sense of suitability to task. This can be significant when we interconnect networks and systems of differing noise floors and IMD dynamic ranges—as we do, for example, when connecting a receiver to an antenna. As a means of better evaluating a receiver's front-end performance, the authors propose a new figure of merit, *dynamic measure*, defined as

$$DM = [(IP_{3,\text{in}} + \text{Att}) - (NF_r + \text{Att} + NF_{ant})]\frac{NF_r}{NF_r + \text{Att}} \tag{1-53}$$

where *DM* is the dynamic measure, a dimensionless number; $IP_{3,\text{in}}$ is the receiver's third-order input intercept in dBμV (instead of dBm) without any added attenuation; NF_r is the receiver noise figure in dB without any added attenuation; NF_{ant} is the antenna-system noise figure in dB as determined by Eq. (1-40); and Att is the decibel value of any added attenuation, which is sometimes used to shift the upper and lower limits of a system's dynamic range to higher absolute signal levels, increasing $IP_{3,\text{in}}$ at the expense of making the system more "deaf." Our dynamic measure figure of merit takes this trade-off into account to allow more straightforward comparisons of systems that have the same dynamic range but different sensitivities. Setting NF_{ant} equal to zero allows evaluation of receiver performance under laboratory conditions; setting NF_{ant} equal to the measured or expected NF of the antenna system to which the receiver will be connected allows evaluation of receiver performance in a system sense. Rather than express IP_3 in dBm, we are going to

express it in dBμV. This is necessary because for negative values of the difference between IP_3 and the added noise figure, the multiplier term will make the result larger when its numerical value is less than 1.

The usefulness of the dynamic measure figure of merit is evident when we consider several examples.

Example 1. Here we evaluate a receiver that exhibits the characteristics $IP_{3,\text{in}} = 16$ dBm(123 dBμV), $NF_r = 8$ dB, and $NF_{\text{ant}} = 10$ dB. We calculate its laboratory DM by setting NF_{ant} equal to 0:

$$DM = [123 - (8 + 0)]\frac{8}{8 + 0} = [115]1 = 115 \tag{1-54}$$

Setting NF_{ant} equal to 10 returns the receiver's system DM:

$$DM = [123 - (8 + 10)]\frac{8}{8 + 0} = [115]1 = 105 \tag{1-55}$$

Example 2. This setup consists of the Example 1 receiver with a 10-dB pad switched in, resulting in the characteristics $IP_{3,\text{in}} = 26$ dBm(133 dBμV), $NF_r = 18$ dB, and $NF_{\text{ant}} = 10$ dB. Setting NF_{ant} equal to 0, we calculate its laboratory DM as

$$DM = [133 - (8 + 10 + 0)]\frac{8}{8 + 10} = [115]0.4444 = 51.1 \tag{1-56}$$

A comparison with Eq. (1-54) shows that, although inserting the 10-dB pad has shifted $IP_{3,\text{in}}$ from 16 dBm (123 dBμV) to 26 dBm (133 dBμV), it has also increased the resulting NF from 8 to 18 dB, reducing the DM from 115 to 51.1. Inserting the pad also degrades the receiver's system DM compared to Example 1, as we see by setting NF_{ant} equal to 10:

$$DM = [133 - (8 + 10 + 10)]\frac{8}{8 + 10} = [105]0.4444 = 46.7 \tag{1-57}$$

Example 3. On the other hand, had we inserted 2 dB of attenuation rather than 10 dB [$IP_{3,\text{in}} = 18$ dBm (125 dBμV), $NF_r = 10$ dB, and $NF_{\text{ant}} = 10$ dB], the laboratory DM would have been

$$DM = [125 - (8 + 2 + 0)]\frac{8}{8 + 2} = [115]0.8 = 92 \tag{1-58}$$

and the system DM would have been

$$DM = [125 - (8 + 2 + 10)]\frac{8}{8 + 2} = [105]0.8 = 84 \tag{1-59}$$

Table 1-12 shows these and several additional examples for comparison.

The interested reader will notice the following: The receiver in Example 1 (base NF 8 dB) was unnecessarily sensitive compared to the antenna NF of 10 dB, so adding a 2-dB pad increased the intercept point by 2 dB and matched the antenna noise floor. Yet, the equation rightfully indicates a loss of dynamic measure. The relationship between NF_{ant} and NF_r is not taken

Table 1-12 Eight dynamic measure cases compared

Example	System	Base $IP_{3,in}$ (dBm)	Base $IP_{3,in}$ (dBμV)	Attenuation (dB)	Resulting $IP_{3,in}$ (dBm)	Resulting $IP_{3,in}$ (dBμV)	Base NF_r (dB)	Resulting NF_r (dB)	NF_{ant} (dB)	DM_{lab}	DM_{system}
1	HF receiver	16.0	123.0	0.0	16.0	123.0	8.0	8.0	10.0	115.0	105.0
2	HF receiver with 10-dB pad	16.0	123.0	10.0	26.0	133.0	8.0	18.0	10.0	51.1	46.7
3	HF receiver with 2-dB pad	16.0	123.0	2.0	18.0	125.0	8.0	10.0	10.0	92.0	84.0
4	Rohde & Schwarz XK 2100 shortwave transceiver	40.0	147.0	0.0	40.0	147.0	10.0	10.0	10.0	137.0	127.0
5	XK 2100 with 10 dB of unnecessary attenuation	40.0	147.0	10.0	50.0	157.0	10.0	20.0	10.0	68.5	63.5
6	Satellite receiver	–2.2	104.8	0.0	–2.2	104.8	0.9	0.9	0.2	103.9	103.7
7	Wireless front end without duplexer/filter losses	–12.0	95.0	0.0	–12.0	95.0	2.6	2.6	2.0	92.4	90.4
8	Wireless front end with 3-dB duplexer/filter losses	–12.0	95.0	3.0	–9.0	98.0	2.6	5.6	2.0	42.9	42.0

into consideration. Also, we have decided to leave it to the reader to develop a formula that deals with the addition of a preamplifier rather than an attenuator. This will complicate the formula because the preamplifier's gain and noise figure are not necessarily related.

Triple-Beat Distortion and Cross-Modulation. P_{-1dB} is a single-tone figure of merit; blocking, intercept point, and dynamic range evaluate two-tone behavior. For networks that must handle AM and composite (AM and angle modulation) signals very linearly, such as television transmitters and cable TV distribution systems, a three-tone figure of merit called *triple-beat distortion* has gained acceptance. Signals at ω_1 and ω_2 (closely spaced) and ω_3 (positioned far away from ω_1 and ω_2) are applied to the network under test, at levels, frequencies, and spacings that vary with the application. One triple-beat distortion figure of merit is the ratio, expressed in decibels, of the IM product at $\omega_3 + (\omega_2 - \omega_1)$ to one of the network's linear outputs at a specified output level. Alternatively, the triple-beat figure of merit may express the network output level at which a specified triple-beat ratio occurs.

Triple-beat distortion is the mechanism underlying cross-modulation, a form of intermodulation in which one or more AM signals present in a network amplitude-modulate all

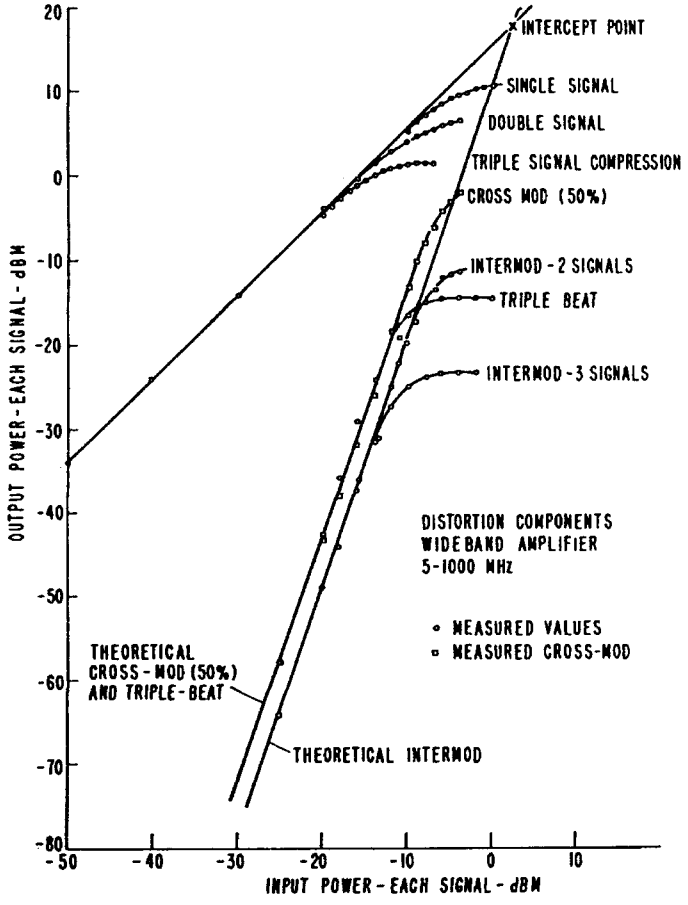


Figure 1-107 Measured distortion components in a wideband (5–1000 MHz) amplifier. Figure 1-108 shows a magnified view of the gain-compression region [15].

signals present in the network [14]. Angle-modulation-based wireless systems are largely immune to such effects.

Figures 1-107 and 1-108 graph the results of gain compression, two-tone intermodulation, cross-modulation, and triple-beat testing on a wideband (5–1000 MHz) amplifier.

Noise Power Ratio. Triple-beat testing is one way of improving on two-tone testing as a means of evaluating a network's intermodulation behavior in the presence of multiple signals. Another figure of merit, *noise power ratio* (*NPR*), uses thermal noise as a test signal. The test measures the introduction, by IM, of noise into a quiet slot created by the insertion of a bandstop filter, equal in stopband width to the width of the measurement channel, between the noise generator and the network under test (Figure 1-109).

Large-Signal Effects. Except for P_{-1dB}, the figures of merit discussed so far evaluate amplitude nonlinearity under small-signal conditions. At input powers that drive a network into gain compression and saturation, IM products of odd orders higher than 3 become significant, and curves, dips, and nulls appear in their characteristics (Figure 1-110). Phase shifts related to nonlinear (primarily voltage-dependent) capacitances in solid-state devices are one cause of these effects. Under such conditions, a network may exhibit hysteresis, with its behavior at any given instant depending not only on the voltage or current applied to it but also on its recent history [16].

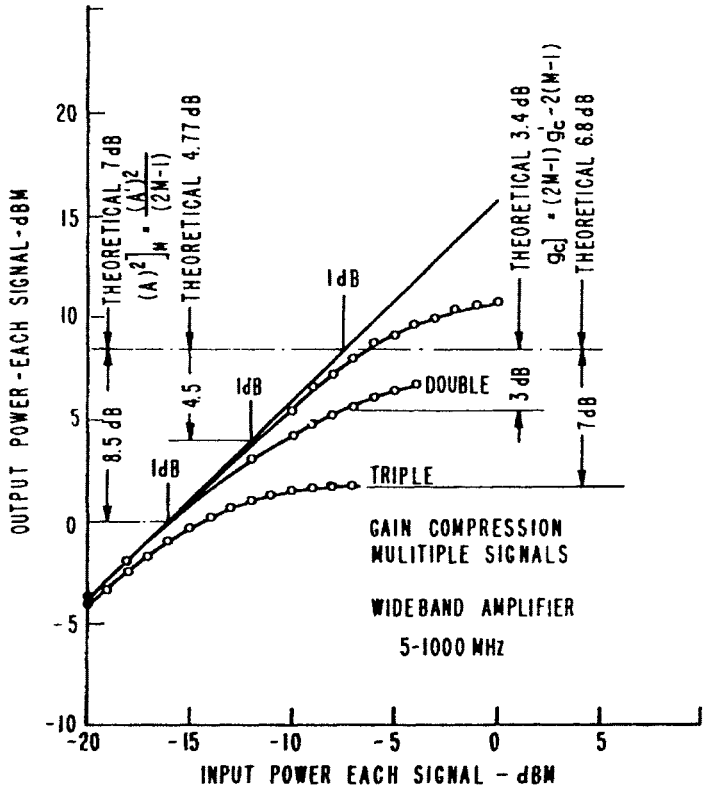


Figure 1-108 Measured multiple-signal gain compression of the 5- to 1000-MHz amplifier [15].

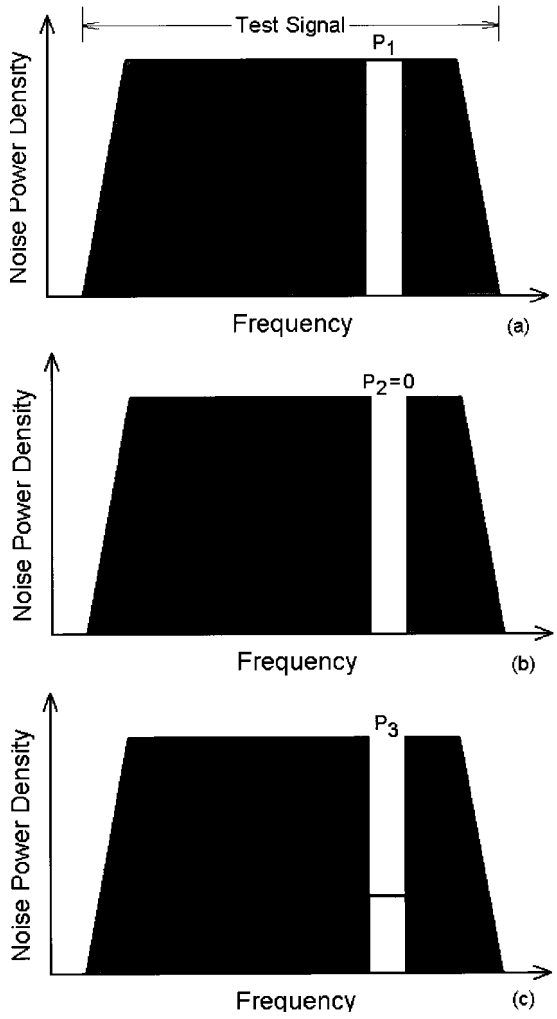


Figure 1-109 Determining a network's *noise power ratio* (*NPR*) involves the application of a test signal consisting of thermal noise [15]. (a) The reference measurement-channel noise power, P_1, is then measured. (b) Next, a bandstop filter is placed between the noise generator and network under test to keep the test signal out of the measurement channel. Assuming sufficient filter attenuation, if the network were absolutely noiseless and linear, the ideal noise power in the measurement channel, P_2, would then be zero. (c) In practice, the network's own thermal noise and intermodulation between noise components outside the measurement channel result in an actual measurement-channel noise power (P_3) greater than zero. The noise power ratio equals P_1/P_3.

AM-to-PM Conversion. The nonlinear distortion effects we've discussed so far can be termed *AM-to-AM distortion*—distortion that, to a degree that depends on the amplitude of the signal(s) applied to the network, results in changes in the network's gain and/or production of signals at new frequencies. AM-to-PM distortion can also occur. As a network nears saturation, part of its driving signal goes into shifting the bias point(s) of its active device(s), changing their drive-dependent reactances and shifting the phase of the output signal relative to its value at input levels below compression (Figures 1-111 and 1-112). This

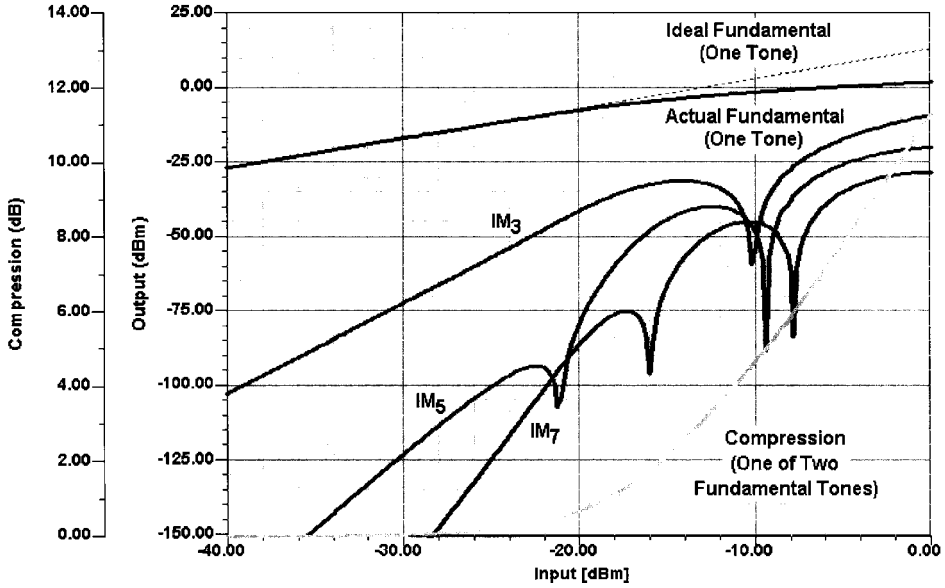


Figure 1-110 As a network is driven into compression, IM products at odd orders higher than 3 become significant, and phase shifts in power-dependent device capacitances cause curves and dips in the IM characteristics. The onset of these departures from IM-response linearity occurs at generally lower input power levels for higher IM orders; their severity, and their position on the IM curves, differs among the various products of a given order and varies with network topology and tone spacing. Figures of merit based on straight-line IM responses fail to usefully predict nonlinear network behavior under these conditions. This graph shows the simulated performance of a single-BJT broadband amplifier driven by two equal-amplitude tones at 10 and 11 MHz.

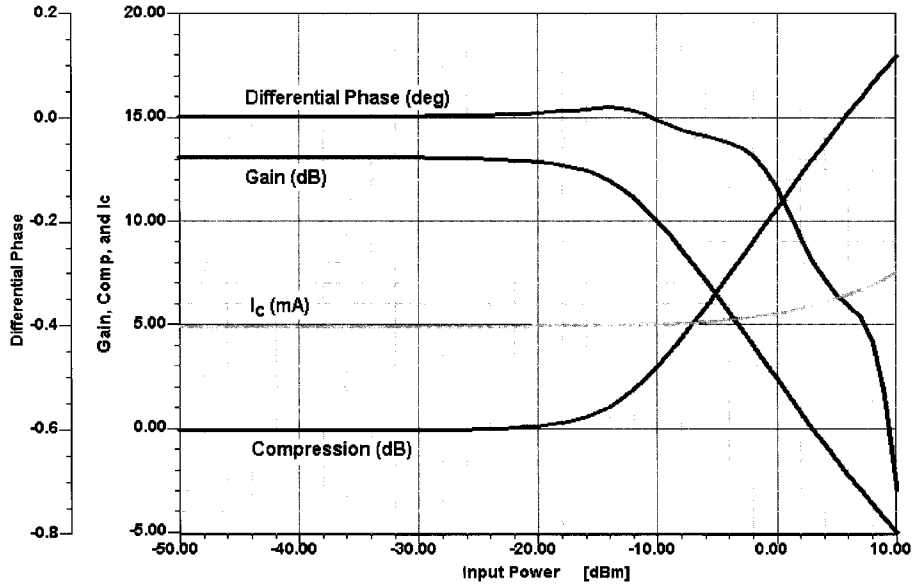


Figure 1-111 Driving a network into compression and saturation shifts the bias point(s) of its active device(s), changing their drive-dependent reactances and shifting the phase of the output signal relative to its value at input levels below compression. This graph shows the simulated performance of a single-BJT broadband amplifier driven by a single tone at 10 MHz.

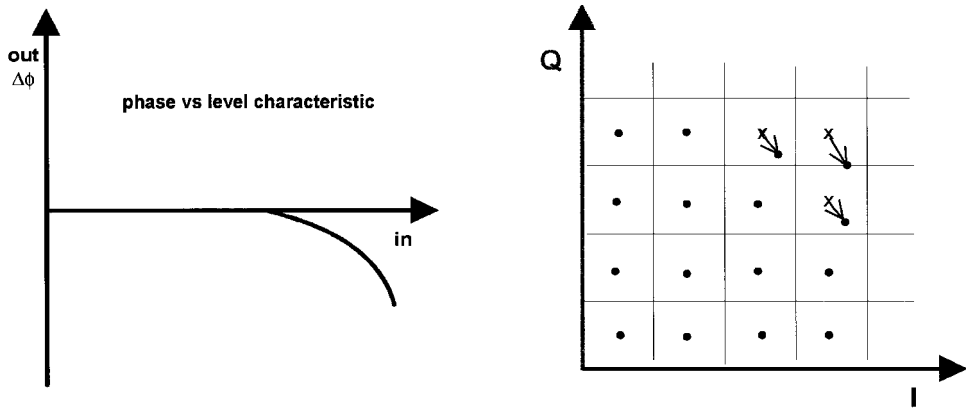


Figure 1-112 Influence of differential phase error (AM-to-PM conversion) on a QAM constellation.

effect, *AM-to-PM conversion*, can cause incidental phase modulation that degrades the performance of digital communication systems.

Spectral Regrowth and Adjacent-Channel Power Ratio. Spectral regrowth occurs largely as a result of third-, fifth-, and seventh-order IMD in power amplifiers operated near or in compression—at power levels where hysteretic IM effects result in poor agreement between measured behavior and predictions based on small-signal IM figures of merit [17]. We therefore evaluate the impact of spectral regrowth more directly, using a figure of merit called *adjacent-channel power ratio* (ACPR). ACPR measurement techniques that incorporate memory can be used to increase ACPR predictions for networks that exhibit saturation hysteresis [18]. Figure 1-113 shows the critical relationship between compression, power-added efficiency, and ACPR in a MESFET power amplifier.

Phase Response Issues and Figures of Merit. We have already seen how large-signal nonlinear distortion can result in amplitude-dependent phase shifts through AM-to-PM conversion. Because phase linearity is critical at *all* signal levels in PM systems, especially those using digital modulation, we must also consider linear distortion in evaluating networks used in wireless systems.

Differential Group Delay. Every frequency-selective network subjects signals passing through it to some degree of time delay. Ideally, this delay, also known as *group* or *envelope delay*, does not vary with frequency; that is, the network's phase shift versus frequency response is monotonic and linear. In practice, a network's time delay varies across its passband, transition bands, and stopbands, exhibiting curvature, ripple, and transition-band peaks (Figure 1-114). The network's *differential* group delay—its group-delay spread—is therefore of considerable importance. This is especially so in digitally modulated systems, where the resulting phase distortion can cause errors in modulation and demodulation.

Effects of Phase Noise. As Chapter 5 will discuss in detail, the phase of an oscillator's output signal is subject to random phase variations (Figures 1-115 and 1-116). Called *phase noise*, this effect is often quantified as the decibel ratio of the phase-noise power in a single

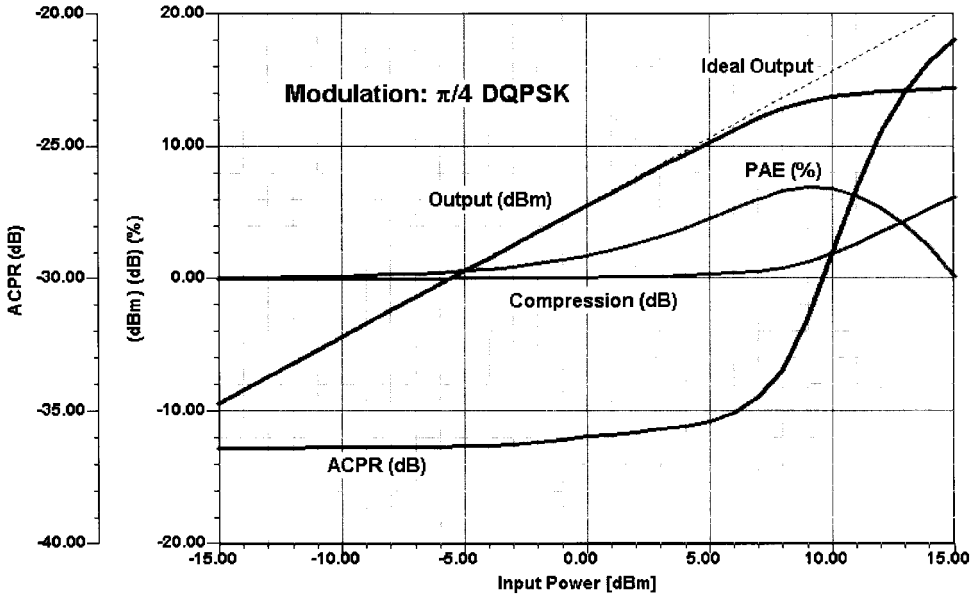


Figure 1-113 Keeping adjacent-channel interference under control can involve a critical trade-off between a wireless transmitter's power-added efficiency (PAE) and ACPR, as shown in this graph of the simulated behavior of the 1-GHz MESFET power amplifier introduced into Figure 1-101. As the amplifier is driven into compression, a peak occurs in the PAE response—in this case, at P_{-1dB}—and the ACPR rises sharply.

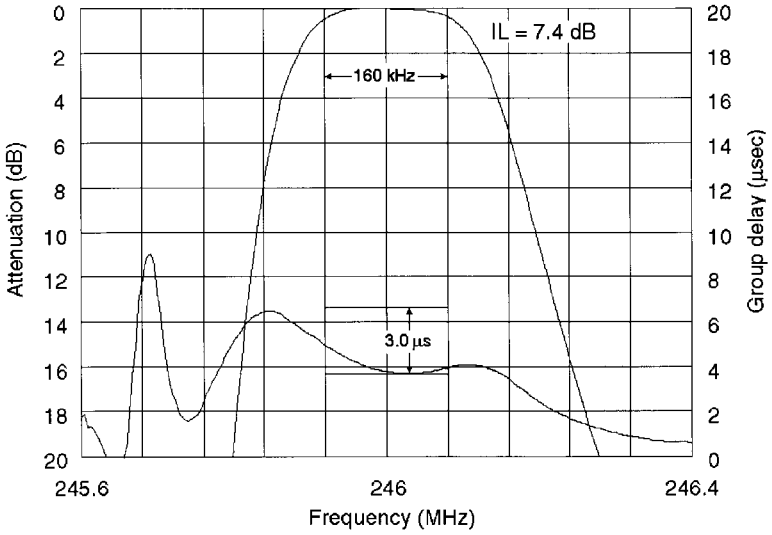


Figure 1-114 Close-in amplitude and group-delay responses for a 246-MHz SAW filter designed for GSM applications [19]. This filter is well within its 3.0-μs differential group-delay specification across its passband (160 kHz at –3 dB); the peaks just outside the passband limits are characteristic of a network's transition-band phase response.

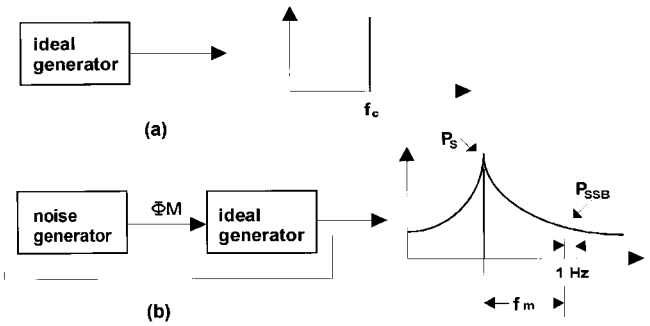


Figure 1-115 SSB phase noise. (a) An ideal signal generator would produce an absolutely pure carrier. (b) A real signal generator acts like an ideal generator driven by a noise generator, producing a noise-modulated carrier.

(the upper or lower) phase-noise sideband, in a 1-Hz bandwidth centered at a specified frequency offset from the oscillator carrier, to the carrier power (Figure 1-117); alternatively, it may be specified in degrees rms. A microwave voltage-controlled oscillator, for instance, might exhibit an SSB phase noise of –95 dBc/Hz at 10 kHz. Oscillator phase noise may manifest itself, through a mechanism known as *reciprocal mixing*, as the emission of unacceptably strong noise outside a transmitter's occupied bandwidth or as an increase in receiver noise floor. Phase noise may also directly introduce phase errors that result in modulation and demodulation errors.

Because the oscillators used for frequency translation in wireless systems are usually embedded in phase-locked loops, their phase-noise characteristics differ from those of "bare" oscillators as shown in Figures 1-118 and 1-119. Figures 1-120 and 1-121 show the measured phase noise of the Rohde & Schwarz SMY and SMIQ signal generators. The SMY is a low-cost signal source, while the SMIQ is a very high performance signal generator capable of being programmed for all digital modulations; therefore, their PLL systems exhibit different phase-noise versus frequency responses as the measured results show.

Reciprocal Mixing. In reciprocal mixing, incoming signals mix with LO sideband energy to produce IF output (Figure 1-122). Because one of the two signals is usually noise, the resulting IF output is usually noise. (Reciprocal mixing effects are not limited to noise;

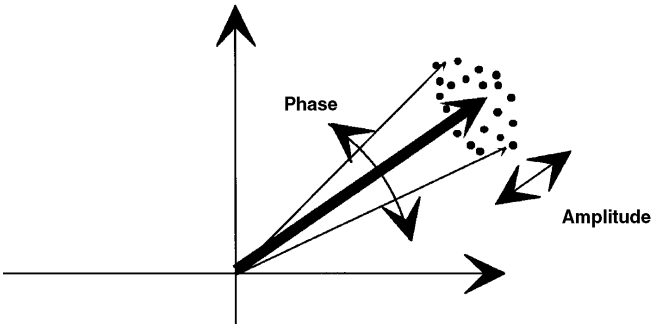
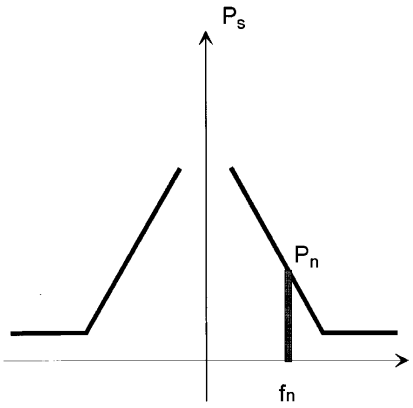


Figure 1-116 Oscillator noise can be split into amplitude and phase components.



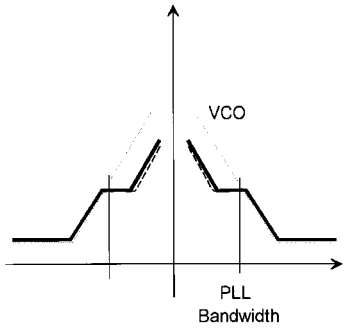
$$L(f) = 10 * \log \frac{P_n}{P_s}$$

P_n = Sideband noise in 1-Hz bandwidth at offset frequency f_n

P_S = Total signal power

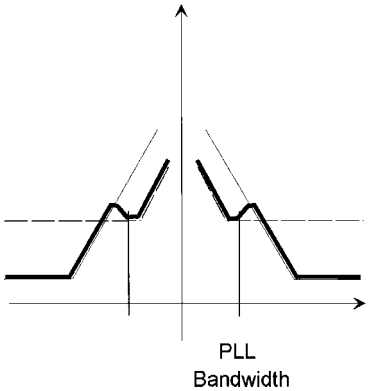
Single-sideband phase noise is expressed in dBc measured in a bandwidth of 1 Hz [dBc(Hz)] at offset frequency f_n.

Figure 1-117 Phase noise calculation.



Phase Noise is controlled by reference phase noise within PLL bandwidth and follows VCO noise outside PLL bandwidth

Figure 1-118 Phase noise of an oscillator controlled by a phase-locked loop.



Improper designed PLL bandwidth

Phase noise increases outside of PLL bandwidth

Figure 1-119 Effect of improper loop-filter design.

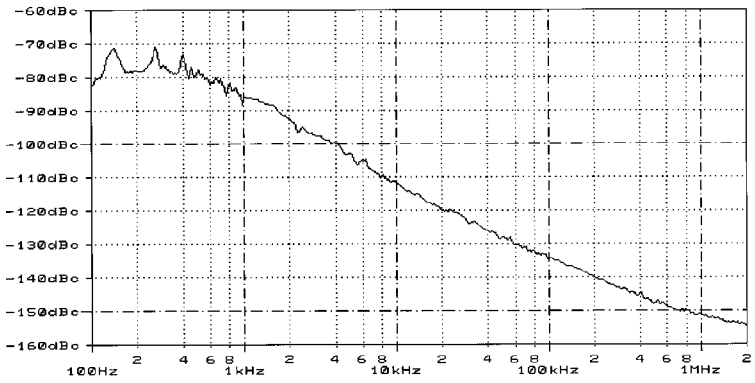


Figure 1-120 Measured phase noise of the Rohde & Schwarz SMY signal generator at 1 GHz. This signal generator has no provision for digital modulation and therefore shows the best possible phase noise in its class.

discrete-frequency oscillator sideband components, such as those resulting from crosstalk to or reference energy on a VCO's control line, or the discrete-frequency spurious signals endemic to direct digital synthesis, can also mix incoming signals to IF.) In practice, the resulting noise-floor increase can compromise the receiver's ability to detect weak signals and achieve a high IMD dynamic range; on the test bench, noise from reciprocal mixing may invalidate desensitization, cross-modulation, and IM testing by obscuring the weak signals that must be measured in making these tests.

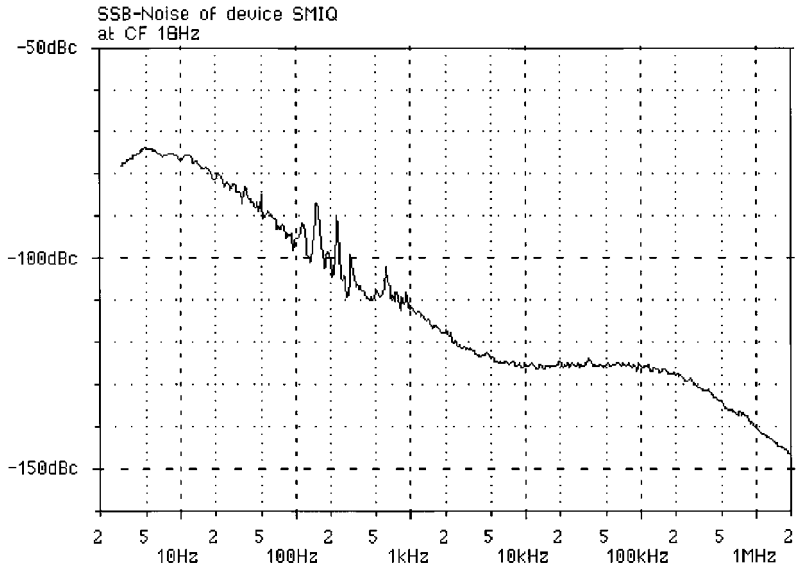


Figure 1-121 Measured phase noise of the Rohde & Schwarz SMIQ signal generator at 1 GHz. This signal generator is optimized for all digital modulation capabilities and can be configured via appropriate programming. Above 10 kHz, the influence of the wideband loop becomes noticeable; above 200 kHz, the resonator *Q* takes over.

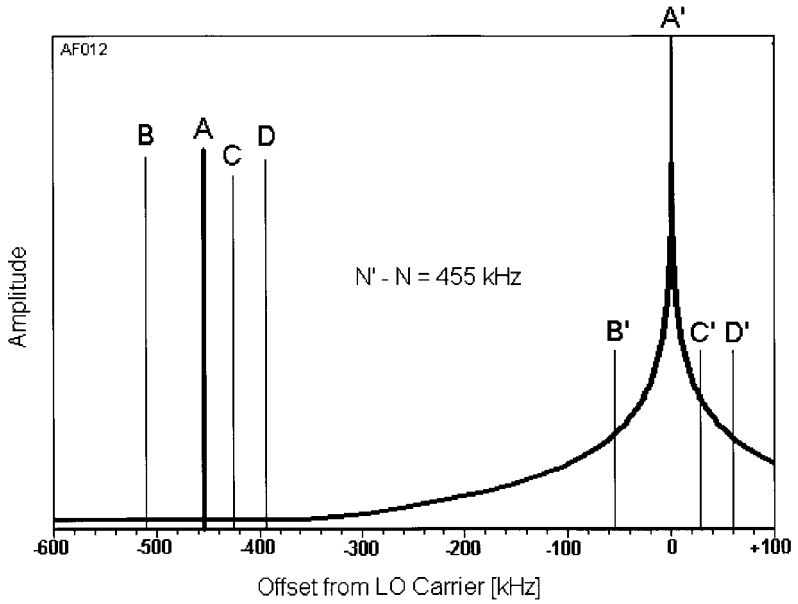


Figure 1-122 Reciprocal mixing occurs when incoming signals mix energy from an oscillator's sidebands to the IF. In this example, the oscillator is tuned so that its carrier, at A′, heterodynes the desired signal, A, to the 455 kHz as intended; at the same time, the undesired signals B, C, and D mix the oscillator noise sideband energy at B′, C′, and D′, respectively, to the IF. Depending on the levels of the interfering signals and the noise sideband energy, the result may be a significant rise in the receiver noise floor.

Figure 1-123 shows a typical arrangement of a dual-conversion receiver with local oscillators. The signal coming from the antenna is filtered by an arrangement of tuned circuits referred to as providing *input selectivity*. For a minimum attenuation in the passband, an operating bandwidth of

$$B = \frac{f}{\sqrt{2}Q_L}$$

is needed. This approximate formula is valid for the insertion loss of about 1 dB due to loaded Q.

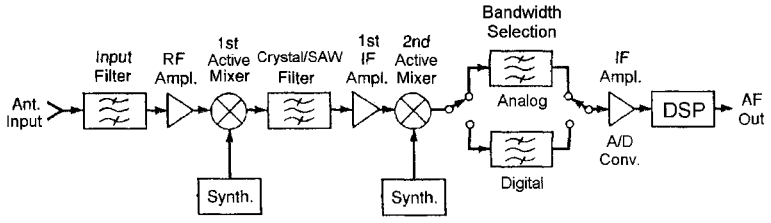


Figure 1-123 Block diagram of an analog/digital receiver showing the signal path from antenna to audio output. No AGC or other auxiliary circuits are shown. This receiver principle can be used for all types of modulation, since the demodulation is done in the DSP block.

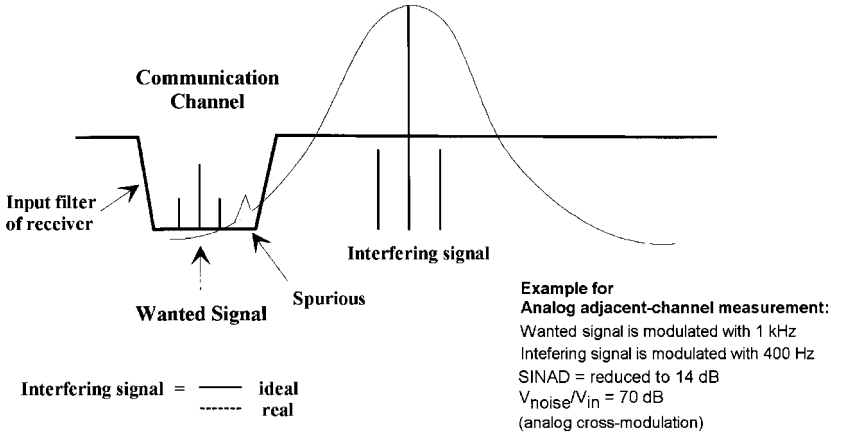
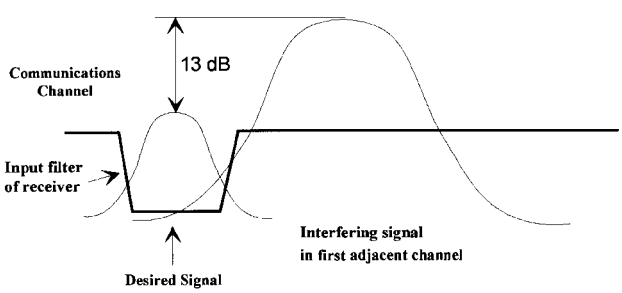


Figure 1-124 Principle of selectivity measurement for analog receivers.

The filter in the first IF is typically either a SAW filter (in the frequency range from 500 MHz to 1 GHz) or a crystal filter (45–120 MHz). Typical insertion loss is 6 dB. Since these resonators have significantly higher Q than LC circuits, the bandwidth for the first IF will vary from ±5 to ±500 kHz. It is now obvious that the first RF filter does not protect the first IF because of its wider bandwidth. For typical communication receivers, IF bandwidths from 150 Hz to 1 MHz are found; for digital modulation, the bandwidth varies roughly from 30 kHz to 1 MHz. Therefore, the IF filter in the second IF has to accommodate these bandwidths; otherwise, the second mixer easily gets into trouble from overloading. This also means that both synthesizer paths must be of low-noise and low-spurious design. The second IF of this arrangement (Figure 1-123) can be either analog or digital, or even zero-IF. In practical terms, there are good reasons for using IFs like 50/3 kHz, as in hi-fi audio equipment such as the Walkman, with DSP processing at this frequency (50/3 kHz) using the low-cost modules found in mass-market consumer products.

Figures 1-124 and 1-125 show the principle of selectivity measurement for both analog and digital signals. The main difference is that the occupied bandwidth for the digital system can be significantly wider, and yet both signals can be interfered with by either a noise synthesizer/first LO or a synthesizer that has unwanted spurious frequencies. Such a spurious signal is shown in Figure 1-124. In the case of Figure 1-124, the analog adjacent-channel



Example of ACP suppression:

	Requirement	Good Generator
NADC		
30kHz from carrier	<-26 dBc	-51 dBc
60kHz "	<-45 dBc	-70 dBc
90kHz "	<-49 dBc	-75 dBc
PDC		
50kHz from carrier	<-45 dBc	-65 dBc
100kHz "	<-60 dBc	-76 dBc

Figure 1-125 Principle of selectivity measurement for digital receivers.

measurement has some of the characteristics of cross-modulation and intermodulation, while in the digital system the problem with the adjacent-channel power suppression is more obvious. Rarely has the concept of adjacent-channel power (ACP) been used with analog systems. Also, to meet the standards, signal generators have to be used that are better, with some headroom, than the required dynamic measurement. We have therefore included in Figure 1-125 the achievable performance for a practical signal generator—in this case, the Rohde & Schwarz SMHU58.

Because reciprocal mixing produces the effect of noise leakage around IF filtering, it plays a role in determining a receiver's dynamic selectivity (Figure 1-126). There is little value in using IF filtering that exhibits more stopband rejection than a value 3–10 dB greater than that which results in an acceptable reciprocal mixing level.

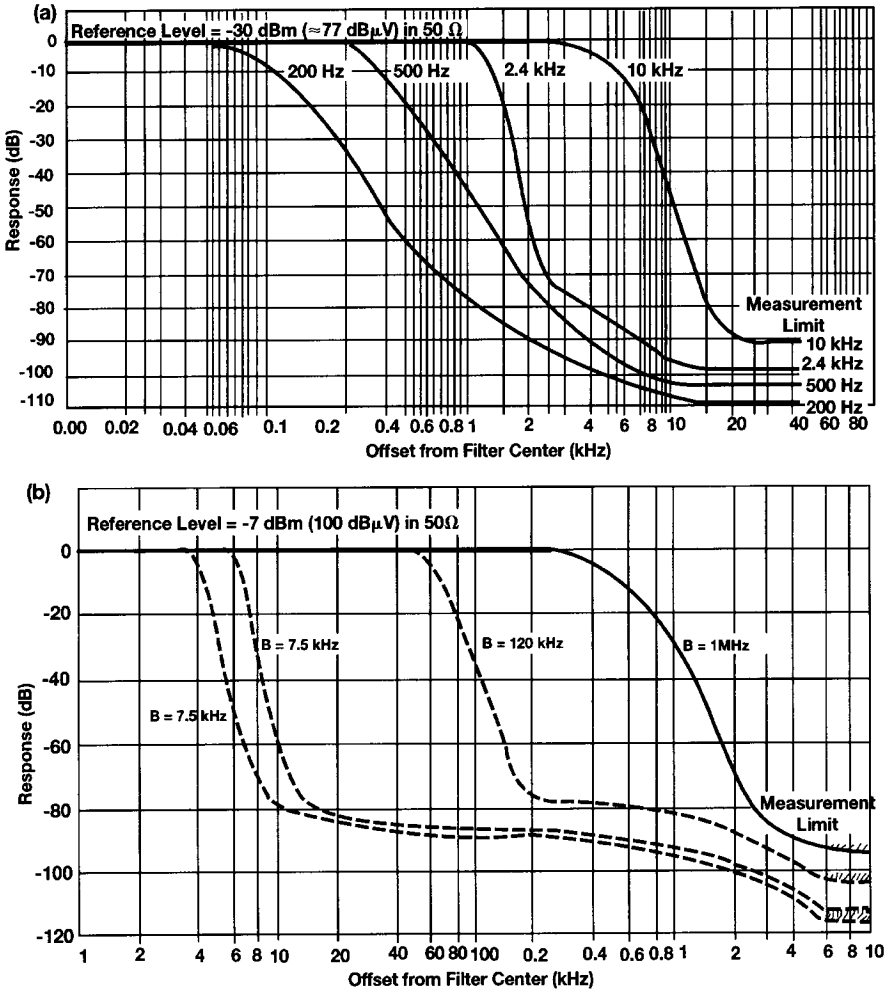


Figure 1-126 Dynamic selectivity versus IF bandwidth for (a) the Rohde & Schwarz ESH2 test receiver (9 kHz to 30 MHz) and (b) the Rohde & Schwarz ESV test receiver (10 MHz to 1 GHz). Reciprocal mixing widens the ESH2's 2.4-kHz response below –70 dB (–100 dBm) at (a) and the ESV's 7.5-, 12-, and 120-kHz responses below approximately –80 dB (–87 dBm) at (b).

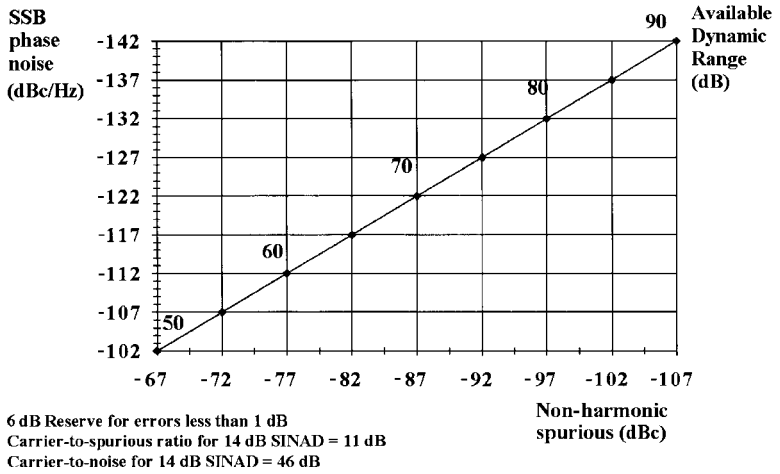


Figure 1-127 This graph shows the available dynamic range, which is determined either by the masking of the unwanted signal by phase noise or by discrete spurious signals. As far as the culprit synthesizer is concerned, it can be either the local oscillator or the effect of a strong adjacent-channel signal that takes over the function of the local oscillator.

Although additional RF selectivity can reduce the number of signals that contribute to the noise, improving the LO's spectral purity is the only effective way to reduce reciprocal mixing noise from *all* signals present at a mixer's RF port.

Factoring in the effect of discrete spurious signals with that of oscillator phase noise can give us the useful dynamic range of which an instrument or receiver is capable (Figure 1-127).

In evaluating the performance of digital wireless systems, we are interested in determining a receiver's resistance to adjacent-channel signals.

Phase Errors. In PM systems, especially those employing digital modulation, oscillator phase noise contributes directly to modulation and demodulation errors by introducing random phase variations that cannot be corrected by phase-locking techniques (Figure 1-128). The greater the number of phase states a modulation scheme entails, the greater its sensitivity to phase noise.

When an output signal is produced by mixing two signals, the resulting phase noise depends on whether the input signals are *correlated*—all referred to the same system clock—or uncorrelated, as shown in Figure 1-129.

Error Vector Magnitude. Several earlier figures (1-104, 1-112, and 1-128) have shown how particular sources of amplitude and/or phase error can shift the values of a digital emission's data states toward decision boundaries, resulting in increased BER due to intersymbol interference. Figures 1-130, 1-131, and 1-132 show three additional sources of such errors.

A figure of merit known as *error vector magnitude* (*EVM*) has been developed as a sensitive indicator of the presence and severity of such errors. An emission's error vector magnitude is the magnitude of the phasor difference as a function of time between an ideal reference signal and the measured transmitted signal after its timing, amplitude, frequency, phase, and dc offset have been modified by circuitry and/or propagation. Figure 1-133 illustrates the EVM concept.

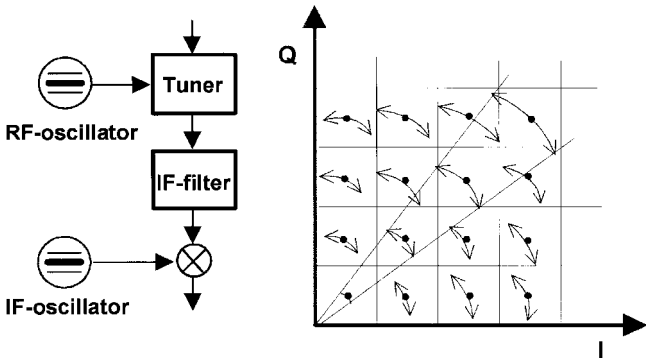


Figure 1-128 Phase noise is critical to digitally modulated communication systems because of the modulation errors it can introduce. Intersymbol interference (ISI), accompanied by a rise in BER, results when state values become so badly error-blurred that they fall into the regions of adjacent states. This drawing depicts only the results of phase errors introduced by phase noise; in actual systems, thermal noise, AM-to-PM conversion, differential group delay, propagation, and other factors may also contribute to the spreading of state amplitude and phase values.

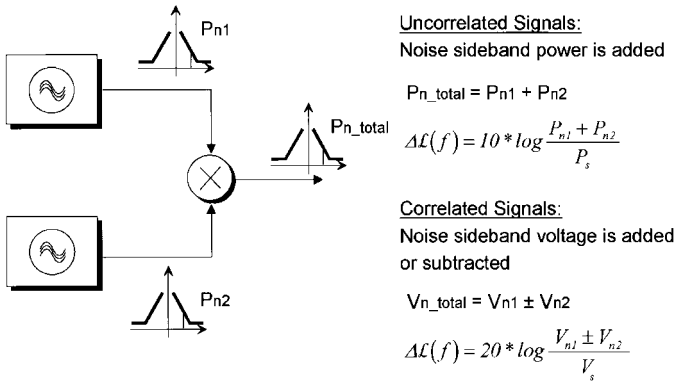


Figure 1-129 The noise sideband power of a signal that results from mixing two signals depends on whether the signals are correlated—referenced to the same system clock—or uncorrelated.

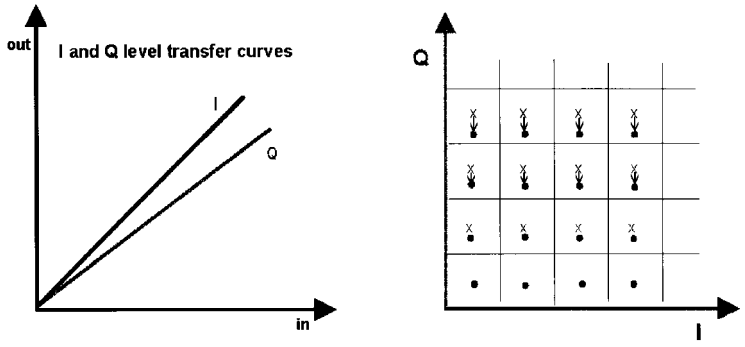


Figure 1-130 Effect of gain imbalance between I and Q channels on a data signal's phase constellation.

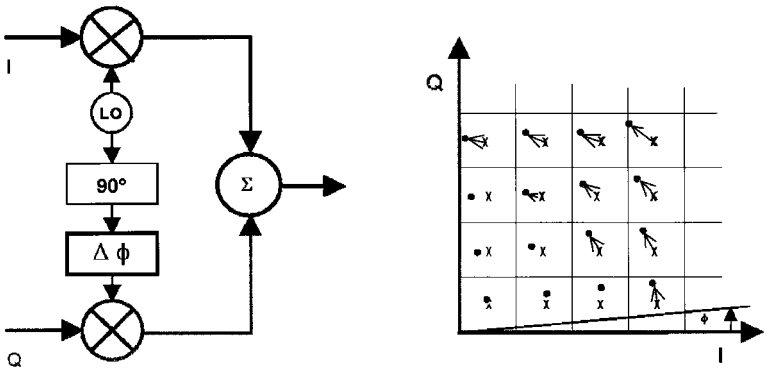


Figure 1-131 Effect of quadrature offset on a data signal's phase constellation.

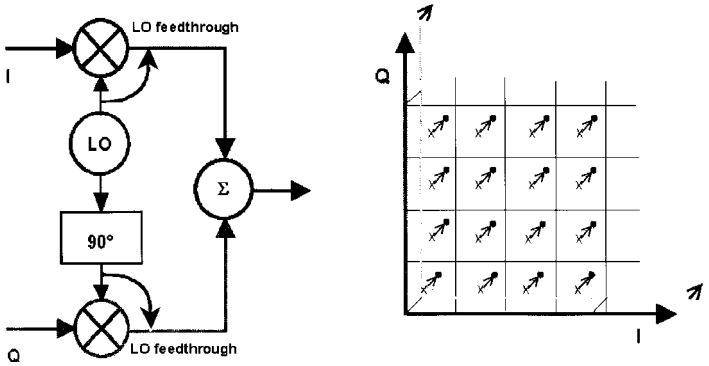


Figure 1-132 Effect of LO-feedthrough-based IQ offset on a data signal's phase constellation.

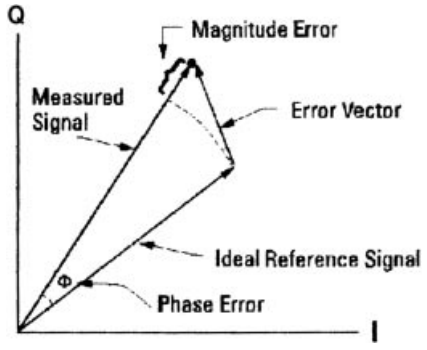


Figure 1-133 Representing errors in a digitally modulated signal's in-phase (I) and quadrature (Q) values relative to the ideal as a single error vector allows us to derive the resulting *error vector magnitude* (*EVM*), a sensitive indicator of transmission quality [20].

1-8 TESTING

1-8-1 Introduction

Testing of digital circuits deviates from the typical analog measurements, and yet the analog measurements are still necessary and related. As an example, the second-generation (2G) cellular telephone standards really have addressed only the transfer of voice and are just beginning to look into the transfer of data. Some services allow limited "e-mails," which are really messages that appear on the handset LCD after a delay. The future 3G standards with adaptive bandwidth will address high-speed data and video.

Since the major use of 2G phones is voice, information such as signal-to-noise ratio as a function of many things is important. In particular, because of the Doppler effect and the use of digital rather than analog signals, where the phase information is significant, the designer ends up using coding schemes for error correction—specifically, forward error correction (FEC). The signal-to-noise ratio as we know it from analog circuits now determines the bit error rate (BER), and its tolerable values depend on the type of modulation used. Because of correlation, it is possible to "rescue" a voice with a bit error rate of 1E – 4, but for higher data rates with little, if any, correlation, we are looking for BERs down to 1E – 9. The actual bit error rate depends on the type of filtering, coding, modulation, and demodulation. In several earlier figures (1-53 and 1-57) we related BER to signal-to-noise ratio; this is a key issue in receiver design.

Another problem in receivers that has to do with the transmitter of a second station is adjacent-channel power ratio (ACPR). Given the fact that a transmitter handling digital modulation delivers its power in pulses, its transmission may affect adjacent channels by producing transient spurious signals similar to what we call *splatter* in analog SSB systems. This is a function of the linearity of the transmitter system all the way out to the antenna and forces most designers to resort to less-efficient Class A operation. As possible alternatives, some researchers are looking into Class D and E modulation, more about which is found in the power amplifier section of Chapter 3.

It is actually not uncommon to do many linear measurements, and then by using correlation equations translate these measured results into their digital equivalents. Therefore, the robustness of the signal as a function of antenna signal at the receiver site, constant or known phase relationships, and high adjacent power ratios will provide good transfer characteristics.

1-8-2 Transmission and Reception Quality

The acoustic transmission and reproduction quality of a mobile phone is the most important characteristic in everyday use. Accurate and reproducible measurements of a cellular phone's frequency response cannot be achieved with static sinewave tones (SINAD measurements) because of coder and decoder algorithms. In this case, test signals simulating the characteristic of the human voice—that is, tones that are harmonic multiples of the fundamental—are required. A so-called vocoder is used to produce the lowest possible data rate. Instead of the actual voice, only the filter and fundamental parameters required for signal reconstruction are transmitted. Particularly in the medium and higher audiofrequency ranges, the static sinusoidal tones become a more or less stochastic output signal. For example, if a tone of approximately 2.5 kHz is applied to the telephone at a constant sound pressure, the amplitude of the signal obtained at the vocoder output varies by approximately 20 dB. In type-approval

tests, where highly accurate measurements are required, the coder and decoder are excluded from the measurement.

Whether the results obtained for the fundamental frequency are favorable depends on how far the values coincide with the clock of the coding algorithm. Through a skillful choice of fundamental frequencies, test signals with overlapping spectral distribution can be generated, giving a sufficient number of test points in subsequent measurements at different fundamental frequencies so that a practically continuous frequency response curve is obtained. Evaluation can be done by means of FFT analysis with a special window function and selection of result bins. The results are sorted and smoothed by the software and display in the form of a frequency response curve. Depending on the measurement, a program calculates the sending or receiving loudness rating in line with CCITT P.79 and shows the result.

Acoustic measurements relevant to GSM 11.10 include:

- Sending frequency response
- Sending loudness rating
- Receiving frequency response
- Receiving loudness rating
- Sidetone masking rating
- Listener sidetone rating
- Echo loss
- Stability margin
- Distortion sending
- Distortion receiving
- Idle channel noise receiving
- Idle channel noise sending

There are two categories of testing. One is a full-compliance test for all channels and all combinations of capabilities; the other is a production tester with evaluation of the typical characteristic data. Both Hewlett-Packard and Rohde & Schwarz are the leading companies in this area, and as the technology develops further, different equipment will be made available. Figure 1-134 shows a Rohde & Schwarz radiocommunication tester capable of evaluating cellular systems. Table 1-13 shows the typical specifications required to make the appropriate cellular telephone measurements.

The test setup must be capable of measuring key IS-95 parameters, using the following:

- Frame errors
- Waveform quality
- Error vector magnitude
- Phase error
- Magnitude error
- Carrier feedthrough
- Frequency accuracy
- Power measurements
- Base-station signaling for mobile testing

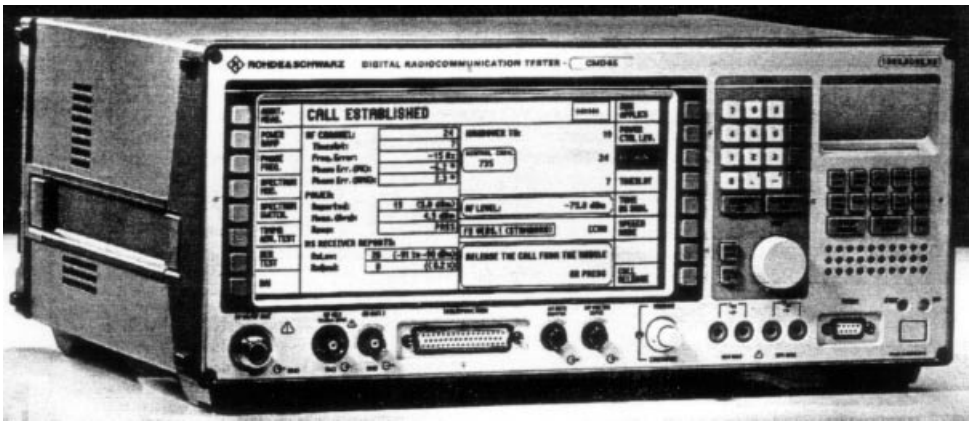


Figure 1-134 The Rohde & Schwarz CMD 65 digital radiocommunication tester can be used to measure all the relevant specifications of cellular telephones.

Table 1-13 Digital radiocommunication tester model CMD 80

SIGNAL GENERATOR	
Frequency	
Range	869–894 MHz
Resolution	1 Hz
Error	Same as timebase
Output level	
RF IN/OUT	–20 to –124 dBm
RF OUT2	0 to –105 dBm
Resolution	0.1 dB
Error (RF IN/OUT)	≤1.5 dB
Modulation	
Carrier suppression	–30 dBc
ANALYZER	
Frequency	
Range	824–849 MHz
Resolution	1 Hz
Error	Same as timebase
Power measurement	
Reference level range	
RF IN/OUT (full scale)	+41 to –28 dBm
RF IN2 (full scale)	0 to –69 dBm
Measurement error, absolute	±1.5 dB
Measurement error, relative	≤0.4 dB (ref. level to –40 dB)
Dynamic range	50 dB below reference level

Table 1-13 (Continued)

Modulation analyzer	
ρ error (25 ± 10) °C	0.003 (for ρ = 0.9–1)
Frequency measurement range	± 3 kHz
Frequency measurement error	Reference ±30 Hz
Timing measurement error	±60 ns
TIMEBASE	
Standard timebase	
Nominal frequency	10 MHz
Frequency drift in temperature range 0–35 °C	$\leq 1.5 \times 10^{-6}$
Frequency aging	$\leq 0.5 \times 10^{-6}$/yr (at 35 °C)
OCXO reference oscillator	
Nominal frequency	10 MHz
Frequency drift in temperature range 0–45 °C	$\leq 1 \times 10^{-7}$
Frequency aging	$\leq 2 \times 10^{-7}$/yr $\leq 0.5 \times 10^{-9}$/day after 30 days of operation
Warm-up time (at 25 °C)	approx. 5 min
Reference frequency inputs/outputs	
Synchronization input (frequency selectable)	100 kHz to 40 MHz, step 10 kHz
Impedance	approx. 100 Ω
Level	0 dBm to TTL
Synchronization output	
Frequency	10 MHz or frequency at synchronization input
Level	TTL, R_{out} = 50 Ω
AWGN generator	
Equivalent noise bandwidth	1.8 MHz typ.
Gain adjustment range	–20 to +6 dB of forward channel power
dc Voltage measurements	
Range	0 to ±30 V
Resolution	10 mV
Error	2% + resolution
dc Current measurements	
Mode	averaging, +peak, –peak
Range	0 to ±10 A
Common mode rejection	±30 V
Shunt resistance	50 mΩ
Resolution for averaging	1mA/10 mA
Resolution for peak	10 mA
Residual indication	<10 mA
Error	<2% + resolution + residual indication

1-8-3 Base-Station Simulation

Simulating a CDMA (IS-95) base station involves:

- Synchronization of mobile
- Registration of mobile
- Incoming/outgoing call origination
- Handoff

During the call, the tester must verify the handset's RF performance and check the correct operation of the basic signaling features. The best testers do not rely on any special test modes in the mobile station, performing their measurements under conditions almost identical to those in a real network. A voice loop-back allows quick verification of the performance of a mobile as it is perceived by the user.

1-8-4 GSM

Measurement, test, and adjustment capabilities for GSM should include:

- Synchronization of mobile phone with base station (which is simulated by CTS)
- Location update
- Call setup (incoming/outgoing)
- Call release (incoming/outgoing)
- Control and measurement of transmitter power
- Handover (channel change)
- Sensitivity, including bit error rate (BER) and raw bit error rate (RBER); limit sensitivity via search routine; RxLev and RxQual
- Phase and frequency error
- Power ramp versus time
- Timing error
- AFC (automatic frequency correction) and RSSI (radio signal strength indication)
- I/Q modulator adjustment
- Echo test (voice test, includes also testing of loudspeaker and microphone)
- Functional test of mobile's keypad through display of dialed number
- Display of IMSI (international mobile subscriber identity), IMEI (international mobile equipment identity), power class, and revision level
- Short message service (SMS)

1-8-5 DECT

Measurement, test, and adjustment capabilities for DECT should include:

- Synchronization of DUT with the CTS
- Call setup
- Call release
- Echo test

Often BER diagrams do not have C/N as abscissa (FIG) but E_b/N_0, which is the energy per information bit E_b referred to the normalized noise power N_0 [1]. C/N describes the ratios in the transmission channel, SNR the signals after the √cos receive filter. The following applies:

$$C/N = SNR + k_{roll\ off}\ [dB]$$

In converting the two quantities, some factors have to be taken into account as shown by equations 1 and 2 on the right.

To determine C/N [dB] or E_b/N_0 [dB], the logarithmic ratios have to be corrected using the following factors:

- Factor for Reed-Solomon FEC*

$$k_{FEC} = 10 \times \lg \frac{188}{204}$$

$$k_{FEC} = -0.3547\ [dB]$$

- Factors for QPSK/QAM modulation

$$k_{QPSK/QAM} = 10 \times \lg(m)$$

Modulation	m	$k_{QPSK/QAM}$ [dB]
QPSK	2	3.0103
16QAM	4	6.0206
64QAM	6	7.7815
256QAM	8	9.0309

- Factor for coding rate (P = 1 for QAM)*

$$k_P = 10 \times \lg(P)$$

Modulation	P	k_P [dB]
QPSK	1/2	–3.0103
	2/3	–1.7609
	3/4	–1.2494
	5/6	–0.7918
	7/8	–0.5799
QAM	1	0

- Factor for √cos roll-off filtering in demodulator/receiver

$$k_{roll\ off} = 10 \times \lg\left(1 - \frac{\alpha}{4}\right)$$

Modulation	α	$k_{roll\ off}$ [dB]
DVB-C	0.15	–0.1660
DVB-S	0.35	–0.3977

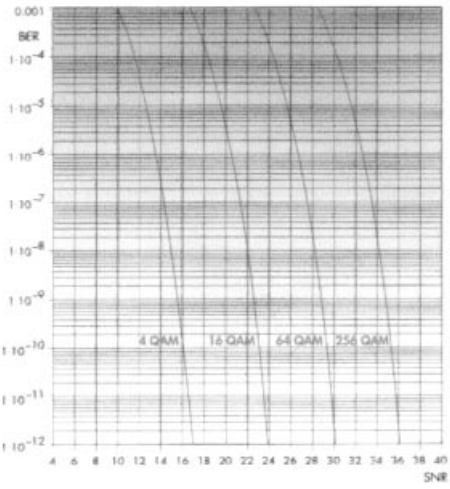
* To be taken into account if E_b only refers to the information bits; not to be taken into account if E_b refers to all bits transmitted (information plus error control bits)

$$C/N = E_b/N_0 + 10 \times \lg\frac{188}{204} + 10 \times \lg(m) + 10 \times \lg(P) - 10 \times \lg\left(1 - \frac{\alpha}{4}\right)\ [dB]$$

Equation 1

$$E_b/N_0 = C/N - 10 \times \lg\frac{188}{204} - 10 \times \lg(m) - 10 \times \lg(P) + 10 \times \lg\left(1 - \frac{\alpha}{4}\right)\ [dB]$$

Equation 2



BER referred to SNR for 4, 16, 64 and 256QAM

The types of correction factor required depend on whether measurement is made

- in the transmission channel,
- before or after Viterbi correction,
- with QAM or QPSK modulation.

Examples of conversion equations –

For **in-channel** measurements with QAM transmission, the following applies:

$$E_b/N_0 = C/N - 10 \times \lg\frac{188}{204} - 10 \times \lg(m)\ [dB]$$

The factors for √cos roll-off filtering and puncturing rate are not needed.

For measurements in the **QAM demodulator**, √cos roll-off filtering filtering has to be taken into account.

$$E_b/N_0 = C/N - 10 \times \lg\frac{188}{204} - 10 \times \lg(m) + 10 \times \lg\left(1 - \frac{\alpha}{4}\right)\ [dB]$$

For measurements in the **satellite demodulator** with QPSK, the equation for determining the BER as a function of E_b/N_0 after Viterbi FEC is as follows:

$$E_b/N_0 = C/N - 10 \times \lg\frac{188}{204} - 10 \times \lg(m) - 10 \times \lg(P) + 10 \times \lg\left(1 - \frac{\alpha}{4}\right)\ [dB]$$

In the latter case all correction factors are included.

Reference:
[1] News From Rohde & Schwarz, Number 163 (1999/II)

Figure 1-135 Conversion of C/N or BER to E_b/N_0.

- Detection and display of RFPI (FP)
- Normal transmit power (NTP)
- Power ramp versus time
- Modulation characteristics versus time
- Frequency offset
- Maximum modulation deviation

- Frequency drift
- Timing (jitter, packet delay)
- Bit error rate (BER), frame error rate (FER)

1-9 CONVERTING C/N OR SNR TO E_b/N_0

Figure 1-135 shows an application note for converting carrier-to-noise ratio (C/N) or SNR to energy per bit/normalized noise power (E_b/N_0).

REFERENCES

1. Zoran Kostic, "Digital Signal Processors in Cellular Radio Communications," *IEEE Commun. Mag.*, **35**(12), 22–35, December 1997.
2. Behzad Razavi, "Recent Advances in RF Integrated Circuits," *IEEE Commun. Mag.*, **35**(12), 36–43, December 1997.
3. William Reuter, "Source and Synthesizer Phase Noise Requirements for QAM Radio Applications," Part 1, *Microwave Prod. Digest*, 36, 38, 50–51, 54–56, January 1999.
4. Peter Vizmuller, *RF Design Guide: Systems, Circuits, and Equations* . Boston: Artech House, 1995.
5. Bernard Sklar, *Digital Communications: Fundamentals and Applications* . New York: Prentice-Hall, 1995.
6. Lin Nan Lee and Jim Mullen, "Third-Generation Cellular Advances Toward Voice and Data," *Wireless Syst. Des.*, **2**(12), 32–38, December 1997.
7. Philips RF Application Note AN1981, "New Low-Power Single-Sideband Circuits."
8. Philips RF Application Note AN1982, "Apply the Oscillator of the SA602 in Low-Power Mixer Applications."
9. Ulrich L. Rohde, Jerry C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*. New York: McGraw-Hill, 1997, p. 58.
10. William E. Sabin and Edgar O. Schoenike, eds., *Single Sideband Systems and Circuits* , 2nd ed. New York: McGraw-Hill, 1995, pp. 123–124.
11. H. T. Friis, "Noise Figures of Radio Receivers," *Proc. IRE*, **32**(7), 419–422, July 1944.
12. Mark Williams, Fred Bonn, Curtis Gong, and Tina Quach, "GaAs RF ICs Target 2.4-GHz Frequency Band," *Microwaves & RF*, **33**(7), 111–118, July 1994.
13. Jack Smith, *Modern Communication Circuits* , 2nd ed. New York: McGraw-Hill, 1997, pp. 93–99.
14. S. M. Perlow, "Third-Order Distortion in Amplifiers and Mixers," *RCA Rev.*, 234–265, June 1976.
15. Edward C. Jordan, ed., *Reference Data for Engineers* , 7th ed. Indianapolis: Howard W. Sams & Co, 1985, pp. 12-33–12-35.
16. Stephen A. Maas, *Nonlinear Microwave Circuits* . Norwood, MA: Artech House, 1988, pp. 18 and 169.
17. Joseph Staudinger, "Specifying Power Amplifier Linearity via Intermodulation Distortion and Channel Spectral Regrowth," *Appl. Microwave & Wireless* , **9**(4), 62, 64, 66, 68, 70, July/August 1997.

18. Michael S. Heutmaker, Eleanor Wu, and John R. Welch, "Envelope Distortion Models with Memory Improve Some Amplifier Spectral Regrowth Predictions," *Appl. Microwave & Wireless*, **9**(4), 72, 74, 76, 78, July/August 1997.
19. Toko SF246ZA-019 filter specifications, available via the Web address http://www.tokoam.com/saw/applications.html.
20. *Using Vector Modulation Analysis in the Integration, Troubleshooting and Design of Digit al RF Communications Systems*, Product Note HP89400-8 (Hewlett-Packard Co., 1994, available via the Web address http://www.tmo.hp.com/tmo/Notes/English/5091-8687E.html).
21. George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990, pp. 275–278.
22. William G. Scanlon, "Health Aspects of Low-Level Exposure to RF Electromagnetic Waves," *RF Des.*, **22**(7), 40, 44, 46, 48, 50, 52, 54, 56, 64, 66, 68, July 1999.
23. Van Valkenburg, editor-in-chief, *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, 8th ed. Carmel, CA: Prentice-Hall Computer Publishing, 1993, pp. 25-19 –25-20.

Further Reading

Walid Y. Ali-Ahmad, "Improving the Receiver Intercept Point Using Selectivity," *RF Des.*, **20**(12), 22, 24, 26, 28, 30, December 1997.

"AM Signal Analysis in a Dual Down-Conversion Receiver," *Symphony Examples Volume*, Ansoft Corporation, 1999.

Leon W. Couch II, *Digital and Analog Communications Systems*. New York: Macmillan, 1990.

Paul Danzer, ed., *The ARRL Handbook for Radio Amateurs*, 75th ed. Newington, CT: American Radio Relay League, 1997, p. 26.44.

Nikhil Deshpande, Steven Stanton, and Michael Hurst, "ACPR Specs Place Demands on WCDMA Base-Station Amplifiers," *Wireless Systems Design*, **4**(8), 15–22, August 1999.

Wing S. Djen and Prasanna M. Shah, "Implementation of a 900 MHz Transmitter System Using Highly Integrated ASIC," *Proceedings of the 44th IEEE Vehicular Technology Conference*, June 1994, pp. 1341–1345.

Dimitrios Efstathiou, "π/4-DPQSK Data-Modulation Schemes Help Enable 3G Systems," *Wireless Syst. Des.*, **4**(1), 27–32, January 1999.

"Double Down Conversion Communications Receiver," *Symphony Examples Volume*, Ansoft Corporation, 1999.

Kamilo Feher, "FQPSK Transceivers Double the Spectral Efficiency of Wireless and Telemetry Systems," *Appl. Microwave & Wireless*, **10**(5), 73–79, June 1998.

David K. Lovelace, Jesus L. Finol, and Jeffrey C. Durec, "Sub-Micron Silicon RF IC Technologies: An Overview," 1998 IEEE MTT-S International Microwave Symposium, 1859–1862, June 1998.

William Gosling, ed., *Radio Receivers*. London: Peter Peregrinus Ltd., 1986, p. 29.

Helge O. Granberg, "Power MOSFETs Versus Bipolar Transistors," Motorola Semiconductor Application Note AN860.

GSM Standard, "Physical Layer on the Radio-Path," July 1988.

W. H. Hayward, *Introduction to Radio Frequency Design*. Newington, CT: American Radio Relay League, 1994 (original edition: 1982).

Gene Heftman, "Cellular ICs Move Toward 3G Wireless—Gingerly," *Microwaves & RF*, **38**(2), 31–38, February 1999.

Roy Hejhall, "VHF MOS Power Applications," Motorola Semiconductor Application Note AN878.

IS-54-B, EIA/TIA, "Cellular System Dual-Mode Mobile Station-Base Station Compatibility Standard," April 1992.

M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems*. New York: Plenum Press, 1992.

Lawrence E. Larson, ed., *RF and Microwave Circuit Design for Wireless Communications*. Boston: Artech House, 1996.

Chia-Liang Liu and Kamilo Feher, "π/4-QPSK MODEMS for Satellite Sound/Data Broadcast Systems," *IEEE Trans. Broadcasting*, 1–8, March 1991.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

David Lovelace, Jeff Durec, and Mike McGinn, "A Silicon Bipolar Transmitter IC for 900MHz ISM Band Applications," *Proceedings of the Sixth Annual WIRELESS Symposium*, Santa Clara, CA, February 9–13, 1998, pp. 480–484.

Brian J. Minnis, *Designing Microwave Circuits by Exact Synthesis*. Boston: Artech House, 1996.

Simon Morris, "DSP and ASIC Partitioning Offer Options for cdma2000 Base-Station Designs," *Wireless Syst. Des.*, **3**(11), 17–25, November 1998.

Mohamed K. Nezami, "Evaluate the Impact of Phase Noise on Receiver Performance," *Microwaves & RF*, **37**(5), 165–172, May 1998.

Philips Semiconductors, "PCD5070 GSM Baseband Interface," preliminary specification, September 1992.

Ulrich L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design*. New York: Wiley, 1997.

William E. Sabin, "Use of Mixers in HF Upconversion Receivers/Exciters," *Professional Program Session Record 24*, Wescon/81, San Francisco, CA, September 15–17, 1981, pp. 24/4-3 to 24/4-4.

Jan Sevenhans and Bart Verstraeten, "10 Years of Silicon Integration," *Microwave Eng. Eur.*, 35–40, May 1999.

Mark Simpson, "Low-Noise Synthesizers Applied to QAM-Modulated Digital Radios," *Proceedings of the Sixth Annual Wireless Symposium*, Santa Clara, CA, February 9–13, 1998, pp. 182–190.

Raymond Steele, *Mobile Radio Communications*. New York: IEEE Press, 1992.

Tim Stroud, "Tutorial on All-Digital Modulation Techniques," *RF Des.*, **20**(6), 52–58, June 1997.

Jan Uddenfeldt, "Digital Cellular—Its Roots and Its Future," *Proc. IEEE*, **86**(7), 1319–1324, July 1998.

Robert T. Van Brunt, "The Next Generation of Wireless: Coming Soon, If Not By 2000," *Microwave Prod. Digest*, 10, 30, 34, April 1999.

Peter Vizmuller, *RF Design Guide: Systems, Circuits and Equations*. Norwood, MA: Artech House, 1995. A good general RF engineering handbook in compact format.

Dave Vondrian, "Noise Figure Measurement: Corrections Related to Match and Gain," *Microwave J.*, **42**(3), 22–38, March 1999.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

1

INTRODUCTION TO WIRELESS CIRCUIT DESIGN

1-1 OVERVIEW

Wireless circuits are not that different from commonly known two-way radio, television, and broadcast arrangements. Some of them require high linearity in modulation (TV picture); some work via relay stations (two-way radio). The real differences lie in the fact that the cell sizes are much smaller, and that in most cases we attempt multiple channel use (reuse) using time-division multiplex, spread spectrum, or some other efficient means of reducing the bandwidth required for communication. One can argue that the wireless circuits include simple devices, such as garage-door openers and wireless keys for automobiles (we have seen many cases where strong interfering signals prevented the car owners from reclaiming their cars until the interfering signal disappeared). Another longtime favorite is cordless telephones: initially, 50-MHz models with essentially no privacy protection; later, more sophisticated models that operate at 900 MHz; and now, dual-band designs that use 900 MHz and 2.4 GHz.

The largest wireless growth area is probably the cellular telephones. The two major applications are the handsets, commonly referred to as cell phones or occasionally as "handies," and the base stations. The base stations have many more problems with large-signal-handling linearity at high power, although handset users may run into similar problems. An example of this is the waiting area of an airport, where many travelers are trying to conduct last-minute business: In one instance, we concluded that about 30% of all the people present were on the air! It would have been fun to evaluate this receiver-hostile environment with a spectrum analyzer.

From such use comes anxiety factors, the lesser of which is "When will my battery die?"—a spare battery tends to help—and the greater of which the ongoing question, "Will this cell-phone transmitter harm my body?" [22]. A brief comment for the self-proclaimed experts in this area: A 50–100-kW TV transmitter, specifically its video or picture portion, connected to a high-gain antenna, emits levels of energy in line-of-sight paths that by far exceed the pulsed energy from a cell phone. Specifically, the duration of energy is signifi-

cantly smaller, and the absolute energy is more than a thousandfold higher, than the radio frequency (RF) supposedly harming us from the cellular phone. Handheld two-way radios have been used for the last 30 years or so by police and other security interests, operating in the frequency range from 50 to 900 MHz with antennas close to the users' heads, and there are no known cases of cancer or any other illnesses caused by these handheld radios. Recent studies in England, debatably or not, showed that the reaction-time level of people using cell phones actually *increased*—but then there are always the skeptics and politically motivated who ignore the facts, try to influence the media, and have their 15 minutes of fame (as Andy Warhol used to say).

As to the "harmful" radiation, Figure 1-1 shows the simulated radiation of a Motorola flip phone. While there are no absolute values attached to the pattern colors, it is interesting to see that the antenna extension inside the plastic casing also radiates, but most of the energy definitely is emitted by the top of the antenna. It seems to be a good idea to hold the telephone in such a way that the antenna points away from the head, "just in case." The user will find a "warm" sensation that will have more to do with the efficiency of the RF power amplifier heating up the case than the effect of radiation.

With this introduction in place, we will first take a look at a typical ultra-high-frequency/super-high-frequency (UHF/SHF) transceiver and explain the path from the microphone to the antenna and back. After this, we will inspect the radio channel and its effect on various methods of digital modulation. Analysis of wireless receivers and transmitters will be next, followed by a look at available building blocks and how they affect the overall system. To validate proper system operation, a fairly large number of measurements and tests must be performed, and conveying their purpose and importance will necessitate the definition of a number of system characteristics and concepts, such as dynamic range. Finally, after this is done, we will look at the issue of wireless system testing. Again, we intend to

Figure 1-1 Simulated antenna radiation of a Motorola flip phone.

give guidance applicable to battery-operated, handheld operation as well as high-powered base stations.

1-2 SYSTEM FUNCTIONS

A cellular telephone is a hybrid between a double-sideband and frequency-modulated [FM; or phase-modulated (PM)] transceiver. The actual transmission is not continuous but is pulsed, and because of the pulse spectrum there is a signal bandwidth concern due to keying transients, not unlike intermodulation products of a single-sideband (SSB) transceiver cluttering up adjacent channels. The cellular telephone is also a linear transceiver in the sense that its signal-handling circuitry must be sufficiently amplitude- and phase-linear to preserve the modulation characteristics of the AM/PM hybrid emissions it transmits and receives. Containing such an emission's spectral regrowth, which affects operation on adjacent channels, is not unlike the linearity requirements we encounter in SSB transceivers—requirements so stringent that amplifiers must be run nearly in Class A to meet them. The time-division multiple access (TDMA) operating mode, which allows many stations to use the same frequency through the use of short, precisely timed transmissions, requires a system that transmits with a small duty cycle, putting much less thermal stress on a power amplifier than continuous operation. Power management, including a sleep mode, is another important issue in handset design.

Figure 1-2 shows the block diagram of a handheld transceiver. This is applicable for cellular telephones and other systems that allow full duplex. For those not too familiar with transceivers, here is a "walk" through the block diagram. The RF signal intercepted by the antenna is fed through a duplex filter into a front end consisting of a preamplifier, an additional filter, and a mixer. The duplexer is optimized more for separating transmit and receive frequencies than extreme selectivity, but because of the typical low field strengths of incoming signals, it provides enough selectivity to guard the receiver path against overload and intermodulation products. The preamplifier is either a single transistor or a cascode arrangement with a filter following it. These high-band filters, mostly supplied by Murata, are typically surface acoustic wave (SAW) filters with very small dimensions. We would already like to point out in this part of the block diagram that these filters typically have high-impedance inputs and outputs (somewhere between 200 Ω and 1 kΩ), therefore eliminating the nice test-setup possibilities typically provided in a 50-Ω system. Generally, integrated circuit (IC)-type mixers also operate at high impedances, which makes matching easier. The filter following the mixer is responsible for reducing the image, and then we go to the intermediate frequency (IF) and demodulation. The particular chip or chips mentioned here, supplied by Philips, are set out for a double-conversion receiver, and the demodulation is accomplished with a quadrature detector for FM analog modulation. The rest of the circuitry on the horizontal path does digital signal processing (DSP) and overall control functions. The four blocks at the far right refer to the central processor, which handles such things as display, power management, and information storage (such as frequently used telephone numbers). A nice overview about DSP in "readable" form is given by Kostic [1].

The transmit portion consists of an independent synthesizer that is modulated. There are dual synthesizer chips available to accommodate this. Both receive and transmit frequencies are controlled by a miniature temperature-compensated crystal oscillator (TCXO). One of its outputs is also used as the system master clock for all the digital activities. The output of the voltage-controlled oscillator (VCO) is then amplified and fed to the antenna through the

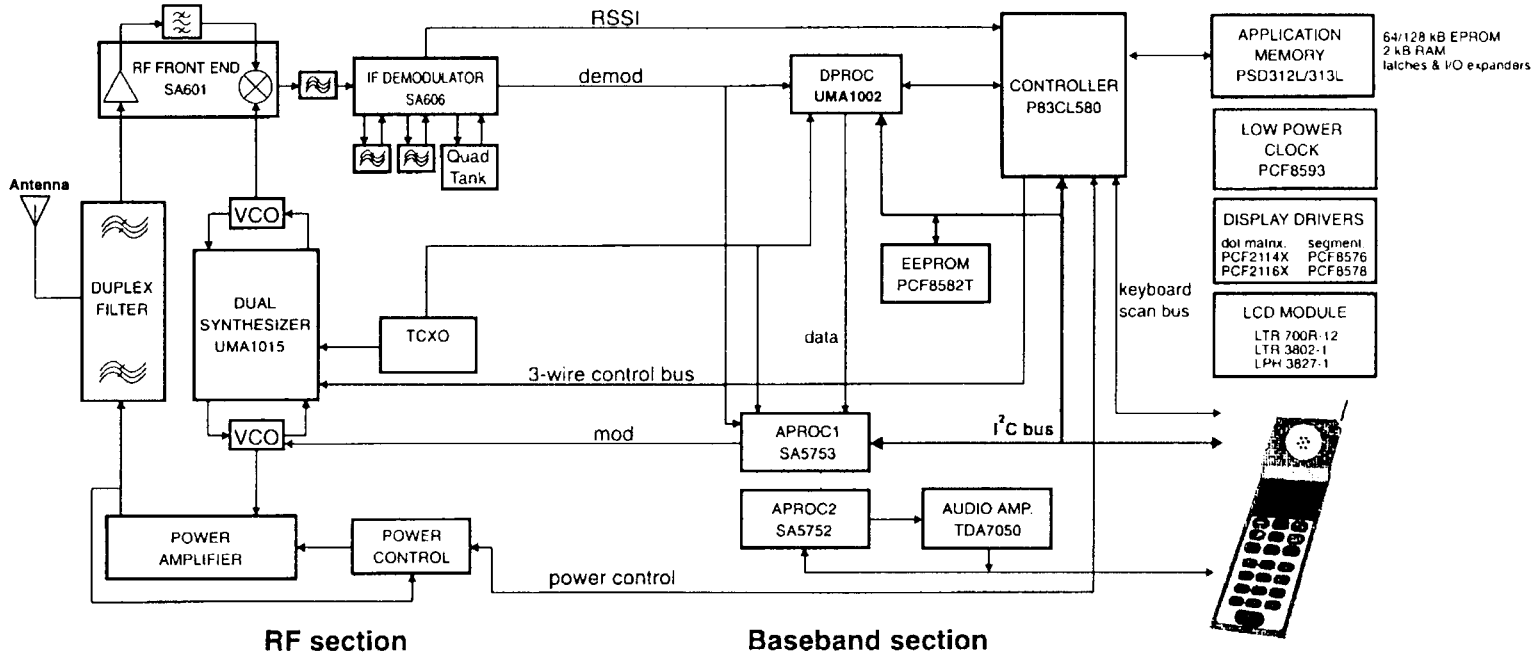


Figure 1-2 Block diagram of a handheld cellular telephone transceiver.

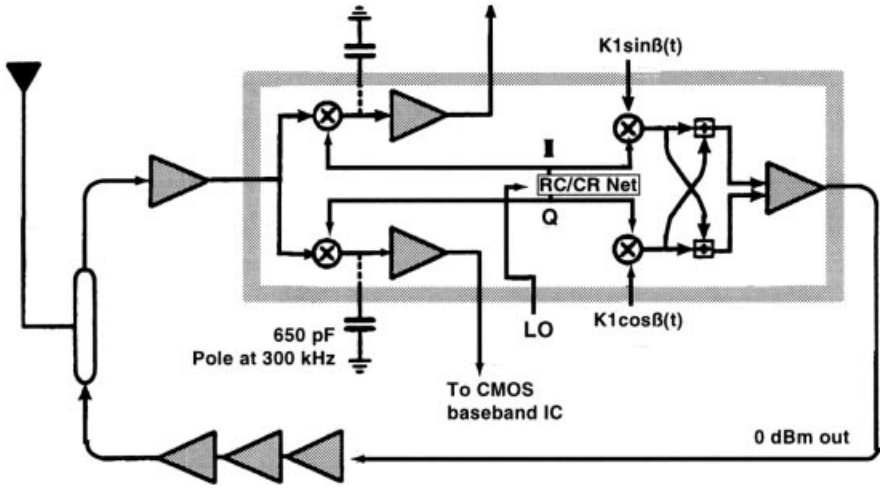


Figure 1-3 Single-chip direct-conversion transceiver by Alcatel. Channel selection is accomplished at baseband by low-pass switched-capacitor filters in a companion mixed-signal complementary metal-oxide semiconductor (CMOS) IC. A trimmed resistance-capacitance/capacitance-resistance (RC/CR) network generates the necessary quadrature signals for the chip's mixers.

same duplex filter as the receive portion. There are also schemes available for advanced modulation methods, specifically, code- , frequency-, and time-division multiple access (CDMA, FDMA, and TDMA, respectively). In these cases, the transmitter is not active all the time, and the duplexer can be replaced with a diode switch using a quarter-wavelength transmission line together with a PIN diode for the required switching.

Many modern devices use "zero IF" or direct conversion, which simplifies the IF or modulation portion of the unit significantly. Figure 1-3 shows an Alcatel single-chip direct-conversion transceiver. The signal is fed to an image-reject mixer with the local oscillator (LO) in quadrature, and the selectivity is obtained by manipulating the "audio bandwidth." Today we have a large number of implementations using different schemes that are beyond the scope of this book; therefore, we have decided to limit ourselves to a basic introduction because most of the relevant demodulation and coding are done in DSP, for which we will give appropriate references. A nice overview of different architectures is found in Razavi [2].

1-3 THE RADIO CHANNEL AND MODULATION REQUIREMENTS

1-3-1 Introduction

The transmission of information from a fixed station to a mobile is considerably influenced by the characteristics of the radio channel. The RF signal not only arrives at the receiving antenna on the direct path but is normally reflected by natural and artificial obstacles in its way. Consequently the signal arrives at the receiver several times in the form of echoes, which are superimposed on the direct signal (Figure 1-4). This superposition may be an advantage as the energy received in this case is greater than in single-path reception. This feature is made use of in the digital audio broadcasting (DAB) single-frequency network. However,

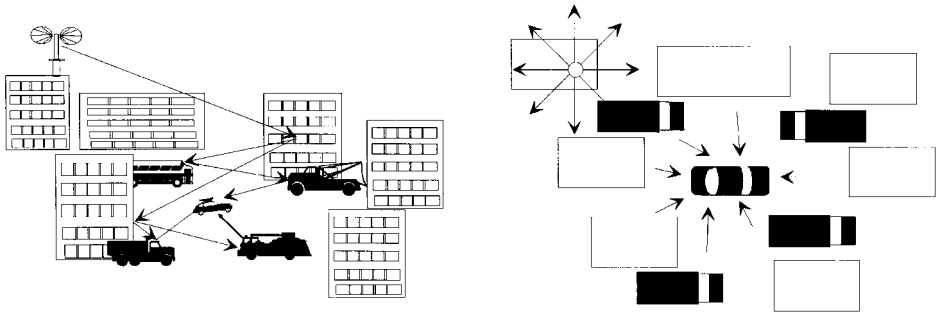


Figure 1-4 Mobile receiver affected by fading.

this characteristic may be a disadvantage when the different waves cancel each other under unfavorable phase conditions. In conventional car radio reception this effect is known as fading. It is particularly annoying when the vehicle stops in an area where the field strength is reduced because of fading (e.g., at traffic lights). Additional difficulties arise when digital signals are transmitted. If strong echo signals (compared to the directly received signal) arrive at the receiver with a delay on the order of a symbol period or more, time-adjacent symbols interfere with each other. In addition, the receive frequency may be falsified at high vehicle speeds because of the Doppler effect so that the receiver may have problems in estimating the instantaneous phase in the case of angle-modulated carriers. Both effects lead to a high symbol error rate even if the field strength is sufficiently high. Radio broadcasting systems using conventional frequency modulation are hardly affected by these interfering effects. If an analog system is replaced by a digital one that is expected to offer advantages over the previous system, it has to be ensured that these advantages—for example, better audiofrequency signal/noise (AF S/N) and the possibility of supplementary services for the subscriber—are not at the expense of reception in hilly terrain or at high vehicle speeds because of extreme fading.

For this reason a modulation method combined with suitable error protection has to be found for mobile reception in a typical radio channel, which is immune to fading, echo, and Doppler effects.

With a view to this, more detailed information on the radio channel is required. The channel can be described by means of a model. In the worst case, which may be the case for reception in built-up areas, it can be assumed that the mobile receives the signal on several indirect paths but not on a direct one. The signals are reflected, for example, by large buildings; the resulting signal delays are relatively long. In the vicinity of the receiver these paths are split up into a great number of subpaths; the delays of these signals are relatively short. These signals may again be reflected by buildings but also by other vehicles or natural obstacles like trees. Assuming the subpaths are statistically independent of each other, the superimposed signals at the antenna input cause considerable time- and position-dependent field-strength variations with an amplitude obeying the Rayleigh distribution (Figures 1-5 and 1-6).

If a direct path is received in addition, the distribution changes to the Rice distribution, and finally, when the direct path becomes dominant, the distribution follows the Gaussian distribution with the field strength of the direct path being used as the center value.

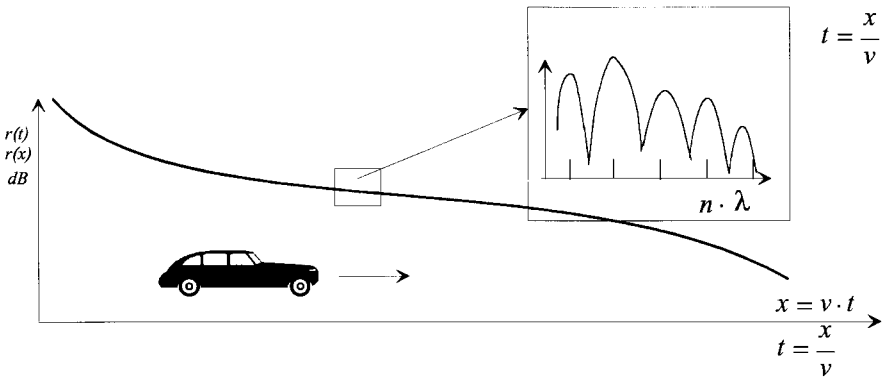


Figure 1-5 Receive signal as a function of time or position.

In a Rayleigh channel the bit error rate (BER) increases dramatically compared to the BER in an additive white Gaussian noise (AWGN) channel (Figure 1-7).

1-3-2 Channel Impulse Response

This scenario can be demonstrated by means of the channel impulse response. Let's assume that a very short pulse of extremely high amplitude [in the ideal case, a Dirac pulse $\delta(t)$] is sent by the transmitting antenna at a time $t_0 = 0$. This pulse arrives at the receiving antenna direct and in the form of reflections with different delays τ_i and different amplitudes because of path losses. The impulse response of the radio channel is the sum of all received pulses (Figure 1-8). Since the mobile receiver and also some of the reflecting objects are moving, the channel impulse response is a function of time and of delays τ_i; that is, it corresponds to

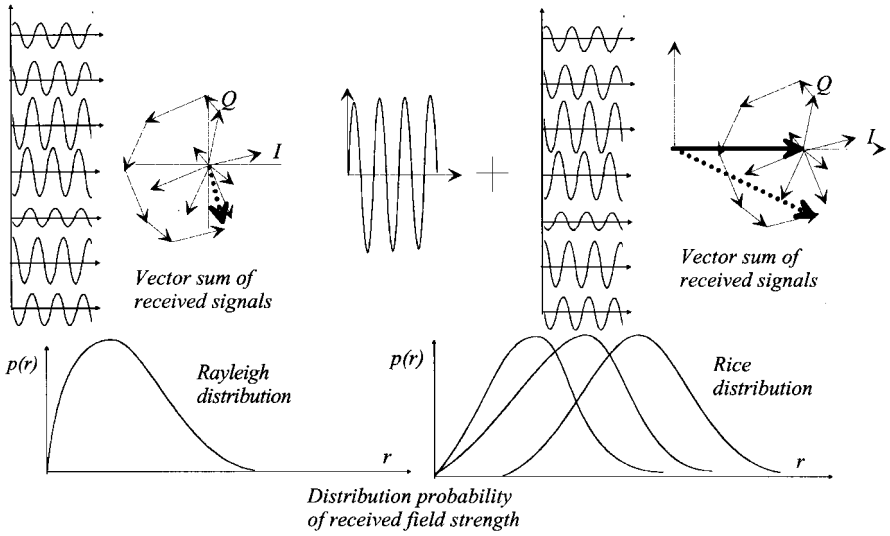


Figure 1-6 Rayleigh and Rice distributions.

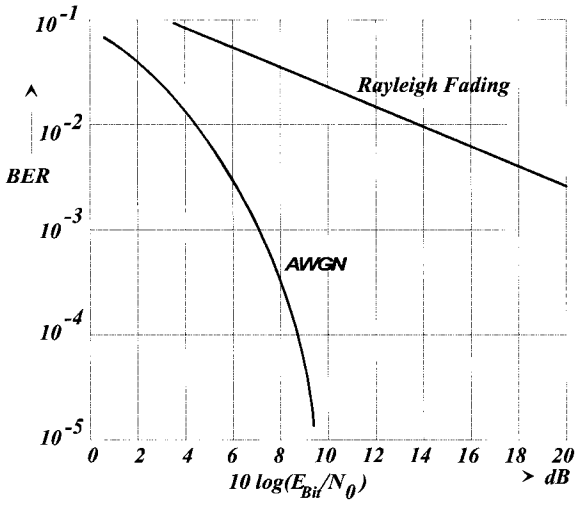


Figure 1-7 BER in a Rayleigh channel.

$$h(t,\tau) = \sum_{N} a_i \delta(t - \tau_i) \tag{1-1}$$

This shows that delta functions sent at different times t cause different reactions in the radio channel.

In many experimental investigations different landscape models with typical echo profiles were created. The most important are:

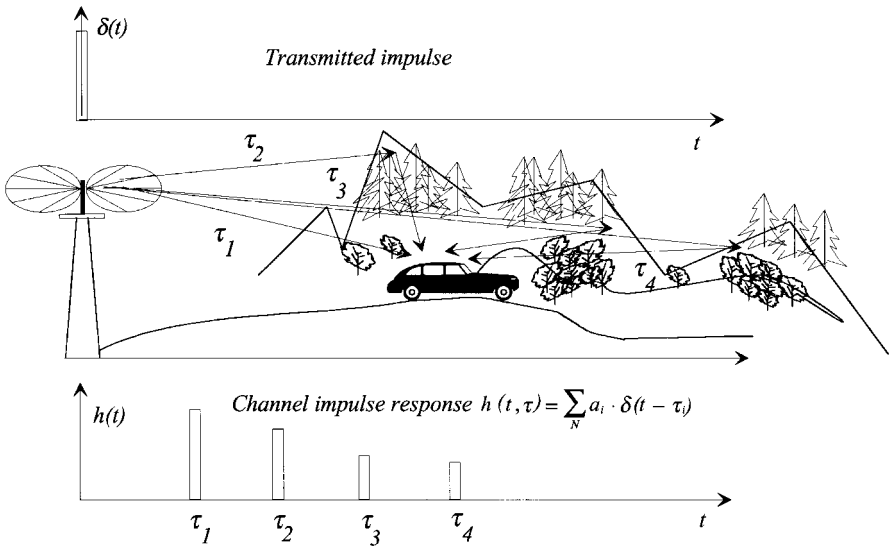


Figure 1-8 Channel impulse response.

- Rural area (RA)
- Typical urban area (TU)
- Bad urban area (BA)
- Hilly terrain (HT)

The channel impulse response informs on how the received power is distributed to the individual echoes. A parameter, the "delay spread," can be calculated from the channel impulse response, permitting an approximate description of typical landscape models (Figure 1-9).

The delay spread also roughly informs on the modulation parameters, carrier frequency, symbol period, and duration of guard interval, which have to be selected in relation to each other. If the receiver is located in an area with a high delay spread (e.g., in hilly terrain), echoes of the symbols sent at different times are superimposed when broadband modulation methods with a short symbol period are used. In the case of DAB, this problem is aggravated by the use of single-frequency networks. An adjacent transmitter emitting the same information on the same frequency has the effect of an artificial echo (Figure 1-10).

A constructive superposition of echoes is only possible if the symbol period is much greater than the delay spread. The following holds:

$$T_s > 10T_d \tag{1-2}$$

This has the consequence that relatively narrowband modulation methods have to be used. If this is not possible, channel equalizing is required.

For channel equalizing, a continuous estimation of the radio channel is necessary. The estimation is performed with the aid of a periodic transmission of data known to the receiver. In networks according to the Groupe Speciale Mobile (GSM) standards a midamble consisting of 26 bits—the training sequence—is transmitted with every burst. The training

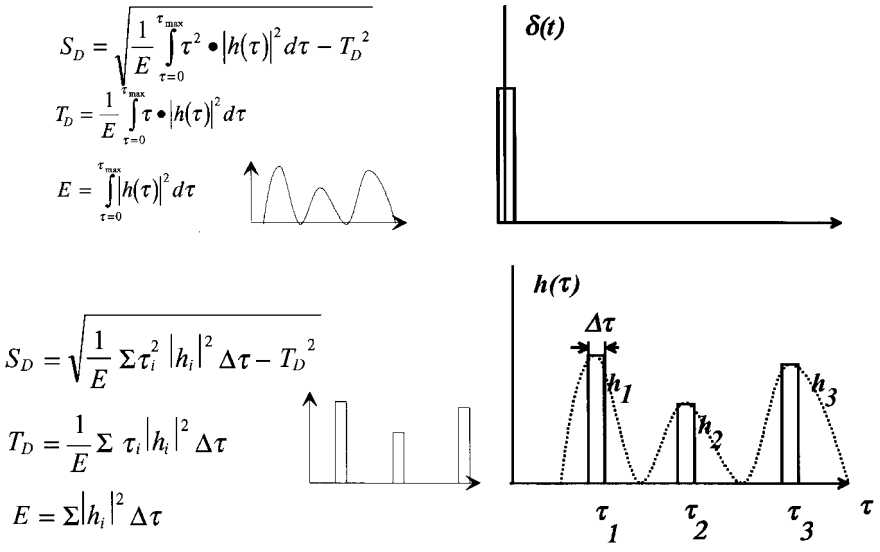


Figure 1-9 Calculation of delay spread.

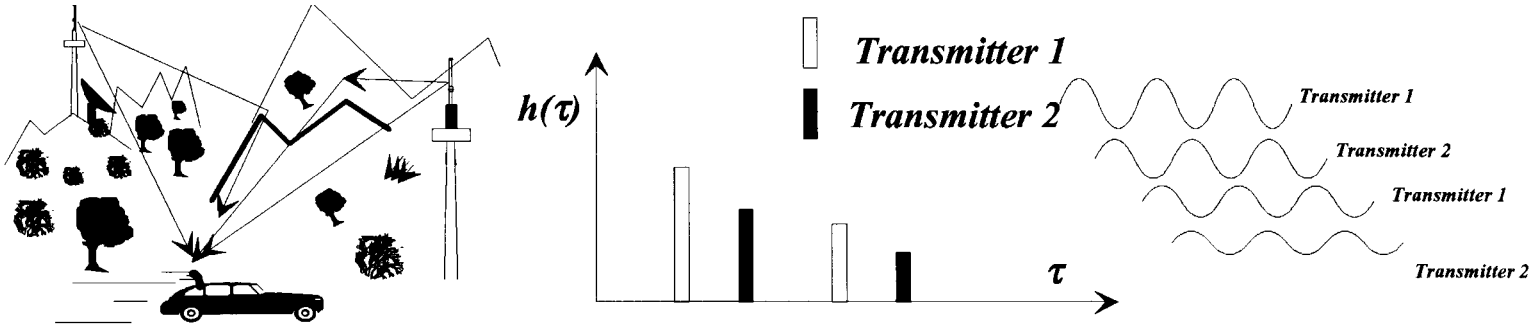


Figure 1-10 Artificial and natural echoes in the single-frequency network.

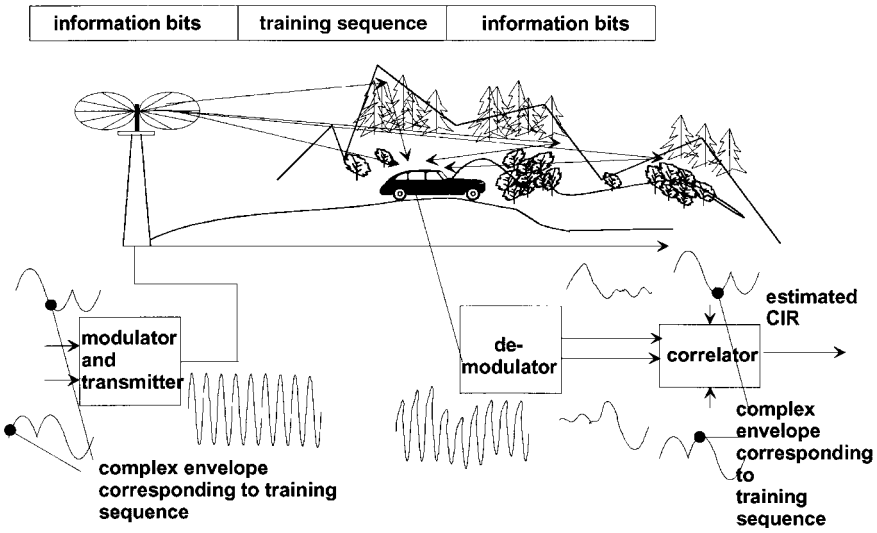


Figure 1-11 Channel estimation.

sequence corresponds to a characteristic pattern of I/Q signals that is kept in a memory in the receiver. The baseband signals of every received training sequence are correlated with the stored ones. From this correlation the channel can be estimated, and the properties of the estimated channel will then be fed to the equalizer (Figure 1-11).

The equalizer uses the Viterbi algorithm (maximum sequence likelihood estimation) for the estimation of the phases, which most likely have been sent at the sampling times. From these phases the information bits are calculated (Figure 1-12). A well-designed equalizer then will superimpose the energies of the single echoes constructively, so that the result in an area where the echoes are not too much delayed—delay times up to 16 μs have to be tolerated by a receiver—is better than in an area with no significant echoes (Figure 1-13).

Remaining bit errors are eliminated using another Viterbi decoder at the transmitter for the convolutionally encoded data sequences.

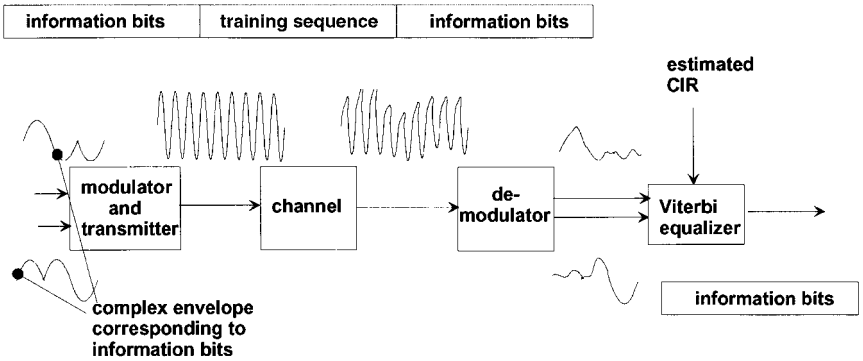


Figure 1-12 Channel equalization.

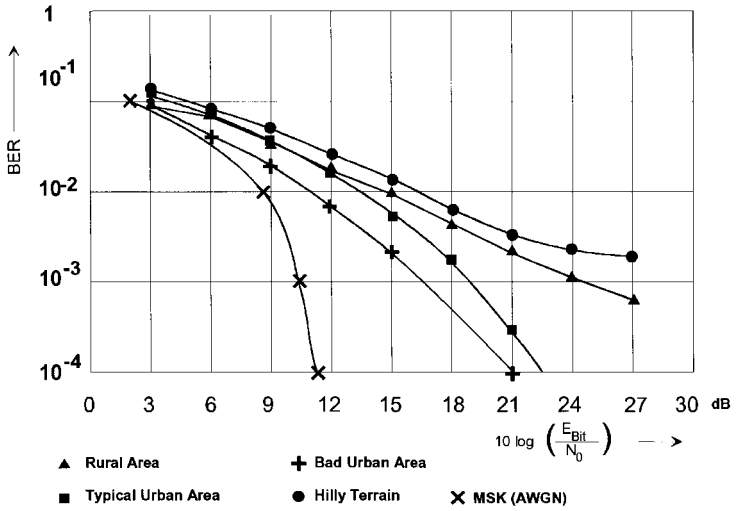


Figure 1-13 BERs after the channel equalizer in different areas.

The ability of a mobile receiver to work in a hostile environment, such as the radio channel with echoes, must be proved. The test is performed with the aid of a fading simulator. The fading simulator simulates different scenarios with different delay times and different Doppler profiles. A signal generator generates undistorted I/Q modulated RF signals, which are downconverted into the baseband. Here the I/Q signals are digitized and split into different channels, where they are delayed and attenuated, and where Doppler effects are superimposed. After combination of these distorted signals at the output of the baseband section of the simulator these signals modulate the RF carrier, which is the test signal for the receiver under test (Figure 1-14).

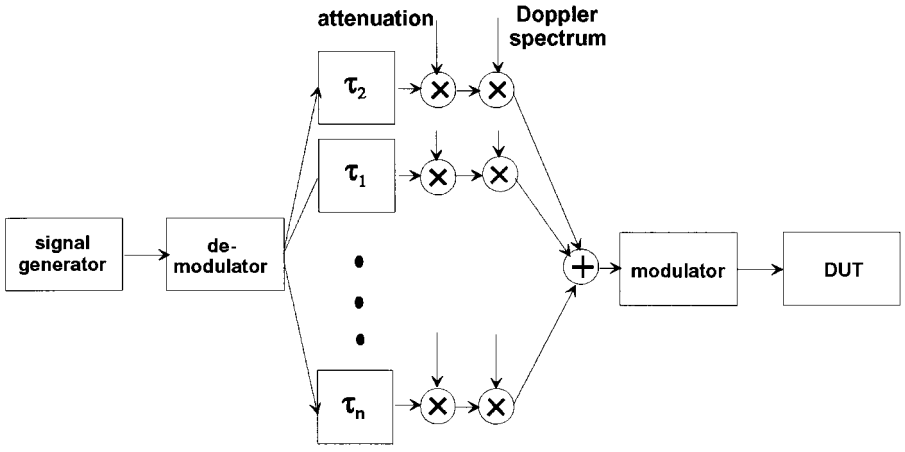


Figure 1-14 Fading simulator.

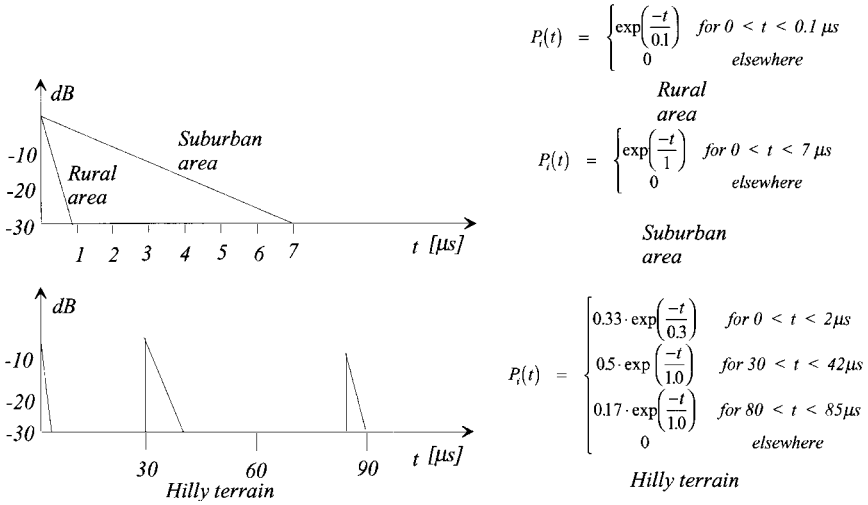


Figure 1-15 Typical landscape profiles.

To make the tests comparable, the Groupe Speciale Mobile (GSM) recommends typical profiles, for example,

- Rural area (RAx)
- Typical urban (TUx)
- Hilly terrain (HTx)

where number and strengths of the echoes and the Doppler spectra are prescribed (Figure 1-15).

1-3-3 Doppler Effect

Since the mobile receiver and some of the reflecting objects are in motion, the receive frequency is shifted because of the Doppler effect. In the case of single-path reception, this shift is calculated as follows:

$$f_d = \frac{v}{c} f_c \cos \alpha \tag{1-3}$$

where v = speed of vehicle
c = speed of light
f = carrier frequency
α = angle between v and the line connecting transmitter and receiver

In the case of multipath reception the signals on the individual paths arrive at the receiving antenna with different Doppler shifts because of the different angles α_i, and the receive spectrum is spread. Assuming an equal distribution of the angles of incidence, the power density spectrum can be calculated as follows:

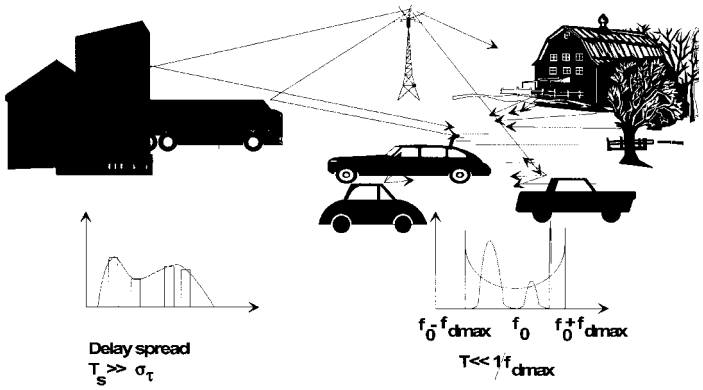


Figure 1-16 Doppler spread.

$$P(f) = \frac{1}{\pi} \frac{1}{\sqrt{f_d^2 - f^2}} \quad \text{for } |f| < |f_d| \tag{1-4}$$

where f_d = maximum Doppler frequency

Of course, other Doppler spectra are possible in addition to the pure Doppler shift described above: for example, spectra with a Gaussian distribution using one or several maxima. A Doppler spread can be calculated from the Doppler spectrum analogously to the delay spread (Figure 1-16).

1-3-4 Transfer Function

The fast Fourier transform (FFT) value of the channel impulse response is the transfer function $H(f, t)$ of the radio channel, which is also time dependent. The transfer function describes the attenuation of frequencies in the transmission channel. When examining the frequency dependence it will be evident that the influence of the transmission channel on two sine-wave signals of different frequencies becomes greater with increasing frequency difference. This behavior can adequately be described by the coherence bandwidth, which is approximately equal to the reciprocal delay spread; that is,

$$(\Delta f)_c = \frac{1}{T_d} \tag{1-5}$$

If the coherence bandwidth is sufficiently wide and, consequently, the associated delay spread is small, the channel is not frequency selective. This means that all frequencies are subject to the same fading. If the coherence bandwidth is narrow and the associated delay spread wide, even very close adjacent frequencies are attenuated differently by the channel. The effect on a broadband-modulated carrier with respect to the coherence bandwidth is obvious. The sidebands important for the transmitted information are attenuated to a different degree. The result is a considerable distortion of the receive signal combined with a high bit error rate even if the received field strength is high. This characteristic of the radio channel again speaks for the use of narrowband modulation methods (Figure 1-17).

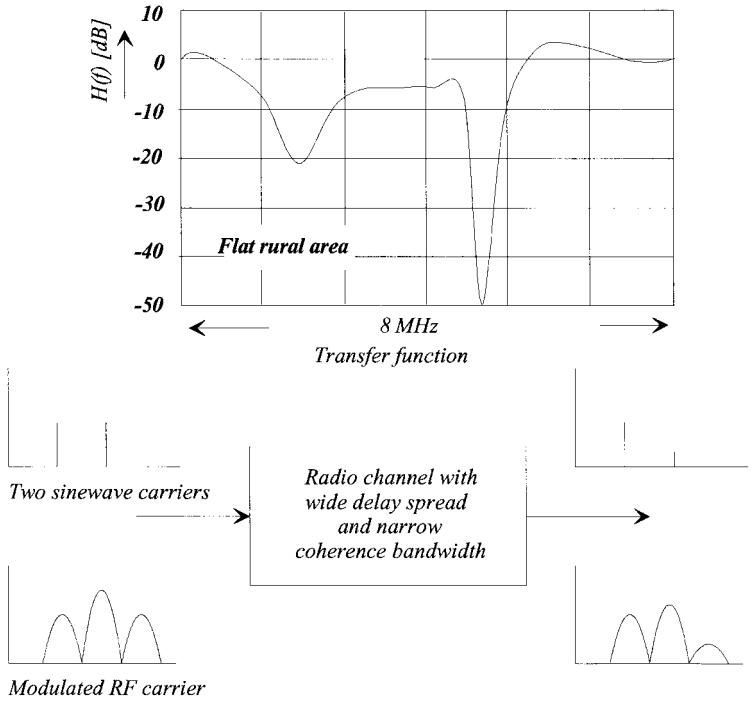


Figure 1-17 Effect of transfer function on modulated RF signals.

1-3-5 Time Response of Channel Impulse Response and Transfer Function

The time response of the radio channel can be derived from the Doppler spread. It is assumed that the channel rapidly varies at high vehicle speeds. The time variation of the radio channel can be described by a figure, the coherence time, which is analogous to the coherence bandwidth. This calculated value is the reciprocal bandwidth of the Doppler spectrum. A wide Doppler spectrum therefore indicates that the channel impulse response and the transfer function vary rapidly with time (Figure 1-18). If the Doppler spread is reduced to a single

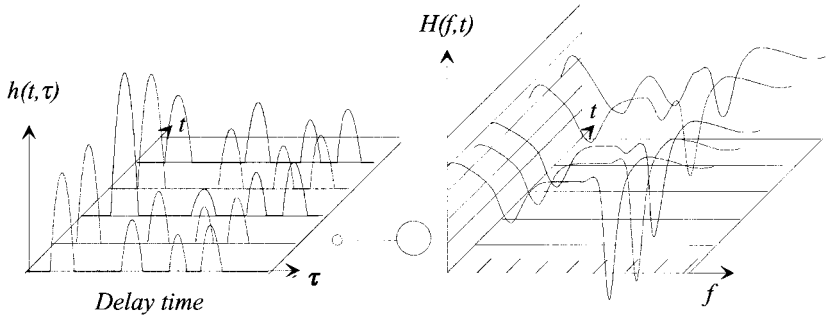


Figure 1-18 Channel impulse response and transfer function as a function of time.

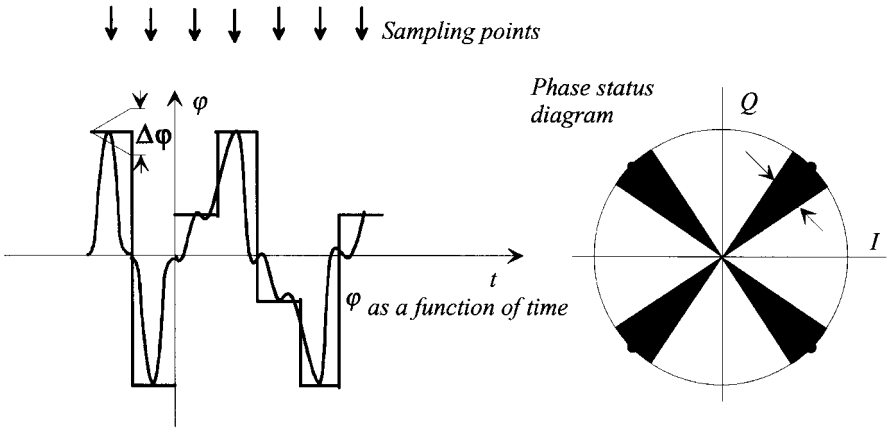


Figure 1-19 Phase uncertainty caused by Doppler effect.

line, the channel is time invariant. In other words, if the vehicle has stopped or moves at a constant speed in a terrain without reflecting objects, the channel impulse response and the transfer function measured at different times are the same.

The effect on information transmission will be illustrated in an example. In the case of *M*-ary phase shift keying (MPSK) modulation using hard keying, the transmitter holds the carrier phase for a certain period of time; that is, for the symbol period *T*. In the case of soft keying with low-pass-filtered baseband signals for limiting the modulated RF carrier, the nominal phase is reached at a specific time, the sampling time. In both cases the phase error $\varphi_f = f_d T_S$ is superimposed onto the nominal phase angle, which yields a phase uncertainty of $\Delta\varphi = 2\varphi_f$ at the receiver. The longer the symbol period the greater the angle deviation (Figure 1-19). Considering this characteristic of the transmission channel, a short symbol period of $T_S << (\Delta t)_c$ should be used. However, this requires broadband modulation methods.

Figure 1-20 shows the field strength or power arriving at the mobile receiver if the vehicle moves in a Rayleigh distribution channel. Since the phase depends on the vehicle position,

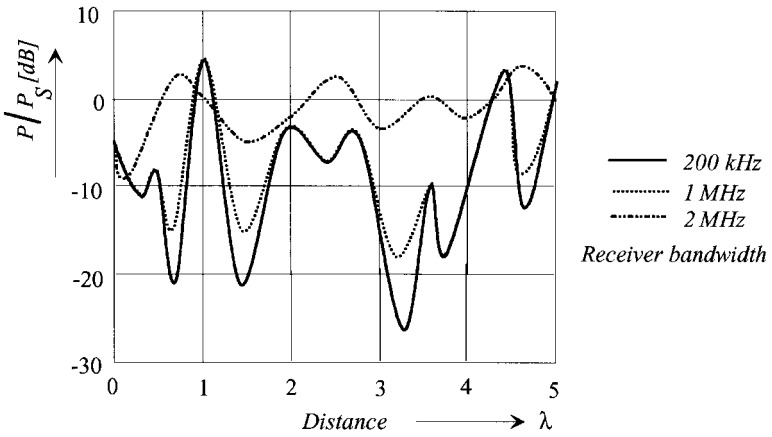


Figure 1-20 Effect of bandwidth on fading.

the receiver moves through positions of considerably differing field strength at different times (time dependence of radio channel). In the case of frequency-selective channels this applies to one frequency only; that is, to a receiver using a narrowband IF filter for narrowband emissions. As Figure 1-20 shows, this effect can be reduced by increasing the bandwidth of the emitted signal and consequently the receiver bandwidth.

1-3-6 Lessons Learned

The strongly frequency-selective channel causing inadmissible distortion of the broadband-modulated carrier and channel impulse responses like those expected in a hilly terrain speak in favor of narrowband modulation methods with long symbol periods. In hilly terrain extensively delayed echoes cause intersymbol interference when broadband modulation with short symbol periods is used. On the other hand, narrowband modulation has the disadvantage that the signals arrive at the receiver considerably attenuated and reception may be interrupted for an indefinite period of time. The burst errors occurring in digital information transmission cannot be corrected even with the most elaborate error protection methods.

These transmission interruptions can be avoided by using broadband modulation methods, but these are sensitive to strongly frequency-selective channels. Since broadband modulation is obtained through the use of short symbol periods, broadband modulation is unsuitable when greatly delayed echoes are expected.

It remains to be defined when a signal is considered a narrowband and when a broadband signal. This question shall be answered with the aid of an example. Apart from extremely narrowband analog modulation methods—as are used for sound broadcasting in the long-wave, mediumwave, and shortwave bands—FM sound broadcasting transmissions in the VHF bands are narrowband. In the case of digital modulation this means that transmissions with a rate of 400 kbps modulated onto a carrier with a bandwidth efficiency of 1.5 (bps)Hz over a bandwidth of approximately 300 kHz can be regarded as narrowband transmissions according to the definition above. Consequently, a DAB signal with this gross bit rate would be a narrowband signal and suitable for transmission on the radio channel with restrictions only. Based on the experience with conventional FM broadcasting systems in large cities and hilly terrain, this was obvious from the very beginning.

Consequently, it is necessary to find ways for spreading the band artificially without reducing the bandwidth efficiency. This means that a large band must be available for the transmission of several programs, the full bandwidth being used by all the programs without mutual interference.

Several approaches can be adopted to tackle the problem (Figure 1-21). One way would be a continuous change of the transmit and receive frequency according to a defined pattern (frequency hopping). This method is used in mobile radio, for instance, but only marginal investigations have been made in this respect for DAB.

Another possibility is to multiply the symbols of the individual programs with digital signals (pseudo-noise function) using a much higher bit rate so that a higher symbol rate is obtained. In this case the different programs are assigned different functions, which must be orthogonal to each other (code-division multiple access, CDMA). The “chopped” bit streams of the individual programs are modulated onto carriers of identical frequency and the modulated carriers are added. A correlation receiver knowing the pseudo-noise function divides the incoming CDMA signal into the individual programs. The disadvantage is obvious. Symbol periods are very short and elaborate means will be required for compensating the intersymbol interference.

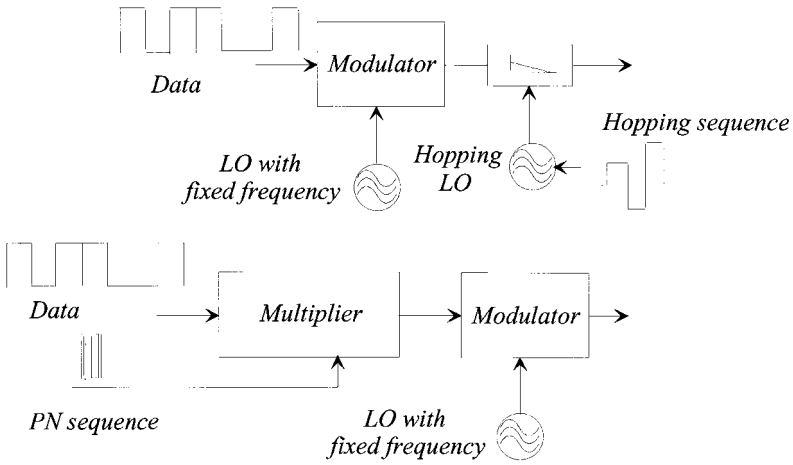


Figure 1-21 Band spreading by frequency hopping and CDMA.

A different approach has been chosen for DAB, which does not involve continuous and elaborate channel measurements, so that it would be possible to use favorably priced receivers as are demanded in the field of consumer electronics. The method used for DAB is a multicarrier method, where the information to be transmitted is spread onto many carriers using time and frequency interleaving. The terms time and frequency interleaving will be explained in the course of this discussion. The result is a broadband transmission method with long symbol periods. However, certain limitations caused by the Doppler effect will have to be accepted particularly at high carrier frequencies.

1-3-7 Wireless Signal Example: The TDMA System in GSM

Frequency-Division Multiple Access (FDMA). In analog radio systems the trend has always been toward a more efficient utilization of the available frequency spectrum by reducing the channel spacing. The number of radio channels obtained at a channel spacing of 12.5 kHz is of course twice that obtained at 25 kHz. However, any improvement brings about its disadvantage: the narrower the channel spacing, the higher the required frequency accuracy and the lower the possible maximum deviation of the frequency modulation. The latter leads to a poorer transmission quality due to the lower S/N ratio. Furthermore, the gaps between the channels, which must be a number of kilohertz wide for safety reasons, also reduce the available system bandwidth (see Figures 1-22 and 1-23).

The use of an available system spectrum divided into individual frequency channels enables the user to simultaneously access a multitude of different frequencies. This multiple access is called *frequency-division multiple access* (FDMA). Consequently, all radio systems with a spectrum divided into channels are FDMA systems. At present, the technically useful limit is reached with a channel spacing of 10–12.5 kHz.

Advantages of FDMA

- Simultaneous access to a given bandwidth by many subscribers
- Increase in the number of channels through reduction of channel spacing

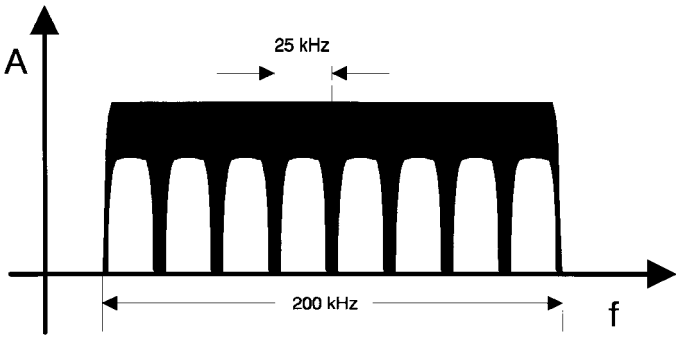


Figure 1-22 Channel spacing in broadband/narrowband systems.

Disadvantages of FDMA

- Higher frequency accuracy required
- Transmission quality decreasing with reduction of channel bandwidth
- Better rejection filters required
- One transmitter/receiver required per channel

Time-Division Multiple Access (TDMA). With TDMA systems, the available bandwidth is divided into considerably fewer, and therefore wider, channels than in FDMA systems. Each of these channels is available to several subscribers quasi-simultaneously. (See Figure 1-24.) However, a given subscriber can use the whole channel for a very short period (timeslot) only; for the rest of the time, no access is available. This serial access of several users is repeated within a fixed time frame.

Advantages of TDMA

- Simultaneous use of a specific bandwidth by a great number of subscribers
- Depending on the number of available timeslots, several subscribers can be served by one transmitter/receiver

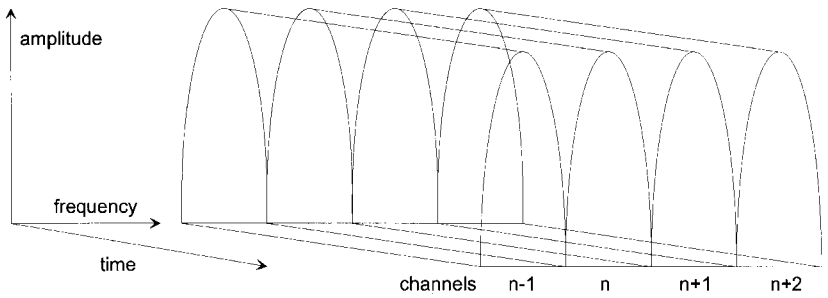


Figure 1-23 Frequency-division multiple access (FDMA).

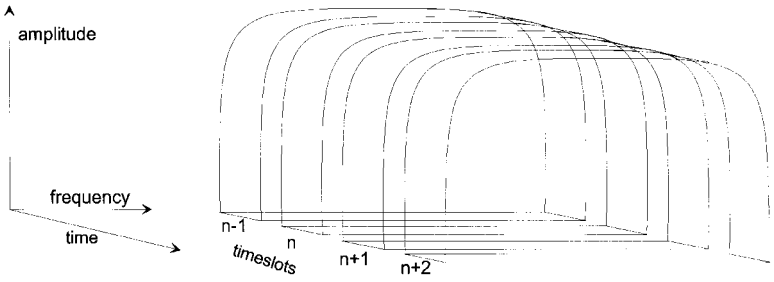


Figure 1-24 Time-division multiple access (TDMA).

- Transmitter and receiver are not permanently switched on (saves battery power)
- The RF section may carry out other tasks in the intervals between transmission and reception
- Reduced susceptibility to frequency-selective fading in the case of larger channel bandwidths

Disadvantages of TDMA

- Accurate time synchronization of subscribers required
- Higher processor capacity required
- Broadband modulators required

Code-Division Multiple Access (CDMA). The increasing use of low-priced and powerful signal processors allows a less common technique of multiple access to be employed in mass communication systems. In the case of *code-division multiple access* (CDMA), the whole system bandwidth is available to all subscribers at any time; that is, all send and receive simultaneously, with each using a specific code (Figure 1-25).

Logic "1" represents a certain bit sequence; logic "0" is the inversion of this sequence. The different signals are distinguished in the receiver by means of a cross-correlation of the received signal, which comprises a great number of codes, with the bit sequence expected so that the desired transmission signal can be detected.

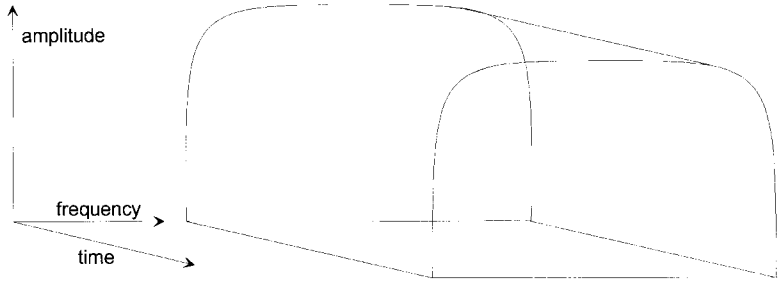


Figure 1-25 Code-division multiple access (CDMA).

Advantages of CDMA

- Simultaneous use of a specific channel or subband by many subscribers
- Several signals can be received simultaneously by one receiver
- Reduced susceptibility to frequency-selective fading in the case of large channel bandwidths
- More subscribers can be served
- Reduced costs for radio network planning

Disadvantages of CDMA

- Accurate time synchronization of subscribers required
- Fast transmitter power control over a wide dynamic range
- No experience with use in mass communication

TDMA in GSM

RF Data. In spite of the competition with other mobile radio systems, a common frequency band could be defined for GSM worldwide. All operators who signed the GSM memorandum of understanding committed themselves to install their GSM systems within the standardized frequency range. The competition for frequencies mainly affects countries using NMT900, the frequency range of which corresponds to the GSM P band. TACS also partly overlaps the GSM P band; the G1 band is completely within the TACS range. Cordless telephones operating in accordance with the CT1 standard also use the upper end of the GSM P band. CT1+ telephones, which had been assigned a frequency range below the P band years ago to protect them against GSM, have now been ousted by the G1 band. See Table 1-1.

Table 1-1 RF data for GSM900 and GSM1800

Variable	GSM900		GSM1800
Frequency range	P band	G1 band	
Uplink (MHz)			
(MS transmitting)	890–915	880–890	1710–1785
Downlink (MHz)			
(BTS transmitting)	935–960	925–935	1805–1880
Duplex spacing (MHz)	45	45	95
Spectrum (MHz)	2×25	2×10	2×75
Frequency channels	124	49	374
Channel numbers (ARFCN)	1–124	975–1023	512–885
Channel spacing	200 kHz		
Modulation	GMSK with BxT = 0.3		
Data transmission rate	270.833 kbps		
Bit duration	3.69 µs		

Since each frequency channel is divided into eight timeslots, transmitter and receiver operate in an intermittent mode, with the receive and transmit time in the upper and lower channels shifted by three timeslots (Figure 1-26). Although this alternative sending and receiving scheme operates in half-duplex rather than full-duplex operation, the received signal sounds continuous to the user.

TDMA Structure

FRAME AND MULTIFRAME. All GSM radio channels are organized in frames of approximately 4.62-ms duration. The frames are continuously repeated. Each frame is divided into 8 timeslots of approximately 577 μs each. A timeslot contains an information packet, the burst. The 26-type multiframes are used on all timeslots containing a traffic channel (voice and/or data); the 51-type multiframes on all timeslots are reserved for control channels. See Figure 1-27.

The TDMA structure uses other frame types above the multiframe level, as shown in Table 1-2.

TDMA TIMERS. The frame number within the hyperframe is counted continually so that counting of the TDMA clock restarts after approximately 3.5 hours. The frame number therefore represents a time unit in the GSM system. Similar to time counting, where the seconds are combined into minutes, hours, and days, GSM does not count the absolute frame numbers but uses timers instead. These timers are structured as shown in Table 1-3.

The absolute frame number is obtained by a multiplication of the three timers. However, on certain occasions a short version of the timers is used.

Burst Structures. Information between base station and mobile is sent in the timeslots. In each slot a certain amount of information—that is, a burst—can be transmitted. Normally the timeslot is occupied by the normal burst (Figure 1-28), which is used for signaling as well as for voice and data transmission.

Each part of the burst serves a specific purpose as described below.

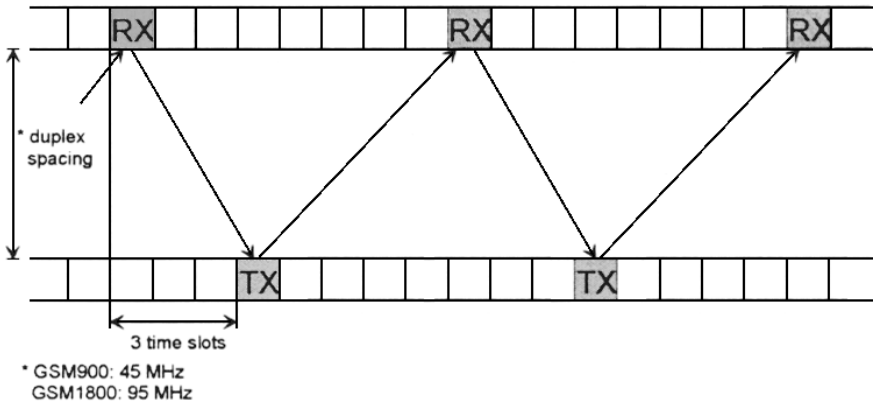


Figure 1-26 Duplex spacing of transmission and reception.

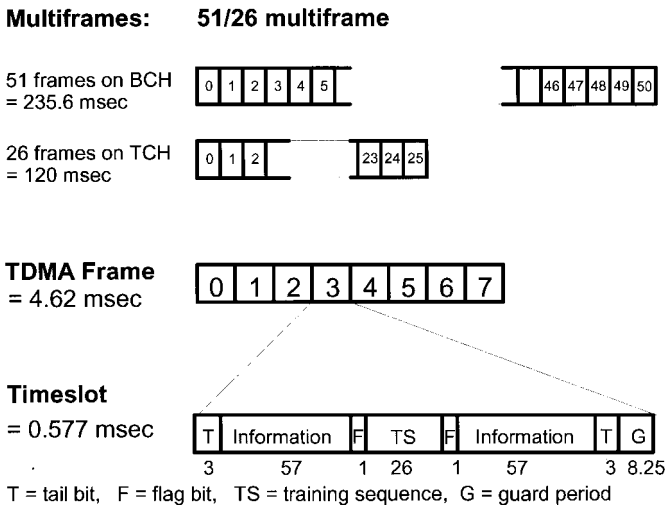


Figure 1-27 Timeslot, frame, and multiframe.

Information Bits. The normal burst is able to transmit 2×57 information bits. Since this information can be replaced every 4.62 ms during an ongoing call, the average theoretical transmission rate is

$$114 \text{ bits} \div 4.62 \text{ ms} \approx 24.7 \text{ kbps} \tag{1-6}$$

At the same time, this rate also represents the maximum transmission rate that can be obtained in the GSM system using one timeslot per transmission per time frame. Consequently, the transmission rate could be increased only if more than one timeslot is used for the transmission.

The bit rate is much lower in the control channels; that is, the above transmission rate is only attained by the mobile station if a traffic channel has been set up. In this case, the base station and the mobile station use a signaling channel, which also uses up capacity, in addition to the voice or data channel. Table 1-4 shows the assignment of the theoretically available capacity with a traffic channel setup.

Table 1-2 Superframes and hyperframes

Superframe	= 51 × 26 frames
	= 1326 frames
	= 6.12 s
Hyperframe	= 2048 frames
	= 2.715.648 frames
	= 3 h, 29 min, 3.5 s

Table 1-3 TDMA timers

T1:	= FN div. (26 × 51)	Range:	0–2047
T2:	= FN mod 26	Range:	0–25
T3:	= FN mod 51	Range:	0–50
FN (Frame Number)		Range:	0 2715647
FN_{max}	$= 51 \times 26 \times 2048 - 1$		

Training Sequence. In the middle of the normal burst, a 26-bit training sequence, the bit sequence of which is known to the receiver, is sent. The *training-sequence code* (TSC) can be one of eight different sequences. These sequences are stored in all receivers, and at the beginning of a transmission, the base transceiver station (BTS) decides on the TSC to be used. The training sequence serves two main purposes: bit synchronization and estimation of channel impulse response.

BIT SYNCHRONIZATION. Data transmitted via the air interface are in the asynchronous mode; that is, the receiver has to regenerate the bit clock from the data stream. To enable synchronization in the receiver, the transmitter adds synchronization bits to the information stream. Therefore, in normal data transmission, data telegrams start with a "...10101010..." sequence so that the receiver can regenerate the bit clock. A predefined bit word informs the receiver when the actual information (block synchronization) starts. A receiver synchronized in this way is able to decode the data stream online. The training sequence of the burst has to assume both synchronization tasks. Since it is in the middle of the burst, direct decoding is not possible.

Each burst must first be stored in the receiver and then decoded by postprocessing. The reason for using this method is the second task of the training sequence. Synchronization itself is carried out by means of cross-correlation; that is, the expected training sequence is compared (correlated) to the center of the received burst and to the beginning and end of the training sequence so that the bit clock is also known. A burst containing other than the expected training sequence cannot be synchronized and decoded.

COMPENSATION OF MULTIPATH RECEPTION. The signal from the transmitter (in Figure 1-29, BTS → MS; the same applies also in the opposite direction) arrives at the receiver not only along the direct path but also via various other paths as a result of reflection and diffraction caused by obstacles in the signal path.

Propagation conditions on these additional paths differ from those on the direct path. For instance, we can expect signals traveling via additional paths to exhibit:

- Longer travel times because of increased path length

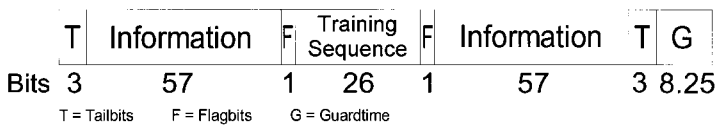


Figure 1-28 Normal burst.

Table 1-4 Transmission bit rates

	Information	Error Protection	Total
Traffic channel			22.8 kbps
Voice (full rate)	13.0 kbps	9.8 kbps	
Data	2.4 kbps	20.4 kbps	
	4.8 kbps	18.0 kbps	
	9.6 kbps	13.2 kbps	
Control channels			0.95 kbps
Idle frame			0.95 kbps
Total			**24.7 kbps**

- Various strengths
- Different Doppler shifts

Because of the different travel times, the signals arrive with a different phase at the receiving antenna. Depending on this phase, components may be canceled—that is, they may totally disappear—or added so that a high-quality signal is received for only a short period of time. RF level variations are statistically distributed; level shifts due to fading may be as great as 40 dB.

In addition to RF level fading, another annoying effect is encountered that, uncompensated, would make correct signal decoding rather difficult. Because of the additional distance the signal travels via the indirect path, the signal arriving at the receiving antenna exhibits time delay of its modulation in addition to variable phase shifts. The total of all channel responses to a single transmitted pulse is called *channel impulse response* (CIR, Figure 1-30). If the indirect path is only 1 kilometer longer, the GSM echo bit reaches the receiver later than the directly received bit and thus interferes with the next bit received. This *intersymbol interference* (ISI) may occur over several bits in succession. With delays of up to 15 µs,

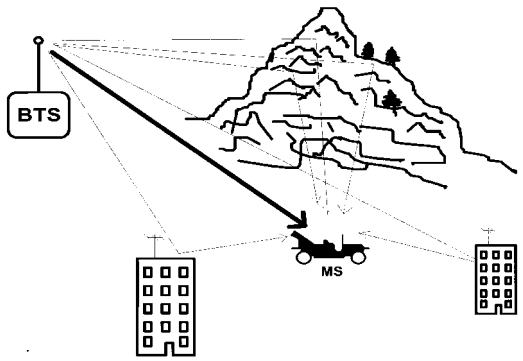


Figure 1-29 Multipath reception due to reflection and diffraction. The *base transceiver station* (BTS) is transmitting to the *mobile station* (MS).

differentiating the desired-signal components from echoes becomes more and more difficult. This problem can also be solved with the aid of the training sequence. The echoes on the delayed paths also contain the training sequence. The correlation used for detecting the original training sequence may also be used for detecting the training-sequence echoes as well as their delay and loss. With the aid of this information, the received signal can be corrected by a channel equalizer.

Guard Period. Transmission in each timeslot is terminated with a guard period (Figure 1-31) of 8.25 bit periods (≈ 30 μs). During this time the level of the burst must be reduced from nominal to a minimum value (by up to 70 dB) and the burst is modulated with so-called dummy bits (logic 1); that is, no information is transmitted. The user of the next timeslot should start sending during this guard period so that his burst has reached nominal power when the actual transmission in the timeslot starts. This means that the switching time that cannot be used for transmitting information is used twice.

Delay Correction. The integrity of a timeslot depends on whether the subscribers send only during the period assigned to them and otherwise keep quiet. This is only possible when all subscribers are accurately synchronized. For practical reasons, the clock signal is generated by the BTS and the mobile stations synchronize to it. Conflicts with adjacent timeslots may occur in the uplink where several subscribers have to share the same channel. Although several subscribers are addressed in the downlink, signals can only be transmitted by the BTS.

We will now examine the effect of a distance of 10 km between an MS and a BTS (Figure 1-32). Synchronization of the MS is as follows:

Delay over 10 km = 33.3 μs (distance ÷ velocity of light).
The sync signals from the BTS require this time for transmission.
→ MS is synchronized by 33.3 μs too late.

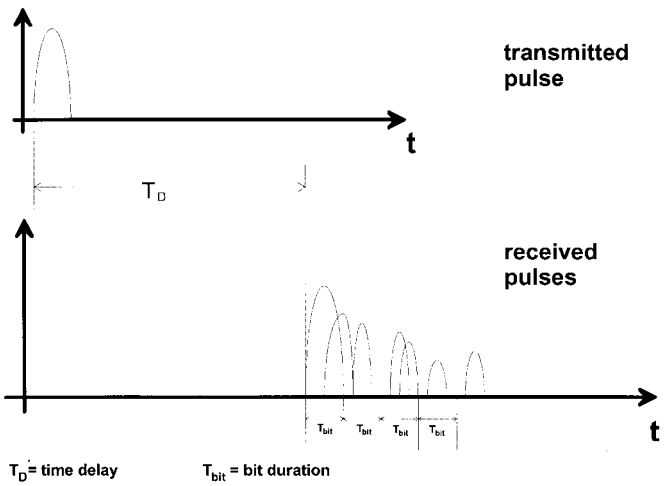


Figure 1-30 Channel impulse response.

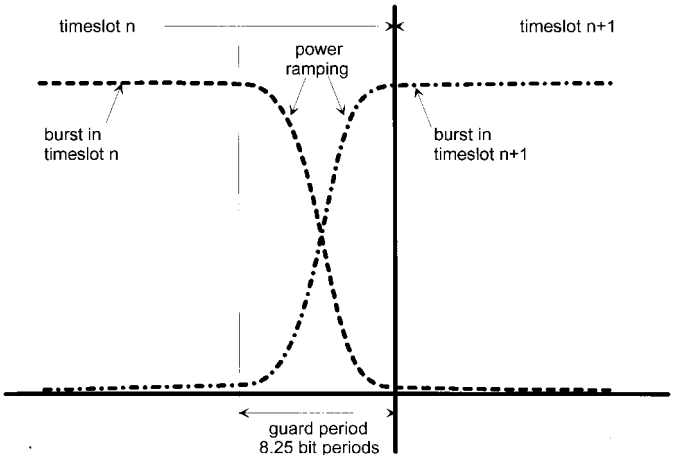


Figure 1-31 Guard period at the end of each timeslot.

MS sends a burst at the correct time from its point of view.
→ The burst is sent 33.3 µs too late.
Signal delay over a distance of 10 km → the burst requires another 33.3 µs. From the point of view of the BTS, the burst arrives with a delay of twice the delay time. Transmission cannot be made in the assigned timeslot and interferes with the next one.

The guard period at the end of each burst is only approximately 30 µs long and fully used in the example above. The greater the distance between the MS and BTS, the greater the effect of the signal delay. The only way to solve this problem is to make the MS send the burst at an earlier time. To do so, the distance between the MS and BTS must be known. The BTS determines the distance by means of a delay measurement and informs the MS of the actual delay. As a result, the MS corrects its transmission time so that the signals again arrive time-synchronized with the other mobiles at the BTS antenna.

In GSM, this procedure is called *timing advance* (TA) and is carried out continuously for all active mobile stations. Every 480 ms, a new TA value is sent to all active mobiles. A few limiting values of the GSM system can be deduced from the TA:

1. The TA is transmitted as a 6-bit word. Numerals from 0 to 63 can be represented with 6 bits. For instance, a TA of 10 means that for an MS from which bursts arrive with

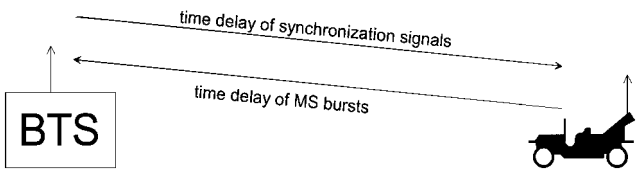


Figure 1-32 Signal delay and its effect.

a delay of 10 bit periods, transmission must be advanced by this amount. Since the TA is transmitted as an absolute value, a maximum delay of 232.5 μs (63 × 3.69 μs) can be signaled and corrected. This maximum double delay corresponds to a distance of approximately 34.9 km.

2. With the aid of the TA, the distance between the MS and the BTS can be determined. Since the smallest delay increment is a bit period, the resolution for a distance of one half of the delay corresponds to one bit period and therefore to approximately 550 m.

This synchronization scheme cannot solve an inherent problem: The delay of *the very first burst* sent by an MS cannot be corrected because, with contact between the MS and BTS yet to be established, the MS has not yet received an appropriate TA value from the BTS. Yet, if the mobile sends a normally timed burst, it may spill over into the next timeslot and interfere with transmission from another MS. To avoid causing interference in such cases, the mobile uses a special burst at the beginning of a transmission or when no valid TA has been received. Called the *access burst*, it is considerably shorter than a normal burst. The next section describes the access burst and other additional burst types in greater detail.

Burst Types. In addition to the normal burst described before, other bursts, including the access burst, are available for special purposes (Figure 1-33).

FREQUENCY CORRECTION BURST. The 142 "fixed bits" of the *frequency correction burst* (FCB) are all set to logic 0. With Gaussian minimum shift keying (GMSK), the type of modulation used in GSM, a stationary carrier frequency deviation, in this case approximately 67.7 kHz, is generated with this burst. The FCB is sent by the BTS only and used by the mobile for synchronization to the carrier frequency and for compensating a possible Doppler shift. It is sent by the BTS every 10 frames (approximately every 46 ms), but only in the timeslot 0 and only on a single (the C0) carrier.

SYNCHRONIZATION BURST. One frame after the frequency correction burst, but also in timeslot 0, the synchronization burst (SB) is transmitted. This burst is sent only by the BTS, and only on the C0 carrier. A particular difference between it and the normal burst is its considerably longer training sequence. Like the 26-bit training sequence of a normal burst, the SB's training sequence is also used for bit synchronization. Due to the great length of the sequence, synchronization can be more exact.

The twice 39 encrypted bits comprise timers T1, T2, and T3 in a coded form, and also the *base-station identification code* (BSIC). When this "GSM time" is received, the MS is synchronized to the BTS.

DUMMY BURST. A BTS must send continuously—that is, in all timeslots—on its C0 carrier because this carrier is used by the MS to find the nearest BTS and for evaluating reception quality. If no normal burst is available for transmission in a timeslot, the BTS sends dummy bursts instead, as the carrier cannot be transmitted without a modulation signal. Only the BTS sends dummy bursts, and only on the C0 carrier.

ACCESS BURST. As already pointed out, the access burst is sent when the MS first calls the BTS as a means of minimizing interference to other mobile transmitters (MTs) while initiating a delay measurement and the determination of the TA. In most cases, the access

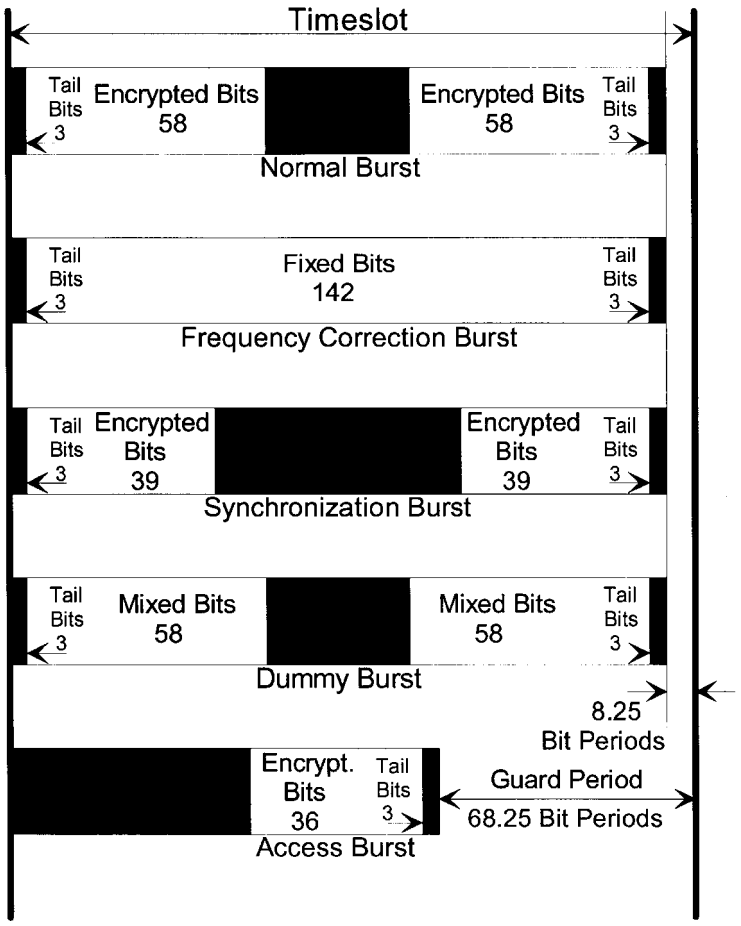


Figure 1-33 Burst types.

burst is also sent on the C0 carrier in the uplink direction, but in the case of a handover it may be transmitted on any carrier.

1-4 ABOUT BITS, SYMBOLS, AND WAVEFORMS

1-4-1 Introduction

Digital modulation of an RF carrier is the allocation of physically existing RF waveforms to the single elements of an alphabet of logical symbols where the number of allowed waveforms is equal to the number of logical elements of the alphabet (Figure 1-34). The most common alphabet is the binary one with the two logical symbols "0" and "1," but we will also deal with quaternary, octernary, and hexadecimal alphabets or more generally with *M*-ary alphabets comprising many more elements when discussing the signal generation with

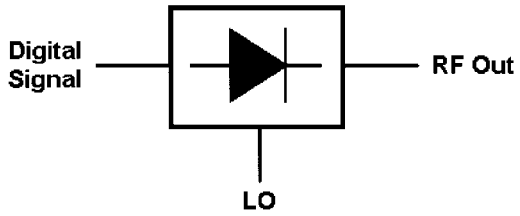


Figure 1-34 At base, digital modulation involves frequency shifting a baseband digital signal to RF. In practice, the process is more complicated than this because of bandwidth constraints on the resulting RF signal.

signal generators and dedicated software packages. The waveforms representing these symbols differ from each other by their parameter's amplitude $a(t)$, their frequency $f(t)$, and their phase $\varphi(t)$.

A modulator therefore is nothing more than a device by which this allocation is performed (Figure 1-35). From a coder it receives the logical symbols and emits at its output the corresponding waveforms $s_i(t)$. The waveform generation may be done by using a set of distinct generators (e.g., two oscillators to generate two signals with different frequencies in the case of binary frequency shift keying), by classical amplitude or frequency modulators, or by more sophisticated equipment such as I/Q modulators for M-ary modulations.

On their way across the RF channel from the transmitter to the receiver, these waveforms are distorted by noise and other disturbing properties of the RF channel.

The task of the receiver is to interpret the received waveforms $r_i(t)$ and to reallocate the proper logical symbols to them. For this purpose it is not necessary to reconstruct the original waveforms from the distorted ones (Figure 1-36). The important thing is to find out which

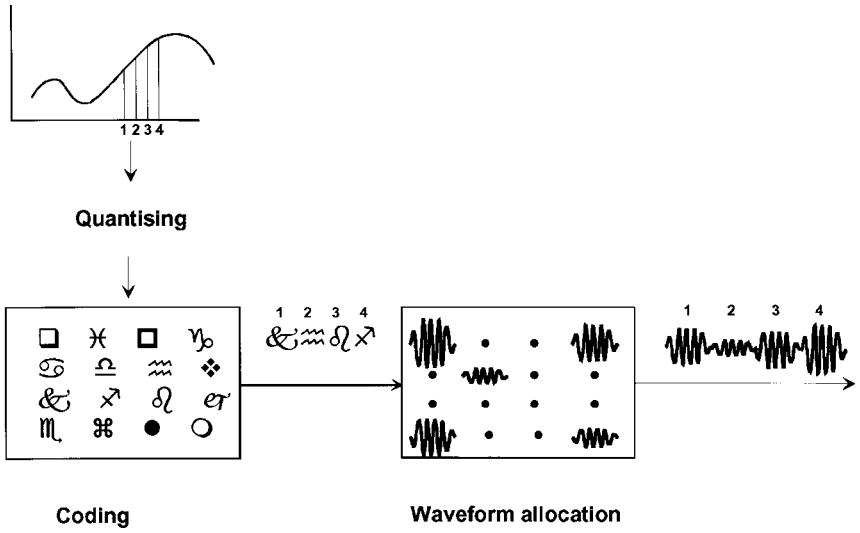


Figure 1-35 Digital modulator.

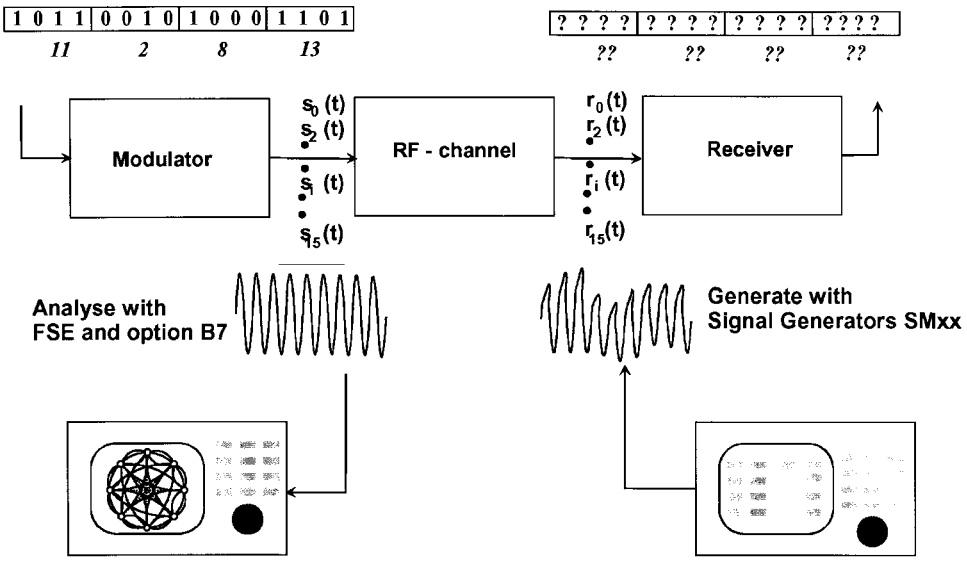


Figure 1-36 The information channel.

symbol has most probably been sent when a certain signal $r_i(t)$ has been received, a process that is known as maximum likelihood estimation.

For meaningful receiver tests, therefore, waveforms have to be generated that mimic real, distorted signals to prove the ability of a receiver to tolerate waveform distortions to a certain extent.

Representation of a Modulated RF Carrier. The waveform of a modulated RF carrier can be expressed as

$$s(t) = a(t)\cos[2\pi f_c(t)t + \varphi(t)] \tag{1-7}$$

and is defined by its amplitude $a(t)$, its carrier frequency $f_c(t)$, and its phase $\varphi(t)$. All three parameters are time variant and may be altered to generate different waveforms to represent logical symbols. If the occupied bandwidth of this modulated carrier is narrow compared to the carrier frequency f_c, we call this signal the *RF-bandpass signal* (Figure 1-37).

As any frequency variation causes a phase variation and vice versa a phase variation always causes a frequency variation, we can replace any frequency modulation by a corresponding phase modulation. Therefore we simplify the above equation to

$$s(t) = a(t)\cos[2\pi f_c t + \varphi(t)] \tag{1-8}$$

that is, we consider the carrier frequency as a constant and concentrate all frequency and phase variations into the parameter $\varphi(t)$.

For our purposes, a more suitable representation is

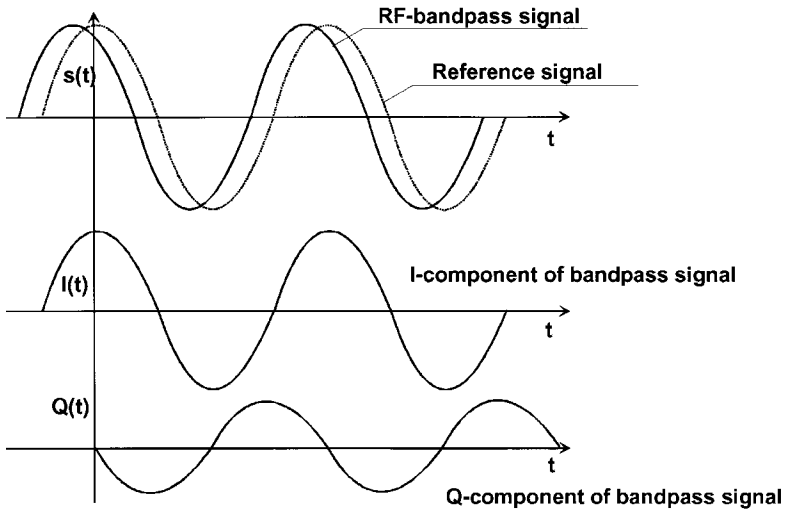


Figure 1-37 The bandpass signal and the I/Q representation of a carrier.

$$a(t)\cos[2\pi f_c t + \varphi(t)] = \cos[\varphi(t)]a(t)\cos(2\pi f_c t)$$

$$-\sin[\varphi(t)]a(t)\sin(2\pi f_c t) \tag{1-9}$$

which we call the *I/Q* representation of the RF signal. I/Q means that we have an I (in-phase) signal, namely, $\cos[\varphi(t)]a(t)\cos(2\pi f_c t)$, and a Q (quadrature) signal, namely, $-\sin[\varphi(t)]a(t)\sin(2\pi f_c t)$. These equations help us a lot in understanding an I/Q modulator. Because of the phase difference of 90° between the two carrier components, these are said to be *orthogonal* to each other.

All the information about the (modulated) carrier with the carrier frequency f_c is contained in the terms

$$c_I(t) = a(t)\cos[\varphi(t)] \tag{1-10}$$

$$c_Q(t) = a(t)\sin[\varphi(t)] \tag{1-11}$$

and, lazy as we are, we therefore disregard the terms $\cos(2\pi f_c t)$ and $-\sin(2\pi f_c t)$ from further consideration and denote the above signals $c_I(t)$ and $c_Q(t)$ as the components of the *complex baseband waveform* or *baseband signal*.

This leads us immediately to the *vector representation* of the signal, where we consider the two components $c_I(t)$ and $c_Q(t)$ of the complex baseband signal as the time-variant components of a time-variant vector with the vector length $a(t)$ and the angle to the I-axis $\varphi(t)$.We also get

$$a(t) = \sqrt{c_I^2(t) + c_Q^2(t)} \tag{1-12}$$

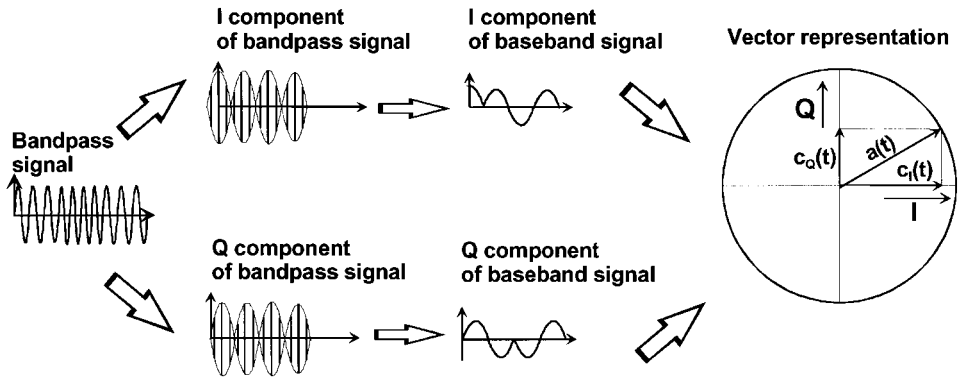


Figure 1-38 Different forms of signal representation.

$$\varphi(t) = \arctan\left(\frac{c_Q}{c_I}\right) \tag{1-13}$$

The vector can be depicted in the I/Q area (Figure 1-38).

Generation of the Modulated Carrier. Once we have realized that the modulated carrier can be represented as the sum of its I and Q components, which are the product of the two baseband components with two orthogonal RF carriers of the same frequency, it is easy to understand the hardware of the modulator (Figure 1-39). An unmodulated RF carrier is split up into two equal oscillations $\cos(2\pi ft)$, one of the two is then shifted by 0.5π and therefore is described by $-\sin(2\pi ft)$. The component $\cos(2\pi ft)$ is multiplied with the I component of the baseband signal $c_I(t)$; the other one, $-\sin(2\pi ft)$, is multiplied with the Q component $c_Q(t)$ of the baseband signal. Each multiplication may be performed using a double-balanced mixer. Afterward the two RF components are added in a simple power combiner. As it is difficult to shift the carrier by 90° over a broad frequency range, the modulated carrier is generated at an intermediate frequency and then upconverted to the wanted output frequency in a second mixer stage.

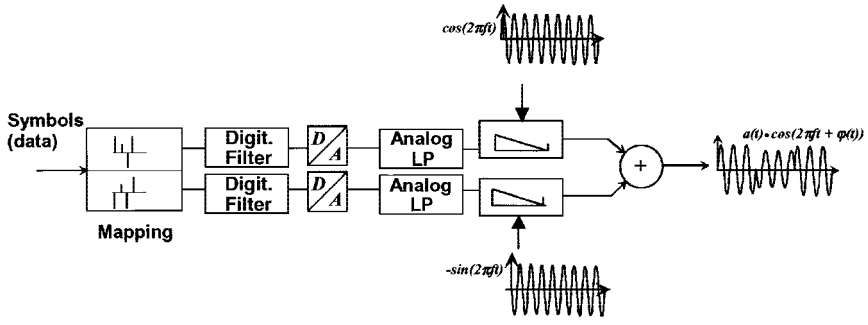


Figure 1-39 Principle of a digital I/Q modulator.

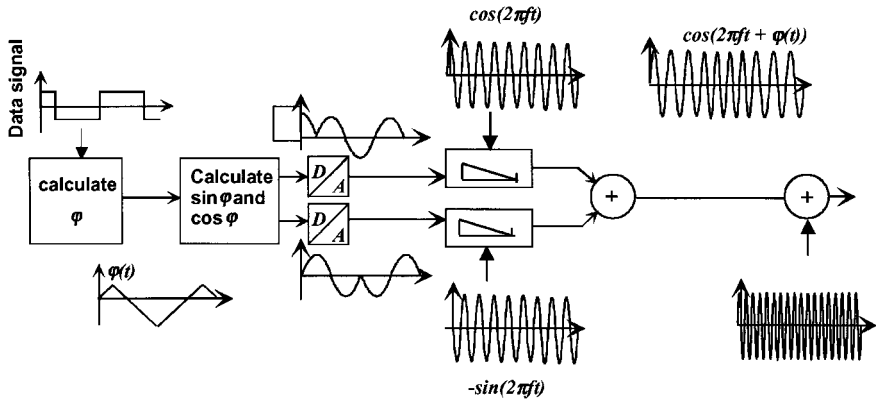


Figure 1-40 I/Q modulation (MSK and GMSK).

The baseband signals are generated by mapping every digital symbol onto a pair of digital pulses, which are fed to digital baseband filters. The output signal of these filters is D/A converted and smoothed by analog low-pass filters.

Figure 1-40 shows another example where, for a given modulation (MSK or GMSK), the instantaneous phase and then the corresponding cosinusoid and sinusoid, which modulate the two carrier components, are calculated from the data signal.

Digital designs of the modulator also exist, in which the IF-carrier generation, the time-variant phase shift, the multiplication with the baseband signals, and the sum of the components are calculated in a digital signal processor, the output of which is D/A converted and upconverted to the output frequency in the classical way. A further possibility is the generation of the modulated carrier with direct digital synthesis (DDS), as used in the Rohde & Schwarz SME signal generator.

Mapping the Data onto the Baseband Waveforms. The next question is: "How do we generate the baseband waveforms $c_I(t)$ and $c_Q(t)$?" There is no general answer to this question, as the generation of the baseband waveforms depends on the type of modulation. The following short descriptions will suffice for the moment.

Linear Modulations (All Kinds of Amplitude and Phase Shift Keying, and *M*-ary QAM)

- For binary amplitude and phase shift keying (ASK and BPSK), the data signal itself represented as a unipolar (ASK) or bipolar (BPSK) nonreturn-to-zero (NRZ) signal is the baseband waveform $c_I(t)$; the component $c_Q(t)$ does not exist.
- For *M*-ary phase shift keying and *M*-ary quadrature amplitude modulation, *N* bits are combined to form new symbols that are elements of an alphabet with $M = 2N$ elements. In the simplest case, every symbol is allocated an I and a Q amplitude during the symbol duration, which is *N* times the bit duration. The modulating signals $c_I(t)$ and $c_Q(t)$ are then staircase functions, and the modulated carrier has a time-varying envelope with the instantaneous amplitude $a(t)$ (Figure 1-41). Because the steps of the envelope cause unwanted side lobes of the RF spectrum, the baseband signals are filtered to smooth the shape of the RF envelope and reduce the occupied bandwidth of the modulated RF signal.

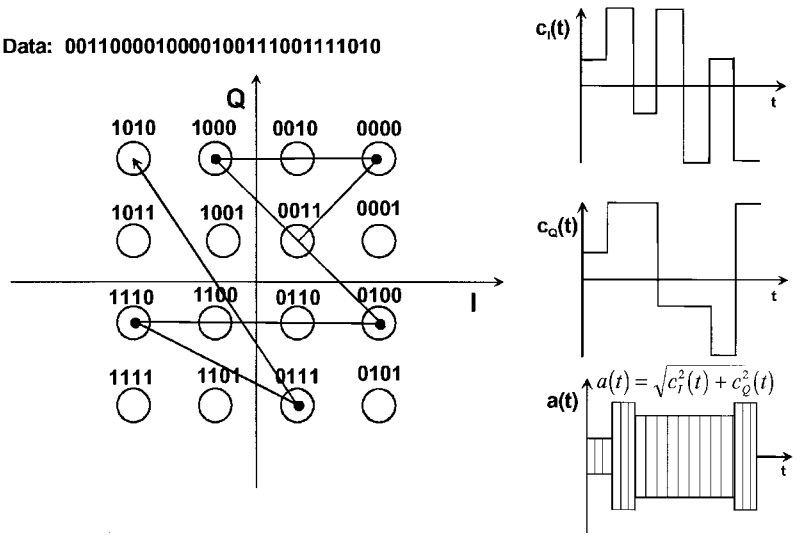


Figure 1-41 Constellation diagram, and baseband and RF signals of 16-QAM.

Nonlinear Modulations (Frequency Shift Keying, Minimum Shift Keying, and Gaussian Minimum Shift Keying)

- Despite the fact that *M*-ary frequency shift keying (FSK) could be performed using an I/Q modulator, for this type of modulation much simpler equipment such as a voltage-controlled oscillator is used as a frequency modulator. Figure 1-42 shows an example of quaternary frequency shift keying (4FSK), which also is known as 4 pulse-amplitude modulation/frequency modulation (4PAM/FM). This term indicates that every two bits are combined to a dibit that is mapped onto a baseband pulse with an amplitude taking on one of four possible levels. The pulse then is shaped by a baseband filter before being fed to the frequency modulator.
- If more precise modulations are required (e.g., MSK and GMSK, which also turn out to be frequency modulations), first the instantaneous phase of the modulated RF carrier is calculated from the data. The corresponding sine and cosine values that form the modulating baseband signals $c_I(t)$ and $c_Q(t)$ are determined from a lookup table. This operation is the reason for the fact that frequency modulation is called a nonlinear modulation.

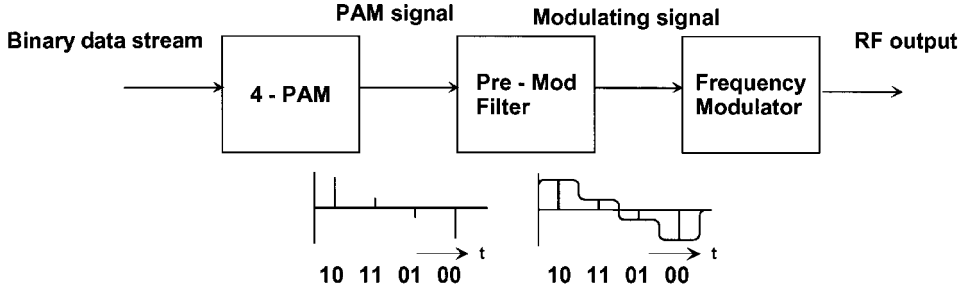


Figure 1-42 4PAM/FM.

The Spectrum of a Digitally Modulated Carrier. It is the task of any transmission process to occupy as little bandwidth as possible. The absolute lower limit in the baseband is half the symbol rate of the baseband signal, where for M-ary modulation the symbol rate r_{Symbol} is equal to the bit rate divided by $ld\,(M)$. This lower limit is only theoretical as ideal rectangular filters, which cannot be realized, were necessary. Therefore, in practice, a minimum baseband bandwidth of about $0.75r_{\text{Symbol}}$ has to be taken into account.

With linear modulation, the occupied bandwidth in the RF range is twice the occupied baseband bandwidth. This follows from the lag theorem, according to which the double-sided spectrum of a time function is shifted from $f = 0$ to the frequency $f = f_c$ when the time function is multiplied with $\cos(2\pi f_c t)$ (Figure 1-43).

Expressing this with formulas, we find

$$c(t) \quad \circ\!\!-\!\!\bullet \quad C(f) \tag{1-14}$$

$$e^{j2\pi f_0 t}c(t) \quad \circ\!\!-\!\!\bullet \quad C(f - f_0) \tag{1-15}$$

Therefore, if the baseband spectrum is limited by a low-pass filter, the RF spectrum is limited as if it were filtered by an RF bandpass filter with twice the bandwidth of the baseband filter.

Demodulation of Digitally Modulated Carriers. The demodulation process is an estimation process. The receiver compares the received waveform with all the possible waveforms and decides which symbol has been received. The waveform allocated is the one that is most similar to the received signal. This process is also called *maximum likelihood estimation*.

Figure 1-44 shows a possible demodulator. The received signals are cross-correlated with the stored ones in M correlators. High output from a given correlator indicates a match between its stored symbol and a symbol in the incoming signal.

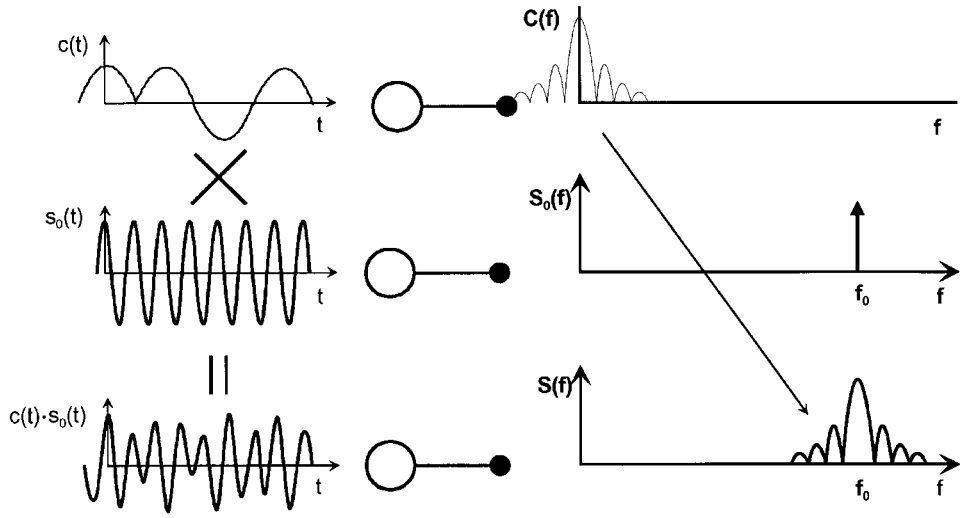


Figure 1-43 Occupied bandwidth in the baseband and the RF range.

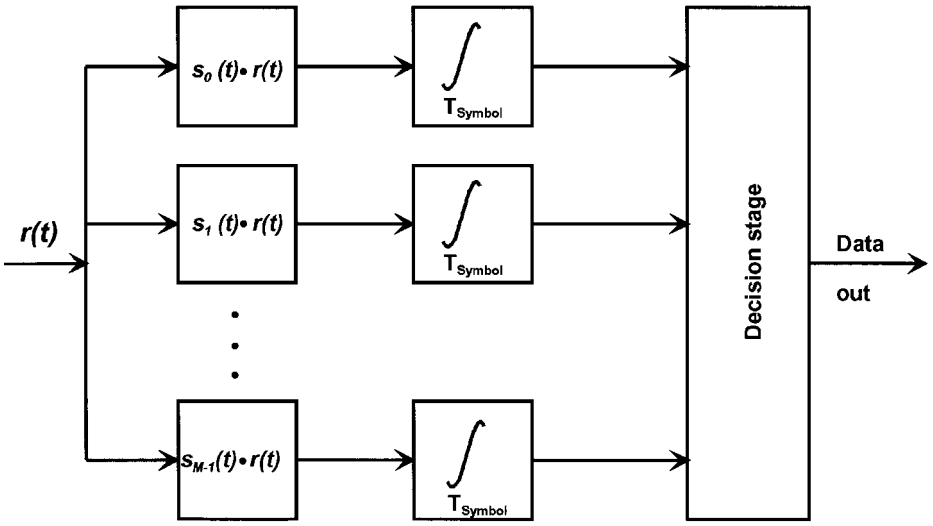


Figure 1-44 Correlation receiver.

As can easily be seen, such a receiver is quite complex. Therefore, other principles of receivers have been studied. The most often used receiver is based on an I/Q demodulator, the principle of which is shown in Figure 1-45.

The received signal is split into two equal components. One component is multiplied with the reconstructed carrier signal, the other with its orthogonal counterpart; that is, the $\pi/2$-shifted reconstructed carrier. After low-pass filtering, this multiplication delivers the I and the Q components of the baseband signal, which is more or less distorted by interference, multipath, and noise in the radio channel. These demodulated baseband signals then can be A/D converted and treated using simple or sophisticated algorithms to decide which symbols were originally sent. This type of demodulation requires an exact reconstruction of the unmodulated carrier, with respect not only to frequency but also to phase, and is called *coherent demodulation*. While the frequency recovery is in principle quite simple, the

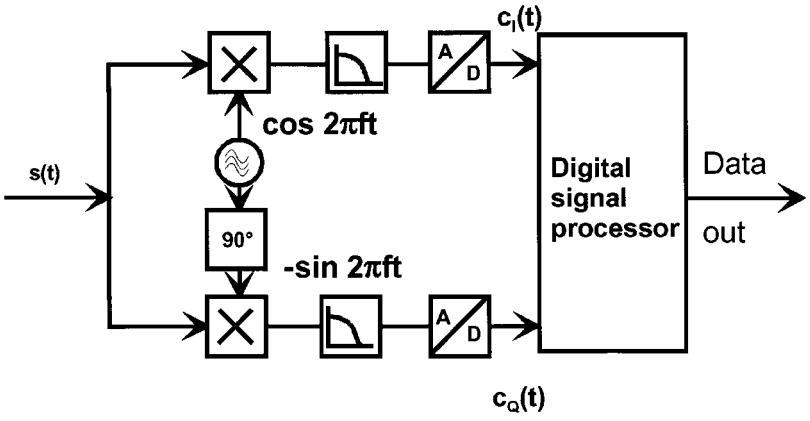


Figure 1-45 Coherent demodulation.

acquisition and tracking of the signal phase require complex signal processing of the baseband signal.

Coherent demodulation and its accompanying costly signal acquisition and tracking are necessary for all types of phase and quadrature amplitude modulation in which any part of the transmitted information is coded in the absolute phase of the carrier. It only can be avoided when the information is coded in the phase difference of two sequential waveforms. Such modulation schemes are called *differential phase modulation*. Examples of such modulation include differential binary phase shift keying (DBPSK) and differential quadriphase shift keying (DQPSK) and their derivatives.

The following list shows common digital modulation types and the telecommunication systems that use them:

BPSK—Satellite radio links, GPS, Inmarsat
QPSK—Satellite radio links, GPS, IS-95
OQPSK—IS-95 (reverse link)
π/4-QPSK—NADC, PHS, PDC. TETRA
GMSK—GSM, DCS 1800
GFSK (Gaussian FSK)—DECT
COFDM (coded orthogonal frequency-division multiplexing)—DAB, DVB

1-4-2 Some Fundamentals of Digital Modulation Techniques

The first waveforms that were used for "digital modulation" were amplitude shift keying (ASK), similar to on/off-keyed Morse code, and phase shift keying (similar to FM). Figure 1-46 shows the resulting waveforms in the time domain. It is also useful to look at them in the phase domain—the in-phase (I) and quadrature (Q) planes. From Figure 1-47 we see that ASK has only two states: no signal and signal at I (in-phase), while phase shift keying can have two states: +I and –I. In the frequency domain, ASK and BPSK have the output spectra

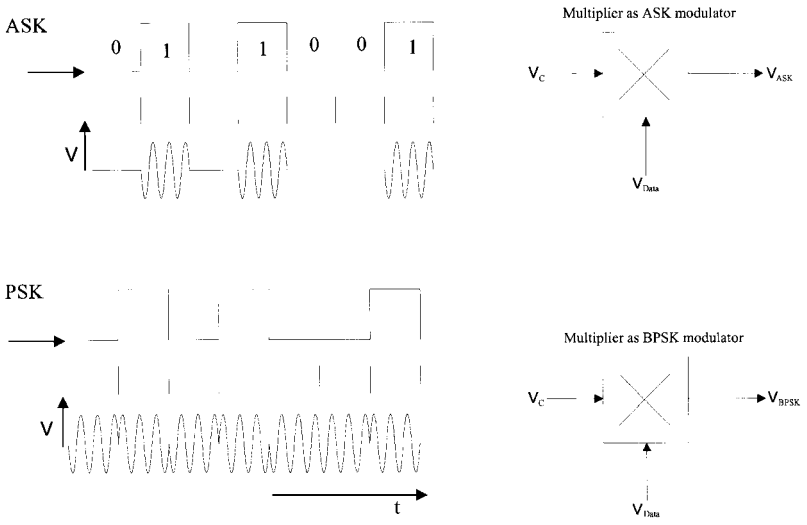


Figure 1-46 Amplitude shift keying (ASK) and phase shift keying (PSK) in the time domain.

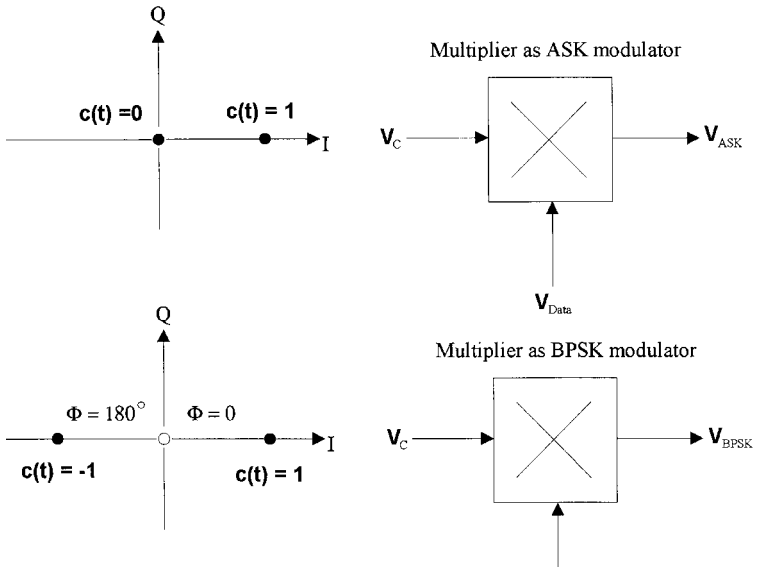


Figure 1-47 ASK and PSK in the I/Q plane.

shown in Figure 1-48. These output spectra, with their (sin *X*)/*X* appearance, depend on the rise and decay times and duty cycle of the modulating signal.

BPSK sends one data bit per signal state. Another way to put this is that each of BPSK's two signal states is a *symbol* that stands for just one bit—0 or 1. Transmitting more than one bit per symbol would allow us to improve our data throughput per unit of time, and this is

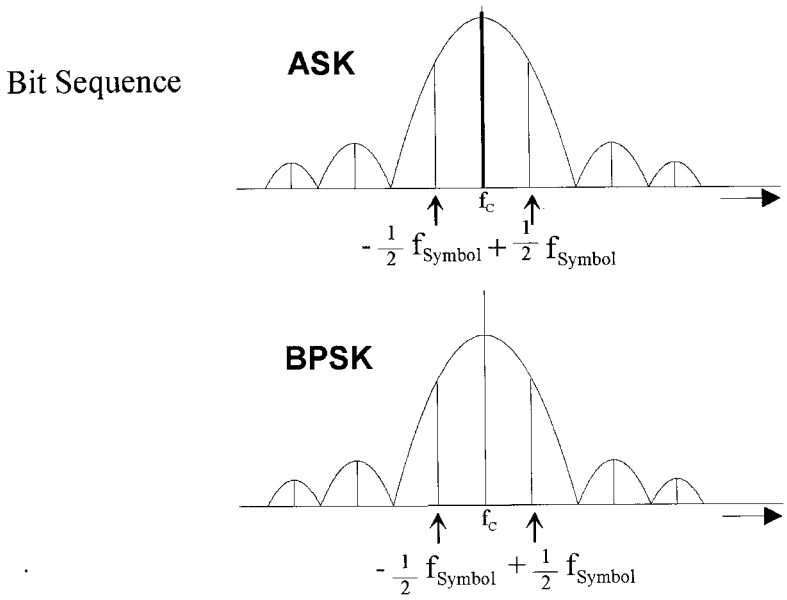
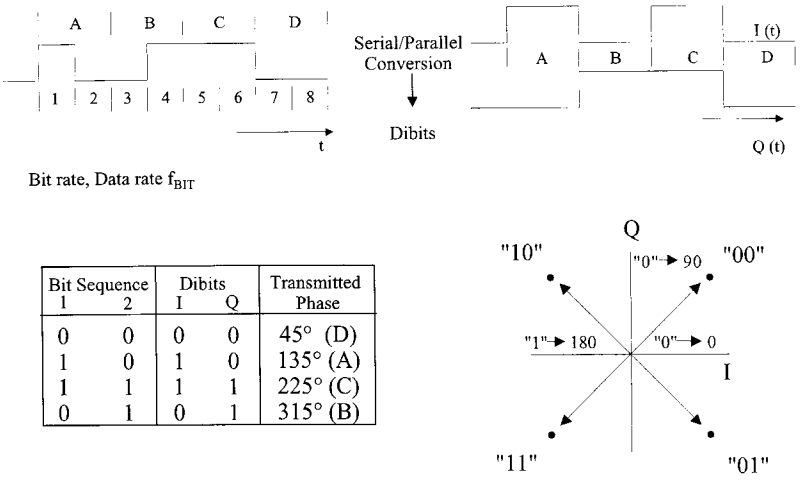


Figure 1-48 ASK and BPSK in the frequency domain.



Bit Sequence 1	2	Dibits I	Q	Transmitted Phase
0	0	0	0	45° (D)
1	0	1	0	135° (A)
1	1	1	1	225° (C)
0	1	0	1	315° (B)

Figure 1-49 Quadrature PSK (QPSK) modulator (I/Q modulator).

exactly what's done in quadrature PSK (QPSK), which uses *four* possible signal states as symbols for 2-bit sequences called *dibits* (Figure 1-49). Figure 1-50 shows a QPSK modulator and the QPSK constellation diagram; Figure 1-51 shows the result of QPSK in the time and frequency domains; Figure 1-52 shows a spectrogram of an actual QPSK emission.

Moving from ASK to BPSK or QPSK results in increased resistance to noise. Figure 1-53 illustrates this by graphing bit error rate versus signal-to-noise ratio (here expressed as 10 log (E_{bit}/N_0), where E_{bit} is energy per bit and N_0 is the noise [23]). Figure 1-54 compares the maximum interference voltages for BPSK and QPSK.

Figure 1-55 shows bandwidth requirements and constellation diagrams for BPSK and QPSK. Unfiltered BPSK and QPSK are purely angle-modulated emissions; for each, all symbols are transmitted at the same amplitude. As we'll see, however, bandwidth-limiting a BPSK or QPSK signal to minimize adjacent-channel interference results in envelope variations that must be preserved through careful circuit design if bit errors are to be minimized.

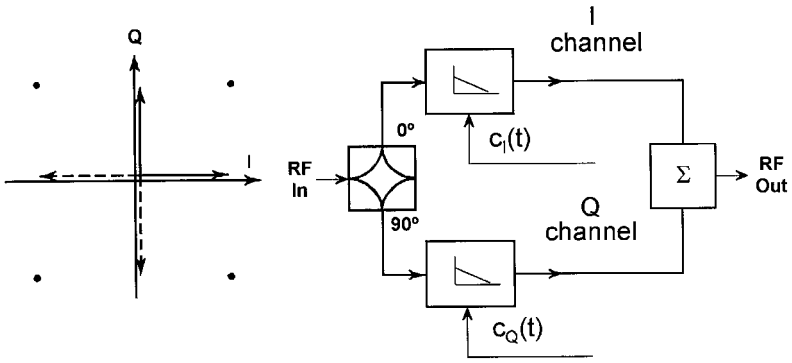


Figure 1-50 QPSK constellation diagram (*left*) and modulator (*right*).

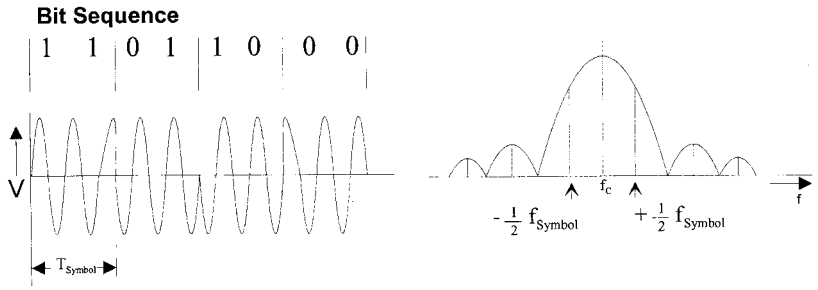


Figure 1-51 Result of QPSK modulation in the time and frequency domains.

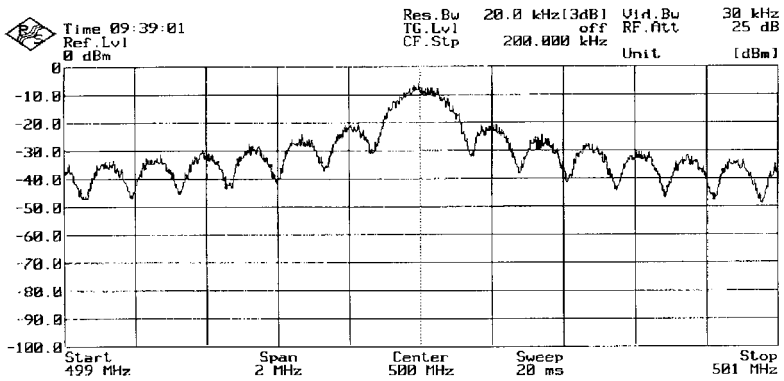


Figure 1-52 QPSK spectrum.

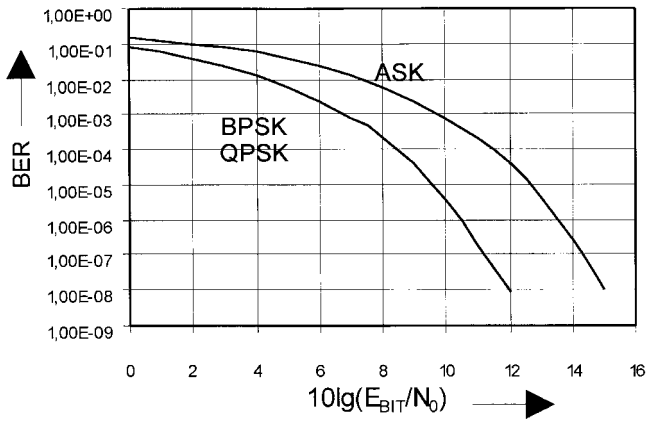


Figure 1-53 Bit error rate (BER) in terms of E_{bit}/N_0 for BPSK, QPSK, and ASK.

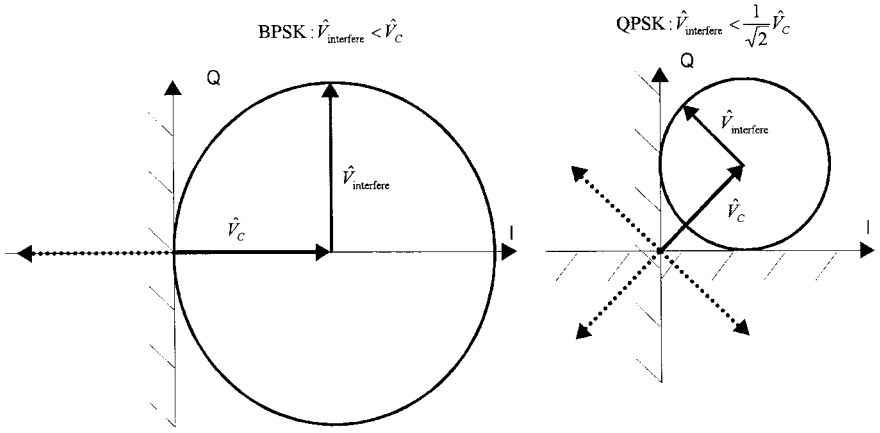


Figure 1-54 Maximum interference voltages for BPSK and QPSK, where $\hat{V}_C$ is the carrier voltage of the desired signal and $\hat{V}_{interfere}$ is the carrier voltage of the interfering signal.

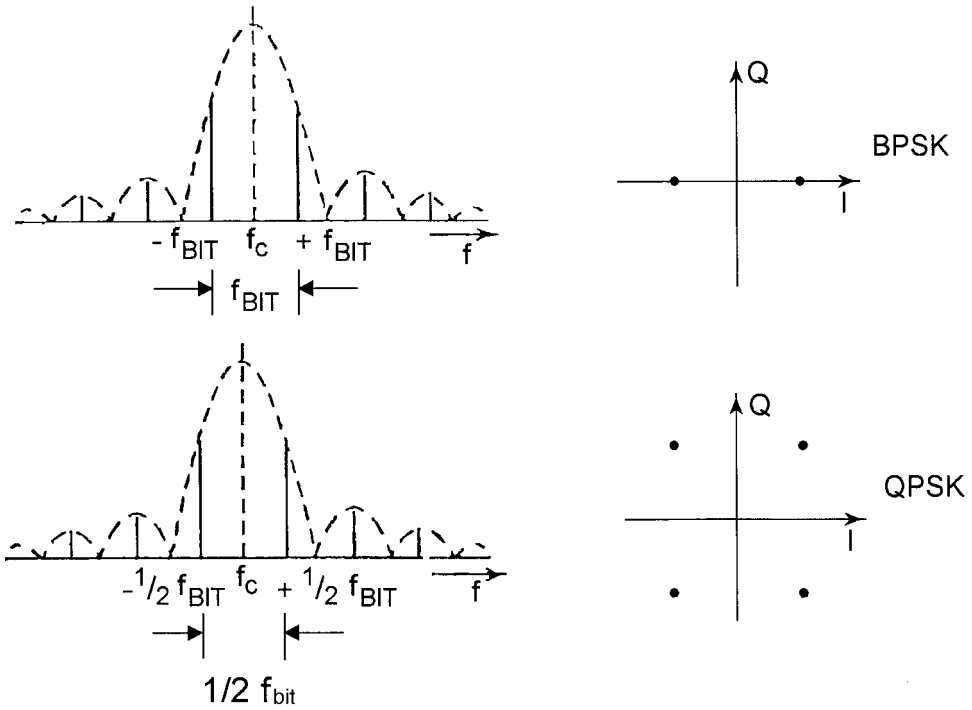


Figure 1-55 Bandwidth requirements (*left*) and constellation diagrams (*right*) for BPSK and QPSK. For each emission, all symbols are transmitted at the same amplitude, as is indicated by their equidistance from the constellation origin.

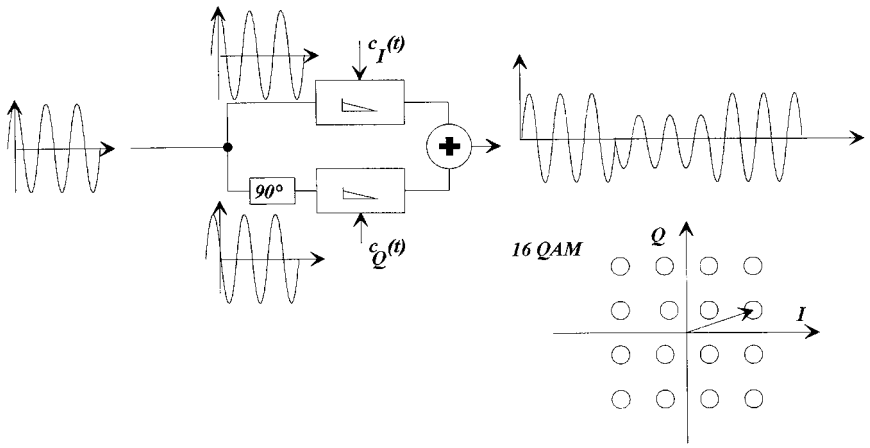


Figure 1-56 Quadrature amplitude modulation (QAM), with constellation diagram for 16-state QAM (16-QAM). Because transitions between QAM symbols involve shifts in phase *and amplitude*; even unfiltered QAM is a combination of amplitude and angle modulation.

Further increases in bit rate per symbol are possible with phase-modulation schemes that divide the 360° of each carrier cycle into increasingly smaller segments. Transmitting 3 bits per symbol, with each symbol spaced from its neighbors by 45°, gives 8-PSK; transmitting four bits per symbol, with each symbol spaced from its neighbors by 22.5°, results in 16-PSK. Increasing the data rate by increasing the number of bits per symbol doesn't give us something for nothing, however: The smaller the phase difference between adjacent symbols, the greater the likelihood that phase shifts resulting from phase noise, multipath reception, and other sources of phase perturbation will cause demodulation errors.

Modulation schemes that differentiate their data symbols through a combination of phase *and amplitude* shifts are susceptible to disturbances in amplitude in addition to phase. Figure 1-56 shows the modulator and constellation diagram for 16-state quadrature amplitude

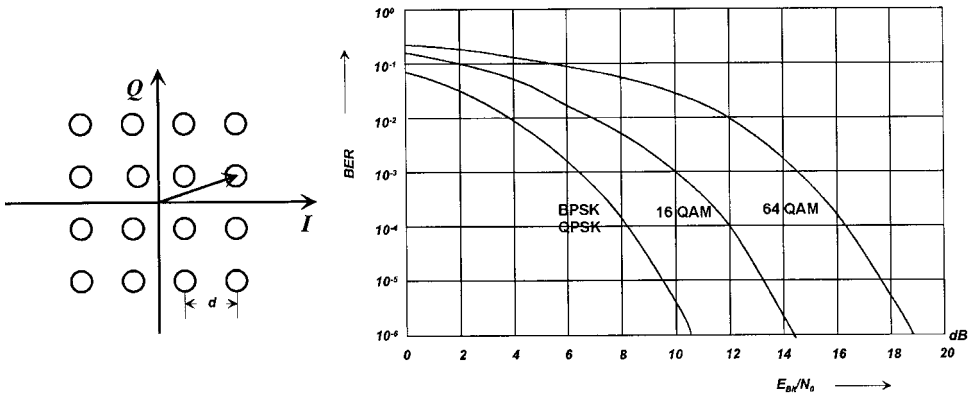


Figure 1-57 Bit error rate versus E_{bit}/N_0 for BPSK/QPSK, 16-QAM, and 64-QAM, showing the significantly greater SNR necessary for a given BER as the number of signal states is increased. In 64-QAM, each symbol represents 6 bits.

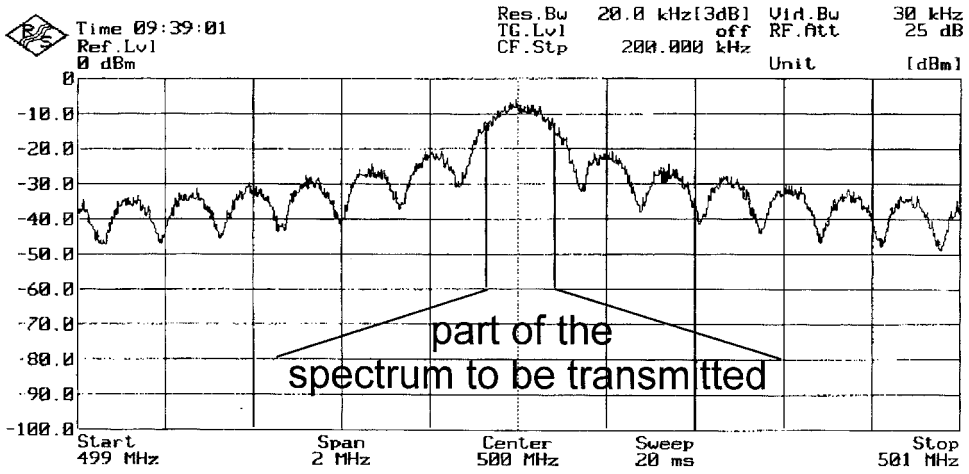


Figure 1-58 QPSK spectrum resulting from pseudorandom binary sequence (PRBS) data. Most of an angle-modulated emission's infinite sidebands are unnecessary for communication and can be removed by filtering—at the cost of introducing amplitude variations that must be sufficiently preserved to keep the bit error rate low.

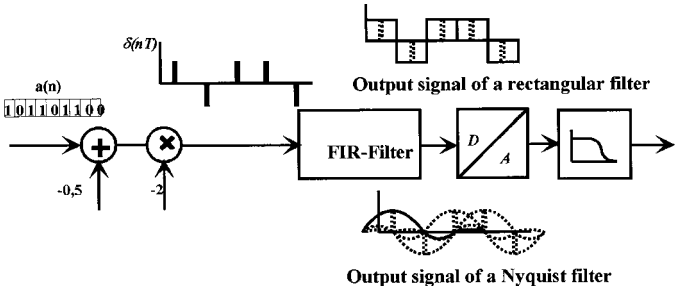


Figure 1-59 Baseband filtering. Figure 1-60 shows the result for QPSK.

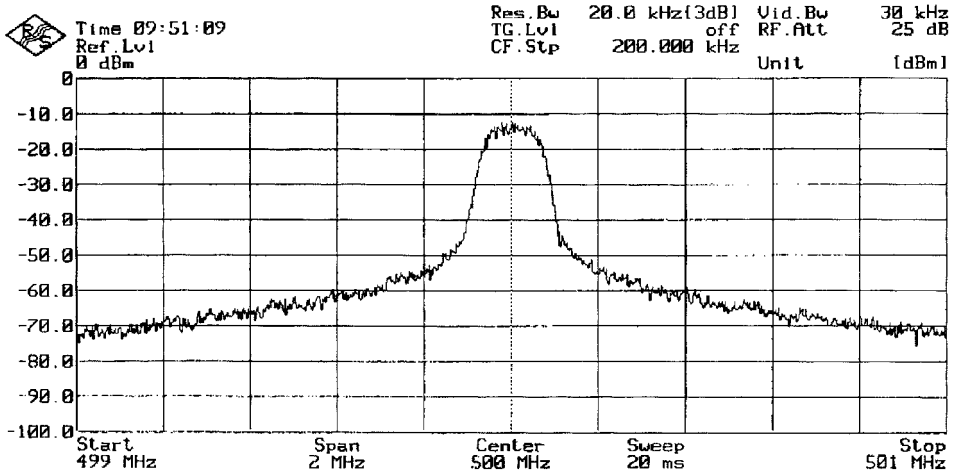


Figure 1-60 Spectrum of a band-limited QPSK signal.

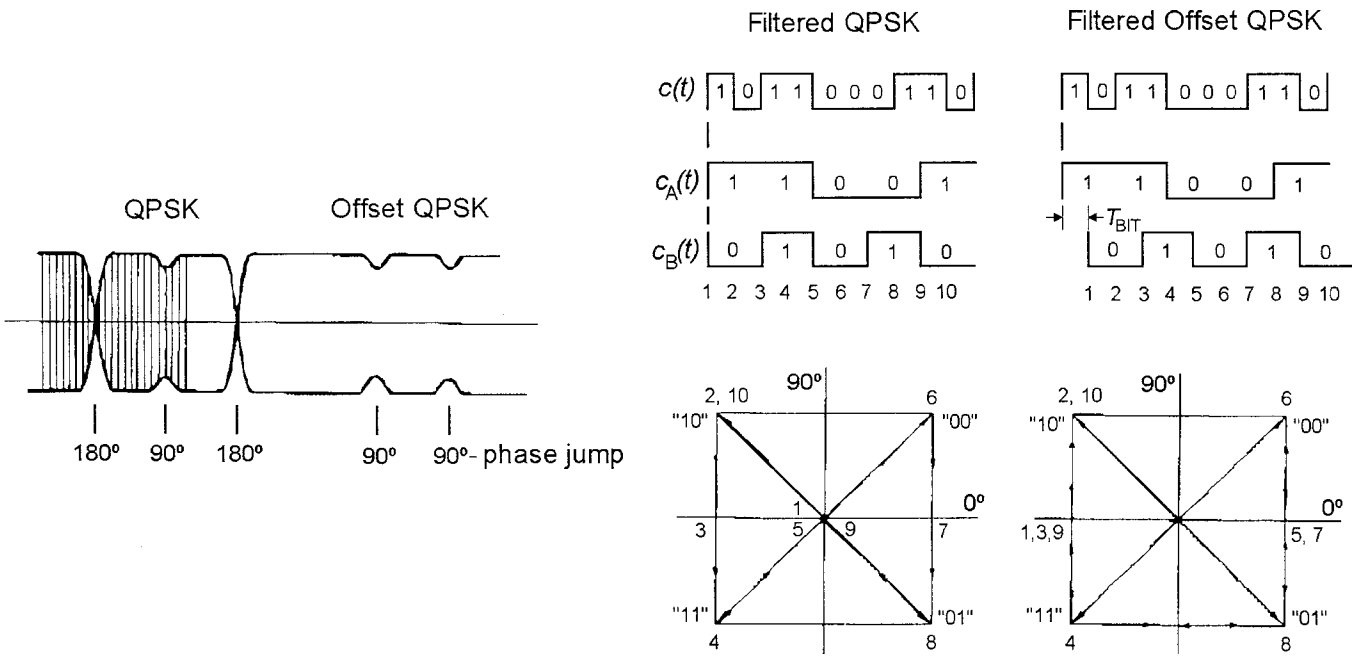


Figure 1-61 Offset QPSK (OQPSK) exhibits reduced envelope variations compared to standard QPSK by time-offsetting the I and Q channels.

modulation (16-QAM), in which each symbol represents 4 bits. Figure 1-57 shows bit error rate versus SNR for BPSK/QPSK, and 16- and 64-QAM. QAM systems can support very high bit rates—*if* sufficient SNR can be maintained and disturbances in signal amplitude and phase can be kept under control [3].

The infinite number of sidebands produced by angle modulation must be reduced by filtering to minimize the spectrum occupied by an emission to just that necessary for communication (Figure 1-58). Considerations of filter realizability and transmitter frequency agility dictate that this filtering must be done at baseband (Figure 1-59). The filter characteristics are optimized in accordance with the modulation scheme used (Figure 1-60). The filtering is done in using DSP, minimizing component count and resulting in characteristics that are essentially temperature independent and identical from unit to unit.

Fundamental angle-modulation theory tells us that the envelope of an angle-modulated signal does not vary in amplitude. Such an emission could be amplified in highly nonlinear stages without distortion. But because angle-modulated signals must be band-limited for practical use, and because band-limiting an angle-modulated signal strips away sideband energy that would contribute to its envelope constancy, the emissions used in modern wireless systems, even those that do not involve intentional amplitude modulation, vary in amplitude with modulation. How much a wireless emission's amplitude varies depends on the degree of filtering and the particular modulation scheme used. Various schemes, many of which rely on particular synergies of coding, modulation, and filtering, have been devised for minimizing amplitude variation in digitally modulated signals. Figure 1-61 shows one [offset QPSK (OQPSK)]; Figure 1-62 compares the spectra of three [QPSK, minimum shift keying (MSK), and Gaussian MSK (GMSK)]. Detailed discussion of such techniques is more concerned with coding, logic circuitry, and software than with radio hardware and is therefore beyond the scope of this book. What is important to the circuit designer is that stages handling the emission(s) on which a system depends must be sufficiently amplitude-linear to maintain the modulation integrity of the signal and, in the case of transmitters, to keep adjacent-channel interference and other spurious emissions within acceptable levels.

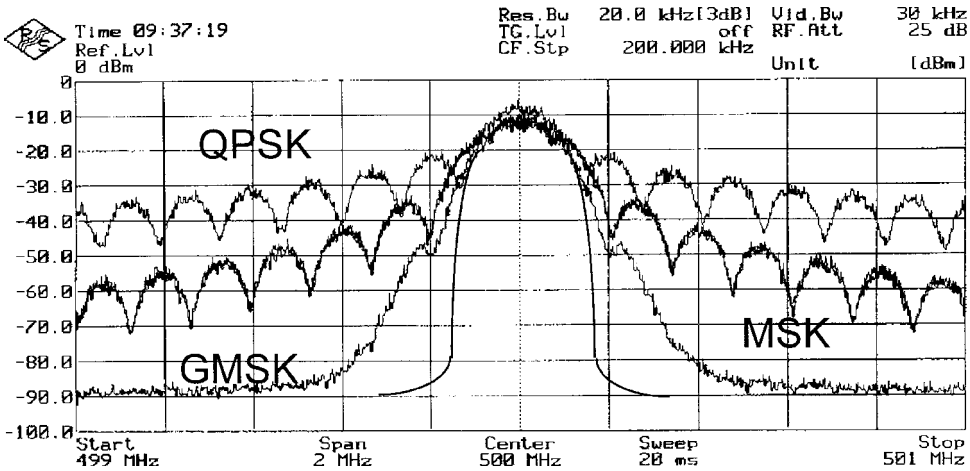


Figure 1-62 Comparison of QPSK, MSK, and GMSK.

1-5 ANALYSIS OF WIRELESS SYSTEMS

1-5-1 Analog and Digital Receiver Designs

For the purpose of showing the capability of modern computer-aided design (CAD), we now show first an analog receiver and following this an analog/digital receiver with its functionality and performance.

Receiver Design Examples

Analog Receiver Design. A transmitter modulates information signals onto an RF carrier for the purpose of efficient transmission over a noise-filled radio channel. The RF receiver's job is to demodulate that information signal while maintaining a sufficient signal-to-noise ratio (SNR). This must be done for widely varying input RF power level and with the presence of noise and interferers.

Modern communication standards place requirements on key system specifications, such as RF sensitivity and spurious response rejection. These system specifications must then be separated into individual circuit specifications via an accurate overall system model. CAD can play a major role in modeling systems and in determining the individual component requirements. A 2.4-GHz dual-down conversion system (see Figure 1-63) will be used as an example. The first IF is 200 MHz and the second IF is 45 MHz.

As a signal propagates from the transmitter to the receiver it is subject to path loss and multipath, resulting in extremely low signal levels at the receive antenna. RF sensitivity is a measure of how well a receiver can respond to these weak signals. It is specified differently for analog and digital receivers. For analog receivers there are several sensitivity measures, including minimum discernible signal (MDS), signal-to-noise plus and distortion ratio (SINAD), and noise figure. For digital receivers the typical sensitivity measure is maximum bit error rate (BER) at a given RF level. Typically, a required SINAD at the baseband demodulator output is specified over a given RF input power range. For example, audio measurements may require 12-dB SINAD at the audio output over RF input powers ranging from −110 to −35 dBm. This can then be translated to a minimum carrier-to-noise (C/N) ratio at the demodulator input to achieve a 12-dB SINAD at the demodulator output (see Figure 1-64).

Determining receiver sensitivity requires accurate determination of each component's noise contribution. Modern CAD software, such as Symphony by Ansoft, can model noise from each stage in the system including oscillator phase noise. The C/N ratio can then be plotted as a function of input power (see Figure 1-65).

This plot enables straightforward determination of the minimum input power level to achieve a certain C/N ratio. Once the sensitivity has been specified the necessary gain, or

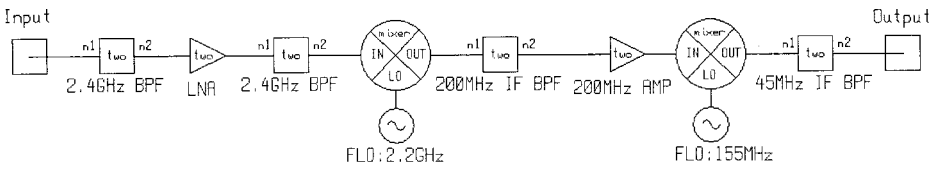


Figure 1-63 Dual-downconversion receiver schematic.

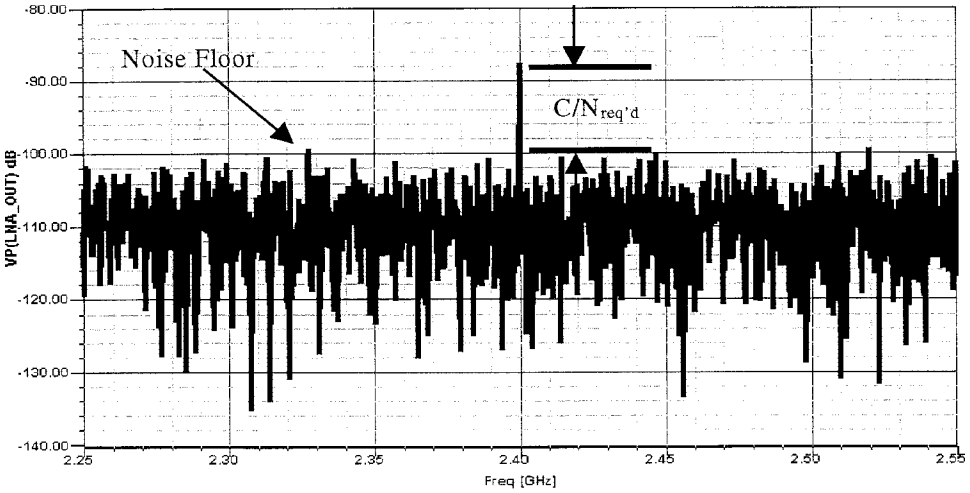


Figure 1-64 Receiver sensitivity measure showing required C/N.

loss, of each component can be determined. A budget calculation (see Table 1-5) can be used to examine the effects of each component on a particular system response.

Another key system parameter is receiver spurious response. The receiver's mixers typically cause the spurious responses. The RF and LO harmonics mix and create spurious responses at the desired IF frequency. These spurious responses can be characterized [4] by the equation

$$\pm mf_{\mathrm{RF}} \pm nf_{\mathrm{LO}} = \pm f_{\mathrm{IF}} \tag{1-16}$$

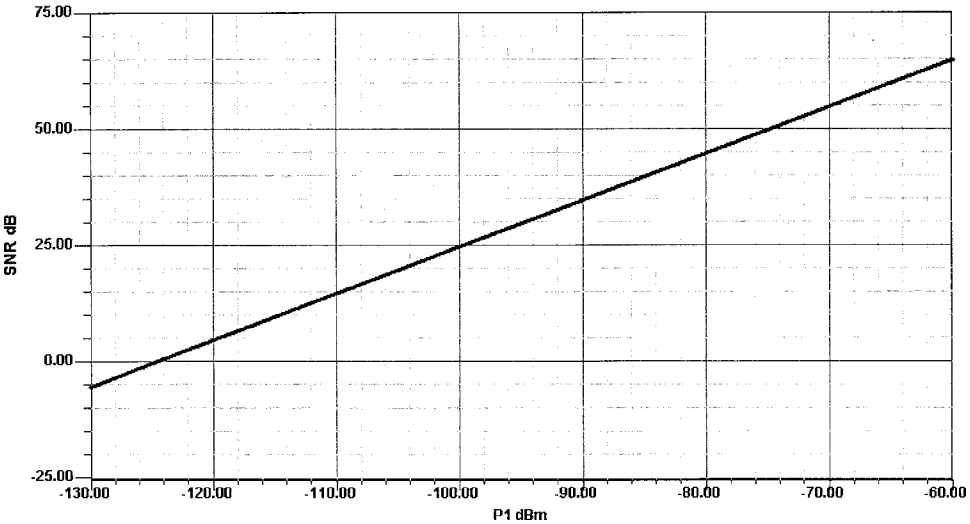


Figure 1-65 C/N ratio versus input RF power level.

Table 1-5 Budget calculation for the ΔS_{21} across each system component accounting for impedance mismatches

Names	ΔS_{21} (dB)
2.4 GHz BPF	–0.82
LNA	22.04
2.4 GHz BPF	–0.93
MIXER_1	–5.34
200 MHz IF BPF	–3.85
200 MHz AMP	22.21
MIXER_2	–8.43
45 MHz IF BPF	–0.42

Some spurious responses, or spurs, can be especially problematic because they may be too close to the intended IF to filter, thus masking the actual information-bearing signal. These spurious responses and their prediction can be especially troublesome if the receiver is operated near saturation. The analysis can then be refined by including a spur table that predicts the spur level relative to the IF signal. The spur's power at the output of a mixer is calculated using the spur table provided for this mixer. Once generated, each spur is carried through the remainder of the system with all mismatches accounted for. The output power level of each spur is shown in Table 1-6. Another useful output is the RF and LO indices, which indicate the origin of each spur.

In addition to sensitivity and spurious response calculations, an analog system can be analyzed in a CAD tool such as Symphony for gain, output power, noise figure, third-order intercept point, intermodulation distortion (IMD; due to multiple carriers), system budget, and dynamic range.

Table 1-6 Spur output powers and the corresponding RF and LO harmonic indices

Frequency (MHz)	P_{out} (dBm)
45.00	–26
20.00	–106
65.00	–86
90.00	–76
110.00	–66
135.00	–96
155.00	–44
175.00	–96
180.00	–106
200.00	–56
220.00	–96
245.00	–76
245.00	–46
265.00	–76

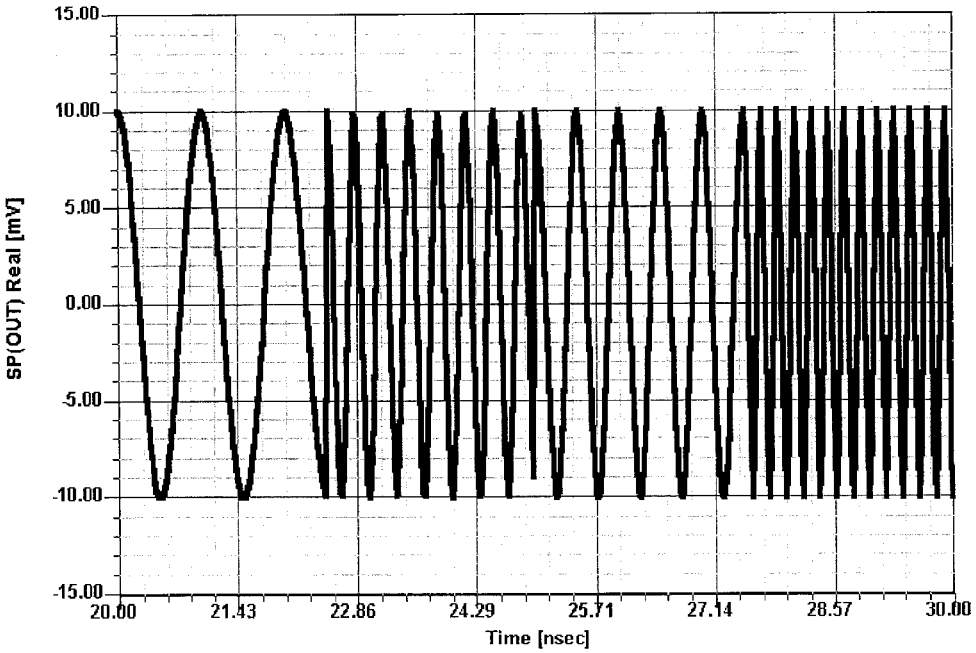


Figure 1-66 4FSK signal generated in Symphony.

Mixed-Mode MFSK Communication System. Next, a mixed-mode (digital and RF) communication system will be described and simulated. In this example digital symbols will be used to modulate an 8-GHz carrier. The system uses *M*-ary frequency shift keying (MFSK) bandpass modulation with a data rate of 40 Mbps. Convolutional coding is employed as a means of forward error correction. The system includes digital signal processing sections as well as RF sections and channel modeling. Several critical system parameters will be examined, including BER. FSK modulation can be described [5] by the equation

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos(\omega_i t + \phi) \tag{1-17}$$

where $i = 1, \ldots, M$. So the frequency term will have M discrete values with almost instantaneous jumps between each frequency value (see Figure 1-66). These rapid jumps between frequencies in an FSK system lead to increased spectral content.

Figure 1-67 shows a block diagram of the complete system. The system can be split into several major functional subsystems—the baseband modulator, the RF transmitter section, channel model, RF receiver, clock recovery circuitry, and baseband demodulator. Looking specifically at the baseband modulator circuitry (Figure 1-68), a pseudorandom bit source is used and the bit rate is set to 40 MHz.

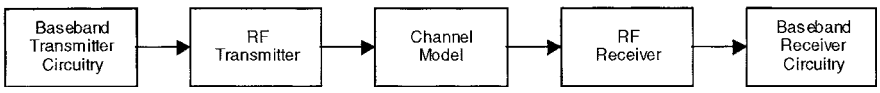


Figure 1-67 Mixed RF/DSP MFSK communication system simulated in Symphony.

A convolutional encoder then produces two coded bits per data bit and increases the bit rate to 80 MHz. The purpose of the convolutional encoder is to add redundancy to improve the received BER. A binary-to-*M*-ary encoder then assigns one symbol to every 2 bits, creating the four levels for the 4FSK, effectively halving the bit rate down to 40 MHz again. The signal is then scaled and upsampled. To decrease the bandwidth, a root-raised cosine filter is used to shape the pulses. The filtered signal then serves as the input to the frequency modulator. The RF section (see Figure 1-69) includes the transmitter that modulates the baseband signal onto an 8-GHz carrier.

That signal is amplified and then filtered to remove any harmonics. The signal is then passed through an additive white Gaussian noise model. The received signal is filtered, amplified, and downconverted twice to baseband. After carrier demodulation the signal is then sent through the clock recovery circuitry. Clock recovery is employed to ensure that sampling of the received signal is executed at the correct instances. This recovered timing information is then used as a clock signal for sample and hold circuitry. Clock recovery in this system was achieved using a phase-locked loop (PLL) configuration. The schematic for the clock recovery circuit is shown in Figure 1-70.

At the heart of the clock recovery circuit is the phase comparator element and the frequency modulator. The inputs to the phase comparator consist of a sample of the received data signal and the output of the feedback path that contains the frequency modulator. The frequency modulator reacts to phase differences in its own carrier and the received data signal. The output of the frequency modulator is fed back into the phase comparator, whose output is dependent on the phase differential at its inputs. The phase of the frequency modulator continually reacts to the output of the phase comparator and eventually lock is achieved. Once the timing of the received signal is locked onto, the clock that feeds the sample and hold will be properly aligned and correct signal sampling will be assured.

Figure 1-71 shows the received signal (filtered and unfiltered) as well as the recovered clock and the final data. Equalizer circuitry is then used to compensate for channel effects, which degrade the transmitted signal. The equalizer acts to undo or adapt the receiver to the effects of the channel. The equalizer employed in this MFSK system is the recursive least squares equalizer. The equalizer consists of a filter of *N* taps that undergoes an optimization in order to compensate for the channel effects. The equalizer depends on a known training sequence in order to adapt itself to the channel. The equalizer model updates the filter coefficients based on the input signal and the error signal (i.e., the difference between the output of the equalizer and the actual desired output). The update (optimization) is based on the recursive least squares algorithm. Several equalizers are available in Symphony, including complex least mean square equalizer, complex recursive least squares equalizer, least mean square equalizer, recursive least squares equalizer, and the Viterbi equalizer. After equalization, the BER of the system is analyzed versus SNR (see Figure 1-72).

PLL CAD SIMULATION. The Symphony system simulator also allows us to evaluate the performance of phase-locked loops. Figure 1-73 shows the block diagram of a PLL in which the VCO is synchronized against a reference. For the purpose of demonstrating the capability, we have selected a 1:1 loop with a crossover point of about 100 kHz, meaning that the loop gain is 1 at this frequency.

Figure 1-74 shows the system's simulated open- and closed-loop phase-noise performance. At the crossover point, the loop is running out of gain. The reason why the closed-loop noise increases above 2 kHz has to do with the noise contribution of various components of the loop system. The highest improvement occurs at 100 Hz, and as the loop gain decreases,

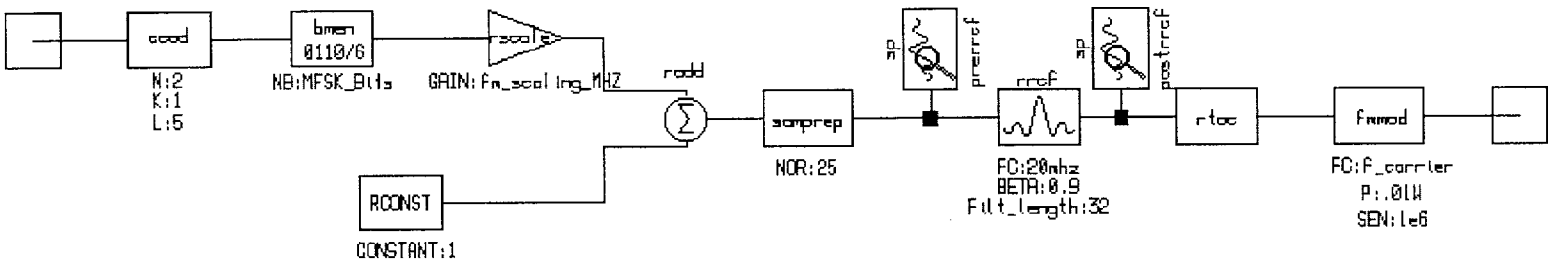


Figure 1-68 MFSK baseband circuitry including frequency modulator.

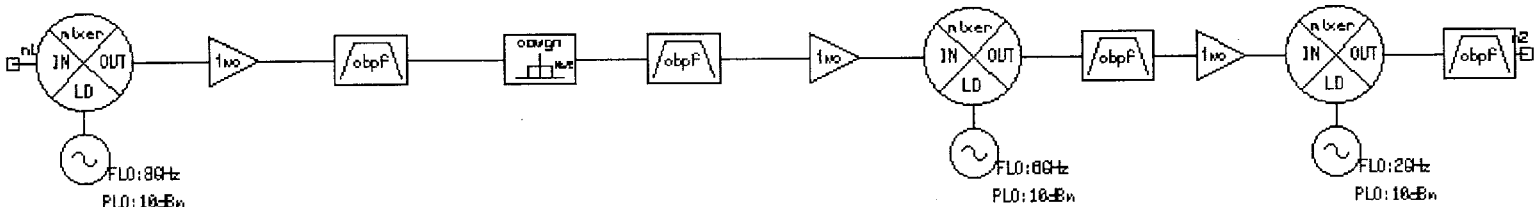


Figure 1-69 RF section including Gaussian noise channel model.

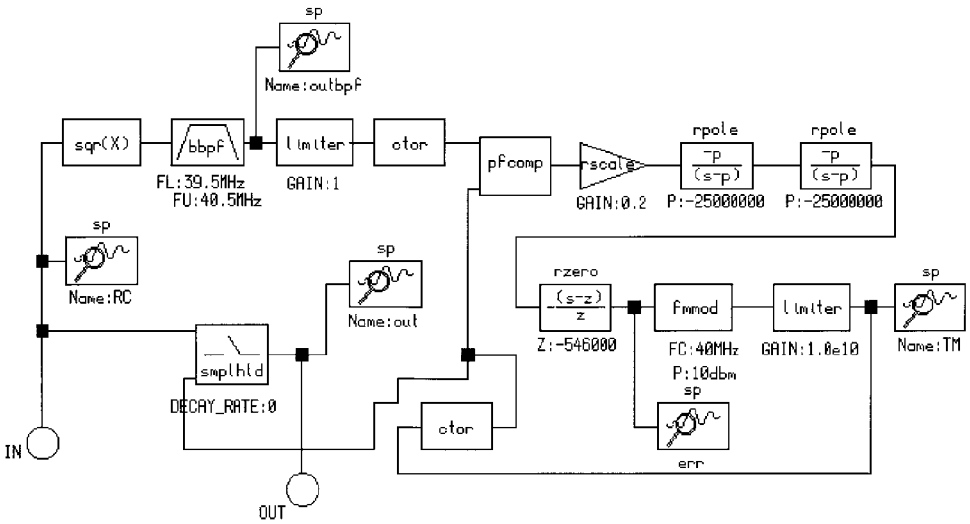


Figure 1-70 Clock recovery circuitry.

the improvement vanishes. Therefore it is desirable to make the loop bandwidth as wide as possible, as this also improves the loop's switching speed. On the other hand, as the phase noise of the free-running oscillator crosses over the phase noise of the reference noise, divider noise, and other noise contributors, one can make the noise actually worse than that of the free-running state. In a single-loop system, a compromise must be made between phase noise, switching speed, and bandwidth. A first-order approximation for switching speed is $2/f_L$, where f_L is the loop bandwidth. Assuming a loop bandwidth of 100 kHz, the switching speed

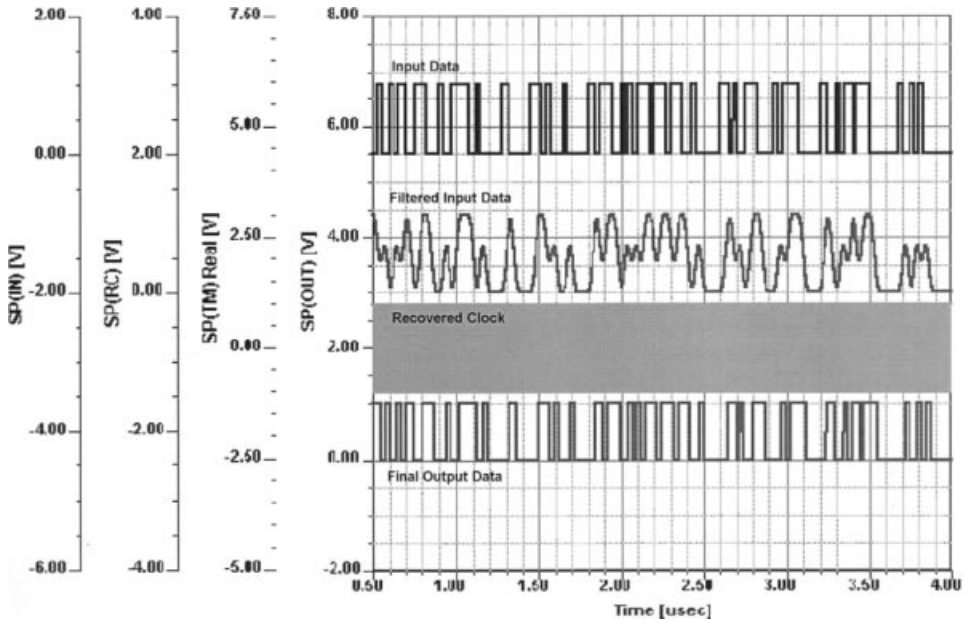


Figure 1-71 Input data, filtered input data, recovered clock, and final output data for the MFSK communication system.

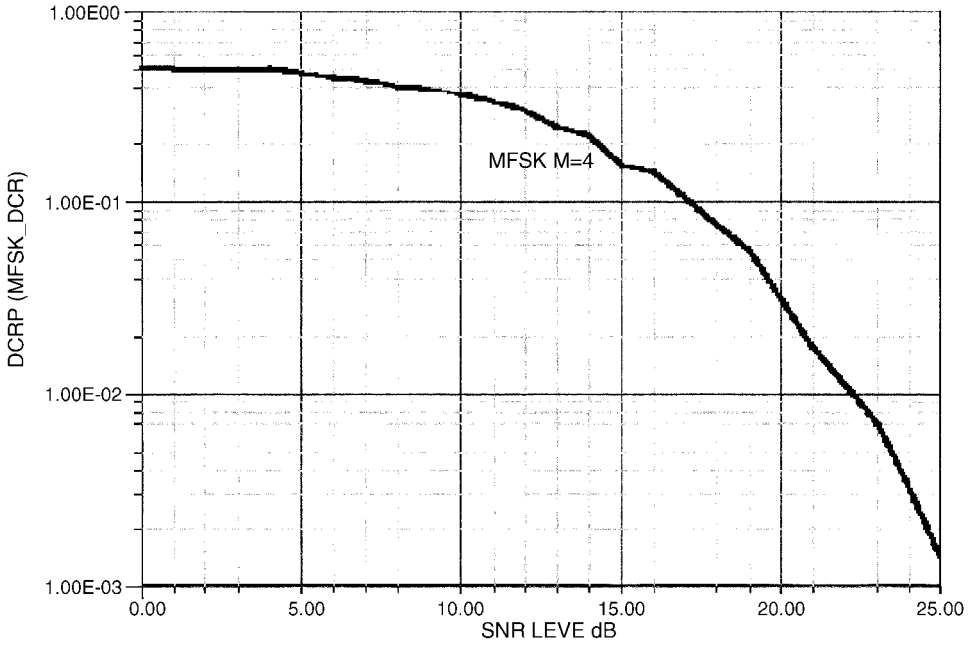


Figure 1-72 BER versus SNR for the MFSK system.

will be 20 μs. Looking at Figure 1-74, we can clearly see the trend from 200 Hz to 100 kHz. Because of the resolution of the sampling time (computation time), the open-loop phase noise below 150 Hz is too low and could be corrected by a straight-line extrapolation from 500 Hz toward 100 Hz. We have left the drawing uncorrected to show the effect of not-quite-adequate resolution.

Since we have already referenced analog and digital waveforms, we need to mention that the current system in place worldwide is referenced as the second-generation (2G) system.

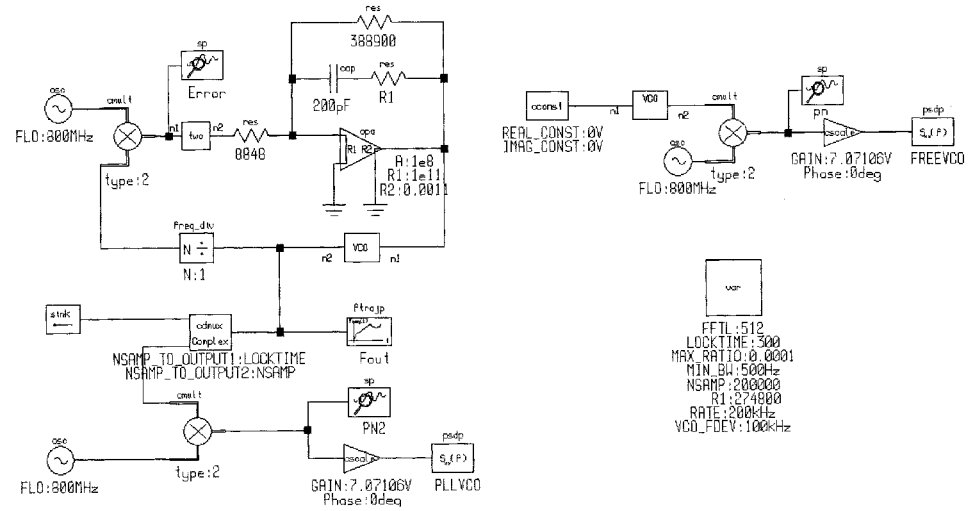


Figure 1-73 Block diagram of a CAD-based PLL system.

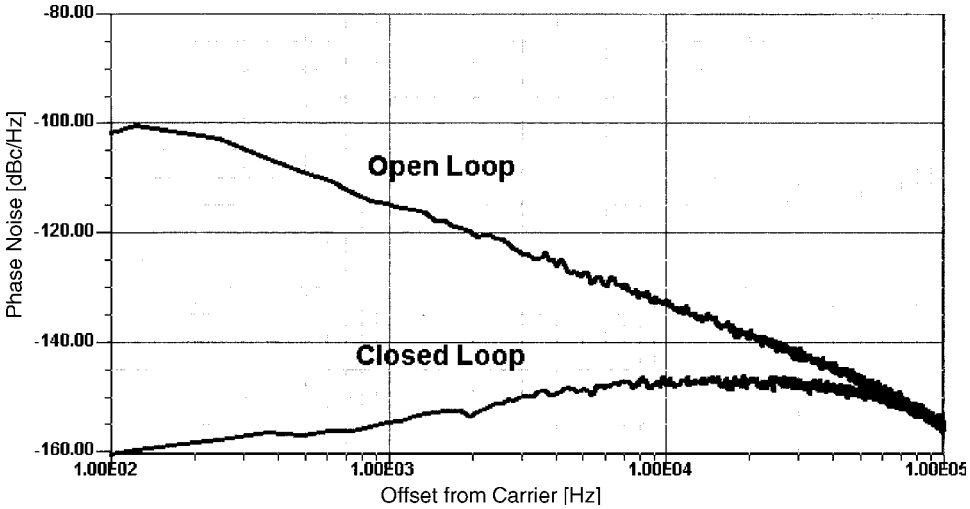


Figure 1-74 Open- and closed-loop phase noise for a CAD-based test phase-locked loop.

Table 1-7 Cellular and cordless standards

Cellular Telephone Networks	
Analog	Digital
AMPS	**IS-54**
Advanced Mobile Phone System	North American Digital Cellular
TACS	**IS-95**
Total Access Communication System	North American Digital Cellular
NMT	**GSM**
Nordic Mobile Telephone	Global System for Mobile Communications
	PDC
	Personal Digital Cellular
Cordless Telephone Networks	
Analog	Digital/PCN
CTO	**CT2/CT2+**
Cordless Telecom 0	Cordless Telecom 2
JCT	**DECT**
Japanese Cordless Telecom	Digital European Cordless Telecom
CT1/CT1+	**PHS**
Cordless Telecom 1	Personal Handy Phone System
	DCS-1800

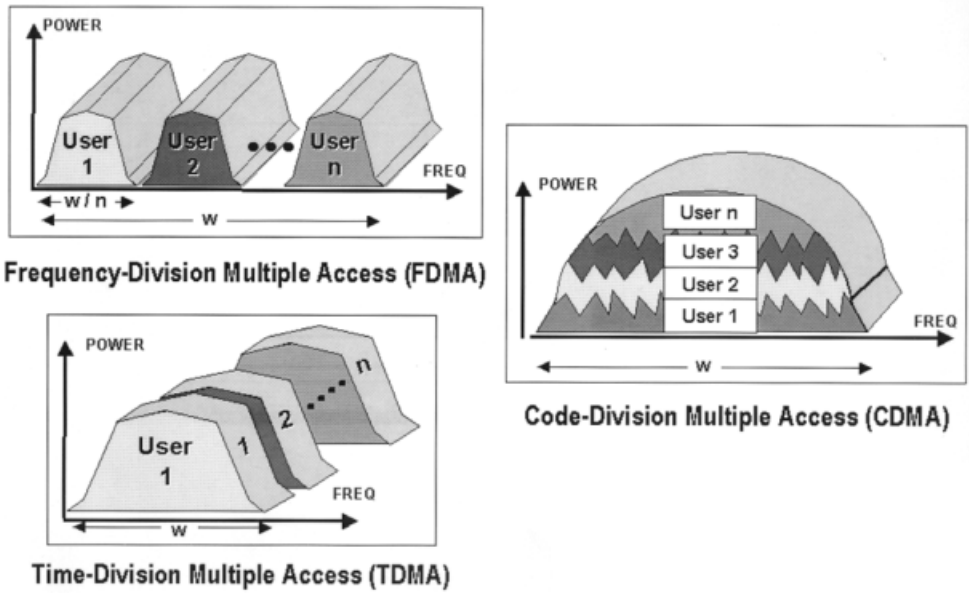


Figure 1-75 Multiplexing schemes.

Table 1-8 Parameters of cellular and cordless systems

	Cellular Telephones			Cordless Telephones		
Characteristics	IS-54	IS-95	GSM	CT2	DCS1800	DECT
Band (downlink)	869-894	869-894	935-960	864-868	1805-1880	1880-1900
Band (uplink)	824-849	824-849	890-915	864-868	1710-1785	1880-1900
Bandwidth	50 MHz	50MHz	50MHz	2 MHz	150 MHz	20 MHz
Channelization	TDMA/ FDMA	CDMA/ FDMA	TDMA/ FDMA	FDMA	TDMA/ FDM	TDMA
Channel spacing	30 kHz	1250 kHz	200 kHz	100 kHz	200 kHz	1728 kHz
Channels/carrier	3	55–62?	8	1	16	12
Number of channels	832 (3 users/ch.)	20 (798 users/ch.)	124 (8 users/ch.)	40	750 (16 users/ch.)	10 (12 users/ch.)
Duplex method	FDD	FDD	FDD	TDD	FDD	TDD
Channel bit rate	48.6 kbps	1.2288 Mbps	271 kbps	72 kbps	271 kbps	1152 kbps
Speech codec	VSELP	CELP	RPE-LTP	ADPCM	RPE-LTP	ADPCM
Bit rate (voice)	8 kbps	1.2–9.6 kbps	13 kbps	32 kbps	13 kbps	32 kbps
Modulation	π/4-DQPSK	BPSK/ OQPSK	GMSK	FSK	GMSK	GMSK
Mobile peak power	0.6–3 W	0.2–2 W	2–20 W	10 mW	0.25–2 W	250 mW
Mobile average power	0.6–3 W	0.2–2 W	0.25–2.5 W	5 mW	0.03–0.25 W	10 mW
Cell radius	30 miles	30 miles	1–5 miles	?	?	40–140 m
Cluster size (min)	7	1	3	N/A	3	N/A

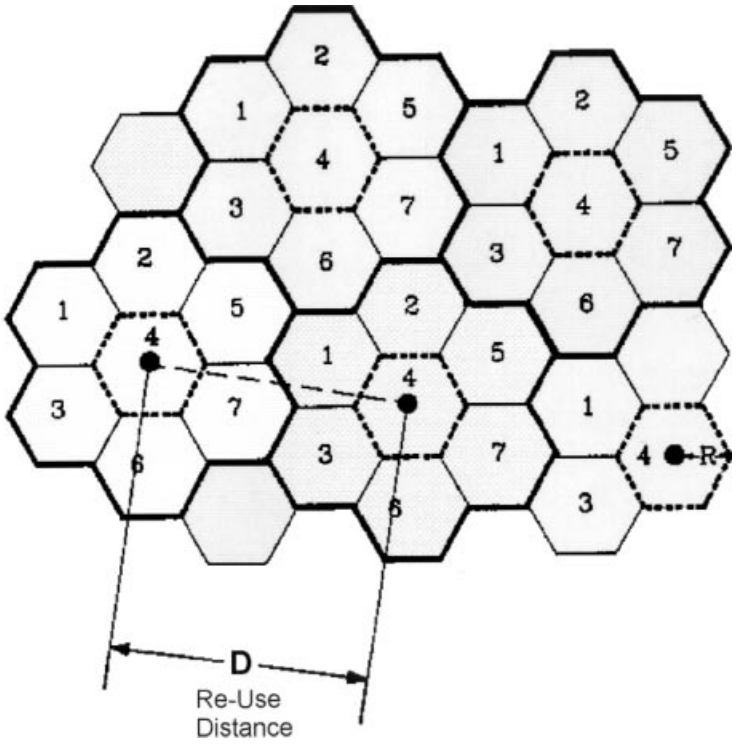


Figure 1-76 Cellular structure. $W = N \times B_C$, where W is the full bandwidth and B_C is the bandwidth per cell. The number of channels/cell = B_C/B_U, where B_U is bandwidth per user (FDMA) or N users (TDMA). $D = R\sqrt{3N}$.

The all-digital, adaptive-bandwidth third-generation (3G) system is being discussed in Europe, Japan, and the United States [6]. This is not without headaches. The Groupe Speciale Mobile (GSM) system in Europe still has potential to grow and is not under any urgent pressure to be upgraded. The opposite is true in Japan, because Japan is running out of capacity and needs to go to wideband CDMA. The important perceived advantage of this is that the number of simultaneous users can be extremely high: there is much less interference because of orthogonal modulation types with extremely low correlation and adaptive bandwidth. The actual definitions are still in flux, with the discussions dominated by the United States. This domination stems from the National Security Agency being determined to be able to eavesdrop on conversations that are none of their business (drug smuggling is always used as an excuse) but invites invasion of monitoring without proper court authorization. This is a very hot issue. The other similar problem has to do with patents. This type of technology is somewhat dominated by Qualcomm, a U.S. company, which likes to send a bill for patent infringement to everybody, possibly even to users. A similar speculation took place with the introduction of digital audio tapes and their duplication and is occurring nowadays with CDs carrying either music recordings or movies. This issue is driven by money and not by technology.

That said, Table 1-7 shows cellular and cordless standards, everything referring to 2G. Figure 1-75 shows the multiplexing schemes used. Table 1-8 shows the parameters of cellular and cordless systems. Figure 1-76 shows the cellular structure.

1-5-2 Transmitters

Since it is possible to generate any type of modulation using DSP, including those described in this chapter, a DSP-based transmitter can also be built. A good example of this is the Philips SA900, which is truly universal.

Introduction to the SA900.* The SA900 (Figure 1-77) is a truly universal in-phase and quadrature (I/Q) radio transmitter that can perform many types of analog and digital modulation including AM, FM, SSB, QAM, BPSK, QPSK, and FSK. It is a highly integrated system that saves space and cost for the manufacturers producing cellular and wireless products. The device allows baseband signals to directly modulate the I/Q carriers, which are generated by an internal phase shift network, in the 1-GHz range, and to maintain good linearity required for a linear modulation scheme (e.g., $\pi/4$-DQPSK). It contains an on-chip frequency divider, phase detector, and VCO, which can be built into a phase-locked loop (PLL) frequency synthesizer to create a transmit offset frequency. Its unique internal design allows frequency conversion without having an external image rejection filter for eliminating the sum term after mixing. The SA900 meets the specifications required by the IS-54, the industry standard for North America Digital Cellular (NADC) system. Here we review the basic concept of I/Q modulation and discuss the key points when designing the SA900 for an RF transmitter.

I/Q Modulation. Any bandpass RF signals can be represented in polar form by

$$s(t) = A(t)\cos[\omega_c t + \phi(t)] \tag{1-18}$$

where $A(t)$ is the signal envelope and $\phi(t)$ is the phase. By using the trigonometric identities, we can represent Eq. (1-18) in rectangular form by

$$s(t) = \mathrm{I}(t)\cos(\omega_c t) - \mathrm{Q}(t)\sin(\omega_c t) \tag{1-19}$$

where $\mathrm{I}(t) = A(t)\cos[\phi(t)]$
$\mathrm{Q}(t) = A(t)\sin[\phi(t)]$

Since the baseband signals $\mathrm{I}(t)$ and $\mathrm{Q}(t)$ modulate two exactly 90° out-of-phase carriers $\cos(\omega_c t)$ and $-\sin(\omega_c t)$, respectively, we call the system implementing Eq. (1-19) an in-phase and quadrature (I/Q) modulator. Figure 1-78 shows the mathematics and hardware implementation of an I/Q modulator.

The local oscillator, usually a VCO within a PLL, generates the carrier and is split into two equal signals. One goes directly into a double-balanced mixer to form the I-channel and the other one goes into the other mixer via a 90° phase shifter (realized by passive elements) to provide the Q-channel. The baseband signals $\mathrm{I}(t)$ and $\mathrm{Q}(t)$, either analog or digital in nature, modulate the carrier to produce the I and Q components, which are finally combined to form the desired RF transmitting signal. Since any RF signal can be represented in the I/Q form, any modulation scheme can be implemented by an I/Q modulator.

*Based on portions of the Philips Semiconductors/Signetics RF Communications Products Application Note AN1892, "SA900 I/Q Transmit Modulator for 1 GHz Applications," August 20, 1997. Used with permission.

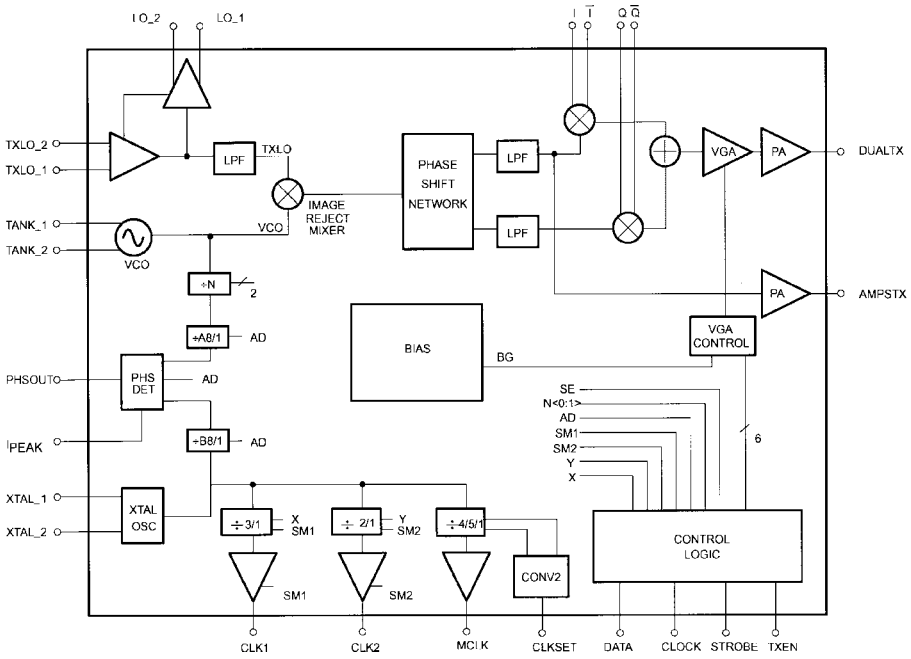


Figure 1-77 SA900 transmit modulator.

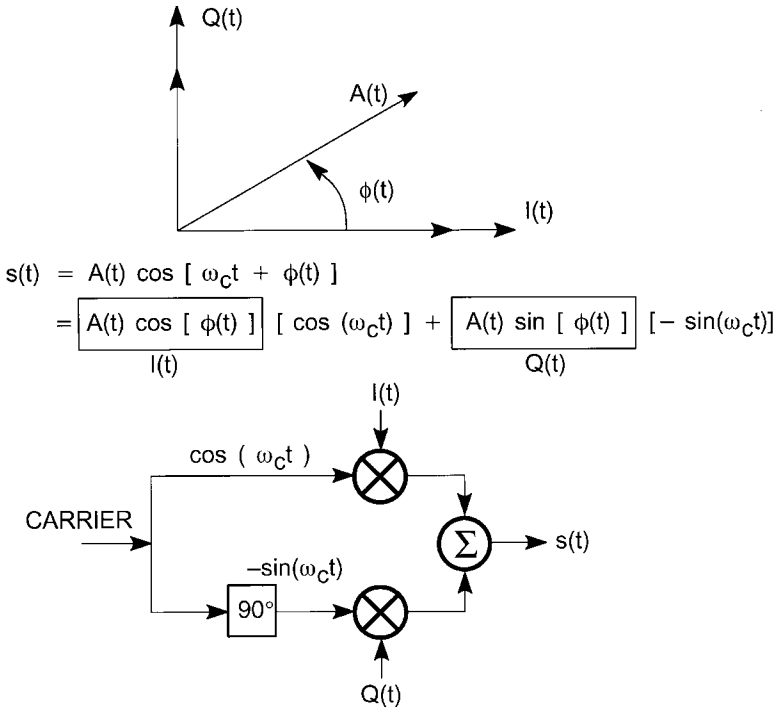


Figure 1-78 Mathematical representation and hardware implementation of I/Q modulator.

Linear Digital Modulation. Linear digital modulation techniques depend on varying the phase and/or magnitude of an analog carrier according to digital information: ones and zeros. This digital information can be the output of an analog-to-digital converter (e.g., voice codec), or it can be digital data in some standard formats (e.g., ASCII). The most popular digital signaling format is nonreturn-to-zero (NRZ), where 1's and 0's are converted into signals with amplitudes of 1 and –1, respectively, in a symbol duration. Since an NRZ signal has infinite bandwidth, transmit filters have to be used to limit the spectral spreading. To ensure each NRZ symbol does not smear into its neighbors due to low-pass filtering and channel distortion, causing intersymbol interference (ISI), the frequency response of the low-pass filter has to satisfy Nyquist criteria. One example of this type of filter is the linear phase square-root-raised cosine filter. Together with the same type of filter for receive low-pass filtering, the signal is guaranteed ISI-free in a Gaussian environment. One straightforward technique of transmitting such a band-limited signals through communication channels would be to apply it directly to the mixer of the I channel to generate an RF signal. This is known as binary phase shift keying (BPSK), where the phase of the carrier is shifted 180(to transmit a data change from 0 to 1 or 1 to 0.

Quadrature or quaternary phase shift keying (QPSK) is a much more common type of modulation scheme used in mobile and satellite communications. It has four possible states (90° apart) and each of them represents 2 bits of data. Figure 1-79 shows the baseband generator for QPSK (without the differential phase encoder). NRZ data bits go through the serial-to-parallel converter (see Figure 1-80) and are mapped in accordance to some rules to generate I and Q values. The generic rule states that the values of I and Q components are 1 and 1 for the data bits "11" (45°) and –1 and –1 for the data bits "00" (–135°). These discrete signals have to be band-limited by Nyquist low-pass filters to be ISI-free.

A more sophisticated way of mapping results in π/4-DQPSK (D for differential encoding), which is chosen for North America Digital Cellular (IS-54), Personal Digital Cellular (PDC) in Japan, and Personal Handy Phone System (PHS) in Japan. In this scheme, consecutive pairs of bits are encoded into one of the four possible phases: π/4 for "11," 3π/4 for "01," –3π/4 for "00," and –π/4 for "10." However, unlike the previous case that "11" is always π/4 and "00" is always –3π/4, the encoded phases are the degrees that the carrier has to shift at each sampling instances. Thus the information is contained in the phase difference (differential) instead of absolute phase for π/4-DQPSK.

A better way to tell the difference between QPSK and π/4-DQPSK is by looking at the signal constellation diagram, shown in Figure 1-81, which displays the possible values of I and Q vectors and changes of state. A constellation diagram is also known as a phase diagram because it shows the phase of the carrier at the sampling point. Note that the phases of QPSK are assigned for every 2 bits of data; therefore it can transmit twice as much information as BPSK in a given bandwidth; that is, it is more bandwidth-efficient. The 8-PSK is another type of modulation used for high-efficiency requirements. It maps 3 bits into eight phases, 45° apart, in the constellation. More spectrally efficient modulation can be created by

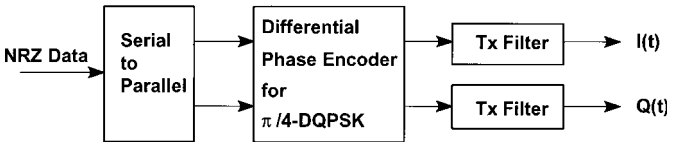


Figure 1-79 QPSK and π/4-DQPSK baseband generator.

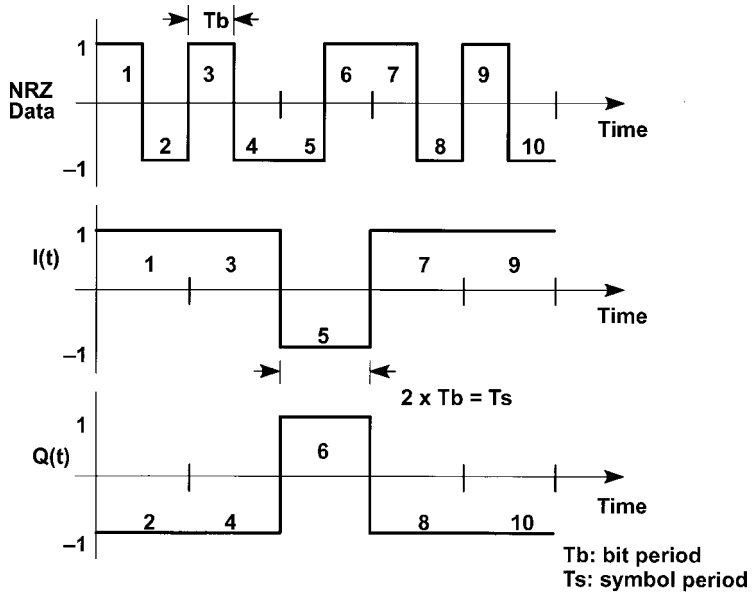


Figure 1-80 Serial-to-parallel conversion.

mapping more bits into one phase at each sampling point. However, as we add states to the signal constellation, the signal susceptibility to noise increases because the difference between states decreases. Then, a higher carrier-to-noise (C/N) ratio is required to maintain the same bit error rate.

One common misconception is that since π/4-DQPSK has eight states, it is just another type of 8-PSK. Note, however, that at every sampling instant, the carrier of π/4-DQPSK is allowed to switch to only one of the four possible states (see Figure 1-81). So we still have two data bits that get encoded into four phases. Thus π/4-DQPSK has the same spectral efficiency as QPSK for the same carrier power. The reason for using this modulation scheme is twofold. First, the envelope fluctuation, which causes spectral spreading due to nonlinearity of transmitter and amplifier, is reduced because the maximum phase shift is 135° instead

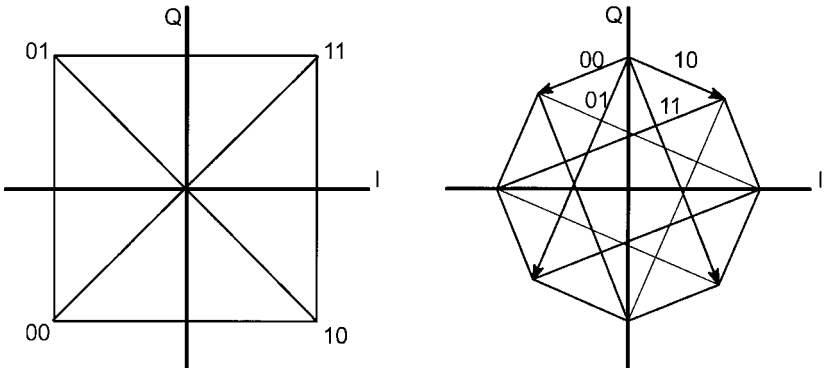


Figure 1-81 Signal constellation of QPSK and π/4-DQPSK.

of 180°. Second, the signal can be demodulated noncoherently, which simplifies the receiver circuitry by eliminating the need for carrier recovery.

Digital and Analog FM. Another family of digital modulation is categorized by frequency change of the carrier instead of phase and/or amplitude change. One member of this family is frequency shift keying (FSK), where the carrier switches between two frequencies. FSK is also known as digital FM because it can be generated by feeding the NRZ data stream into an analog VCO. FSK appears as a unit circle in the signal constellation because the RF signal envelope is constant and the phase is continuous. Baseband filtering is usually applied for FSK to limit the RF bandwidth of the signal so that more channels can fit into a given frequency band.

One common modulation of this type is known as Gaussian minimum shift keying (GMSK), which is used for GSM and some other wireless applications. GMSK can be generated by following its definition: band-limit the NRZ data stream by a Gaussian low-pass filter, then modulate a VCO with modulation index (2 × frequency deviation/bit rate) set to 0.5. In other words, the single-sided frequency deviation is one-fourth of the bit rate ($\Delta f = R/4$).

Another way of generating GMSK is by I/Q modulator. Referring back to Eq. (1-19), any RF signal can be split into I and Q components. Unlike the QPSK mentioned before, baseband I(t) and Q(t) are not discrete points for FM signals; rather, they are continuous functions of time. The way to produce FM is shown in Figure 1-82. We first store all the possible values of $\cos[\phi(t)]$ and $\sin[\phi(t)]$ in a ROM lookup table, which will be addressed by the incoming data to generate the I and Q samples. The output data from the ROM is then applied to D/A converters, after low-pass filtering for signal smoothing, to produce the analog baseband I and Q signals. This method guarantees the modulation index to be exactly 0.5, which is required for coherent detection of GMSK (e.g., the GSM system). The same I/Q principle can also be applied to generating analog FM signals.

Single-Sideband AM (SSB AM). AM signals can be divided into three types: the conventional AM, double sideband suppressed carrier AM (DSBSC AM), and SSBSC AM. The first type is not attractive because for 100% modulation, two-thirds of the transmit signal power appears in the carrier, which itself conveys none of the information added by modulation. By using a balanced mixer (e.g., a Gilbert cell), one can generate DSBSC AM, where the carrier is totally suppressed and only the upper and lower sidebands are present. However, this is still not the best because the information is transmitted twice, once in each sideband. To further increase the efficiency of transmission, only one sideband is needed to deliver the information. The SSBSC AM can be generated by an I/Q modulator with the baseband information feeding the modulator (by quadrature), as shown in Figure 1-83. This modulation technique can greatly reduce the bandwidth of the signal and allows more signals to be transmitted in a given frequency band. This topic is discussed in detail in [7].

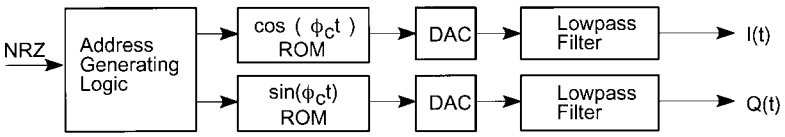


Figure 1-82 Digital FM (e.g., GMSK) baseband I/Q generator.

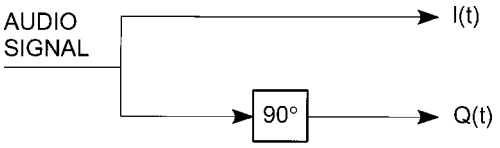


Figure 1-83 Baseband processing for SSBSC AM.

System Architecture. There are usually two schemes, the dual conversion and direct conversion, used for implementing transmit modulators. Dual conversion is simpler to implement by modulating an oscillator at lower frequency and then upconverting to the carrier frequency. This scheme, however, is more expensive due to the need for additional filtering and more printed circuit (PC) board space. By using only one mixer, direct conversion requires fewer components but is harder to implement.

The problems that direct conversion suffers are carrier leakage and modulated signal coupling. Poor RF isolation of the surface mount packages will allow the carrier to be present at the transmitter output, thus making it difficult to have –40 dBc carrier suppression. In addition, modulated RF signal would couple back to the oscillator (usually a VCO in a PLL synthesizer loop) and cause modulation distortion.

Based on the concept of dual conversion, the SA900 uses an image-reject mixer to eliminate the need for IF filtering and allow monolithic integration. The transmit carrier (LO) is downconverted by the frequency synthesized by the on-chip VCO, which operates from 90 to 140 MHz. This LO is then modulated by the baseband I/Q signals to obtain a complex modulation scheme. The image (sum term) after mixing and LO are sufficiently suppressed by the image-reject mixer. Any residual amounts can be further suppressed by an external duplex filter.

Figures 1-84 and 1-85, respectively, show how the SA900 can be used in frequency-division duplex (FDD) and time-division duplex (TDD) transceivers. Note that the LO for both systems is running at a frequency that is higher than the transmit frequency, thus minimizing carrier leakage. In the FDD system only one external VCO is required for generating both transmit and receive LO signals when using the SA900.

Figure 1-86 shows the IS-54 front-end chipset that consists of the SA601, SA7025, SA900, and SA637. This receiver architecture (SA637) supports a digital magnitude/phase baseband

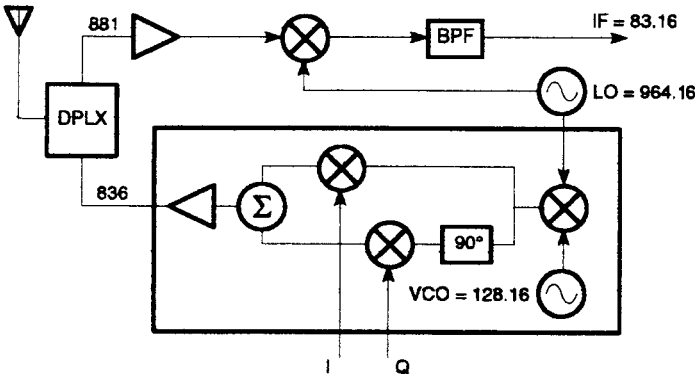


Figure 1-84 FDD system using SA900.

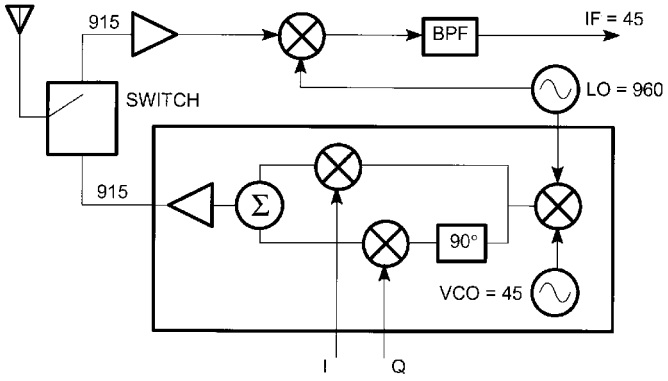


Figure 1-85 TDD system using SA900.

demodulator. An alternate configuration will be using the SA606 FM/IF receiver in conjunction with an external I/Q demodulator IC. Table 1-9 shows the possible configurations for the IS-54 handsets using the SA900 as transmitters.

Designing with the SA900

Baseband I/Q Inputs. The baseband modulation inputs are designed to be driven differentially for the SA900 to operate at its best. The I and Q inputs should have a dc offset of $V_{CC}/2$, which is externally provided by common DSP chips. If all four inputs are biased from the same source, the device can tolerate ±0.5-V dc error; however, inaccuracy of dc bias between I1/I2 or Q1/Q2 causes reduced suppression of the carrier. Thus, it is important to have a well-regulated dc supply for I and Q signal biasing. The bandwidth of the inputs is much higher than the specified 2 MHz. Approximately 2 dB of power loss will be experienced if the I and Q inputs are 50 MHz.

The SA900 generates a minimum of 0 dBm of power to a 50-Ω load when the amplitude of the I and Q signals are 400 mV P-P. The output power will decrease by 6 dB for every 50% decrease in I/Q amplitude. Single-ended I and Q sources can be used but are not recommended due to the degradation in carrier suppression (more than 10 dB compared to differential). In addition, the entire noise performance of the device will suffer. $V_{CC}/2$ should be applied to I2 and Q2 pins if the part is driven single-endedly.

Transmit Local Oscillator. The transmit local oscillator path consists of a TXLO input buffer, LO output buffer, VCO, image-reject mixer, and phase shift network. Together with a few external components, this section provides the I and Q carriers for modulation.

The TXLO inputs and LO outputs are designed to be used in an external PLL that synthesizes different frequencies for channel selection. The RF signal being generated is fed into TXLO inputs and then comes out of LO outputs to complete the system synthesizer loop. The TXLO inputs are differential in nature and have a voltage standing wave ratio (VSWR) of 2:1 with input impedance of 50 Ω. Single-ended sources can be used by ac grounding the TXLO_2 input, as done on the demo board. This signal should also be ac coupled into the TXLO_1 input. The frequency range for these inputs is from 900 to 1040 MHz while the input power should be between −10 and −13 dBm. The output level will be changed significantly if the input level is below −25 dBm.

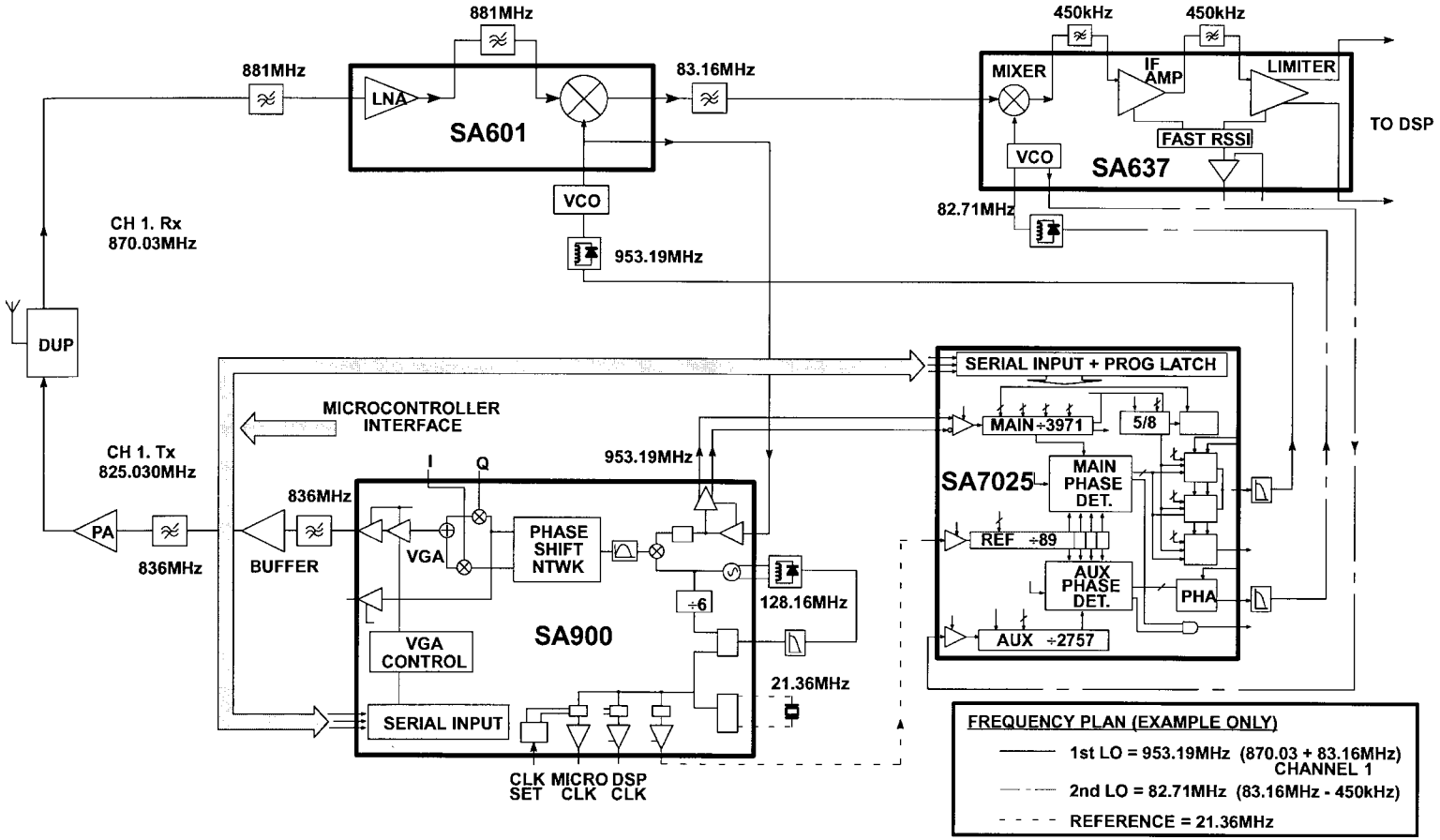


Figure 1-86 IS-54 front-end chipset from Philips.

Table 1-9 Possible configurations for IS-54 handsets

Rx First IF	On-Chip VCO Frequency	On-Chip ÷ *N* Value	Crystal Frequency
83.16 MHz	128.16 MHz	6	21.36 MHz
71.64 MHz	116.64 MHz	6	19.44 MHz
45 MHz	90 MHz	6	15 MHz
84.6 MHz	129.6 MHz	9	14.4 MHz

The output power of the LO buffered signal changes by about 2 dB when the SA900 is in a different mode of operation. Typical values are –13.5 dBm and –15.5 dBm for DUAL mode and STANDBY mode, respectively.

The 90° phase shift network, realized by resistor–capacitor (RC) networks, is capable of operating over a wide frequency range. Even though their frequency characteristics are optimized for cellular band, the part can also be used in other applications in a different band. In such cases, designers have to test the part experimentally to find out the performance, such as sideband suppression, carrier suppression, and image rejection.

Crystal Oscillator. The crystal oscillator (XTAL_1 and XTAL_2 pins) is used to provide a reference frequency between 10 and 45 MHz for the phase detector and the three on-chip clocks. It can be configured as a crystal oscillator using an external crystal and capacitors, or it can be driven by an external source. In the latter case, pin XTAL_2 can be left floating. Information regarding crystal oscillator design can be found in [8].

VCO. The VCO, together with the phase detector, the divider, and an external low-pass filter, can form a PLL for the transmit offset frequency. The image-reject mixer downconverts the TXLO signal to the RF carrier by the amount of VCO frequency. Thus the TXLO frequency should be the desired channel frequency plus the IF offset generated by the VCO. Note that the part will not function if the VCO section is not used.

The VCO is designed for generating IF frequency between 90 and 140 MHz. Together with an external varactor diode and resonator, it can be configured as an oscillator as shown in Figure 1-87.

The resonant frequency of such a circuit is

$$f_{\text{VCO}} = \frac{1}{2\pi\sqrt{LC_T}} \tag{1-20}$$

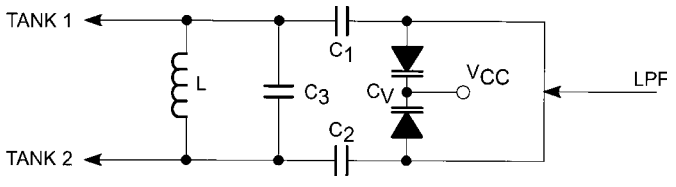


Figure 1-87 VCO tank circuit.

where $C_T = (C_1 \parallel C_2 \parallel C_V) + C_3$. C_V is a varactor diode, which changes capacitance as the voltage across it varies.

Calculation

$$C_1 = C_2 = 33 \text{ pF}$$

$$C_3 = 5.6 \text{ pF}$$

$$C_V = 33.5 \text{ pF @ } 2.5 \text{ V}$$

$$L = 100 \text{ nH}$$

$$C_T = 5.6 + (1/33 + 1/33 + 1/33.5)^{-1} = 16.7 \text{ pF}$$

$$f_{\text{VCO}} = \frac{1}{2\pi(100\text{E} - 9 \times 16.7\text{E} - 12)^{0.5}} = 123\text{MHz}$$

On the demo board, a 1:1 ratio RF transformer is also included to allow single-ended external source driving differential inputs when the VCO is not used.

When designing the VCO, careful printed circuit board (PCB) layout has to be made. Traces have to be short to avoid the parasitic capacitance and inductance that may cause unwanted oscillation. Referring to Eq. (1-20), there are a large number of combinations of L and C_T values that will give the same resonant frequency. If undesired spurs are found in the design due to PCB layout, experimenting with a different set of LC values may sometimes solve the problem.

Output Impedance Matching. The equivalent output impedance at the DUALTX pin is approximately equal to 600 Ω in parallel with 2 pF at 830 MHz. It has to be matched properly to generate maximum power into a 50-Ω load (e.g., a SAW filter). Figure 1-88 shows the recommended matching network. The shunt inductor (L_1) is used to provide maximum swing at the output (short at dc) and also provide reactance to make the real impedance 50 Ω looking

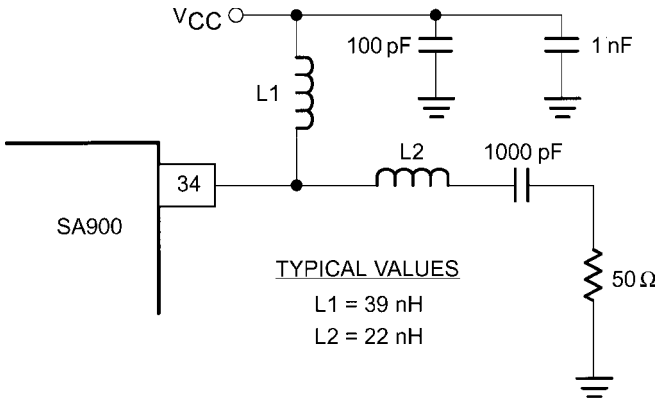


Figure 1-88 DUALTX output matching network.

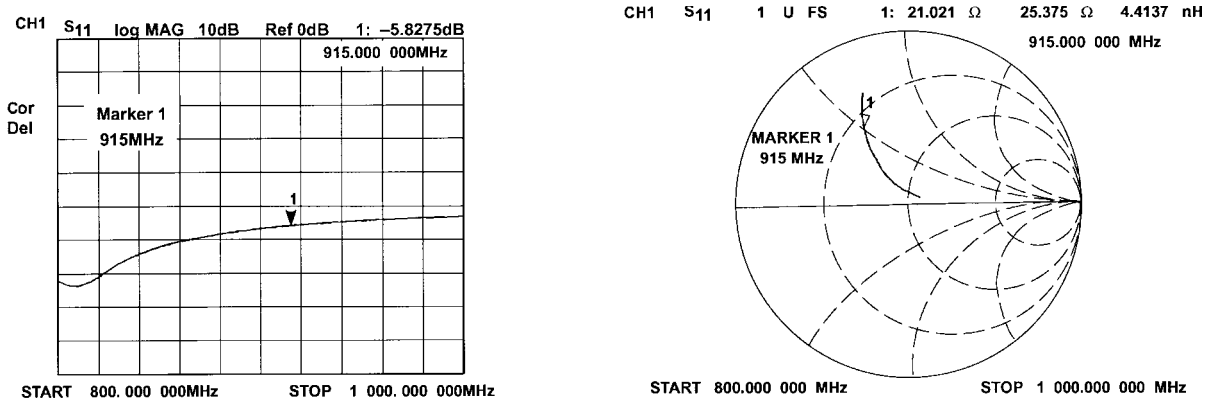
into the matching network. The remaining negative reactance is canceled by the series inductor (L_2). The values used on the demo board can be used as a reference but may not be suitable if a different layout is implemented. The two shunt capacitors are included to bypass the high-frequency RF signal, avoiding direct coupling into V_{CC}. The series ac coupling capacitor is used to maintain the proper bias for the output stage. Their values are big enough to be left out in the impedance matching calculation.

Using a network analyzer to measure the SA900's output *S* characteristic is necessary for obtaining optimum matching, which generates maximum output power. Figure 1-89 shows how to match the output impedance to a 50-Ω load at 915 MHz. First, calibrate the network analyzer to the DUALTX SMA connector on the demo board. Then, short the point where the series inductor is located and use the DELAY feature of the network analyzer to move the point of reference in the Smith Chart to the leftmost point. Now the network analyzer is calibrated to the beginning of the matching network, not just the SMA connector. The frequency response (Figure 1-89a) shows that the "dip" is around 830 MHz, the frequency where the board was originally matched. The Smith Chart shows that it requires less inductance to bring the marker to the center of the chart (50 Ω). By using a 15-nH series inductor, the "dip" was moved closer to 915 MHz (–15 dB) and a better match is achieved (Figure 1-89b).

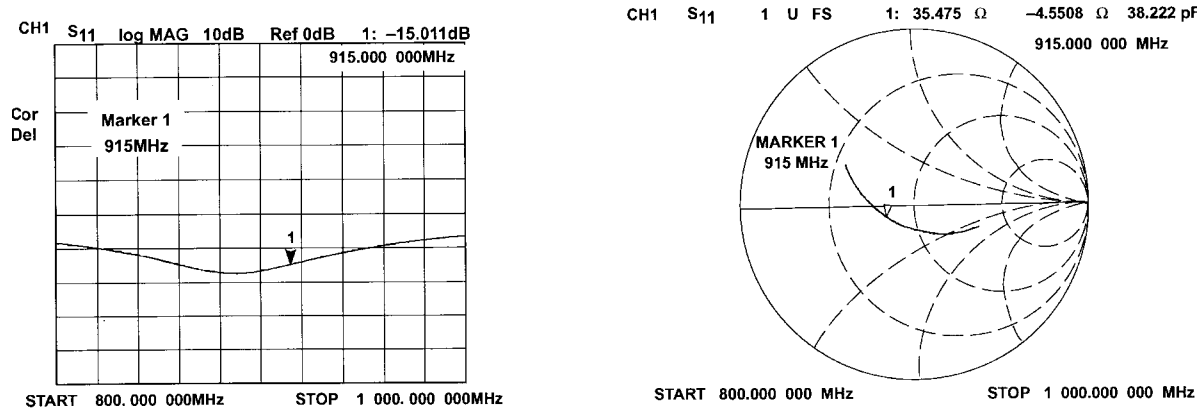
On-Chip Clocks. The crystal oscillator is buffered to provide three external clock signals: CLK1, CLK2, and MCLK. Table 1-10 shows the divide ratio and the controlling mechanism. CLK1 is usually used for the system synthesizer (e.g., SA7025) reference. Since MCLK is active all the time, it is ideal for providing the master clock for the microcontroller. When the device is in STANDBY mode, CLK1 and MCLK provide the clock signals necessary for receiving RF signals. CLK2 can also be used as a clock for the digital signal processing (DSP) chip.

Modes of Operation. The SA900 is intended for either AMPS mode (analog cellular) or DUAL mode (digital cellular, IS-54) operation. When the device is running in AMPS mode, the I/Q modulator, variable gain amplifier (VGA), and phase shifter are disabled. The fixed-gain amplifier is powered up during AMPS mode operation. However, since the divide ratio is too low (6, 7, or 8), the comparison frequency of the on-board PLL is too high, making it very difficult for the loop bandwidth to be less than 300 Hz for analog frequency modulation. The device includes two power-saving modes of operation that disable partial circuitry to reduce the power consumption of the overall chip. The SLEEP mode disables all the circuitry except the master clock (MCLK pin) of the SA900. The STANDBY mode shuts down everything except the TXLO buffer, MCLK, and CLK1, which allows the system synthesizer (e.g., SA7025) to continue running. These two power-saving modes are common to both AMPS and DUAL mode operation. The SA900 draws 60 mA in the DUAL mode, reduced to 3 mA and 8 mA, respectively, in the SLEEP and STANDBY modes.

The TXEN pin is for hardware powering down the modulator and synthesizer. The falling edge of the signal disables the modulator and synthesizer while the rising edge enables the modulator. To power down the synthesizer using software, send a data word with the SE bit set to "0" ("1" for enable). The synthesizer will be disabled right after the strobe signal is transmitted. Either SE or TXEN going low will turn off the synthesizer. This operation is common to both the AMPS and DUAL modes.



a. S_{11} Response and Smith Chart when L2 = 22nH



b. S_{11} Response and Smith Chart when L2 = 15nH

Figure 1-89 SA900 output matching using *S* parameters.

Table 1-10 Possible SA900 clock outputs

CLK1	divide by 3	X (bit 18) = 1
	divide by 1	X (bit 18) = 0
CLK2	divide by 2	Y (bit 19) = 1
	divide by 1	Y (bit 19) = 0
MCLK	divide by 4	CLKSET pin = V_{CC}
	divide by 5	CLKSET pin = $V_{CC}/2$
	divide by 1	CLKSET pin grounded

Performance of the SA900

Performance Criteria. Since the I/Q modulator is a universal transmitter, measuring only the frequency stability and modulation index of a generated FM signal would not be useful for other modulation schemes. Measurement parameters should be general enough so that they can represent the performance of modulators when applying different types of modulation and allow fair comparisons among different I/Q modulators. Based on this idea, two measurement techniques—in-phase modulation and quadrature modulation—are used for evaluating I/Q modulators.

The in-phase modulation relies on injecting two equal frequencies and phase signals at f_{mod} into the I and Q inputs. The result of this modulation is two sidebands appearing at f_{mod} offset from the carrier, with the carrier totally suppressed. This is also known as double-sideband (DSB) conversion. The quadrature modulation requires two equal frequencies (but 90° out-of-phase signals) being injected into the I and Q inputs. The result is a single-sideband suppressed carrier (SSBSC) signal with either the upper or lower sideband at f_{mod} carrier offset being suppressed. This is also known as single-sideband (SSB) upconversion. Figure 1-90 summarizes these two tests.

In a practical system, imperfection of an I/Q modulator is directly related to these two measurements. Sideband and carrier suppression from the quadrature modulation test will show the amount of gain imbalance, phase imbalance, and dc offset. On the other hand, intermodulation product suppression from the in-phase modulation test will show the linearity of an I/Q modulator. When making measurements, it is important to have well-balanced I and Q baseband modulating signals for measurement since the signal imperfection will translate into degraded sideband and carrier suppression.

Performance Graphs. In making those measurements for the demo board, the following parameters were used:

In-Phase Modulation

PIN 43 I1 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 0°
PIN 42 I2 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 180°
PIN 41 Q1 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 0°
PIN 40 Q2 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 180°

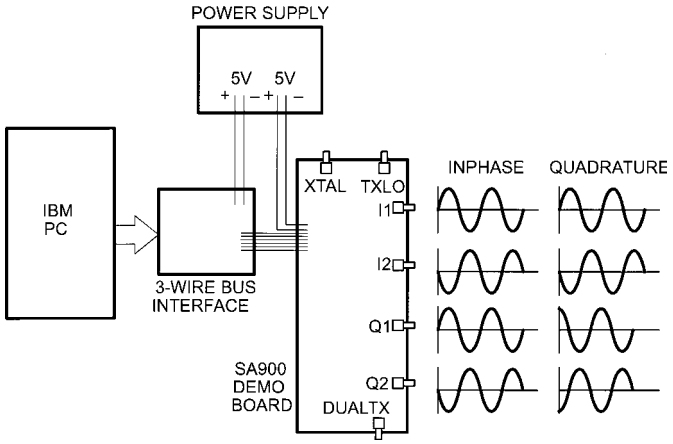


Figure 1-90 In-phase and quadrature modulation test.

Quadrature Modulation

PIN 43 I1 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 0°
PIN 42 I2 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 180°
PIN 41 Q1 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 90°
PIN 40 Q2 = 400 mV P-P, dc = $V_{CC}/2$ at 200 kHz, Phase = 270°

Figure 1-91 illustrates what the typical output spectra would be if in-phase and quadrature modulation were applied to an I/Q modulator. Quadrature modulation will produce lower sideband (LSB) or upper sideband (USB) signal, depending on the phase angle between the I and Q signals. The SA900 was designed to have USB suppressed when the I signal leads the Q signal. The undesired signals are carrier breakthrough and the harmonic products of the baseband modulating signals sitting at $f_c \pm nf_{\text{mod}}$, where n is an integer ≥ 2.

Referring to Figure 1-91a, the output power is 1.3 dBm (cable loss = 0.7 dB) for the LSB while better than 38 dBc of carrier, sideband, and harmonics suppression is measured. The USB suppression of better than 26 dBc implies the residual AM of the transmit signal is better than 5%, a requirement of the IS-54 specification.

The in-phase modulation test will generate both LSB and USB. Beside these two tones, the carrier breakthrough and the harmonics, intermodulation (IM) products will all appear at the output. The odd IM products are dominant, and they satisfy the following rules: Let $f_1 = f_c - f_{\text{mod}}$, and $f_2 = f_c + f_{\text{mod}}$.

$$\text{Third-order IM:} \quad 2f_1 - f_2 = f_c - 3f_{\text{mod}}, \quad 2f_2 - f_1 = f_c + 3f_{\text{mod}}$$

$$\text{Fifth-order IM:} \quad 3f_1 - 2f_2 = f_c - 5f_{\text{mod}}, \quad 3f_2 - 2f_1 = f_c + 5f_{\text{mod}}$$

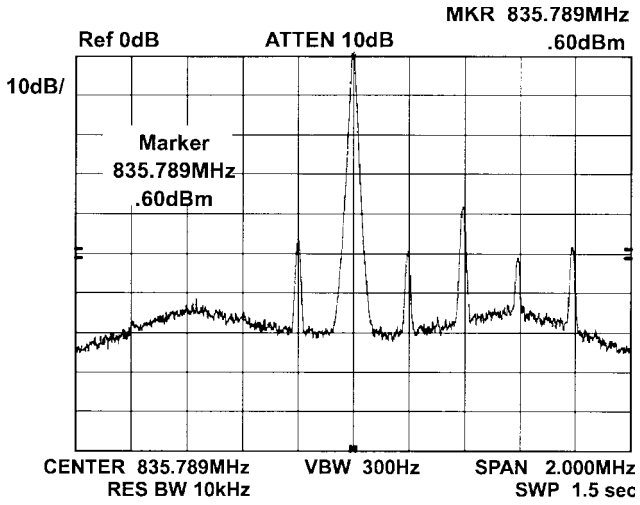
$$\text{Seventh order IM:} \quad 4f_1 - 3f_2 = f_c - 7f_{\text{mod}}, \quad 4f_2 - 3f_1 = f_c + 7f_{\text{mod}}$$

Referring to Figure 1-91b, the LSB and USB are –1.6 dBm (cable loss = 0.7 dB), which is 3 dB less than the measured power for the quadrature modulation test. The IM3 is better than –35 dBc. IM products of much higher orders are totally suppressed.

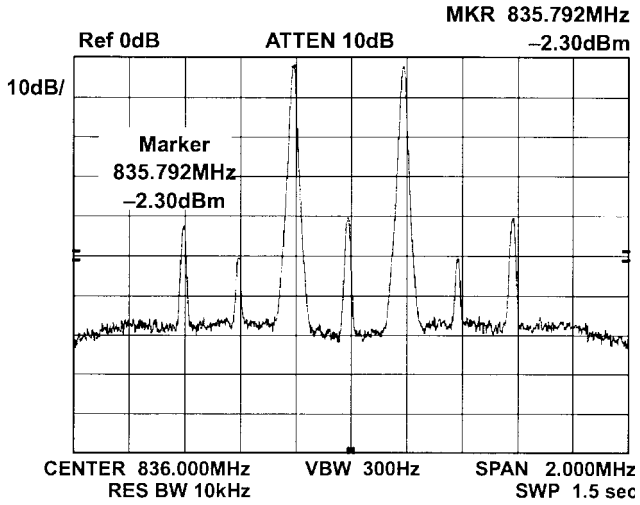
Amplitude and Phase Imbalance. Both amplitude and phase imbalance (error) of an I/Q modulator can be calculated directly from the SSB performance plots. Assuming that phase error equals ϕ radians and amplitude error equals K. The sideband suppression, X, in dBc can be expressed as (see the section entitled "I/Q Modulator Equations" for the derivation):

$$\text{SSB suppression,} \quad X(\text{dBc}) = 10 \log\left(\frac{K^2 + 2 \times K \times \cos(\phi) + 1}{K^2 - 2 \times K \times \cos(\phi) + 1}\right) \tag{1-21}$$

Collecting the like terms and expressing ϕ in terms of K and X, we find



a. Quadrature Test (SSB Up-conversion)



b. In-phase Test (DSB Up-conversion)

Figure 1-91 Typical SA900 output spectra.

$$\phi = \cos^{-1}\left(\frac{10^{X/10} \times K^2 + 10^{X/10} - 1 - K^2}{2 \times K + 2 \times K \times 10^{X/10}}\right) \tag{1-22}$$

For a given X, there will be a set of ϕ and K that satisfies Eq. (1-22). We can represent this relationship graphically, as shown in Figure 1-92. The contours show the phase and amplitude errors for SSB suppression, X, from 44 to 26 dBc. When X equals 40 dBc, phase error is less than 1.2° with an amplitude error of 0 dB. By the same token, the amplitude error is less than 0.2 dB with a 0° phase error.

Spectral Mask. To fully characterize the performance of an I/Q modulator, measurements of the power spectral density of various digital modulation schemes have to be made. Figures 1-93a and 1-93b show the measured spectral masks of IS-54 and PDC standards, which designate π/4-DQPSK as the modulation format.

GMSK is a digital modulation scheme widely used for wireless and mobile communications. Figure 1-93c shows the spectral mask of the modulation format required by GSM, the digital cellular standard in Europe. At carrier offsets of 200 and 300 kHz, the power of the signal is suppressed by 46 and 58 dB, respectively, which is well within the GSM specification.

Power ON Time. The power ON time for the SA900 is mainly determined by the loop bandwidth of the on-board PLL frequency synthesizer. It can be measured by using the HP 53310A modulation domain analyzer set to the EXTERNAL TRIGGERED mode. The STROBE signal from a three-wire bus is used to trigger the equipment. Figure 1-94 shows that the part can be powered up and locked in about 62 μs.

ISM Band Application. In the United States the industrial, scientific, and medical (ISM) band from 902 to 928 MHz is attractive because users are allowed, without having a license, to use transmit powers up to 1 watt in frequency hopping or direct sequence CDMA operation. The wide bandwidth nature of the SA900 fits well into this application. Figure 1-95 shows the output spectra of the SA900, illustrating how well the image-reject mixer works. A common IF (45 MHz) was chosen to be the offset frequency, and then injected externally into the VCO pins. The closest images (± 45 MHz), are down more than 36 dBc.

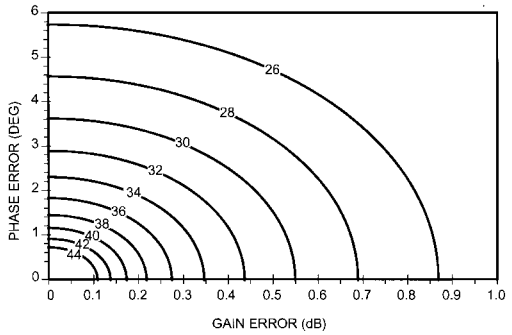
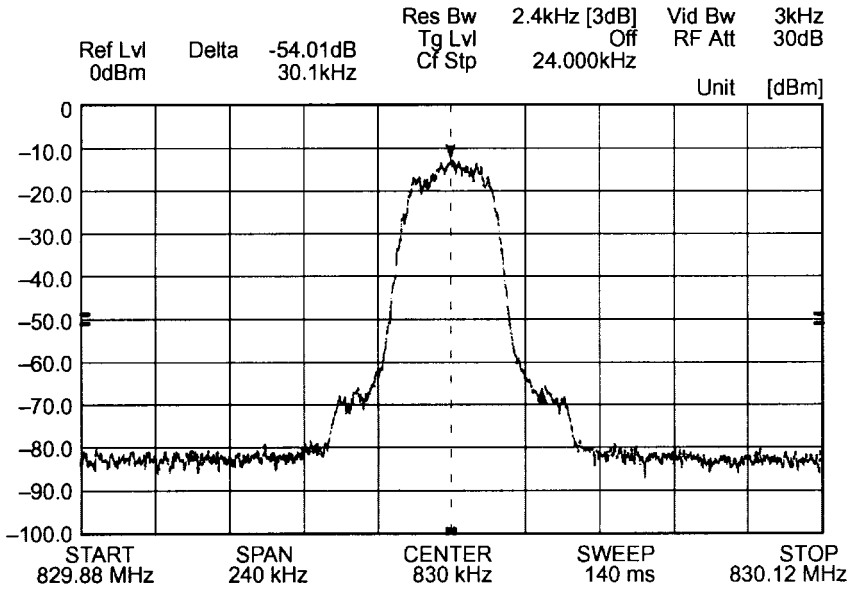
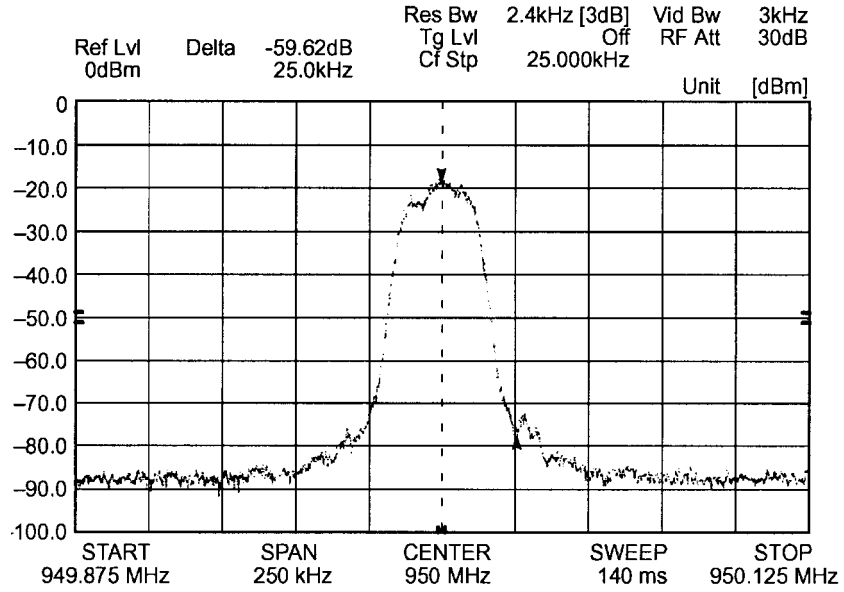


Figure 1-92 SSB suppression contours.

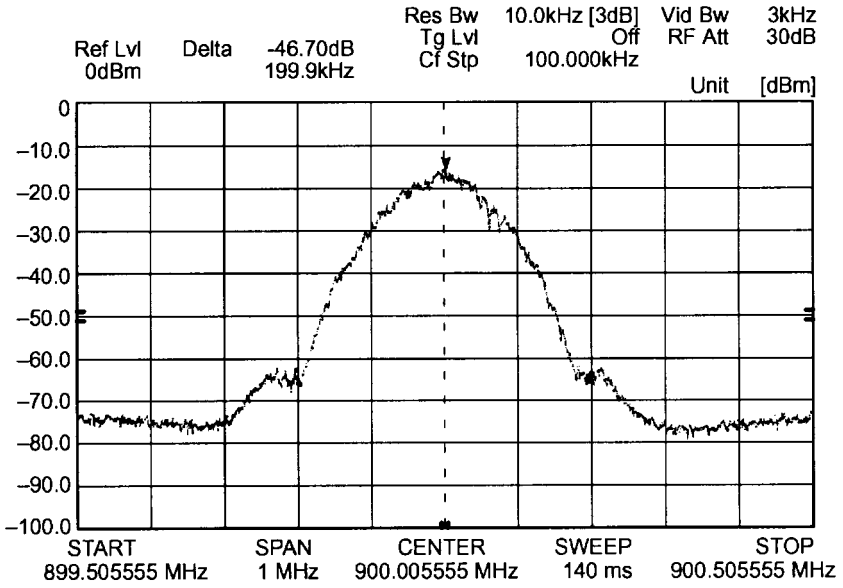


a. Power Spectrum of ADC Modulation
Bit rate = 48.6 kb/s; $\alpha = 0.35$



b. Power Spectrum of PDC Modulation
Bit rate = 42 kb/s; $\alpha = 0.5$

Figure 1-93 SA900 ADC, PDC, and GSM outputs.



c. Power Spectrum of GSM Modulation
Bit rate = 270.833 kb/s; $BT_b = 0.3$

Figure 1-93 Continued

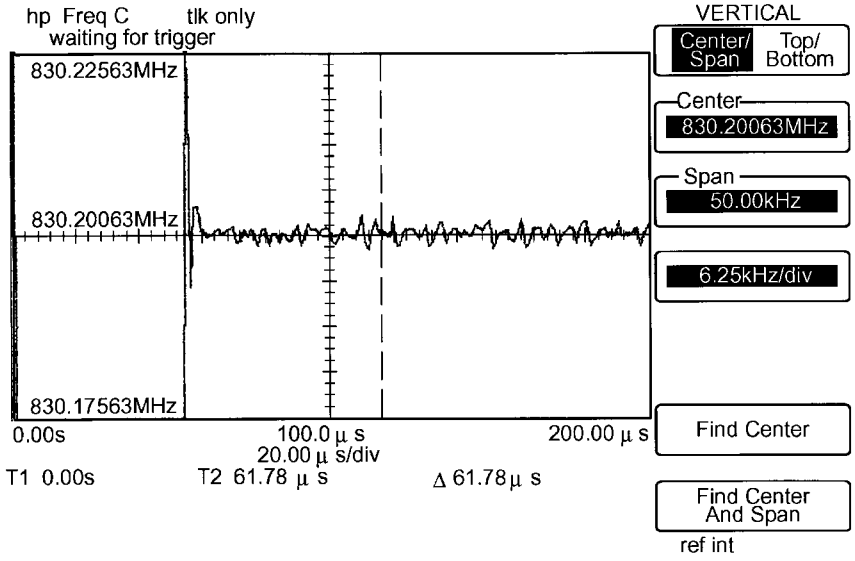


Figure 1-94 Power ON time measurement.

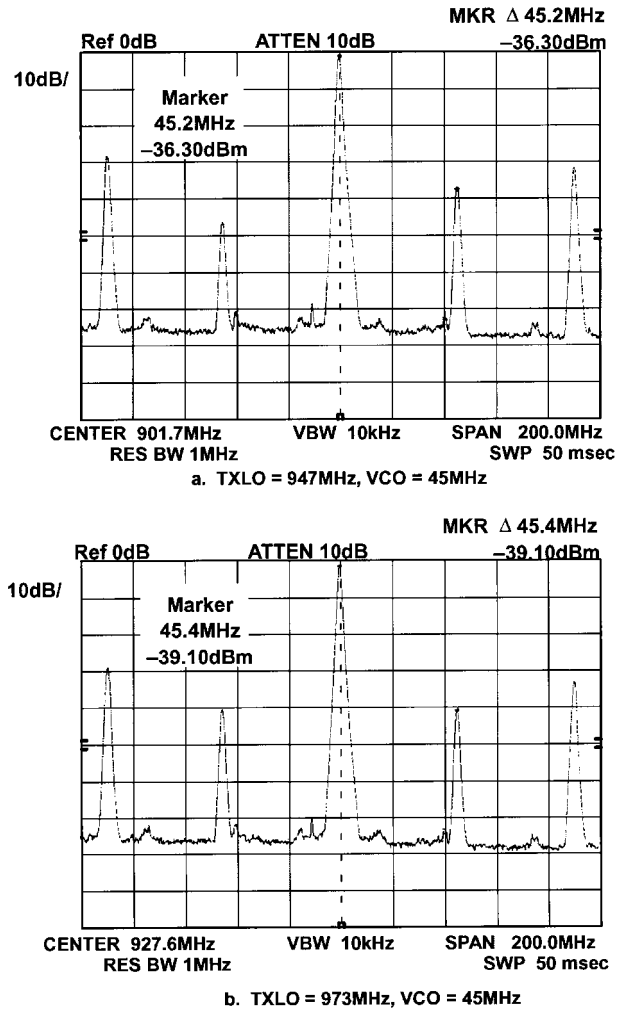


Figure 1-95 Output spectra of the SA900 in the 902-MHz ISM band.

I/Q Modulator Equations. Assume an imperfect I/Q modulator with gain error K and phase error ϕ, modulated by quadrature I/Q signals (SSB upconversion) ω_m. Then the signal, $s(t)$, at the output of the I/Q modulator becomes

$$s(t) = K\cos(\omega_c t + \phi)\cos(\omega_m t) - \sin(\omega_c t)\cos(\omega_m t + 90°) \tag{1-23}$$

Using trigonometric identity and letting $\omega_c - \omega_m = A$ and $\omega_c + \omega_m = B$, we obtain

$$s(t) = \frac{K}{2}\cos(At + \phi) + \frac{K}{2}\cos(Bt + \phi) + \frac{1}{2}\cos(At) - \frac{1}{2}\cos(Bt) \tag{1-24}$$

Assume the information is in the LSB—that is, A—and the spur is the USB—that is, B. We have

$$\text{Signal} = \frac{K}{2}\cos A \cos\phi + \frac{1}{2}\cos A - \frac{K}{2}\sin A \sin\phi \tag{1-25}$$

$$\text{Noise} = \frac{K}{2}\cos B \cos\phi + \frac{1}{2}\cos B - \frac{K}{2}\sin B \sin\phi \tag{1-26}$$

To find the power, we have to evaluate the envelope (amplitude) of these two signals. Recall that for any given bandpass signal in rectangular form, that is, bandpass signal = $X \cos \omega_t - Y \sin \omega_t$, the envelope is

$$\text{Envelope} = (X^2 + Y^2)^{0.5}$$

Therefore, from Eqs. (1-25) and (1-26)

$$\text{Signal} = \left[\left(\frac{K}{2}\cos\phi + \frac{1}{2}\right)^2 + \left(\frac{K}{2}\sin\phi\right)^2\right]^{0.5} \tag{1-27}$$

$$\text{Noise} = \left[\left(\frac{K}{2}\cos\phi - \frac{1}{2}\right)^2 + \left(\frac{K}{2}\sin\phi\right)^2\right]^{0.5} \tag{1-28}$$

Finally, the S/N ratio can be found by taking 10 log the ratio of Eqs. (1-27) and (1-28):

$$S/N = 10\log\left(\frac{K^2 + 2K\cos\phi + 1}{K^2 - 2K\cos\phi + 1}\right) \tag{1-29}$$

1-6 BUILDING BLOCKS

Our earlier block diagram already referred to a variety of integrated circuits that can be used to build a wireless system—in particular, a portable telephone. If we open any of the earlier-mentioned magazines and go through its advertising section, we will find many suppliers of subsystems/components that fit this requirement.

They can be split into various technologies. The following list is a product overview of next-generation advances in wireless technology taken from an ad of Stanford Microdevices, one of the very aggressive component/subsystem companies:

- New SX amplifiers featuring 2.4-GHz MMICs
- GaAs heterojunction bipolar transistor (HBT) discrete transistors (100 mW to 10 W)
- InP/GaAs (indium phosphate/gallium arsenide) high-linearity gain blocks
- Millimeter-wave product line featuring ICs for LMDS
- Silicon–germanium product line (gain blocks and low-noise amplifiers)
- SXQ power modules for cellular and PCS applications
- Power modules featuring 4–25-W power amplifiers for land mobile amplifiers
- Thirty- to 120-W silicon LDMOS power transistors for PCS, CDMA, and W-CDMA
- SAW filters featuring integrated transceiver modules
- Electro-optic and data communications products

More information about Stanford products is available at **http://www.stanfordmicro.com/**. Another interesting company along these lines is Maxim Integrated Products

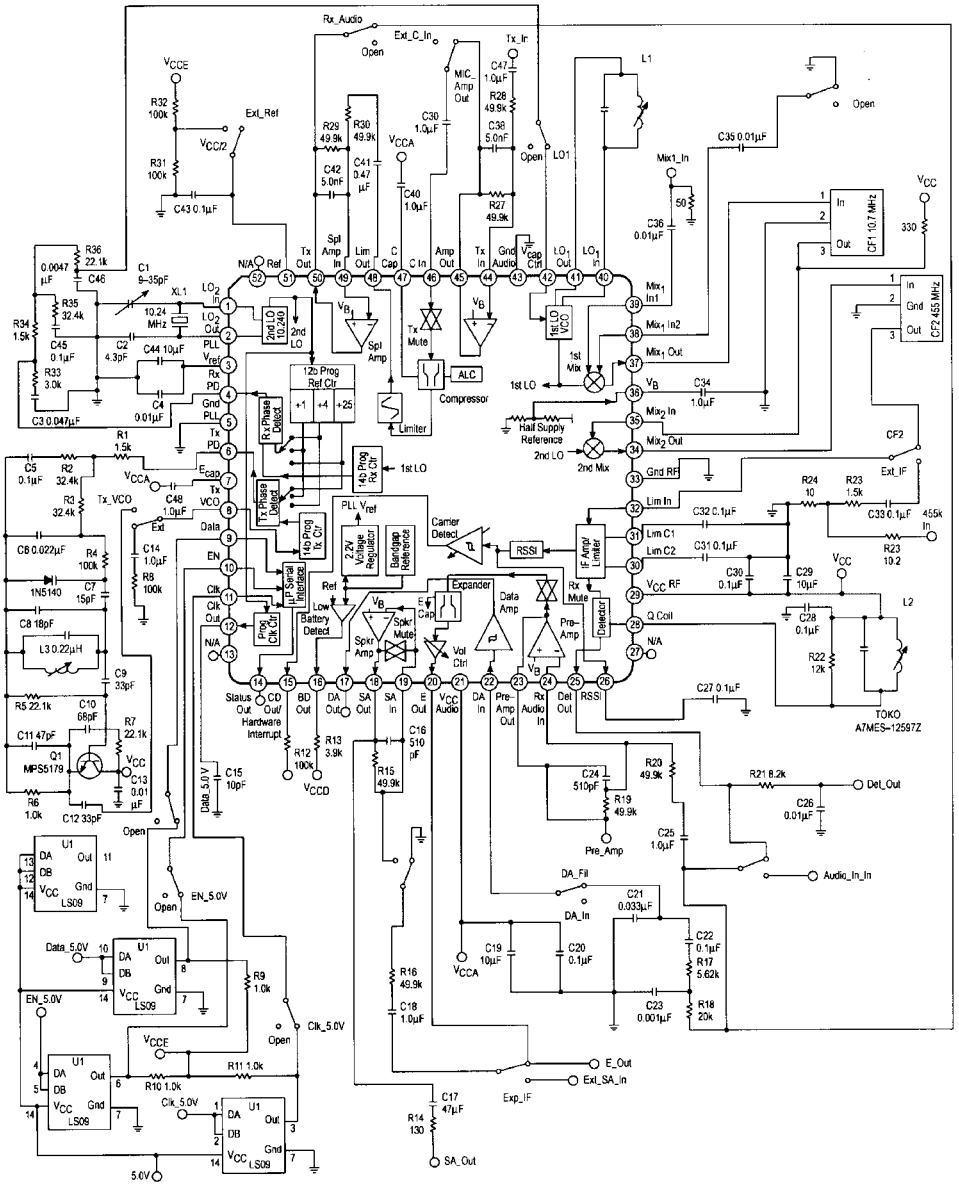


Figure 1-96 MC13109FB universal cordless telephone subsystem IC.

(**http://www.maxim-ic-com/**). They distribute an Analog Design Solutions brochure describing, among other things, the industry's first SiGe RFICs for improved radio front-end performance (1999). On the base-station side, companies such as Hewlett-Packard and Synergy Microwave produce high-performance components/modules as they are needed in the more hostile environment of base-station operation. The offerings in the field are changing very rapidly, and the best way to deal with this is to obtain subscriptions to the magazines mentioned earlier or contact manufacturers directly. Among those not mentioned are Alpha, Ericsson, Siemens (now Infineon Technologies), Motorola, and NEC.

In the beginning of this chapter we looked at block diagrams; here we focus on building blocks. Most manufacturers will try to move to the highest possible integration while avoiding external components. Figure 1-96, a test circuit for Motorola's MC13109FB cordless telephone subsystem IC, shows the level of complexity needed to test such an IC with all its functionality.

Another key issue at those frequencies is to really know a component's frequency dependencies. Figure 1-97 shows a good example. Typical components are capacitors and inductors as reactive elements, and resistors as passive elements. Depending on the type of integration, one may also have to look into the noise performance of integrated resistors. Examination of this issue is best achieved through discussions with the foundries. Along these lines, the dielectric material on which these surface-mounted components are mounted plays a big role and must be addressed individually.

Besser Associates of Los Altos, California, among other firms, offers a very nice one-day short course titled "Wireless Circuit Components: Measurements, Models and Data Extraction."

Table 1-11 presents a sampling of RFICs for wireless applications. These types of ICs, which were taken from Motorola, are made in similar form by many companies. The selection guide is useful in starting off with a design that is less integrated, therefore allowing the designer to have access to portions of the circuit that later "disappear" inside the IC.

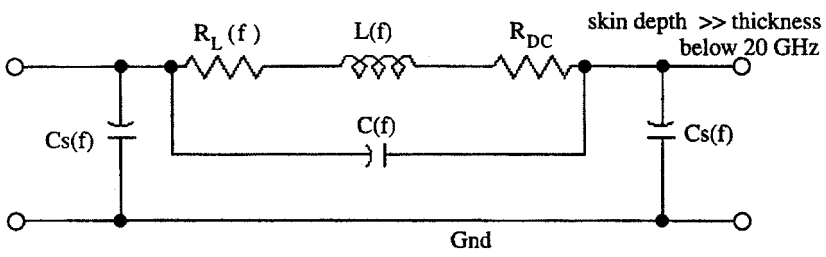


Figure 1-97 Equivalent model of a film resistor above a ground plane.

Table 1-11 Sample RF component/IC selector guide

RFICs

DOWNCONVERTERS

Device	RF Frequency Range (MHz)	Supply Voltage Range (V/dc)	Supply Current (mA) (Typ)	LNA Gain (dB) (Typ)	LNA NF (dB) (Typ)	Mixer Converter Gain (dB) (Typ)	Mixer NF (dB) (Typ)	Package	System Applicability
MC13142	dc to 1800	2.7 to 6.5	13.5	17	1.8	–3.0	12	SO-8	ISM, Cellular, PCS
MRFIC1502	1575	4.5 to 5.5	52	20	—	45	9.5	LQFP-48	GPS
MRFIC1814	1800 to 2000	2.7 to 4.5	10	17	2.5	8.0	10	TSSOP-16	CDS1800, PCS PHS

UPCONVERTERS/EXCITERS

Device	RF Frequency Range (MHz)	Supply Voltage Range (V/dc)	Supply Current (mA) (Typ)	Standby Current (mA) (Typ)	Converter Gain (dB) (Typ)	Output IP3 (dB) (Typ)	Package	System Applicability
MRFIC0954	800 to 1000	2.7 to 5.0	65	5.0	31	28	TSSOP-20EP	CDMA, TDMA, ISM
MRFIC1813	1700 to 2000	2.7 to 4.5	25	0.1	15	11	TSSOP-16	DCS1800, PCS
MRFIC1854	1700 to 2000	2.7 to 5.0	70	5.0	31	23	TSSOP-20EP	CDMA, TDMA, PCS

POWER AMPLIFIERS

Device	Frequency Range (MHz)	Supply Voltage Range (V/dc)	Saturated P_{out} (dBm) (Typ)	PAE (%) (Typ)	Gain P_{out}/P_{in} (dB) (Typ)	Package	System Applicability
MRFIC0913	800 to 1000	2.7 to 7.5	35	50	25	PFP-16	GSM
MRFIC0917	800 to 1000	2.7 to 5.5	34.5	45	22.5	PFP-16	GSM
MRFIC0919	800 to 1000	3.0 to 5.5	35.3	48	32.3	TSSOP-16EP	GSM
MRFIC1805	1500 to 2200	2.7 to 5.0	25	28	20	TSSOP-16	PHS, DECT, PCS
MRFIC1807	1500 to 2200	3.0 to 5.0	26.8	35	8.0	SO-16	DECT, PCS

MRFIC1817	1700 to 2000	2.7 to 5.0	33.5	42	30.5	PFP-16	DCS1800, PCS
MRFIC1818	1700 to 2000	2.7 to 6.0	34.5	42	31.5	PFP-16	DCS1800, PCS
MRFIC1819	1700 to 2000	3.0 to 5.0	33	40	27	TSSOP-16EP	DCS1800, PCS
MRFIC1856	800 to 1000	3.0 to 5.6	32	50	32	TSSOP-20EP	TDMA, CDMA, AMPS
	1700 to 2000		30	35	30		TDMA, CDMA, PCS
MRFIC2006	500 to 1000	1.8 to 4.0	15.5	25	23	SO-8	Cellular, ISM, CT

RF Building Blocks

AMPLIFIERS

Device	RF Frequency Range (MHz)	Supply Voltage Range (V/dc)	Supply Current (mA) (Typ)	Standby Current (μA) (Typ)	Small Signal Gain (dB) (Typ)	Output IP3 (dBm) (Typ)	NF (dB) (Typ)	Package	System Applicability
MC13144	100 to 2000	1.8 to 6.0	8.5	1	17	–5.0	1.4	SO-8	ISM, PCS, Cellular
MRFIC0915	100 to 2500	2.7 to 5.0	2.0	—	16.2	4.0	1.9	SOT-143	ISM, PCS, Cellular
MRFIC0916	100 to 2500	2.7 to 5.0	4.7	—	18.5	11	1.9	SOT-143	ISM, PCS, Cellular
MRFIC0930	800 to 1000	2.7 to 4.5	8.5	20	19	10	1.7	SO-8	GSM, AMPS, ISM
MRFIC1808DM	1700 to 2100	2.7 to 4.5	5.0	8.0	18	13	1.6	Micro-8	DCS1800, PCS
MRFIC1830	1700 to 2100	2.7 to 4.5	9.0	20	18.5	8.5	2.1	Micro-8	DCS1800, PCS
MRFIC1501	1000 to 2000	3.0 to 5.0	5.9	—	18	10	1.1	SO-8	GSP

MIXERS

Device	RF Frequency Range (MHz)	Supply Voltage Range (V/dc)	Supply Current (mA) (Typ)	Standby Current (μA) (Typ)	Converter Gain (dB) (Typ)	Input IP3 (dBm) (Typ)	Package	System Applicability
MC13143	DC to 2400	1.8 to 6.0	4.1	—	–2.6	16	SO-8	ISM, PCS, Cellular

(*continued*)

Table 1-11 Sample RF component/IC selector guide (Continued)

Frequency Synthesis

PLL SYNTHESIZERS

Frequency (MHz)	Supply Voltage (V/dc)	Nominal Supply Current (mA)	Phase Detector	Standby	Interface	Device	Suffix/Case
20 @ 5.0 V	3.0 to 9.0	7.5 @ 5 V	Single-ended three-state, double-ended	No	Parallel	MC145151-2	DW/751F
			Double-ended			MC145152-2	DW/751F
			Single-ended three-state, double-ended		Serial	MC145157-2	DW/751G
						MC145158-2	DW/751G
60 @ 3.0 V	2.5 to 5.5	3 @ 3 V	Two single-ended three-state	Yes		MC145162[a]	P/648, D/751B
85 @ 3.0 V	2.5 to 5.5	3 @ 3 V			Serial	MC145162-1[a]	D/751B
100 @ 3.0 V 185 @ 5.0 V	2.7 to 5.5	2 @ 3 V 6 @ 5 V	Single-ended three-state, double-ended	No		MC145170-2	P/648, D/751B, DT/948C
550, 60	1.8 to 3.6	3	Loop 1 = current source/sink/float Loop 2 = three-state	Yes		MC145181[a]	FTA/873C
1000	2.7 to 5.5	4.25	Current source/sink/float	No	Parallel	MC12181	D/751B
1100	4.5 to 5.5	7 @ 5 V	Current source/sink/float, double-ended	Yes	Serial	MC145191	F/751J, DT/948D
1100	2.7 to 5	6 @ 2.7 V				MC145192	F/751J, DT/948D
1100	2.7 to 5.5	12	Two current source/sink/float, double-ended			MC145220[a]	F/803C, DT/948D
1200, 550	1.8 to 3.6	4	Loop 1 = current source/sink/float Loop 2 = three-state			MC145225[a]	FTA/873C
2000	4.5 to 5.5	12 @ 5 V	Current source/sink, double-ended			MC145201	F/751J, DT/948D
2000	2.7 to 5.5	4 @ 3 V				MC145202	F/751J, DT/948D
2200, 550	1.8 to 3.6	5	Loop 1 = current source/sink/float Loop 2 = three-state			MC145230[a]	FTA/873C
2500	2.7 to 5.5	9.5	Current source/sink/float with dual outputs	No		MC12210	D/751B, DT/948E
2800	4.5 to 5.5	3.5	Current source/sink/float		None	MC12179	D/751

[a]Dual PLL.

1-7 SYSTEM SPECIFICATIONS AND THEIR RELATIONSHIP TO CIRCUIT DESIGN

Wireless communication involves a large range of signal powers—from levels on the order of 10^{-18} watt (at the receiver input) to 10^{2} watts (at a base-station transmitter output). A receiver must be able to demodulate signals that have been attenuated billions of times through propagation; a transmitter must be able to produce a properly modulated signal at a frequency suitable for propagation, at a level high enough to overcome worst-case propagation losses and provide a useful signal at the receiver. Gain is therefore an essential attribute of wireless systems. Because no single active device can provide all the gain required for transmission or reception, we must distribute the gain among multiple stages, designing each for optimum performance across the power span it bridges.

Two inescapable realities impose limits on the gain and absolute power output we may achieve with a given circuit: All real electrical and electronic networks generate noise to some degree, and all real electrical and electronic networks distort the signals applied to them to some degree. A signal weaker than a circuit's inherent noise cannot be amplified by that circuit because it remains indistinguishable from the noise. A signal that exceeds the power-handling capability of the circuit to which it is applied may be degraded, even rendered useless, by the resulting distortion. The following sections examine issues particular to system noise and linearity performance.

1-7-1 System Noise and Noise Floor

Assuming that a system's gain is sufficient, the weakest signal that may be processed satisfactorily, a figure of merit referred to as *noise floor* or (in receivers) *minimum detectable signal* (*MDS*), is limited by thermal noise, assumed to be equal to the noise power available from a resistor at 290 K (about 17 °C or 62 °F), an arbitrary reference value near standard room temperature. The noise power is equal to

$$P_n = kTB \tag{1-30}$$

where P_n is the noise power, k is Boltzmann's constant (1.38×10^{-23} watts per kelvin), T is the temperature in kelvins, and B is the bandwidth (in hertz) in which the noise appears. For $T = 290$ K, P_n is therefore 4.00×10^{-18} watts, or –174 dBm in a 1-Hz bandwidth. Increasing B to a value suitable for digital communications, such as 160 kHz for a GSM system, admits more noise to the network, raising the minimum noise power against which an incoming signal must compete to –122 dBm.

If the noise figure and bandwidth are known, the noise floor can be calculated using the equation

$$\text{Noise floor} = -174 \text{ dBm} + NF + 10 \log B \tag{1-31}$$

The trouble with this equation is that the "integrated" bandwidth depends so much on the selectivity shape factor, which is not always known. Figure 1-98 shows the translation of the bandwidth of a single tuned circuit with its Gaussian shape into its rectangular equivalent. The transformation is done by sizing the rectangle such that A′ = A and B′ = B; when this is true, the area of the rectangle equals the area under the curve.

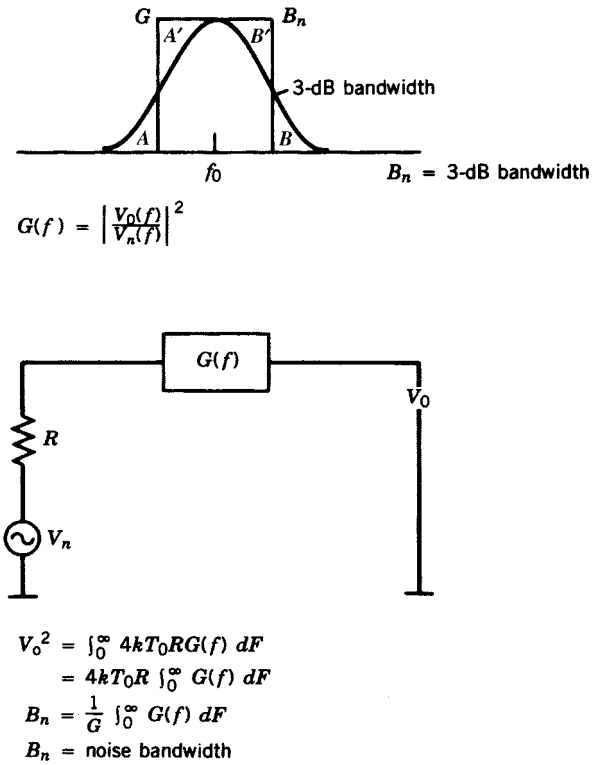


Figure 1-98 Graphical and mathematical explanation of the noise bandwidth from a comparison of the Gaussian-shaped bandwidth to the rectangular filter response.

Signal-to-Noise Ratio (S/N, SNR) and Sensitivity. Successful radiocommunication depends on the achievement of a specified minimum ratio of signal power to noise power, expressed in decibels, at the output of the receiver. The input voltage, expressed in absolute units or decibels relative to a microvolt (dBμV), necessary to achieve a particular signal-to-noise ratio in a particular bandwidth may be specified as a figure of merit called *sensitivity*. Because techniques used to measure S/N actually measure the ratio of signal-*plus-noise* to noise, specifications may refer to (or imply) S+N/N or (S+N)/N rather than S/N. The difference between (S+N)/N and S/N becomes negligible at high ratios of signal to noise; even at an SNR of 10 dB—a common value—the difference is only 0.46 dB [9].

Most receivers are designed to operate optimally when connected to an antenna system of a specified impedance (commonly 50 Ω), but relatively few receivers exhibit this design load impedance at their input terminals; that is, they are not designed for a conjugate input match. It is therefore customary to specify sensitivity in terms of "open circuit" voltage—the signal voltage that, with the receiver's antenna input terminated in its design antenna impedance, results in the desired ratio of signal to noise. If, as is commonly the case, the input voltage for a given SNR is determined using instrumentation calibrated in terms of *closed*-circuit voltage—that is, in terms of voltage across a load resistance equal to the instrument's source resistance—the voltage indicated will be one-half the open-circuit value for the SNR specified [10]. By convention, the open-circuit measurement condition is indicated by a sensitivity specification in volts of electromotive force (EMF). Specifying sensitivity in

terms of available signal power (usually decibels relative to 1 mW, or dBm) eliminates this open/closed-circuit confusion.

SINAD Ratio. Extending the measurement of signal-plus-noise to noise to include distortion results in a figure of merit called *SINAD* (signal-plus-noise-and-distortion), commonly applied to FM receivers:

$$SINAD = 10 \log_{10}\left(\frac{S+N+D}{N+D}\right) \tag{1-32}$$

where *SINAD* is in decibels, S is signal power, N is noise power, and D is distortion power. At a SINAD ratio of 12 dB—a common specification—the noise-and-distortion power is 25% that of the desired signal. As is true of (S+N)/N, SINAD closely approximates S/N at high ratios of signal to noise.

Bit Error Rate and Noise. For digital systems, signal-to-noise ratio and bit error rate are related. As introduced earlier, depending on the waveform, coding, and filtering, different BERs are related to particular SNRs (Figure 1-99). The SNR also depends on the interference from the synthesizer, as will be seen in the section on adjacent-channel power ratio (ACPR). Since a digital transmitter is always involved, in the on-the-air testing the laboratory measurement of SINAD cannot quite be correlated. At the IF level, we find the handover point between the transceiver's analog front end and its digital portion, and a standard noise-figure test setup, such as the one by Hewlett-Packard, is still a valuable instrument for evaluating noise figure and signal-to-noise ratio.

Here is an example for a digital measurement: Assume that we use a test generator like the SMIQ to evaluate a front end having an 8.8-dB noise figure. A 384-kbps π/4-DQPSK digital modulation signal (–5 dBm) is applied to the transceiver front end. Testing at 2.47 GHz, we find that the error vector magnitude (EVM; defined later in this section) is about 3.89% in transmit and about 1.37% in receive. To determine the receive sensitivity of the RF front end, the signal from the digitally modulated generator (at –95 dBm) is applied to the

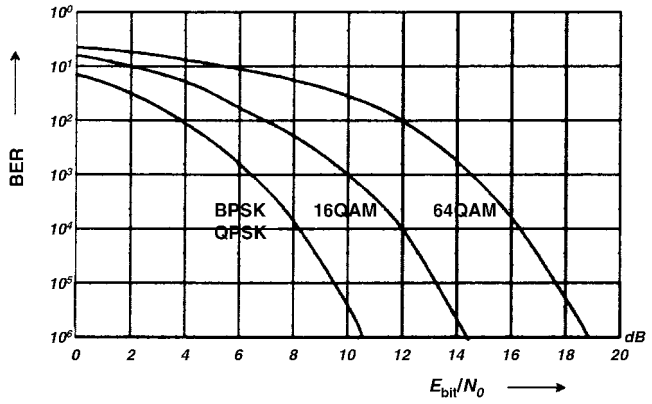


Figure 1-99 Bit error rate versus E_{bit}/N_0 for BPSK/QPSK, 16-QAM, and 64-QAM, showing the significantly greater SNR necessary for a given BER as the number of signal states is increased.

RF front end in receive mode. The measured EVM at 2.47 GHz is now 20%. By using the equation

$$S/N = -20 \log (EVM) \tag{1-33}$$

we see that the ratio that corresponds to an EVM of 20% is 14 dB, since

$$S/N = \frac{(E_S R_{S)}}{N_0 B} \tag{1-34}$$

where E_S is the energy of a symbol, R_S is the symbol rate, N_0 is the noise power density, and B is the bandwidth. Assuming a π/4-DQPSK signal and ratio of 14 dB, it can be determined that

$$E_S/N_0 = (S/N)\ (B/R_S) = 17 \text{ dB} \tag{1-35}$$

Assuming that the BER is 1E – 5, the required E_S/N_0 is 10 dB. Hence, from the definition of the sensitivity,

$$\begin{aligned} \text{Sensitivity} &= -144 + NF + R_S + E_S/N_0 \\ &= -144 + 8.8 + 10 \log(384/2) + 10 \\ &= -102 \text{ dBm} \end{aligned} \tag{1-36}$$

The value –144 is derived from kT_0 by taking the bandwidth into consideration.

The result obtained in Eq. (1-36) is consistent with the obtained E_S/N_0 of 17 dB from the measurement at –95 dBm input power [since –102 + (17 – 10) = –95]. Hence, it was determined that the receive sensitivity of the RF front end is approximately –102 dBm. We can then determine the dynamic range (DR) from

$$DR = \text{P}_{-1\text{dB}} - \text{Sensitivity} = -21 - 102 = 81 \text{ dB} \tag{1-37}$$

Noise Factor and Noise Figure. The noise floor predicted by Eq. (1-30) cannot be achieved and maintained in any real network or system of networks because all real networks generate noise. Determining how closely the SNR achieved at a given input level approaches the SNR achievable at that input level in a noiseless network is therefore of high interest to the circuit and system designer. The degree to which a network's noise contribution degrades the noise floor predicted by Eq. (1-30) is evaluated by its *noise factor* (*F*), which is expressed as the ratio

$$F = \frac{N_{\text{in}} + N_{\text{added}}}{N_{\text{in}}} \tag{1-38}$$

where F is the noise factor, N_{in} is the noise power available from the source, and N_{added} is the noise power added by the network, with both powers determined in the same bandwidth. Expressing this ratio in decibels ($10 \log_{10} F$) returns the *noise figure* (*NF*), a bandwidth-

independent figure of merit of great value in evaluating the noise performance of networks and communication systems. We can also express NF as the ratio of the network's input SNR to its output SNR:

$$NF = 10 \log_{10} \left[\frac{(S_{in}/N_{in})}{(S_{out}/N_{out})} \right] \tag{1-39}$$

where NF is the noise figure in decibels, S is signal power, and N is noise power, with the input and output values of these quantities signified by the subscripts and all powers determined in the same bandwidth. The noise figure of an ideal noiseless network is 0 dB; for all real, noisy networks, NF is positive. The NF of a lossy passive device is equal to its insertion loss.

Noise figure can also be defined for antennas and antenna systems:

$$NF_{ant} = 10 \log_{10} \frac{N_t + N_{ant}}{N_t} \tag{1-40}$$

where NF_{ant} is the antenna system noise figure in decibels, N_t is the antenna system's thermal noise power, as defined in Eq. (1-30), and N_{ant} is the total noise power picked up by the antenna system. From the lower end of the radio spectrum, and decreasingly up to approximately 400 MHz, noise intercepted by an antenna system from atmospheric, human-generated, and galactic sources will dominate NF_{ant} (Figure 1-100), and N_t can be considered as equivalent to the noise power of a resistor at 290 K. Atmospheric noise subsides above 40 MHz; from this region to perhaps 700 MHz, N_t is still largely negligible, with noise from human-generated and/or sky sources largely determining an antenna's noise figure. At these frequencies, an antenna's directivity and orientation relative to noise sources play a critical role in determining its noise figure; a strongly directive antenna located in a city suburb, for instance, exhibits a significantly higher noise figure when pointed toward the city center than

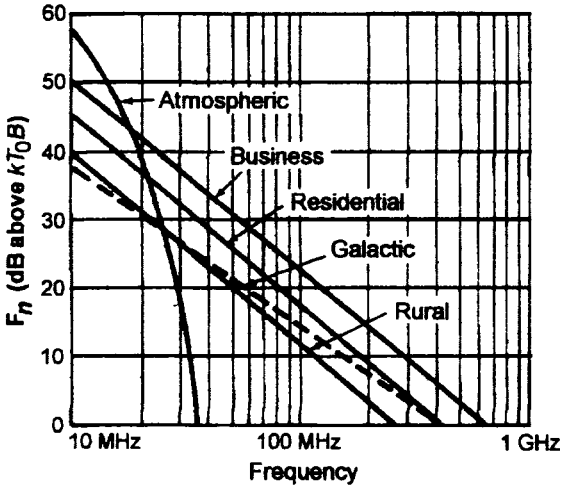


Figure 1-100 The higher the frequency, the more quiet the RF environment becomes, although the noise profile of specific sources may contradict this general rule.

it exhibits when pointed toward a more sparsely populated area. At frequencies above about 700 MHz, the RF environment is generally quieter, allowing receiver NF, not antenna NF, to more routinely limit a wireless receiver's sensitivity. At these frequencies, the concept of *noise temperature* is commonly used to evaluate the quietness of a receiver, its antenna system and its RF environment. Noise temperature is particularly useful in designing systems for space communication an antenna for which may be cooled, through radiation of a portion its own noise into the RF-cold sky, to a noise temperature an order of magnitude below 290 K.

Noise Figure of Cascaded Networks. The noise figure of two networks in cascade may be determined from

$$NF_{\text{total}} = 10 \log_{10}\left(F_1 + \frac{F_2 - 1}{G_1}\right) \tag{1-41}$$

where NF is the noise figure in dB, F_1 is the noise factor of the first network, F_2 is the noise factor of the second network, and G_1 is the gain (as a numerical ratio, *not* in dB) [11]. The noise figure of a system with more than two stages can be evaluated through repeated iterations of Eq. (1-41). Note that Eq. (1-41) assumes two conditions: (1) that F_1 and F_2 are determined in the same bandwidth, and (2) that the networks' input and output terminations are resistive—a condition that is commonly *not* true of RF amplifiers optimized for lowest noise. Equation (1-41) can be expanded to account for bandwidth, but accounting for complex terminations requires the use of *noise correlation matrix* techniques as described by Rohde et al. [9].

1-7-2 System Amplitude and Phase Behavior

If we could build electronic systems that were absolutely amplitude- and phase-linear, radiocommunication system design would be greatly simplified. An amplifier designed for a power gain of 10 dB, for example, would merely increase the power of all signals at its input by a factor of 10, regardless of their frequencies, while perfectly maintaining their relative phases. But all real electrical and electronic networks, even those designed (or supposed) to be amplitude- and phase-linear, exhibit amplitude and phase nonlinearity to some degree, just as they all generate noise to some degree.

The effects of amplitude nonlinearity, generically referred to as *nonlinear distortion*, include the generation, through *harmonic distortion* and/or *intermodulation distortion* (IMD), of output signals at frequencies not present at a system's input. Nonlinear distortion also results in *gain compression*—changes in system gain with changes in input-signal level. By convention, when workers in electronics refer to or consider a network's "linearity," they usually mean its *amplitude* linearity; likewise, by "distortion" they usually mean *nonlinear* distortion.

The effects of phase or frequency nonlinearity are generically referred to as *linear distortion* because they occur independently of signal amplitude and polarity. We often intentionally apply linear distortion through *filtering*, which modifies the amplitude relationships among existing spectral components of a signal without creating any new frequencies. Another linear-distortion effect, *phase* or *delay distortion*, results in the delay of signals of differing frequencies by differing amounts of time. In a system where signal phase conveys

information, as is true of most wireless links, phase distortion can seriously degrade communication.

The fact that all real networks are amplitude- and angle-nonlinear to some degree means that all real networks modify the amplitude and angle characteristics of the signals they handle. What is perhaps less obvious is that subjecting a signal to amplitude and angle nonlinearities causes "crosstalk" between its amplitude and angle characteristics. For example, through a nonlinear distortion effect called AM-to-PM conversion, changes in a signal's amplitude result in changes in its phase. Filtering an angle-modulated signal produces the reverse effect: Deprived by the filter's selectivity of spectral components necessary to maintain its envelope constancy, the signal emerges from the filter as a combination of angle and amplitude modulation.

Spectral Considerations of Analog and Digitally Modulated Signals. Many wireless modulation schemes exist and even more are proposed; all use angle (usually phase) modulation or a combination of phase and amplitude modulation using an emission format engineered to achieve multiple goals of information throughput, robustness, spectral and hardware efficiency, and reproducibility. Digital modulation is standard in mainstream wireless applications because it allows increased channel capacity, and immunity to noise and distortion, compared to analog systems. At the circuit-design level, the particulars of whether, or to what degree, a modulation scheme is AM, PM, analog, or digital matters less than the actual angle and amplitude characteristics of the signal(s) involved, and the tolerances within which these characteristics can be expected and allowed to vary.

A system's amplitude linearity is of major concern because of its relation to energy efficiency, receiver dynamic range, and transmitter spectral purity. Energy efficiency, important because many wireless applications use battery power, generally decreases inversely with amplitude linearity. Yet, sufficient amplitude linearity must be guaranteed in the front-end circuitry of wireless receivers subjected to multiple strong signals, and in all wireless receiver and transmitter stages handling variable-envelope signals.

Some digital modulation schemes, designed to be distortion-tolerant, produce constant-envelope signals; Gaussian minimum shift keying (GMSK) and Feher's quadrature phase shift keying (FQPSK) are examples of these. Such signals can be processed in highly nonlinear circuitry—for example, an RF power amplifier operated in saturation to maximize efficiency—without degrading their spectral composition. Other digital modulation schemes result in the emission—ideally—of a single PM sideband, with the carrier and all other sideband components suppressed. Such a signal exhibits, of necessity, phase and amplitude variations; the IS-54 cellular standard's $\pi/4$ differential quadrature phase shift keying ($\pi/4$-DQPSK) scheme, which results in envelope fluctuations of 3–6 dB, exemplifies this [12]. Intermodulation distortion among such a signal's spectral components can generate products that fall outside the bandwidth painstakingly controlled in the modulation process. This *spectral regrowth* must be minimized to prevent adjacent-channel interference (Figure 1-101).

For a better understanding, Figure 1-102 shows the various channels and the energy levels associated with them.

Amplitude Linearity Issues and Figures of Merit. A network's amplitude nonlinearity can be characterized by the expansion

$$y = k_1 f(x) + k_2[f(x)]^2 + k_3[f(x)]^3 + \text{higher-order terms} \tag{1-42}$$

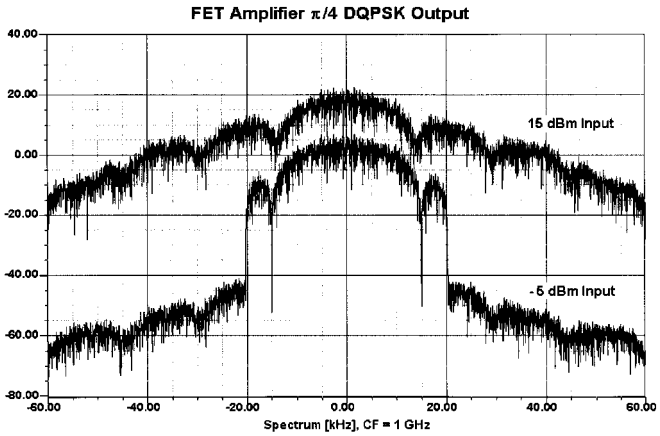


Figure 1-101 *Spectral regrowth* results when a variable-envelope emission, π/4-DQPSK in this case, is subjected to significant nonlinear distortion. This graph shows the simulated performance of a MESFET power amplifier operating at 1 GHz; the amplifier is 6 dB into compression when driven at 15 dBm.

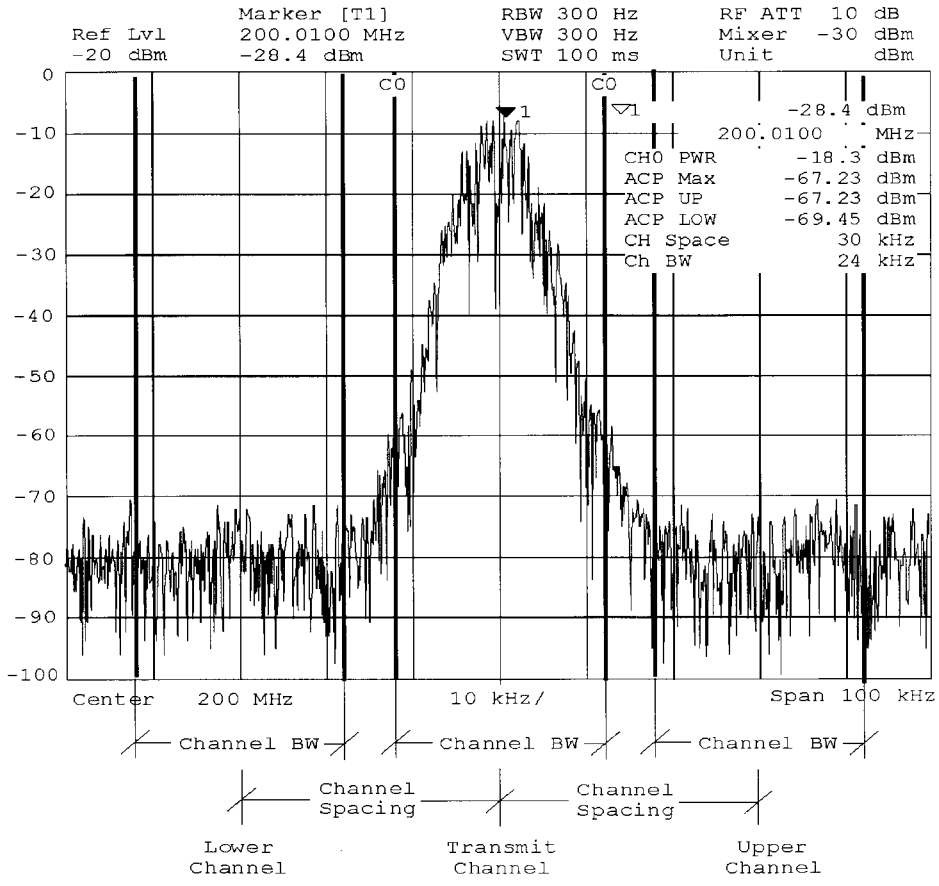


Figure 1-102 NADC signal and parameters, including channel spacing and channel bandwidth.

where y represents the output, the coefficients k_n represent complex quantities whose values can be determined by an analysis of the output waveforms, and $f(x)$ represents the input. Even though all practical networks exhibit amplitude nonlinearity, we can (and often do) refer to many networks as "linear." We say this of networks that are *sufficiently amplitude-linear for our purposes*—for example, weakly nonlinear networks in which small-signal operation is assumed even though the signal levels involved are sufficient to cause slight distortion. For many practical purposes, the first three terms of Eq. (1-42) adequately describe such a network's nonlinearity:

$$y = k_1 f(x) + k_2[f(x)]^2 + k_3[f(x)]^3 \tag{1-43}$$

In adopting this simplification, we assume also that the nonlinearity is frequency independent—that is, the network has sufficient bandwidth to allow all of the products predicted by Eq. (1-43) to appear at its output terminals unperturbed [13].

When multiple signals are present in a network, even weak nonlinearity can result in profound consequences. To illustrate this, we'll let $f(x)$ consist of two sinusoidal signals:

$$f(x) = A_1 \cos\omega_1 t + A_2 \cos\omega_2 t \tag{1-44}$$

We'll assume that ω_1 and ω_2 are close enough so that the coefficients k_i can be considered equal for both signals. We'll also assume for simplicity that all of the k_i are real. If Eq. (1-43) describes the network's response to an input $f(x)$, the response will be

$$\begin{aligned} y &= k_1(A_1 \cos\omega_1 t + A_2 \cos\omega_2 t) + k_2(A_1 \cos\omega_1 t + A_2 \cos\omega_2 t)^2 + k_3(A_1 \cos\omega_1 t + A_2 \cos\omega_2 t)^3 \\ &= k_1(A_2 \cos\omega_1 t + A_2 \cos\omega_2 t) \\ &\quad + k_2\left[A_1^2 \frac{1+\cos 2\omega_1 t}{2} + A_2^2 \frac{1+\cos 2\omega_2 t}{2} + A_1 A_2 \frac{\cos(\omega_1+\omega_2)t + \cos(\omega_1-\omega_2)t}{2}\right] \\ &\quad + k_3 \left\{ \begin{array}{l} \left[\left[A_1^3\left(\dfrac{\cos\omega_1 t}{2} + \dfrac{\cos\omega_2 t}{4} + \dfrac{\cos 3\omega_1 t}{4}\right) + A_2^3\left(\dfrac{3\cos\omega_2 t}{4} + \dfrac{\cos 3\omega_2 t}{4}\right)\right]\right] \\ + A_1^2 A_2[\tfrac{3}{2}\cos\omega_2 t + \tfrac{3}{4}\cos(2\omega_1 - \omega_2)t + \tfrac{3}{4}\cos(2\omega_1 - \omega_2)t] \\ + A_2^2 A_1[\tfrac{3}{2}\cos\omega_1 t + \tfrac{3}{4}\cos(2\omega_2 + \omega_1)t + \tfrac{3}{4}\cos(2\omega_2 - \omega_1)t] \end{array} \right\} \end{aligned} \tag{1-45}$$

The k_1 term of Eq. (1-45) represents the results of amplitude-linear behavior. No new frequency components have appeared; the two sine waves have merely been "rescaled" by k_1.

The second- and third-order terms of Eq. (1-45) represent the effects of harmonic distortion and intermodulation distortion. Second-order effects include second-harmonic distortion (the production of new signals at $2\omega_1$ and $2\omega_2$) and IMD (the production of new signals at $\omega_1 + \omega_2$ and $\omega_1 - \omega_2$). Third-order effects include gain compression, third-harmonic distortion (the production of new signals at $3\omega_1$ and $3\omega_2$), and IMD (the production of new signals at $2\omega_1 \pm \omega_2$ and $2\omega_2 \pm \omega_1$).

Gain Compression. Gain compression occurs when a network cannot increase its output amplitude in linear proportion to an amplitude increase at its input; gain *saturation* occurs when a network's output amplitude stops increasing (in practice, it may actually decrease) with increases in input amplitude. We can deduce from Eq. (1-45) that the amplitude of the $\cos\omega_1 t$ signal has become

$$A_1' = k_1 A_1 + k_3\left(\tfrac{3}{4}A_1^3 + \frac{3}{2}A_1 A_2^2\right) \tag{1-46}$$

Because k_3 will normally be negative, a large signal $A_2 \cos\omega_2 t$ can effectively mask a smaller signal $A_1 \cos\omega_1 t$ by reducing the network's gain. This third-order effect, known as *blocking* or *desensitization* when it occurs in a receiver, is a special case of gain compression. The presence of additional signals results in a greater reduction in gain; the gain reduction for each signal is a function of the relative levels of all signals present. A receiver's blocking behavior may be characterized in terms of the level of off-channel signal necessary to reduce the strength of an in-passband signal by a specified value, typically 1 dB; alternatively, the decibel ratio of the off-channel signal's power to the receiver's noise-floor power may be cited as *blocking dynamic range*. Desensitization may also be characterized in terms of the off-channel signal power necessary to degrade a system's SNR by a specified value.

Multiple signals need not be present for gain compression to occur. If only one signal is present, the ratio of gain with distortion to the network's idealized (linear) gain is

$$A_1' = \frac{k_1 + k_3\left(\tfrac{3}{4}A_1^2\right)}{k_1} \tag{1-47}$$

This is referred to as the *single-tone gain-compression factor*. Figure 1-103 shows how the k_3 term causes a network's gain to deviate from the ideal. The point at which a network's power gain is down 1 dB from the ideal for a single signal is a figure of merit known as the *1-dB compression point* (P_{-1dB}). Many networks (including many receiving and low-level transmitting circuits, such as low-noise amplifiers, mixers, and IF amplifiers) are usually operated under small-signal conditions—at levels sufficiently below P_{-1dB} to maintain high linearity. As we'll see, however, some networks (including power amplifiers for wireless systems) may be operated under large-signal conditions—near or in compression—to achieve optimum efficiency at some specified level of linearity. Figure 1-104 shows what happens when a digital emission that uses amplitude to convey information is subjected to amplitude compression.

Intermodulation. The new signals produced through intermodulation distortion (IMD) can profoundly affect the performance even of systems operated far below gain compression (Figure 1-105). IMD products of significant power can appear at frequencies remote from, in, and/or near the system passband, resulting in demodulation errors (in reception) and interference to other communications (in transmission). Where an IMD product appears relative to the passband depends on the passband width and center frequency, the frequencies of the signals present at the system input, and the order of the nonlinearity involved. These factors also determine the strength of an IMD product relative to the desired signal.

Second-order IMD (IM_2) results, for an input consisting of two signals ω_1 and ω_2, in the production of new signals at $\omega_1 + \omega_2$ and $\omega_1 - \omega_2$; third-order IMD (IM_3) results, for an input

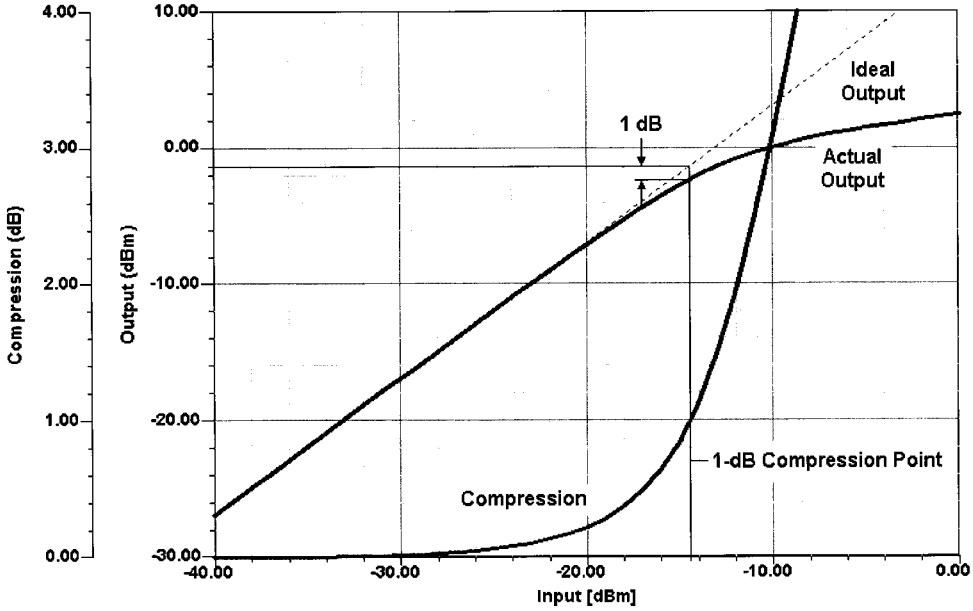


Figure 1-103 The power level at which a network's power output is down 1 dB relative to that of its ideally linear equivalent is a figure of merit known as the *1-dB compression point* (P_{-1dB}). The 1-dB compression point can be expressed relative to input power ($P_{-1dB,in}$) or output power ($P_{-1dB,out}$). For the amplifier simulated here, $P_{-1dB,in} \approx -14.5$ dBm and $P_{-1dB,out} \approx -1.3$ dBm.

consisting of two signals ω_1 and ω_2, in the production of new signals at $2\omega_1 \pm \omega_2$ and $2\omega_2 \pm \omega_1$.

Under small-signal conditions—that is, at levels well below compression—the power of an IM_2 product varies by 2 dB, and the power of an IM_3 product varies by 3 dB, per decibel change in input power level. This allows us to derive a network figure of merit, the *intermodulation intercept point* (*IP*), for a given IM order by extrapolating a network's linear and IM responses to their point of intersection (Figure 1-106)—the point at which their powers would be equal if compression did not occur. The intercept point for a given IM order

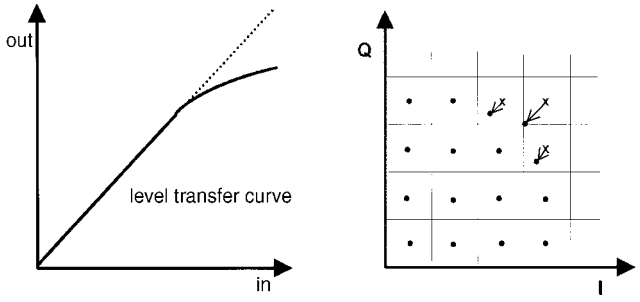


Figure 1-104 Influence of differential amplitude error (compression) on a QAM constellation.

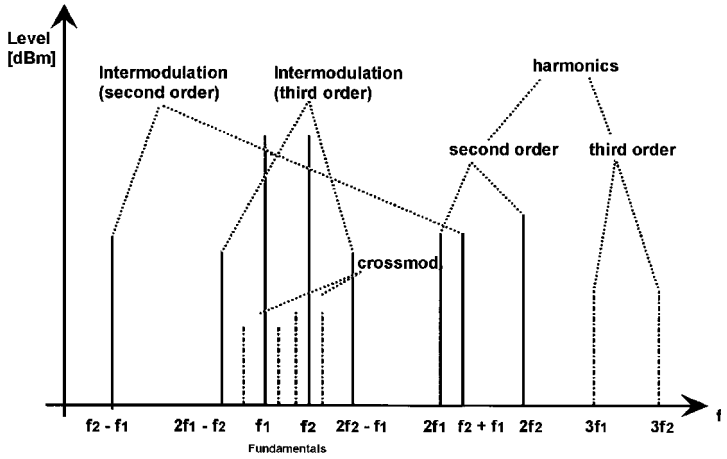


Figure 1-105 Relationships between fundamental and spurious signals, including harmonics and products of intermodulation.

n can be expressed, and should always be characterized, relative to input power ($IP_{n,\text{in}}$) or output power ($IP_{n,\text{out}}$); the IP_{in} and IP_{out} values differ by the network's linear gain. For equal-level test tones, $IP_{n,\text{in}}$ can be determined by

$$IP_{n,\text{in}} = \frac{nP_A - P_{\text{IM}_n}}{n-1} \tag{1-48}$$

where n is the order, P_A is the input power (of one tone), P_{IM_n} is the power of the IM product, and IP is the intercept point. The intercept point for cascaded networks can be determined from

$$IP_{2,\text{in}} = \frac{1}{\left(\frac{1}{\sqrt{IP1}} + \frac{G}{\sqrt{IP2}}\right)^2} \tag{1-49}$$

for IP_2 and from

$$IP_{3,\text{in}} = \frac{1}{\frac{1}{IP1} + \frac{G}{IP2}} \tag{1-50}$$

for IP_3. In both equations, $IP1$ is the input intercept of Stage 1 in watts, $IP2$ is the input intercept of Stage 2 in watts, and G is the gain of Stage 1 (as a numerical ratio, *not* in decibels). Both equations assume the worst-case condition, in which the distortion products of both stages add in-phase.

The ratio of the signal power to the IM-product power, the *distortion ratio*, can be expressed as

$$R_{dn} = (n-1)\left[IP_{n,\text{in}} - P_{\text{in}}\right] \tag{1-51}$$

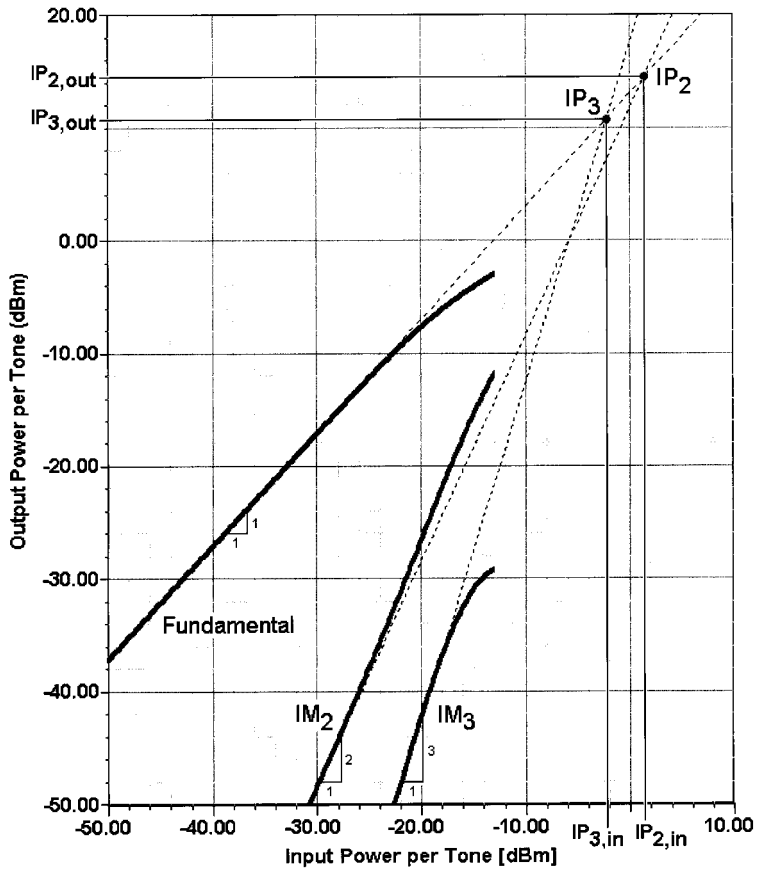


Figure 1-106 The level at which the power of one of a network's IM products equals that of the network's linear output is a figure of merit known as the *intermodulation intercept point* (*IP*). The intercept point for a given IM order *n* can be expressed, and should always be characterized, relative to input power ($IP_{n,in}$) or output power ($IP_{n,out}$); the IP_{in} and IP_{out} values differ by the network's linear gain. For the amplifier simulated here, $IP_{2,in} \approx 1.5$ dBm, $IP_{2,out} \approx 14.5$ dBm, $IP_{3,in} \approx -2.3$ dBm, and $IP_{3,out} \approx 10.7$ dBm. Each curve depicts the power in one tone of the response evaluated.

where n is the order, R_{dn} is the distortion ratio, $IP_{n,in}$ is the input intercept point, and P_{in} is the input power of one tone.

Discussions of IMD have traditionally downplayed the importance of IM_2 because the incidental distributed filtering contributed by the tuned circuitry once common in radiocommunication systems was usually enough to render out-of-passband IM_2 products caused by in-passband signals, and in-passband IM_2 products caused by out-of-passband signals, vanishingly weak compared to fundamental and IM_3 signals. In broadband systems that operate at bandwidths of an octave or more, however, in-passband signals may produce significantly strong in-passband IM_2 and second-harmonic products. In such applications, balanced circuit structures (such as push–pull amplifiers and balanced mixers) can be used to minimize IM_2 and other even-order nonlinear products.

As with IM_2, which IM_3 products are important depends on the spacing of the signals involved and the relative width of the system passband. If ω_1 and ω_2 are of approximately

the same frequency, the additive products $2\omega_1 + \omega_2$ and $2\omega_2 + \omega_1$ will be outside the passband of a narrowband system. The subtractive products $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$, however, will likely appear near or within the system passband. The IM_3 performance of any network subjected to multiple signals is therefore of critical importance, and an array of IM_3-related, sometimes application-specific, figures of merit has evolved as a result.

Dynamic Range. As we have seen, thermal noise sets the lower limit of the power span over which a network can operate. Distortion—that is, degradation by distortion of the signal's ability to convey information—sets the upper limit of a network's power span. Because the power level at which distortion becomes intolerable varies with signal type and application, a generic definition has evolved: The upper limit of a network's power span is the level at which the power of one IM product of a specified order is equal in power to the network's noise floor. The ratio of the noise-floor power to the upper-limit signal power is referred to as the network's *dynamic range* (*DR*), often more carefully characterized as *two-tone IMD dynamic range*, which, when evaluated with equal-power test tones, is a figure of merit commonly used to characterize receivers. When $IP_{n,\text{in}}$ and MDS are known, IMD DR can be determined from

$$DR_n = \frac{(n-1)[IP_{n,\text{in}} - MDS]}{n} \tag{1-52}$$

where *DR* is the dynamic range in decibels, *n* is the order, IP_{in} is the input intercept power in dBm, and *MDS* is the minimum detectable signal power in dBm.

Dynamic Measure: A New Receiver Figure of Merit. Taken by themselves, the IP_n and DR figures of merit can be misleading in the sense that broadband, resistive attenuation directly improves them by the amount of attenuation added while directly degrading NF by the same amount. The NF and MDS figures of merit can mislead us because they convey no sense of suitability to task. This can be significant when we interconnect networks and systems of differing noise floors and IMD dynamic ranges—as we do, for example, when connecting a receiver to an antenna. As a means of better evaluating a receiver's front-end performance, the authors propose a new figure of merit, *dynamic measure*, defined as

$$DM = [(IP_{3,\text{in}} + \text{Att}) - (NF_r + \text{Att} + NF_{ant})]\frac{NF_r}{NF_r + \text{Att}} \tag{1-53}$$

where *DM* is the dynamic measure, a dimensionless number; $IP_{3,\text{in}}$ is the receiver's third-order input intercept in dBμV (instead of dBm) without any added attenuation; NF_r is the receiver noise figure in dB without any added attenuation; NF_{ant} is the antenna-system noise figure in dB as determined by Eq. (1-40); and Att is the decibel value of any added attenuation, which is sometimes used to shift the upper and lower limits of a system's dynamic range to higher absolute signal levels, increasing $IP_{3,\text{in}}$ at the expense of making the system more "deaf." Our dynamic measure figure of merit takes this trade-off into account to allow more straightforward comparisons of systems that have the same dynamic range but different sensitivities. Setting NF_{ant} equal to zero allows evaluation of receiver performance under laboratory conditions; setting NF_{ant} equal to the measured or expected NF of the antenna system to which the receiver will be connected allows evaluation of receiver performance in a system sense. Rather than express IP_3 in dBm, we are going to

express it in dBμV. This is necessary because for negative values of the difference between IP_3 and the added noise figure, the multiplier term will make the result larger when its numerical value is less than 1.

The usefulness of the dynamic measure figure of merit is evident when we consider several examples.

Example 1. Here we evaluate a receiver that exhibits the characteristics $IP_{3,\text{in}} = 16$ dBm(123 dBμV), $NF_r = 8$ dB, and $NF_{\text{ant}} = 10$ dB. We calculate its laboratory DM by setting NF_{ant} equal to 0:

$$DM = [123 - (8 + 0)]\frac{8}{8 + 0} = [115]1 = 115 \tag{1-54}$$

Setting NF_{ant} equal to 10 returns the receiver's system DM:

$$DM = [123 - (8 + 10)]\frac{8}{8 + 0} = [115]1 = 105 \tag{1-55}$$

Example 2. This setup consists of the Example 1 receiver with a 10-dB pad switched in, resulting in the characteristics $IP_{3,\text{in}} = 26$ dBm(133 dBμV), $NF_r = 18$ dB, and $NF_{\text{ant}} = 10$ dB. Setting NF_{ant} equal to 0, we calculate its laboratory DM as

$$DM = [133 - (8 + 10 + 0)]\frac{8}{8 + 10} = [115]0.4444 = 51.1 \tag{1-56}$$

A comparison with Eq. (1-54) shows that, although inserting the 10-dB pad has shifted $IP_{3,\text{in}}$ from 16 dBm (123 dBμV) to 26 dBm (133 dBμV), it has also increased the resulting NF from 8 to 18 dB, reducing the DM from 115 to 51.1. Inserting the pad also degrades the receiver's system DM compared to Example 1, as we see by setting NF_{ant} equal to 10:

$$DM = [133 - (8 + 10 + 10)]\frac{8}{8 + 10} = [105]0.4444 = 46.7 \tag{1-57}$$

Example 3. On the other hand, had we inserted 2 dB of attenuation rather than 10 dB [$IP_{3,\text{in}} = 18$ dBm (125 dBμV), $NF_r = 10$ dB, and $NF_{\text{ant}} = 10$ dB], the laboratory DM would have been

$$DM = [125 - (8 + 2 + 0)]\frac{8}{8 + 2} = [115]0.8 = 92 \tag{1-58}$$

and the system DM would have been

$$DM = [125 - (8 + 2 + 10)]\frac{8}{8 + 2} = [105]0.8 = 84 \tag{1-59}$$

Table 1-12 shows these and several additional examples for comparison.

The interested reader will notice the following: The receiver in Example 1 (base NF 8 dB) was unnecessarily sensitive compared to the antenna NF of 10 dB, so adding a 2-dB pad increased the intercept point by 2 dB and matched the antenna noise floor. Yet, the equation rightfully indicates a loss of dynamic measure. The relationship between NF_{ant} and NF_r is not taken

Table 1-12 Eight dynamic measure cases compared

Example	System	Base $IP_{3,in}$ (dBm)	Base $IP_{3,in}$ (dBμV)	Attenuation (dB)	Resulting $IP_{3,in}$ (dBm)	Resulting $IP_{3,in}$ (dBμV)	Base NF_r (dB)	Resulting NF_r (dB)	NF_{ant} (dB)	DM_{lab}	DM_{system}
1	HF receiver	16.0	123.0	0.0	16.0	123.0	8.0	8.0	10.0	115.0	105.0
2	HF receiver with 10-dB pad	16.0	123.0	10.0	26.0	133.0	8.0	18.0	10.0	51.1	46.7
3	HF receiver with 2-dB pad	16.0	123.0	2.0	18.0	125.0	8.0	10.0	10.0	92.0	84.0
4	Rohde & Schwarz XK 2100 shortwave transceiver	40.0	147.0	0.0	40.0	147.0	10.0	10.0	10.0	137.0	127.0
5	XK 2100 with 10 dB of unnecessary attenuation	40.0	147.0	10.0	50.0	157.0	10.0	20.0	10.0	68.5	63.5
6	Satellite receiver	–2.2	104.8	0.0	–2.2	104.8	0.9	0.9	0.2	103.9	103.7
7	Wireless front end without duplexer/filter losses	–12.0	95.0	0.0	–12.0	95.0	2.6	2.6	2.0	92.4	90.4
8	Wireless front end with 3-dB duplexer/filter losses	–12.0	95.0	3.0	–9.0	98.0	2.6	5.6	2.0	42.9	42.0

into consideration. Also, we have decided to leave it to the reader to develop a formula that deals with the addition of a preamplifier rather than an attenuator. This will complicate the formula because the preamplifier's gain and noise figure are not necessarily related.

Triple-Beat Distortion and Cross-Modulation. P_{-1dB} is a single-tone figure of merit; blocking, intercept point, and dynamic range evaluate two-tone behavior. For networks that must handle AM and composite (AM and angle modulation) signals very linearly, such as television transmitters and cable TV distribution systems, a three-tone figure of merit called *triple-beat distortion* has gained acceptance. Signals at ω_1 and ω_2 (closely spaced) and ω_3 (positioned far away from ω_1 and ω_2) are applied to the network under test, at levels, frequencies, and spacings that vary with the application. One triple-beat distortion figure of merit is the ratio, expressed in decibels, of the IM product at $\omega_3 + (\omega_2 - \omega_1)$ to one of the network's linear outputs at a specified output level. Alternatively, the triple-beat figure of merit may express the network output level at which a specified triple-beat ratio occurs.

Triple-beat distortion is the mechanism underlying cross-modulation, a form of intermodulation in which one or more AM signals present in a network amplitude-modulate all

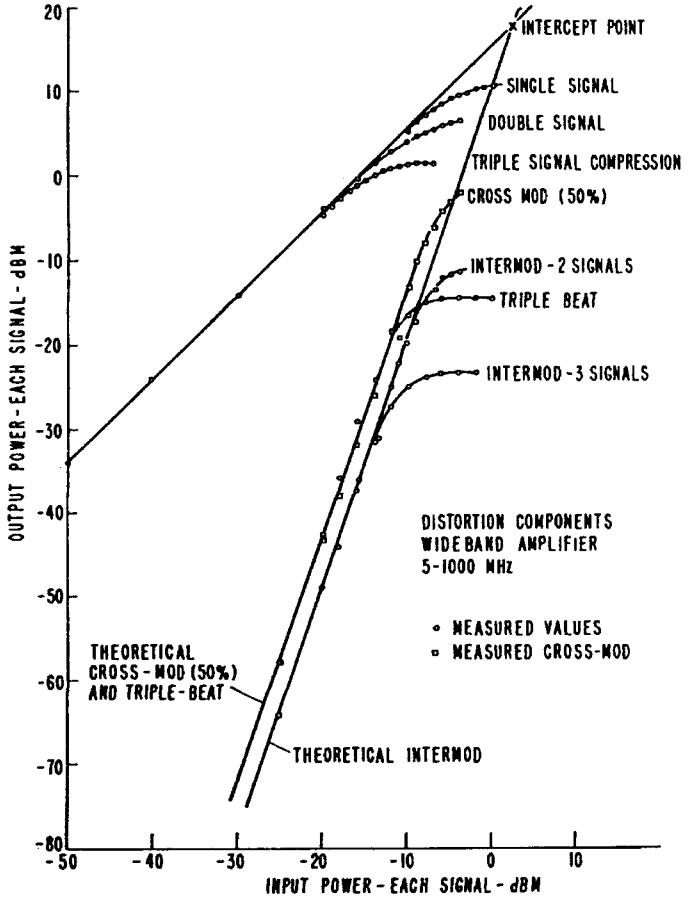


Figure 1-107 Measured distortion components in a wideband (5–1000 MHz) amplifier. Figure 1-108 shows a magnified view of the gain-compression region [15].

signals present in the network [14]. Angle-modulation-based wireless systems are largely immune to such effects.

Figures 1-107 and 1-108 graph the results of gain compression, two-tone intermodulation, cross-modulation, and triple-beat testing on a wideband (5–1000 MHz) amplifier.

Noise Power Ratio. Triple-beat testing is one way of improving on two-tone testing as a means of evaluating a network's intermodulation behavior in the presence of multiple signals. Another figure of merit, *noise power ratio* (*NPR*), uses thermal noise as a test signal. The test measures the introduction, by IM, of noise into a quiet slot created by the insertion of a bandstop filter, equal in stopband width to the width of the measurement channel, between the noise generator and the network under test (Figure 1-109).

Large-Signal Effects. Except for P_{-1dB}, the figures of merit discussed so far evaluate amplitude nonlinearity under small-signal conditions. At input powers that drive a network into gain compression and saturation, IM products of odd orders higher than 3 become significant, and curves, dips, and nulls appear in their characteristics (Figure 1-110). Phase shifts related to nonlinear (primarily voltage-dependent) capacitances in solid-state devices are one cause of these effects. Under such conditions, a network may exhibit hysteresis, with its behavior at any given instant depending not only on the voltage or current applied to it but also on its recent history [16].

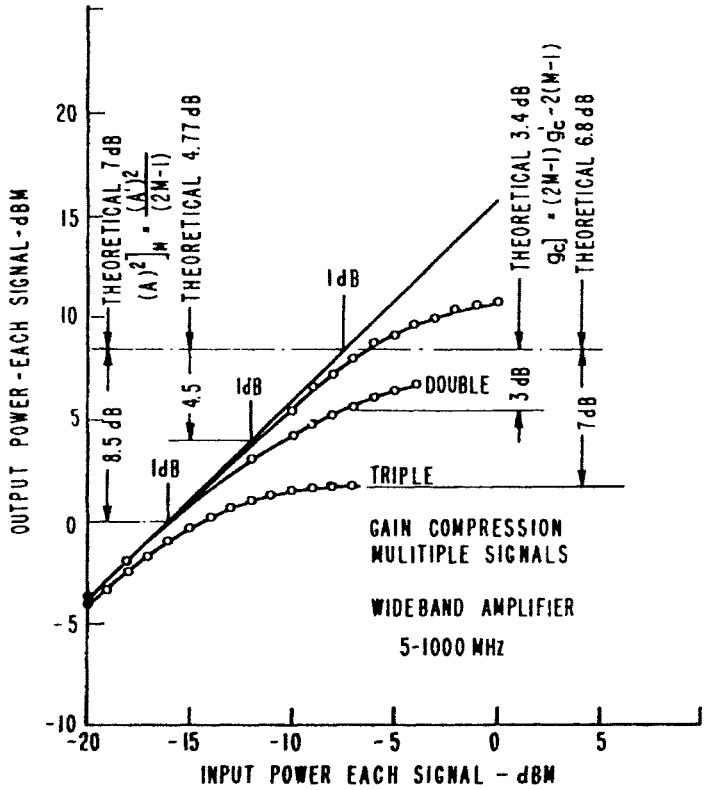


Figure 1-108 Measured multiple-signal gain compression of the 5- to 1000-MHz amplifier [15].

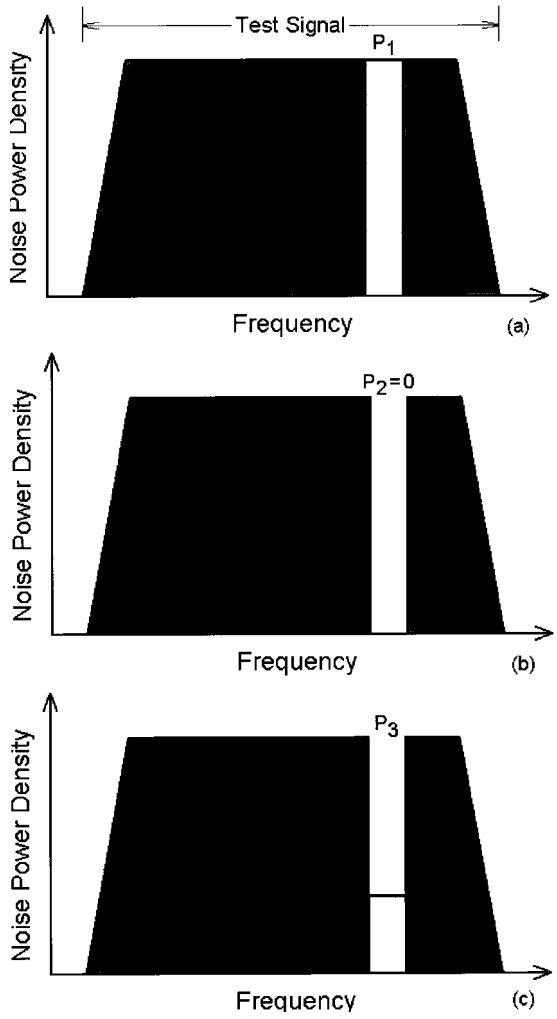


Figure 1-109 Determining a network's *noise power ratio* (*NPR*) involves the application of a test signal consisting of thermal noise [15]. (a) The reference measurement-channel noise power, P_1, is then measured. (b) Next, a bandstop filter is placed between the noise generator and network under test to keep the test signal out of the measurement channel. Assuming sufficient filter attenuation, if the network were absolutely noiseless and linear, the ideal noise power in the measurement channel, P_2, would then be zero. (c) In practice, the network's own thermal noise and intermodulation between noise components outside the measurement channel result in an actual measurement-channel noise power (P_3) greater than zero. The noise power ratio equals P_1/P_3.

AM-to-PM Conversion. The nonlinear distortion effects we've discussed so far can be termed *AM-to-AM distortion*—distortion that, to a degree that depends on the amplitude of the signal(s) applied to the network, results in changes in the network's gain and/or production of signals at new frequencies. AM-to-PM distortion can also occur. As a network nears saturation, part of its driving signal goes into shifting the bias point(s) of its active device(s), changing their drive-dependent reactances and shifting the phase of the output signal relative to its value at input levels below compression (Figures 1-111 and 1-112). This

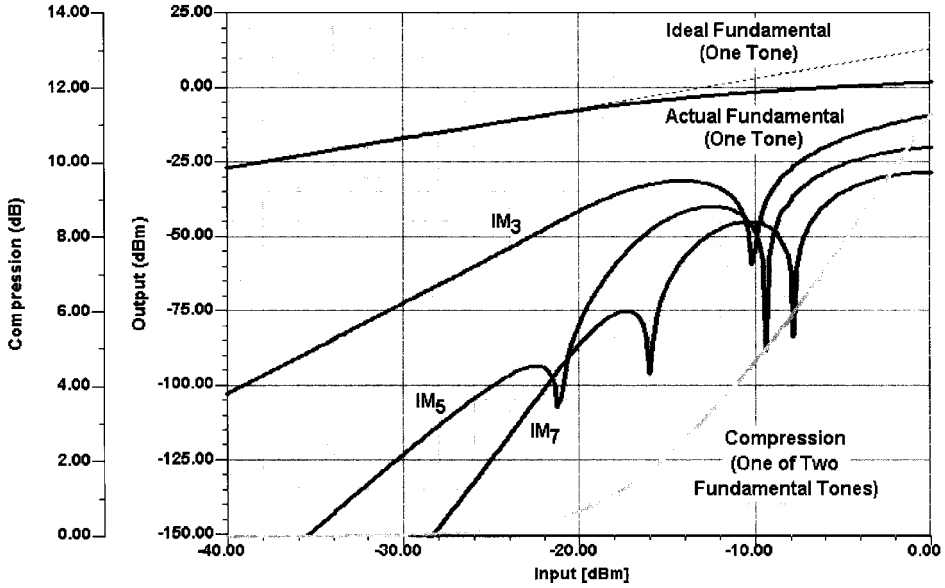


Figure 1-110 As a network is driven into compression, IM products at odd orders higher than 3 become significant, and phase shifts in power-dependent device capacitances cause curves and dips in the IM characteristics. The onset of these departures from IM-response linearity occurs at generally lower input power levels for higher IM orders; their severity, and their position on the IM curves, differs among the various products of a given order and varies with network topology and tone spacing. Figures of merit based on straight-line IM responses fail to usefully predict nonlinear network behavior under these conditions. This graph shows the simulated performance of a single-BJT broadband amplifier driven by two equal-amplitude tones at 10 and 11 MHz.

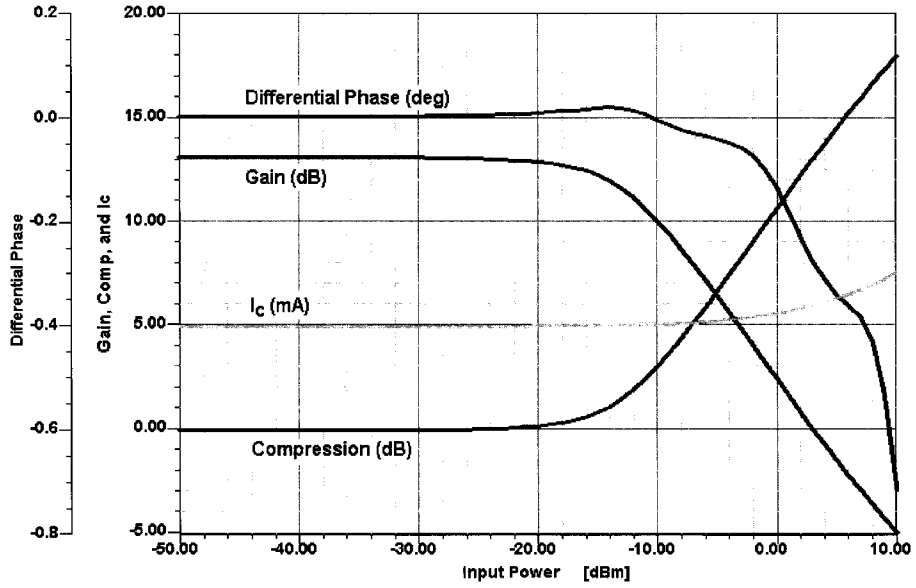


Figure 1-111 Driving a network into compression and saturation shifts the bias point(s) of its active device(s), changing their drive-dependent reactances and shifting the phase of the output signal relative to its value at input levels below compression. This graph shows the simulated performance of a single-BJT broadband amplifier driven by a single tone at 10 MHz.

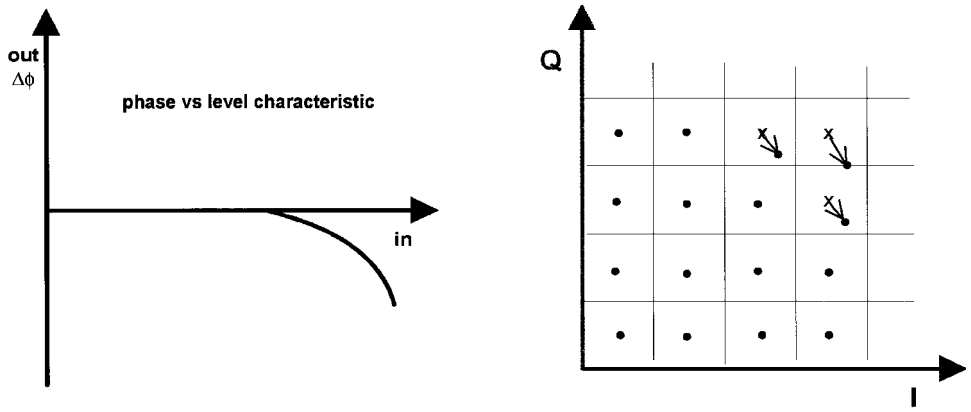


Figure 1-112 Influence of differential phase error (AM-to-PM conversion) on a QAM constellation.

effect, *AM-to-PM conversion*, can cause incidental phase modulation that degrades the performance of digital communication systems.

Spectral Regrowth and Adjacent-Channel Power Ratio. Spectral regrowth occurs largely as a result of third-, fifth-, and seventh-order IMD in power amplifiers operated near or in compression—at power levels where hysteretic IM effects result in poor agreement between measured behavior and predictions based on small-signal IM figures of merit [17]. We therefore evaluate the impact of spectral regrowth more directly, using a figure of merit called *adjacent-channel power ratio* (ACPR). ACPR measurement techniques that incorporate memory can be used to increase ACPR predictions for networks that exhibit saturation hysteresis [18]. Figure 1-113 shows the critical relationship between compression, power-added efficiency, and ACPR in a MESFET power amplifier.

Phase Response Issues and Figures of Merit. We have already seen how large-signal nonlinear distortion can result in amplitude-dependent phase shifts through AM-to-PM conversion. Because phase linearity is critical at *all* signal levels in PM systems, especially those using digital modulation, we must also consider linear distortion in evaluating networks used in wireless systems.

Differential Group Delay. Every frequency-selective network subjects signals passing through it to some degree of time delay. Ideally, this delay, also known as *group* or *envelope delay*, does not vary with frequency; that is, the network's phase shift versus frequency response is monotonic and linear. In practice, a network's time delay varies across its passband, transition bands, and stopbands, exhibiting curvature, ripple, and transition-band peaks (Figure 1-114). The network's *differential* group delay—its group-delay spread—is therefore of considerable importance. This is especially so in digitally modulated systems, where the resulting phase distortion can cause errors in modulation and demodulation.

Effects of Phase Noise. As Chapter 5 will discuss in detail, the phase of an oscillator's output signal is subject to random phase variations (Figures 1-115 and 1-116). Called *phase noise*, this effect is often quantified as the decibel ratio of the phase-noise power in a single

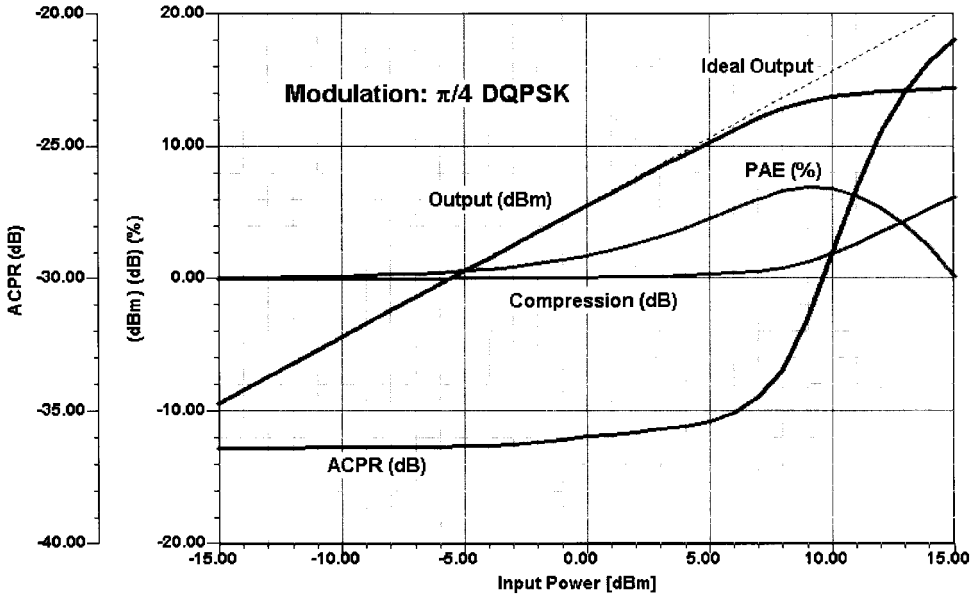


Figure 1-113 Keeping adjacent-channel interference under control can involve a critical trade-off between a wireless transmitter's power-added efficiency (PAE) and ACPR, as shown in this graph of the simulated behavior of the 1-GHz MESFET power amplifier introduced into Figure 1-101. As the amplifier is driven into compression, a peak occurs in the PAE response—in this case, at P_{-1dB}—and the ACPR rises sharply.

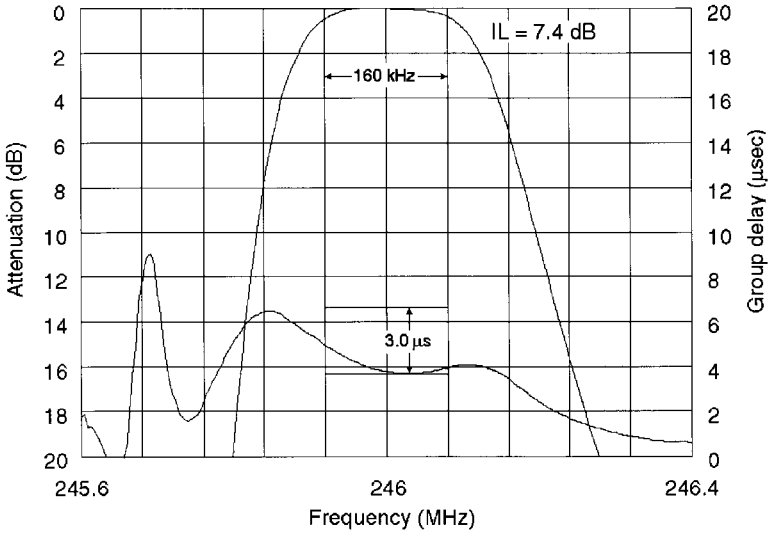


Figure 1-114 Close-in amplitude and group-delay responses for a 246-MHz SAW filter designed for GSM applications [19]. This filter is well within its 3.0-μs differential group-delay specification across its passband (160 kHz at –3 dB); the peaks just outside the passband limits are characteristic of a network's transition-band phase response.

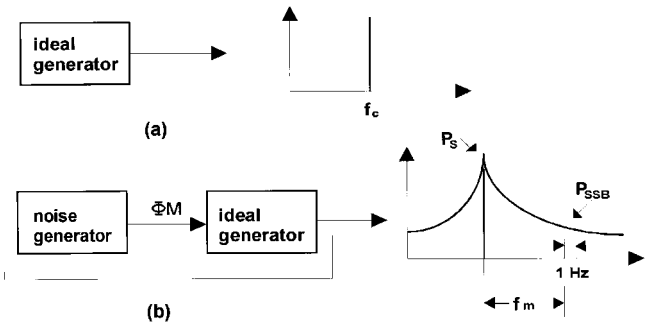


Figure 1-115 SSB phase noise. (a) An ideal signal generator would produce an absolutely pure carrier. (b) A real signal generator acts like an ideal generator driven by a noise generator, producing a noise-modulated carrier.

(the upper or lower) phase-noise sideband, in a 1-Hz bandwidth centered at a specified frequency offset from the oscillator carrier, to the carrier power (Figure 1-117); alternatively, it may be specified in degrees rms. A microwave voltage-controlled oscillator, for instance, might exhibit an SSB phase noise of –95 dBc/Hz at 10 kHz. Oscillator phase noise may manifest itself, through a mechanism known as *reciprocal mixing*, as the emission of unacceptably strong noise outside a transmitter's occupied bandwidth or as an increase in receiver noise floor. Phase noise may also directly introduce phase errors that result in modulation and demodulation errors.

Because the oscillators used for frequency translation in wireless systems are usually embedded in phase-locked loops, their phase-noise characteristics differ from those of "bare" oscillators as shown in Figures 1-118 and 1-119. Figures 1-120 and 1-121 show the measured phase noise of the Rohde & Schwarz SMY and SMIQ signal generators. The SMY is a low-cost signal source, while the SMIQ is a very high performance signal generator capable of being programmed for all digital modulations; therefore, their PLL systems exhibit different phase-noise versus frequency responses as the measured results show.

Reciprocal Mixing. In reciprocal mixing, incoming signals mix with LO sideband energy to produce IF output (Figure 1-122). Because one of the two signals is usually noise, the resulting IF output is usually noise. (Reciprocal mixing effects are not limited to noise;

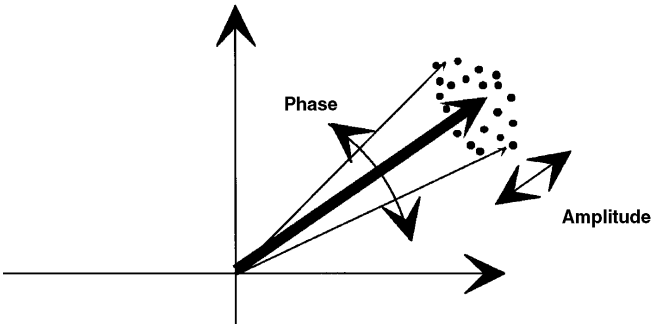
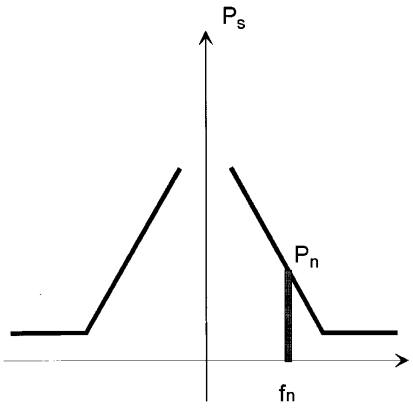


Figure 1-116 Oscillator noise can be split into amplitude and phase components.



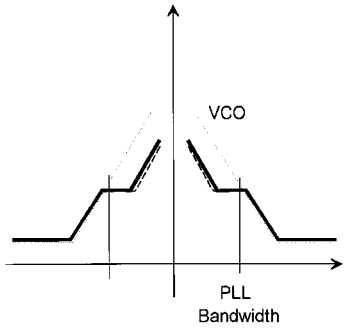
$$L(f) = 10 * \log \frac{P_n}{P_s}$$

P_n = Sideband noise in 1-Hz bandwidth at offset frequency f_n

P_S = Total signal power

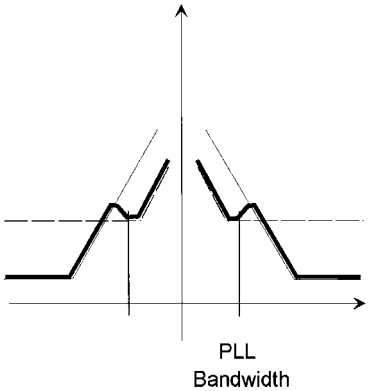
Single-sideband phase noise is expressed in dBc measured in a bandwidth of 1 Hz [dBc(Hz)] at offset frequency f_n.

Figure 1-117 Phase noise calculation.



Phase Noise is controlled by reference phase noise within PLL bandwidth and follows VCO noise outside PLL bandwidth

Figure 1-118 Phase noise of an oscillator controlled by a phase-locked loop.



Improper designed PLL bandwidth

Phase noise increases outside of PLL bandwidth

Figure 1-119 Effect of improper loop-filter design.

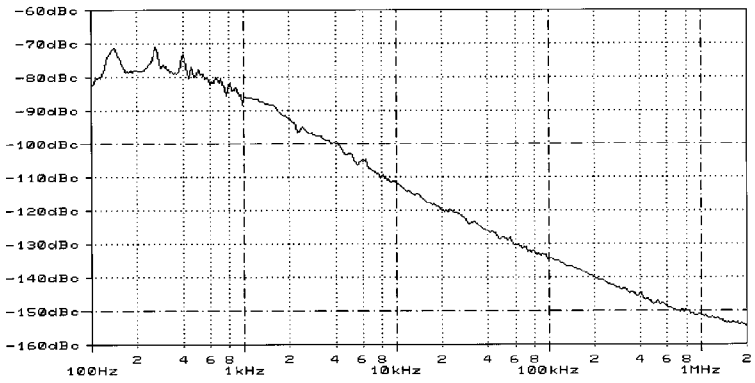


Figure 1-120 Measured phase noise of the Rohde & Schwarz SMY signal generator at 1 GHz. This signal generator has no provision for digital modulation and therefore shows the best possible phase noise in its class.

discrete-frequency oscillator sideband components, such as those resulting from crosstalk to or reference energy on a VCO's control line, or the discrete-frequency spurious signals endemic to direct digital synthesis, can also mix incoming signals to IF.) In practice, the resulting noise-floor increase can compromise the receiver's ability to detect weak signals and achieve a high IMD dynamic range; on the test bench, noise from reciprocal mixing may invalidate desensitization, cross-modulation, and IM testing by obscuring the weak signals that must be measured in making these tests.

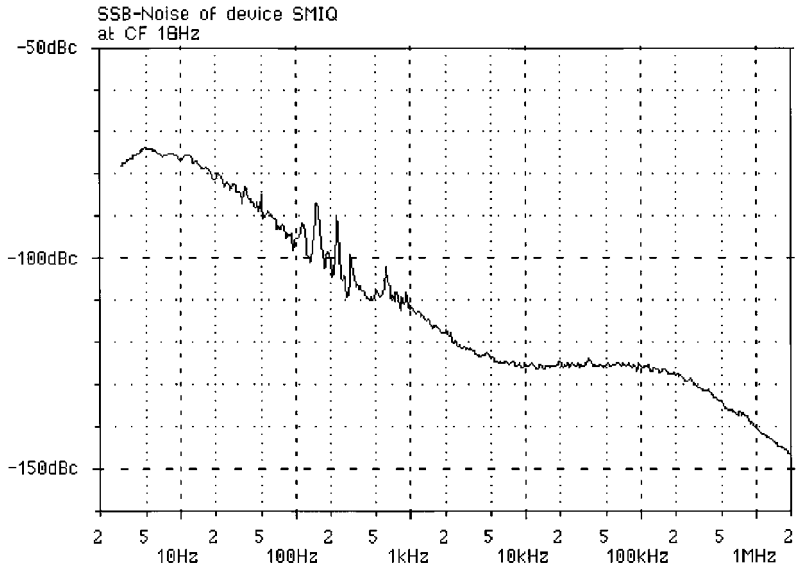


Figure 1-121 Measured phase noise of the Rohde & Schwarz SMIQ signal generator at 1 GHz. This signal generator is optimized for all digital modulation capabilities and can be configured via appropriate programming. Above 10 kHz, the influence of the wideband loop becomes noticeable; above 200 kHz, the resonator *Q* takes over.

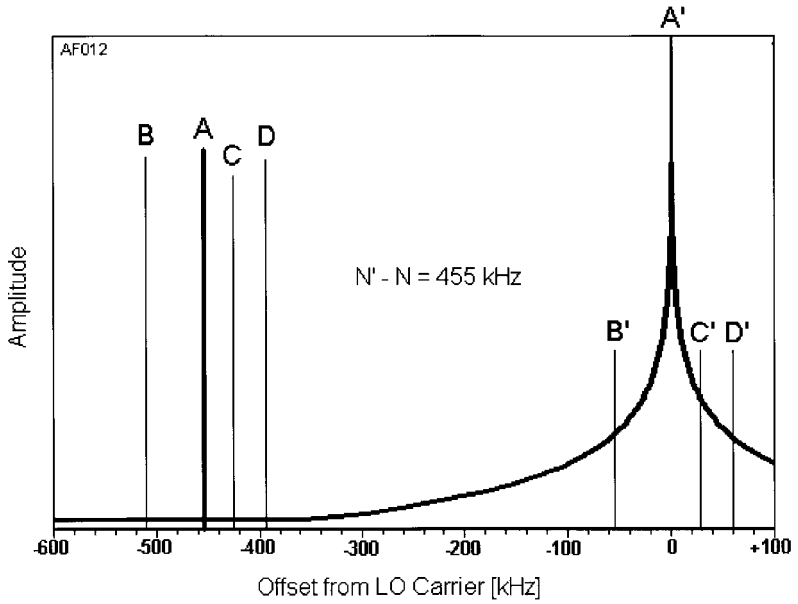


Figure 1-122 Reciprocal mixing occurs when incoming signals mix energy from an oscillator's sidebands to the IF. In this example, the oscillator is tuned so that its carrier, at A′, heterodynes the desired signal, A, to the 455 kHz as intended; at the same time, the undesired signals B, C, and D mix the oscillator noise sideband energy at B′, C′, and D′, respectively, to the IF. Depending on the levels of the interfering signals and the noise sideband energy, the result may be a significant rise in the receiver noise floor.

Figure 1-123 shows a typical arrangement of a dual-conversion receiver with local oscillators. The signal coming from the antenna is filtered by an arrangement of tuned circuits referred to as providing *input selectivity*. For a minimum attenuation in the passband, an operating bandwidth of

$$B = \frac{f}{\sqrt{2}Q_L}$$

is needed. This approximate formula is valid for the insertion loss of about 1 dB due to loaded Q.

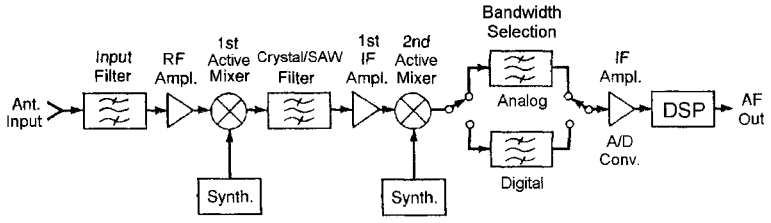


Figure 1-123 Block diagram of an analog/digital receiver showing the signal path from antenna to audio output. No AGC or other auxiliary circuits are shown. This receiver principle can be used for all types of modulation, since the demodulation is done in the DSP block.

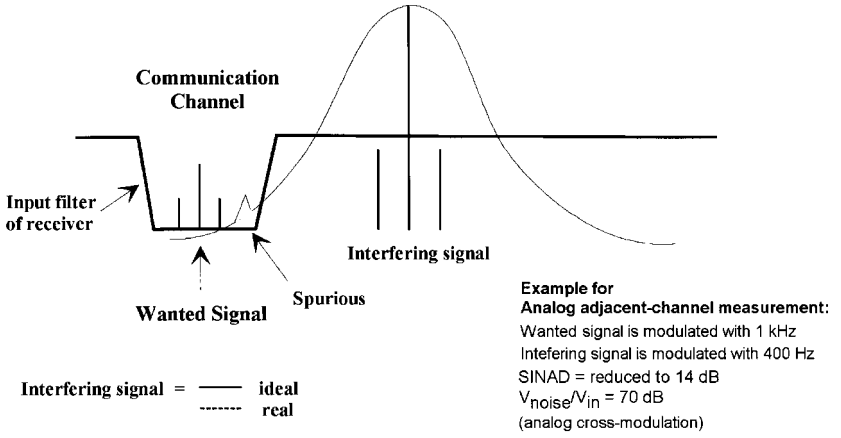
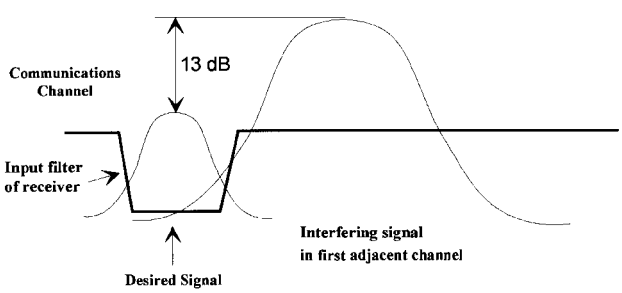


Figure 1-124 Principle of selectivity measurement for analog receivers.

The filter in the first IF is typically either a SAW filter (in the frequency range from 500 MHz to 1 GHz) or a crystal filter (45–120 MHz). Typical insertion loss is 6 dB. Since these resonators have significantly higher Q than LC circuits, the bandwidth for the first IF will vary from ±5 to ±500 kHz. It is now obvious that the first RF filter does not protect the first IF because of its wider bandwidth. For typical communication receivers, IF bandwidths from 150 Hz to 1 MHz are found; for digital modulation, the bandwidth varies roughly from 30 kHz to 1 MHz. Therefore, the IF filter in the second IF has to accommodate these bandwidths; otherwise, the second mixer easily gets into trouble from overloading. This also means that both synthesizer paths must be of low-noise and low-spurious design. The second IF of this arrangement (Figure 1-123) can be either analog or digital, or even zero-IF. In practical terms, there are good reasons for using IFs like 50/3 kHz, as in hi-fi audio equipment such as the Walkman, with DSP processing at this frequency (50/3 kHz) using the low-cost modules found in mass-market consumer products.

Figures 1-124 and 1-125 show the principle of selectivity measurement for both analog and digital signals. The main difference is that the occupied bandwidth for the digital system can be significantly wider, and yet both signals can be interfered with by either a noise synthesizer/first LO or a synthesizer that has unwanted spurious frequencies. Such a spurious signal is shown in Figure 1-124. In the case of Figure 1-124, the analog adjacent-channel



Example of ACP suppression:

	Requirement	Good Generator
NADC		
30kHz from carrier	<-26 dBc	-51 dBc
60kHz "	<-45 dBc	-70 dBc
90kHz "	<-49 dBc	-75 dBc
PDC		
50kHz from carrier	<-45 dBc	-65 dBc
100kHz "	<-60 dBc	-76 dBc

Figure 1-125 Principle of selectivity measurement for digital receivers.

measurement has some of the characteristics of cross-modulation and intermodulation, while in the digital system the problem with the adjacent-channel power suppression is more obvious. Rarely has the concept of adjacent-channel power (ACP) been used with analog systems. Also, to meet the standards, signal generators have to be used that are better, with some headroom, than the required dynamic measurement. We have therefore included in Figure 1-125 the achievable performance for a practical signal generator—in this case, the Rohde & Schwarz SMHU58.

Because reciprocal mixing produces the effect of noise leakage around IF filtering, it plays a role in determining a receiver's dynamic selectivity (Figure 1-126). There is little value in using IF filtering that exhibits more stopband rejection than a value 3–10 dB greater than that which results in an acceptable reciprocal mixing level.

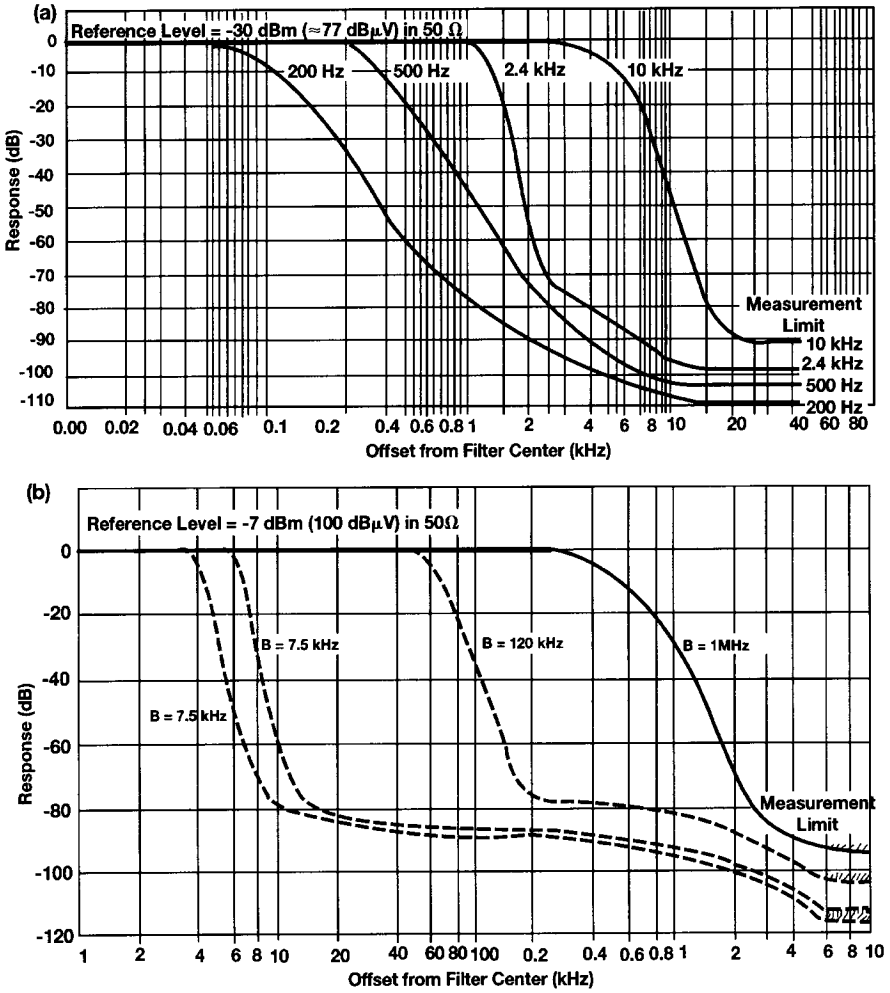


Figure 1-126 Dynamic selectivity versus IF bandwidth for (a) the Rohde & Schwarz ESH2 test receiver (9 kHz to 30 MHz) and (b) the Rohde & Schwarz ESV test receiver (10 MHz to 1 GHz). Reciprocal mixing widens the ESH2's 2.4-kHz response below –70 dB (–100 dBm) at (a) and the ESV's 7.5-, 12-, and 120-kHz responses below approximately –80 dB (–87 dBm) at (b).

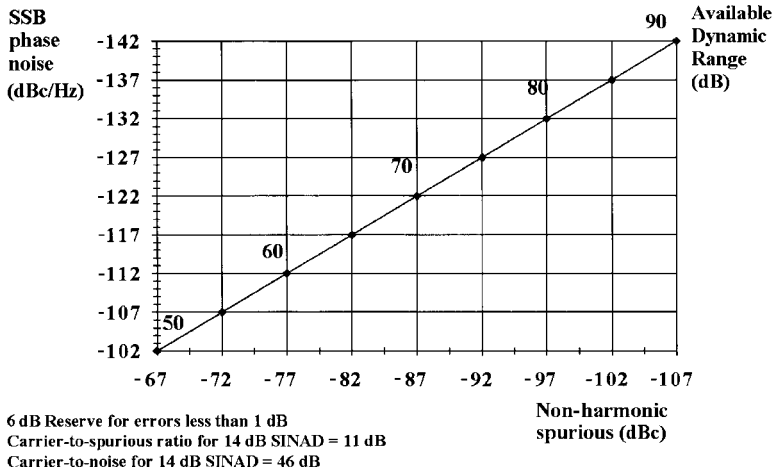


Figure 1-127 This graph shows the available dynamic range, which is determined either by the masking of the unwanted signal by phase noise or by discrete spurious signals. As far as the culprit synthesizer is concerned, it can be either the local oscillator or the effect of a strong adjacent-channel signal that takes over the function of the local oscillator.

Although additional RF selectivity can reduce the number of signals that contribute to the noise, improving the LO's spectral purity is the only effective way to reduce reciprocal mixing noise from *all* signals present at a mixer's RF port.

Factoring in the effect of discrete spurious signals with that of oscillator phase noise can give us the useful dynamic range of which an instrument or receiver is capable (Figure 1-127).

In evaluating the performance of digital wireless systems, we are interested in determining a receiver's resistance to adjacent-channel signals.

Phase Errors. In PM systems, especially those employing digital modulation, oscillator phase noise contributes directly to modulation and demodulation errors by introducing random phase variations that cannot be corrected by phase-locking techniques (Figure 1-128). The greater the number of phase states a modulation scheme entails, the greater its sensitivity to phase noise.

When an output signal is produced by mixing two signals, the resulting phase noise depends on whether the input signals are *correlated*—all referred to the same system clock—or uncorrelated, as shown in Figure 1-129.

Error Vector Magnitude. Several earlier figures (1-104, 1-112, and 1-128) have shown how particular sources of amplitude and/or phase error can shift the values of a digital emission's data states toward decision boundaries, resulting in increased BER due to intersymbol interference. Figures 1-130, 1-131, and 1-132 show three additional sources of such errors.

A figure of merit known as *error vector magnitude* (*EVM*) has been developed as a sensitive indicator of the presence and severity of such errors. An emission's error vector magnitude is the magnitude of the phasor difference as a function of time between an ideal reference signal and the measured transmitted signal after its timing, amplitude, frequency, phase, and dc offset have been modified by circuitry and/or propagation. Figure 1-133 illustrates the EVM concept.

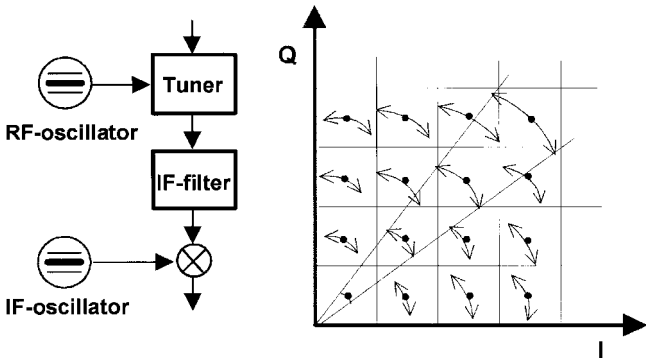


Figure 1-128 Phase noise is critical to digitally modulated communication systems because of the modulation errors it can introduce. Intersymbol interference (ISI), accompanied by a rise in BER, results when state values become so badly error-blurred that they fall into the regions of adjacent states. This drawing depicts only the results of phase errors introduced by phase noise; in actual systems, thermal noise, AM-to-PM conversion, differential group delay, propagation, and other factors may also contribute to the spreading of state amplitude and phase values.

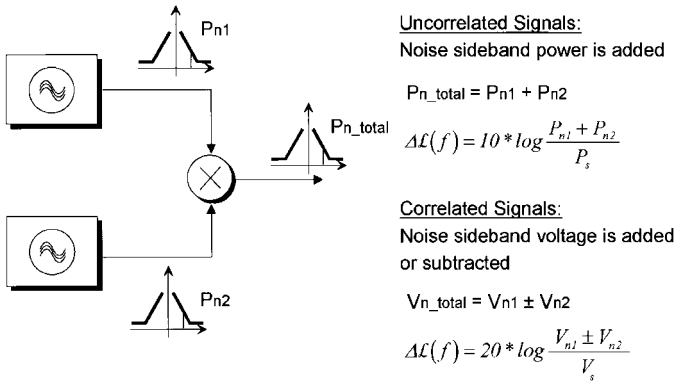


Figure 1-129 The noise sideband power of a signal that results from mixing two signals depends on whether the signals are correlated—referenced to the same system clock—or uncorrelated.

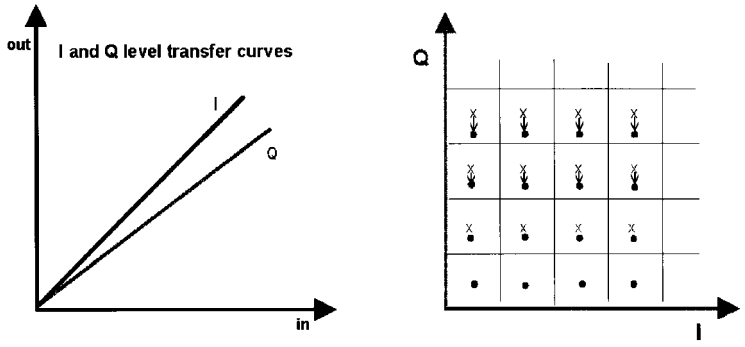


Figure 1-130 Effect of gain imbalance between I and Q channels on a data signal's phase constellation.

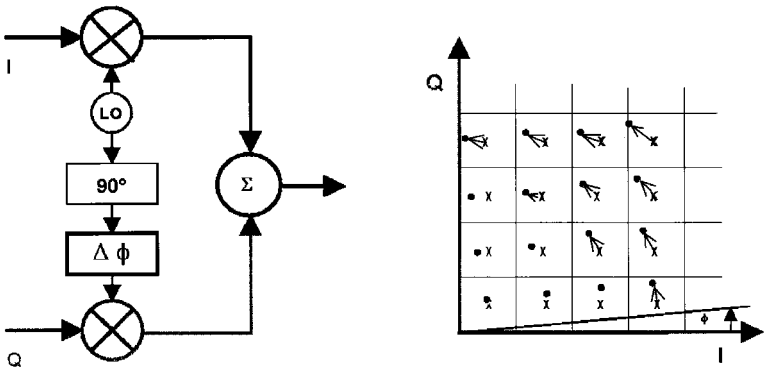


Figure 1-131 Effect of quadrature offset on a data signal's phase constellation.

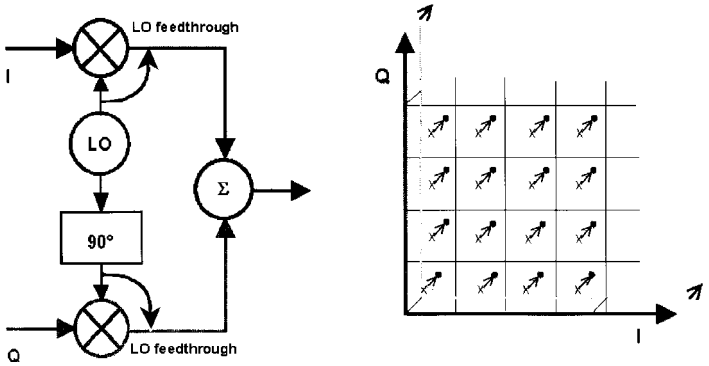


Figure 1-132 Effect of LO-feedthrough-based IQ offset on a data signal's phase constellation.

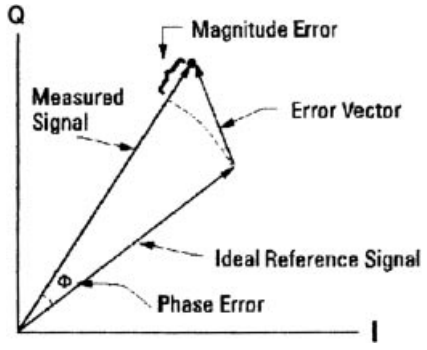


Figure 1-133 Representing errors in a digitally modulated signal's in-phase (I) and quadrature (Q) values relative to the ideal as a single error vector allows us to derive the resulting *error vector magnitude* (*EVM*), a sensitive indicator of transmission quality [20].

1-8 TESTING

1-8-1 Introduction

Testing of digital circuits deviates from the typical analog measurements, and yet the analog measurements are still necessary and related. As an example, the second-generation (2G) cellular telephone standards really have addressed only the transfer of voice and are just beginning to look into the transfer of data. Some services allow limited "e-mails," which are really messages that appear on the handset LCD after a delay. The future 3G standards with adaptive bandwidth will address high-speed data and video.

Since the major use of 2G phones is voice, information such as signal-to-noise ratio as a function of many things is important. In particular, because of the Doppler effect and the use of digital rather than analog signals, where the phase information is significant, the designer ends up using coding schemes for error correction—specifically, forward error correction (FEC). The signal-to-noise ratio as we know it from analog circuits now determines the bit error rate (BER), and its tolerable values depend on the type of modulation used. Because of correlation, it is possible to "rescue" a voice with a bit error rate of 1E – 4, but for higher data rates with little, if any, correlation, we are looking for BERs down to 1E – 9. The actual bit error rate depends on the type of filtering, coding, modulation, and demodulation. In several earlier figures (1-53 and 1-57) we related BER to signal-to-noise ratio; this is a key issue in receiver design.

Another problem in receivers that has to do with the transmitter of a second station is adjacent-channel power ratio (ACPR). Given the fact that a transmitter handling digital modulation delivers its power in pulses, its transmission may affect adjacent channels by producing transient spurious signals similar to what we call *splatter* in analog SSB systems. This is a function of the linearity of the transmitter system all the way out to the antenna and forces most designers to resort to less-efficient Class A operation. As possible alternatives, some researchers are looking into Class D and E modulation, more about which is found in the power amplifier section of Chapter 3.

It is actually not uncommon to do many linear measurements, and then by using correlation equations translate these measured results into their digital equivalents. Therefore, the robustness of the signal as a function of antenna signal at the receiver site, constant or known phase relationships, and high adjacent power ratios will provide good transfer characteristics.

1-8-2 Transmission and Reception Quality

The acoustic transmission and reproduction quality of a mobile phone is the most important characteristic in everyday use. Accurate and reproducible measurements of a cellular phone's frequency response cannot be achieved with static sinewave tones (SINAD measurements) because of coder and decoder algorithms. In this case, test signals simulating the characteristic of the human voice—that is, tones that are harmonic multiples of the fundamental—are required. A so-called vocoder is used to produce the lowest possible data rate. Instead of the actual voice, only the filter and fundamental parameters required for signal reconstruction are transmitted. Particularly in the medium and higher audiofrequency ranges, the static sinusoidal tones become a more or less stochastic output signal. For example, if a tone of approximately 2.5 kHz is applied to the telephone at a constant sound pressure, the amplitude of the signal obtained at the vocoder output varies by approximately 20 dB. In type-approval

tests, where highly accurate measurements are required, the coder and decoder are excluded from the measurement.

Whether the results obtained for the fundamental frequency are favorable depends on how far the values coincide with the clock of the coding algorithm. Through a skillful choice of fundamental frequencies, test signals with overlapping spectral distribution can be generated, giving a sufficient number of test points in subsequent measurements at different fundamental frequencies so that a practically continuous frequency response curve is obtained. Evaluation can be done by means of FFT analysis with a special window function and selection of result bins. The results are sorted and smoothed by the software and display in the form of a frequency response curve. Depending on the measurement, a program calculates the sending or receiving loudness rating in line with CCITT P.79 and shows the result.

Acoustic measurements relevant to GSM 11.10 include:

- Sending frequency response
- Sending loudness rating
- Receiving frequency response
- Receiving loudness rating
- Sidetone masking rating
- Listener sidetone rating
- Echo loss
- Stability margin
- Distortion sending
- Distortion receiving
- Idle channel noise receiving
- Idle channel noise sending

There are two categories of testing. One is a full-compliance test for all channels and all combinations of capabilities; the other is a production tester with evaluation of the typical characteristic data. Both Hewlett-Packard and Rohde & Schwarz are the leading companies in this area, and as the technology develops further, different equipment will be made available. Figure 1-134 shows a Rohde & Schwarz radiocommunication tester capable of evaluating cellular systems. Table 1-13 shows the typical specifications required to make the appropriate cellular telephone measurements.

The test setup must be capable of measuring key IS-95 parameters, using the following:

- Frame errors
- Waveform quality
- Error vector magnitude
- Phase error
- Magnitude error
- Carrier feedthrough
- Frequency accuracy
- Power measurements
- Base-station signaling for mobile testing

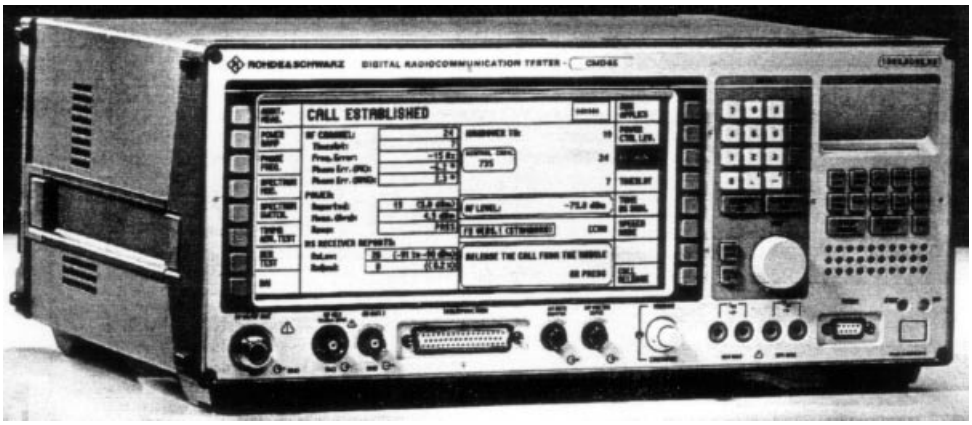


Figure 1-134 The Rohde & Schwarz CMD 65 digital radiocommunication tester can be used to measure all the relevant specifications of cellular telephones.

Table 1-13 Digital radiocommunication tester model CMD 80

SIGNAL GENERATOR	
Frequency	
Range	869–894 MHz
Resolution	1 Hz
Error	Same as timebase
Output level	
RF IN/OUT	–20 to –124 dBm
RF OUT2	0 to –105 dBm
Resolution	0.1 dB
Error (RF IN/OUT)	≤1.5 dB
Modulation	
Carrier suppression	–30 dBc
ANALYZER	
Frequency	
Range	824–849 MHz
Resolution	1 Hz
Error	Same as timebase
Power measurement	
Reference level range	
RF IN/OUT (full scale)	+41 to –28 dBm
RF IN2 (full scale)	0 to –69 dBm
Measurement error, absolute	±1.5 dB
Measurement error, relative	≤0.4 dB (ref. level to –40 dB)
Dynamic range	50 dB below reference level

Table 1-13 (Continued)

Modulation analyzer	
ρ error (25 ± 10) °C	0.003 (for ρ = 0.9–1)
Frequency measurement range	± 3 kHz
Frequency measurement error	Reference ±30 Hz
Timing measurement error	±60 ns
TIMEBASE	
Standard timebase	
Nominal frequency	10 MHz
Frequency drift in temperature range 0–35 °C	$\leq 1.5 \times 10^{-6}$
Frequency aging	$\leq 0.5 \times 10^{-6}$/yr (at 35 °C)
OCXO reference oscillator	
Nominal frequency	10 MHz
Frequency drift in temperature range 0–45 °C	$\leq 1 \times 10^{-7}$
Frequency aging	$\leq 2 \times 10^{-7}$/yr $\leq 0.5 \times 10^{-9}$/day after 30 days of operation
Warm-up time (at 25 °C)	approx. 5 min
Reference frequency inputs/outputs	
Synchronization input (frequency selectable)	100 kHz to 40 MHz, step 10 kHz
Impedance	approx. 100 Ω
Level	0 dBm to TTL
Synchronization output	
Frequency	10 MHz or frequency at synchronization input
Level	TTL, R_{out} = 50 Ω
AWGN generator	
Equivalent noise bandwidth	1.8 MHz typ.
Gain adjustment range	–20 to +6 dB of forward channel power
dc Voltage measurements	
Range	0 to ±30 V
Resolution	10 mV
Error	2% + resolution
dc Current measurements	
Mode	averaging, +peak, –peak
Range	0 to ±10 A
Common mode rejection	±30 V
Shunt resistance	50 mΩ
Resolution for averaging	1mA/10 mA
Resolution for peak	10 mA
Residual indication	<10 mA
Error	<2% + resolution + residual indication

1-8-3 Base-Station Simulation

Simulating a CDMA (IS-95) base station involves:

- Synchronization of mobile
- Registration of mobile
- Incoming/outgoing call origination
- Handoff

During the call, the tester must verify the handset's RF performance and check the correct operation of the basic signaling features. The best testers do not rely on any special test modes in the mobile station, performing their measurements under conditions almost identical to those in a real network. A voice loop-back allows quick verification of the performance of a mobile as it is perceived by the user.

1-8-4 GSM

Measurement, test, and adjustment capabilities for GSM should include:

- Synchronization of mobile phone with base station (which is simulated by CTS)
- Location update
- Call setup (incoming/outgoing)
- Call release (incoming/outgoing)
- Control and measurement of transmitter power
- Handover (channel change)
- Sensitivity, including bit error rate (BER) and raw bit error rate (RBER); limit sensitivity via search routine; RxLev and RxQual
- Phase and frequency error
- Power ramp versus time
- Timing error
- AFC (automatic frequency correction) and RSSI (radio signal strength indication)
- I/Q modulator adjustment
- Echo test (voice test, includes also testing of loudspeaker and microphone)
- Functional test of mobile's keypad through display of dialed number
- Display of IMSI (international mobile subscriber identity), IMEI (international mobile equipment identity), power class, and revision level
- Short message service (SMS)

1-8-5 DECT

Measurement, test, and adjustment capabilities for DECT should include:

- Synchronization of DUT with the CTS
- Call setup
- Call release
- Echo test

Often BER diagrams do not have C/N as abscissa (FIG) but E_b/N_0, which is the energy per information bit E_b referred to the normalized noise power N_0 [1]. C/N describes the ratios in the transmission channel, SNR the signals after the √cos receive filter. The following applies:

$$C/N = SNR + k_{roll\ off}\ [dB]$$

In converting the two quantities, some factors have to be taken into account as shown by equations 1 and 2 on the right.

To determine C/N [dB] or E_b/N_0 [dB], the logarithmic ratios have to be corrected using the following factors:

- Factor for Reed-Solomon FEC*

$$k_{FEC} = 10 \times \lg \frac{188}{204}$$

$$k_{FEC} = -0.3547\ [dB]$$

- Factors for QPSK/QAM modulation

$$k_{QPSK/QAM} = 10 \times \lg(m)$$

Modulation	m	$k_{QPSK/QAM}$ [dB]
QPSK	2	3.0103
16QAM	4	6.0206
64QAM	6	7.7815
256QAM	8	9.0309

- Factor for coding rate (P = 1 for QAM)*

$$k_P = 10 \times \lg(P)$$

Modulation	P	k_P [dB]
QPSK	1/2	–3.0103
	2/3	–1.7609
	3/4	–1.2494
	5/6	–0.7918
	7/8	–0.5799
QAM	1	0

- Factor for √cos roll-off filtering in demodulator/receiver

$$k_{roll\ off} = 10 \times \lg\left(1 - \frac{\alpha}{4}\right)$$

Modulation	α	$k_{roll\ off}$ [dB]
DVB-C	0.15	–0.1660
DVB-S	0.35	–0.3977

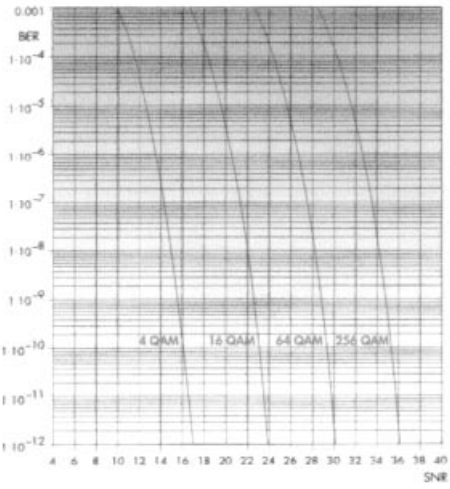
* To be taken into account if E_b only refers to the information bits; not to be taken into account if E_b refers to all bits transmitted (information plus error control bits)

$$C/N = E_b/N_0 + 10 \times \lg\frac{188}{204} + 10 \times \lg(m) + 10 \times \lg(P) - 10 \times \lg\left(1 - \frac{\alpha}{4}\right)\ [dB]$$

Equation 1

$$E_b/N_0 = C/N - 10 \times \lg\frac{188}{204} - 10 \times \lg(m) - 10 \times \lg(P) + 10 \times \lg\left(1 - \frac{\alpha}{4}\right)\ [dB]$$

Equation 2



BER referred to SNR for 4, 16, 64 and 256QAM

The types of correction factor required depend on whether measurement is made

- in the transmission channel,
- before or after Viterbi correction,
- with QAM or QPSK modulation.

Examples of conversion equations –

For **in-channel** measurements with QAM transmission, the following applies:

$$E_b/N_0 = C/N - 10 \times \lg\frac{188}{204} - 10 \times \lg(m)\ [dB]$$

The factors for √cos roll-off filtering and puncturing rate are not needed.

For measurements in the **QAM demodulator**, √cos roll-off filtering filtering has to be taken into account.

$$E_b/N_0 = C/N - 10 \times \lg\frac{188}{204} - 10 \times \lg(m) + 10 \times \lg\left(1 - \frac{\alpha}{4}\right)\ [dB]$$

For measurements in the **satellite demodulator** with QPSK, the equation for determining the BER as a function of E_b/N_0 after Viterbi FEC is as follows:

$$E_b/N_0 = C/N - 10 \times \lg\frac{188}{204} - 10 \times \lg(m) - 10 \times \lg(P) + 10 \times \lg\left(1 - \frac{\alpha}{4}\right)\ [dB]$$

In the latter case all correction factors are included.

Reference:
[1] News From Rohde & Schwarz, Number 163 (1999/II)

Figure 1-135 Conversion of C/N or BER to E_b/N_0.

- Detection and display of RFPI (FP)
- Normal transmit power (NTP)
- Power ramp versus time
- Modulation characteristics versus time
- Frequency offset
- Maximum modulation deviation

- Frequency drift
- Timing (jitter, packet delay)
- Bit error rate (BER), frame error rate (FER)

1-9 CONVERTING C/N OR SNR TO E_b/N_0

Figure 1-135 shows an application note for converting carrier-to-noise ratio (C/N) or SNR to energy per bit/normalized noise power (E_b/N_0).

REFERENCES

1. Zoran Kostic, "Digital Signal Processors in Cellular Radio Communications," *IEEE Commun. Mag.*, **35**(12), 22–35, December 1997.
2. Behzad Razavi, "Recent Advances in RF Integrated Circuits," *IEEE Commun. Mag.*, **35**(12), 36–43, December 1997.
3. William Reuter, "Source and Synthesizer Phase Noise Requirements for QAM Radio Applications," Part 1, *Microwave Prod. Digest*, 36, 38, 50–51, 54–56, January 1999.
4. Peter Vizmuller, *RF Design Guide: Systems, Circuits, and Equations* . Boston: Artech House, 1995.
5. Bernard Sklar, *Digital Communications: Fundamentals and Applications* . New York: Prentice-Hall, 1995.
6. Lin Nan Lee and Jim Mullen, "Third-Generation Cellular Advances Toward Voice and Data," *Wireless Syst. Des.*, **2**(12), 32–38, December 1997.
7. Philips RF Application Note AN1981, "New Low-Power Single-Sideband Circuits."
8. Philips RF Application Note AN1982, "Apply the Oscillator of the SA602 in Low-Power Mixer Applications."
9. Ulrich L. Rohde, Jerry C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*. New York: McGraw-Hill, 1997, p. 58.
10. William E. Sabin and Edgar O. Schoenike, eds., *Single Sideband Systems and Circuits* , 2nd ed. New York: McGraw-Hill, 1995, pp. 123–124.
11. H. T. Friis, "Noise Figures of Radio Receivers," *Proc. IRE*, **32**(7), 419–422, July 1944.
12. Mark Williams, Fred Bonn, Curtis Gong, and Tina Quach, "GaAs RF ICs Target 2.4-GHz Frequency Band," *Microwaves & RF*, **33**(7), 111–118, July 1994.
13. Jack Smith, *Modern Communication Circuits* , 2nd ed. New York: McGraw-Hill, 1997, pp. 93–99.
14. S. M. Perlow, "Third-Order Distortion in Amplifiers and Mixers," *RCA Rev.*, 234–265, June 1976.
15. Edward C. Jordan, ed., *Reference Data for Engineers* , 7th ed. Indianapolis: Howard W. Sams & Co, 1985, pp. 12-33–12-35.
16. Stephen A. Maas, *Nonlinear Microwave Circuits* . Norwood, MA: Artech House, 1988, pp. 18 and 169.
17. Joseph Staudinger, "Specifying Power Amplifier Linearity via Intermodulation Distortion and Channel Spectral Regrowth," *Appl. Microwave & Wireless* , **9**(4), 62, 64, 66, 68, 70, July/August 1997.

18. Michael S. Heutmaker, Eleanor Wu, and John R. Welch, "Envelope Distortion Models with Memory Improve Some Amplifier Spectral Regrowth Predictions," *Appl. Microwave & Wireless*, **9**(4), 72, 74, 76, 78, July/August 1997.
19. Toko SF246ZA-019 filter specifications, available via the Web address http://www.tokoam.com/saw/applications.html.
20. *Using Vector Modulation Analysis in the Integration, Troubleshooting and Design of Digit al RF Communications Systems*, Product Note HP89400-8 (Hewlett-Packard Co., 1994, available via the Web address http://www.tmo.hp.com/tmo/Notes/English/5091-8687E.html).
21. George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990, pp. 275–278.
22. William G. Scanlon, "Health Aspects of Low-Level Exposure to RF Electromagnetic Waves," *RF Des.*, **22**(7), 40, 44, 46, 48, 50, 52, 54, 56, 64, 66, 68, July 1999.
23. Van Valkenburg, editor-in-chief, *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, 8th ed. Carmel, CA: Prentice-Hall Computer Publishing, 1993, pp. 25-19 –25-20.

Further Reading

Walid Y. Ali-Ahmad, "Improving the Receiver Intercept Point Using Selectivity," *RF Des.*, **20**(12), 22, 24, 26, 28, 30, December 1997.

"AM Signal Analysis in a Dual Down-Conversion Receiver," *Symphony Examples Volume*, Ansoft Corporation, 1999.

Leon W. Couch II, *Digital and Analog Communications Systems*. New York: Macmillan, 1990.

Paul Danzer, ed., *The ARRL Handbook for Radio Amateurs*, 75th ed. Newington, CT: American Radio Relay League, 1997, p. 26.44.

Nikhil Deshpande, Steven Stanton, and Michael Hurst, "ACPR Specs Place Demands on WCDMA Base-Station Amplifiers," *Wireless Systems Design*, **4**(8), 15–22, August 1999.

Wing S. Djen and Prasanna M. Shah, "Implementation of a 900 MHz Transmitter System Using Highly Integrated ASIC," *Proceedings of the 44th IEEE Vehicular Technology Conference*, June 1994, pp. 1341–1345.

Dimitrios Efstathiou, "π/4-DPQSK Data-Modulation Schemes Help Enable 3G Systems," *Wireless Syst. Des.*, **4**(1), 27–32, January 1999.

"Double Down Conversion Communications Receiver," *Symphony Examples Volume*, Ansoft Corporation, 1999.

Kamilo Feher, "FQPSK Transceivers Double the Spectral Efficiency of Wireless and Telemetry Systems," *Appl. Microwave & Wireless*, **10**(5), 73–79, June 1998.

David K. Lovelace, Jesus L. Finol, and Jeffrey C. Durec, "Sub-Micron Silicon RF IC Technologies: An Overview," 1998 IEEE MTT-S International Microwave Symposium, 1859–1862, June 1998.

William Gosling, ed., *Radio Receivers*. London: Peter Peregrinus Ltd., 1986, p. 29.

Helge O. Granberg, "Power MOSFETs Versus Bipolar Transistors," Motorola Semiconductor Application Note AN860.

GSM Standard, "Physical Layer on the Radio-Path," July 1988.

W. H. Hayward, *Introduction to Radio Frequency Design*. Newington, CT: American Radio Relay League, 1994 (original edition: 1982).

Gene Heftman, "Cellular ICs Move Toward 3G Wireless—Gingerly," *Microwaves & RF*, **38**(2), 31–38, February 1999.

Roy Hejhall, "VHF MOS Power Applications," Motorola Semiconductor Application Note AN878.

IS-54-B, EIA/TIA, "Cellular System Dual-Mode Mobile Station-Base Station Compatibility Standard," April 1992.

M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems*. New York: Plenum Press, 1992.

Lawrence E. Larson, ed., *RF and Microwave Circuit Design for Wireless Communications*. Boston: Artech House, 1996.

Chia-Liang Liu and Kamilo Feher, "π/4-QPSK MODEMS for Satellite Sound/Data Broadcast Systems," *IEEE Trans. Broadcasting*, 1–8, March 1991.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

David Lovelace, Jeff Durec, and Mike McGinn, "A Silicon Bipolar Transmitter IC for 900MHz ISM Band Applications," *Proceedings of the Sixth Annual WIRELESS Symposium*, Santa Clara, CA, February 9–13, 1998, pp. 480–484.

Brian J. Minnis, *Designing Microwave Circuits by Exact Synthesis*. Boston: Artech House, 1996.

Simon Morris, "DSP and ASIC Partitioning Offer Options for cdma2000 Base-Station Designs," *Wireless Syst. Des.*, **3**(11), 17–25, November 1998.

Mohamed K. Nezami, "Evaluate the Impact of Phase Noise on Receiver Performance," *Microwaves & RF*, **37**(5), 165–172, May 1998.

Philips Semiconductors, "PCD5070 GSM Baseband Interface," preliminary specification, September 1992.

Ulrich L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design*. New York: Wiley, 1997.

William E. Sabin, "Use of Mixers in HF Upconversion Receivers/Exciters," *Professional Program Session Record 24*, Wescon/81, San Francisco, CA, September 15–17, 1981, pp. 24/4-3 to 24/4-4.

Jan Sevenhans and Bart Verstraeten, "10 Years of Silicon Integration," *Microwave Eng. Eur.*, 35–40, May 1999.

Mark Simpson, "Low-Noise Synthesizers Applied to QAM-Modulated Digital Radios," *Proceedings of the Sixth Annual Wireless Symposium*, Santa Clara, CA, February 9–13, 1998, pp. 182–190.

Raymond Steele, *Mobile Radio Communications*. New York: IEEE Press, 1992.

Tim Stroud, "Tutorial on All-Digital Modulation Techniques," *RF Des.*, **20**(6), 52–58, June 1997.

Jan Uddenfeldt, "Digital Cellular—Its Roots and Its Future," *Proc. IEEE*, **86**(7), 1319–1324, July 1998.

Robert T. Van Brunt, "The Next Generation of Wireless: Coming Soon, If Not By 2000," *Microwave Prod. Digest*, 10, 30, 34, April 1999.

Peter Vizmuller, *RF Design Guide: Systems, Circuits and Equations*. Norwood, MA: Artech House, 1995. A good general RF engineering handbook in compact format.

Dave Vondrian, "Noise Figure Measurement: Corrections Related to Match and Gain," *Microwave J.*, **42**(3), 22–38, March 1999.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

2

MODELS FOR ACTIVE DEVICES

Because we are dealing with amplifiers, both small-signal and power, oscillators, and active mixers, active devices will play a major role in circuit design. At our frequency range of interest (about 1000 MHz and above), distributed elements predominate, and even in device modeling we must consider some distributed effects such as packaging.

As we dive into providing functional explanations of the active devices used in practical circuits, there is some temptation not to provide enough information on semiconductor behavior. Because manufacturers of microwave products will use the services of a foundry or specify semiconductors from their suppliers, we will cover the device operation in depth. That said, for the beginner looking into designing circuits from high frequencies to low microwave frequencies, we will provide examples of a variety of devices unique to this frequency range. Some treatment of device physics will therefore be useful in providing a basic understanding of the nonlinear behavior of active devices and modeling in preparing for practical circuits using a computer-aided-design (CAD) tool. For CAD applications, we assume the user will have access to nonlinear CAD software, such as a high-performance SPICE program or a harmonic-balance program, such as Ansoft Serenade 8.0 or its derivatives. It is interesting that most semiconductor houses today still only offer very limited nonlinear data sets, while in the linear area a large body of useful data is made available, mostly on CDs. What CAD does for us—and this applies to SPICE and harmonic-balance programs alike—is predict the dc bias point of a device in the context of its surrounding circuitry. Configuring the simulation to drive the device into large-signal operation provides us with insight into the various forms of distortion and compression.

In evaluating active devices for RF circuit design, we have a few different technologies to consider. First, we will examine the semiconductor junction diode—a basic *pn* junction—and venture from there to the bipolar transistor. An investigation of metal-oxide semiconductor (MOS) devices, another important set of semiconductors for medium frequencies, will follow. We will also examine the gallium arsenide metal-gate Schottky field-effect transistor (GaAs MESFET) and members of its family. Since critical large-signal parameters for some devices may not be available from their manufacturers, we will examine

parameter-extraction techniques and problems. For various device types, we will provide insight into the noise properties of the device.

2-1 DIODES*

The diode model contains a nonlinear current source that follows the Shockley equation:

$$\text{Current} = \text{IS}\left(e^{V_j/\text{N}V_t} - 1\right) \tag{2-1}$$

where V_j = voltage across the junction
V_t = thermal voltage (= kT/q)

These values, with the model parameters IS and N, are used to model the current–voltage effects of the semiconductor junction. This does not include the nonideal operation of real diodes. For example, at low currents (less than 1 nA), other semiconductor processes that increase the flow of currents become noticeable.

By setting IS to different values, we can obtain the characteristics of other devices, such as a Schottky barrier diode, or a silicon diffused-junction diode. High-current effects are modeled, grossly, by including a series resistance that is intended to combine the effects of bulk resistance (the material on each side of the junction) and high-level injection. At high currents, the observed diode current stops following the Shockley form

$$I_{\text{forward}} = \text{IS} \times e^{V_j/\text{N}V_t} \tag{2-2}$$

and approaches a modified form

$$I_{\text{forward}} = \text{IS} \times e^{V_j/2\text{N}V_t} \tag{2-3}$$

2-1-1 Large-Signal Diode Model

Three diode models are used in the industry:

- Microwave diode model
- PIN diode model
- Enhanced SPICE diode model

Figure 2-1 shows the large-signal microwave diode model. Its keywords appear in Table 2-1. This model can also be used to simulate varactor diodes.

Table 2-2 lists SPICE parameters for a selection of Schottky mixer diodes by Alpha.

In most cases, the diode capacitance is modeled by a voltage-dependent capacitor, which is connected in parallel with the nonlinear current generator described previously, to represent the charge-storage effects of the junction. There are two components to this charge

- The reverse-voltage capacitive effect of the depletion region
- The forward-voltage charge represented by mobile carriers in the diode junction

*Portions of this chapter's diode coverage are based on material in Alpha Industries Semiconductor Application Reports 80800, 80200, and 80500. Used with permission.

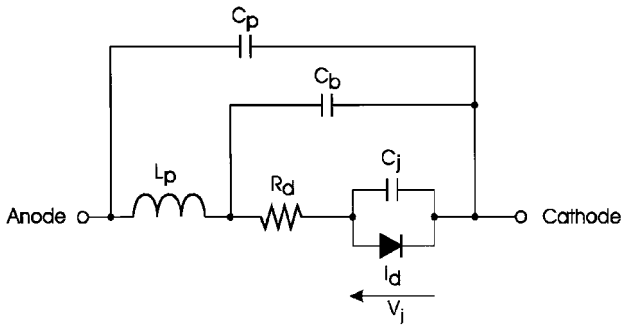


Figure 2-1 The large-signal microwave diode model. This model is temperature dependent.

Reverse-voltage capacitance follows the simple approximation that the depletion region (the area of the junction that is depleted of carriers) serves as the gap between the "plates" of a capacitor. This region varies in thickness, and therefore the capacitance varies with applied voltage. For a step (abrupt) junction, or linearly graded junction, the capacitance approximation is

Table 2-1 Large-signal microwave diode model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
JS	Saturation current	ampere	0
ALFA	Slope factor of conduction current	/volt	38.696
JB	Breakdown saturation current	ampere	10 mA
VB	Breakdown voltage	volt	$-\infty$
E	Power-law parameter of breakdown current	—	10.0
CT0	Zero-bias depletion capacitance	farad	0
FI	Built-in barrier potential	volt	0.8
GAMA	Capacitance power-law parameter	0.5	
GC1	Varactor capacitance polynomial coefficient 1	/volt	0.0
GC2	Varactor capacitance polynomial coefficient 2	/volt^2	0.0
GC3	Varactor capacitance polynomial coefficient 3	/volt^3	0.0
CD0	Zero-bias diffusion capacitance (*pn* diodes)	farad	0
AFAC	Slope factor of diffusion capacitance	/volt	38.696
R0	Bias-dependent part of series resistance in forward-bias condition	ohm	0
T	Intrinsic time constant of depletion layer for abrupt-junction diodes	second	0
KF	Flicker noise coefficient	—	0.0
AF	Flicker noise exponent	—	1.0
FCP	Flicker noise frequency shape factor	—	1.0
AREA	Area multiplier	—	1.0
	Extrinsic Model		
CP	Package parasitic capacitance	farad	0.0
CB	Beam-lead parasitic capacitance	farad	0.0
LP	Package parasitic inductance	henry	0.0

Table 2-2 Diode SPICE parameters

Parameter	Unit	SMS1546	SMS3922	SMS3923	SMS3924	SMS3926	SMS3927	SMS3928	SMS7621	SMS7630
I_S	A	3E–7	3E–8	5E–9	2E–11	2.5E–07	1.3E–09	9E–13	4E–8	5E–06
R_S	Ω	4	9	11	11	4	4	4	12	30
n	—	1.04	1.08	1.05	1.08	1.04	1.04	1.04	1.05	1.05
T_d	S	1E–11	8E–11	8E–11	8E–11	1E–11	1E–11	1E–11	1E–11	1E–11
C_{J0}	PF	0.38	0.9	0.93	1.6	0.42	0.39	0.39	0.1	0.14
m	—	0.36	0.26	0.24	0.4	0.32	0.37	0.42	0.35	0.4
E_G	EV	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
X_{ti}	—	2	2	2	2	2	2	2	2	2
F_C	—	0.5	0.5	0.5	0.5	0.5	0.5	0.5	5	0.5
B_J	V	3	20	46	100	2	3	4	3	2
I_{BV}	A	1E–5	1E–5	1E–5	1E–5	1.00E–05	1.00E–05	1.00E–05	1E–5	0.0001
V_J	—	0.51	0.65	0.15	0.84	0.495	0.595	0.800	0.51	0.34

$$\text{Capacitance} = \frac{C_{J0}}{\left(1 - V_j/\phi\right)^M} \tag{2-4}$$

where C_{J0} is the zero-bias value, ϕ (phi) is the junction barrier potential, and M is the grading coefficient that varies ($\frac{1}{2}$ is used for step junctions and $\frac{1}{3}$ is used for linearly graded junctions, and most junctions are somewhere in between).

There is often confusion about the barrier potential, which appears in the capacitance equation. From capacitance measurements, ϕ (model parameter VJ, and not to be confused with V_j in the equations) takes on a value of nearly 0.7 volt for regular (silicon) junction diodes and a range of 0.58–0.85 volt for various Schottky barrier diodes. This value is sometimes confused with the forward-current voltage drop of the diode or the energy gap of the material; it is *neither* of these.

Varying M generates a variety of reverse-bias capacitance characteristics. Inspection of the capacitance formula reveals that it predicts infinite capacitance for a forward bias, which is not the case for a real junction. Several depletion-capacitance formulas have been proposed that more correctly fit observed operation; however, SPICE uses a simple approach: For forward biases beyond some fraction (set by the parameter FC) of the value for ϕ, the capacitance is calculated as the linear extrapolation of the capacitance at the departure. This provides a continuous numerical result and does not affect circuit operation significantly because, for forward bias, the device capacitance is normally dominated by diffusion capacitance.

The diffusion charge (and therefore the capacitance) varies with forward current and is simply modeled as a transit time (model parameter TT) for the carriers to cross the diffusion region of the junction. The total charge is

$$\text{Diffusion charge} = \text{Device current} \times \text{Transit time} \tag{2-5}$$

and capacitance is the derivative, with respect to bias, of this:

$$\text{Diffusion capacitance} = \text{TT}\,\frac{\text{IS}}{\text{N}V_t}\,e^{V_j/\text{N}V_t} \tag{2-6}$$

Diffusion charge manifests itself as the *storage time* of a switching diode, which is the time required to discharge the diffusion charge in the junction, which must happen before the junction can be reverse-biased (switched off). Storage time is normally specified as the time to discharge the junction so that it is supporting only a fraction (typically 10%) of the initial reverse current. First, a forward current is supplied to the device to charge the junction. Then, as quickly as possible, a reverse current is supplied to the device. Internally, the junction is still forward-biased to a voltage nearly the same as before the switch in current; the junction is still conducting at the forward-current rate. This internal current adds to the external current as the total current discharging the junction. As the junction voltage decreases, the internal current falls off exponentially (according to the Shockley equation). This system is a relatively simple differential equation that can be solved to an explicit equation for the TT parameter (assuming complete discharge) as follows:

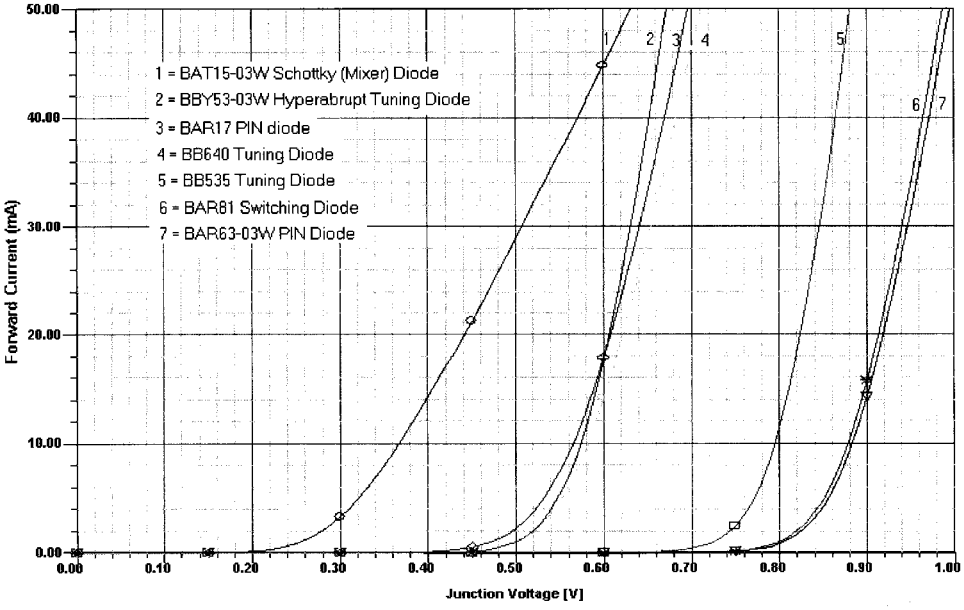


Figure 2-2 The dc *I–V* curves for seven diodes, showing the various barrier voltages that result from different doping profiles.

$$\text{Transit time} = \frac{\text{Storage time}}{\ln\left(\dfrac{I_F - I_R}{-I_R}\right)} \tag{2-7}$$

The diffusion charge dominates the reverse recovery characteristic of the diode. During the last part of the recovery, as the junction becomes reverse-biased, the depletion capacitance dominates. This causes the small tail at the end of the discharge cycle. Total capacitance is taken to be the sum of these capacitances: the depletion approximation dominates for reverse bias as the device current is small, and the diffusion proximity dominates for forward bias as the device current is large.

A special case for diode application is the switching diode, and its description and application will be part of a later chapter. Figure 2-2 shows the dc *I–V* curves, which indicate the different voltage potentials that are a result of the different doping profiles.

2-1-2 Mixer and Detector Diodes

Electrical Characteristics and Physics of Schottky Barriers. Schottky barrier diodes differ from junction diodes in that current flow involves only one type of carrier instead of both types. That is, in *n*-type Schottky barrier diodes the forward current consists of holes flowing from the *n*-type material into the metal.

Diode action results from a contact potential set up between the metal and the semiconductor, similar to the voltage between the two metals in a thermocouple. When metal is brought into contact with an *n*-type semiconductor (during fabrication of the chip), electronics diffuse out of the semiconductor into the metal, leaving a region under the contact that

has no free electrons ("depletion layer"). This region contains donor atoms that are positively charged (because each lost its excess electron), and this charge makes the semiconductor positive with respect to the metal. Diffusion continues until the semiconductor is so positive with respect to the metal that no more electrons can go into the metal. The internal voltage difference between the metal and the semiconductor is called the contact potential and is usually in the range of 0.3–0.8 volt for typical Schottky diodes. A cross section is shown in Figure 2-3.

When a positive voltage is applied to the metal, the internal voltage is reduced, and electrons can flow into the metal. The process is similar to thermionic emission of electrons from the hot cathode of a vacuum tube, except that the electrons are "escaping" into a metal instead of into a vacuum. Unlike the vacuum tube case, room temperature is "hot" enough for this to happen if enough voltage is applied. However, only those electronics whose thermal energy happens to be many times the average can escape, and these "hot electrons" account for all the forward current from the semiconductor into the metal.

One important thing to note is that there is no flow of minority carriers from the metal into the semiconductor and thus no neutral plasma of holes and electrons is formed. Therefore, if the forward voltage is removed, current stops "instantly," and reverse voltage can be established in a few picoseconds. There is no delay effect to charge storage as in junction diodes. This accounts for the exclusive use of barrier diodes in microwave mixers, where the diode must switch conductance states at microwave oscillator rates.

The current–voltage relationship for a barrier diode is described by the Richardson equation (which also applies to thermionic emission from a cathode). The derivation is given in many textbooks (e.g., Sze).

$$I = AA_{\mathrm{RC}}T^2 \exp\left(-\frac{q\phi_B}{kT}\right)\left[\exp\left(\frac{qV_J}{kT}\right) - M\right] \tag{2-8}$$

where A = area (cm^2)

A_{RC} = modified Richardson constant ($\mathrm{A}/K^2 \cdot \mathrm{cm}^2$)

k = Boltzmann's constant

T = absolute temperature (K)

ϕ_{B} = barrier height in volts

V_J = external voltage across the depletion layer (positive for external voltage) = $V - IR_S$

R_S = series resistance

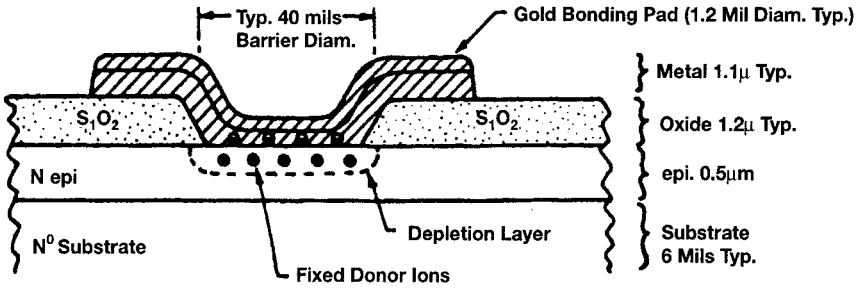


Figure 2-3 Schottky diode chip cross section.

M = avalanche multiplication factor
I = diode current in amperes (positive forward current)

The barrier height, ϕ_B, is typically a few tenths of a volt higher than the contact potential, ϕ_C (about 0.15 V higher than ϕ_C for silicon). This equation agrees well with experimental data for diodes without surface leakage, but it is difficult to use because A_{RC}, ϕ_B, and M are all dependent on applied voltage.

The major cause for variation in ϕ_B with voltage is the "image effect," in which the barrier height is lowered as the electric field near the metal is increased, especially at the edges. A better equation for circuit designers to use is one in which all parameters are independent of voltage and current. The simplest one that agrees reasonably well with Richardson's equation is

$$I = I_S\left[\exp\left(\frac{V_J}{0.028}\right) - 1 + \frac{K}{\left(1 - V_B/V\right)}\right] \tag{2-9}$$

where I_S = saturation current (a temperature-dependent quantity)
0.028 = nkT/q at room temperature (n = 1.08)
n = forward slope factor (derived from the variation of ϕ_B with forward voltage)
K = reverse slope factor (expressing the variation of ϕ_B with reverse voltage)
V_B = breakdown voltage (the voltage at which M = 1)

As before, V and I are considered positive for forward bias and negative for reverse bias.

Typical ranges for these parameters for microwave Schottky and point-contact mixer diodes are:

I_S = 10^{-12} to 10^{-5} A
n = 1.04–1.10
R_S = 2–20 Ω
K = 8–100
V_B = 2–20 V

The quantities I_S and 0.028 are strongly temperature dependent, while both R_S and V_B increase with temperature to a slight degree. R_S increases with current at high current levels (due to carrier velocity saturation) but is essentially independent of current at 10 mA and below for mixer diodes. Thus, for normal mixer and detector operation, R_S can be considered constant.

Agreement between Eqs. (2-8) and (2-9) is not perfect but Eq. (2-9) is much easier to use and is preferred by most circuit designers. A comparison of the two equations near zero bias gives the following relationship between zero-bias barrier height, ϕ_0, and saturation current:

$$I_S = AA_{RC}T^2\exp - \frac{q\phi_0}{kT}$$

$$\cong \left(\frac{10^7 A}{\text{cm}^2}\right)A\exp\left(-\frac{\phi_0}{0.026}\right) \quad \text{(for } n \text{ silicon at room temperature)} \tag{2-10}$$

Small-Signal Parameters. By combining Eqs. (2-9) and (2-10), the values of the parameters in Eq. (2-9) can be derived from a few simple measurements. Many specific equations can be derived, but the following are commonly used for production measurements:

$$R_S = \frac{V_{F10} - V_{F1} - 0.065}{0.009} \quad (\text{for } n = 1.08) \tag{2-11}$$

$$\phi_0 = \frac{V_{F1} - 0.001R_S + 0.28 + 0.12\log_{10} D}{1.08} \tag{2-12}$$

$$n = \frac{V_{F1} - V_{F0.1} - 0.0009R_S}{0.060} \tag{2-13}$$

$$K = \left(\frac{I_{R1}}{I_S} - 1\right)V_B \tag{2-14}$$

$$I_S = \exp\left(-\frac{V_{F1}}{0.028} + \frac{R_S}{28}\right) \quad (\text{in mA}) \tag{2-15}$$

where $V_{F0.1}$, V_{F1}, and V_{F10} are the forward voltages at 0.1, 1, and 10 mA, respectively, and I_{R1} is the reverse current at 1 volt. (The derivation of these equations requires that I_S be small compared to 0.1 mA.) The quantity D is the diameter of the metal–silicon contact in mils.

Measuring V_{F1} at 1 mA and 10 mA instead of some other current levels leads to the best accuracy for typical mixer diodes.

The total dynamic resistance for a forward-biased diode is given by

$$R_T = \frac{dV}{dI} = R_S + \frac{nkT}{q(I + I_S)} = R_S + R_B \tag{2-16}$$

and

$$R_B = \frac{28}{I + I_S} \quad \text{at room temperature (with } I \text{ and } I_S \text{ in mA, } n = 1.08) \tag{2-17}$$

This equation is also good at zero bias (unless K is very large or there is significant surface leakage). That is,

$$R_0 = R_S + \frac{28}{I_S} \tag{2-18}$$

For reverse voltages of a few volts, the dynamic resistance is dominated by the K term:

$$R_R = \text{Reverse resistance} = \frac{dV}{dI} \cong \frac{V_B}{KI_S} \tag{2-19}$$

For typical values of I_S, R_0 is larger than 5000 Ω and R_R is larger than 100 kΩ. For some zero-bias Schottky applications, it is desirable for R_0 to be made smaller than this.

The factors that determine R_S are (1) the thickness of the epitaxial layer, (2) the epitaxial layer doping level (N_D), (3) the barrier diameter, (4) the substrate resistivity ("spreading resistance"), (5) the contact resistances of the metals used for the barrier and the substrate contact, and (6) the resistance associated with the bonding wire or whisker. The barrier height is about 0.15 volt higher than the contact potential between the barrier metal and the semiconductor and is influenced by the method used to apply the metal, conditions at the edge of the junction, and the doping level. Saturation current depends on barrier height, junction area, and temperature; and the slope factors, n and K, depend on doping level, punchthrough voltage, and edge conditions.

Junction Capacitance. The capacitance of a Schottky barrier chip results mainly from two sources: the depletion layer under the metal–semiconductor contact and the capacitance of the oxide layer under the bonding pad (the so-called overlay capacitance). The bonding pad is required because the typical Schottky barrier diameter is so small that it is impractical to bond directly to the metal on the junction. If the semiconductor epitaxial layer is uniformly doped, the capacitance–voltage characteristic is similar to that of a textbook "abrupt-junction" diode.

$$C_J = \frac{\varepsilon_S \varepsilon_0 A'}{X_D} + C_0 \tag{2-20}$$

$$X_D = \sqrt{\frac{2\varepsilon_S \varepsilon_0 (\phi_C - V)}{qN}} \tag{2-21}$$

where ϕ = contact potential
C_0 = overlay (bonding pad) capacitance
ε_S = dielectric constant of the semiconductor ($\cong$ 12 for Si or GaAs)
N = doping level for the epitaxial layer
A' = effective contact area, including fringing corrections

In practical terms, the capacitance can be related to the 0-V barrier capacitance defined by

$$C_{B0} = \frac{\varepsilon_S \varepsilon_0 A'}{X_{D0}} \tag{2-22}$$

where

$$X_{D0} = \sqrt{\left(\frac{1.3 \times 10^{15}}{N_D}\right)\phi_C} \quad \text{(in } \mu\text{m)} \tag{2-23}$$

The resulting C–V relationship can be written

$$C_J = \frac{C_{B0}}{\sqrt{1 - V/\phi_C}} + C_0 \tag{2-24}$$

The contact potential, ϕ_C, is related to the barrier height as follows:

$$\phi_C = \phi_B - 0.026\left[1 + L_n\left(\frac{N_C}{N}\right)\right] \tag{2-25}$$

The theoretical meaning of these terms can be clarified by looking at Figure 2-4.

Parameter Trade-Offs

Barrier Height. The barrier height of a Schottky diode is important because it directly determines the forward voltage. In order to get a good noise figure, the LO drive voltage, V_L, must be large compared to V_T, which is essentially V_{F1}. Normally, it is best to have a low

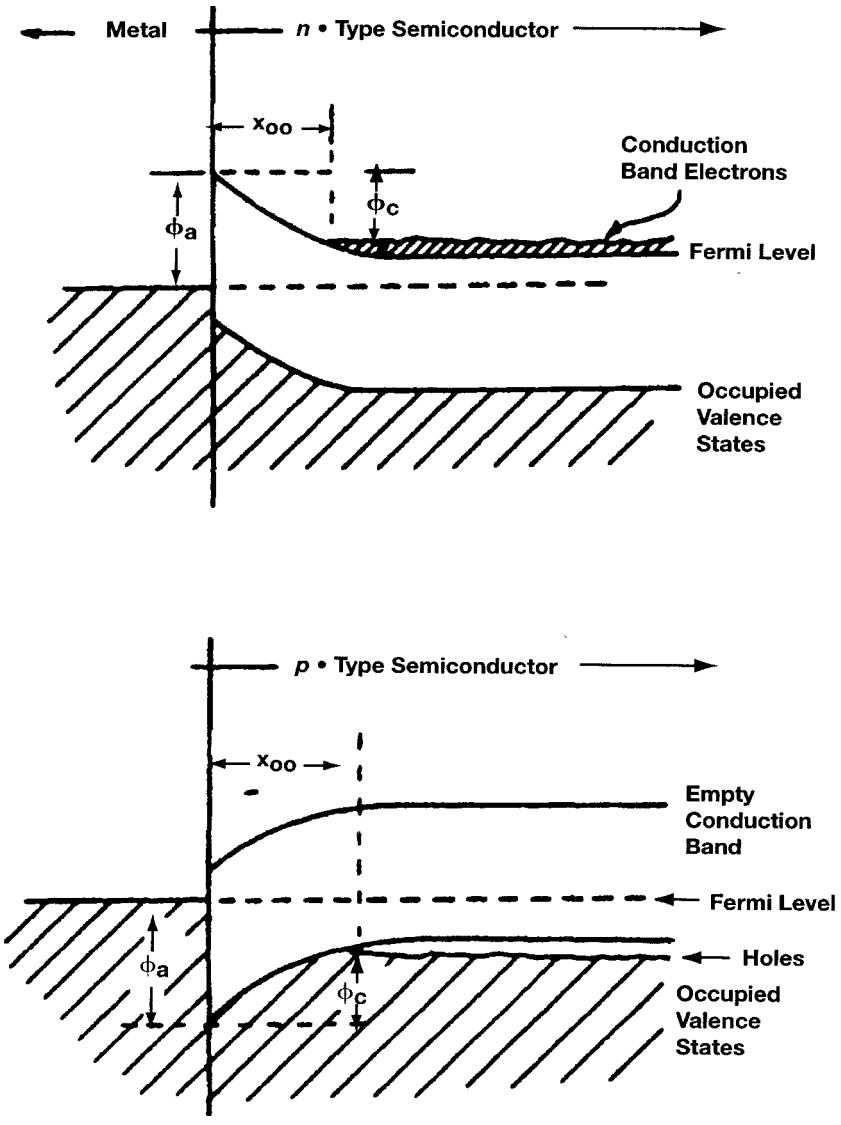


Figure 2-4 Schottky diode band diagrams.

Table 2-3 LO power and V_{F1} for various applications

Type	Typical V_{F1}	LO Power	Application
"Zero bias"	0.10–0.25	< 0.1 mW	Mainly for detectors
Low barrier	0.25–0.35	0.2–2 mW	Low-drive mixers
Medium barrier	0.35–0.50	0.5–10 mW	General purpose
High barrier	0.50–0.80	> 10 mW	High dynamic range

forward voltage (low V_{F1}), or low drive) diode, to reduce the amount of LO power needed. However, if high dynamic range is important, high LO power is needed, and the diode can have a higher V_F and should also have a high V_B (see Table 2-3).

Noise Figure Versus LO Power. At low LO drive levels, the noise figure is poor because of poor conversion loss, due to a too-low conduction angle. At very high LO drive levels, the noise figure again increases due to diode heating, excess noise, and reverse conduction.

If a high LO drive level is needed, for example, to get higher dynamic range, a high V_B (> 5 V) should be specified. However, Nature requires that you pay for this with higher R_S (lower f_C), so the noise figure will be degraded compared to what could be obtained with diodes designed for lower LO drive. Forward voltage and breakdown are basically independent parameters, but high breakdown is not needed or desirable unless high LO power is used.

Such a high-breakdown diode will have low reverse current (which is important only if the diode has to run hot).

Silicon Versus GaAs. Typical silicon Schottky diodes have cutoff frequencies in the 800–200-GHz range, which is good enough through the Ku band.

At the Ku band and above, or for image-enhanced mixers, higher f_C may be needed, which calls for the use of GaAs diodes. These have lower R_S due to higher mobility, which translates to cutoff frequencies in the 400–1000-GHz range.

However, if your IF is low, be careful; GaAs diodes have high 1/*f* noise. They also have high V_{F1}, so more LO power is required.

C_J Versus Frequency. There is quite a lot of latitude in choosing C_J. However, in general, the capacitive reactance should be a little lower than the transformed line impedance (Z_0). If Z_0 is not known, a good way to start is to use $X_C = 100\ \Omega$. Experience has shown that most practical mixers use an X_C near this value (a little higher in waveguide, and lower in 50-Ω systems). This translates to the following rule of thumb for choosing the junction capacitance of a diode for operation at frequency *f* (in GHz):

$$C_{J0} \approx \frac{100}{\omega}$$

$$\approx \frac{1.6}{f} \quad \text{(in pF)} \tag{2-26}$$

In order to evaluate possible tolerances, we show the range of forward current as a function of diode voltage (Figure 2-5), the junction capacitance as a function of the bias voltage

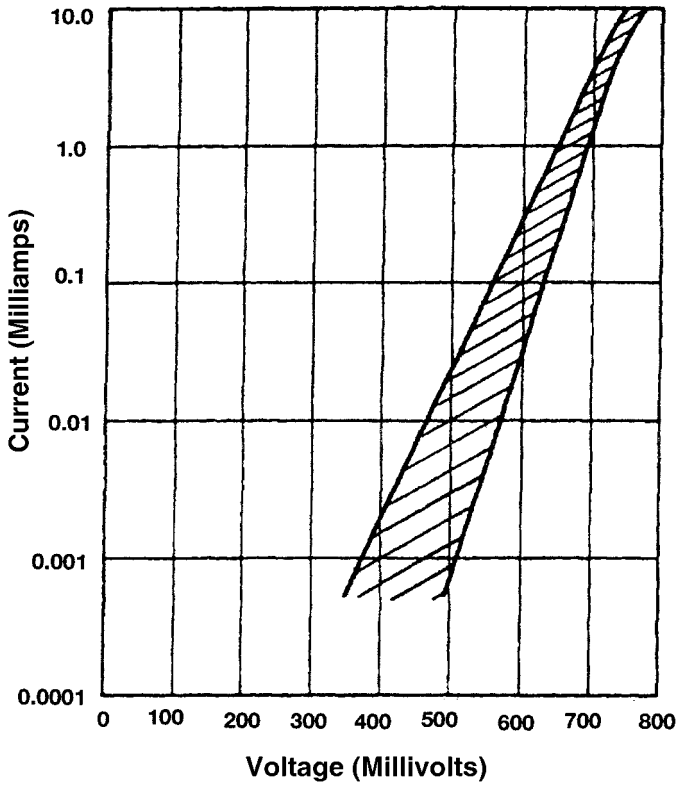


Figure 2-5 Forward dc characteristic curve range—current versus voltage.

(Figure 2-6), and finally some important RF parameters, such as noise figure and IF impedance, as a function of the local-oscillator drive (Figure 2-7).

Mixer Diodes. As an example of some of the parameters for mixer diodes, Table 2-4 gives data on some of the X-band mixer diodes. NF is measured at 9.375 GHz.

Linear Diode Model. Figure 2-8 shows the linear diode model. Its keywords appear in Table 2-5.

The circuit simulators, such as those supplied by Ansoft and Hewlett-Packard, provide a model library that has SPICE-type parameters for diodes (regular diodes, varactor diodes, and PIN diodes) as well as bipolar transistors and FETs, which will be discussed later.

2-1-3 PIN Diodes

Introduction. The PIN diode, in comparison with other microwave semiconductor devices, is fairly easy to understand. This makes possible the reduction of its complex behavior to relatively simple terms.

We do not attempt to describe the many possible microwave circuits in which PIN diodes are used. Rather, we attempt to explain the behavior of the diode in all aspects, giving the facts and some of the theory behind the facts. We offer the circuit designer the opportunity

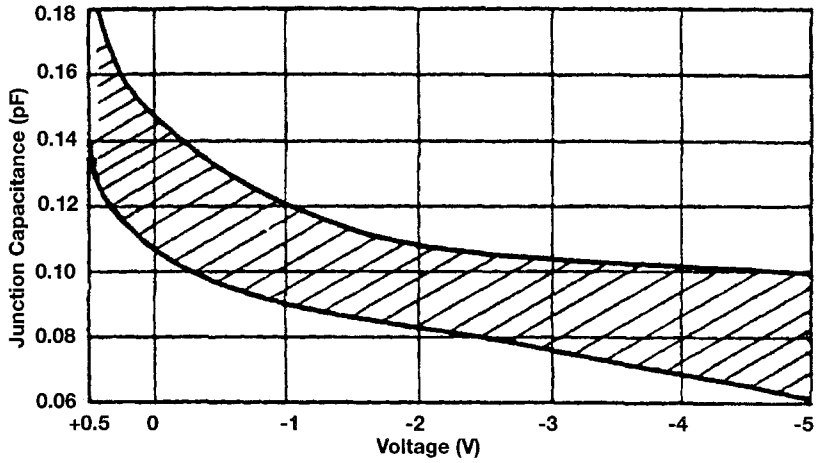


Figure 2-6 Junction capacitance range versus voltage.

to understand the PIN, so that he/she can understand its behavior in circuits. We assume the reader knows the circuit equations; to that knowledge we hope to add diode equations.

Most of the material presented consists of generalized data and explanations of the behavior of PIN diodes; we conclude with a brief description of circuit performance, test methods, and some hints on proper PIN specification writing.

The user can then evaluate the trade-offs involved in diode design and performance and be able to select the most nearly optimum diode from the wide range of diodes offered.

Large-Signal PIN Diode Model. Figure 2-9 shows the large-signal model for a PIN diode. Table 2-6 lists its keywords.

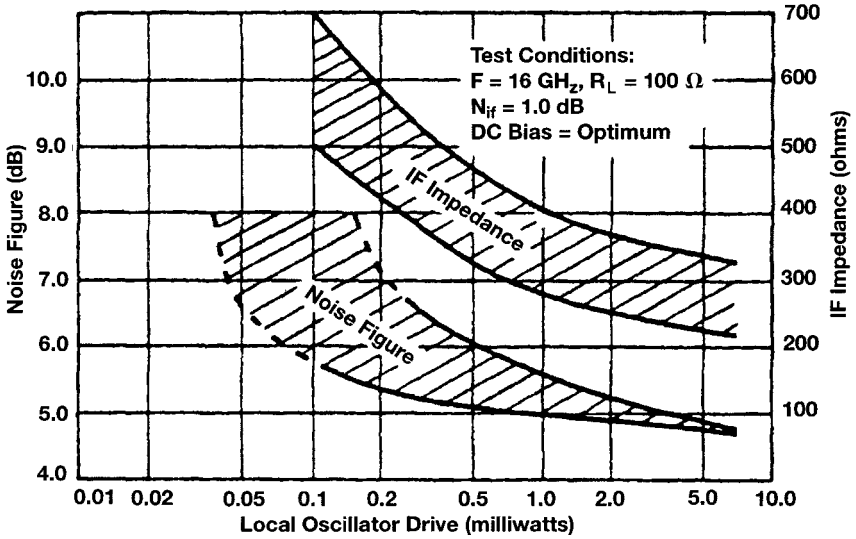


Figure 2-7 RF parameters versus local-oscillator drive level.

Table 2-4 X-band mixer diode data

Material	Barrier	Typical V_F (@1 mA)	Typical F_{C0} (GHz)	Typical R_S (Ω)	Typical C_{J0} (pF)	Maximum NF (dB)
n GaAs	High	0.70	1000	—	0.15	5.0[a]
n GaAs (BL)	High	0.70	500	—	0.15	6.0[a]
n GaAs (chip)	High	0.70	1000	—	0.15	5.3[a]
n Silicon (BL)	Low	0.28	150	6	0.20	6.5
n Silicon (quad)	Low	0.28	150	6	0.20	6.5
p Silicon (BL)	Low	0.28	150	12	0.20	6.5
n Silicon (BL)	High	0.60	100	8	0.20	6.5
n Silicon (quad)	High	0.60	100	8	0.20	6.5
n Silicon	Low	0.28	200	6	0.15	5.5
p Silicon	Low	0.28	200	18	0.14	6.0
p Silicon	Medium	0.40	150	150	0.12	6.5
n Silicon	Low	0.28	150	8	0.18	6.5
p Silicon	Low	0.28	150	12	0.18	6.5

[a]Specified for N_F = 1.0 dB.

Basic Theory: Variable Resistance. Intrinsic or "pure" silicon, as it can be grown in a laboratory, is an almost lossless dielectric. Some of is physical properties include:

Dielectric constant (relative)	12
Dielectric strength	400 V/mil (approximate)
Specific density	2.3
Specific heat	0.72 J/g · °C
Thermal conductivity	1.5 W/cm · °C
Resistivity	300,000 Ω/cm

Since a PIN diode is valuable essentially because it is a variable resistor, let us concentrate initially on the resistivity. Consider a volume comparable to a typical PIN diode chip, say, 20 mils in diameter and 2 mils thick. This chip has a dc resistance of about 0.75 MΩ. High

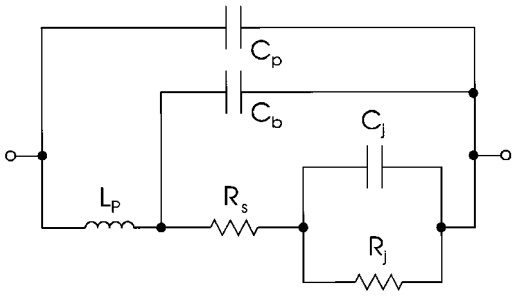


Figure 2-8 The linear diode model. This model is temperature dependent.

Table 2-5 Linear diode model keywords

Keyword	Description	Unit	Default
LP	Package inductance	henry	0.0
CB	Beam-lead capacitance	farad	0.0
CP	Package capacitance	farad	0.0
RS	Contact resistance	ohm	0.0
RJ	Junction resistance	ohm	
CJ	Junction capacitance	farad	

resistivity in any material indicates that most of the likely carriers of electric charge, electrons and holes, are tightly held in the crystal lattice and cannot "conduct."

In real life there are impurities, typically boron, that cannot be segregated out of the crystal. Such impurities contribute carriers, holes or electrons, that are not very tightly bound to the lattice and therefore lower the resistivity of the silicon.

Through various techniques we can adjust the level of impurities, called *dopants*, to produce resistivities ranging from 10 kΩ/cm (for good PIN diodes) to 0.001 Ω/cm (for substrates).

If the impurity adds "electrons" to the crystal, it is called a *donor*; if it adds holes it is called an *acceptor*. Boron adds holes, hence it is an acceptor, and the silicon plus boron combination is called *p*-type, or *positive*, because it has an excess of positive carriers. Phosphorus, on the other hand, is a *donor*, adding electrons, and the corresponding mix is *n*-type, or *negative*.

Notes on the PIN Diode Model

1. The PIN diode model is used to model a bias-dependent RF resistance for use in PIN diode circuits such as attenuators and switches. The resistance varies from RMAX to RS using the R function above. A typical R versus *I* characteristic is shown in Figure 2-10 with parameters IS = 5.96 nA, RS = 2.016, RMAX = 6500, K1 = 0.1272, K2 = 1.0, N = 2.077.
2. The transit-time parameter, TT, can also be used to approximately model a switching PIN diode's reverse-recovery time—a value often provided by diode manufacturers.
3. Diode breakdown can be modeled by specifying IBV and BV parameters.

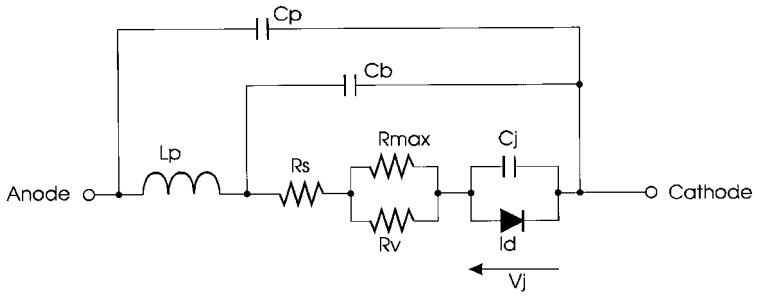


Figure 2-9 The large-signal PIN diode model. This model is temperature dependent.

Table 2-6 PIN diode model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
IS	Saturation current	ampere	1.0E–14
N	Emission coefficient		1.0
IBV	Magnitude of current at the reverse breakdown voltage	ampere	1.0E–10
BV	Magnitude of the reverse breakdown voltage	volt	∞
FC	Coefficient for forward-bias depletion capacitance		0.5
CJ0	Zero-bias *pn* junction capacitance	farad	0.0
VJ	Built-in junction potential	volt	1.0
M	*pn* Junction grading coefficient		0.5
GC1	Varactor capacitance polynomial coefficient 1	/volt	0.0
GC2	Varactor capacitance polynomial coefficient 2	/volt^2	0.0
GC3	Varactor capacitance polynomial coefficient 3	/volt^3	0.0
TT	Transit time	second	0.0
K1	Variable resistance coefficient	volt	0.0
K2	Variable resistance current exponent		1.0
RMAX	Maximum resistance of PIN intrinsic region	ohm	0.0
KF	Flicker noise coefficient		0.0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
AREA	Area multiplier		1.0
	Extrinsic Model		
RS	Series resistance (minimum resistance of PIN diode)	ohm	0.0
CP	Package parasitic capacitance	farad	0.0
CB	Beam-lead parasitic capacitance	farad	0.0
LP	Package parasitic inductance	henry	0.0

4. The reverse-bias capacitance characteristics can be more accurately modeled than the common expression derived from *pn* junction theory. The capacitance grading coefficient exponent can be expressed as a polynomial function of voltage by specifying values for GC1, GC2, and GC3.
5. The PIN diode model was derived from J. Walston, "SPICE Circuit Yields Recipe for PIN Diode," *Microwaves & RF*, November 1992.
6. Following Simon M. Sze, *Physics of Semiconductor Devices* (New York: John Wiley, 1982), the variable resistance may be modeled by setting

$$\text{K1} = \frac{3}{8} V_T \frac{W^2}{D_a \tau_a} \qquad \text{K2} = 1.0$$

where W is the width of the intrinsic region, D_a is the ambipolar diffusion coefficient, τ_a is the ambipolar lifetime, and V_T is the thermal voltage.

There are many concepts important to the physicist, but not to the diode user, that elaborate on the impact of impurities on the behavior of silicon. These can be studied in Watson [1]. The more carriers added, the lower the resistivity.

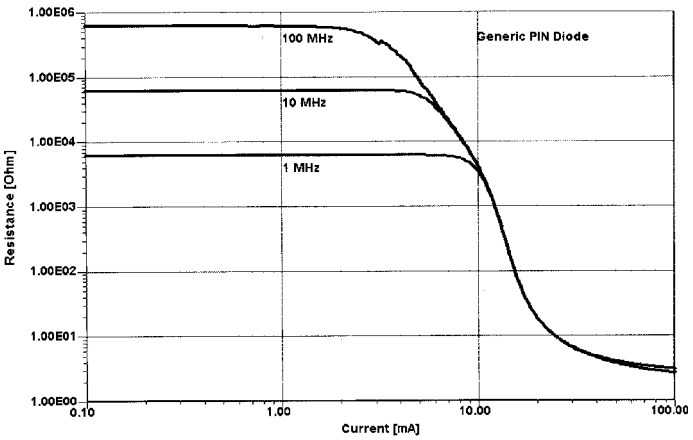


Figure 2-10 Simulated PIN diode resistance as a function of dc at 1, 10, and 100 MHz.

If one wished to vary the resistance of a given diode, in principle he/she could bring it into a semiconductor laboratory, add or subtract carriers as desired, and perhaps even make the process reversible. However, this is a slow, expensive, and impractical way to make a variable resistor; one would be better advised to take a wrench and a soldering iron and replace a component.

The PIN diode derives its value from the fact that the free-charge-carrier concentration in silicon, and hence its resistance, can be varied electronically by means of current from a simple bias supply. This can be done rapidly (in nanoseconds in some cases), reversibly, repeatably, and accurately. The thing that makes this possible is called a *junction*, the interface between the relatively pure silicon in the middle of the PIN diode (the I stands for *intrinsic*) and the heavily doped layers on either end, p^+ and n^+. The p^+ region is rich in holes; the n^+ region is rich in electrons. Both of these regions have low resistance. The I region is the variable resistive element in the diode (see Figure 2-11). In the absence of any external bias, internal effects within the crystal keep the charges fixed; the resistance of the I region is high.

When the p^+ region (anode) is biased positively with respect to the n^+ region (cathode), the interface potential "barrier" is overcome, and direct current flows in the form of holes

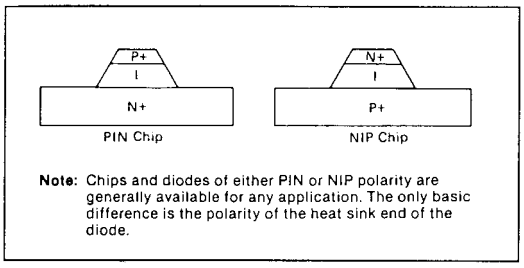


Figure 2-11 General outline of PIN diode construction.

streaming from p^+ toward n^+, with electrons moving in the opposite direction; we say that free carriers have been injected into the I region. The resistance of the I region becomes low.

The number of free carriers within the I region determines the resistivity of the region and thus the resistance of the diode.

Consider "one hole" and "one electron" drifting in opposite directions in the I region under the impetus of the applied field. Under certain conditions, imperfections in the silicon may cause these carriers to recombine. They are no longer available to constitute current or to lower the resistivity of the I region.

It can be shown that the amount of "recombination" between holes and electrons that continuously takes place in a semiconductor is governed by a property of the lattice called *lifetime*. In fact, lifetime is defined as the reciprocal of recombination rate.

Thus, $Q_S = Q_0 \exp(-t/T_L)$, where Q_S is the total amount of free charge "stored" in the I region, and T_L is the lifetime, or the mean time between recombination events. In the steady-state condition, the bias supply must deliver current to maintain a constant Q_S. The required current is

$$I_{\text{dc}} = \frac{Q_S d}{dt} = \frac{-Q_S}{T_L} \tag{2-27}$$

or $Q_S = I_{\text{dc}} T_L$, dropping the minus sign.

Ignoring some details that are not crucial to this section, we can now calculate the resistance of a given diode, of area A and thickness W. (W stands for *base width*, the width or thickness of the intrinsic layer.) The p^+ and n^+ regions have essentially zero resistance, as they are very heavily doped.

The resistivity of a given material is inversely proportional to the number of free carriers, N, and the mobility (not quite the same as velocity) of the carriers. Thus,

$$\rho = \frac{1}{q(\mu_n N + \mu_p P)} \tag{2-28}$$

where μ_n and μ_p are mobilities of electrons and holes, and N and P are numbers of electrons and holes. Simplifying, we find

$$\rho = \frac{C}{Q_S d} \tag{2-29}$$

where C is a collection of constants, and $Q_S d$ is the stored charge density (numbers per unit volume). For our piece of silicon, volume is WA, and

$$Q_S d = \frac{Q_S}{WA} \tag{2-30}$$

The resistivity is

$$\rho = \frac{CWA}{Q_S} \tag{2-31}$$

and the resistance is

$$R_S = \rho \frac{W}{A} = \frac{CW^2}{I_{\text{dc}} T_L} \tag{2-32}$$

This is a fundamental equation in PIN diode theory and design.

Rigorous analysis shows that

$$R = \frac{2kT/q}{I_f} \sinh\left(\frac{W}{2\sqrt{DT_L}} \tan^{-1}\left[\sinh\frac{W}{2\sqrt{DT_L}}\right]\right) \tag{2-33}$$

where k = Boltzmann's constant
T = temperature (in kelvins)
D = diffusion coefficient = $\mu kt/q$

For most PIN diodes, W/DT_L is less than unity, and the equation simplifies to the simple equation above.

Typical data on R_S as a function of bias current are shown in Figure 2-12. Many design choices are available, as the data indicate. Many combinations of W and T_L have been developed to satisfy the full range of applications.

Breakdown Voltage, Capacitance, Q Factor. The previous section on R_S explained how a PIN can become a low resistance, or a "short." This paragraph will describe the other state: a high impedance, or an "open." Clearly, the better PIN diode is the one that has the better on/off ratio at the frequency and power level of interest.

If we return to the undoped or intrinsic I region, we note that it is an almost lossless dielectric. As such, it has a dielectric strength of about 400 V/mil, and all PIN diodes have

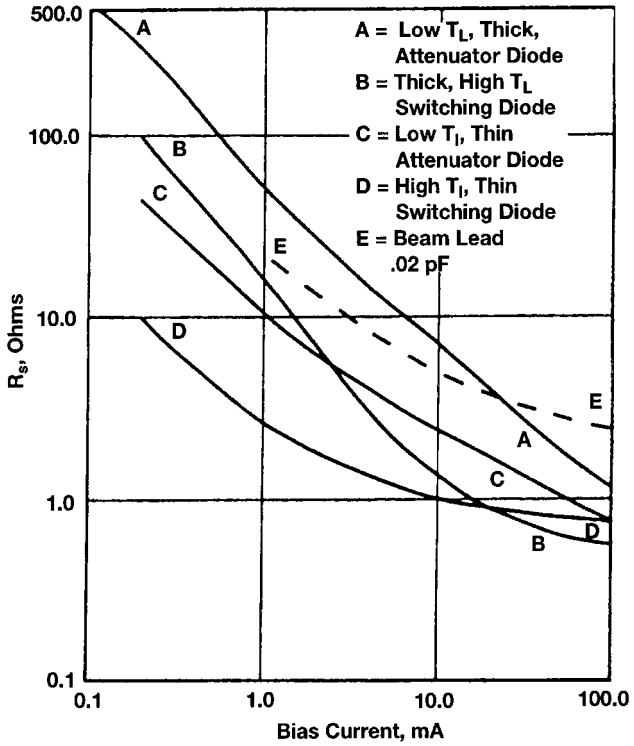


Figure 2-12 Typical series resistance as a function of bias (1 GHz).

a parameter called V_b, breakdown voltage, which is a direct measure of the width of the I region. Voltage in excess of this parameter results in a rapid increase in current flow (called avalanche current); see Figure 2-13. When the negative bias voltage is below the bulk breakdown of the I region, a few nanoamperes will be drawn. As V_b is approached, the leakage current increases often gradually, as is exaggerated in the curve. This current is primarily caused by less-than-perfect diode fabrication, although there is some contribution from temperature. Typically, the leakage current occurs at the periphery of the I region. For this reason, various *passivation* materials (silicon dioxide, silicon nitride, hard glass) are grown or deposited to protect and stabilize this surface and minimize leakage. These techniques have been well advanced over the years, and PIN diode reliability has improved as a result.

Most diodes are specified in terms of minimum V_b for a nominal leakage, usually 10 μA.

It will be noted later that RF voltage swings in excess of the rated V_b are permitted, for the mechanisms causing leakage current do not always respond at radiofrequencies. However, bulk breakdown is effectively instantaneous, and that voltage should never be exceeded.

The next characteristic of our "open" circuit is the capacitance. In simplest form, the capacitance of a PIN diode is determined by the area and width of the I region and the dielectric constant of silicon; however, we have discussed the fact that intrinsic material does contain some carriers and therefore has some conductivity. An *E* field could not exist unless all these carriers were swept out, or depleted.

Application of a reverse bias accomplishes this. At zero bias, the excess carriers on either side of the junction are separated, held apart, by "built-in" fields. This is the contact potential (about 0.5 V for silicon). If there are only a few excess carriers in the I region, this "potential" can separate the charges more easily. The junction "widens" in the sense that, starting at the p^+ and I interface, there is a region of no free carriers called the *depletion zone*. Beyond this depletion zone, the I region still contains the free charges with which it started. With the application of reverse bias, the depletion zone widens.

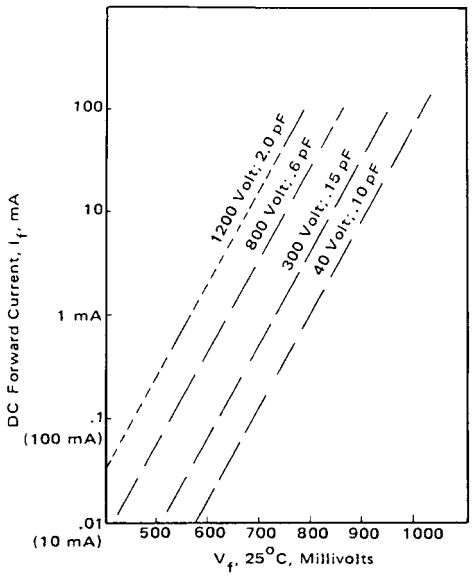


Figure 2-13 Current versus voltage for various PIN diodes.

Eventually, at a bias equal to a so-called *punchthrough* voltage (V_{PT}), the depletion zone fills the entire I region. At this voltage, the 1-MHz capacitance bottoms out and the diode Q reaches its maximum. Figure 2-14 illustrates the equivalent circuit of the I region before punchthrough.

Some very interesting facts can be derived from this model. Consider the undepleted region; this is a lossy dielectric consisting of a volume (area A, length ℓ) of silicon of permittivity 12 and resistivity ε. The capacitance is

$$12\frac{\varepsilon_0 A}{\ell} \tag{2-34}$$

and the admittance is

$$2\pi\frac{12\varepsilon_0 A}{\ell} \tag{2-35}$$

The resistance is $\varepsilon\ell/A$ and the conductance is $A/\varepsilon\ell$.

At very low frequencies, the undepleted zone looks like a pure resistor. At very high frequencies it looks like a lossy capacitor. The "crossover" frequency depends on the resistivity of the I-region material. For ε of 160 Ω/cm, the frequency is 1 GHz. Higher resistivity is generally used for PINs—say, 1000 Ω/cm, and the crossover frequency is 160 MHz.

Diode manufacturers measure junction capacitance at 1 MHz; clearly, what is measured is the depletion-zone capacitance.

If the I region thickness is W and the depletion is X_d, the undepleted region is $(W - X_d)$.

The capacitance of the depleted zone is, proportionally,

$$\frac{1}{X_d} \tag{2-36}$$

and of the undepleted zone,

$$\frac{1}{W - X_d} \tag{2-37}$$

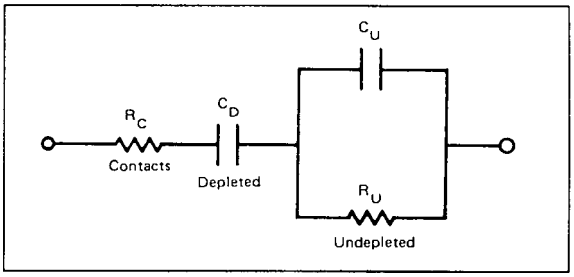


Figure 2-14 Equivalent circuit of I region before punchthrough.

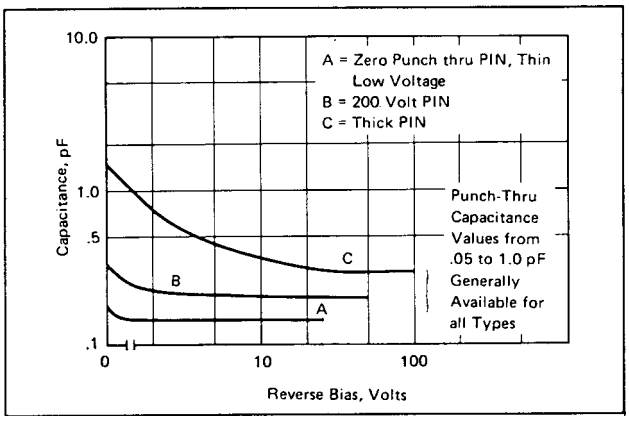


Figure 2-15 Typical 1-MHz capacitance.

The 1-MHz capacitance as a function of reverse bias is seen in Figure 2-15. The 1-MHz capacitance decreases with bias until punchthrough, where $X_d = W$. However, at microwave frequencies well above the crossover, the junction looks like two capacitors in series:

$$C_T = \frac{C_d C_u}{C_d + C_u} \propto \frac{1}{W} \tag{2-38}$$

That is, the microwave capacitance tends to be constant, independent of X_d and bias voltage.

However, since the undepleted zone is lossy, an increase in bias up to the punchthrough voltage reduces the loss.

At any given frequency, the equivalent network can now be drawn as Figure 2-16. R_v is now the equivalent series resistance of the undepleted region. Typical R_v data are shown in Figure 2-17. An alternate equivalent network is shown in Figure 2-18, and typical R shunt data are shown in Figure 2-19.

A good way to understand the effects of series resistance is to observe the insertion loss of a PIN chip shunt mounted in a 50-Ω line, as shown in Figure 2-20. An accepted way to include reverse loss in the figure of merit of a PIN diode is to write

$$Q = \frac{1}{2\pi + \sqrt{R_s R_v C}} \tag{2-39}$$

where R_s and R_v are measured under the expected forward- and reverse-bias conditions at the frequency of interest.

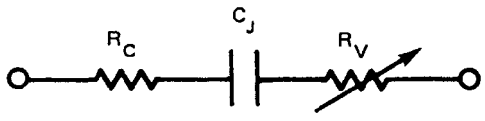


Figure 2-16 Simplified equivalent series circuit.

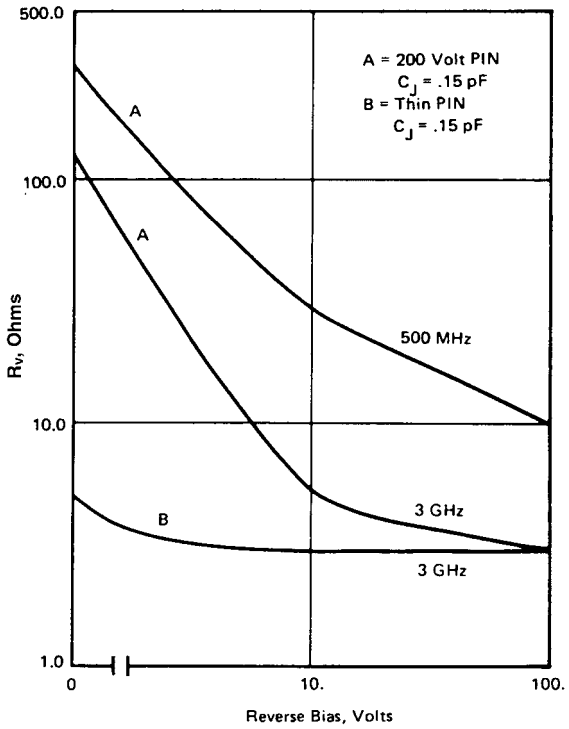


Figure 2-17 Reverse series resistance R_V.

The punchthrough voltage is a function of the resistivity and thickness of the I region. It is advisable to measure loss as a function of bias voltage and RF voltage to determine if the correct diode has been selected for your application. (*Note*: At frequencies below crossover, and diodes with thin I regions, the effective junction capacitance can increase substantially at low forward bias, on the order of 1–200 μA.)

Incidentally, if you are working with PIN or NIP chips that do not have an opaque covering, note that PIN diodes are photosensitive. Incident light causes photogeneration of carriers in the I region, increasing the chip's insertion loss.

PIN Diode Applications. If the intrinsic zone is thick (10–100 μm), we then have a high-reverse-voltage rectifier with a low forward-voltage drop at high current or, in other

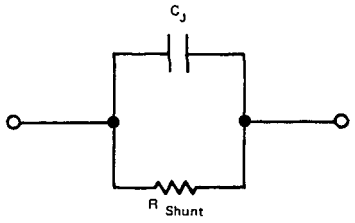


Figure 2-18 Simplified equivalent circuit, shunt.

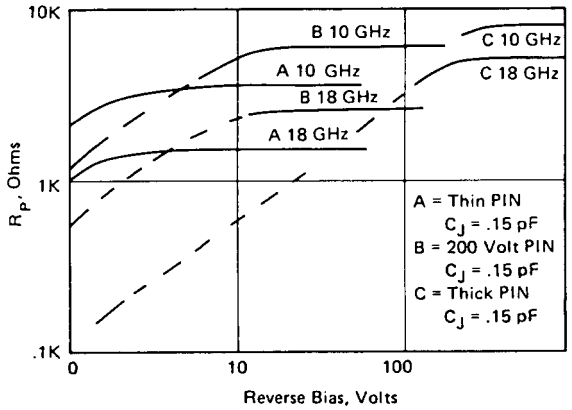


Figure 2-19 Reverse shunt resistance, R_p.

words, a highly efficient rectifier. The low forward voltage results from the fact that the conductivity of the I zone can be modulated by large amounts of charge carriers injected from the *p* and the *n* zones.

Another application of PIN diodes is the high-frequency (HF) field. Here, the fact is exploited that, due to the long carrier lifetime at frequencies beyond approximately 10 MHz, a rectifying effect will no longer occur and the PIN diode rather behaves like a real resistance, the magnitude of which depends on the forward direct current passed by the device and

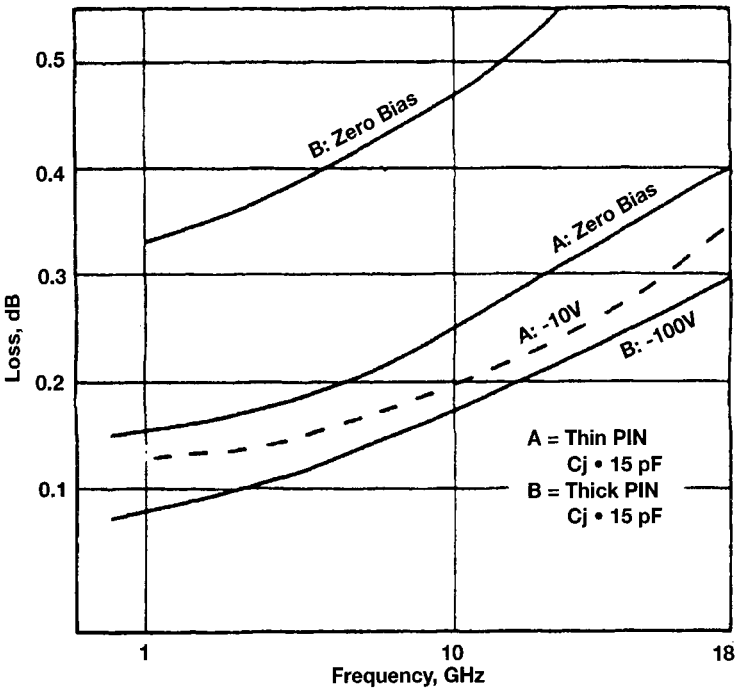


Figure 2-20 Insertion loss versus frequency.

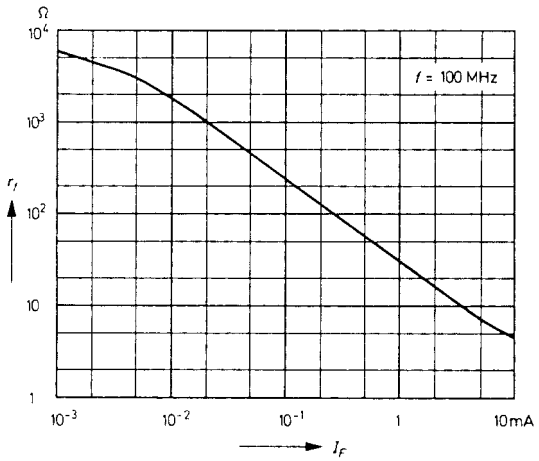


Figure 2-21 Forward resistance versus forward current.

produces an equal effect on both half waves of the HF signal. In view of this behavior, the PIN diode can be used as a switch or a variable resistor for HF signals. Thus it becomes possible, for example, to subject an HF signal to amplitude modulation by means of an AF-controlled PIN diode.

An important application of PIN diodes that has found favor in recent times is for dc-operated attenuators in TV tuners and antenna distribution amplifiers. Figure 2-21 shows the real HF forward resistance r_f as a function of the forward current I_f, measured at 100 MHz. Figures 2-22 and 2-23 show second-order IMD and cross-modulation for PIN diodes.

Applying the PIN in Amplitude Control of High-Frequency Signals. In conventional transistorized TV tuners, automatic gain control (AGC) is usually achieved by varying the emitter current of the input transistor. This method exploits the fact that the input transistor exhibits maximum gain at a certain level of emitter current and that this gain decreases when the emitter current is either increased or decreased relative to this point. Since modulation capacity and cross-modulation resistance grow with the emitter current, "upward" gain control has lately become the preferred solution, and according to it, gain reduction is achieved by an increase of the input transistor's emitter current. An input stage based on this design, therefore, is least resistant to cross-modulation when receiving a signal from a weak transmitter because it is operating at or below the point at which AGC action begins. This effect is particularly disadvantageous when the signal of a weak transmitter is to be received in the presence of a strong local transmitter. The unsatisfactory cross-modulation properties of such an input stage manifest themselves more and more frequently in the form of image perturbations as the TV reception band is crowded by an increasing number of transmitters. Another disadvantage is the variation in the transistor parameters that results from the control action. The variations affect antenna matching at the input and the response of the RF filter at the output.

To obviate these problems, methods being adopted today provide for the input stage to be equipped with cross-modulation-resistant transistors with fixed bias, gain control being provided by a variable attenuator preceding the transistor(s). The input-filter termination

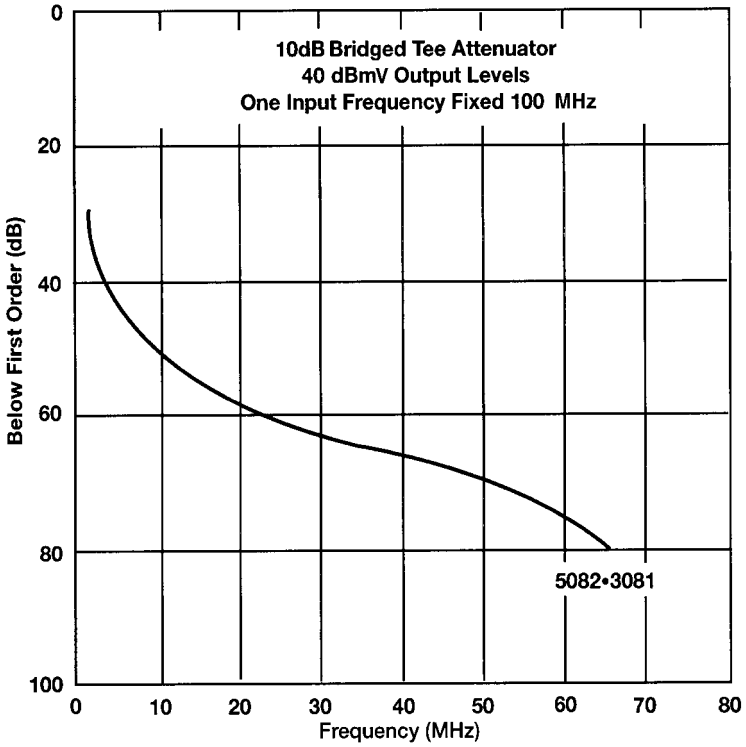


Figure 2-22 Second-order IMD in a PIN diode. Equation (1-48) can be used to determine the diode's IP_2 from the absolute values of its fundamental and IM_2 outputs.

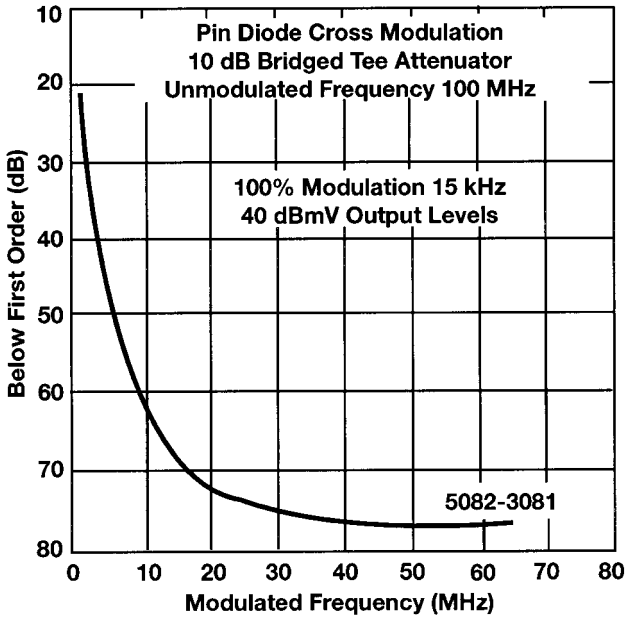


Figure 2-23 Cross-modulation in a PIN diode. Equation (2-61) relates cross-modulation level to IMD.

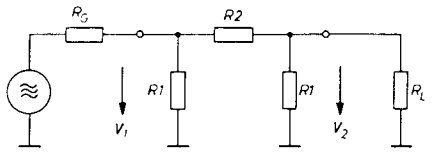


Figure 2-24 Basic circuit of a π-mode filter attenuator.

provided by this attenuator must be independent of the degree of attenuation to preserve the filter's response characteristics.

The first two requirements are met by the properties of the PIN diode and the latter requirement can be met by designing a PIN diode control in the form of a π network, a schematic of which is shown in Figure 2-24. The attenuation of a matched π network is expressed by the following equations:

$$\frac{a}{\text{dB}} = 20 \log \frac{Z + R_2}{R_2 - Z} \tag{2-40}$$

$$\frac{a}{\text{dB}} = 20 \log \left(\frac{R_1}{Z} + \sqrt{\frac{R_1^2}{Z^2} + 1} \right) \tag{2-41}$$

Figure 2-25 plots the relationship between the resistance levels of R_1 and R_2 and the input and output voltages of the π network. Figure 2-26 shows the circuit diagram of an implementation of the π network composed of PIN diodes. The resistance characteristic required according to Figure 2-25 is achieved approximately by varying the control current I_{co}.

If the control current I_{co} in Figure 2-26 equals zero, then—due to the fact that $R_1 = R_2$ and $R_3 = R_4$—the auxiliary voltage V_h causes forward currents of equal magnitude to flow through diodes D_1 and D_3. The voltage drops across R_1 and R_2 are of equal magnitude, so that no current passes through diode D_2. Therefore, the latter presents a high resistance, whereas D_1 and D_3 present a low resistance. Under these conditions, the network produces maximum attenuation. A control current I_{co} reduces the forward currents through D_1 and D_3 and allows a forward current to be passed by diode D_2. When it has attained its maximum value, diodes

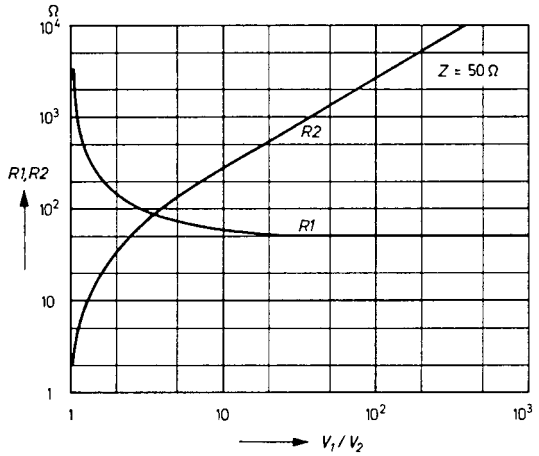


Figure 2-25 Values of R_1 and R_2 in Figure 2-24 versus relation of input and output voltage.

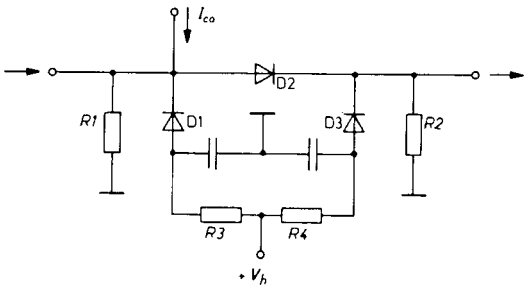


Figure 2-26 PIN diode π-mode attenuator.

D_1 and D_3 are blocked—that is, they are high-ohmic—and D_2 is low-ohmic. Under these conditions, the network produces its minimum attenuation.

If the maximum attenuation of the π network is to be increased, it is possible to make the two shunt diodes D_1 and D_3 more low-ohmic than the impedance Z, although this would reduce the reflection coefficient of attenuation. For example, if a reflection coefficient of about 7 dB is accepted over the entire control range, then a maximum attenuation of 25–30 dB could be achieved in the VHF range. The series inductances of the transverse diodes and the shunt capacitance of the series diode are responsible for this relatively unsatisfactory value. Improved attenuation is brought about with the integrated π network described in the next section.

Example: A PIN Diode π Network for TV Tuners. The since-discontinued TDA1053 is an example of an integrated PIN π network attenuator intended for TV-tuner use. It comprises three silicon planar PIN diodes connected to form a π network (Figure 2-27) and serves for the electronic amplitude control of the input signals of TV tuners and antenna distribution amplifiers in the 40–1000-MHz range. Its input and output impedances remain constant over the entire control range. This can also be achieved with discrete diodes.

The TDA1053 was normally supplied with vertical leads. The characteristics stated below apply to devices of this configuration.

MAXIMUM RATINGS OF INDIVIDUAL DIODES

Reverse voltage	V_R	30	V
Forward current at $T_{amb} = 25$ °C	I_F	50	mA
Junction temperature	T_j	125	°C
Storage temperature range	T_s	−55 to +125	°C

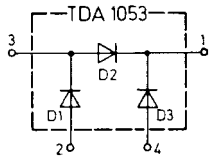


Figure 2-27 Internal circuitry of the TDA1053.

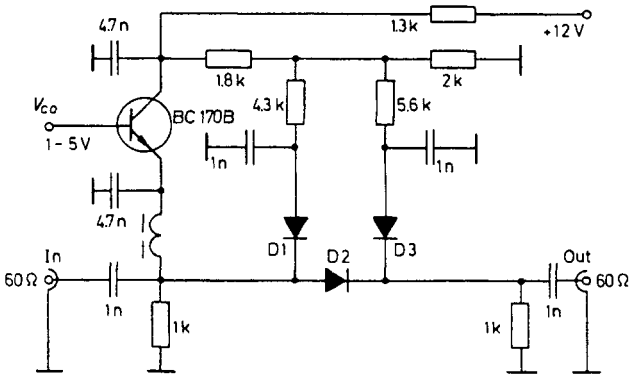


Figure 2-28 Test and application circuit.

MAXIMUM RATINGS OF THE π NETWORK

Ambient operating temperature range when operating according to Figure 2-28	T_{amb}	100	°C

CHARACTERISTICS OF INDIVIDUAL DIODES AT T_{amb} = 25 °C

Forward voltage at I_F = 50 mA	V_F	< 1.2	V
Forward current at V_R = 15 V	I_R	< 500	nA
Differential forward resistance:			
At I_F = 10 mA, f = 100 MHz	r_f	5	Ω
At I_F = 10 µA, f = 100 MHz	r_f	1.4	kΩ

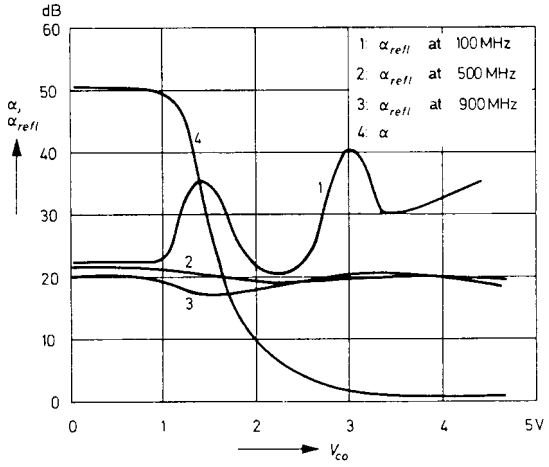


Figure 2-29 Attenuation and reflection attenuation versus control voltage.

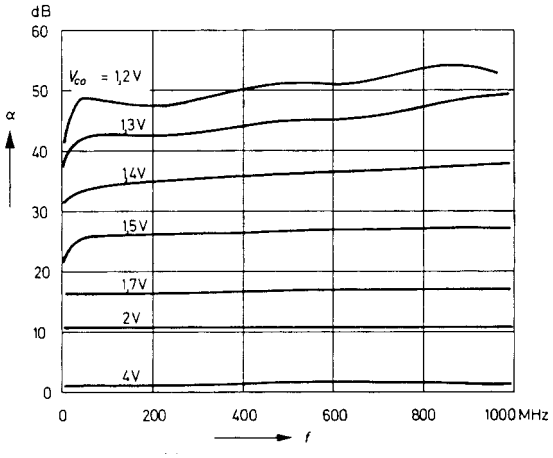


Figure 2-30 Attenuation versus frequency.

CHARACTERISTICS IN THE TEST CIRCUIT (FIGURE 2-28) AT T_{amb} = 25 °C

Voltage for 1% cross-modulation	V_{cr}	1	V
Attenuation in the 40–1000-MHz range:			
At V_{co} = 1.5 V (1–2 V)	a_{max}	45 (> 36)	dB
At V_{co} = 5 V (4–5 V)	a_{min}	1.5 (< 2)	dB
Reflection coefficient in the 40–1000-MHz range over the entire control range, depending on circuit design	a_{refl}	20 (> 16)	dB

These data reveal that the compact design of the three PIN diodes in a common 50B4 plastic package guarantees favorable values for minimum and maximum attenuation, as well as reflection attenuation. The test and application circuit shown in Figure 2-28 also comprises the transistor control-signal amplifier. The typical characteristic of the attenuation and the reflection attenuation for this circuit are shown in Figure 2-29, as a function of the control voltage V_{co}. Figure 2-30 shows the attenuation at different control voltages, as a function of frequency.

2-1-4 Tuning Diodes

Introduction. In recent years, continuous development of tuning diodes—also known as *varactors* or *varicaps*—together with increased commercial and military use has led to substantial improvement in Q, reproducibility, and reliability. Concurrently, new techniques for producing and controlling a hyperabrupt dopant profile in the semiconductor permit the capacitance–voltage law to be much faster than the classical square root or cube root behavior.

Current tuning diode materials include silicon and gallium arsenide; silicon is favored for low cost and lower Q applications from HF through microwave frequencies. Hyperabrupt varactors, also of silicon, are finding many applications in commercial television tuners, where their high tuning ratios, linear tuning, and low cost are needed. New developments include low-capacitance hyperabrupts for microwave and wireless applications.

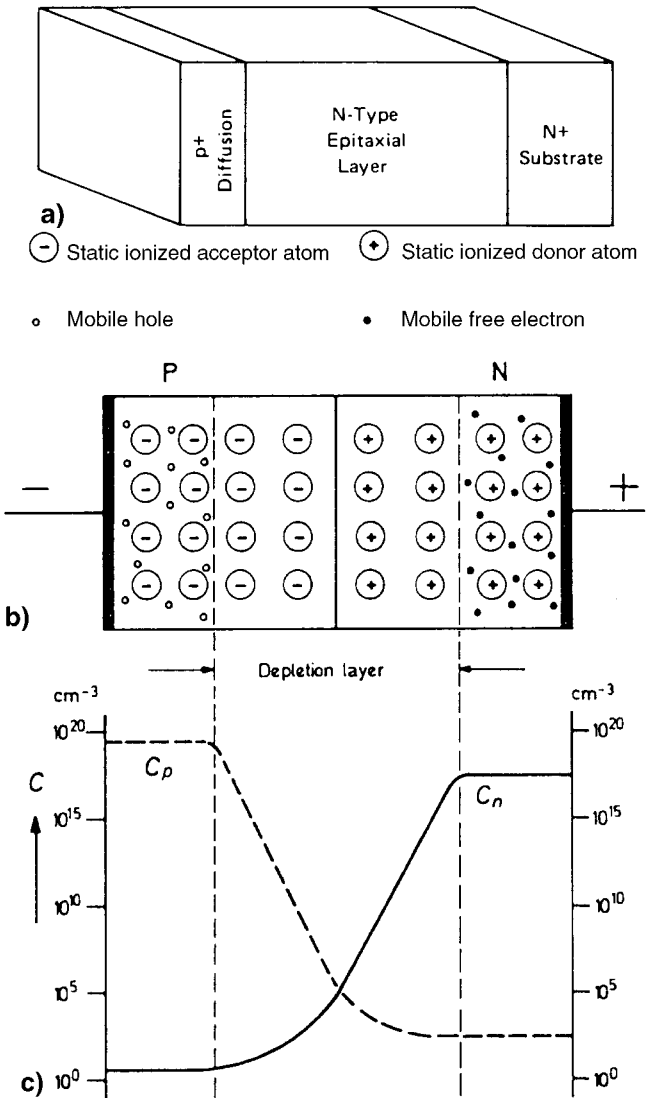


Figure 2-31 (a) Basic PIN structure; (b) cross section of a reverse-biased *pn* junction; and (c) density distribution of free charge carriers.

Gallium arsenide is used with high operating frequency and dictates the highest Q possible, as in parametric amplifiers and millimeter multipliers.

This section will acquaint the reader with tuning diodes: how they work, and what they can or cannot be expected to do in an electronic circuit. The basic properties of a tuning diode will be described in terms of the parameters that manufacturers use in characterizing them. The following topics will also be addressed:

- Capacitance ratio with respect to voltage and voltage breakdown
- Q as a function of design and operating conditions

- Stability—leakage current, temperature coefficient, and post-tuning drift
- Distortion products
- Packaging parasitics
- Applications—suggestions on how to specify a varactor

Tuning Diode Physics

Introduction. All junction diodes are made up of the same physical parts: a *pn* junction, a carefully controlled epitaxial layer, and a very-low-resistance substrate. These parts are shown in Figure 2-31.

No matter what type of junction device we are discussing—a tuning diode, a step-recovery diode, or a PIN diode—these parts are all present; the main difference between these devices is the resistivity and thickness of the epitaxial layer. Tuning diodes and multiplier diodes need epitaxial layers where both the resistivity and thickness are carefully controlled.

Abrupt Junction. An abrupt-junction diode is one in which the p^+ diffused region of the diode is much more highly doped than the epitaxial layer. Also, the high doping drops to the doping level of the epitaxial layer in a distance that is short compared to the epitaxial layer thickness, and the doping level of the epitaxial layer is constant over its thickness. This is shown in Figure 2-32, with the corresponding C–V curve shown in Figure 2-33. When these requirements are satisfied, the diode capacity, diode area, epitaxial layer doping level, and diode voltage are related by

$$\frac{C(V)}{A} = K\left(\frac{N}{V+\varphi}\right)^n \tag{2-42}$$

where $C(V)$ = capacitance of the diode at voltage V
A = area of the diode
N = doping level of the epitaxial layer

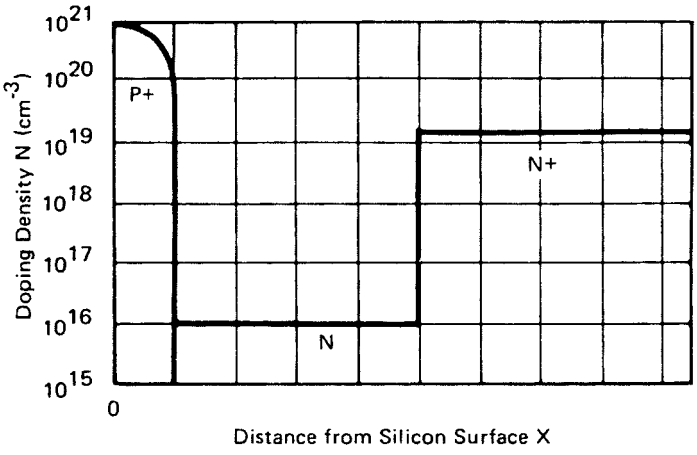


Figure 2-32 *N–X* abrupt junction.

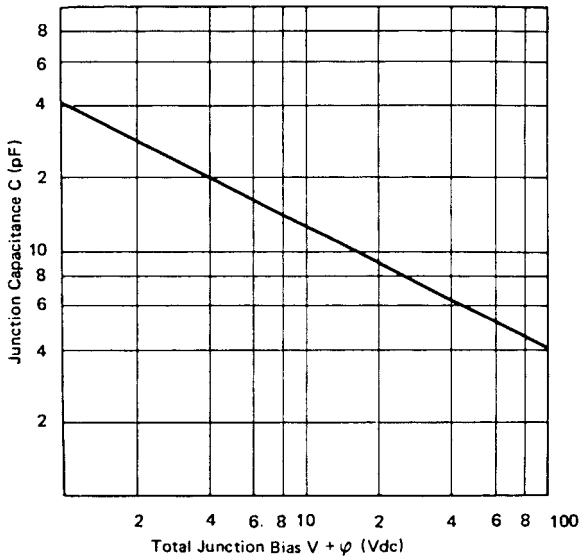


Figure 2-33 Capacitance versus total junction bias for abrupt-junction diode.

V = voltage applied to the diode
φ = built-in potential of the diode (0.6–0.8 V)
n = slope of diode C–V curve; $n \approx 0.5$ for an abrupt-junction diode
K = constant

As a consequence of the physical properties of a *pn* junction, a depletion layer is formed between the *p* and *n* regions whose width depends on the voltage applied to the diode. The capacitance of the diode is inversely proportional to the width of the depletion layer; that is, $C \sim 1/\ell_0$. In addition, the series resistance of the diode is proportional to the width of the undepleted epitaxial layer. Thus, as diode reverse bias is increased, the depletion layer increases, causing a decrease in capacitance and a decrease in series resistance. As the diode reverse bias is increased further, a point is reached where the electric field caused by the reverse bias reaches a critical level, and current through the diode increases rapidly; this is

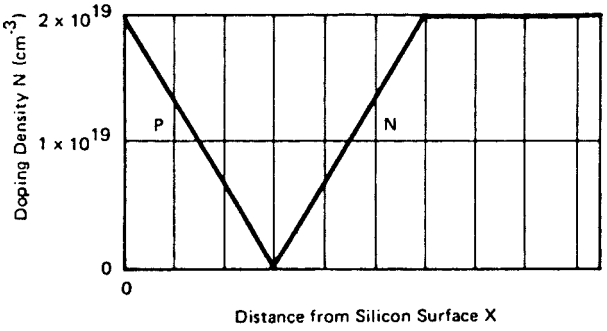


Figure 2-34 *N–X* linearly graded junction.

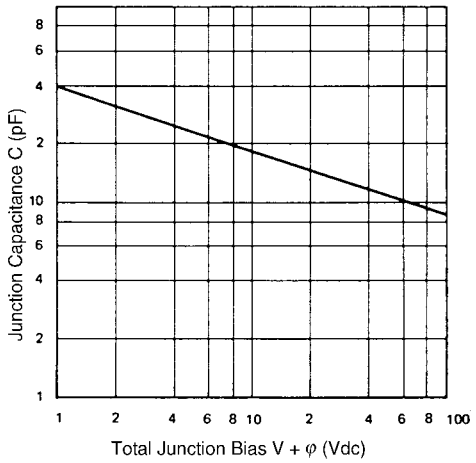


Figure 2-35 Capacitance versus total junction bias for linearly graded junction.

breakdown voltage of the diode. If, at the breakdown voltage, the epitaxial layer is not completely depleted, the diode will have excessive series resistance. Conversely, if the epitaxial layer is depleted before the breakdown voltage is reached, no further capacitance decrease occurs after the total depletion, and a condition called *punchthrough* occurs.

While, in the ideal case, voltage breakdown will occur just as the epitaxial layer is totally depleted, this seldom occurs in practice, and we generally have a condition of either punchthrough or excess series resistance.

Linearly Graded Junction. If, instead of the junction profile shown in Figure 2-32, we have a p^+ region and an n region whose doping levels increase linearly with distance from the pn junction as shown in Figure 2-34, with its corresponding C–V curve in Figure 2-35, then we have what is called a linearly graded junction diode. This diode follows Eq. (2-42) with the exception that the exponent n is equal to $\frac{1}{3}$. This means that, for a given voltage change, the linearly graded junction will have a smaller capacitance change than an abrupt-junction diode. Since, in most cases, the designer is looking for the maximum

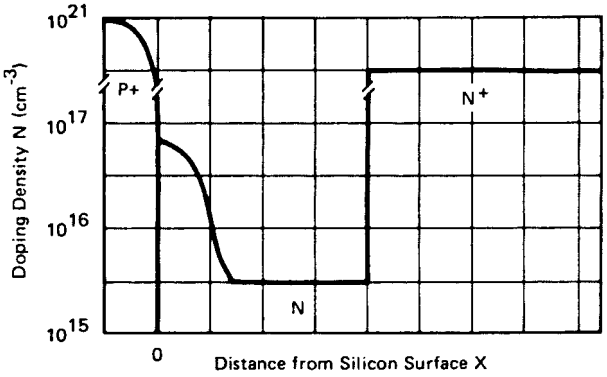


Figure 2-36 *N–X* hyperabrupt-junction diode.

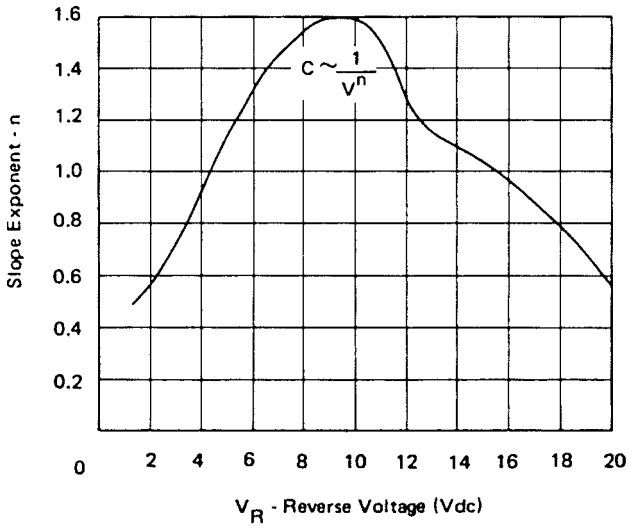


Figure 2-37 Typical *n* versus reverse voltage characteristic for hyperabrupt-junction diode.

capacitance change obtainable, the linearly graded junction is not used as a tuning diode. This structure found its greatest use several years ago as a "cube law" multiplier, but even this use has decreased as new structures have been developed.

Hyperabrupt Junction. The hyperabrupt-junction diode provides a greater capacitance change than the abrupt-junction diode for a given voltage change, as well as a linear frequency versus voltage characteristic over a limited voltage range. The structure of the hyperabrupt-junction diode is shown in Figure 2-36 and can be seen to be an abrupt-junction diode with an additional, increased doping level at the *pn* junction. This diode also follows Eq. (2-42) with the exception that n is not a function of voltage and is generally in the range of 0.5–2. A typical curve of n versus voltage is shown in Figure 2-37.

The C–V curve in a hyperabrupt-junction diode is shown in Figure 2-38 and is seen to start at a high value of capacitance per unit area at low bias (high epitaxial doping) and change to a lower value of capacitance per unit area (low epitaxial doping) at high bias. The details of the curve depend on details of the shape of the more highly doped region near the *pn* junction.

Unfortunately, with a hyperabrupt-junction diode, you must settle for a lower Q than an abrupt-junction diode with the same breakdown voltage and same capacitance at 4 V.

It should be noted that any diode that has an n value that exceeds 0.5 at any bias voltage is, by definition, a hyperabrupt-junction diode. Thus, the hyperabrupt-junction diode family can have an infinite number of different C–V curves. Since the abrupt-junction diode has a well-defined C–V curve, the capacitance value at one voltage is sufficient to define the C–V capacitance at any other voltage. This is not the case for the hyperabrupt-junction diode. In order to adequately define the C–V characteristics of a hyperabrupt-junction diode, two and sometimes three points on the curve must be specified.

Silicon Versus Gallium Arsenide. Everything mentioned so far applies to both silicon and gallium arsenide (GaAs) diodes. The main difference between silicon and GaAs from a

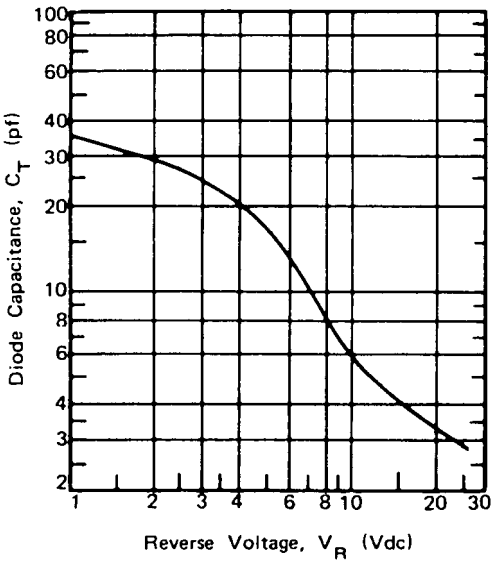


Figure 2-38 Capacitance versus junction bias for hyperabrupt-junction diode.

user's point of view is that higher Q can be obtained from GaAs devices. This is due to the lower resistivity of GaAs from a given doping level N. The resistivity of the epitaxial layer, or substrate, of a diode is given by

$$\rho = \frac{1}{Ne\mu} \tag{2-43}$$

where ρ = the resistivity
N = the doping level of the layer
e = the charge on an electron
μ = is the mobility of the charge carriers in the layer

Gallium arsenide has a mobility about four times that of silicon and, thus, a lower resistivity and higher Q for a given doping level N. Since diode capacitance is proportional to $\sqrt{N}$, independent of resistivity, a silicon diode and a GaAs diode of equal area and doping will have a capacitance difference proportional to the square root of the dielectric constant ratio. This gives the GaAs diode a 5% higher capacitance and is thus of little practical significance. The penalty paid for using GaAs is an unpassivated diode and a more expensive diode due to higher material and processing costs. If the higher Q of the GaAs device is not really needed, a substantial price saving will be obtained by using a silicon device.

Planar Versus Mesa Construction. The two basic construction techniques used to manufacture tuning diodes are planar and mesa; cross sections for both of these devices are shown in Figure 2-39. The planar process, which is the backbone of the integrated circuit industry, lends itself to large-volume production techniques and is the one used for the 1N series of tuning diodes. Mesa processing, on the other hand, requires more

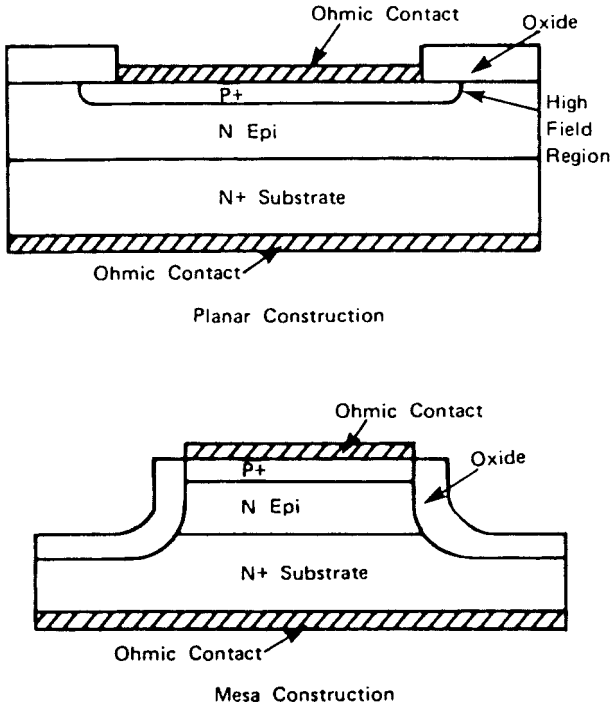


Figure 2-39 Cross sections of planar and mesa devices.

processing steps and is generally done on a wafer-by-wafer basis. This results in a more costly process and thus a more expensive diode. All microwave tuning diodes are of mesa design because of greatly higher Q. Due to the relatively small radius of curvature at the junction edge of a planar diode, the electric field in this area is greater than the electric field in the center (flat) portions of the junction. As a result, the breakdown voltage of the diode is determined by both the epitaxial resistivity and the radius of curvature of the junction edge. Thus, for a given breakdown voltage, a planar diode must use higher resistivity epitaxial material than a mesa diode, which has a completely flat junction. The end result is that the planar diode has a greater series resistance than a mesa diode for the same capacitance and breakdown voltage, and thus lower Q.

Capacitance

Capacitance Ratio. From the user's point of view, the capacitance ratio is simply the capacitance change available in the circuit. Thus, a user tuning from, say, –4 to –45 volts defines capacitance ratio as

$$R = \frac{C_T(-4)}{C_T(-45)} \tag{2-44}$$

where C_T includes C_J plus C_P plus C_F. The manufacturer, however, defines C_T as $C_J + C_P$.

To explore the significance of this difference, let's take two examples, one with a large C_J and another with a small C_J, in chip, package, and "typical" fringe situations. Both are 45-V tuning varactors.

Device	C_{J0}	C_{J45}	Ratio
A	0.6 pF	0.1 pF	6.0
B	15.0 pF	2.5 pF	6.0

Put both devices in a standard 023 package with C_P (strap and ceramic) of 0.18 pF:

Device	C_{T0}	C_{T45}	Ratio
C	0.78 pF	0.28 pF	2.75
D	15.18 pF	2.68 pF	5.67

Notice the drop in capacitance ratio, especially for the low-C_J diode. If we now add a typical 0.04 pF for external fringe capacitance, we get the following results:

Device	C_{T0}	C_{T45}	Ratio
C	0.82 pF	0.32 pF	2.56
D	15.22 pF	2.72 pF	5.6

The reduction in capacitance ratio, and thus circuit tuning capability, by the fringing fields is quite obvious and amounts to 7% in this example.

Because of the often stringent specifications on tuning ratio, it is mandatory that the manufacturer and customer clearly agree on the exact design of the holder used to measure the varactor in question.

Having described how to measure capacitance, it is relatively easy to describe the results. The section on diode physics described the various types of "laws," or C–V curves, and we will not repeat them here. Nonetheless, several important points must be covered.

The first is "available capacitance swing." The laws indicate a steadily decreasing capacitance with voltage, which indicates that the epitaxial region is widening and the electric field is increasing. (For an abrupt junction, since $C \propto 1/\sqrt{V}$, the depletion zone with W is increasing as $\sqrt{V}$, and the electric field V/W increases as $\sqrt{V}$.)

Two things can happen:

1. The junction width widens so that the entire intrinsic region is depleted. The capacitance bottoms out, resulting in voltage punchthrough.
2. The electric field exceeds the dielectric strength of silicon (or GaAs), and "solid-state discharge" or "avalanche" current is drawn.

The diode impedance drops, the varactor no longer "varacts," and circuit operation ceases. Moreover, if more than a few milliamperes of current is drawn, localized overheating may destroy the diode, resulting in breakdown voltage. All varactors are characterized for breakdown voltage—for example, 45 V minimum.

The theoretical tuning varactor is designed so that the punchthrough occurs at a voltage equal to the voltage breakdown of the diode. Logically, this means that in order to obtain

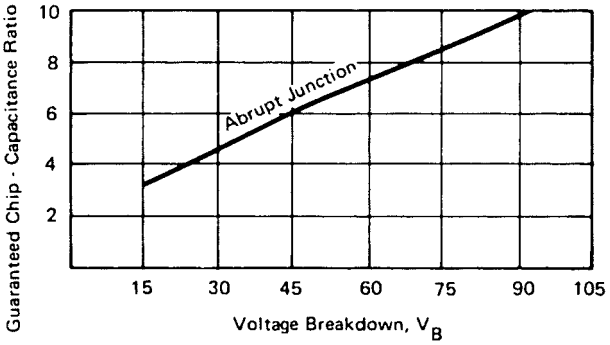


Figure 2-40 Capacitance ratio versus breakdown voltage.

greater tuning ratios, it is necessary to be able to increase the depletion-layer width without reaching punchthrough or breakdown. You must have a thicker epitaxial region to make this possible.

Figure 2-40 shows catalog ratio values, from zero bias to breakdown, as a function of breakdown voltage necessary. In the next section, on Q, we will discuss other elements in your choice of V_B.

To complete this section, we should mention that semiconductor processing control has been refined so well that capacitance tracking to within ±1% over the full range from zero to breakdown is now readily obtainable in production quantities.

Temperature Coefficient of Capacitance (T_{CC}). Unfortunately, since most datasheets give the value of T_{CC} at 4 V, it is sometimes assumed that this value applies at all bias voltages. This is not the case. Consider Eq. (2-45), a rewritten form of Eq. (2-42):

$$C(V) = \frac{C(O)}{(V + \varphi)^n} \tag{2-45}$$

Taking the derivative of this with respect to temperature T we have

$$\frac{dC(V)}{dT} = \frac{+nC(O)}{(v + \varphi)(v + \phi)^n} \frac{d\varphi}{dT} \tag{2-45a}$$

or, after substituting Eq. (2-45):

$$T_{CC} = \frac{1}{C(V)} \frac{dC(V)}{dT} = \frac{-n}{(V + \varphi)} \frac{d\varphi}{dT} \tag{2-46}$$

As a first approximation, we can say that $d\varphi/dT = -2.3$ mV/°C over the temperature range of interest.

From Eq. (2-46) we can draw the following conclusions:

1. The temperature coefficient is inversely proportional to the applied voltage.
2. The temperature coefficient is directly proportional to the diode slope, n.

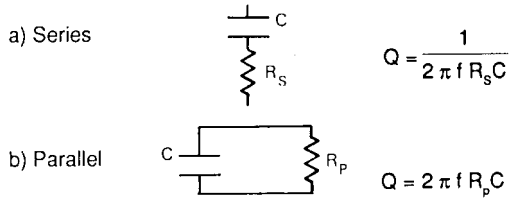
For an abrupt-junction diode that has a constant value of n (0.5), the temperature coefficient has a smooth curve in the form $K/(v + \varphi)$. However, in the case of hyperabrupt-junction diodes, n is a function of voltage, and the shape of the T_{CC} curve depends on the details of the $n(V)$ curve. A typical T_{CC} curve for an abrupt-junction diode is shown in Figure 2-41 and for an Alpha DKV6520 series hyperabrupt-junction diode in Figure 2-42. The inflection in the hyperabrupt T_{CC} is due to the fact that in this voltage range $n(V)$ is increasing faster than $1/V$, giving an increase in T_{CC}. It should also be noted, however, that over the range of the T_{CC} minimum the temperature coefficient is relatively constant, and operation in this area may be advantageous in some applications where a restricted tuning range can be used.

Q Factor or Diode Loss

Definitions. The classical definition of the Q of any device or circuit is

$$Q = \frac{2\pi \times \text{Energy stored}}{\text{Energy dissipated per cycle}} \tag{2-47}$$

For a capacitor, two formulations are possible:



Clearly, the two definitions must be equal at any frequency, which establishes

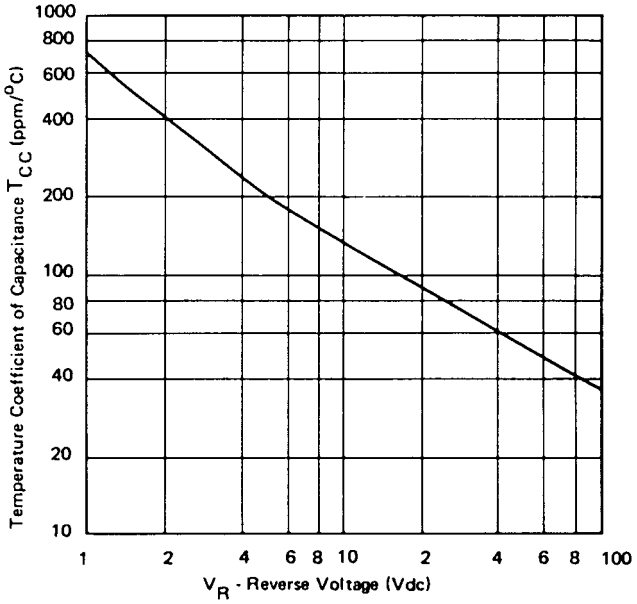


Figure 2-41 Temperature coefficient of capacitance versus tuning voltage for an abrupt-junction diode.

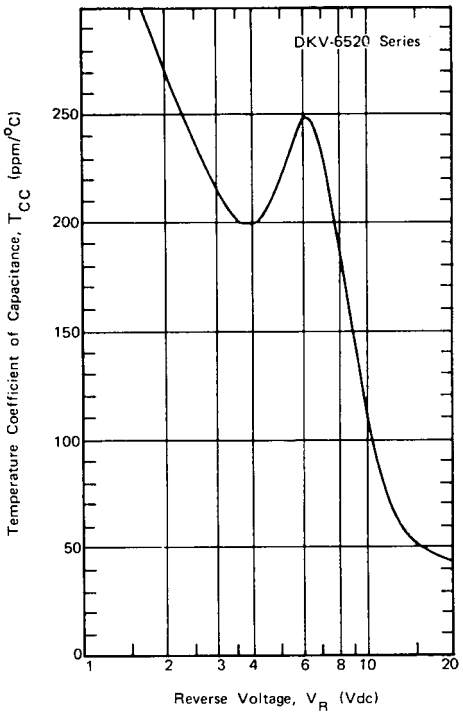


Figure 2-42 Temperature coefficient of capacitance versus tuning voltage ($T_A = 25$ °C) hyperabrupt-junction diode.

$$R_P = \frac{1}{(2\pi f)^2 C^2 R_S} \tag{2-48}$$

In the case of a high-Q tuning diode, the proper physical model is the series configuration, because the depleted region is almost perfectly pure capacitance, and the undepleted region, due to its relatively low resistivity, is almost a pure resistor in series with the capacitance. Furthermore, the contact resistances are also clearly in series. Q, then, for a tuning diode, is given by

$$Q_{(-v)} = \frac{1}{2\pi f_o R_{(-v)} C_{(-v)}} \tag{2-49}$$

where f_o is the operating frequency, $C_{(-v)}$ is the junction capacitance, and $R_{(-v)} = R(\text{epi}) + R_{C'}$, the sum of the resistance of the undepleted epitaxial layer and the fixed contact resistance.

Cutoff frequency, f_c, is defined as that frequency at which Q equals unity. Thus,

$$f_{c_{(-v)}} = \frac{1}{2\pi R_{(-v)} C_{(-v)}} \tag{2-50}$$

Historically, the tuning diode business has specified Q at 50 MHz, despite the fact that Q values of microwave diodes are so high that it is almost impossible to measure them at 50

Table 2-7 Parameters for 0.6-pF diode

Breakdown Voltage V_B	Resistivity ($\Omega \cdot$ cm)	Junction Diameter (mills)	Epitaxial Region (μm)	Depleted Epitaxial $V = -4$ V (μm)	Undepleted Epitaxial $V = -4$ V (μm)	R Undepleted (Ω)	R_{sp} (Ω)	Q_4
30	0.31	2.4	1.37	0.54	0.83	0.81	0.18	5200
45	0.52	2.8	2.25	0.73	1.52	1.86	0.15	2600
60	0.74	3.2	3.20	0.90	2.30	3.24	0.13	1500
90	1.25	3.7	5.27	1.21	4.06	7.17	0.10	700

MHz. Instead, as discussed below, Q is measured at microwave frequencies (e.g., 1–3 GHz) and related to 50 MHz by the relationship

$$Q_{(f1)} = Q_{(f2)} \frac{f_2}{f_1} \tag{2-51}$$

which derives quickly from the assumption that f_c is independent of the measuring frequency.

Since both junction capacitance and epitaxial layer resistance are functions of the applied bias, it is not possible to calculate Q as a function of bias from a measurement of capacitance alone. Catalog specifications typically show Q at –4 V, together with the capacitance at two or more voltages.

Relative to $Q_{(-4)}$, Q increases faster than the reduction in capacitance for bias greater than 4 V and, conversely, decreases faster for bias less than 4 V.

In the following section, we will discuss the diode design parameters that determine Q. Following this, we will describe some elementary Q measurement techniques.

Causes. In the discussion of device physics, the resistivity of the epitaxial region was discussed, together with its impact on punchthrough and breakdown. For example, Table 2-7 supplies typical resistivity and relative parameters of 0.6 pF ($C_{J(4)}$) diodes of different breakdowns. If we remember that if at any bias lower than breakdown the epitaxial region is not completely depleted, it follows that the undepleted portion presents a resistance in series with the pure capacitance of the depleted zone. The magnitude of this "undepleted" resistance is also shown in the table.

The entry R_{sp} ($R_{spreading}$) is the series resistance between the epitaxial region and the low-resistivity substrate. The calculations are for idealized cylindrical epitaxial regions of uniform resistivity, low-resistivity contact on the anode (top), and low-resistivity substrate on the cathode. This resistance is constant, independent of bias; also shown are epitaxial thickness region and width of the depletion zone at –4 V bias.

Note the substantial reduction in Q for high-voltage diodes caused by the increased epitaxial layer resistance; this is true for any value of capacitance or any type of junction. For greater voltage breakdown, the epitaxial region thickness must be increased, which requires an increase in epitaxial layer resistivity; the higher the resistivity of the undepleted zone, multiplied by the fact that it is much wider for high-voltage diodes, means the resistance increases substantially.

Consequently, a rule of thumb emerges: For maximum Q, never choose a tuning diode with a voltage breakdown in excess of what is needed for the necessary tuning range. If the

Table 2-8 ***Q*** **versus bias for 0.6-pF diode**

Breakdown Voltage V_B	C_{J0}	$Q(0)$	$C_J(-4)$	$Q(-4)$	$C_J(-10)$	Q (−10)	$C_J(-30)$	$Q(-30)$	$C_J(-45)$	$Q(-45)$	$C_J(-60)$	$Q(-60)$	$C_J(-90)$	Q(−90)
30	1.43	1700	0.6	5200	.4	10k	0.23	64k						
45	1.43	850	0.6	2600	.4	5k	0.23	20k	0.19	90k				
60	1.43	550	0.6	1500	.4	3k	0.23	9k	0.19	23k	0.17	170k		
90	1.43	270	0.6	700	.4	1.3k	0.23	3.5	0.19	6k	0.17	10k	0.14	220k

Table 2-9 Capacitance ratios (C_{J0}/C_JV_B)

Breakdown Voltage V_B	Optimum Ratio	Minimum Guaranteed Ratio	$Q_{(-4)}$ Typical
30	6.2	4.5	3000
45	7.5	6.0	2500
60	8.4	7.5	1400
90	10.2	8.7	650

required tuning range is an octave, requiring a 4:1 ratio, the selection of a 30-V diode will result in diode losses half those of a 60-V diode.

Table 2-8 lists capacitance and Q for each of these chips as a function of bias. Please remember that Q is calculated at 50 MHz.

Table 2-9 rewrites the data of Table 2-8 to show available capacitance ratios between zero bias and breakdown. The first column is the theoretical optimum, as tabulated. The second column is the typical catalog specification. The reduction in tuning ratio below theoretical optima is caused by nonideal junction fabrication. The junctions are never perfectly abrupt.

Although the tables and numbers given refer to abrupt-junction silicon devices, the principle applies without exception to all types of tuning diodes. For comparison, Table 2-10 lists available ratios and Q values for a number of different varactors. The high Q values for GaAs and the low values for hyperabrupts are apparent.

One last point: The Q values and series resistance refer to chips only. The effects of package parasitics will be discussed later, but it is important to consider circuit contact losses here. For low-capacitance diodes—for example, $C_{J4} = 0.6$ pF—the epitaxial region contributes a high value of resistance and dominates Q except at punchthrough. Diode contact losses are less significant.

Post-Tuning Drift. Post-tuning drift (PTD) is the change in oscillator frequency with time after the tuning voltage has stabilized. The minimization of PTD has assumed greater importance with the design of more sophisticated electronic countermeasure systems, where rapid, accurate frequency changes are required.

Table 2-10 Comparative Tuning Diodes

Type	C_{J0}	Breakdown Voltage V_B	$Q_{(-4)}$[a]	Ratio C_{J0}/C_JV_B[a]
Silicon abrupt	1.0	30	5,000	4.5
Silicon abrupt	2.5	30	4,600	4.5
Silicon abrupt	5.0	30	3,800	4.5
Gallium arsenide abrupt	1.0	25	10,000	3.6
Gallium arsenide abrupt	0.5	10	17,000	2.5
Silicon hyperabrupt	50.0	22	300	17.0
Silicon hyperabrupt	2.5	22	500	14

[a] Minimum guaranteed.

Post-tuning drift can be characterized as short-term and long-term. Short-term PTD occurs in the time range of tens of nanoseconds to a few seconds, while long-term PTD occurs in the time range of seconds to minutes, hours, or days.

Short-term PTD is mainly dependent on the thermal properties of the diode and is improved by high Q (low power loss) and flip chip construction. Long-term PTD depends on oxide stability and freedom of mobile charge in the oxide. It should be noted that actual oscillation frequency change may occur even with a perfect tuning diode because of variation with frequency in the power dissipated by the diode, changes in the diode heat-sink temperature, and frequency changes due to other circuit elements. Less than 0.01% short- and long-term PTD can be obtained.

Distortion Products. Inasmuch as nonlinear components generate harmonics and other distortion products, an understanding of this mechanism is of prime interest to the circuit designer. In some instances, the distortion products are the desired end result of the circuit design, as in frequency multipliers, where harmonics of the input signal frequency are the required output signal. For other applications, such as tuning-diode-tuned linear circuits, distortion products are extremely undesirable, and in some instances the end product specification may set a maximum limit on the distortion products allowed.

Cross-Modulation. Cross-modulation is the transfer of the modulation on one signal to another signal and is caused by third-order and higher odd-order nonlinearities in the behavior of the device. Rewriting Eq. (2-42) we have

$$C(V) = \frac{C_0}{(1 + V/\varphi)^n} \tag{2-52}$$

where C_0 = capacitance
V = applied voltage = $V_0 + v$
V_0 = dc applied voltage
v = ac applied voltage

Then, for a desired signal of

$$S_1 = v_1 \sin \omega_1 t \tag{2-53}$$

and a second, amplitude-modulated signal of

$$S_2 = v_2(1 + m \cos \omega_m t)\sin \omega_2 t \tag{2-54}$$

it can be shown that the cross-modulation, γ, defined by

$$\text{Output signal} \sim v_1 \sin \omega_1 t + \gamma \sin (\omega_1 \pm \omega_m)t \tag{2-55}$$

is found to be

$$\gamma = \frac{n(n+1)mv_2^2}{4(V_0+\varphi)^2} \tag{2-56}$$

From this equation it can be seen that cross-modulation is:

- Proportional to the square of the interfering signal
- Directly proportional to the interfering signal's modulation index, m
- Independent of the strength of the desired signal
- Independent of the frequencies of the desired and interfering signals (assuming that the nonlinearity to which both signals are subjected is sufficiently frequency-indiscriminate so this is the case)
- Present for all values of n; that is, no value of n gives zero cross-modulation

Solving Eq. (2-56) for the signal level v_2 required to produce cross-modulation of value γ, we have

$$v_2 = \frac{2(V_0+\phi)\gamma}{n(n+1)^m} \tag{2-57}$$

The interfering signal levels required to produce 1% cross-modulation from a 30%-modulated interfering signal applied to an abrupt-junction diode and a hyperabrupt-junction diode are shown in Figure 2-43.

From this figure, it can be seen that the hyperabrupt-junction diode is more susceptible to cross-modulation than the abrupt-junction diode in the region of maximum slope of the

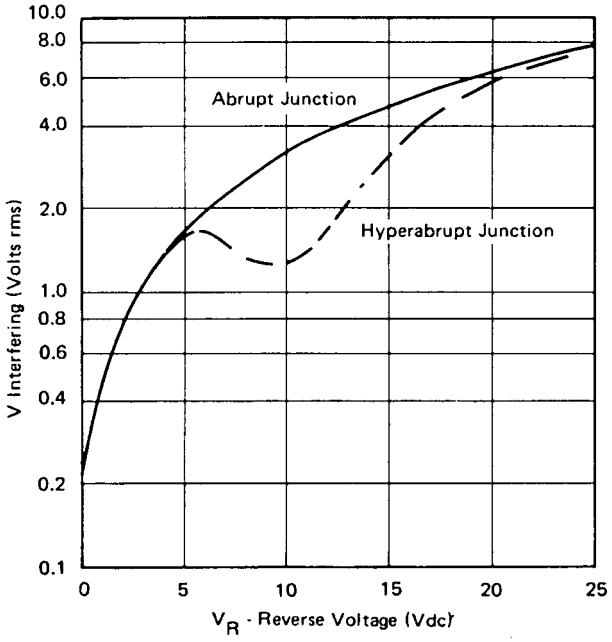


Figure 2-43 Interfering signal (30% amplitude modulated) level versus bias for 1% cross-modulation for abrupt-junction and hyperabrupt-junction diodes.

hyperabrupt-junction diode. For many applications, however, distortion products will be generated by other devices, such as transistors, at signal levels considerably below those given in Figure 2-43.

Intermodulation. Intermodulation is the production of undesired frequencies in the form

$$\sin(2\omega_1 t - \omega_2 t) \quad \text{and} \quad \sin(\omega_1 t - 2\omega_2 t) \tag{2-58}$$

from an input signal in the form of

$$v(\cos \omega_1 t + \cos \omega_2 t) \tag{2-59}$$

From an analysis similar to that done for cross-modulation, it can be shown that

$$\text{Intermodulation} = \frac{n(n+1)v^2}{8(V_0 + \varphi)^2} \tag{2-60}$$

or

$$\text{Cross-modulation} = (2m) \times \text{Intermodulation} \tag{2-61}$$

Harmonic Distortion. Harmonic distortion products are integral multiples of the signal frequencies and decrease in amplitude as the harmonic number decreases. Due to passband considerations and amplitude decrease with harmonic number, the second harmonic is the one of prime concern. Again, it can be shown that the second harmonic, v_2, of a signal of amplitude v_1 is

$$v_2 = \frac{n}{3(V_0 + \phi)} v_1^2 \tag{2-62}$$

Figure 2-44 shows the signal level required to produce 10% second-harmonic distortion in an abrupt-junction and a hyperabrupt-junction diode. Again, as in the case of cross-modulation, the hyperabrupt-junction diode is slightly worse than the abrupt-junction diode in the region of maximum slope of the hyperabrupt-junction diode.

Reduction of Distortion Products. In some cases, the signal levels applied to the diode generate distortion products larger than desirable for the circuit application. In this case, significant reduction in the distortion products can be achieved by using two diodes in a back-to-back configuration, as shown in Figure 2-45. Analysis shows that the fundamental signal components through the diodes are in phase and add, while some distortion products are out of phase and cancel, thus improving distortion performance.

Since the gradient of the electrical field produced in the depletion layer by a reverse bias applied to the device is proportional to the space charge density, the following equations can be written for the junction width W as a function of the reverse bias V_R:

For an Abrupt *pn* Junction (Alloyed Diodes)

$$W = {}_2\sqrt{2\frac{\varepsilon_r \varepsilon_0}{q}\left(\frac{1}{C_p} + \frac{1}{C_n}\right)(V_R + V_D)} \tag{2-63}$$

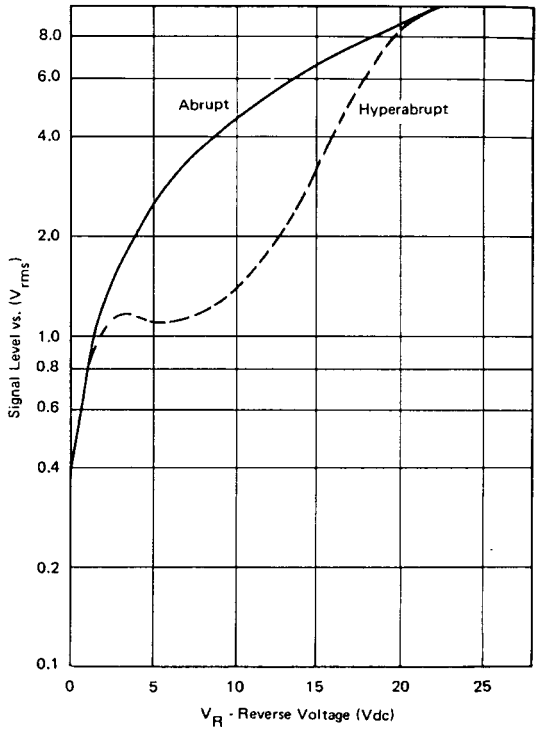


Figure 2-44 Signal level versus reverse voltage for 10% harmonic distortion.

For Linear *pn* Junctions (Single-Diffused Diodes, such as BA110–BA112)

$$W = \sqrt[3]{12\frac{\varepsilon_r \varepsilon_0}{aq}(V_R + V_D)} \tag{2-64}$$

where a is the impurity gradient within the depletion layer, ε_0 is the absolute dielectric constant ($8.85 \times 10^{-14}\ A_s/V_{cm}$, and $\varepsilon_r \approx 12$, the relative dielectric constant of silicon.

The junction capacitance, which is inversely proportional to the junction width, therefore varies in alloyed diodes with the square root, and in single-diffused diodes with the cube root of the externally applied reverse bias, and can be calculated from the general equation

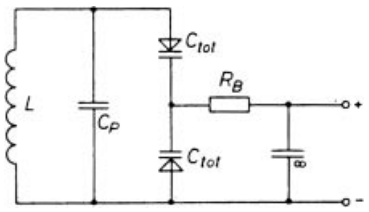


Figure 2-45 Back-to-back diodes. C_P is a fixed parallel capacitance, R_B is a bias decoupling resistor, and the capacitor marked ∞ is a low-impedance bypass.

$$C = \frac{\varepsilon_r \varepsilon_0 S}{W} \tag{2-65}$$

where S is the surface area of the *pn* junction. By way of approximation, we can also use the equation

$$C = \frac{K}{(V_R + V_D)^n} \tag{2-66}$$

where all constants and all parameters determined by the manufacturing process are contained in K. The exponent n is a measure of the slope of the capacitance/voltage characteristics and is 0.5 for alloyed diodes, 0.33 for single-diffused diodes, and (on average) 0.75 for tuner diodes with a hyperabrupt *pn* junction. Figure 2-46 shows the capacitance/voltage characteristics of an alloyed, a diffused, and a tuner diode.

Recently, an equation has been indicated which, although purely formal, describes the practical characteristics better than Eq. (2-66):

$$C = C_0 \left(\frac{A}{A + V_R} \right)^m \tag{2-67}$$

where C_0 is the capacitance at $V_R = 0$, and A is a constant whose dimension is a voltage. The exponent m is much less dependent on voltage than the exponent n in Eq. (2-66).

Equations (2-63) through (2-67) express the pure junction capacitance of the capacitance diode, but to this must still be added a constant capacitance, determined by structure parameters, in order to obtain the diode capacitance C_{tot}, which interests the user. With high inverse voltages—that is, low junction capacitance—a difference will therefore arise between the theoretical capacitance/voltage characteristic according to Eq. (2-66) and the practical characteristic, as shown in Figure 2-47.

The operating range of a capacitance diode or its useful capacitance ratio

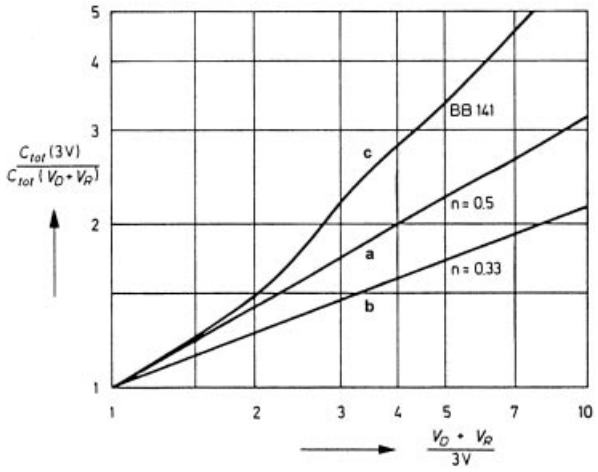


Figure 2-46 Capacitance/voltage characteristic for (a) an alloyed capacitance diode, n = 0.5; (b) a diffused capacitance diode, n = 0.33; and (c) a wide-range tuner diode (BB141).

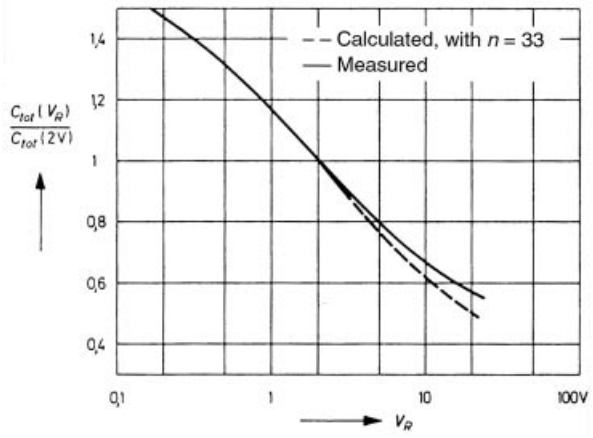


Figure 2-47 Capacitance/voltage characteristic of the BA110 diode.

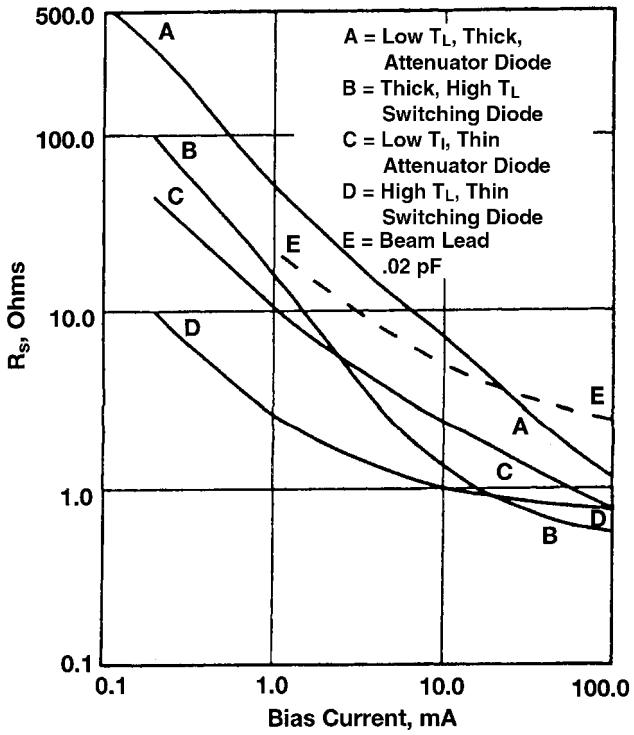


Figure 2-48 Basic current/voltage and capacitance/voltage characteristics.

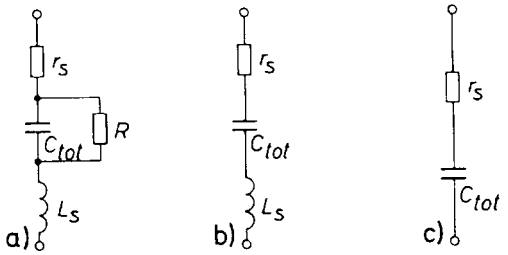


Figure 2-49 Equivalent circuits for capacitance diodes: (a) complete circuit, (b) simplified circuit, and (c) further simplification for low frequencies.

$$\frac{C_{\max}}{C_{\min}} = \frac{C_{\text{tot}}(V_{R\min})}{C_{\text{tot}}(V_{R\max})} \tag{2-68}$$

is limited by the fact that the diode must not be driven by the alternating voltage superimposed on the tuning voltage either into the forward mode or the breakdown mode. Otherwise, rectification, which would shift the diode's bias and considerably affect its figure of merit, would take place. Figure 2-48 plots the capacitance/voltage characteristics of a capacitance diode to clarify the relationship. The useful operating range lies between the voltages

$$V_{\min} > \hat{V} - V_F \quad \text{and} \quad V_{\max} < V_{(\text{BR})R} - \hat{V} \tag{2-69}$$

As has already been indicated, the exponent n for large-capacitance-ratio or tuning diodes used today for TV tuners is not constant but voltage-dependent and subject to manufacturing tolerances. This means that the capacitance/voltage characteristic of these diodes is likewise subject to manufacturing tolerances. Since, in a TV tuner, it is necessary for two or three circuits to be tuned uniformly, tuner diodes must be selected empirically for identical characteristics and supplied in equipment lots.

Electrical Properties of Tuning Diodes. In this section, the electrical properties of capacitance diodes are described, with reference to data published on the tuner diodes BB141 and BB142.

Equivalent Circuit. Since a capacitance diode is not an ideal capacitor, it is useful to introduce an equivalent circuit that can serve as a basis for discussing the diode's electrical properties. Figure 2-49 shows various types of equivalent circuits.

The complete equivalent circuit (Figure 2-49a), which conforms closely to physical conditions, comprises, in addition to the diode capacitance C_{tot}, the series resistance r_S and the series inductance L_S, and also the reverse resistance $R = dV_R/dI_R$. At higher frequencies, this resistance can usually be disregarded, so that the equivalent circuit is simplified in terms of Figure 2-49b—the configuration usually employed. In many cases, the series inductance can also be disregarded, in which case we obtain Figure 2-49c.

Capacitance. Figure 2-50 shows the capacitance/voltage characteristic borrowed from the datasheets of diodes BB141 and BB142. As this characteristic, like most characteristics published in datasheets, represents merely a typical curve, it would not convey sufficient

information for the user to dimension a tuner. Therefore, the datasheets contain additional data on capacitance and useful capacitance ratio:

Capacitance		BB141A (pF)	BB141B (pF)	BB142 (pF)
At $V_R = 1$ V	C_{tot}	16	19	17
At $V_R = 3$ V	C_{tot}	11	13	12
At $V_R = 25$ V	C_{tot}	2–2.35	2.25–2.65	2–3

$$\text{Useful capacitance ratio} = \frac{C_{tot}(3\text{ V})}{C_{tot}(25\text{ V})} = 4 \text{ to } 6$$

On the basis of the spread of the diode capacitance C_{tot} guaranteed for $V_R = 25$ V, it is possible to calculate the tuner with the aid of the spread guaranteed for the capacitance ratio.

Three further capacitance graphs are shown in Figures 2-51 through 2-53. Figure 2-51 shows a (typical) curve representing the normalized slope as a function of the reverse voltage, Figure 2-52 the normalized capacitance as a function of the junction temperature, and Figure 2-53 the temperature coefficient of the capacitance as a function of the reverse voltage.

The variation of diode capacitance with ambient temperature, as shown in Figures 2-52 and 2-53, is really only a function of the temperature dependence of the diffusion voltage [see Eq. (2-66)]. The diffusion voltage drops with rising temperatures by about 2 mV/°C , which means that the diode capacitance rises with temperature. The influence of the diffusion voltage, and thus the temperature coefficient of the capacitance, decreases as the reverse voltage rises. It is therefore advisable to run a capacitance diode with as high a reverse voltage as the required capacitance ratio permits. Compensation for the dependence of capacitance on temperature will be discussed later.

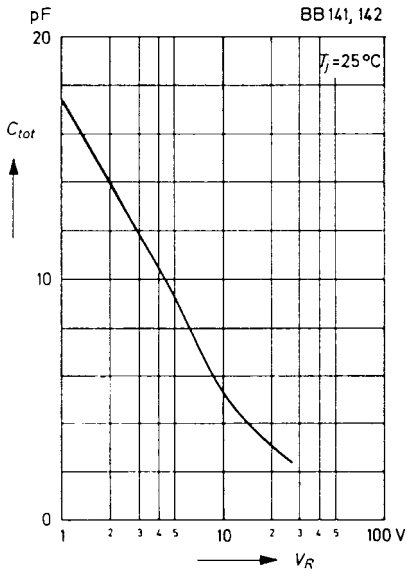


Figure 2-50 Capacitance/voltage characteristic of diodes BB141 and BB142.

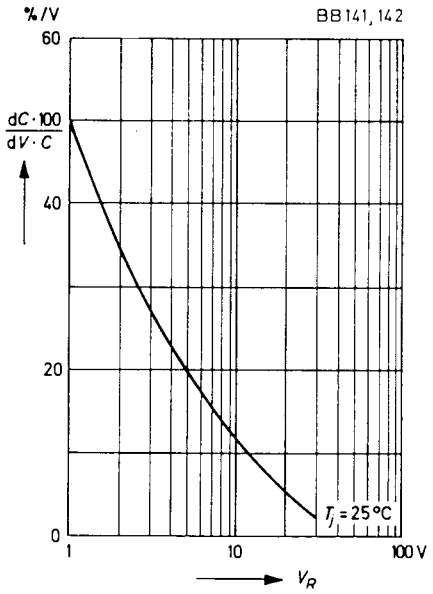


Figure 2-51 Slope (normalized) as a function of the reverse voltage.

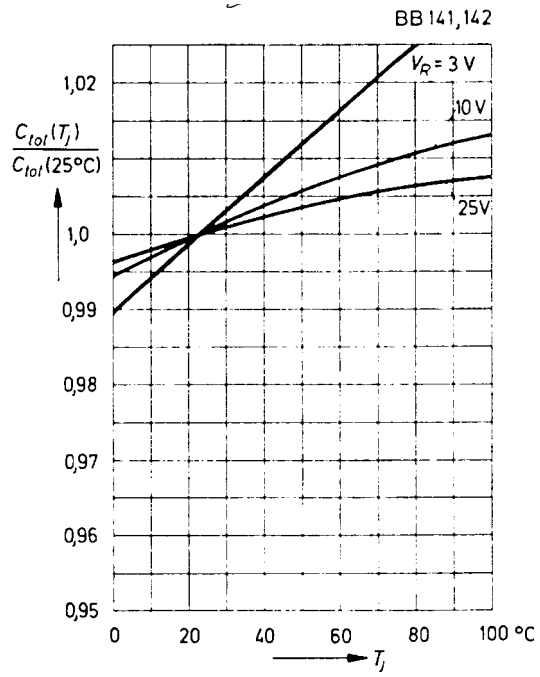


Figure 2-52 Capacitance (normalized) as a function of the junction temperature.

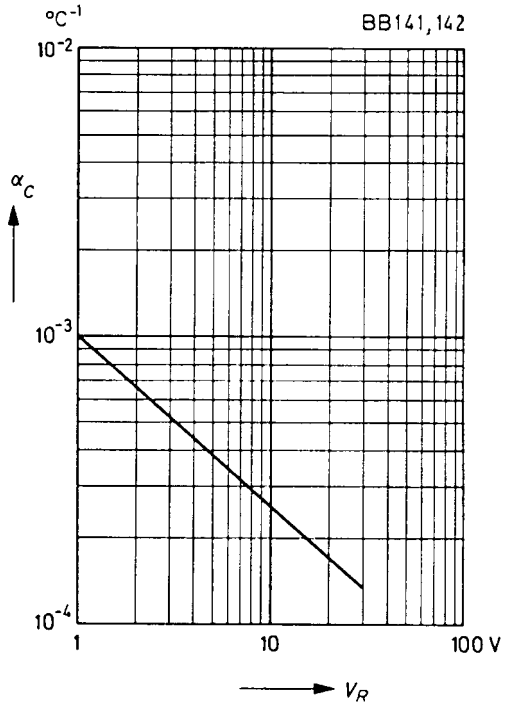


Figure 2-53 Temperature coefficient of capacitance as a function of the reverse voltage.

Series Resistance, Figure of Merit (Q). Since capacitance diodes are intended for employment in resonant circuits, indications are required as to the circuit attenuation that they produce. Usually, the series resistance or the figure of merit that can be calculated therefrom is indicated. With reference to the equivalent circuit shown in Figure 2-49a, the figure of merit (or Q factor) is calculated:

$$Q = \frac{1}{\omega C_{\text{tot}} r_S + 1/\omega C_{\text{tot}} R} \tag{2-70}$$

Figure 2-54 shows the theoretical (normalized) Q as a function of frequency. With the frequency range of 10–1000 MHz, which is essential for tuner diodes, the parallel resistance R caused by the reverse current (see Figure 2-49b) can be disregarded, and Eq. (2-70) can be simplified to read as follows:

$$Q = \frac{1}{\omega C_{\text{tot}} r_S} \tag{2-71}$$

In Figure 2-55, Q is plotted as a function of frequency for the tuner diodes BB141 and BB142. The datasheet indicates, in addition to this curve, the value of the series resistance at $f = 470$ MHz and $C_{\text{tot}} = 9$ pF:

	BB141A, BB141B	BB142
Series resistance (r_S)	0.6 (< 0.8)	0.9 (< 1.2)

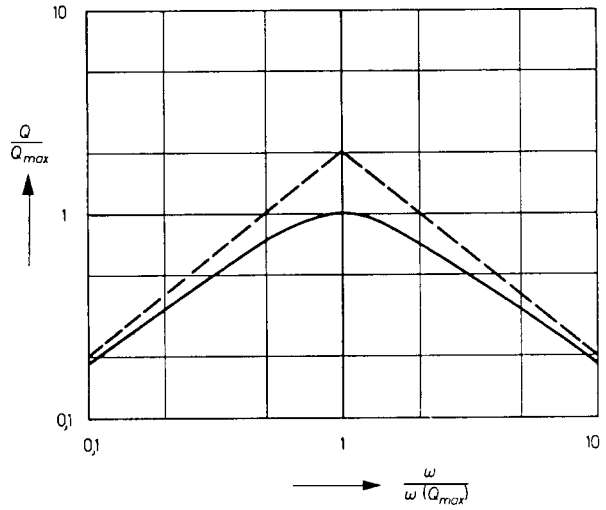


Figure 2-54 *Q* (normalized) as a function of frequency.

Since the diode capacitance and the series resistance decrease as the reverse voltage rises, Q is likewise dependent on the reverse voltage. It is smallest at a low reverse voltage, that is, at a low frequency. The series and parallel capacitances inevitable in a diode-tuned resonant circuit influence the effective Q of the tuning capacitance, as will be described later.

Series Inductance, Series Resonant Frequency, Cutoff Frequency for Q = 1. These three parameters define the behavior of the diode at high frequencies when the influence of

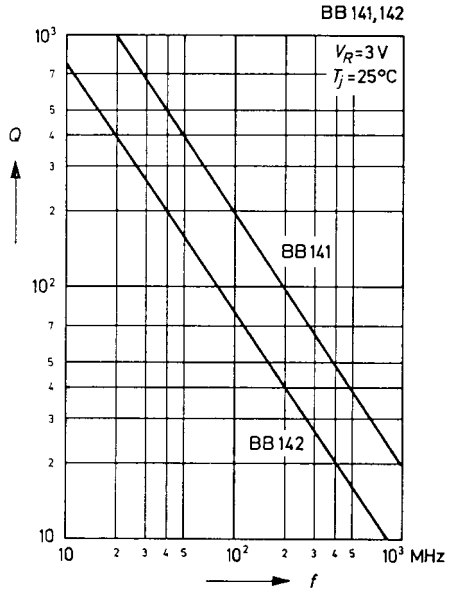


Figure 2-55 *Q* as a function of frequency.

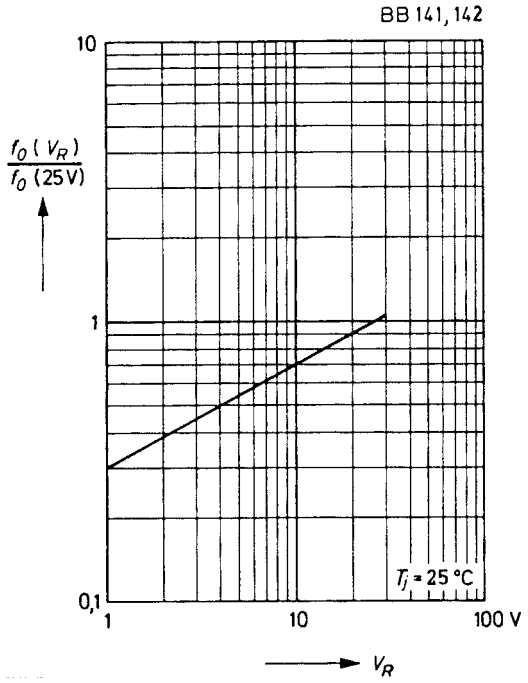


Figure 2-56 Series-resonant frequency (normalized) as a function of the reverse voltage.

the series inductance can no longer be neglected. The parameter on which the calculation is based is the series inductance L_S, which is determined by the connection leads. As indicated in Figure 2-49a, it is in series with the capacitance and, given a suitable frequency, forms a series-resonant circuit therewith. In that case

$$\omega L_S = \frac{1}{\omega C_{\text{tot}}} \tag{2-72}$$

and the series-resonant frequency is

$$f_0 = \frac{1}{2\pi\sqrt{C_{\text{tot}} L_S}} \tag{2-73}$$

Above this frequency, the impedance of the diode is inductive, but it depends on the reverse voltage. Figure 2-56 shows how the (normalized) series-resonant frequency depends on the reverse voltage for the tuner diodes BB141 and BB142.

The cutoff frequency for $Q = 1$ is defined as the frequency at which Q assumes the value of unity. In a rough calculation, the reactive impedance of the diode capacitance can be neglected and the frequency f_{Q1} obtained for the condition in which the inductive reactance equals the series resistance:

$$\omega L_S = r_S \tag{2-74}$$

$$f_{Q1} = \frac{r_S}{2\pi L_S} \tag{2-75}$$

For the tuner diodes BB141 and BB142, Figure 2-57 shows how the (normalized) cutoff frequency depends on the reverse voltage. The datasheets contain the following numerical information:

		BB141	BB142	
Series inductance measured 1.5 mm from the case	L_S	2.5	2.5	nH
Series-resonant frequency at $V_R = 25$ V	f_0	2	1.8	GHz
Cutoff frequency for $Q = 1$ at $V_R = 3$ V	f_{Q1}	24	16	GHz

Depending on the intended application, either the series-resonant frequency or the cutoff frequency for $Q = 1$ determines the maximum useful frequency range.

Leakage Current, Breakdown Voltage. These two parameters mark the dc performance of the diode when a reverse bias is applied to it. In the databook for the tuner diodes BB141 and BB142, the breakdown voltage is quoted as follows:

$$\text{Reverse breakdown voltage at } I_R = 100\ \mu\text{A}: \quad V_{(BR)R} > 30\ \text{V}$$

This means that the maximum reverse voltage that may be applied to the diode as a tuning voltage is 30 V.

The leakage current for the above-mentioned tuner diodes is guaranteed as follows:

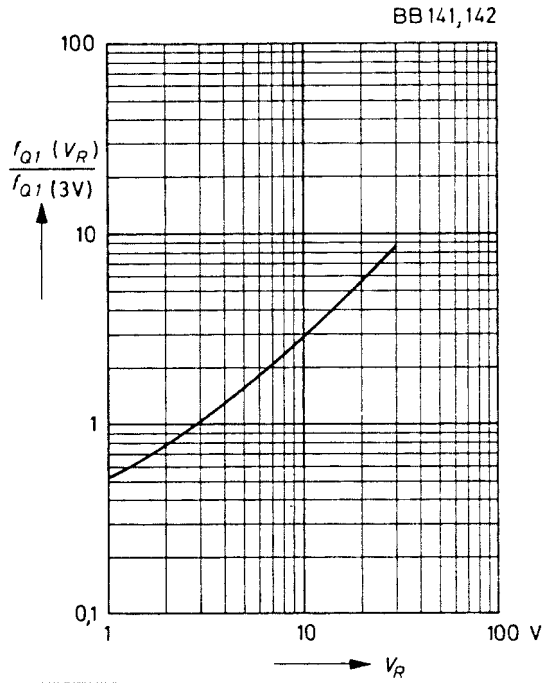


Figure 2-57 Cutoff frequency for $Q = 1$ (normalized) as a function of the reverse voltage.

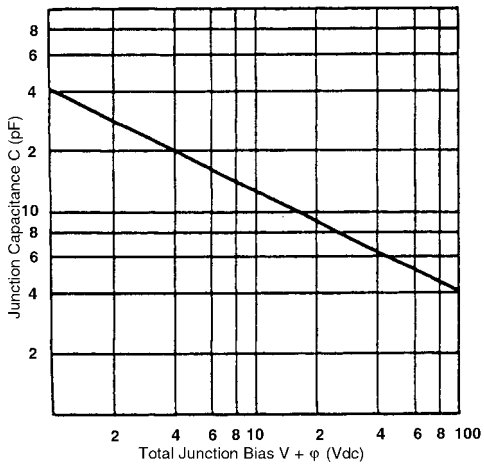


Figure 2-58 Leakage current as a function of reverse voltage.

$$\text{Leakage current at } V_R = 28 \text{ V:} \quad I_R < 50 \text{ nA}$$

Moreover, Figure 2-58 shows a graph that illustrates the leakage current as a function of the reverse voltage. Since this current is temperature dependent (as with every silicon diode)—it doubles with each temperature rise of about 10 °C—care should be taken when dimensioning the tuner circuit that the leakage current does not cause inadmissible voltage variations at increased ambient temperatures.

Matching of Tuner Diodes, Uniform Parameters. As has already been mentioned, the capacitance versus voltage characteristics of modern tuner diodes are subject to a certain scatter, so that it becomes necessary to test these diodes to obtain equipment lots empirically.

Diode-Tuned Resonant Circuits

Tuner Diode in the Parallel-Resonant Circuit. Figures 2-59, 2-60, and 2-61 illustrate three basic circuits for the tuning of parallel-resonant circuits by means of capacitance diodes.

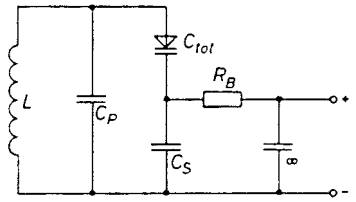


Figure 2-59 Parallel-resonant circuit with tuner diode, and bias resistor parallel to the series capacitor.

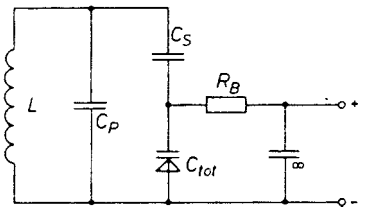


Figure 2-60 Parallel-resonant circuit with tuner diode, and bias resistor parallel to the diode.

In the circuit diagram of Figure 2-59, the tuning voltage is applied to the tuner diode via the input coil and the bias resistor R_B. Series-connected to the tuner diode is the series capacitor C_S, which completes the circuit for ac but isolates the cathode of the tuner diode from the coil and thus from the negative terminal of the tuning voltage. Moreover, a fixed parallel capacitance C_P is provided. The decoupling capacitor preceding the bias resistor is large enough for its value to be disregarded in the following discussion. Since for high-frequency purposes the biasing resistor is connected in parallel with the series capacitor, it is transformed into the circuit as an additional equivalent shunt resistance R_C. We have the equation

$$R_C = R_B\left(1 + \frac{C_S}{C_{\text{tot}}}\right)^2 \tag{2-76}$$

If in this equation the diode capacitance is substituted by the resonant circuit frequency ω, we obtain

$$R_C = R_B\left(\frac{\omega^2 LC_S}{1 - \omega^2 LC_P}\right)^2 \tag{2-77}$$

The resistive loss R_C caused by the bias resistor R_B is seen to be highly frequency dependent, and this may result in the bandwidth of the tuned circuit being independent of frequency if the capacitance of the series capacitor C_S is not chosen sufficiently high.

Figure 2-60 shows that the tuning voltage can also be applied directly and in parallel to the tuner diode. For the parallel loss resistance transformed into the circuit, we have the expression

$$R_C = R_B\left(\frac{1 + C_{\text{tot}}}{C_S}\right)^2 \tag{2-78}$$

and

$$R_C = R_B\left(\frac{\omega^2 LC_S}{\omega^2 L(C_S + C_P) - 1}\right)^2 \tag{2-79}$$

The influence of the bias resistor R_B in this case is larger than in the circuit of Figure 2-60, provided that

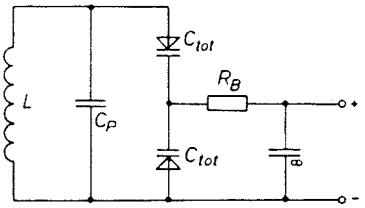


Figure 2-61 Parallel-resonant circuit with two tuner diodes.

$$C_S^2 > C_S(C_{\text{tot}} + C_P) + C_{\text{tot}}C_P \tag{2-80}$$

This is usually the case because the largest possible capacitance will be preferred for the series capacitor C_S, and the smallest for the shunt capacitance C_P. The circuit of Figure 2-59 is therefore normally preferred to that of Figure 2-60. An exception would be the case in which the resonant circuit is meant to be additionally damped by means of the bias resistor at higher frequencies.

In the circuit of Figure 2-61, the resonant circuit is tuned by two tuner diodes that are connected in parallel via the coil for tuning purpose, but series-connected in opposition for high-frequency signals. This arrangement has the advantage that the capacitance shift caused by the ac modulation (see section entitled "Modifying the Diode Capacitance by the Applied ac Voltage") takes effect in opposite directions in these diodes and therefore cancels itself. The bias resistor R_B, which applies the tuning voltage to the tuner diodes, is transformed into the circuit at a constant ratio throughout the whole tuning range. Given two identical, loss-free tuner diodes, we obtain the expression

$$R_C = 4R_B \tag{2-81}$$

Capacitances Connected in Parallel or Series with the Tuner Diode. Figures 2-59 and 2-60 show that a capacitor is usually in series with the tuner diode, in order to close the circuit for alternating current and, at the same time, to isolate one terminal of the tuner diode from the rest of the circuit with respect to direct current, so as to enable the tuning voltage to be applied to the diode. As far as possible, the value of the series capacitor C_S will be chosen such that the effective capacitance variation is not restricted. However, in some cases, as, for example, in the oscillator circuit of receivers whose intermediate frequency is of the order of magnitude of the reception frequency, this is not possible and the influence of the series capacitance will then have to be taken into account. By connecting the capacitor C_S, assumed to be lossless, in series with the diode capacitance C_{tot}, the tuning capacitance is reduced to the value

$$C^* = C_{\text{tot}} \frac{1}{1 + C_{\text{tot}}/C_S} \tag{2-82}$$

The Q of the effective tuning capacitance, taking into account the Q of the tuner diode, increases to

$$Q^* = Q\left(1 + \frac{C_{\text{tot}}}{C_S}\right) \tag{2-83}$$

The useful capacitance ratio is reduced to the value

$$\frac{C^*_{\max}}{C^*_{\min}} = \frac{C_{\max}}{C_{\min}} \frac{1 + C_{\min}/C_S}{1 + C_{\max}/C_S} \tag{2-84}$$

where $C_{\max}$ and $C_{\min}$ are the maximum and minimum capacitances of the tuner diode.

On the other hand, the advantage is gained that, due to capacitive potential division, the amplitude of the alternating voltage applied to the tuning diode is reduced to

$$\hat{v}^* = \hat{v} \frac{1}{1 + C_{\text{tot}}/C} \tag{2-85}$$

so that the lower value of the tuning voltage can be smaller, and this results in a higher maximum capacitance $C_{\max}$ of the tuner diode and a higher useful capacitance ratio. The influence exerted by the series capacitor, then, can actually be kept lower than Eq. (2-83) would suggest.

The parallel capacitance C_P that appears in Figures 2-59 to 2-61 is always present, since wiring capacitances are inevitable and every coil has its self-capacitance. By treating the capacitance C_P, assumed to be lossless, as a shunt capacitance, the total tuning capacitance rises value and, if C_S is assumed to be large enough to be disregarded, we obtain

$$C^* = C_{\text{tot}} \left(1 + \frac{C_P}{C_{\text{tot}}} \right) \tag{2-86}$$

The Q of the effective tuning capacitance, as derived from the Q of the tuner diode, is

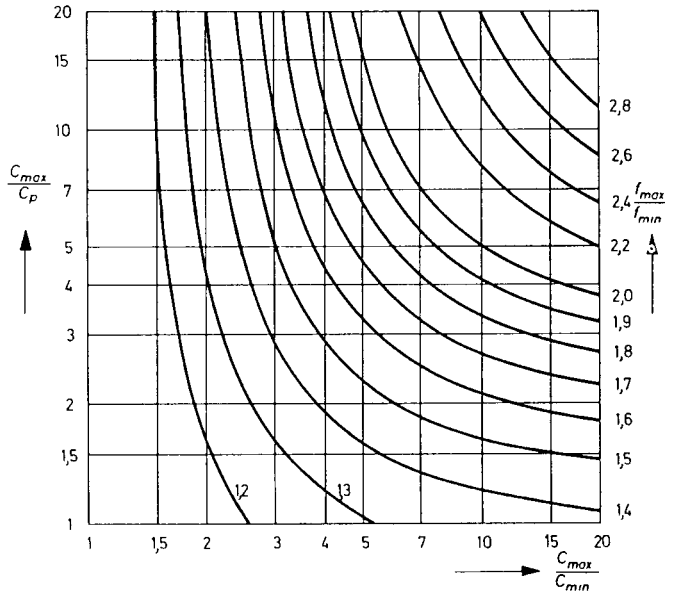


Figure 2-62 Diagram for determining the capacitance ratio and minimum capacitance.

$$Q^* = Q\left(1 + \frac{C_P}{C_{\text{tot}}}\right) \tag{2-87}$$

or, in other words, it rises with the magnitude of the parallel capacitance. The useful capacitance ratio is reduced:

$$\frac{C^*_{\text{max}}}{C^*_{\text{min}}} = \frac{C_{\text{max}}}{C_{\text{min}}} \frac{1 + C_P/C_{\text{max}}}{1 + C_P/C_{\text{max}}} \tag{2-88}$$

In view of the fact that even a comparatively small shunt capacitance reduces the capacitance ratio considerably, it is necessary to ensure low wiring and coil capacitances in the circuit-design stage.

TUNING RANGE. The frequency range over which a parallel-resonant circuit according to Figure 2-59 can be tuned by means of the tuner diode depends on the useful capacitance ratio of the diode and on the parallel and series capacitances present in the circuit.

The ratio can be found from

$$\frac{f_{\text{max}}}{f_{\text{min}}} = \sqrt{1 + \frac{\dfrac{C_{\text{max}}}{C_P\,(1 + C_{\text{max}}/C_S)}}{1 + \dfrac{C_{\text{max}}}{C_P\,(C_{\text{max}}/C_{\text{min}} + C_{\text{max}}/C_S)}}} \tag{2-89}$$

In many cases, the series capacitor can be chosen large enough for its effect to be negligible. In that case, Eq. (2-89) is simplified as follows:

$$\frac{f_{\text{max}}}{f_{\text{min}}} = \sqrt{\frac{1 + C_{\text{max}}/C_P}{1 + C_{\text{min}}/C_P}} \tag{2-89a}$$

From this equation, the diagram shown in Figure 2-62 is computed. With the aid of this diagram, the tuner diode parameters required for tuning a resonant circuit over a stipulated frequency range—that is, the maximum capacitance and the capacitance ratio—can be determined.

Whenever the series capacitance C_S cannot be disregarded, the effective capacitance ratio is reduced according to Eq. (2-84).

TRACKING. Some applications require the maintenance of a fixed frequency relationship between two or more tuned circuits as their tuning is simultaneously adjusted. Referred to as *tracking*, this technique requires narrow tolerances of capacitance versus tuning voltage. Minimizing tracking error requires special care if the tracking circuits must cover the same frequency span beginning at different start and end frequencies, as is necessary when simultaneously tuning oscillator and mixer/RF circuitry in a superheterodyne receiver. Then, tracking error must be minimized by means of series and shunt capacitances in accordance with methods known from variable capacitors. The frequency deviations that must be anticipated are summarized in the equation

$$\frac{df}{f} = -\frac{1}{2}\frac{dC_0}{C_0} - \frac{1}{2}\frac{d(L-L_0)}{L} - \frac{1}{2}\frac{dL_0}{L_0} + \frac{n}{2}\frac{dV_R}{V_R + V_D} \tag{2-90}$$

The spread of parameters dC_0/C_0 and dL_0/L_0 can only be compensated by varying the circuit inductances $d(L–L_0)/L$ or by varying the bias $dV_R/(V_R + V_D)$.

Modulating the Diode Capacitance by the Applied ac Voltage. In normal operation, the sum of the tuning voltage and the alternating signal voltage of the resonant circuits is applied to the tuner diode. The bias, and thus the capacitance, of the tuner diode therefore varies at the rhythm of the alternating voltage. Due the nonlinear character of the capacitance versus voltage curve, voltage distortions and capacitance shifts are inevitable, and these must be kept within adequate limits. This is done by maintaining the ac applied to the diode(s) at sufficiently low ac amplitude and by choosing an adequate minimum value for the tuning voltage. In the resonant circuit, a tuner diode is modulated predominantly by a current free from harmonics, according to the equation

$$i = \hat{i}\cos \omega t \tag{2-91}$$

The alternating voltage across the diode is

$$v = (V_R + V_D)\left[\left(1 + \frac{\hat{i}(1-n)}{\omega C_{\text{tot}} V_R}\sin \omega t\right)^{1/(1-n)} - 1\right] \tag{2-92}$$

An evaluation of this equation shows that especially the first harmonic makes its appearance. The capacitance shift caused by the alternating voltage superimposed on the tuning voltage is shown in Figure 2-63. However, the voltage distortion, and thus the capacitance shift, can be largely avoided if two tuner diodes are used, as in Figure 2-61.

COMPENSATING TEMPERATURE DEPENDENCE. As has already been mentioned, variations are caused in the capacitance of the diode mainly by the dependence of the diffusion voltage on

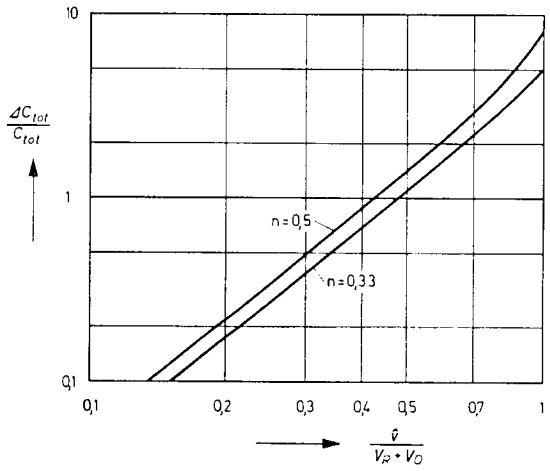


Figure 2-63 Capacitance increase as a function of the ac voltage drop across the tuner diode.

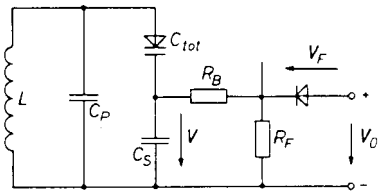


Figure 2-64 Temperature-compensation circuit with diode.

temperature. In a diode-tuned resonant circuit, the temperature coefficient of the resonant frequency, therefore, depends on the tuning voltage and thus on the resonant frequency. It is therefore impossible to compensate the temperature dependence of the resonant frequency by means of temperature-dependent capacitors—the method usually adopted for mechanically tuned circuits.

To achieve satisfactory compensation, the tuning voltage should be increased by an amount equal to that by which the diffusion voltage of the diode is reduced with rising temperature; that is, by approximately 2 mV/°C. This can be done easily by connecting a forward-biased silicon diode in series with the source of the tuning voltage, as shown in Figure 2-64. Since the forward voltage of the diode varies by –2 mV/°C with ambient temperature, the voltage

$$V = V_0 + V_D - V_F$$

which determines the bias of the tuner diode, and therefore also the diode capacitance, is almost temperature independent. In practice, it is advisable to feed this diode with additional current via the resistor R_F to keep its differential forward resistance sufficiently low and thus to prevent capacitive noise voltages from leaking into the tuning circuit.

To prevent the load placed by the resistor R_F on the source of the tuning voltage from fluctuating in a manner depending on the tuning voltage, a circuit such as shown in Figure 2-65 may be used in which the emitter diode of the transistor effects the compensation, and the only load placed on the source of the tuning voltage is the base current of this transistor.

Dynamic Stability. The resonant frequency of a diode-tuned RF circuit follows the equation

$$f = f_0 \sqrt{\left(\frac{V_R + V_D}{V_{R0} + V_D}\right)^n} \tag{2-93}$$

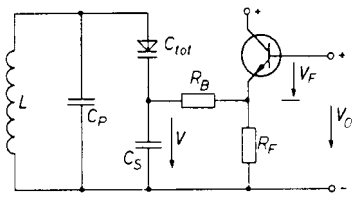


Figure 2-65 Temperature-compensation circuit with transistor.

The frequency f_0 is set with the aid of the circuit inductance when the tuning voltage equals V_{R0}. Since the signal amplitude in the circuit depends on the field strength of the received transmitter and since the resonant frequency of the circuit depends on the amplitude, undesirable feedback between the tuning voltage and the received alternating voltage may occur at the tuner diode. A disturbance that has a similar effect is the microphonic effect, which can sometimes be observed with mechanically variable capacitors. In an investigation of the feedback in diode-tuned resonant circuits, the following assumptions were made:

- The impedance of the antenna (or of the signal generator) is negligible.
- The full alternating voltage is applied to the diode (whereas, in practice, the series capacitor brings about a potential division).
- The capacitance versus voltage characteristic is an exponential function.
- The Q of this circuit is independent of the amplitude of the alternating voltage.

As a result of a high alternating voltage at the resonant circuit, three undesirable effects—harmonic generation, frequency shift, and cross-modulation—are encountered.

HARMONIC GENERATION. The nonlinearity of the capacitance versus voltage characteristic produces harmonics. However, the selectivity of the receiver reduces the effects of these harmonics to noise level.

FREQUENCY SHIFT. When a sinusoidal voltage v is applied to the tuner diode of the RF circuit, the variation of the diode capacitance does not follow the sine law, but the equation

$$\frac{C_\omega}{C} = \left(1 + \frac{v}{V_R + V_D} \sin \omega t\right)^{-n} \tag{2-94}$$

This results in a change of the resonant frequency of the circuit. The inductive slope of the resonant curve steepens and may even turn back. This could lead to bistable behavior of the resonant circuit. The change of frequency takes place as if the tuning voltage, and thus the resonant frequency, has decreased. The order of magnitude of this change of resonant frequency was determined with the aid of a computer from Fourier analysis of

$$\frac{\Delta f}{f} = 1 + \frac{1}{2\pi} \int_0^{2\pi} \left(1 + \frac{v}{V_R + V_D} \sin \omega t\right)^{n/2} d(\omega t) \tag{2-95}$$

Figure 2-66 shows the admissible alternating voltage at the diode, computed for a 2% frequency shift, as a function of the tuning voltage. The measured values differ by not more than 10% from the computed values. Since the Q of the circuits is about 50 for the VHF as well as the UHF range, a 2% frequency shift is permissible.

CROSS-MODULATION. Cross-modulation is a disturbance caused by nonlinearities of the characteristics of the components concerned. The modulation of an undesired amplitude-modulated signal is here transferred to the carrier of the amplitude-modulated intelligence signal. Cross-modulation is virtually independent of the amplitude of the intelligence signal.

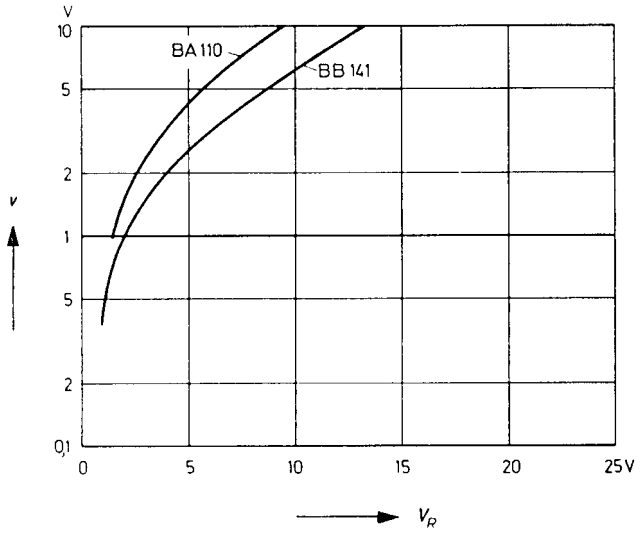


Figure 2-66 Admissible alternating voltage for a 2% frequency shift as a function of tuning voltage.

The cross-modulation factor for tuner diodes is

$$k = \frac{v^2}{(V_R + V_D)^2} m_v \frac{n(n+1)}{2} \tag{2-96}$$

Resolving this equation in terms of v, we obtain

$$v = \frac{V_R + V_D}{n} \sqrt{\frac{2kn}{m_v (n+1)}} \tag{2-97}$$

The usual definition of the cross-modulation percentage for measurement purposes is as follows: A 100% modulated noise signal ($m_v = 1$) of amplitude v is received simultaneously with the intelligence signal in the diode-tuned RF circuit. If a 1% modulation of the signal carrier by the noise signal takes place, this amounts to a cross-modulation factor $k = 0.01$. We thus obtain for tuner diodes

$$V = 0.14 \frac{V_R + V_D}{n} \sqrt{\frac{n}{n+1}} \tag{2-98}$$

Figure 2-67 gives the graphs for this equation. In practical operation, FM and VHF/UHF tuners equipped with the tuner diodes BB121 or BB141 exhibit merely the cross-modulation caused by the BJTs or FETs employed. For bipolar transistors, $v \approx 15$–100 mV; for FETs, 100–300 mV.

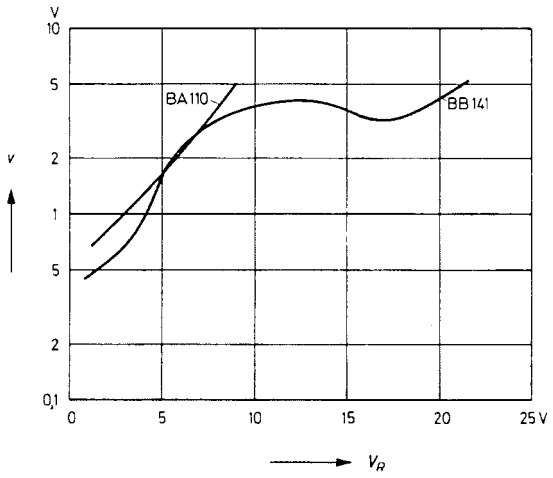


Figure 2-67 Graphical representation of Eq. (2-98).

Static Stability. To ensure trouble-free operation, the static parameters of the tuner diodes, especially the leakage current, must be taken into account. When differentiating the expression for the tuning voltage

$$V_R = V_B - I_R R_B \tag{2-99}$$

we obtain for a known leakage current I_R the tuning voltage variation

$$dV_R = dV_B - R_B I_R \left(\frac{dR_B}{R_B} + \frac{dI_R}{I_R} + dV_B \right) \tag{2-100}$$

This leads to a change in the resonant frequency

$$\frac{df}{f} = \frac{n}{2(V_R + V_D)} \left[dV_D - R_B I_R \left(\frac{dR_B}{R_B} + \frac{dI_R}{I_R} \right) + dV_B \right] \tag{2-101}$$

The resonant frequency is more stable the smaller the exponent n and the higher the tuning voltage V_R. The influence of the temperature dependence of the diffusion voltage V_D and its compensation have already been discussed. If the series bias resistance R_B is not too high, the effect of its variation ($dR_B/R_{B)}$ and the effect of the temperature dependence of the leakage current—doubling for every 10 °C temperature rise—may be neglected.

Generating the Tuning Voltage. The supply voltage V_B in Eq. (2-101) exerts an influence in many respects. The temperature coefficient of the voltage source—usually, for economic reasons, only a Zener diode is provided for stabilization—affects conditions not only with ambient temperature variations, but also when the Zener diode heats up after the set is switched on, or when its operating current varies. Because external temperature compensation of the operating voltage of the Zener diode cannot cover all detrimental effects and, moreover, increases the series bias resistance R_B, as well as any variations the latter may

undergo, the use of a temperature-compensated Zener diode, such as the 1N4065A (nominal Zener voltage, 33 V; temperature coefficient, 0.002 °C/W), is recommended.

Diode Switches. The diode switches described here differ somewhat from the switching diodes used in computer and pulse technology. Whereas in the case of switching diodes the signal itself triggers the switching operation—current does or does not pass through the switching diode depending on the signal level—diode switches allow an alternating current to be switched on or off by means of a direct voltage or a direct current. Figure 2-68 shows an example of the use of diode switches.

Diode Switch Technology. Special manufacturing techniques are needed to attain the low junction capacitance, low differential forward resistance r_F, and the low-inductance structure required for the special applications envisaged. Two conflicting requirements have to be met: On the one hand, a high-resistivity basic material should be used and the diode area should be kept as small as possible to minimize the junction capacitance for a given reverse voltage; on the other hand, the bulk resistance should be kept low. The size of the junction area, therefore, can represent only a compromise between these two requirements. To obtain a low bulk resistance R_B despite the use of high-resistivity silicon around the junction, the diodes are manufactured by the epitaxial process, characterized by a thin high-resistivity layer on a low-resistivity substrate. If the forward current I_F is sufficiently high, then the specific resistance of the epitaxial layer does not affect the bulk resistance, because under these conditions the resistivity of this layer is considerably reduced by being flooded with carriers.

The junction impedance Z_j is somewhat more critical (Figure 2-69). It is composed of the junction capacitance C_S, which does not depend on current, the current-dependent diffusion capacitance

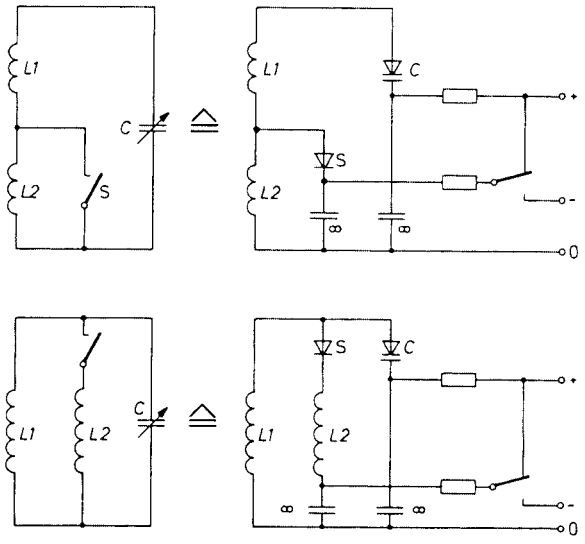


Figure 2-68 Comparison of mechanically and electronically tuned and switched resonant circuits.

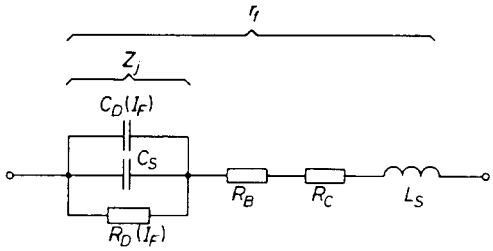


Figure 2-69 Equivalent circuit of diode switches in the frequency range from 1 MHz to 1 GHz.

$$C_D = \frac{\mathfrak{t}_p I_F}{V_T} \tag{2-102}$$

and the diffusion resistance

$$R_D = \frac{V_T}{I_F} \tag{2-103}$$

where $\mathfrak{t}_p$ is the lifetime of the minority charge carriers and

$$V_T = \frac{kT}{q} = 26 \text{ mV} \tag{2-104}$$

is the voltage equivalent of thermal energy. The resistive component of the impedance Z_j is calculated as follows:

$$\text{Re}(Z_j) = \frac{R_D}{1 + \omega^2 R_D^2 (C_S + C_D)^2} \tag{2-105}$$

As can be seen from the following approximations, this resistance becomes rather dependent on current at higher frequencies:

$$\text{For } I_F \to 0 \quad \text{Re}(Z_j) \approx \frac{I_F}{\omega^2 C_S^2 V_T} \tag{2-106}$$

$$\text{For } I_F \to \infty \quad \text{Re}(Z_j) \approx \frac{I_F}{\omega t_p I_F} \tag{2-107}$$

In Figure 2-70 the differential forward resistance is plotted as a function of the forward current. The only parameter by which the differential forward resistance can be reduced for a given current and a given frequency according to Eq. (2-107) is the carrier lifetime. Steps will have to be taken, therefore, to make the carrier lifetime as long as possible.

Because diode switches are designed for applications up to the gigahertz range, a geometry resulting in a low inherent capacitance and inductance had to be used. Furthermore, the device had to be assembled in such a way that virtually no additional contact resistance R_C was introduced. For this reason, a double-plug construction (hard-glass pressure contacts)

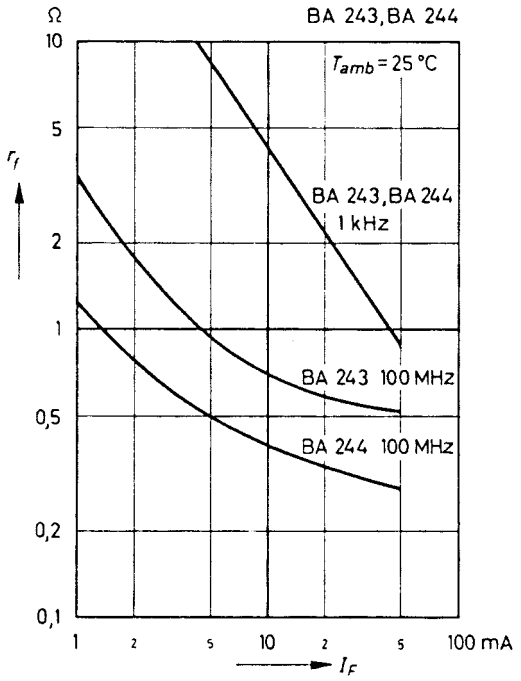


Figure 2-70 Differential forward resistance as a function of forward current.

was chosen to avoid the S- or U-bends customary with other encapsulations. The system is pressure-clamped between two stud-shaped enlargements of the connecting leads, the necessary high pressure being produced as the glass shrinks during cooling. This construction results in an encapsulation of short length (only 4 mm) and low series inductance, the advantages of which can be fully exploited because soldered connections to the leads can be made directly at the glass body.

Diode Switch Data. The silicon planar diode switches BA243 and BA244 are designed for electronic band selection in tuners in the frequency range from 10 to 1000 MHz. The differential forward resistance remains constant and is very low over a wide frequency and current range. The diode capacitance is likewise small and independent of voltage through a wide range. The devices are characterized by low-inductance geometry. The BA243 is intended for the VHF range, and the BA244 for the UHF range. See Table 2-11.

Resonant Circuits Incorporating Diode Switches. Figure 2-71 illustrates the function of the diode switch with reference to a parallel-resonant circuit whose resonant frequency can be switched by short-circuiting part of the circuit inductance. Such a switching procedure may be necessary for a VHF tuner, for example, to switch from band I to band III, because the frequencies of these bands are too far apart.

Capacitor C_K represents the capacitance of the resonant circuit. The symbol ∞ at capacitor C is meant to indicate that the value of this capacitor should be chosen to be much higher than the circuit capacitance. C completes the ac circuit when the inductor L_2 is shorted and

Table 2-11 BA243/244 specifications

Maximum Ratings			
Reverse voltage	V_R	20	V
Forward current at $T_A = 25$ °C	I_F	100	mA
Junction temperature	T_j	150	°C
Storage temperature range	T_s	–55 to +150	°C
Characteristics at $T_A = 25$ °C			
Forward voltage at $I_F = 100$ mA	V_F	<1	V
Leakage current at $V_R = 15$ V	I_R	< 100	nA
Dynamic forward resistance at $f = 50$–1000 MHz, $I_F = 10$ mA			
BA243	r_F	0.7 (< 1)	Ω
BA244	r_F	0.4 (< 0.5)	Ω
Relative variation of dynamic forward resistance with the variation of forward current in the range of $I_F + 2$–40 mA	$\frac{\Delta r_F 100}{r_F \Delta I_F}$	5	%/mA
Capacitance at $V_R = 15$ V, $f = 1$ MHz	C_{tot}	1.3 (< 2)	pF
Relative variation of capacitance with the variation of reverse voltage in the range of $V_R = 7$–20 V, $f = 100$ MHz	$\frac{\Delta C_{tot} 100}{C_{tot} \Delta V_R}$	1	%/V
Series inductance across case	L_S	2.5	nH

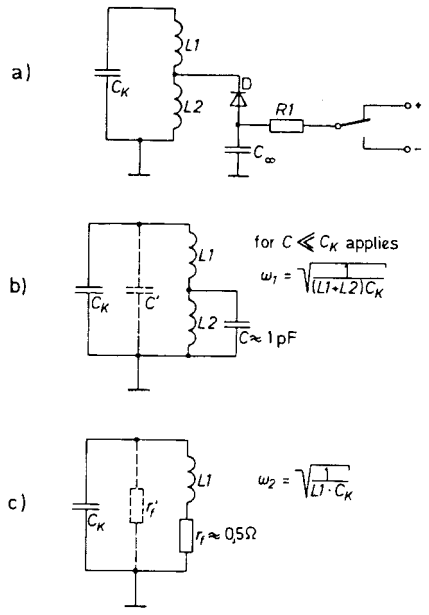


Figure 2-71 (a) Diode switch as bandswitch, shorting a partial inductance; (b) equivalent circuit with blocked diode switch; (c) equivalent circuit with conducting diode switch.

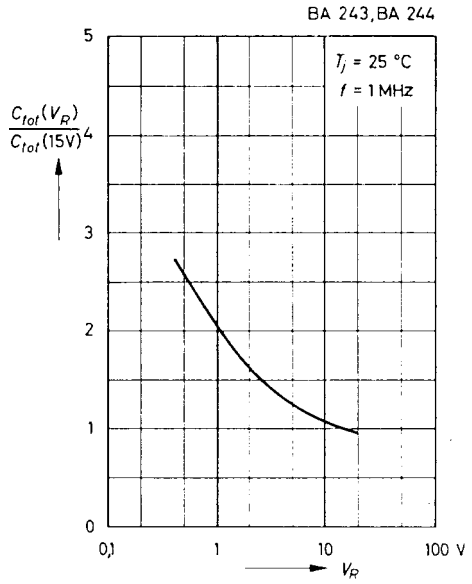


Figure 2-72 Capacitance as a function of the reverse voltage.

enables the switching or reverse voltage to be applied at a point of virtually zero RF potential. The resistor R_1 limits the diode current in the forward direction.

At a reverse voltage of $V_R = 15$ V, for example, the diode switch represents a 1-pF capacitor, as shown in Figure 2-71b. This represents parasitic capacitance C' in the circuit, which reduces the effective capacitance variation of the resonant circuit in the case of capacitive tuning. The capacitance C' is small enough to be negligible in most cases relative to the wiring and winding capacitances. Figure 2-72 shows the relative capacitance of the blocked-diode switches, as a function of the reverse voltage.

At a forward current of, for example, 10 mA, the diode has a differential resistance of approximately 0.5 Ω (see Figure 2-70). The effect of this resistance r_F on the resonant circuit is illustrated in Figure 2-71c. Conversion of a series loss resistance r_F into a parallel loss resistance r'_F is based on the following considerations: The equivalent parallel resistance r'_F affects the tuned circuit in the same way as the series resistance r_F; that is, the Q of the circuit should be the same in both cases. If one assumes that $\omega L >> r_F$, the equivalent of parallel resistance can be expressed as follows:

$$Q = \frac{\omega L_1}{r_F} = \frac{r'_F}{\omega L_1} \tag{2-108}$$

$$r'_f = \frac{(\omega L_1)^2}{r_F} \tag{2-109}$$

An evaluation of this equation shows that a differential resistance $r_F \approx 0.5$ Ω can be tolerated in most applications as regards amplification and selectivity.

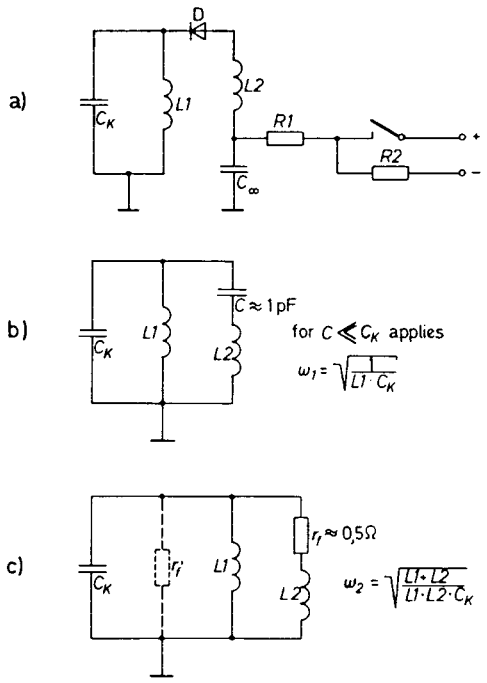


Figure 2-73 (a) Diode switch as a bandswitch, with an inductance in parallel; (b) equivalent circuit with blocked-diode switch; (c) equivalent circuit with conducting diode switch.

Instead of altering the resonant frequency of a tuned circuit by short-circuiting part of its inductance, the same result can be obtained by connecting a second inductance in parallel. The use of the diode switch for this purpose is illustrated in Figure 2-73. This differs from the one shown in Figure 2-71 in that the circuit inductance consists of either L_1 or of L_1 and L_2 in parallel. By applying a permanent negative bias to the diode via a high-value resistor R_2, it is possible to dispense with the use of one switch contact. Figure 2-73b shows the equivalent circuit for the diode under reverse-bias conditions (corresponding to a low resonant frequency), and Figure 2-73c shows the equivalent circuit under forward-bias conditions (corresponding to a higher resonant frequency).

In radio and television receivers, it usually is possible to derive the necessary reverse bias for the diode by rectification of the local oscillator signal, since the power required for blocking the diodes is extremely small. Note that under reverse-bias conditions the inductance L_2 and the capacitance C form a series-tuned acceptor circuit. However, this capacitance being extremely small, any undesirable effects due to its presence can normally be avoided. The same diode switching arrangements described for single-tuned circuits can also be employed in multisection bandpass and broadband filters.

In the following paragraph, a practical example for the use of diode switches will be given, and it will be seen that it is possible, in principle, to replace all of the mechanical RF switch contacts normally employed in receivers by diode switches. In cases where application of the diode dc potentials to a "cold" point in the circuit is not feasible, it may be necessary to apply them via *LC* or *RC* isolating networks.

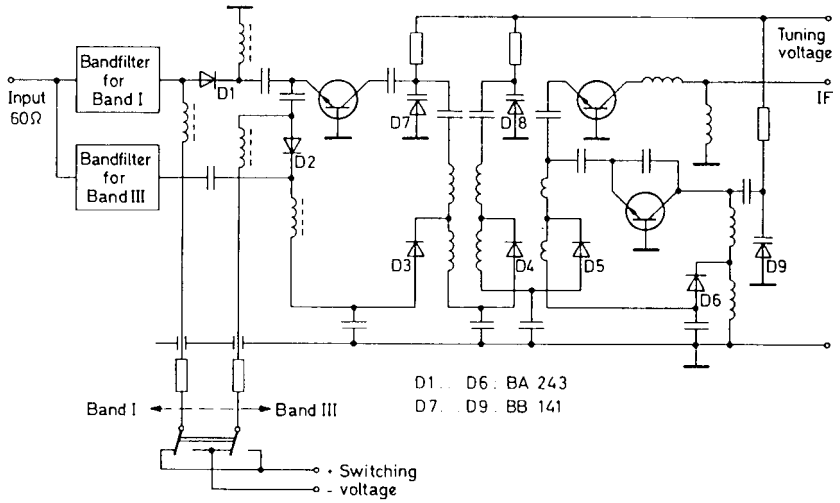


Figure 2-74 Schematic of a VHF television tuner with electronic band selection.

Use of the Diode Switch in a Television Receiver. Figure 2-74 shows the schematic of an electronically tuned and switched band I (48–62 MHz)/band III (175–224 MHz) TV tuner that requires switching of the input filter, bandpass filter, and oscillator circuit. These switching functions are performed by BA243 diode switches. As is customary in most European countries, in this example a broadband input filter between the antenna and the first RF transistor is used rather than a selective front-end circuit. In view of the need to correct noise and power matching over a relatively large frequency range, a switchable filter is employed in most tuners, switching being effected in this case by means of diodes D_1 and D_2.

Diodes D_3–D_6 are used for band-filter switching, for switching the coupling to the mixer transistor, and for switching the oscillator frequency. To minimize the bias-current requirements (each diode requires a bias current of approximately 10 mA), all diodes that are to be activated at the same time are connected in series. This arrangement requires only a few additional RF chokes, which are easily provided because of the high frequencies involved. The direct current required for the diode switches could be obtained by simple rectification, for example, from the heater circuit of the television receiver CRT, and the diode reverse bias could be obtained by rectification of the local oscillator signal, in which case the minute additional power requirement would have to be allowed for in the oscillator design.

A technically sophisticated design for a television tuner takes the form of a combined VHF/UHF unit incorporating electronically switched circuits in which several elements would be common. Such an all-band tuner can be assembled from only three transistors and a combined VHF/UHF band filter in $\lambda/2$ or $\lambda/4$ technology for UHF reception. The VHF circuits are short-circuited when the UHF position is selected.

Multistandard TV receivers require a large number of switch contacts in the tuner, the IF amplifier, the video amplifier, and the time base. The use of diode switches is particularly advantageous in this situation because it renders the physical layout of the equipment virtually independent of the location of the controls and obviates the need for long RF cable runs, which may be prone to interference.

2-2 BIPOLAR TRANSISTORS

2-2-1 Transistor Structure Types

All current RF silicon transistors are of the bipolar *npn* planar epitaxial type. Briefly, the significance of each of these terms is as follows.

Bipolar. In its broadest sense, it is the basic structure shown schematically in Figure 2-75, that is, the familiar three-semiconductor-region structure. *Bipolar* specifically means that the charge carriers of both negative (electron) and positive (hole) polarities are involved in the transistor action. In way of contrast, unipolar types include the junction-gate and insulated-gate field-effect transistors, which are basically one- or two-semiconductor-region structures in which carriers of a single polarity dominate.

npn. This abbreviation is for *negative–positive–negative* and identifies the regions of the structure as to polarity of the dominant or majority carrier in each region. The other polarity type is *pnp*. (See Figure 2-75.)

Silicon. Silicon (Si) is one of two elements from the fourth column of the periodic table that are in widespread use for transistor fabrication [the other is germanium (Ge)]. Other materials used include the compound gallium arsenide (GaAs). Although silicon is in predominant use because it results in the most favorable compromise among high-frequency, high-temperature, high-reliability, and ease-of-use attributes of the usable semiconductor material, silicon–germanium (SiGe) technology is increasingly important because it affords high-frequency performance unachievable with silicon alone.

Planar. The term planar denotes that both emitter–base and base–collector junctions of the transistor intersect the device surface in a common plane (hence, a better term might be *coplanar*). However, the real significance of the so-called planar structure is that the technique of diffusing dopants through an oxide mask, used in fabricating such a structure, results in junctions being formed beneath a protective oxide layer. These protected junctions are less prone to the surface problems sometimes associated with other types of structures, such as the mesa.

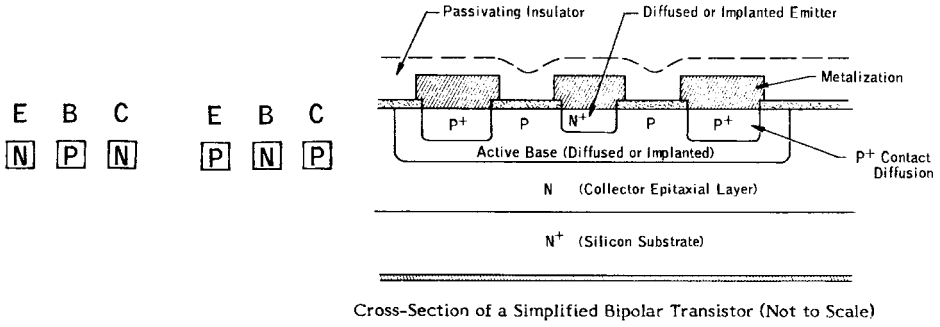


Figure 2-75 Transistor structure schematic.

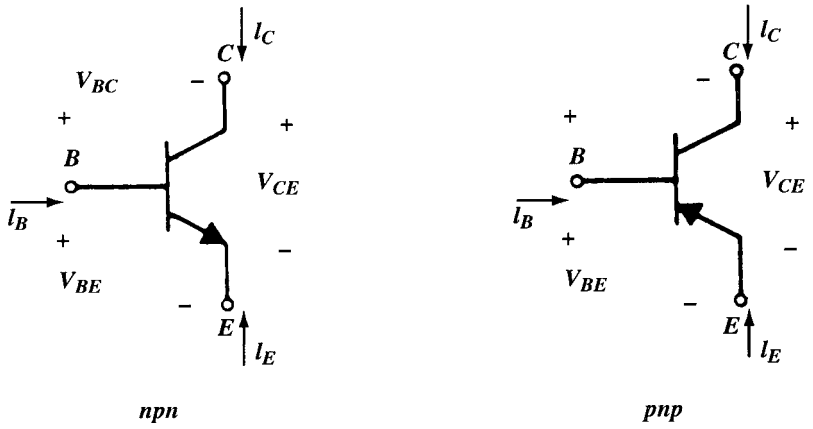


Figure 2-76 Bipolar transistor sign convention.

Epitaxial. This term, as it is commonly used, is actually a shortening of the term *epitaxial-collector*. That is, the collector region of the transistor is formed by the epitaxial technique, rather than by diffusion, which is commonly used to form the base and emitter regions. The epitaxial layer is formed by condensing a single-crystal film of semiconductor material on a wafer or substrate that is usually of the same material. Thus, an epitaxial (collector) transistor is one in which the collector region is formed on a low-resistivity substrate. Subsequently, the base and emitter regions are diffused into the "epi" layer. The epitaxial technique lends itself to precise tailoring of collector-region thickness and resistivity with consequent improved device performance and uniformity.

2-2-2 Large-Signal Behavior of Bipolar Transistors*

In this section, the large-signal or dc behavior of bipolar transistors is considered. Large-signal models are developed for the calculation of total currents and voltages in transistor circuits, and such effects as breakdown-voltage limitations, which are usually not included in models, are also considered. Second-order effects, such as current-gain variation with collector current and Early voltage, can be important in many circuits and are treated in detail.

The sign conventions used for bipolar transistor currents and voltages are shown in Figure 2-76. All bias currents for both *npn* and *pnp* transistors are assumed positive going into the device.

Electrical and Performance Characteristics and Specifications. Electrical characteristics may be described as uniquely defined, measurable electrical transistor properties that are not a function of the measuring circuit or apparatus (except insofar as standard terminations and measurement accuracy are concerned). Performance or operating characteristics are also electrical properties but they are, in general, not unique because their values depend on the measuring circuit (in particular, source and load impedance, which may be arbitrary). As might be expected, the terms are often used somewhat loosely (and sometimes

* Portions of this section, Sections 2-2-3 through 2-2-5, and Sections 2-3-3 through 2-3-6 are based on Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed., ©1993 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

interchangeably), especially in cases where only subtle differences are involved. The terms are generally used on transistor datasheets to segregate (for emphasis) under performance or operating characteristics those properties most directly applicable to the primary intended application.

Performance (Operating) Characteristics. Of the numerous performance characteristics that can be specified for high-frequency transistors, perhaps the most fundamental and pertinent characteristics are:

- Power gain and noise figure, for small-signal applications
- Power gain, power output, and efficiency, for large-signal applications

All of these characteristics are, of course, functions of frequency, bias temperature, and so on, and to completely characterize a transistor over its full frequency, bias, and temperature ranges would be prohibitively costly. Consequently, characterization data are given only for restricted ranges of these variables. These data should portray sufficiently the capabilities of a particular device for its intended applications. In the case of maximum ratings, some applications may require additional characterization by the user or through applications assistance from the manufacturer.

POWER GAIN

G_{max}. Of the various definitions for the measurement of power flow in an active two-port device, such as a transistor, two are unique enough to allow specification without recourse to specifying the complete measuring circuit in detail. One of these definitions is termed *maximum available gain* (G_{max}) and is the power gain obtained when the input and output ports are simultaneously conjugately matched to source and load impedances, respectively. Implicit in this definition is the assumption that the two-port is unconditionally stable; that is, that no combination of input/output tuning can result in increasing gain to the point of oscillation.

$|S_{21}|^2$. The other unique power gain is the gain realized when the transistor is inserted between a source and load with identical impedances (in practice, usually $50 + j0\ \Omega$). This particular insertion of transducer gain happens to coincide with the usual definition of the two-port forward scattering parameter, S_{21}. More precisely, it is equal to the magnitude squared of this parameter and is therefore often identified by the symbol $|S_{21}|^2$. For wideband applications, $|S_{21}|^2$ is important because wideband terminations "not too different from 50 Ω" are more easily realized than are wideband transforming networks that provide the matching required for G_{max}.

NOISE FIGURE. A common measure of the noise generated by an active two-port device—noise that sets a lower limit on amplifier sensitivity—is noise factor (F). This is defined as

$$F = \frac{\text{Input signal-to-noise ratio}}{\text{Output signal-to-noise ratio}} \tag{2-110}$$

or more generally as

$$F = \frac{\text{Total output noise power}}{\text{Output power due to source resistance}} \tag{2-111}$$

At high frequencies, spot noise factor or noise factor for a small fractional bandwidth (say, 1%) is used and is usually expressed as *noise figure* (*NF*), in decibels:

$$NF = 10 \log F \tag{2-112}$$

As already discussed, noise figure is a function of source impedance (as well as functions of frequency, bias, etc.) and hence, there is an infinity of noise figures associated with a given device corresponding to the infinity of possible impedances that may be presented to the device input. The only unique one, in the sense that it does not involve arbitrary source impedances, is $NF_{\min}$, the minimum noise figure obtained (at a given bias and frequency) when the input is tuned to optimize this parameter. This is the noise figure value commonly given on datasheets.

In practical amplifiers that involve more than one stage, *overall noise factor* (F_O) is given by

$$F_O = F_1 + \frac{F_2 - 1}{G_1} + \cdots + \frac{F_n - 1}{G_{(n-1)}} \tag{2-113}$$

where n = number of stages
G_n = gain of the nth stage
F_n = noise factor of the nth stage

This expression emphasizes the important fact that for low-noise amplifiers, the first stage must be designed for the lowest noise figure and highest gain possible. [Note that the noise contribution of the second stage is divided (reduced) by the gain of the first stage.] Since the optimum source impedance and bias currents for optimum gain and noise figure do not often coincide, very careful circuit design is required to minimize overall noise figure. For more accuracy, we recommend the noise correlation approach.

POWER OUTPUT. This characteristic is important for both amplifier and oscillator transistors. In both cases, it is extremely circuit sensitive. For amplifiers, maximum useful output is often limited to that power output level at which gain has compressed 1 dB, an indicator of the upper limit of the device's linearity range. For oscillators, it is merely a quantitative measure of RF power for given dc input power.

EFFICIENCY. In the most general sense, this characteristic expresses as a percentage the ratio of RF power output to the total circuit input power, both dc and RF. That expression is efficiency, η, defined as

$$\eta = \frac{P_O}{P_i + P_{\text{dc}}} \times 100 \tag{2-114}$$

where P_O = RF output power
P_i = RF input power
P_{dc} = total dc power input

Since for oscillator transistors there is no RF power input, and for amplifier transistors the maximum input RF power is calculable from the power gain and power output specifications,

the inclusion of P_i in efficiency is redundant. Moreover, since the major portion of the dc power is dissipated by the transistor collector, a more restricted definition of efficiency is pertinent. This parameter, termed *collector efficiency* (η_C), is given by

$$\eta_C = \frac{P_O}{P_{CC}} \times 100 \tag{2-115}$$

where $P_{CC} = V_{CC} \times I_{CC}$
V_{CC} = collector supply voltage
I_{CC} = collector supply current

Electrical Characteristics. Electrical characteristics may be conveniently classified into two main types, dc and ac.

dc CHARACTERISTICS. The importance of dc characteristics of high-frequency transistors lies primarily in biasing and reliability considerations. However, certain dc characteristics are also directly related to high-frequency performance. For example, high-frequency noise figure is affected by dc current gain. The dc characteristics discussed here are those usually found on high-frequency transistor datasheets.

$V_{(BR)CBO}$, I_{CBO}. These two parameters serve to characterize the reverse-biased collector–base *pn* junction and are defined as follows (with the aid of Figure 2-77a). The collector–base breakdown voltage, $V_{(BR)CBO}$, identifies the voltage at which collector current tends to increase without limit, usually due to the high electric field developed across the junction. This voltage sets a limit on the maximum transistor operating voltage and, as mentioned before under maximum ratings, usually is the basis for the collector–base maximum voltage rating. $V_{(BR)CBO}$ should be specified at a value of $I_C = I_{C1}$ in the figure, which is within the avalanche (or high slope) region of the device's reverse characteristic. Typical values of I_{C1} are in the 1–10-μA region for high-frequency transistors.

To further define the quality of the reverse *I–V* characteristic, a specification is usually placed on the collector cutoff current, I_{CBO}, which is measured at some value of collector–base voltage less than $V_{(BR)CBO}$. For a good-quality silicon junction ("sharp" instead of soft; see Figure 2-77a), I_{CBO} is in the nanoampere range.

$V_{(BR)EBO}$, I_{EBO}. These two parameters characterize the reverse-biased emitter–base *pn* junction in an analogous manner to the collector–base junction parameters $V_{(BR)CBO}$ and I_{CBO}, given above, and are shown in Figure 2-77b. No further discussion of these parameters will be given here.

$V_{(BR)CEO}$, I_{CEO}. The collector–emitter breakdown voltage and cutoff current are somewhat more complex in nature than the collector–base and emitter–base parameters. In the latter two, only a single *pn* diode is involved. In the collector–emitter case, two diodes are involved. Moreover, each is influenced by the other through transistor action, since the reverse current of the collector–base diode flows through the emitter–base junction as forward current. Thus, the collector–base reverse current is amplified by the dc current gain of the transistor, resulting in:

- I_{CEO} being greater than I_{CBO} (for a given voltage)
- Typically, the familiar negative-resistance region in the *I–V* characteristic as shown in Figure 2-77c.

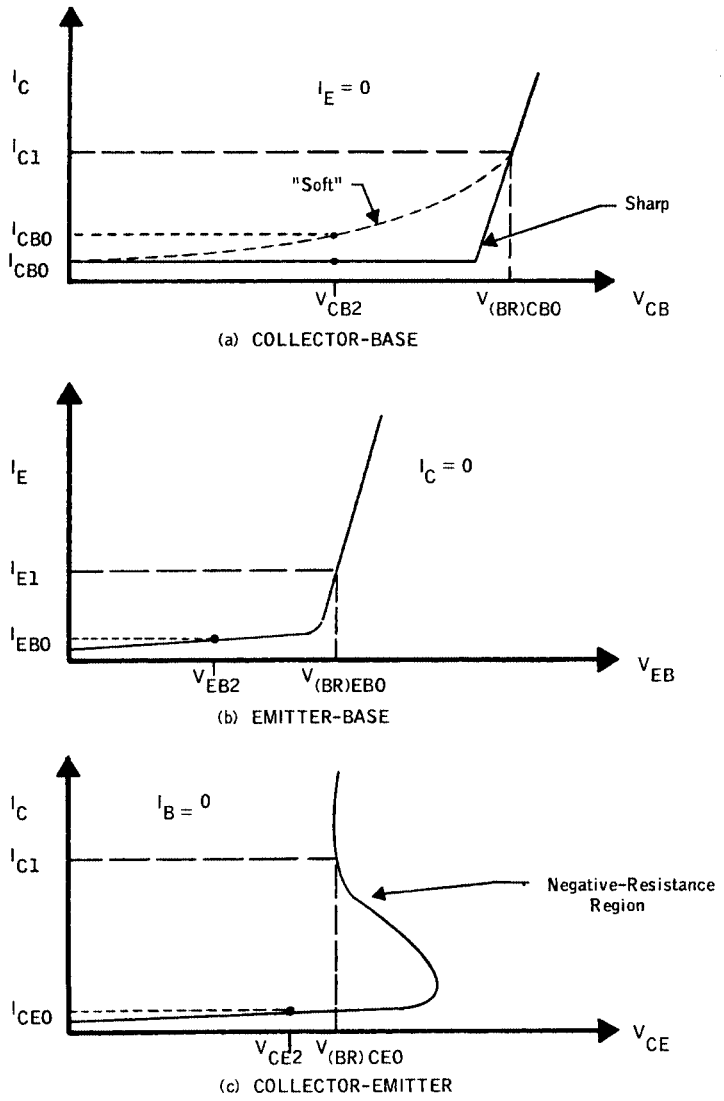


Figure 2-77 Transistor reverse *I–V* characteristics.

Consequently, $V_{(BR)CEO}$ is typically specified at collector currents one to three orders of magnitude higher than in the case of $V_{(BR)CBO}$ and $V_{(BR)EBO}$ to establish the minimum value of this characteristic.

h_{fe}. This parameter is simply the dc common-emitter current gain; that is, the ratio of collector current to base current at some specified collector voltage and current.

ac CHARACTERISTICS. Of the numerous ac characteristics defined for transistors, only relatively few are commonly used in characterizing high-frequency transistors. Some of the more pertinent parameters are briefly covered here.

S Parameters. By far the most useful and conveniently measured set of two-port parameters for transistor high-frequency (roughly 100 MHz and above) characterization is

the *S*-parameter or scattering-matrix set. These parameters completely and uniquely define the small-signal gain and input/output immittance properties of any linear "black box." (By definition, a transistor or any active device is linear under small-signal conditions.) However, these parameters reveal nothing (except possibly indirectly and approximately) about large-signal or noise behavior. Simply interpreted (more general definitions and other interpretations abound in the technical literature), the *S* parameters are merely insertion gains, forward and reverse, and reflection coefficients, input and output, with driven and nondriven ports both terminated in equal impedances, usually $50 + j0\ \Omega$. Such an interpretation tends to make *S* parameters very attractive, once some familiarity is gained, at high (especially microwave) frequencies, since the power flow or gain and reflection-coefficient concepts are more intuitively meaningful than voltage and current conceptual schemes. It should also be mentioned that *S*-parameters can be converted through straightforward matrix transformations to other two-port parameter sets, such as *h*-, *y*-, and *z*-parameters.

Proceeding with more specific definitions, the *S*-parameters are defined analytically by

$$b_1 = S_{11}a_1 + S_{12}a_2 \tag{2-116}$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \tag{2-117}$$

or, in matrix form,

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{2-118}$$

where, referring to Figure 2-78

a_1 = (incoming power at Port 1)$^{1/2}$
b_1 = (outgoing power at Port 1)$^{1/2}$
a_2 = (incoming power at Port 2)$^{1/2}$
b_2 = (outgoing power at Port 2)$^{1/2}$
E_1, E_2 = electrical stimuli at Port 1, Port 2

From the figure and defining linear equations, for $E_2 = 0$, then $a_2 = 0$, and (skipping through numerous rigorous steps)

$$S_{11} = \frac{b_1}{a_1}$$

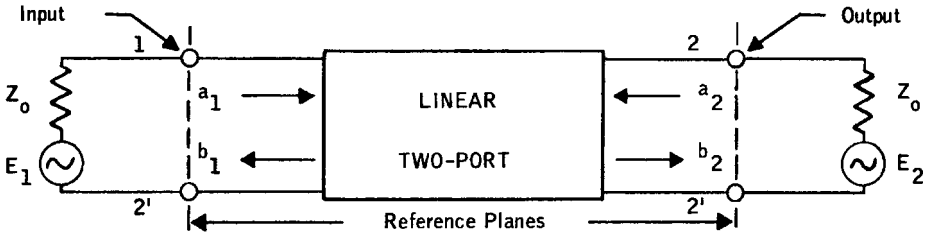


Figure 2-78 Two-port *S*-parameter definition.

$$= \left(\frac{\text{Outgoing input power}}{\text{incoming input power}}\right)^{1/2}$$

$$= \frac{\text{Reflected voltage}}{\text{Incident voltage}}$$

$$= \text{Input reflection coefficient} \tag{2-119}$$

$$S_{21} = \frac{b_2}{a_1}$$

$$= \left(\frac{\text{Outgoing output power}}{\text{incoming input power}}\right)^{1/2}$$

$$= \left(\frac{\text{Output power}}{\text{Available input power}}\right)^{1/2}$$

$$= (\text{Forward transducer gain})^{1/2} \tag{2-120}$$

or, more precisely in the case of S_{21}:

$$\text{Forward transducer gain} = G_{\text{TF}} = |S_{21}|^2 \tag{2-121}$$

$$Z_i = Z_o \tag{2-122}$$

Similarly, at Port 2 for $E_1 = 0$, then $a_1 = 0$, and

$$S_{22} = \frac{b_2}{a_2} = \text{Output reflection coefficient} \tag{2-123}$$

$$S_{12} = \frac{b_1}{a_2} = (\text{Reverse transducer gain})^{1/2} \tag{2-124}$$

$$G_{\text{TR}} = |S_{12}|^2 \tag{2-125}$$

Since many measurement systems display *S*-parameter magnitudes in decibels, the following relationships are particularly useful:

$$|S_{11}|_{\text{dB}} = 10\log|S_{11}|^2$$

$$= 20\log|S_{11}| \tag{2-126}$$

$$|S_{22}|_{\text{dB}} = 20\log|S_{22}| \tag{2-127}$$

$$|S_{21}|_{\text{dB}} = 10\log|S_{21}|^2$$

$$= 20 \log|S_{21}|$$

$$= 10 \log|G_{TF}| = |G_{TF}|_{dB} \tag{2-128}$$

$$|S_{12}|_{dB} = 10 \log|S_{12}|^2$$

$$= 20 \log|S_{12}|$$

$$= 10 \log|G_{TR}| = |G_{TR}|_{dB} \tag{2-129}$$

Figure 2-79 shows an example of the use of *S* parameters in transistor stage design.

Transition Frequency. One of the better known, but perhaps least understood, figures of merit for high-frequency transistors is the so-called transition frequency, f_T. Part of the misunderstanding that appears to exist is due to the use of the misleading (but common, for

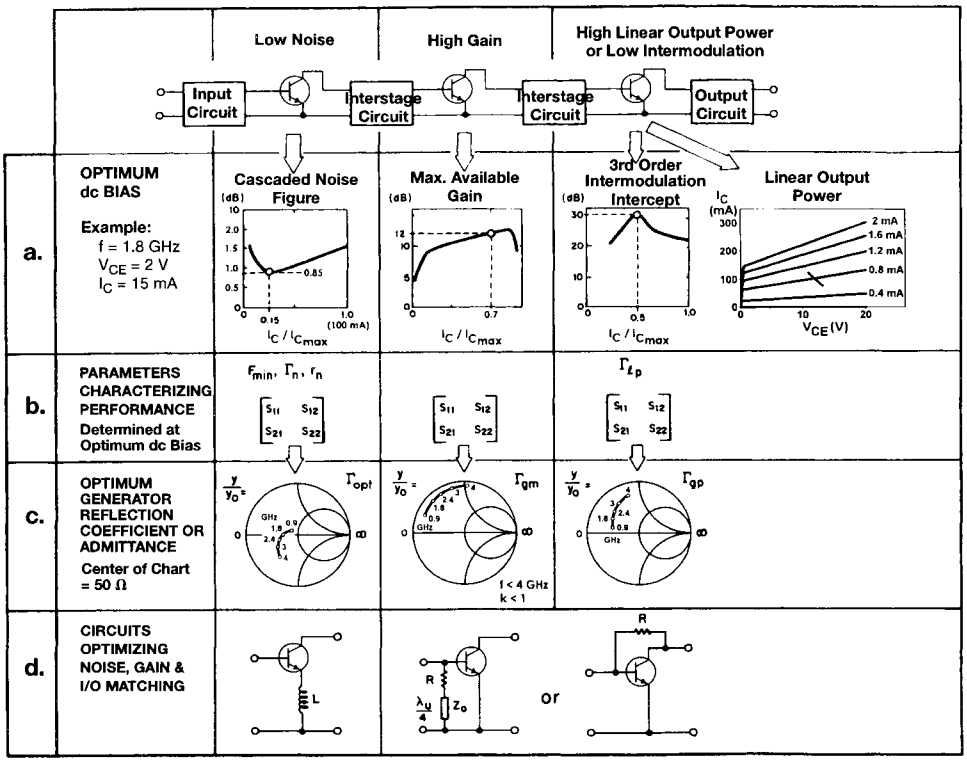


Figure 2-79 Key parameters in applying a BJT in low-noise front-end, high-gain, and linear-power stages. (a) The BJT characteristics that lead to optimum dc bias for a low-noise front end; (b) parameters characterizing device performance; (c) Γ_{opt} for the first stage, and S_{11} and S_{22} for the output stage as simulated by CAD software. The example is based on the Siemens BJTs BFP420 (low-noise stage), BFP450 (high-gain stage), and BFG235 (output stage). Circuits to optimize stage noise, gain, and I/O matching are shown in (d). (Presentation based on Figure 14 in Charles A. Liechti, "Microwave Field-Effect Transistors—1976," *IEEE Trans. Microwave Theory Tech.,* vol. MTT-24, pp. 279–300, June 1976.)

historical reasons) term *short-circuit gain–bandwidth product* for this parameter. By definition, f_T is that characteristic frequency described by the equation

$$f_T = h_{fe} \times f_{\text{meas}} \tag{2-130}$$

where h_{fe} is the magnitude of the small-signal common-emitter short-circuit current gain, and f_{meas} is the frequency of measurement, chosen such that

$$2 \le h_{fe} \le \frac{h_{fe0}}{2} \tag{2-131}$$

where h_{fe0} is the low-frequency value of h_{fe}.

To varying degrees of approximation, depending on the transistor type, f_T is the frequency at which h_{fe} approximates unity. It is not, in general, the frequency at which h_{fe} is precisely equal to unity. To clarify these points further, consider the plot of h_{fe} against frequency sketched in Figure 2-80.

At low frequencies, $f << f_B$, h_{fe} is constant and equal to h_{fe0}.

At f_B, h_{fe} has decreased to 0.707 h_{fe0}; that is, f_B is the 3-dB cutoff frequency for common-emitter short-circuit current gain, h_{fe}.

For frequencies such that

$$2f_B < f < f_T \tag{2-132}$$

h_{fe} varies inversely proportional to frequency. That is, the f_T defining relationship holds:

$$h_{fe} \times f_{\text{meas}} = \text{constant} = f_T \tag{2-133}$$

At frequencies approaching f_T, other parameters, especially package parasitics, can cause $|h_{fe}|$ to depart significantly from this $1/f$ variation. Therefore, the frequency, f_1, at which $|h_{fe}|$ actually equals unity can be somewhat different from f_T.

Applying these frequency-gain characteristics to the common-emitter wideband, low-pass amplifiers gives rise to the terminology of f_T being a "gain–bandwidth product." However,

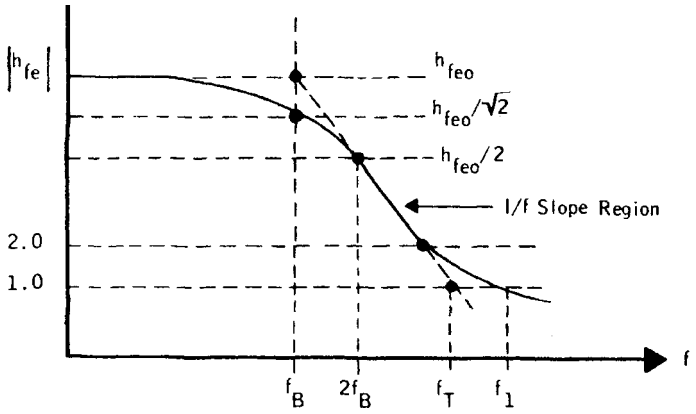


Figure 2-80 $|h_{fe}|$ frequency characteristics.

this is an optimistic approximation at best, since the product of low-frequency circuit gain and the 3-dB cutoff frequency is reduced from f_T by an amount that depends on circuit impedances.

The real significance of f_T lies in the fact that it is a measure of certain internal transistor parameters that do, in fact, affect high-frequency performance: for example, gain (though not in the convenient quantitative manner implied by the term "gain–bandwidth product"). In particular, good high-frequency noise performance requires that f_T be high. Thus, f_T is included on transistor datasheets primarily as a figure of merit, not as a parameter to be used directly in design.

Collector–Base Time Constant, $r_b'C_c$. This is an internal device parameter that relates only indirectly to high-frequency performance. It is primarily a measure of internal feedback within the transistor. It also relates to transistor high-frequency impedance. As the name says (in symbols), it is a measure of transistor base resistance and collector capacitance in combination; however, except for certain low-frequency transistors, it cannot be considered the simple two-element lumped *RC* time constant implied by the terminology. (In high-frequency transistors, base resistance and collector impedance must be considered to be distributed when considered in detail). As a figure of merit, it is included on transistor datasheets to indicate how well base resistance and collector impedance have been minimized. It also allows the estimation of certain gain properties of the transistor (see the $f_{\max}$ parameter, which we'll describe shortly).

Collector–Base Capacitance, C_{cb}. This parameter is simply the total collector–base *pn* junction capacitance measured at a low frequency (typically, 1 MHz), where it can be considered a single lumped element. For high-frequency transistors it is, of course, desirable that C_{cb} be small from bandwidth and stability considerations as well as from gain considerations alone.

Maximum Frequency of Oscillation, $f_{\max}$. This is another figure-of-merit parameter, as opposed to measurable parameters directly usable in the application of transistors. Its importance stems from the following approximate relationships (which will not be derived here):

$$f_{\max} \approx \left(\frac{f_T}{8\pi r_b' C_C} \right)^{1/2} \tag{2-134}$$

$$G_{\max} \approx \left(\frac{f_{\max}}{f_{\text{oper}}} \right)^2 \tag{2-135}$$

These expressions illustrate in a quantitative way the importance and the interrelationship between high f_T and low $r_b'C_C$ insofar as high-frequency gain is concerned. However, since they are approximations and since their derivation involves several assumptions that are not always valid, they must be interpreted with caution. For example, the expression for $G_{\max}$ is obviously not applicable at low frequencies, since as $f \to 0$, $G_{\max} \to \infty$, according to this expression. As a rule of thumb, the $G_{\max}$ expression is a reasonable approximation for frequencies such that

$$5 > \frac{f_{\max}}{f_{\text{oper}}} > 1 \tag{2-136}$$

For accurate analysis of transistor gain and stability, a complete set of two-port parameters must be employed in exact expressions, such as those from which the approximations shown above were derived.

Datasheet Transistor Specifications. The following is a reprint of a Siemens microwave transistor datasheet. This 25-GHz f_T transistor will soon be replaced by a 40-GHz family, based on a silicon–germanium design. Key parameters describing the transistor are:

- The dc parameters, such as maximum current and voltage and dissipation
- Noise figure as a function of generator impedance, bias point, and frequency
- Frequency-dependent behavior, typically expressed in *S* parameters, and also as a function of bias point and frequency
- Large-signal performance to be calculated based on a nonlinear model, such as the Gummel–Poon

The transistor description typically gives details about packaging, as the Gummel–Poon model is intrinsic (without parasitics). Not only does the package determine the parasitics, it also affects the device's ability to transfer heat to its heat sink. The following figures, reproduced with permission from the datasheet for the BFP420 transistor by Siemens (now Infineon Technologies), are a typical representation of the dc-I_C curves, *S* parameters, and noise parameters of a normal, packaged microwave transistor. Besides its dc parameters and capacitance, there are other critical parameters. A BJT's noise figure is a strong function of its dc current gain, and one can determine a 3-dB cutoff frequency referred to as f_T, which shows the gain roll-off as a function of collector current and frequency. Another important measure or figure of merit is $f_{\max}$, the maximum frequency of oscillation:

$$f_{\max} = \sqrt{\frac{|Y_{21}|}{16\pi R_{bb} C_{be} C_{bc}}} \tag{2-137}$$

Figure 2-81 shows a photograph of the die for the BFP420.

Since the intercept point is an important figure of merit, Figure 2-82 shows intercept point as a function of bias point for the BFR93AW BJT. We wish more manufacturers would supply this information.

2-2-3 Large-Signal Transistors in the Forward-Active Region

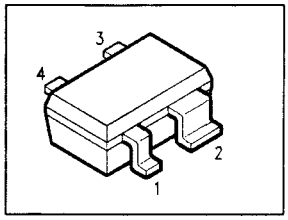
We saw in the diode case that its description is physics-based and it is, therefore, logical to apply the same modeling for the transistors. The first complete modeling for transistors was described in the Gummel–Poon model, but many engineers use the somewhat more simpler Ebers–Moll model. By combining two diodes, the parasitic elements, and the base separating resistor, we can show the "intrinsic" nonlinear model for the bipolar transistor.

Many equations that have been published contain some approximations, but for exact nonlinear analysis, we have to resort to the somewhat more complex or complete equations as shown in Appendix A for the bipolar model.

Recently, some other models have been proposed by researchers, such as the MEXTRAM models (Philips), but the trouble with these models is that while they may give quite good results for RF applications, it is virtually impossible to obtain parameters for them. So far the only industry-standard model is the Gummel–Poon, and to the authors' knowledge there

SIEMENS **SIEGET®25 BFP420**

NPN Silicon RF Transistor



- For High Gain Low Noise Amplifiers
- For Oscillators up to 10 GHz
- Noise Figure F = 1.05 dB at 1.8 GHz
  Outstanding G_{ms} = 20 dB at 1.8 GHz
- Transition Frequency f_T = 25 GHz
- Gold metallization for high reliability
- **SIEGET®25-Line**
  Siemens **G**rounded **E**mitter **T**ransistor-
  25 GHz f_T-**Line**

ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering Code (8-mm taped)	Pin Configuration 1	2	3	4	Package [1]
BFP420	AMs	Q62702-F1591	B	E	C	E	SOT343

Maximum Ratings

Parameter	Symbol		Unit
Collector-emitter voltage	V_{CEO}	4.5	V
Collector-base voltage	V_{CBO}	15	V
Emitter-base voltage	V_{EBO}	1.5	V
Collector current	I_C	35	mA
Base current	I_B	3	mA
Total power dissipation, Ts ≤ 107°C [2)3)]	P_{tot}	160	mW
Junction temperature	T_j	150	°C
Ambient temperature range	T_A	-65...+150°C	°C
Storage temperature range	T_{stg}	-65...+150°C	°C

Thermal Resistance

Junction-soldering point [2)]	$R_{th\ JS}$	270	K/W

1) For detailed information see chapter Package

2) T_S is measured on the emitter lead at the soldering point to the pcb.

3) P_{tot} due to Maximum Ratings.
At typical Ts ≤ 80°C: P_{tot} = 250 mW due to thermical characteristics.

SIEMENS **SIEGET®25 BFP420**

Electrical Characteristics

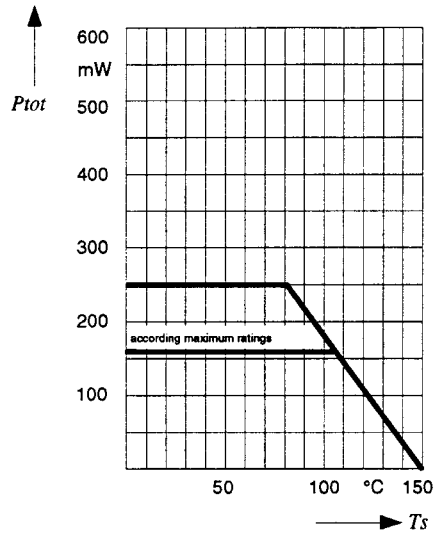
at T_A = 25 °C, unless otherwise specified.

Parameter	Symbol	Value			Unit
		min.	typ.	max.	
DC Characteristics					
Collector-emitter breakdown voltage I_C = 1 mA	$V_{(BR)CEO}$	4.5	5	6.5	V
Collector-cutoff current V_{CB} = 5 V, I_E = 0	I_{CBO}	-	-	200	nA
Emitter base cuttoff current V_{EB} = 1.5 V, I_C = 0	I_{EBO}	-	-	35	µA
DC current gain I_C = 20 mA, V_{CE} = 4 V	h_{FE}	50	80	150	
AC Characteristics					
Transition frequency I_C = 30 mA, V_{CE} = 3 V, f = 2 GHz	f_T	20	25	-	GHz
Collector-base capacitance V_{CB} = 2 V, V_{BE} = v_{be} = 0, f = 1 MHz	C_{cb}	-	0.15	0.24	pF
Collector-emitter capacitance V_{CE} = 2 V, V_{BE} = v_{be} = 0, f = 1 MHz	C_{ce}	-	0.41	-	pF
Emitter-base capacitance V_{EB} = 0.5 V, V_{CB} = v_{cb} = 0, f = 1 MHz	C_{eb}	-	0.55	-	pF
Noise figure I_C = 5 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}	F	-	1.05	1.4	dB
Power gain I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt} , Z_L = Z_{Lopt}	G_{ms} [1]	-	20	-	dB
Insertion power gain I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_L = 50Ω	$\lvert S_{21}\rvert^2$	14	17	-	dB
Third order intercept point at output I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}, Z_L = Z_{Lopt}	IP_3	-	22	-	dBm
1dB Compression point I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}, Z_L = Z_{Lopt}	P_{-1dB}	-	12	-	dBm

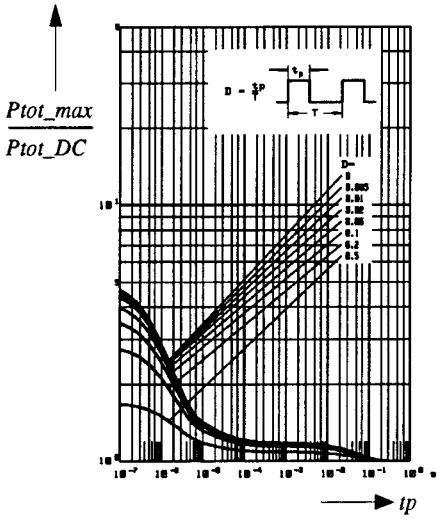
1) $G_{ms} = \left|\frac{S21}{S12}\right|$

SIEMENS **SIEGET®25 BFP420**

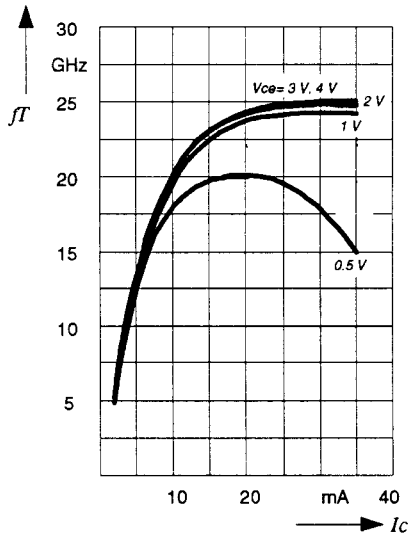
Total Power Dissipation
versus Soldering Point Temperature



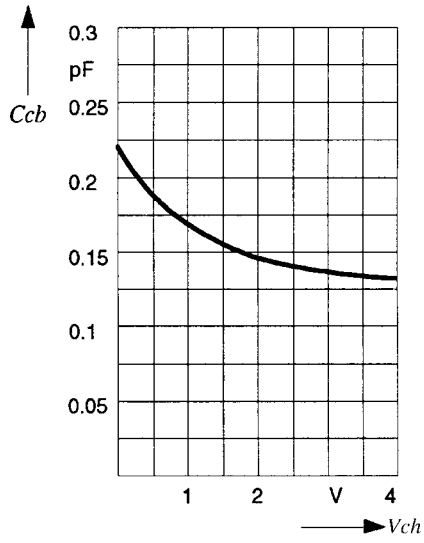
Permissible Pulse Power Dissipation
versus On-Time (V_{CE0max} = 4.5 V)



Transition Frequency
versus Collector Current
f = 2 GHz



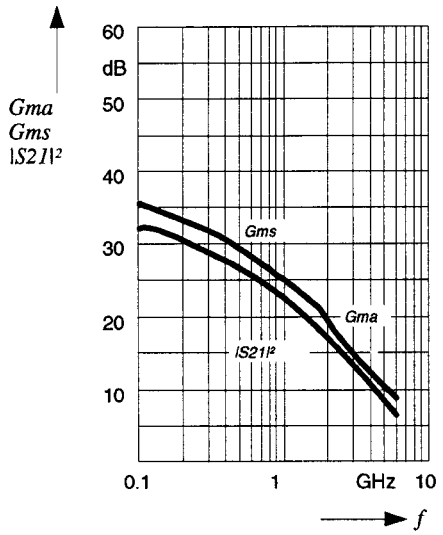
Collector-base Capacitance
versus Collector-base Voltage
V_{BE} = 0 V, f = 1MHz



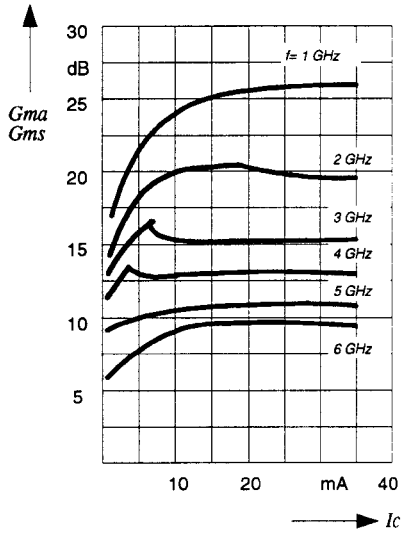
SIEMENS

SIEGET®25 BFP420

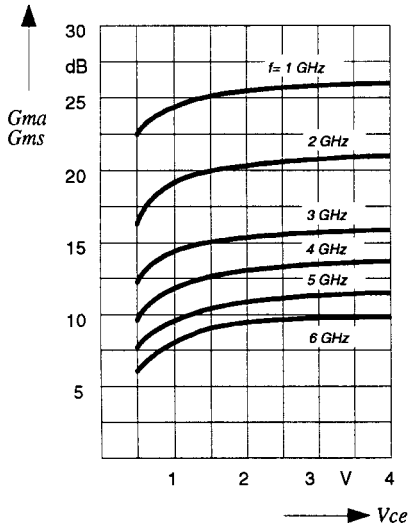
Power Gain
versus Frequency
V_{CE} = 2 V , I_C = 20 mA



Power Gain
versus Collector Current
V_{CE} = 2 V

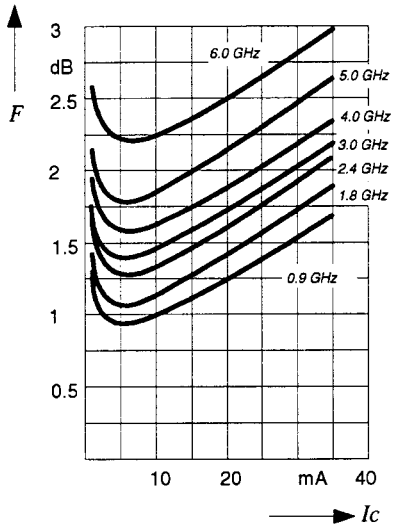


Power Gain
versus Collector Voltage
I_C = 20 mA

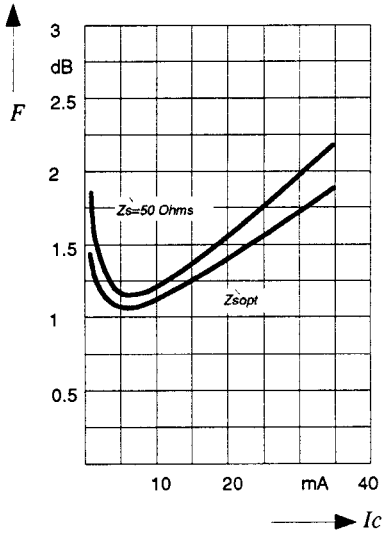


SIEMENS SIEGET®25 BFP420

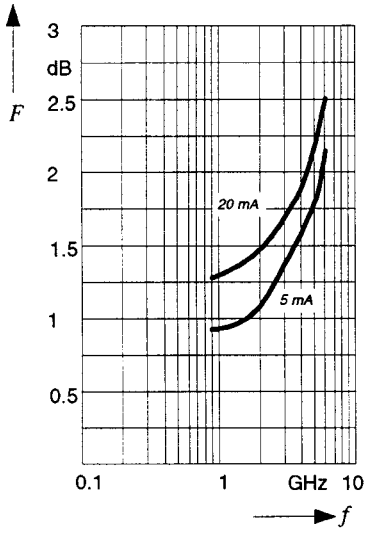
Noise Figure
versus Collector Current
V_{CE} = 2 V, Z_S = Z_{Sopt}



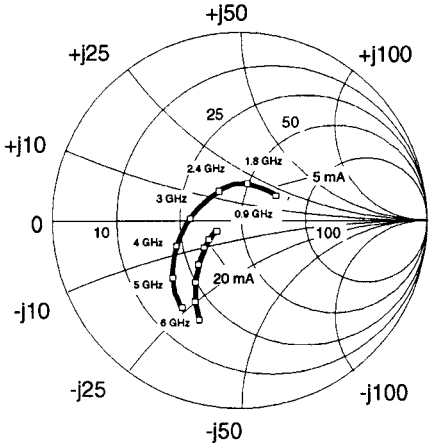
Noise Figure
versus Collector Current
VCE = 2 V, f = 1.8 GHz



Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA,
Z_S = Z_{Sopt}



Source Impedance for min.
Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA



SIEMENS **SIEGET®25 BFP420**

Common Emitter S-Parameters

f	S_{11}		S_{21}		S_{12}		S_{22}	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
V_{CE} = 2 V, I_C = 20 mA								
0.01	0.452	-2.3	37.62	178.3	0.0011	94.4	0.956	-0.6
0.1	0.447	-25.1	36.30	164.7	0.0068	82.5	0.941	-12.4
0.5	0.386	-101.1	23.41	121.0	0.0262	61.7	0.632	-47.2
1.0	0.378	-146.2	13.99	96.0	0.0395	57.8	0.395	-63.9
2.0	0.405	173.5	7.18	70.8	0.0664	54.0	0.222	-87.3
3.0	0.446	149.4	4.77	52.6	0.0949	47.1	0.133	-111.3
4.0	0.501	130.0	3.52	36.8	0.1206	38.5	0.133	-158.5
6.0	0.599	104.8	2.27	8.2	0.1646	18.9	0.196	142.0
8.0	0.700	78.5	1.51	-20.8	0.1800	-2.4	0.289	99.3
9.0	0.758	67.6	1.25	-34.4	0.1820	-13.0	0.379	84.1
10.0	0.800	62.0	1.04	-43.5	0.1800	-19.3	0.465	76.6
V_{CE} = 2 V, I_C = 5 mA								
0.01	0.790	-1.0	15.14	179.2	0.0012	83.4	0.988	-0.7
0.1	0.786	-11.6	14.98	171.8	0.0092	84.1	0.982	-6.5
0.5	0.702	-55.7	12.86	140.1	0.0398	62.8	0.857	-29.8
1.0	0.589	-99.1	9.63	112.6	0.0603	46.5	0.647	-48.6
2.0	0.507	-156.0	5.60	79.4	0.0798	34.6	0.401	-70.3
3.0	0.511	168.5	3.84	57.1	0.0957	29.8	0.267	-84.2
4.0	0.549	142.0	2.87	38.5	0.1121	25.1	0.207	-110.5
5.0	0.604	123.9	2.26	22.1	0.1285	19.4	0.150	-137.3
6.0	0.633	110.0	1.86	6.7	0.1442	13.1	0.173	-169.8

Common Emitter Noise Parameters

f	F_{min} 1)	G_a 1)	Γ_{opt}		R_N	r_n	$F_{50\Omega}$ 2)	$\lvert S_{21}\rvert^2$ 2)
GHz	dB	dB	MAG	ANG	Ω	-	dB	dB
V_{CE} = 2 V, I_C = 5 mA								
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	-175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	-157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	-142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	-123.0	15.0	0.30	3.30	5.3

1) Input matched for minimum noise figure, output for maximum gain 2) Z_S=Z_L=50 Ω

SIEMENS **SIEGET®25 BFP420**

SPICE Parameters:

Transistor Chip Data T502 (Berkeley-SPICE 2G.6 Syntax):

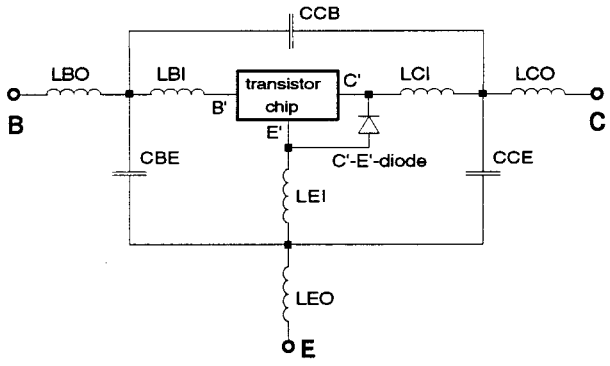
IS =	0.20045	fA	BF =	72.534	-	NF =	1.2432	-
VAF =	28.383	V	IKF =	0.48731	A	ISE =	19.049	fA
NE =	2.0518	-	BR =	7.8287	-	NR =	1.3325	-
VAR =	19.705	V	IKR =	0.69141	A	ISC =	0.019237	fA
NC =	1.1724	-	RB =	3.4849	OHM	IRB =	0.72983	mA
RBM =	8.5757	OHM	RE =	0.31111	OHM	RC =	0.10105	OHM
CJE =	1.8063	fF	VJE =	0.8051	V	MJE =	0.46576	-
TF =	6.7661	ps	XTF =	0.42199	-	VTF =	0.23794	V
ITF =	1.0	mA	PTF =	0	deg	CJC =	234.53	fF
VJC =	0.81969	V	MJC =	0.30232	-	XCJC =	0.3	-
TR =	2.3249	ns	CJS =	0	fF	VJS =	0.75	V
MJS =	0	-	XTB =	0	-	EG =	1.11	eV
XTI =	3.0	-	FC =	0.73234	-	Tnom =	300	K

C'-E'- Diode Data (Berkeley-SPICE 2G.6 Syntax):

IS =	3.5	fA	N =	1.02	-	RS =	10	Ω

All parameters are ready to use, no scaling is necessary. Caution: In these extracted parameters, RBM > the operating RB.

Package Equivalent Circuit:



SOT343-3:

LBI = 0.47 nH
LBO = 0.53 nH
LEI = 0.23 nH
LEO = 0.05 nH
LCI = 0.56 nH
LCO = 0.58 nH
CBE = 136 fF
CCE = 134 fF
CCB = 6.9 fF

valid up to 6 GHz

The SOT343 package has two emitter leads. To avoid high complexity of the package equivalent circuit, both leads are combined in one electrical connection.

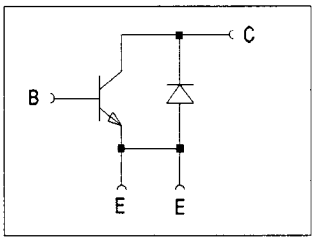
Extracted on behalf of SIEMENS Small Signal Semiconductors by:
Institut für Mobil- und Satellitenfunktechnik (IMST)
© 1996 SIEMENS AG

For non-linear simulation:

- Use transistor chip parameters in Berkeley SPICE 2G.6 syntax for all simulators.
- If you need simulation of the reverse characteristics, add the diode with the C'-E'-diode data between collector and emitter.
- Simulation of the package is not necessary for frequencies < 100 MHz.
  For higher frequencies add the wiring of the package equivalent circuit around the non-linear transistor and diode model.

Note:

- This transistor is constructed in a common emitter configuration. This feature causes an additional, reverse biased diode between emitter and collector, which does not effect normal operation.



Transistor Schematic Diagram

The common emitter configuration shows the following advantages:

- Higher gain because of lower emitter inductance.
- Power is dissipated via the grounded emitter leads, because the chip is mounted on the copper emitter leadframe.

Please note, that the broadest lead is the emitter lead.

- The AC-Characteristics are verified by random sampling.

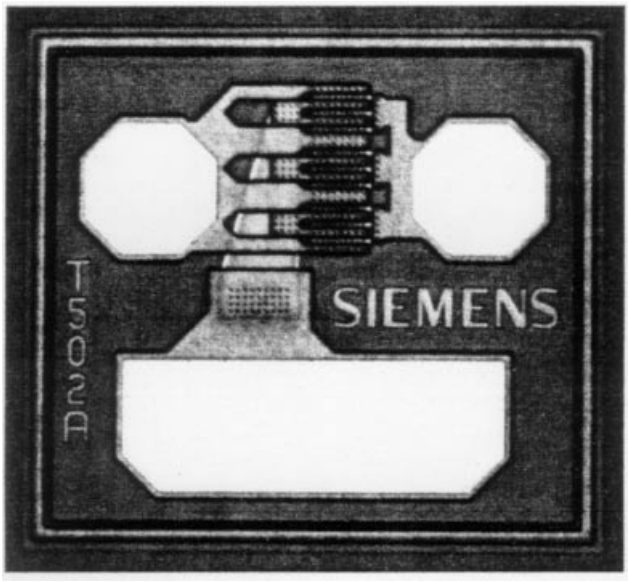


Figure 2-81 Photograph of the BFP420 die (310 × 290 μm). The bonding pad for the base is at the upper right; for the collector at the upper left. The large, lower pad connects to the emitter.

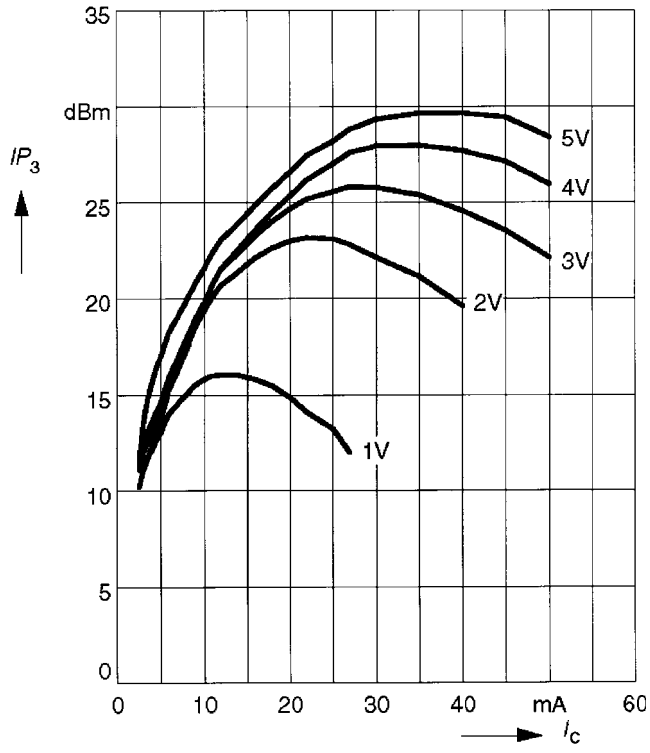


Figure 2-82 $IP_{3,out}$ versus I_C at five values of V_{CE} for the BRF93AW BJT.

are no accepted or affordable parameter extractor programs available, not even to consider the cost of hardware for such extractions. Hewlett-Packard recently published a system with a hardware cost in the vicinity of $1 million.

A typical *npn* planar bipolar transistor structure is shown in Figure 2-83a, where collector, base, and emitter are labeled *C*, *B*, and *E*, respectively. The impurity doping density in the base and the emitter of such a transistor is not constant but varies with distance from the top surface. However, many of the characteristics of such a device can be predicted by analyzing

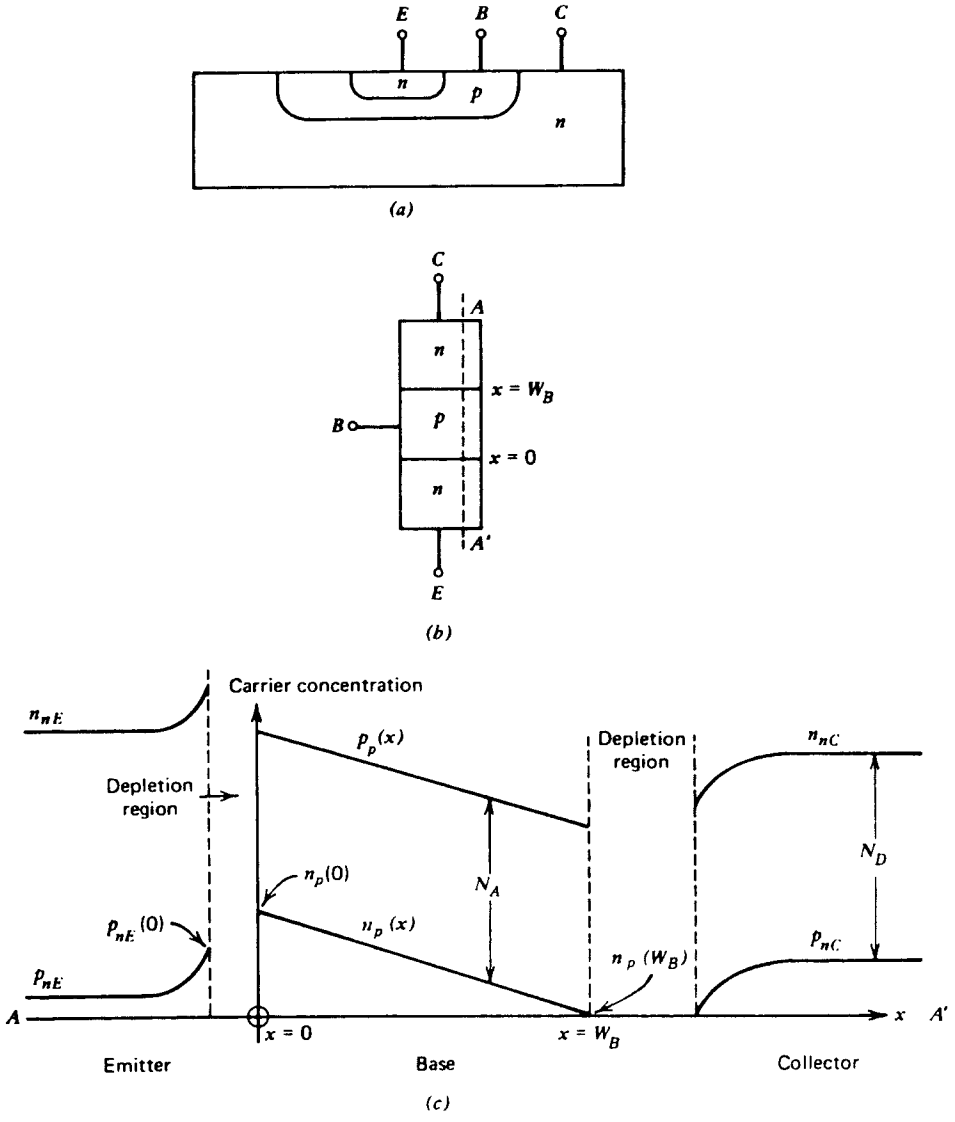


Figure 2-83 (a) Cross section of a typical *npn* planar bipolar transistor structure; (b) idealized transistor structure; (*c*) carrier concentrations along the cross section *AA*′ of the transistor in (b). Uniform doping densities are assumed. (Not to scale)

the idealized transistor structure shown in Figure 2-83b. In this structure, the base and emitter doping densities are assumed constant, and this is sometimes called a "uniform-base" transistor. Where possible in the following analyses, the equations for the uniform-base analysis are expressed in a form that applies also to nonuniform-base transistors.

A cross section AA' is taken through the device of Figure 2-83b and carrier concentrations along this section are plotted in Figure 2-83c. Hole concentrations are denoted by p and electron concentrations by n with subscripts p or n representing p-type or n-type regions. The n-type emitter and collector regions are distinguished by subscripts E and C, respectively. The carrier concentrations shown in Figure 2-83c apply to a device in the *forward-active region*. That is, the base–emitter junction is forward-biased and the base–collector junction is reverse-biased. The minority-carrier concentrations in the base at the edges of the depletion regions can be calculated from a Boltzmann approximation to the Fermi–Dirac distribution function to give [2]

$$n_p(0) = n_{po} \exp \frac{V_{BE}}{V_T} \tag{2-138}$$

$$n_p(W_B) = n_{po} \exp \frac{V_{BC}}{V_T} \simeq 0 \tag{2-139}$$

where W_B is the width of the base from the base–emitter junction depletion layer edge of the base–collector depletion layer edge and n_{po} is the equilibrium concentration of electrons in the base. Note that V_{BC} is negative for an *npn* transistor in the forward-active region and thus $n_p(W_B)$ is very small. Low-level injection conditions are assumed in the derivation of Eqs. (2-138) and (2-139). This means that the minority-carrier concentrations are always assumed much smaller than the majority-carrier concentrations.

If *recombination* of holes and electrons in the base is small, it can be shown that the minority-carrier concentration $n_p(x)$ in the base varies *linearly* with distance [3]. Thus, a straight line can be drawn joining the concentrations at $x = 0$ and $x = W_B$ in Figure 2-83c.

For charge neutrality in the base, it is necessary that

$$N_A + n_p(x) = p_p(x) \tag{2-140}$$

and thus

$$p_p(x) - n_p(x) = N_A \tag{2-141}$$

where $p_p(x)$ is the hole concentration in the base and N_A is the base doping density that is assumed constant. Equation (2-141) indicates that the hole and electron concentrations are separated by a constant amount and thus $p_p(x)$ also varies linearly with distance.

Collector current is produced by minority-carrier electrons in the base diffusing in the direction of the concentration gradient and being swept across the collector–base depletion region by the field existing there. The diffusion current density due to electrons in the base is

$$J_n = qD_n \frac{dn_p(x)}{dx} \tag{2-142}$$

where D_n is the diffusion constant for electrons. From Figure 2-83c,

$$J_n = -qD_n \frac{n_p(0)}{W_B} \tag{2-143}$$

If I_C is the collector current and is taken as positive flowing *into* the collector, it follows from Eq. (2-143) that

$$I_C = qA D_n \frac{n_p(0)}{W_B} \tag{2-144}$$

where A is the cross-sectional area of the emitter. Substitution of Eq. (2-138) into Eq. (2-144) gives

$$I_C = \frac{qAD_n n_{po}}{W_B} \exp \frac{V_{BE}}{V_T} \tag{2-145}$$

$$= I_S \exp \frac{V_{BE}}{V_T} \tag{2-146}$$

where

$$I_S = \frac{qA D_n n_{po}}{W_B} \tag{2-147}$$

and I_S is a constant used to describe the transfer characteristic of the transistor in the forward-active region. Equation (2-147) can be expressed in terms of the base doping density by noting that [4]

$$n_{po} = \frac{n_i^2}{N_A} \tag{2-148}$$

and substitution of Eq. (2-148) in Eq. (2-147) gives

$$I_S = \frac{qAD_n n_i^2}{W_B N_A} = \frac{qA\overline{D}_n n_i^2}{Q_B} \tag{2-149}$$

where $Q_B = W_B N_A$ is the number of doping atoms in the base per unit area of the emitter and n_i is the intrinsic carrier concentration in silicon. In this form Eq. (2-149) applies to both uniform- and nonuniform-base transistors and D_n has been replaced by $\overline{D}_n$, which is an average effective value of the electron diffusion constant in the base. This is necessary for nonuniform-base devices because the diffusion constant is a function of impurity concentration. Typical values of I_S as given by Eq. (2-149) are 10^{-14} to 10^{-16} A.

Equation (2-146) gives the collector current as a function of base–emitter voltage. The base current I_B is also an important parameter and, at moderate current levels, consists of two major components. One of these (I_{B1}) represents recombination of holes and electrons in the base and is proportional to the minority-carrier charge Q_e in the base. From Figure 2-83c, the minority-carrier charge in the base is

$$Q_e = \frac{1}{2} n_p(0) W_B q A \tag{2-150}$$

and we have

$$I_{B1} = \frac{Q_e}{\tau_b} = \frac{1}{2} \frac{n_p(0) W_B q A}{\tau_b} \tag{2-151}$$

where τ_b is the minority-carrier lifetime in the base. I_{B1} represents a flow of majority holes from the base lead into the base region. Substitution of Eq. (2-138) into Eq. (2-151) gives

$$I_{B1} = \frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} \exp \frac{V_{BE}}{V_T} \tag{2-152}$$

The second major component of base current (usually the dominant one in the integrated-circuit *npn* devices) is due to injection of holes from the base into the emitter. This current component depends on the gradient of minority-carrier holes into the emitter and is [5]

$$I_{B2} = \frac{q A D_p}{L_p} p_{nE}(0) \tag{2-153}$$

where D_p is the diffusion constant for holes and L_p is the diffusion length (assumed small) for holes in the emitter. $p_{nE}(0)$ is the concentration of holes in the emitter at the edge of the depletion region and is

$$P_{nE}(0) = p_{nEo} \exp \frac{V_{BE}}{V_T} \tag{2-154}$$

If N_D is the donor atom concentration in the emitter (assumed constant) then

$$p_{nEo} \simeq \frac{n_i^2}{N_D} \tag{2-155}$$

The emitter is deliberately doped much more heavily than the base, making N_D large and p_{nEo} small, so that the base-current component, I_{B2}, is minimized.

Substitution of Eqs. (2-154) and (2-155) into Eq. (2-153) gives

$$I_{B2} = \frac{q A D_p}{L_p} \frac{n_i^2}{N_D} \exp \frac{V_{BE}}{V_T} \tag{2-156}$$

The total base current, I_B, is the sum of I_{B1} and I_{B2}:

$$I_B = I_{B1} + I_{B2} = \left(\frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} + \frac{q A D_p}{L_p} \frac{n_i^2}{N_D} \right) \exp \frac{V_{BE}}{V_T} \tag{2-157}$$

Although this equation was derived assuming uniform base and emitter doping, it gives the correct functional dependence of I_B on device parameters for practical double-diffused

nonuniform-base devices. Second-order components of I_B, which are important at low current levels, are considered later.

Since I_C in Eq. (2-146) and I_B into Eq. (2-147) are both proportional to $\exp(V_{BE}/V_T)$ in this analysis, the base current can be expressed in terms of collector current as

$$I_B = \frac{I_C}{\beta_F} \tag{2-158}$$

where β_F is the forward current gain. An expression for β_F can be calculated by substituting Eqs. (2-145) and (2-157) into Eq. (2-158) to give

$$\beta_F = \frac{qAD_n n_{po}/W_B}{\frac{1}{2}\frac{n_{po}W_B qA}{\tau_b} + \frac{qAD_p n_i^2}{L_p N_D}} = \frac{1}{\frac{W_B^2}{2\tau_b D_n} + \frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \tag{2-159}$$

where Eq. (2-148) has been substituted for n_{po}. Equation (2-159) shows that β_F is maximized by minimizing the base width W_B and maximizing the ratio of emitter to base doping densities N_D/N_A. Typical values of β_F for *npn* transistors in integrated circuits are 50–500, whereas lateral *pnp* transistors have values of 10–100. Finally, the emitter current is

$$I_E = -(I_C + I_B) = -\left(I_C + \frac{I_C}{\beta_F}\right) = -\frac{I_C}{\alpha_F} \tag{2-160}$$

where

$$\alpha_F = \frac{\beta_F}{1+\beta_F} \tag{2-161}$$

The value of α_F can be expressed in terms of device parameters by substituting Eq. (2-159) into Eq. (2-161) to obtain

$$\alpha_F = \frac{1}{1+\frac{1}{\beta_F}} = \frac{1}{1+\frac{W_B^2}{2\tau_b D_n} + \frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \simeq \alpha_T \gamma \tag{2-162}$$

where

$$\alpha_T = \frac{1}{1+\frac{W_B^2}{2\tau_b D_n}} \tag{2-163}$$

$$\gamma = \frac{1}{1+\frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \tag{2-164}$$

The validity of Eq. (2-162) depends on $W_B^2/2\tau_b D_n << 1$ and $(D_p/D_n)(W_B/L_p)(N_A/N_D) << 1$, and this is always true if β_F is large [see Eq. (2-159)]. The term γ in Eq. (2-162) is

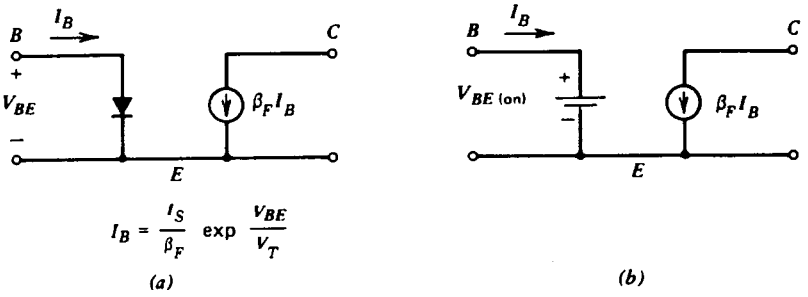


Figure 2-84 Large-signal models of *npn* transistors for use in bias calculations. (a) Circuit incorporating an input diode. (b) Simplified circuit with an input voltage source.

called the *emitter injection efficiency* and is equal to the ratio of the electron current (*npn* transistor) injected into the base from the emitter to the total hole and electron current crossing the base–emitter junction. Ideally $\gamma \to 1$, and this is achieved by making N_D/N_A large and W_B small. In that case very little reverse injection occurs from base to emitter.

The term α_T in Eq. (2-162) is called the *base transport factor* and represents the fraction of carriers injected into the base (from the emitter) that reach the collector. Ideally $\alpha_T \to 1$, and this is achieved by making W_B small. It is evident from the above development that fabrication changes that cause α_T and γ to approach unity also maximize the value of β_F of the transistor.

The results derived above allow formulation of a large-signal model of the transistor suitable for bias-circuit calculations with devices in the forward-active region. One such circuit is shown in Figure 2-84 and consists of a base–emitter diode to model Eq. (2-157) and a controlled collector-current generator to model Eq. (2-158). Note that the collector voltage ideally has no influence on the collector current and the collector node acts as a high-impedance current source. A simpler version of this equivalent circuit, which is often useful, is shown in Figure 2-84b, where the input diode has been replaced by a battery with a value $V_{BE(\text{on})}$, which is usually 0.6–0.7 V. This represents the fact that in the forward-active region the base–emitter voltage varies very little because of the steep slope of the exponential characteristic. In some circuits the temperature coefficient of $V_{BE(\text{on})}$ is important and a typical value for this is –2 mV/°C. The equivalent circuits of Figure 2-84 apply for *npn* transistors. For *pnp* devices the corresponding equivalent circuits are shown in Figure 2-85.

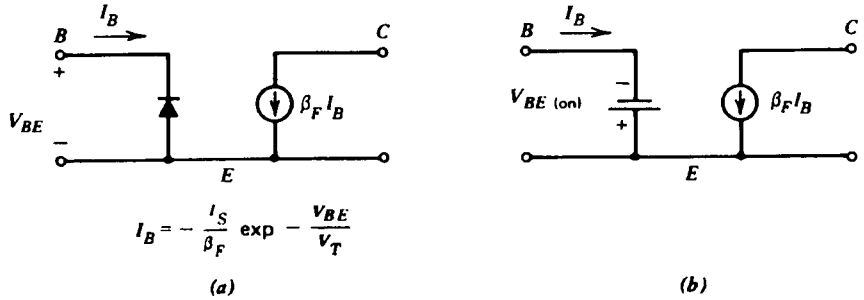


Figure 2-85 Large-signal models of *pnp* transistors corresponding to the circuits of Figure 2-84.

2-2-4 Effects of Collector Voltage on Large-Signal Characteristics in the Forward-Active Region

In the analysis of the previous section, the collector–base junction was assumed reverse-biased and ideally had no effect on the collector currents. This is a useful approximation for first-order calculations, but it is not strictly true in practice. There are occasions where the influence of collector voltage on collector current is important, and this will now be investigated.

The collector voltage has a dramatic effect on the collector current in two regions of device operation. These are the saturation (V_{CE} approaches zero) and breakdown (V_{CE} very large) regions that will be considered later. For values of collector–emitter voltage V_{CE} between these extremes, the collector current increases slowly as V_{CE} increases. The reason for this can be seen from Figure 2-86, which is a sketch of the minority-carrier concentration in the base of the transistor. Consider the effect of changes in V_{CE} on the carrier concentration for constant V_{BE}. Since V_{BE} is constant, the change in V_{CB} equals the change in V_{CE}, and this causes an increase in the collector–base depletion-layer width as shown. The change in the base width of the transistor, ΔW_B, equals the change in depletion-layer width and causes an increase ΔI_C in the collector current.

From Eqs. (2-146) and (2-149) we have

$$I_C = \frac{qA\bar{D}_n n_i^2}{Q_B} \exp \frac{V_{BE}}{V_T} \tag{2-165}$$

Differentiation of Eq. (2-165) yields

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{qA\bar{D}_n n_i^2}{Q_B^2} \left(\exp \frac{V_{BE}}{V_T} \right) \frac{dQ_B}{dV_{CE}} \tag{2-166}$$

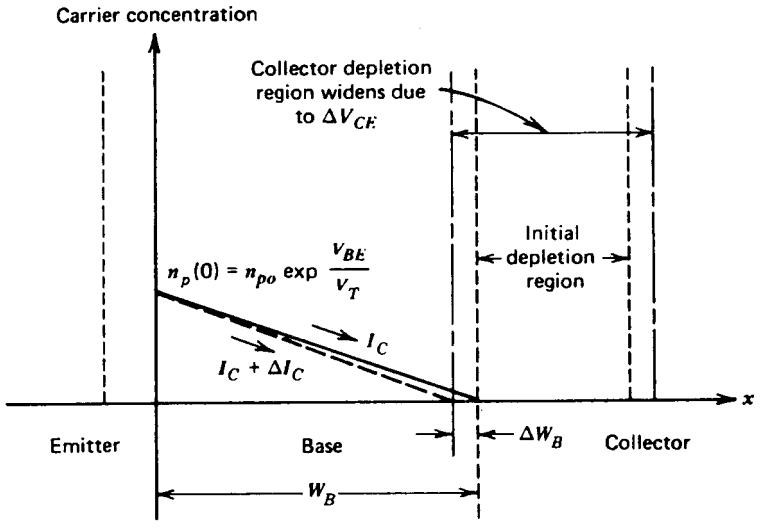


Figure 2-86 Effect of increases in V_{CE} on the collector depletion region and base width of a bipolar transistor.

and substitution of Eq. (2-165) in Eq. (2-166) gives

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{I_C}{Q_B}\frac{dQ_B}{dV_{CE}} \tag{2-167}$$

For a uniform-base transistor $Q_B = W_B N_A$ and Eq. (2-167) becomes

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{I_C}{W_B}\frac{dW_B}{dV_{CE}} \tag{2-168}$$

Note that since the base width *decreases* as V_{CE} increases, dW_B/dV_{CE} in Eq. (2-168) is negative and thus $\partial I_C/\partial V_{CE}$ is positive. dW_B/dV_{CE} is a function of the bias value of V_{CE}, but the variation is typically small for a reverse-biased junction and dW_B/dV_{CE} is often assumed constant. The resulting predictions agree adequately with experimental results.

Equation (2-168) shows that $\partial I_C/\partial V_{CE}$ is proportional to the collector-bias current and inversely proportional to the transistor base width. Thus narrow-base transistors show a greater dependence of I_C on V_{CE} in the forward-active region. The dependence of $\partial I_C/\partial V_{CE}$ on I_C results in typical transistor output characteristics as shown in Figure 2-87. In accordance with the assumptions made in the foregoing analysis, these characteristics are shown for constant values of V_{BE}. However, in most integrated-circuit transistors the base current is dependent only on V_{BE} and not on V_{CE}, and thus constant-base-current characteristics can often be used in the following calculation. The reason for this is that the base current is usually dominated by the I_{B2} component of Eq. (2-156), which has no dependence on V_{CE}. Extrapolation of the characteristics of Figure 2-87 back to the V_{CE} axis gives an intercept V_A called the Early voltage, where

$$V_A = \frac{I_C}{\dfrac{\partial I_C}{\partial V_{CE}}} \tag{2-169}$$

Substitution of Eq. (2-168) in Eq. (2-169) gives

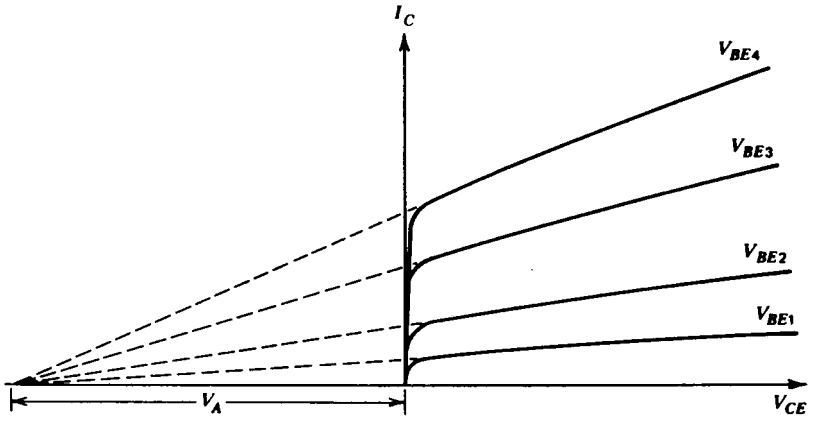


Figure 2-87 Bipolar transistor output characteristics showing the Early voltage, V_A.

$$V_A = -W_B \frac{dV_{CE}}{dW_B} \tag{2-170}$$

which is a constant, independent of I_C. Thus all the characteristics extrapolate to the same point on the V_{CE} axis. The variation of I_C with V_{CE} is called the Early effect and V_A is a common model parameter for circuit-analysis computer programs. Typical values of V_A for integrated-circuit transistors are 15–100 V. The inclusion of Early effect in dc bias calculations is usually limited to computer analysis because of the complexity introduced into the calculation. However, the influence of the Early effect is often dominant in small-signal calculations for high-gain circuits and this point will be considered later.

Finally, the influence of the Early effect on the transistor large-signal characteristics in the forward-active region can be represented approximately by modifying Eq. (2-146) to

$$I_C = I_S \left(1 + \frac{V_{BE}}{V_A}\right) \exp \frac{V_{BE}}{V_T} \tag{2-171}$$

This is a common means of representing the device output characteristics for computer simulation.

2-2-5 Saturation and Inverse Active Regions

Saturation is a region of device operation that is usually avoided in analog circuits because the transistor gain is very low in this region. Saturation is much more commonly encountered in digital circuits, where it provides a well-specified output voltage that represents a logic state. In saturation, both emitter–base and collector–base junctions are forward biased. Consequently, the collector–emitter voltage V_{CE} is quite small and usually in the range 0.05–0.3 V. The carrier concentrations in a saturated *npn* transistor with uniform base doping are shown in Figure 2-88. The minority-carrier concentration in the base of the edge of the depletion region is again given by Eq. (2-139) as

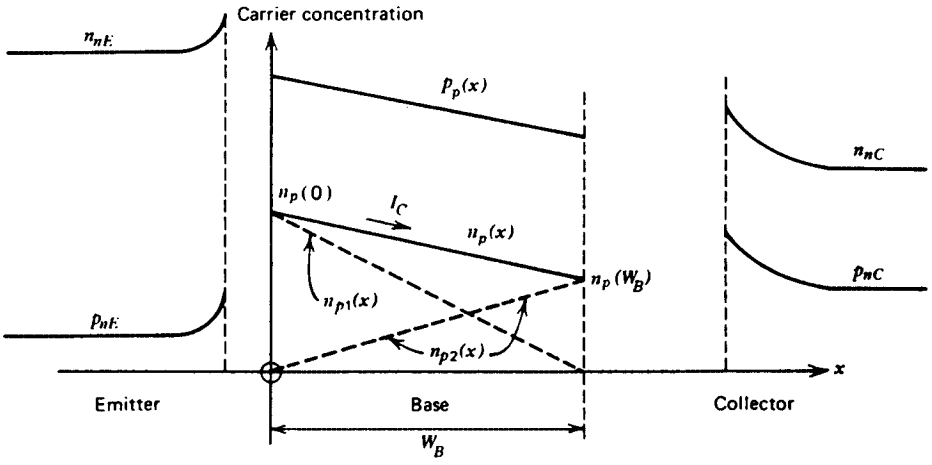


Figure 2-88 Carrier concentrations in a saturated *npn* transistor. (Not to scale)

$$n_p(W_B) = n_{po} \exp \frac{V_{BC}}{V_T} \tag{2-172}$$

but since V_{BC} is now positive, the value of $n_p(W_B)$ is no longer negligible. Consequently, changes in V_{CE} with V_{BE} held constant (which cause equal changes in V_{BC}) directly affect $n_p(W_B)$. Since the collector current is proportional to the slope of the minority-carrier concentration in the base [see Eq. (2-142)], it is also proportional to $[n_p(0) - (n_p(W_B)]$ from Figure 2-88. Thus, changes in $n_p(W_B)$ directly affect the collector current, and the collector node of the transistor appears to have a *low impedance*. As V_{CE} is decreased in saturation with V_{BE} held constant, V_{BC} increases as does $n_p(W_B)$ from Eq. (2-172). Thus, from Figure 2-88 the collector current decreases because the slope of the carrier concentration decreases. This gives rise to the saturation region of the $I_C - V_{CE}$ characteristic shown in Figure 2-89. The slope of the $I_C - V_{CE}$ characteristic in this region is largely determined by the resistance in series with the collector lead due to the finite resistivity of the *n*-type collector material. A useful model for the transistor in this region is shown in Figure 2-90

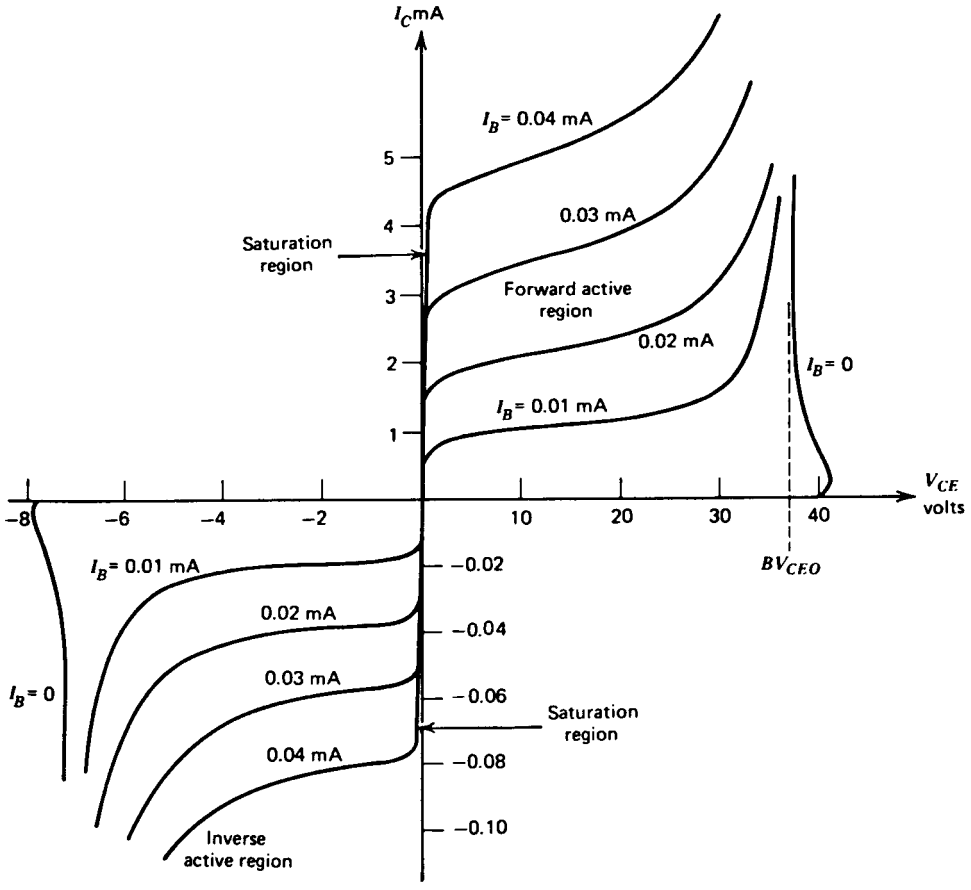


Figure 2-89 Typical $I_C - V_{CE}$ characteristics for an *npn* bipolar transistor. Note the different scales for positive and negative currents and voltages.

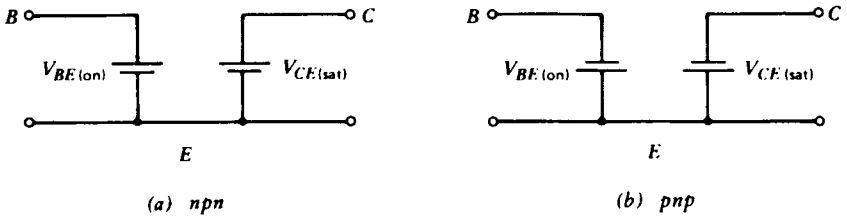


Figure 2-90 Large-signal models for bipolar transistors in the saturation region.

and consists of a fixed voltage source to represent $V_{BE(on)}$, and a fixed voltage source to represent the collector–emitter voltage $V_{CE(sat)}$. A more accurate but more complex model includes a resistor in series with the collector. This resistor can have a value ranging from 20 to 500 Ω, depending on the device structure.

An additional aspect of transistor behavior in the saturation region is apparent from Figure 2-88. For a given collector current, there is now a much larger amount of stored charge in the base than there is in the forward-active region. Thus the base-current contribution represented by Eq. (2-152) will be larger in saturation. In addition, since the collector–base junction is now forward biased, there is a new base-current component due to injection of carriers from the base to the collector. These two effects result in a base current I_B in saturation, which is larger than in the forward-active region for a given collector current I_C. Ratio I_C/I_B in saturation is often referred to as the *forced* β and is always less than β_F. As the forced β is made lower with respect to β_F, the device is said to be more *heavily saturated*.

The minority-carrier concentration in saturation, shown in Figure 2-88, is a straight line joining the two end points, assuming that recombination is small. This can be represented by a linear superposition of the two dotted distributions as shown. The justification for this is that the terminal currents depend *linearly* on the concentrations $n_p(0)$ and $n_p(W_B)$. This picture of device-carrier concentrations can be used to derive some general equations describing transistor behavior. Each of the distributions in Figure 2-88 is considered separately and the two contributions are combined. The *emitter* current that would result from $n_{p1}(x)$ above is given by the classical diode equation

$$I_{EF} = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) \tag{2-173}$$

where I_{ES} is a constant that is often referred to as the "saturation current" of the junction (no connection with the transistor saturation described above). Equation (2-173) predicts that the junction current is given by $I_{EF} \simeq I_{ES}$ with a reverse-bias voltage applied. However, in practice Eq. (2-173) is applicable only in the forward-bias region, since second-order effects dominate under reverse-bias conditions and typically result in a junction current several orders of magnitude larger than I_{ES}. The junction current that flows under reverse-bias conditions is often called the "leakage current" of the junction.

Returning to Figure 2-88, we can describe the *collector* current resulting from $n_{p2}(x)$ alone as

$$I_{CR} = -I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-174}$$

where I_{CS} is a constant. The total collector current I_C is given by I_{CR} plus the fraction of I_{EF} that reaches the collector (allowing for recombination and reverse emitter injection). Thus,

$$I_C = \alpha_F I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) - I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-175}$$

where α_F has been defined previously by Eq. (2-162). Similarly, the total emitter current is composed of I_{EF} plus the fraction of I_{CR} that reaches the emitter with the transistor acting in an inverted mode. Thus,

$$I_E = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + \alpha_R I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-176}$$

where α_R is the ratio of emitter to collector current with the transistor operating *inverted* (i.e., with the collector–base junction forward biased and emitting carriers into the base and the emitter–base junction reverse biased and collecting carriers). Typical values of α_R are 0.5–0.8. An inverse current gain β_R is also defined,

$$\beta_R = \frac{\alpha_R}{1 - \alpha_R} \tag{2-177}$$

and has typical values 1–5. This is the current gain of the transistor when operated inverted and is much lower than β_F because the device geometry and doping densities are designed to maximize β_F. The inverse-active region of device operation occurs for V_{CE} negative in an *npn* transistor and is shown in Figure 2-89. In order to display these characteristics adequately in the same figure as the forward-active region, the negative voltage and current scales have been expanded. The inverse-active mode of operation is rarely encountered in analog circuits.

Equations (2-175) and (2-176) describe *npn* transistor operation in the saturation region when V_{BE} and V_{BC} are both positive, and also in the forward-active and inverse-active regions. These equations are from the *Ebers–Moll* equations. In the forward-active region, they degenerate into a form similar to that of Eqs. (2-146), (2-158), and (2-160) derived earlier. This can be shown by putting V_{BE} positive and V_{BC} negative in Eqs. (2-175) and (2-176) to obtain

$$I_C = \alpha_F I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + I_{CS} \tag{2-178}$$

$$I_E = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) - \alpha_R I_{CS} \tag{2-179}$$

Equation (2-178) is similar in form to Eq. (2-146) except that leakage currents that were previously neglected have now been included. This minor difference is significant only at high temperatures or very low operating currents. Comparison of Eq. (2-178) with Eq. (2-146) allows us to identify $I_S = \alpha_F I_{ES}$ and it can be shown [6] in general that

$$\alpha_F I_{ES} = \alpha_R I_{CS} = I_S \tag{2-180}$$

where this expression represents a reciprocity condition. Using Eq. (2-180) in Eqs. (2-175) and (2-176) allows the Ebers–Moll equations to be expressed in the general form

$$I_C = I_S\left(\exp\frac{V_{BE}}{V_T} - 1\right) - \frac{I_S}{\alpha_R}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-181}$$

$$I_E = -\frac{I_S}{\alpha_F}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + I_S\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-182}$$

This form is often used for computer representation of transistor large-signal behavior.

The effect of leakage currents mentioned above can be further illustrated as follows. In the forward-active region, we have, from Eq. (2-179),

$$I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) = -I_E - \alpha_R I_{CS} \tag{2-183}$$

Substitution of Eq. (2-183) in Eq. (2-178) gives

$$I_C = -\alpha_F I_E + I_{CO} \tag{2-184}$$

where

$$I_{CO} = I_{CS}(1 - \alpha_R\alpha_F) \tag{2-185}$$

and I_{CO} is the collector–base leakage current with the emitter open. Although I_{CO} is given theoretically by Eq. (2-185), in practice, surface leakage effects dominate when the collector–base junction is reverse biased and I_{CO} is typically several orders of magnitude larger than the value given by Eq. (2-185). However, Eq. (2-184) is still valid if the appropriate measured value for I_{CO} is used. Typical values of I_{CO} are 10^{-10} to 10^{-12} A at 25 °C, and the magnitude doubles about every 8 °C. As a consequence, these leakage terms can become very significant at high temperatures. For example, consider the base current I_B. This is

$$I_B = -(I_C + I_E) \tag{2-186}$$

If I_E is calculated from Eq. (2-184) and substituted in Eq. (2-186), the result is

$$I_B = \frac{1 - \alpha_F}{\alpha_F} I_C - \frac{I_{CO}}{\alpha_F} \tag{2-187}$$

But from Eq. (2-161)

$$\beta_f = \frac{\alpha_F}{1 - \alpha_F} \tag{2-188}$$

and use of Eq. (2-188) in Eq. (2-187) gives

$$I_B = \frac{I_C}{\beta_F} - \frac{I_{CO}}{\alpha_F} \tag{2-189}$$

Since the two terms in Eq. (2-189) have opposite signs, the effect of I_{CO} is to *decrease* the magnitude of the external base current at a given value of collector current.

2-2-6 Small-Signal Models of Bipolar Transistors

Figure 2-91 shows the small-signal equivalent BJT model. It consists of an intrinsic model and a package model. Table 2-12 lists its keywords.

In order to better understand the difference between the various technologies, we show Table 2-13, which allows the reader to compare their performances.

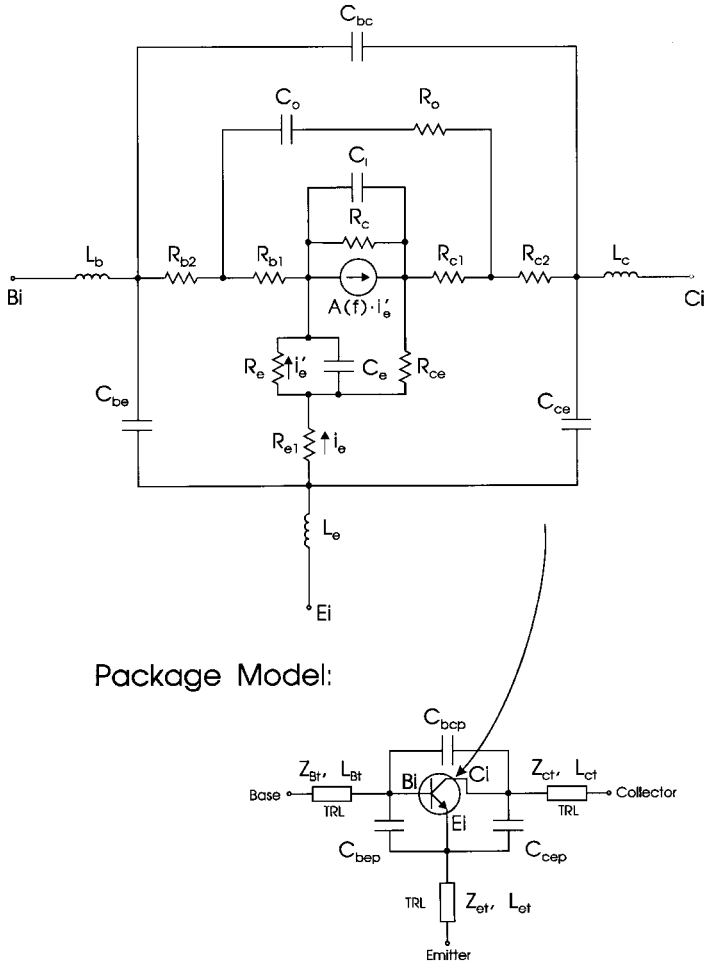


Figure 2-91 Linear BJT model.

Table 2-12 Small-signal equivalent BJT model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
A	Ratio of I_C to I_E at dc		0.0
RE	Emitter resistance	ohm	0.0
F	Current generator roll-off frequency	hertz	∞
T	Time delay	second	0.0
CE	Emitter capacitance	farad	0.0
CI	Collector capacitance	farad	0.0
RCE	Collector–emitter resistance	ohm	∞
RC	Collector resistance	ohm	∞
RO	Extrinsic base–collector resistance	ohm	0.0
CO	Extrinsic base–collector capacitance	farad	0.0
RB1	Intrinsic base resistance (R_{bb})	ohm	0.0
RC1	Parasitic collector resistance	ohm	0.0
RE1	Parasitic emitter resistance	ohm	0.0
RB2	Parasitic base resistance	ohm	0.0
RC2	Parasitic collector resistance	ohm	0.0
CBE	Base-to-emitter package capacitance	farad	0.0
CBC	Base-to-collector package capacitance	farad	0.0
CCE	Collector-to-emitter package capacitance	farad	0.0
LB	Base-lead inductance	henry	0.0
LC	Collector-lead inductance	henry	0.0
LE	Emitter-lead inductance	henry	0.0
TJ	Chip temperature	Kelvin	298
NFAC	Noise factor proportional to drive		1.0
FC	Flicker noise (1/*f* noise) corner frequency	hertz	50
	Package Model		
CBCP	Base-to-collector package capacitance	farad	0.0
CBEP	Base-to-emitter package capacitance	farad	0.0
CCEP	Collector-to-emitter package capacitance	farad	0.0
ZBT	Base transmission line impedance	ohm	50
ZCT	Collector transmission line impedance	ohm	50
ZET	Emitter transmission line impedance	ohm	50
LBT	Base transmission line length at $\varepsilon_r = 1$	meter	0.0
LCT	Collector transmission line length at $\varepsilon_r = 1$	meter	0.0
LET	Emitter transmission line length at $\varepsilon_r = 1$	meter	0.0

Notes:

1. $A \equiv \alpha = I_c/I_e$; β = dc current gain = $\alpha(1 - \alpha)$.
2. The bipolar current gain in this model is described by

$$\mathrm{A} = \mathrm{A}(0) = \frac{e^{-j\omega\mathrm{T}}}{1 + j\,(f/F)}$$

where $\omega = 2\,\pi f$ and f is frequency.

3. The current source is controlled by the current through R_e. The current generator has a cutoff frequency with respect to the total emitter current, I_E:

$$\mathrm{F} = g_m/2\pi C_e$$

where $g_m = 1/R_e$. This frequency becomes infinite for the default value for C_e (0.0). The parammeter F specifies the frequency roll-off for the current generator with respect to the current through R_e. Effectively, this frequency parameter may be used to model additional delays in the device.

Table 2-13 Overview of current BJT technologies including BiCMOS

Parameter	MOSAIC 3	MOSAIC 5	MOSAIC 5SE	RFBICMOS1	TFSOI
Emitter Size	$0.95 \times 3.2\ \mu m^2$	$0.3 \times 0.9\ \mu m^2$	$0.3 \times 0.9\ \mu m^2$	$0.4 \times 1.3\ \mu m^2$	$0.1 \times 1.2\ \mu m^2$
β	150	150	90	80	40–65
F_t (GHz)	14@ $V_{ce} = 1$ V	14@ $V_{ce} = 1$ V	27@ $V_{ce} = 1$ V	16@ $V_{ce} = 1$ V	8@ $V_{ce} = 2$ V
C_{je} (fF)	19.9	4.3	3.8	5.1	1.5
C_{jc} (fF)	25.1	2.8	2.8	3.7	0.34
C_{js} (fF)	49.6	4.3	3.6	7.4	0.95
R_e (Ω)	22	36	50	34	72
R_b (Ω)	462	578	583	462	n/a
BV_{ceo} (V)	6.5	7.0	4.8	6.0	>3.5
BV_{cbo} (V)	20	20	15	18	>10
BV_{ebo} (V)	5.2	4.2	3.8	6.0	4.3
V_A (V)	30	48	20	>25	25–40
Minimum T_{pd}	75 ps @ 1.5 mA	35 ps @ 800 μA	36 ps @ 800 μA	43 ps @ 300 μA	52 ps @ 25 μA

Parameter	TFSOI Ring Device	RFBICMOS3 Analog	RFBICMOS3 Digital	BiGCMOS3 Analog	BiGCMOS3 Digital
Emitter Size	$0.1 \times 7.2\ \mu m^2$	$0.25 \times 0.75\ \mu m^2$	$0.25 \times 0.75\ \mu m^2$	$0.2 \times 0.6\ \mu m^2$	$0.2 \times 0.6\ \mu m^2$
β	40–65	100	150	100	100
F_t (GHz)	8 @ $V_{ce} = 2$ V	15 @ $V_{ce} = 2$ V	25 @ $V_{ce} = 2$ V	15 @ $V_{ce} = 1$ V	35 @ $V_{ce} = 1$ V
C_{je} (fF)	3.4	1.8	1.8	1.7	1.7
C_{jc} (fF)	2.8	1.1	1.1	0.2	0.3
C_{js} (fF)	2.4	12.0	12.0	4	4
R_e (Ω)	26	140	140	113	113
R_b (Ω)	790	n/a	n/a	10 kΩ/sq.	15 kΩ/sq.
BV_{ceo} (V)	>3.5	6.5	4	6.5	4.8
BV_{cbo} (V)	>10	15	15	20	15
BV_{ebo} (V)	4.3	4	4	4	4
V_A (V)	30–40	40	25	40	20
Minimum T_{pd}	75 ps @ 100 μA	n/a	n/a	35 ps @ 50 μA	22 ps @ 200 μA

Parameter	AT&T BEST2[a]	Fujitsu POSET	HP HP-25	HP ISOSAT II[b]	Hitachi SMI
Emitter Size	$1.0 \times 2.0\ \mu m^2$	$0.1 \times 1.7\ \mu m^2$	$0.4 \times 20\ \mu m^2$	$0.6 \times ?\ \mu m^2$	$0.2 \times 2.0\ \mu m^2$
β	100	110	75	100	500
F_t (GHz)	23 @ $V_{ce} = 3$ V	38 @ $V_{ce} = 3$ V	28.5 @ $V_{ce} = 3$ V	14/F_{max} > 20 GHz	82@ $V_{ce} = 2$ V
C_{je} (fF)	5.0	3.5	42	n/a	2.9
C_{jc} (fF)	5.0	2.2	37	n/a	4.8
C_{js} (fF)	8.0	n/a	123	n/a	4.1
R_e (Ω)	25	25	n/a	n/a	30
R_b (Ω)	10 kΩ/sq.	960	100	n/a	240
BV_{ceo} (V)	5.5	5.4	5.2	15	2.5
BV_{cbo} (V)	11	18.2	15	25	6.4
BV_{ebo} (V)	8.0	4.5	>5	2.0	6.2
V_A (V)	24	n/a	n/a	25	8

(continued)

Table 2-13 (Continued)

Parameter	AT&T BEST2[a]	Fujitsu POSET	HP HP-25	HP ISOSAT II[b]	Hitachi SMI
Minimum T_{pd}	37 ps @ 2.1 mA 48 ps @ 600 µA	21.5 ps @ 320 µA	n/a	n/a	16 ps @ $I_g = 1.3$

Parameter	IBM SiGe[c]	Maxim GST-1[d]	Maxim GST-2[e]	Mitsubishi[f]	National ABIC V
Emitter Size	0.5×1.0 µm^2	1.2×2.6 µm^2	0.7×1.5 µm^2	0.5×5.7 µm^2	0.5×0.5 µm^2
β	100	75	190	90	85
F_t (GHz)	41 @ $V_{ce} = 1$ V	12.5	26 @ $V_{ce} = 4$ V	20	13 @ $V_{ce} = 2$ V
C_{je} (fF)	5.2	9	4	n/a	1.9
C_{jc} (fF)	3.4	15	7	9	3.7
C_{js} (fF)	3.6	11	9	n/a	9.6
R_e (Ω)	n/a	n/a	80	n/a	n/a
R_b (Ω)	300	600	710	155	383
BV_{ceo} (V)	3.3	>5	>4	6.2	8.5
BV_{cbo} (V)	9.5	>12	>14	n/a	20
BV_{ebo} (V)	n/a	>4	>4	n/a	7.5
V_A (V)	n/a	20	15	n/a	n/a
Minimum T_{pd}	n/a	n/a	n/a	n/a	38 ps @ 1.8 mA

Parameter	NEC BSA	NEC SiGe[g]	NEC NESAT III	Philips QBiC2	Philips PRET[g]
Emitter	0.4×7.6 µm^2	0.2×1.6 µm^2	0.6×40 µm^2	0.5×1.2 µm^2	0.2×0.9 µm^2
β	n/a	120	80	100	50
F_t (GHz)	43	51	20 @ $V_{ce} = 3$ V	17	14
C_{je} (fF)	20	n/a	120	2.4	1.5
C_{jc} (fF)	16	1.7	140	3.5	1.0
C_{js} (fF)	n/a	n/a	160	6.5	5.0
R_e (Ω)	n/a	80	n/a	120	120
R_b (Ω)	43	220	55	700 Ω/sq.	1 kΩ/sq.
BV_{ceo} (V)	3.3	2.7	6	4.5	7
BV_{cbo} (V)	n/a	n/a	17	13	12.5
BV_{ebo} (V)	2	2.2	4.5	5.5	5.5
V_A (V)	n/a	n/a	n/a	15	20
Minimum T_{pd}	29 ps @ 300 µA	19 ps @ 300 µA	n/a	n/a	n/a

Parameter	Philips SiGe[g]	Siemens B6HF	Siemens B6HF Plus	Siemens SiGe[g]	SONY[h]
Emitter Size	0.4×1.4 µm^2	0.8×3.3 (drawn) µm^2	0.25×2.7 µm^2	0.27×2.5 µm^2	$0.2 \times ?$ µm^2
β	100	>60	115	130	100
F_t (GHz)	45 @ $V_{ce} = 1.5$ V	25 @ $V_{cb} = 0$ V	50 @ $V_{cb} = 2$ V	61 @ $V_{cb} = 1$ V	30
C_{je} (fF)	9.1	11	8.5	7.5	n/a
C_{jc} (fF)	4.6	6.5	8.0	5.5	n/a
C_{js} (fF)	10.8	31	26.5	13	n/a

(continued)

Table 2-13 Overview of current BJT technologies including BiCMOS (Continued)

Parameter	Phillips SiGe[g]	Siemens B6HF	Siemens B6HF Plus	Siemens SiGe[g]	SONY[h]
R_e (Ω)	n/a	n/a	n/a	n/a	n/a
R_b (Ω)	n/a	165	14 kΩ/sq.	8 kΩ/sq.	n/a
BV_{ceo} (V)	3.2	>3.5	7	2.5	4.0
BV_{cbo} (V)	11	n/a	10.8	9.9	18
BV_{ebo} (V)	3.5	n/a	2.0	3.0	3.5
V_A (V)	70–100	n/a	20	130	10
Minimum T_{pd}	14 ps @ 6 mA	25 ps @ 1.2 mA (CML)	16 ps @ 2.0 mA (CML)	11 ps @ 1.9 mA (CML)	21 ps @ I_g = 3.5 mA (CML)

Parameter	STM HSB[i]	STM HSB2A	iSTM HSB3[g,j]	Toshiba	Toshiba Epi Base[g]
Emitter Size	$0.35 \times ? \ \mu m^2$	$0.35 \times ? \ \mu m^2$	$0.2 \times ? \ \mu m^2$	$0.4 \times 5.0 \ \mu m^2$	$0.6 \times 5.0 \ \mu m^2$
β	>50	>50	>50	60	105
F_t (GHz)	n/a	20	~50	22 @ V_{ce} = 2 V	52 @ V_{ce} = 3 V
C_{je} (fF)	n/a	n/a	n/a	9.0	16.1
C_{jc} (fF)	n/a	n/a	n/a	3.6	14.1
C_{js} (fF)	n/a	n/a	n/a	11	n/a
R_e (Ω)	n/a	n/a	n/a	14	22.7
R_b (Ω)	10 kΩ/sq.	n/a	n/a	90	205
BV_{ceo} (V)	4.5	n/a	n/a	6.2	4.1
BV_{cbo} (V)	n/a	n/a	n/a	14.8	9.5
BV_{ebo} (V)	n/a	n/a	n/a	4.2	n/a
V_A (V)	n/a	>30	n/a	n/a	26.5
Minimum T_{pd}	n/a	30 ps	n/a	19 ps @ 1.2 mA	n/a

[a] 1.0 µm Design Rule; NF = 3.25 dB, G_A = 12 dB at I_C = 2 mA, freq. = 2 GHz for $16 \times 1 \times 2 \ \mu m^2$; Resistors: 150 Ω/sq. (mono), 300 Ω/sq. (mono), 500 Ω/sq. (poly); Layers of metal, Schottky diodes, linear capacitor (silicided n^+ poly/n^+Si); Gate length = 0.68 µm (*n*-ch)/0.8 µm (*p*-ch), Gate Oxide = 12.5 nm (3 V)/15 nm (5 V), V_{th} = 0.65 V (*n*-ch)/–0.9 V (*p*-ch), Gm (µS/µm) = 204 (*n*-ch)/65 (*p*-ch).

[b] 0.6 µm Design Rule; Gold metallization, ILD: polyimide; Inductors: $Q > 10$ for $L \leq 10$ nH; Capacitors: $Q > 25$ for $C \leq 20$ pF; Schottky diodes (f_c = 100 GHz), Zener diodes, MIS capacitor, thin film resistors, and LPNP, 10 mask process.

[c] Analog devices using this process.

[d] LPNP F_t = 80 MHz; three-layer metallization; Schottky diodes (R_sC_j ~ 4 ps); Implanted resistors: 200 Ω/sq.; Poly resistors: 400/2000 Ω/sq.; NiCr thin film resistors (50 Ω/sq.) with laser trim option; MOS capacitor.

[e] LPNP F_t = 85 MHz; two-layer metallization; Schottky diodes (R_sC_j ~ 0.8 ps); Implanted resistors: 170/600/1200 Ω/sq.; NiCr thin film resistors (50 Ω/sq.) with laser trim option; MOS capacitor.

[f] L_n = 0.8 µm, T_{ox} = 18 nm, V_{th} = 0.76 V (nmos), V_{th} = –0.78 V (pmos), 2poly-MIM capacitor: 2.7 fF/µm^2, Schottkys, inductors.

[g] Not known if this process is in production.

[h] 0.8 µm Design Rules; 2.5 fF/µm^2 MIS capacitor; 600 MHz PJFET; 12L; F_t = 4 GHz VPNP; poly resistors 80/400/2000 Ω/sq.

[i] Targeted for 900-MHz GSM radio front-end blocks, a CT2 cordless single-chip radio, and DECT (digital enhanced cordless telephones).

[j] Targeted for multimode 900-MHz GSM, 1.8-GHz PCN and 1.9-GHz PCS based systems. Company is also evaluating integrating SiGe in HSB3.

2-3 FIELD-EFFECT TRANSISTORS

Both diodes and *pnp/npn* transistors work on the injection principle, which means the base–emitter junction generates either free electrons or other carriers, which are being "collected" by the collector connection and the base connection controls it. While this is a fairly complex process, its physics is well understood, and its latest spin-offs, heterojunction BJTs and silicon–germanium transistors, seem to offer a bright future.

In opposition to this, the field-effect transistor (FET) controls the electronic conduction in a solid by an electric field, and this concept actually predates the invention of the bipolar transistor. J. E. Lilienfeld filed for a patent on such a device in 1925 (U.S. Patent No. 1,745,175), and W. Shockley presented a comprehensive theory of the FET in 1952. It took until 1960 before the first commercially available FETs came on the market.

Several types of semiconductor materials have been used for making FETs: silicon, germanium, gallium arsenide, and others have been used. By far the most widely used is silicon, followed now by gallium arsenide.

The FET is a class of electronic semiconductor device in which the conduction of a "channel" between source (*S*) and drain (*D*) terminals is controlled by an electric field impressed on the channel via a gate (*G*) terminal. The conduction channel may utilize *n*-type carriers (electrons) or *p*-type carriers (holes). The electric field that controls the channel conduction may be introduced via a *pn* junction [for a "junction" FET (JFET)], a metal plate separated from the semiconductor channel by an oxide dielectric [for a metal-oxide semiconductor (MOS) FET], or a combination of the two methods. The polarity of the controlling electrical field is a function of the type of carriers in the channel. A FET "family tree" is shown in Figure 2-92.

For today's microwave and RF integrated circuits, the JFET with its high-frequency problems (low cutoff frequency) and large variation in parameters, such as transconductance and pinchoff voltage, has fallen out of favor and is mainly used in discrete circuits such as two-way radios or other RF applications that tolerate these wide variations and temperature dependence. The most-used FET types are MOS for small-signal applications, LDMOS (lateral diffused MOS) for power applications, and GaAs MESFET for lower-power applications (low-voltage applications) as well as high-power applications (recent advertisements by NEC have shown output power up to 100 W). Additional information on the FET family is shown in Figure 2-93.

All members of the FET family are high-impedance-input devices. Figure 2-94 shows the key operating parameters for MOSFETs and MESFETs in low-noise stages, high-gain stages, and linearized power stages. The linearization also improves the matching. This figure is similar to its counterpart in Section 2-2-2.

The following several pages present a typical example of a datasheet* for a JFET (U310) and the chip (NZA) on which it is based, followed by an example of a device with a much higher dynamic range (the CP640 family). The CP640 family and devices similar to them are being replaced by low-power LDMOS, which are not yet well-documented. At least one company, InterFET, is devoted entirely to producing discrete JFETs and JFET-related hybrid and small-scale-integration IC products.

* Reproduced with permission of Vishay-Siliconix.

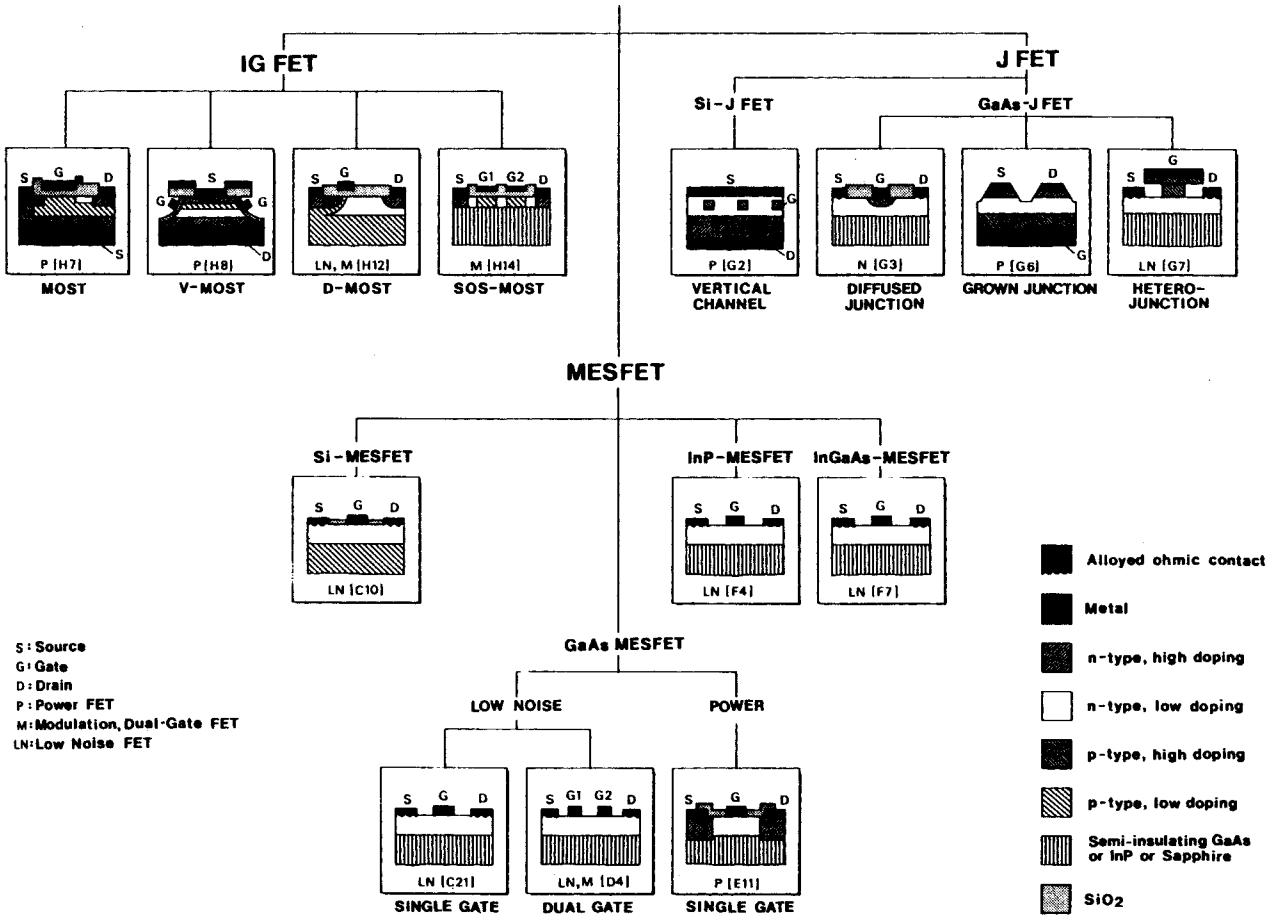


Figure 2-92 "Family tree" of microwave FET types showing cross sections of their structures [7]. Copyright ©1976 IEEE.

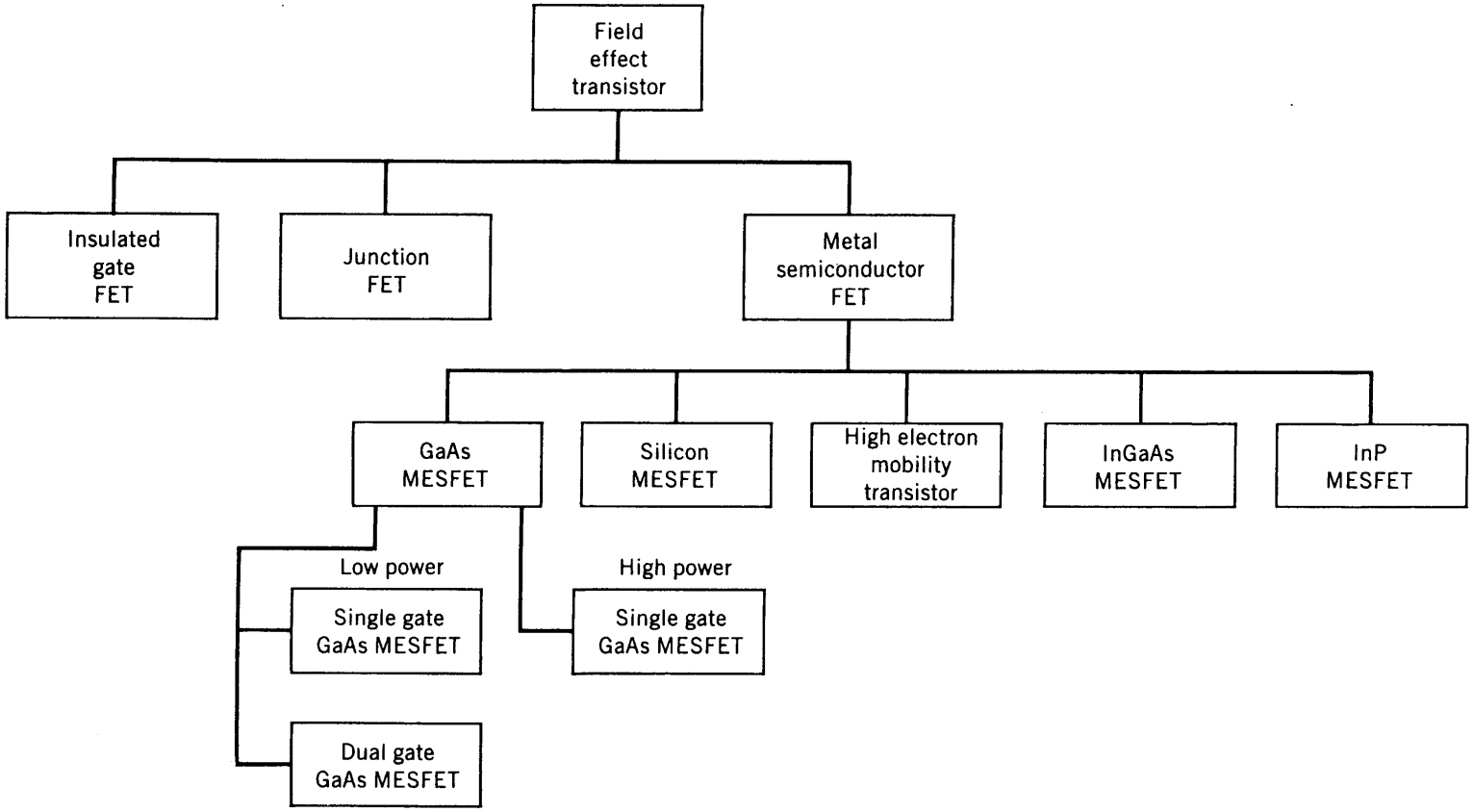


Figure 2-93 Another FET family tree [8].

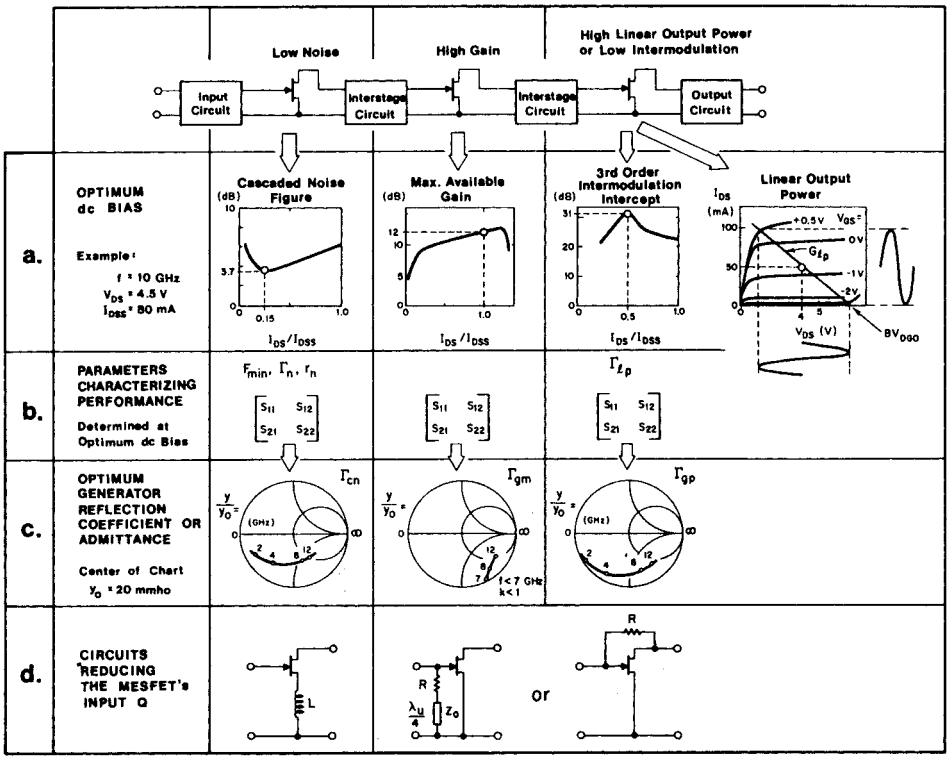


Figure 2-94 Key parameters in operating a MESFET in a low-noise front-end stage, a high-gain intermediate stage, and a linear power stage, showing (a) the device characteristics that lead to optimum dc bias for each service, (b) the parameters that characterize the device performance, (c) the optimum generator reflection coefficient that must be synthesized by each circuit, and (d) circuit arrangements that reduce the transistor's input *Q* or stabilize it at low frequencies. Although this example is based on an HP MESFET with a gate length of 1 μm and a gate width of 500 μm, this principle applies to all members of the FET family [7]. Copyright ©1976 IEEE.

n-channel JFETs *designed for . . .*

Siliconix

Performance Curves NZA See Section 4

- **VHF Amplifiers**
- **Front End High Sensitivity Amplifiers**
- **Oscillators**
- **Mixers**

BENEFITS

- Industry Standard
- High Power Gain
  16 dB at 105 MHz, Common-Gate
  11 dB at 450 MHz, Common-Gate
- Low Noise
  2.7 dB Noise Figure at 450 MHz
- Wide Dynamic Range
  Greater than 100 dB
- 75 Ω Input Match Common Gate

ABSOLUTE MAXIMUM RATINGS (25°C)

Gate-Drain or Gate-Source Voltage . . . . . -25 V
Gate Current . . . . . 20 mA
Total Power Dissipation at T_A = 25°C . . . . . 500 mW
Power Derating to 150°C . . . . . 4.0 mW/°C
Storage Temperature Range . . . . . -65 to +200°C
Lead Temperature
(1/16" from case for 10 seconds) . . . . . 300°C

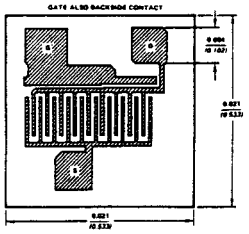
TO-52
See Section 6

ELECTRICAL CHARACTERISTICS (25°C unless otherwise noted)

		Characteristic		U308 Min	U308 Typ	U308 Max	U309 Min	U309 Typ	U309 Max	U310 Min	U310 Typ	U310 Max	Unit	Test Conditions	
1	STATIC	I_{GSS}	Gate Reverse Current			-150			-150			-150	pA	V_{GS} = -15 V,	
2						-150			-150			-150	nA	V_{GS} = 0	T_A = 125°C
3		BV_{GSS}	Gate-Source Breakdown Voltage	-25			-25			-25			V	I_G = -1 μA, V_{DS} = 0	
4		$V_{GS(off)}$	Gate-Source Cutoff Voltage	-1.0		-6.0	-1.0		-4.0	-2.5		-6.0		V_{DS} = 10 V, I_D = 1 nA	
5		I_{DSS}	Saturation Drain Current (Note 1)	12		60	12		30	24		60	mA	V_{DS} = 10 V, V_{GS} = 0	
6		$V_{GS(f)}$	Gate-Source Forward Voltage			1.0			1.0			1.0	V	I_G = 10 mA, V_{DS} = 0	
7	DYNAMIC	g_{fg}	Common-Gate Forward Transconductance (Note 1)	10	17		10	17		10	17		mmho	V_{DS} = 10 V, I_D = 10 mA	f = 1 kHz
8		g_{og}	Common-Gate Output Conductance			250			250			250	μmho		
9		C_{gd}	Drain-Gate Capacitance			2.5			2.5			2.5	pF	V_{GS} = -10 V, V_{DS} = 10 V	f = 1 MHz
10		C_{gs}	Gate-Source Capacitance			5.0			5.0			5.0			
11		$\overline{e_n}$	Equivalent Short Circuit Input Noise Voltage		10			10			10		$\frac{nV}{\sqrt{Hz}}$	V_{DS} = 10 V, I_D = 10 mA	f = 100 Hz
12	HI FREQ	g_{fg}	Common-Gate Forward Transconductance		15			15			15		mmho	V_{DS} = 10 V, I_D = 10 mA	f = 105 MHz
13					14			14			14				f = 450 MHz
14		g_{og}	Common-Gate Output Conductance		0.18			0.18			0.18				f = 105 MHz
15					0.32			0.32			0.32				f = 450 MHz
16		G_{pg}	Common-Gate Power Gain (Note 2)	14	16		14	16		14	16		dB		f = 105 MHz
17				10	11		10	11		10	11				f = 450 MHz
18		NF	Noise Figure		1.5	2.0		1.5	2.0		1.5	2.0			f = 105 MHz
19					2.7	3.5		2.7	3.5		2.7	3.5			f = 450 MHz

NOTES:
1. Pulse test duration = 2 ms.
2. Gain (G_{pg}) measured at optimum input noise match.

NZA



n-channel JFET
designed for . . .

- VHF/UHF Amplifiers
- Front End High Sensitivity Amplifiers
- Oscillators
- Mixers

BENEFITS

- Industry Standard
- High Power Gain
  16 dB at 100 MHz, Common Gate
  11 dB at 450 MHz, Common Gate

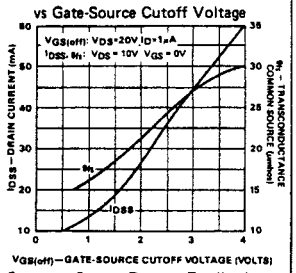
Siliconix

BENEFITS

- Low Noise
  3 dB Noise Figure at 450 MHz
- Wide Dynamic Range
  Greater Than 100 dB
- 75 Ohm Input Match Common Gate

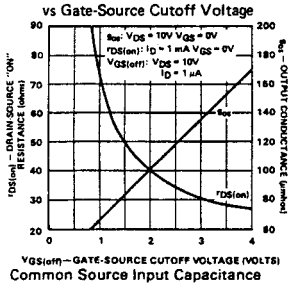
TYPE	PACKAGE	PRINCIPAL DEVICES
Single	TO-52	U308-10
Single	TO-72	U311
Single	TO-92	J308-10
Dual	TO-99	U430-1
Single	Chip	J308CHP-10CHP, U308CHP-10CHP, U311CHP
Dual	Chip	U430CHP-1CHP

Drain Current & Transconductance vs Gate-Source Cutoff Voltage

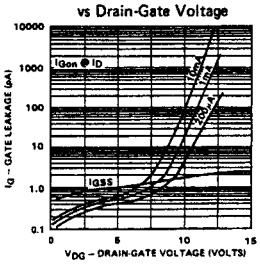


PERFORMANCE CURVES (25°C unless otherwise noted)

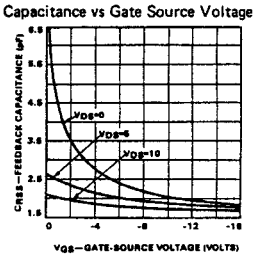
On Resistance & Output Conductance vs Gate-Source Cutoff Voltage



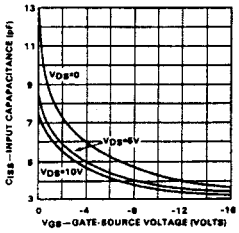
Gate Operating Current vs Drain-Gate Voltage



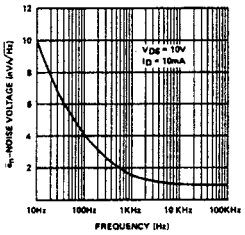
Common Source Reverse Feedback Capacitance vs Gate Source Voltage



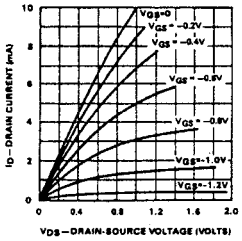
Common Source Input Capacitance vs Gate-Source Voltage



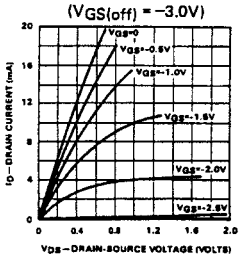
Noise Voltage vs Frequency



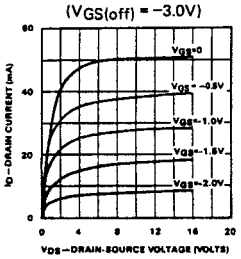
Output Characteristic ($V_{GS(off)} = -1.7V$)



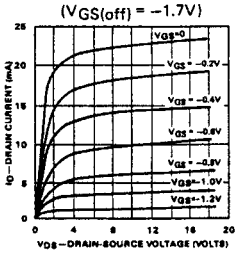
Output Characteristic ($V_{GS(off)} = -3.0V$)



Output Characteristic ($V_{GS(off)} = -3.0V$)



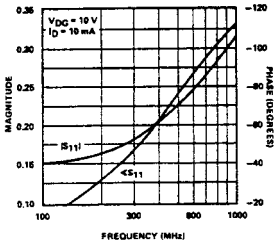
Output Characteristic ($V_{GS(off)} = -1.7V$)



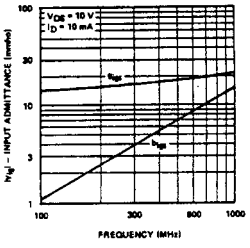
NZA

PERFORMANCE CURVES (Con't) (25°C unless otherwise noted)

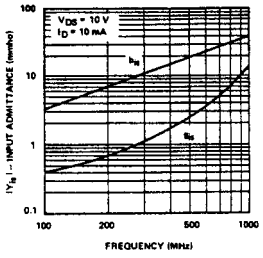
Forward Reflection Coefficient Common Gate



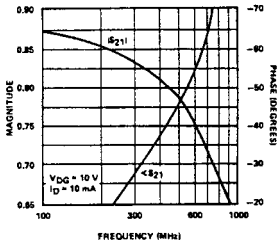
Input Admittance Common Gate



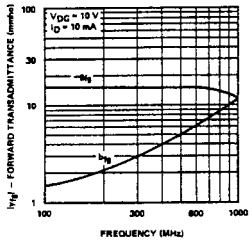
Input Admittance Common Source



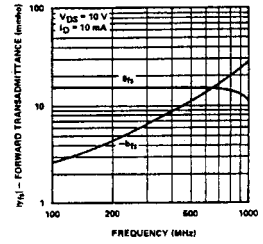
Forward Transmission Coefficient Common Gate



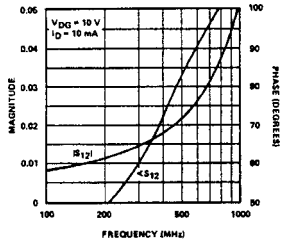
Forward Transfer Admittance Common Gate



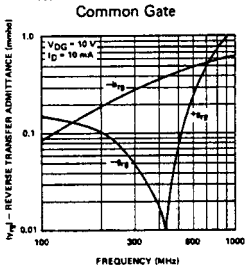
Forward Transfer Admittance Common Source



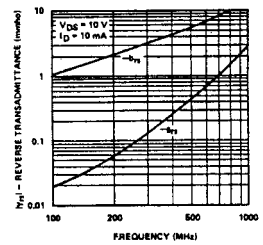
Reverse Transmission Coefficient Common Gate



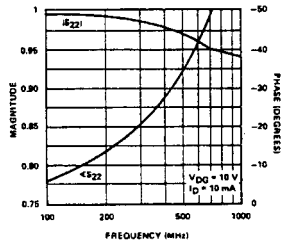
Reverse Transfer Admittance Common Gate



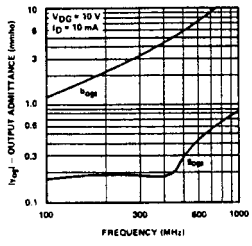
Reverse Transfer Admittance Common Source



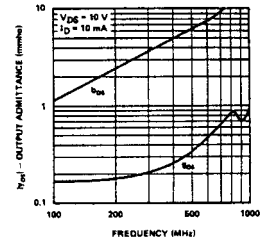
Reverse Reflection Coefficient Common Gate



Output Admittance Common Gate



Output Admittance Common Source



BROADBAND RF FET
SILICON EPITAXIAL JUNCTION N-CHANNEL FIELD EFFECT TRANSISTOR

HIGH DYNAMIC RANGE HF AND VHF AMPLIFIER
FOR USE IN COMMON GATE CONFIGURATION

- USABLE TO OVER 300 MHz
- 50 Ohm VSWR < 1.5:1 0.5-50 MHz (FIG. 1)
- LOW NOISE FIGURE — 2.2 dB TYPICAL @ 50 MHz
- INPUT Z CONSTANT 0.5-50 MHz
- HIGH IM INTERCEPT POINT — > + 40 dBm
- HIGH TRANSCONDUCTANCE — 100,000 μmhos (TYP.)
- 1 dB COMPRESSION POINT > + 20 dBm
- DYNAMIC RANGE > 140 dB (TO 1 dB COMPRESSION)
- HIGH VOLTAGE—TO 50 V.

CP640
CP664
CP665
CP666

ELECTRICAL DATA ABSOLUTE MAXIMUM RATINGS

PARAMETER	SYMBOL	CP 640	CP 664	CP 665	CP 666	UNITS
Drain to Source Voltage	BVDSO	20	30	40	50	Volts
Drain to Gate Voltage	BVDGO	20	30	40	50	Volts
Gate to Source Voltage	BVGSO	−15	−20	−20	−20	Volts
Peak Drain Current	ID	1.2	1.2	1.2	1.2	Amps
Power Dissipation 25 °C CASE	PD	8.0	8.0	8.0	8.0	Watts
Derating Factor (slope)	DF	22	22	22	22	°C/W
Junction Temp.(Oper. & Store)	TJ	−55 °C to +200 °C				

TYPICAL TWO TONE 3rd ORDER IM PRODUCTS — CIRCUIT FIGURE 1

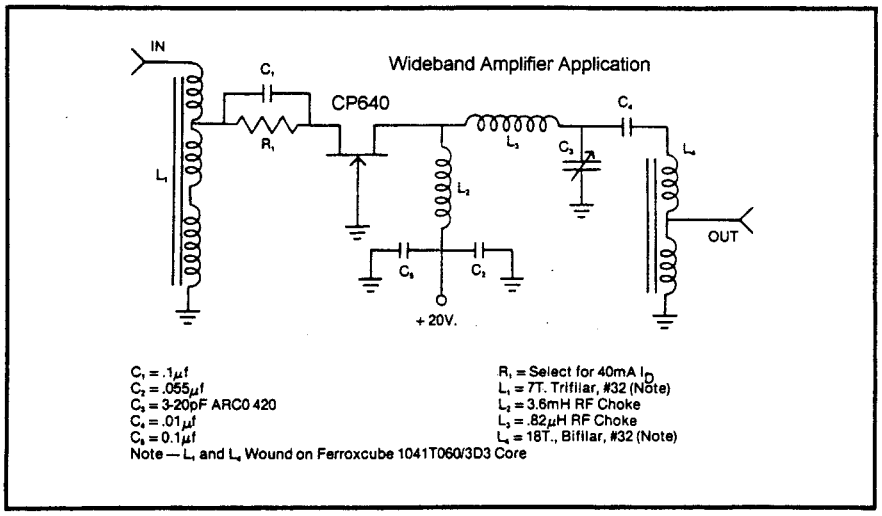
Tones at 3MHz/5MHz

Signal Level EMF	3rd Order Product
1 Volt	-44 dB
0.3 Volt	-75 dB
0.25 Volt (0dBM)	-80 dB

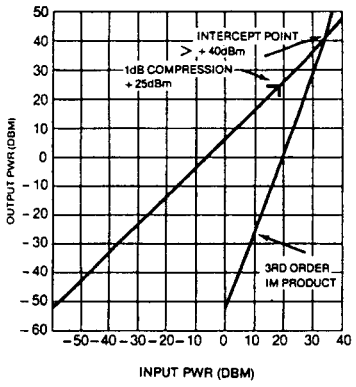
ELECTRICAL CHARACTERISTICS: T_{CASE} = 25 °C (UNLESS OTHERWISE STATED)

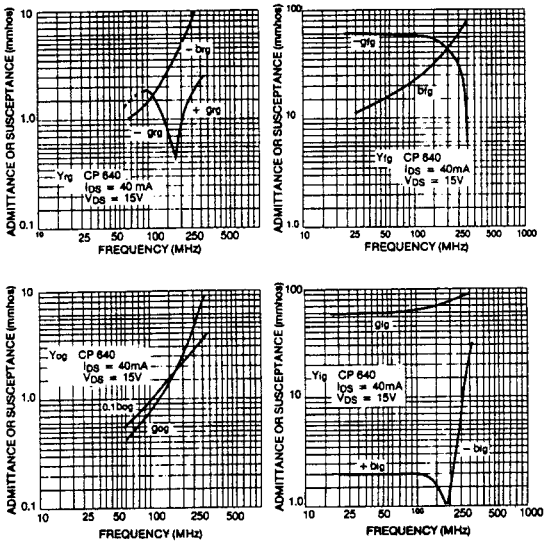
PARAMETERS	CONDITIONS		SYMBOL	Min.	Typ.	Max.	UNITS
Gate Leakage Current	V_{GS} = 15V, V_{DS} = 0	25 °C	I_{GSS}		5	100	nA
		150 °C	I_{GSS}			10	μA
Operating Transconductance	V_{DS} = 15V, I_{DS} = 40 mA		g_{fg}	40	60	80	mmho
Zero Bias Transconductance	V_{DS} = 15V, V_{GS} = 0(1)		g_{fg}	75	100	200	mmho
Gate-Source Cut-Off Voltage	V_{DS} = 5V, I_{DS} = 1.0 mA		$V_{GS(off)}$	2	5	10	Volts
Zero Bias Drain Current	V_{DS} = 15V, V_{GS} = 0 (1)		I_{DSS}	100	400	800	mA
Gate to Source Cap.	V_{GS} = −20V		C_{GS}		15	20	pf
Gate to Drain Cap.	V_{GD} = −20V		C_{GD}		15	20	pf
Power Gain	I_{DS} = 40mA, f = 50MHz, Fig. 1		Gpg	8	8.5	9.5	dB
Noise Figure	I_{DS} = 40mA, f = 30MHz, Fig. 1		N.F.		2.2	3.0	dB
Voltage Standing Wave Ratio	f = 0.5-50MHz, 50 Ω Source, Fig. 1		VSWR			1.5:1	
Common Gate Input Conductance	f = 0.5-50MHz, V_{DS} = 15, I_D = 40mA		g_{igs}		60		mmho
Common Gate Output Conductance	f = −50MHz, V_{DS} = 15, I_D = 40mA		g_{ogs}		0.4		mmho

(1)Pulse Measurement 1% Duty Cycle 10 mS Max.



TYPICAL INTERCEPT AND COMPRESSION POINT





2-3-1 Large-Signal Behavior of JFETs

A section view of a junction field-effect transistor (JFET) is shown in Figure 2-95. This structure contains, between the source and the drain contacts, an *n*-type "channel" embedded in a *p*-type silicon substrate. If it is assumed that the *pn* junction forms a barrier to current flow, then it can be seen that channel conduction is a function of the channel width, length, and thickness and of the density and mobility of the carriers. In this structure, current can flow equally well in either direction through the channel; that is, the drain can be positive or negative with respect to the source.

Since we have found that the SPICE models for JFETs are really incomplete and give poor answers, we will only briefly touch here on the JFET's performance. The modified Materka model, one of the best GaAsFET models, turns out to be best suited for JFETs because both transistors act similarly in the dc area. Both have a gate-to-source diode that becomes conductive above 0.7 V at the input and they are also quite similar in other respects. Their equivalent circuits are also quite similar, as parameter extractions for both approaches have shown.

In general, in a three-terminal device, the drain current I_D is a function of two variables: V_{DS} and V_{GS}. This function is best represented by families of characteristic curves, as shown in Figure 2-96. These curves are for the "common source" configuration, with the drain as the output and the gate as the input. They reveal that for this device, if V_{DS} is greater than about 2 V (but less than the drain breakdown), I_D is primarily determined by the gate voltage V_{GS}. Under these conditions it is valid for small signals to characterize the FET by a single transfer characteristic curve, commonly called the "forward transconductance curve," such as the upper curve shown in Figure 2-96b. Some of the relationships between the forward transfer curve and the output characteristic curves of Figure 2-96a will be examined. The value of V_{GS} that reduces I_D to approximately zero is the gate–source cutoff voltage, $V_{GS(\text{off})}$. With reference to the output curve, Figure 2-96a, note that the drain current at $V_{GS} = 0$ tends to become saturated at a drain voltage approximately equal in magnitude to $-V_{GS(\text{off})}$. This drain voltage is often referred to as the pinchoff voltage V_P; however, in this section pinchoff voltage is used interchangeably with gate–source cutoff voltage. V_P will have the same meaning as $V_{GS(\text{off})}$. The symbol I_{DSS} is commonly used to indicate the value of saturated drain current at $V_{GS} = 0$.

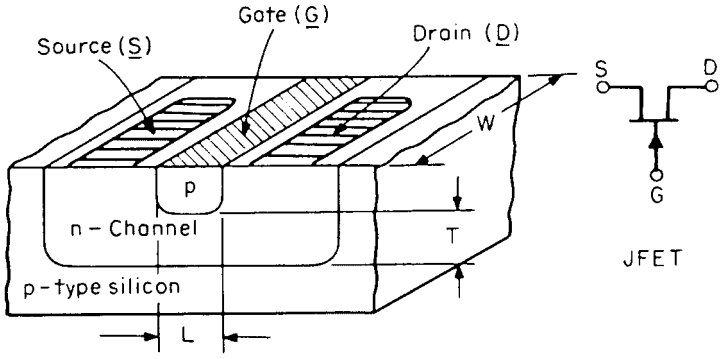


Figure 2-95 Junction field-effect transistor [9].

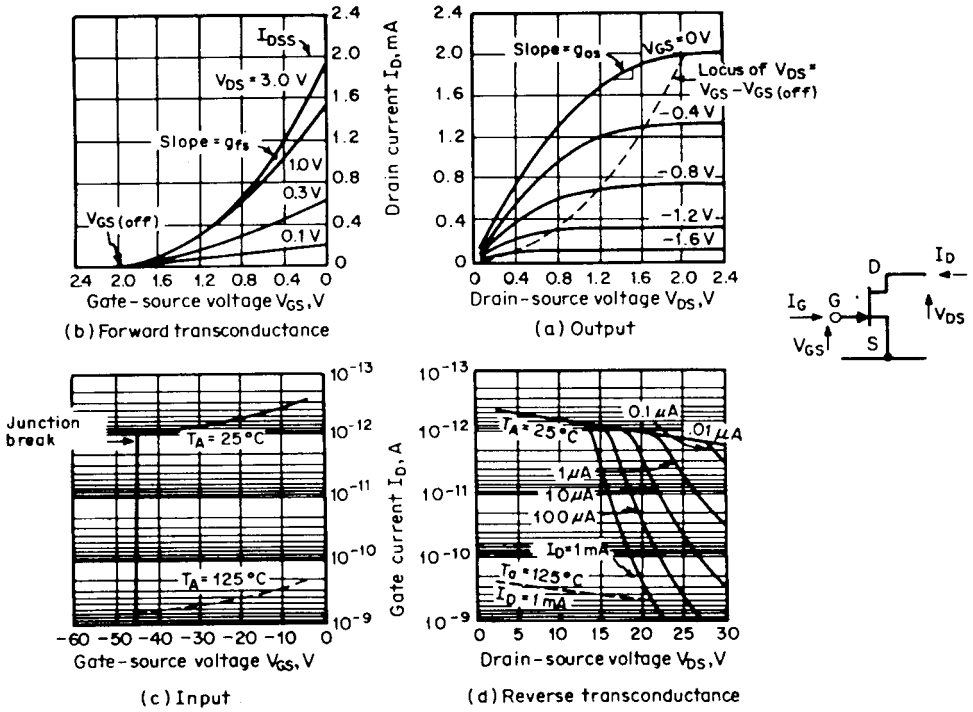


Figure 2-96 Static characteristics of an *n*-channel JFET.

The forward transconductance characteristic of Figure 2-96b can be approximated by a power law relation expressed as

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right)^n \tag{2-190}$$

if $V_{DS} \geq -V_{GS(\text{off})}$. By differentiation the small-signal transconductance g_{fs} is given by

$$g_{fs} = \frac{dI_D}{dV_{GS}} = -n\frac{I_{DSS}}{V_{GS(\text{off})}}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right)^{n-1} \tag{2-191}$$

Some texts indicate a value of $\frac{3}{2}$ for n; however, experimental measurements on a number of n-channel JFET geometries indicate that the exponent n is close to 2, which is the value derived in an approximate treatment by R. D. Middlebrook [10].

A useful relationship between g_{fso}, I_{DSS}, and $V_{GS(\text{off})}$ is derived from the ratio of Eqs. (2-190) and (2-191):

$$\frac{g_{gs}}{I_D} = n(V_{GS} - V_{GS(\text{off})})^{-1} \tag{2-192}$$

At $V_{GS} = 0$, $I_D = I_{DSS}$ and $g_{fs} = g_{fso}$. Using 2 as the value of the constant n leads to

$$g_{fso} = -2\frac{I_{DSS}}{V_{GS(\text{off})}} \tag{2-193}$$

For n-channel FETs I_{DSS} is positive and $V_{GS(\text{off})}$ is negative; for p-channel FETs I_{DSS} is negative and $V_{GS(\text{off})}$ is positive; thus g_{fs} is a positive quantity for both p- and n-channel FETs.

Equation (2-190) indicates that for $V_{GS} = V_{GS(\text{off})}$, $I_D = 0$. In a real device this does not happen. Starting from zero, as the gate voltage is made more negative, the drain current decreases until it reaches a very low value equal to the drain-leakage current. At this value, the source current will consist of source–gate leakage. Any further increase in the magnitude of the negative gate voltage will result in an increase in I_D leakage. For the small-signal device illustrated, this minimum I_D is on the order of 2×10^{-13} A.

For some types of applications of the FET, it is helpful to understand the characteristics at very low values of V_{DS}, such as those shown in Figure 2-97. A "very low value" is one that is small compared to the magnitude of $V_{GS} - V_{GS(\text{off})}$. In this region, V_{DS} is small enough to have little effect on channel thickness, so that the I_D/V_{DS} slope is nearly linear. Since the slope is a function of V_{GS}, the FET can be utilized as a voltage-controlled resistor. The conductance slope ($\Delta I_D/V_{DS}$) at $V_{DS} = 0$ is approximately a linear function of $V_{GS} - V_{GS(\text{off})}$.

If g_{ds} at $V_{GS} = 0$ is given the term g_{dso}, then

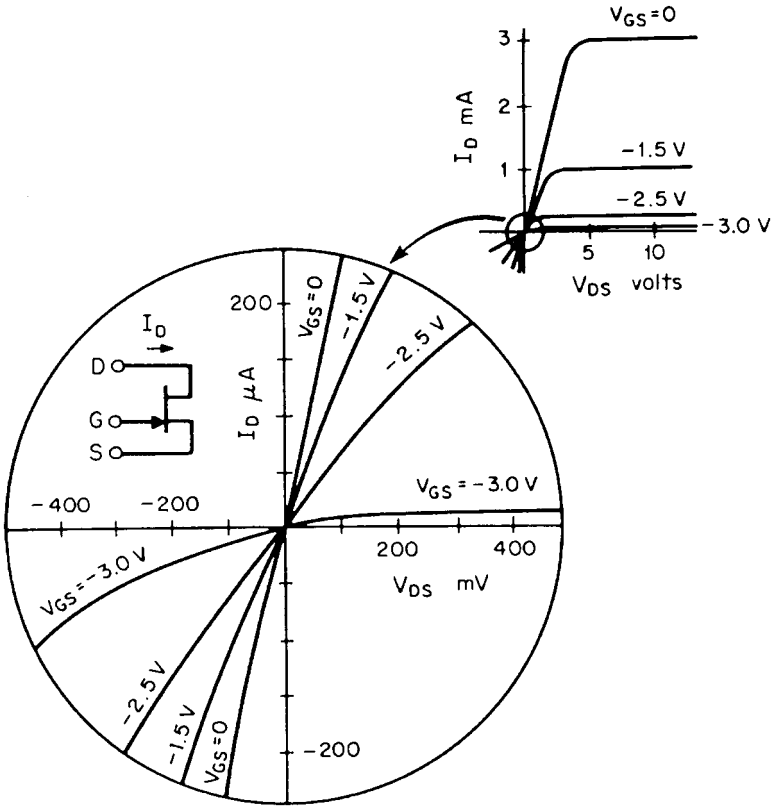


Figure 2-97 Enlargement of n-channel output characteristic around $V_{DS} = 0$.

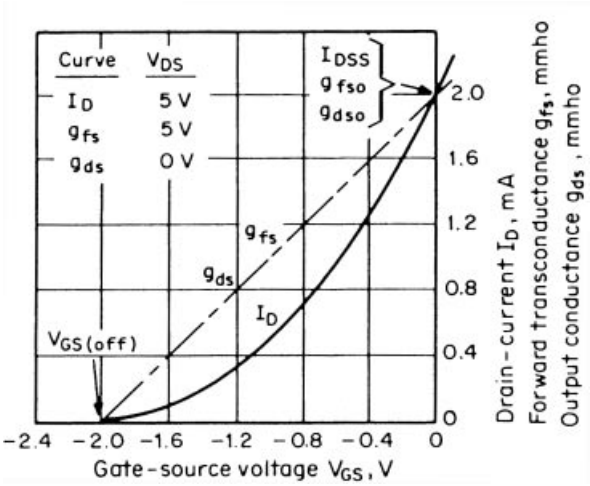


Figure 2-98 JFET I_D, g_{fs}, and g_{ds}.

$$g_{ds} = g_{dso}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right) \tag{2-194}$$

with $V_{DS} = 0$. A plot of this characteristic is shown in Figure 2-98, along with g_{fs} and I_D characteristics.

The relationship between g_{dso}, $V_{GS(\text{off})}$, and I_{DSS} is given by the equation

$$g_{dso} = -2\frac{I_{DSS}}{V_{GS(\text{off})}} \tag{2-195}$$

where I_{DSS} and $V_{GS(\text{off})}$ are as indicated in Figure 2-96. It is important to note that Eqs. (2-194) and (2-195) and the g_{ds} curve of Figure 2-98 are valid only for the case where V_{DS} is very small compared to V_P. Drain–source conductance at higher values of V_{DS} will be discussed later. FETs designed to be used as voltage-controlled resistors typically have a high $V_{GS(\text{off})}$ because the $V_{DS}/(V_{GS} - V_{GS(\text{off})})$ ratio should be low to keep distortion low. The actual large-signal JFET model as used in most simulators is shown in Figure 2-99, including a list of its intrinsic keywords (Table 2-14). Again, we had very little luck using this model as published by several SPICE CAD software manufacturers as offered. While we cannot judge on the quality of the parameters that came with the software, it was not possible to closely match even optimized sets of parameters to this large-signal model and obtain acceptable results. However, since the modified Materka model worked very well for this, we recommend that this popular JFET model not be used at frequencies above 10 MHz.

2-3-2 Small-Signal Behavior of JFETs

Figure 2-100 shows the small-signal equivalent transistor model we would recommend. It consists of an empirical FET model and a package model. Table 2-15 lists its keywords. Its advantage is that its built-in noise model is very accurate, even for JFETs and metal semiconductor FETs.

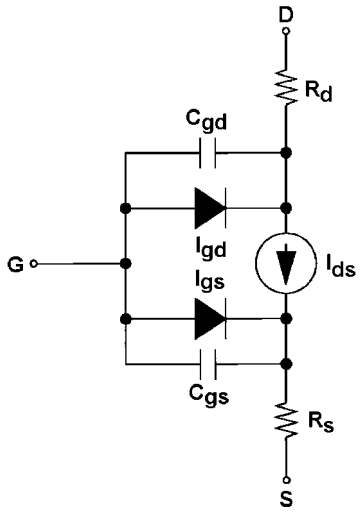


Figure 2-99 The nonlinear JFET model. Table 2-14 lists its intrinsic and extrinsic parameters. In our opinion, this model is limited to frequencies below 10 MHz; above this frequency, the modified Materka model should replace it.

Table 2-14 Large-signal JFET model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
NJF/PJF	Channel-type selection (NJF = *n*-channel, PJF = *p*-channel)		NJF
VT0	Threshold voltage	volt	–2.0
BETA	Transconductance coefficient	ampere/volt^2	1.0E–4
LAMB	Channel-length modulation	/volt	0.0
IS	Gate-junction saturation current	ampere	1.0E–14
PB	Gate-junction potential	volt	1.0
FC	Forward-bias depletion capacitance coefficient		0.5
CGS	Zero-bias gate–source junction capacitance	farad	0.0
CGD	Zero-bias gate–drain junction capacitance	farad	0.0
T	Channel transit-time delay	second	0.0
KF	Flicker noise coefficient		0.0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
SN	Switch to turn device shot noise on (1) or off (0)		1
AREA	Area multiplier		1.0
NOIS	Reference label to a set of noise data		
NAME	Required user-specified name up to 8 characters		
	Extrinsic Model		
RD	Drain ohmic resistance	ohm	0.0
RS	Source ohmic resistance	ohm	0.0

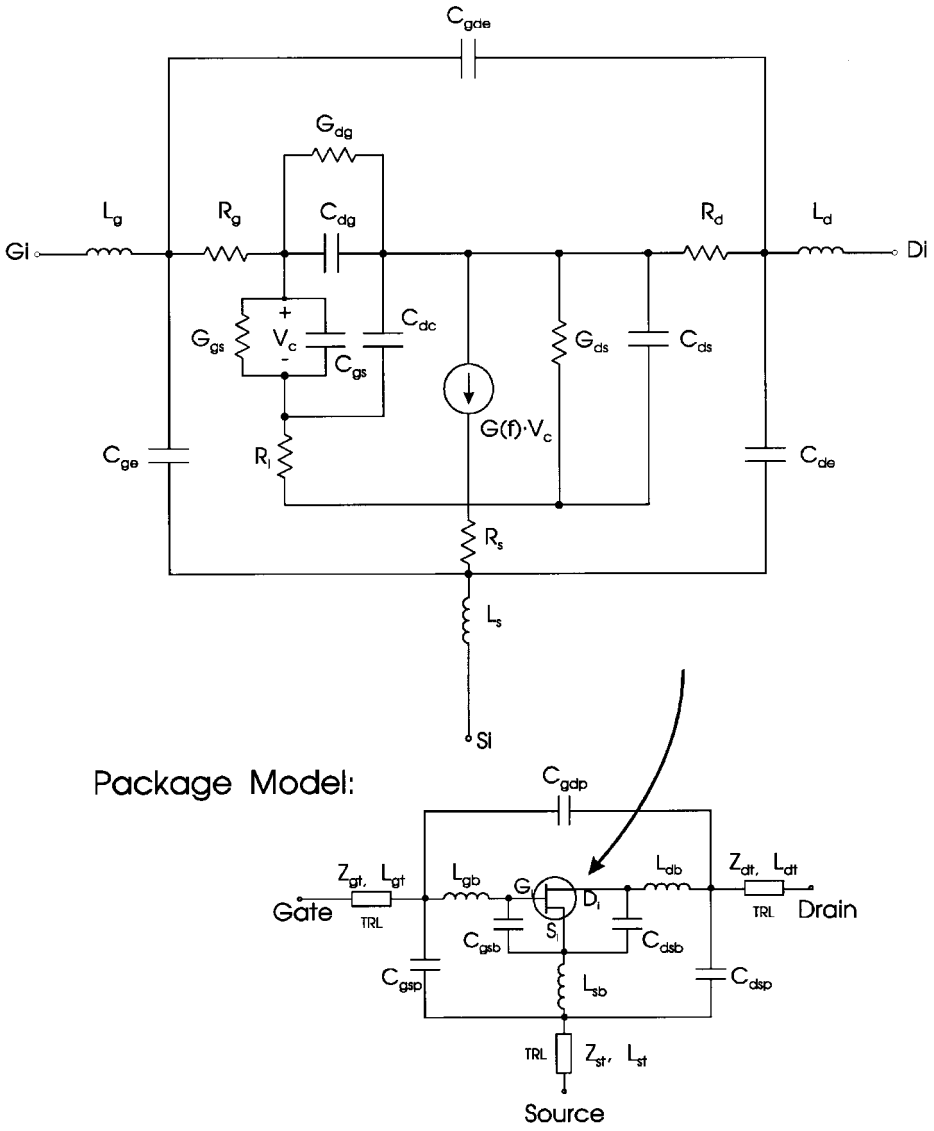


Figure 2-100 Linear FET model.

Electrical noise generated within a JFET is usually represented by equivalent noise sources, $\overline{e}_n$ and $\overline{i}_n$. Both noise voltage $\overline{e}_n$ and noise current $\overline{i}_n$ are frequency dependent and have the characteristics shown in Figure 2-101.

An equivalent noise circuit is shown in Figure 2-102. Above the frequency f_1, $\overline{e}_n$ is approximately given by

$$\overline{e}_n \cong \left(4KTB\,\frac{0.67}{g_{fs}}\right)^{1/2} \tag{2-196}$$

Table 2-15 Small-signal FET model keywords

Keyword	Description	Unit	Default
	Intrinsic model		
G	Transconductance at dc, G_0[a]	/ohm	
CGS	Gate–source capacitance	farad	
F	3-dB roll-off frequency	hertz	∞
T	Time delay	second	0.0
TDS	Drain–source time delay	second	0.0
GGS	Gate–source conductance	/ohm	0.0
CDG	Drain–gate capacitance	farad	0.0
CDC	Dipole layer capacitance	farad	0.0
CDS	Drain–source capacitance	farad	0.0
GDS	Drain–source conductance	/ohm	0.0
RI	Channel resistance	ohm	0.0
RG	Gate resistance	ohm	0.0
RD	Drain resistance	ohm	0.0
RS	Source resistance	ohm	0.0
CGE	External gate capacitance	farad	0.0
CDE	External drain capacitance	farad	0.0
LG	Gate-lead inductance	henry	0.0
LD	Drain-lead inductance	henry	0.0
LS	Source-lead inductance	henry	0.0
CGDE	External gate–drain capacitance	farad	0.0
GDG	Gate–drain conductance	/ohm	0.0
TJ	Chip temperature	kelvin	298
	Package Parasitics		
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate to source package capacitance	farad	0.0
CDSP	Drain to source package capacitance	farad	0.0
CGDP	Gate to drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for at $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for $\varepsilon_r = 1$	meter	0.0
FC	Corner frequency of flicker (1/*f*) noise[b]	hertz	10 MHz
FCP	Shape factor of the 1/*f* noise response		1.0
label	User-defined term that refers to temperature coefficient		

[a]The transconductance of this model may be approximately described by $g_m = G\,\frac{e^{-j\omega T}}{1 + j(f/F)}$ where $\omega = 2\pi f$ and f is frequency.
[b]The flicker noise frequency dependence is given by $1/(f/F_c)^{\mathrm{FCP}}$.

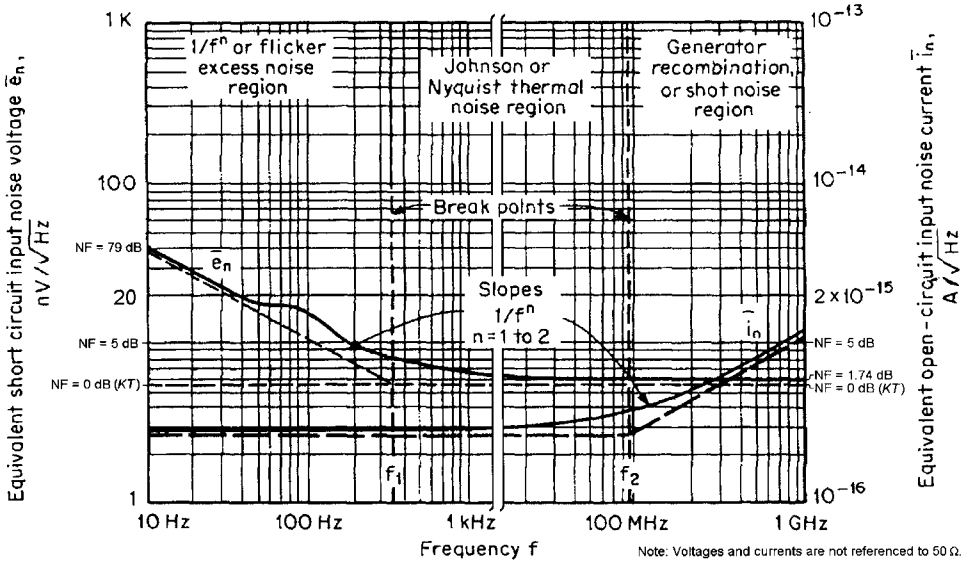


Figure 2-101 JFET noise characteristics.

where $K = 1.374 \times 10^{-23}$ J/K

T = absolute temperature in kelvins (273 K = 0 °C)

B = frequency range in hertz

g_{fs} = transconductance of FET

With the input short-circuited, the noise voltage across the load R_L resulting from the FET is

$$\text{Output noise voltage} = \bar{e}_n A_V \tag{2-197}$$

Below the frequency f_1, $\bar{e}_n$ increases proportional to $1/f^n$ and is expressed as

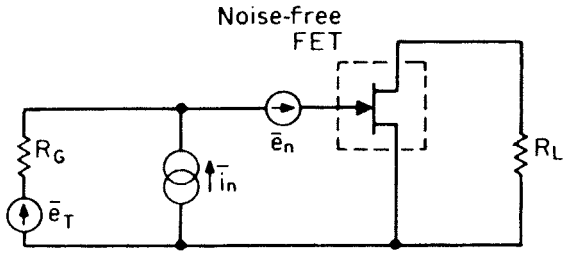


Figure 2-102 Equivalent noise circuit of the FET.

$$\bar{e}_n = \left[4KTB \left(\frac{0.67}{g_{fs}} \right) \left(1 + \frac{f_1}{f^n} \right)^{1/2} \right] \tag{2-198}$$

The low-frequency corner frequency f_1 for JFETs is typically in the 100-Hz to 1-kHz range, and the exponent n is usually between 1 and 2. As indicated by the equations, $\bar{e}_n$ is inversely proportional to the square root of g_{fs}.

The equivalent input noise current $\bar{i}_n$ is caused by the current in the gate-to-channel junction. Its approximate value below f_2 is

$$\bar{i}_n = (2qI_G B)^{1/2} \tag{2-199}$$

where $q = 1.602 \times 10^{-19}$
B = frequency range in hertz
I_G = dc gate current

This expression is fairly accurate when I_G is the result of the active device conductance. Typically, $\bar{i}_n$ will be lower than the calculated value because part of I_G is due to conductance across the device package.

At higher frequencies (above f_2),

$$\bar{i}_n = \left(\frac{4KTB}{R_P} \right)^{1/2} \tag{2-200}$$

where R_P is the real part of gate input impedance. R_P, in terms of Y_{11}, can range from several tens of megohms (at audiofrequencies) to 1 kΩ or less (at VHF/UHF frequencies). The high-frequency corner f_2 is typically in the range of 5–50 kHz and more than 100 MHz for high-frequency devices.

Another form of noise is known as "popcorn" or burst noise, the causes of which have not been completely identified. It shows up as a random short-duration step-function change in drain current, equivalent to an input gate–source voltage change of a few tenths of a microvolt. It is not unlike big bubbles on the surface of boiling water.

As far as the noise is concerned, we can also state that the Materka built-in proprietary noise model is significantly more accurate than the one mentioned above [11–14]. This model is implemented, with minimal documentation of its inner workings because of its proprietary nature, in the linear portion of the circuit simulator in Ansoft's Serenade Design Environment. In many cases, the manufacturer supplies only limited data, so it is most convenient that after SPICE-type parameter extraction is done by using the Scout program, one even gets good noise prediction for JFETs using the modified Materka model.

2-3-3 Large-Signal Behavior of MOSFETs

Metal-oxide-semiconductor field-effect transistors (MOSFETs) are important components in contemporary analog integrated circuits. Whereas initial applications were centered on all-MOS processes, combined bipolar and MOS processes give the designer the best of both worlds. A major advantage of MOS processes for realizing analog functions is that complex, dense digital functions can be realized on the same chip.

Transfer Characteristics of MOS Devices. A cross section of a typical enhancement-mode *n*-channel MOS transistor (NMOS) is shown in Figure 2-103a. Heavily doped *n*-type source and drain regions are fabricated in a *p*-type substrate (often called the body). A thick layer of silicon dioxide is grown over the substrate material and a conductive gate material (metal or polycrystalline silicon) covers this oxide between source and drain. The operation of the device is very similar to that of a JFET in that the gate–source voltage is used to modify the conductance of the region under the gate. This allows the gate voltage to control the current flowing between source and drain, giving gain in analog circuits and switching characteristics in digital circuits.

The enhancement-mode NMOS device of Figure 2-103a shows significant conduction between source and drain only when an *n*-type channel exists under the gate, and this is the origin of the "*n*-channel" designation. The term "enhancement mode" refers to the fact that no conduction occurs for $V_{GS} = 0$, and thus the channel must be "enhanced" to cause conduction. MOS devices can be made equally well by using an *n*-type substrate with a *p*-type conducting channel. Such devices are called enhancement-mode *p*-channel MOS transistors (PMOS). In technologies employing one or the other of these device types, the circuit symbol of Figure 2-103b is commonly used for either one. In complementary MOS technology (CMOS) both device types are present and the circuit symbols of Figure 2-103c will be used to distinguish them. In PMOS and NMOS technologies, the substrate is common to all devices, invariably connected to a dc power supply voltage, and usually not shown on the circuit diagram. In CMOS technology, however, devices of one type or another are fabricated in individual, separate isolation regions, which may or may not be connected to a power supply voltage. If these isolation regions *are* connected to the appropriate power supply, the symbols of Figure 2-103c will be used and the substrate connection will not be shown. If the individual isolation regions are connected elsewhere, however, the devices will be represented by the symbols of Figure 2-103d, where the substrate is labeled *B*. Finally, in NMOS technology an additional device type called a "depletion-mode" device is usually available. This is a conducting channel implanted between the source and drain so that conduction occurs for $V_{GS} = 0$. This device has characteristics that are almost identical to that of a JFET and will be represented by the symbol of Figure 2-103e.

The derivation of the transfer characteristics of the enhancement-mode NMOS device of Figure 2-103a begins by noting that with $V_{GS} = 0$, the source and drain regions are separated by back-to-back *pn* junctions. These junctions are formed between the *n*-type source and drain regions and the *p*-type substrate, resulting in an extremely high resistance (about 10^{12} Ω) between drain and source when the device is off.

Now consider substrate, source, and drain grounded and a positive voltage V_{GS} applied to the gate as shown in Figure 2-104. The gate and substrate then form the plates of a capacitor with the SiO_2 as a dielectric. Positive charge accumulates on the gate and the negative charge in the substrate. Initially, the negative charge in the *p*-type substrate is manifested by creation of a *depletion region* and resulting exclusion of holes under the gate. This is shown in Figure 2-104. The depletion-layer width *X* under the oxide is

$$X = \left(\frac{2\varepsilon\phi}{qN_A} \right)^{1/2} \tag{2-201}$$

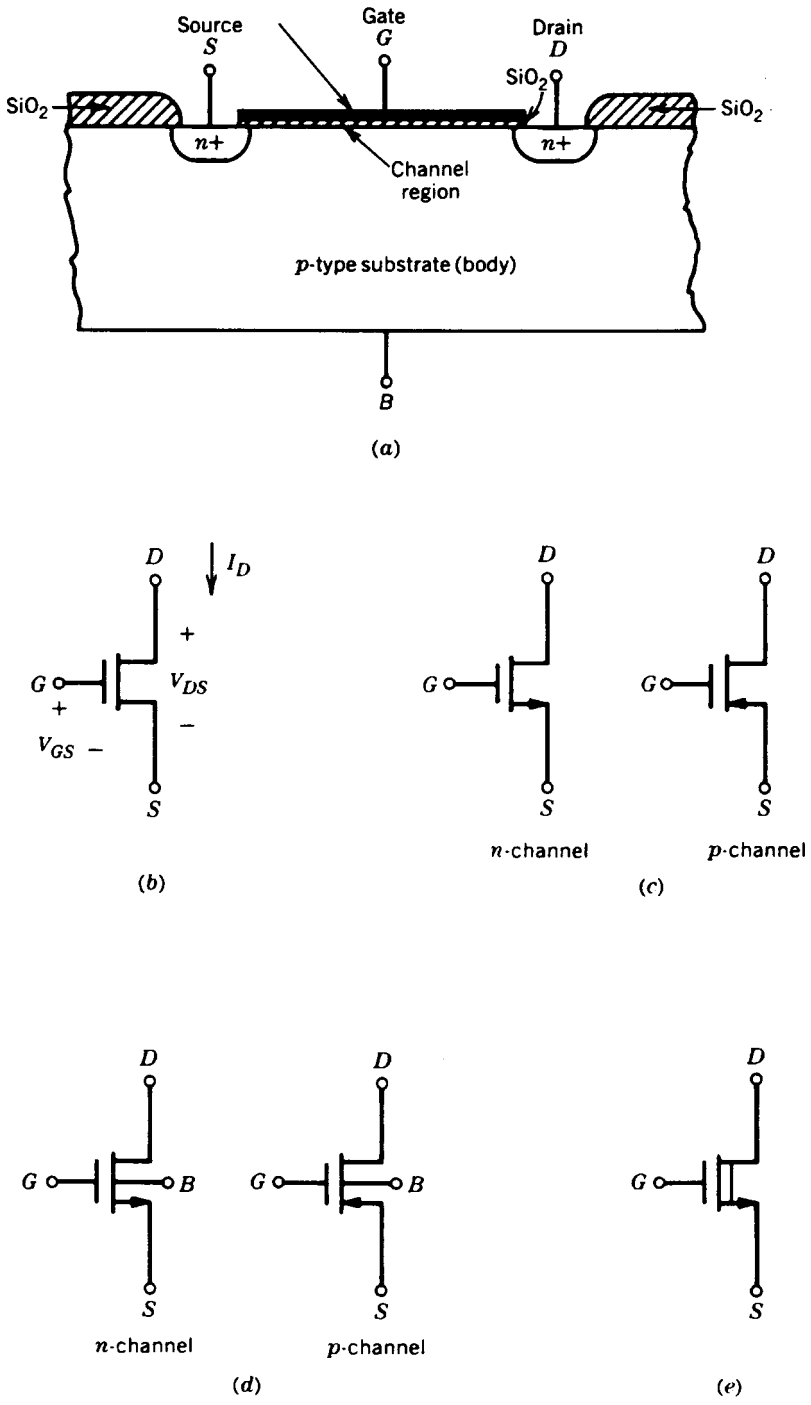


Figure 2-103 (a) Typical enhancement-mode NMOS structure. (b) Enhancement-mode NMOS or PMOS circuit symbol when one device type only is present. (c) NMOS and PMOS symbols used in CMOS circuits. (d) NMOS and PMOS symbols used when the substrate connection is nonstandard. (e) Depletion MOS device symbol.

where ϕ is the potential in the depletion layer at the oxide–silicon interface, N_A atoms/cm^3 is the doping density (assumed constant) of the p-type substrate, and ε is the permittivity of the silicon. The charge per unit area in this depletion region is

$$Q = qN_A X = \sqrt{2qN_A \varepsilon \phi} \tag{2-202}$$

When the potential in the silicon reaches a critical value equal to twice the Fermi level, $\phi_f \simeq 0.3$ V, a phenomenon known as "inversion" occurs [15]. Further increases in gate voltage produce no further changes in the depletion-layer width but instead a thin layer of electronics is induced in the depletion layer directly under the oxide. This produces a continuous n-type region with the source and drain regions and is the conducting channel between source and drain. This channel can then be modulated by increases or decreases in the gate voltage. In the presence of an inversion layer, and with no substrate bias, the depletion region contains a fixed charge

$$Q_{b0} = \sqrt{2qN_A \varepsilon 2\phi_f} \tag{2-203}$$

If a substrate bias voltage V_{SB} (source positive for n-channel devices) is applied between source and substrate, the potential required to produce inversion becomes $(2\phi_f + V_{SB})$ and the charge stored in the depletion region in general is

$$Q_b = \sqrt{2qN_A \varepsilon (2\phi_f + V_{SB})} \tag{2-204}$$

The gate voltage V_{GS}, required to produce an inversion layer, is called the threshold voltage V_t and can now be calculated. This voltage consists of several components. First, a voltage $[2\phi_f + (Q_b/C_{ox})]$ is required to sustain the depletion-layer charge Q_b, where C_{ox} is the gate oxide capacitance per unit area. Second, a work-function difference ϕ_{ms} exists between the

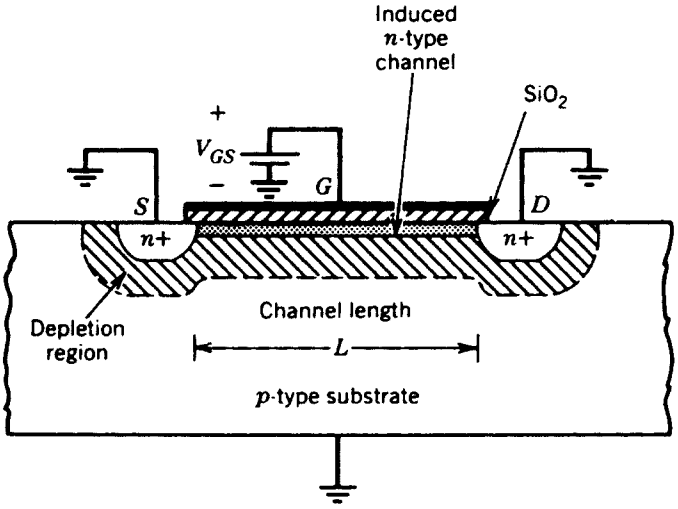


Figure 2-104 Idealized NMOS device cross section with positive V_{GS} applied, showing depletion regions and the induced channel.

gate metal and the silicon. Third, there is always charge density Q_{ss} (positive) in the oxide at the silicon interface. This is caused by crystal discontinuities at the Si–SiO_2 interface and must be compensated by a gate voltage contribution of $-Q_{ss}/C_{ox}$. Thus we have a threshold voltage

$$V_t = \phi_{ms} + 2\phi_f + \frac{Q_b}{C_{ox}} - \frac{Q_{ss}}{C_{ox}}$$

$$= \phi_{ms} + 2\phi_f + \frac{Q_{b0}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} + \frac{Q_b - Q_{b0}}{C_{ox}} \tag{2-205}$$

$$= V_{t0} + \gamma\left(\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f}\right) \tag{2-206}$$

where Eqs. (2-203) and (2-204) have been used and V_{t0} is the threshold voltage with $V_{SB} = 0$. The parameter γ is defined as

$$\gamma = \frac{1}{C_{ox}}\sqrt{2q\varepsilon N_A} \tag{2-207}$$

and

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \tag{2-208}$$

where ε_{ox} and t_{ox} are the permittivity and thickness of the oxide, respectively. A typical value of γ is 0.5 $V^{1/2}$ and $C_{ox} = 3.5 \times 10^{-4}$ pF/μ^2 for $t_{ox} = 0.1$ μm.

In practice, the value of V_{t0} is usually adjusted in processing by implanting additional impurities into the channel region. Extra *p*-type impurities are implanted in the channel to make $V_{t0} \simeq 0.5$–1.5 V for *n*-channel enhancement devices. By implanting *n*-type impurities in the channel region, a conducting channel can be formed, even for $V_{GS} = 0$. This MOS transistor is called a depletion device, with typical values of V_{t0} in the range –1 to –4 V. If Q_i is the charge density per unit area due to the implant, then the threshold voltage given by Eq. (2-205) is shifted by approximately Q_i/C_{ox}.

The preceding equations can now be used to calculate the large-signal characteristics of an NMOS transistor. For purposes of analysis the source is assumed grounded and bias voltages V_{GS}, V_{DS}, and V_{SB} are applied as shown in Figure 2-105. If V_{GS} is greater than V_t, a conducting channel exists and V_{DS} causes a larger reverse bias from drain to substrate than exists from source to substrate, and thus a wider depletion region exists at the drain. However, for simplicity, we assume that the voltage drop along the channel itself is small so that the depletion layer is constant along the channel.

At a distance *y* along the channel the voltage with respect to the source is $V(y)$ and the gate-to-channel voltage at that point is $V_{GS} - V(y)$. We assume this voltage exceeds the threshold voltage V_t, and thus the induced electron charge per unit area in the channel is

$$Q_I(y) = C_{ox}[V_{GS} - V(y) - V_t] \tag{2-209}$$

The resistance *dR* of a length *dy* of the channel is

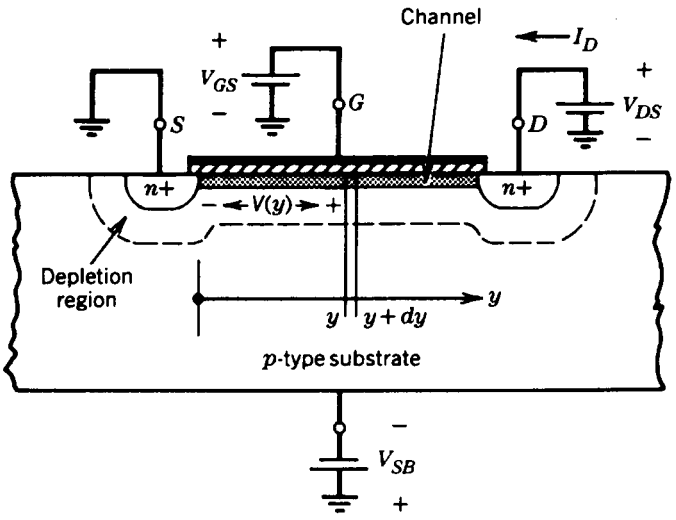


Figure 2-105 NMOS device with bias voltages applied.

$$dR = \frac{dy}{W\mu_n Q_I(y)} \tag{2-210}$$

where W is the width of the device, perpendicular to the plane of Figure 2-105, and μ_n is the average electron mobility of the channel.

The voltage drop dV along the length of the channel dy is

$$dV = I_D\, dR = \frac{I_D}{W\mu_n Q_I(y)}\, dy \tag{2-211}$$

If L is the total channel length, then substitution of Eq. (2-209) into Eq. (2-211) and integration gives

$$\int_0^L I_D\, dy = \int_0^{V_{DS}} W\mu_n C_{\text{ox}}(V_{GS} - V - V_t)\, dV \tag{2-212}$$

This results in

$$I_D = \frac{k'}{2}\frac{W}{L}[2(V_{GS} - V_t)V_{DS} - V_{DS}^2] \tag{2-213}$$

where

$$k' = \mu_n C_{\text{ox}} = \frac{\mu_n \varepsilon_{\text{ox}}}{t_{\text{ox}}} \tag{2-214}$$

Equation (2-213) is important and describes the I–V characteristics of an MOS transistor, assuming a continuous induced channel. A typical value of k' for $t_{\text{ox}} = 0.1$ μm is about 20 μA/V^2 for an n-channel device.

As the value of V_{DS} is increased, the induced conducting channel narrows at the drain end and Eq. (2-209) indicates that Q_I at the drain end approaches zero as V_{DS} approaches $(V_{GS} - V_t)$. This results in the same pinchoff phenomenon as occurs in a JFET, and further increases in V_{DS} produce little change in I_D. Equation (2-213) is thus no longer valid if V_{DS} is greater than $(V_{GS} - V_t)$. The value of I_D in this region is obtained by substituting $V_{DS} = (V_{GS} - V_t)$ in Eq. (2-213), giving

$$I_D = \frac{k'}{2}\frac{W}{L}(V_{GS} - V_t)^2 \tag{2-215}$$

for an MOS transistor in the pinchoff region. As in the case of a JFET, the drain current in the pinchoff region varies slightly as the drain voltage is varied. This is due to the presence of a depletion region between the physical pinchoff point in the channel at the drain end and the drain region itself. If this depletion-layer width is X_d, then the *effective* channel length is given by

$$L_{\text{eff}} = L - X_d \tag{2-216}$$

If L_{eff} is used in place of L in Eq. (2-215), we obtain a more accurate formula for the pinchoff region:

$$I_D = \frac{k'}{2}\frac{W}{L_{\text{eff}}}(V_{GS} - V_t)^2 \tag{2-217}$$

The fact that X_d (and thus L_{eff}) is a function of the drain–source voltage results in a variation of I_D and V_{DS} in the pinchoff region. Using Eqs. (2-216) and (2-217), we obtain

$$\frac{\partial I_D}{\partial V_{DS}} = -\frac{k'}{2}\frac{W}{L_{\text{eff}}^2}(V_{GS} - V_t)^2\frac{dL_{\text{eff}}}{dV_{DS}} \tag{2-218}$$

and thus

$$\frac{\partial I_D}{\partial V_{DS}} = \frac{I_D}{L_{\text{eff}}}\frac{dX_d}{dV_{DS}} \tag{2-219}$$

This equation is analogous to Eq. (2-168) for bipolar transistors. Following a similar procedure, a voltage analogous to the Early voltage can be defined as

$$V_A = \frac{I_D}{\partial I_D/\partial V_{DS}} \tag{2-220}$$

and thus

$$V_A = L_{\text{eff}}\left(\frac{dX_d}{dV_{DS}}\right)^{-1} \tag{2-221}$$

For MOS transistors, the most widely used parameter for the characterization of output resistance is

$$\lambda = \frac{1}{V_A} \tag{2-222}$$

As in the bipolar case, the large-signal properties of the transistor can be approximated by assuming that λ and V_A are constants, independent of bias conditions. Thus we can formulate a better approximation to the I–V characteristics as

$$I_D = \frac{k'}{2}\frac{W}{L}(V_{GS} - V_t)^2 (1 + \lambda V_{DS}) \tag{2-223}$$

In practical MOS transistors variation of X_d with voltage is complicated by the fact that the field distribution in the drain depletion region is not one dimensional. A vertical component in the field distribution is introduced by the potential difference between the gate and channel and the gate and drain. As a result, the calculation of λ from the device structure is quite different [16], and it is usually necessary to develop effective values of λ from experimental data. The parameter λ is a linear function of effective channel length and is an increasing function of the doping level in the channel. Typical values of λ are in the range 0.05–0.005 V^{-1}. Note that, in the case of a JFET, the pinchoff region for MOS devices is often called the *saturation* region.

A plot of I_D versus V_{DS} for an NMOS transistor is shown in Figure 2-106. Below pinchoff, the device behaves as a nonlinear voltage-controlled resistor, which is often called the *ohmic* or *triode* region. Above pinchoff, the device approximates a voltage-controlled current source. Note that for depletion MOS devices, V_t is negative and I_D is finite, even for $V_{GS} = 0$. For PMOS devices, all polarities of voltage and current are reversed.

The results as previously derived can be used to form a large-signal model of an MOS transistor. The model topology in the pinchoff region is the same as that for the JFET, but using Eq. (2-215) for the controlled-current generator.

MOS Device Voltage Limitations. The voltage limitations of MOS transistors depend on the gate length L. For small values of L (less than about 10 μm), the drain-depletion region

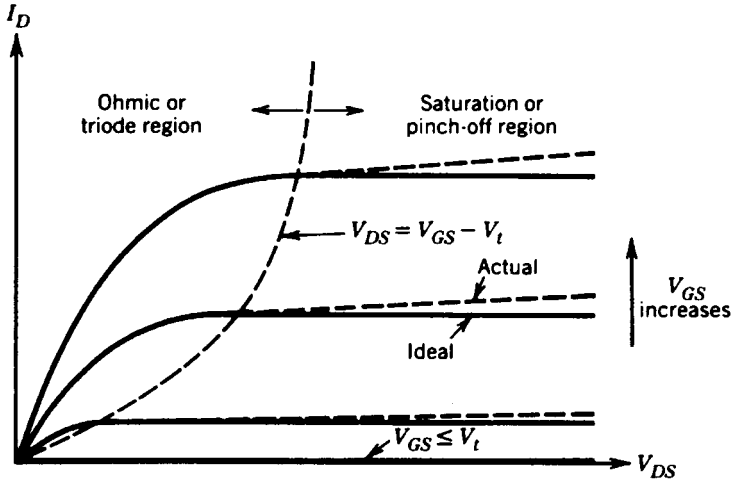


Figure 2-106 NMOS device characteristics.

exerts an appreciable influence on the channel. This causes I_D to rise with increasing V_{DS} in a similar fashion to the bipolar transistor curves of Figure 2-89.

For long channel lengths, the drain-depletion region has little effect on the channel and the I_D-versus-V_{DS} curves follow closely the ideal curves of Figure 2-106. Eventually the drain–substrate *pn*-junction breakdown voltage is exceeded and a sharp breakdown characteristic is obtained, which is similar to the JFET characteristic.

In addition to V_{DS} limitations, MOS devices must also be protected against excessive gate voltages. Typical gate oxides break down with about 25–50 V applied from gate to channel, and this process is destructive to the transistor.

2-3-4 Small-Signal Model of the MOS Transistor in Saturation

The preceding large-signal equations can now be used to derive the small-signal model of the MOS transistor in the saturation or pinchoff region. The important equations are (2-220) and (2-206). Note from Eq. (2-206) that the source–substrate voltage V_{BS} affects V_t, and thus I_D. This is due to the influence of the substrate acting as a second gate and is called *body effect*. As a consequence, I_D is a function of both V_{GS} and V_{BS}, and we require *two* transconductance generators in the small-signal model as shown in Figure 2-107. Variations in the voltage v_{bs} from source to body cause current $g_{mb}v_{bs}$ to flow from drain to source. Note that the body (or substrate) of an NMOS integrated circuit is usually connected to the most negative supply voltage and is thus an ac ground. However, the source connection can have a significant ac voltage impressed on it. Parasitic resistances due to the channel contact regions should be included in series with the source and drain of the model but are usually neglected in hand calculations. These resistances have an inverse dependence on channel width W and have typical values of 50–100 Ω for devices with W of about 1 μm.

The parameters of Figure 2-107 can be determined from Eq. (2-223) by differentiating.

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = k' \frac{W}{L} (V_{GS} - V_t)(1 + \lambda V_{DS}) \tag{2-224}$$

If $\lambda V_{DS} << 1$, this is often approximated as

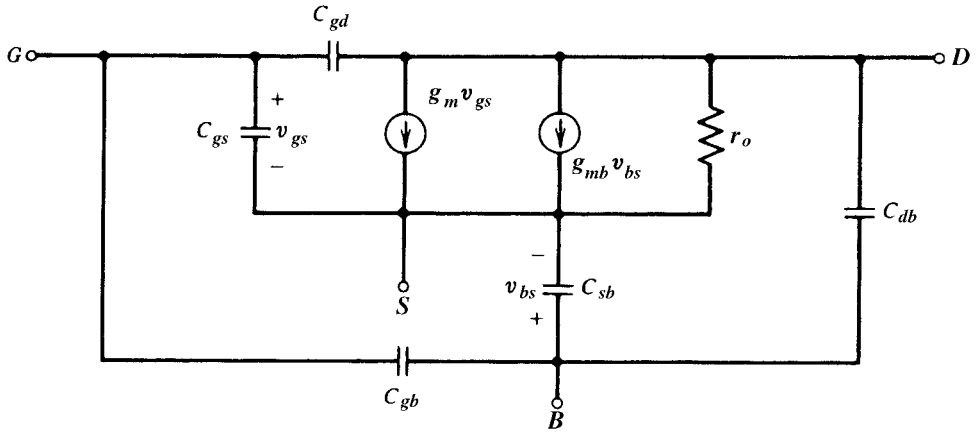


Figure 2-107 Small-signal MOS transistor equivalent circuit.

$$g_m = k' \frac{W}{L} (V_{GS} - V_t) = \sqrt{2k' \frac{W}{L} I_D} \tag{2-225}$$

Like the JFET and unlike the bipolar transistor, the transconductance of the MOSFET depends on both bias current and the *W*/*L* ratio (also on the oxide thickness via *k*′). Similarly,

$$g_{mb} = \frac{\partial I_D}{\partial V_{BS}} = -k' \frac{W}{L} (V_{GS} - V_t) (1 + \lambda V_{DS}) \frac{\partial V_t}{\partial V_{BS}} \tag{2-226}$$

From Eq. (2-206)

$$\frac{\partial V_t}{\partial V_{BS}} = -\frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} = -\chi \tag{2-227}$$

where this equation defines a factor χ. This parameter is equal to the rate of change of threshold voltage with body bias voltage. One obtains

$$\chi = \frac{C_{js}}{C_{\text{ox}}} \tag{2-228}$$

where C_{js} is the capacitance per unit area of the depletion region under the channel, assuming a one-sided step junction with a built-in potential $\psi_0 = 2\phi_f$.

Substitution of Eq. (2-227) into Eq. (2-226) gives

$$g_{mb} = \frac{\gamma k' (W/L) (V_{GS} - V_t) (1 + \lambda V_{DS})}{2\sqrt{2\phi_f + V_{SB}}} \tag{2-229}$$

Again, if $\lambda V_{DS} << 1$, we have

$$g_{mb} = \frac{\gamma \sqrt{k' (W/L) I_D}}{\sqrt{2(2\phi_f + V_{SB})}} \tag{2-230}$$

An important quantity is the ratio g_{mb}/g_m, and from Eqs. (2-224) and (2-229) we find

$$\frac{g_{mb}}{g_m} = \frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} = \chi \tag{2-231}$$

The factor χ is typically in the range 0.1–0.3.

Finally, the small-signal output resistance can be obtained directly from Eq. (2-218):

$$r_o = \left(\frac{\partial I_D}{\partial V_{DS}} \right)^{-1} = \frac{L_{\text{eff}}}{I_D} \left(\frac{dX_d}{dV_{DS}} \right)^{-1} \tag{2-232}$$

and using Eqs. (2-221) and (2-222), we find

$$r_o = \frac{1}{\lambda I_D} = \frac{V_A}{I_D} \tag{2-233}$$

Small-signal model capacitances are also shown in Figure 2-107. Of these, only the gate–source capacitance C_{gs} is intrinsic to the device operation in the saturation region. Capacitances C_{sb} and C_{db} are parasitic depletion-region capacitances between the substrate and the source and drain regions, respectively. These can be expressed as follows:

$$C_{sb} = \frac{C_{sb0}}{(1 + V_{SB}/\psi_0)^{1/2}} \tag{2-234}$$

$$C_{db} = \frac{C_{db0}}{(1 + V_{DB}/\psi_0)^{1/2}} \tag{2-235}$$

These capacitances are proportional to the gate and source region areas (including sidewalls), and C_{sb} also includes depletion-region capacitance from the induced channel in the body.

Capacitance C_{gb} between gate and substrate models parasitic oxide capacitance between the gate contact material and the substrate outside the active-device area. This is a constant capacitance and models coupling between polysilicon and metal interconnects and the underlying substrate. In fact, parasitic capacitance of this type underlies all polysilicon and metal traces on the chip and should be taken into account when simulating and calculating high-frequency circuit and device performance. Typical values depend on oxide thicknesses and range from about 0.04 to 0.15 fF per square micron of interconnect, with fringing effects becoming important for narrow lines (several microns or less in width).

Capacitances C_{gs} and C_{gd} exist from gate to source and drain, respectively. If C_{ox} is the oxide capacitance per unit area from gate to channel then the total capacitance under the gate is $C_{\text{ox}}WL$. This capacitance is intrinsic to the device operation and models the gate control of the channel conductance. In the ohmic region of device operation, this capacitance is split equally between source and drain so that $C_{gs} = C_{gd} = 1/2C_{\text{ox}}WL$. However, in the saturation region, the channel is very narrow at the drain end and the drain voltage exerts little influence on either the channel or the gate charge. As a consequence, the intrinsic portion of C_{gd} is essentially zero in the saturation region and C_{gd} then consists of a constant parasitic oxide-capacitance contribution due to gate overlap of the drain region. This is on the order of 1–10 fF for small devices.

In order to calculate the corresponding value of C_{gs} in the saturation region, we must calculate the total charge Q_T stored in the channel. This can be obtained by integrating Eq. (2-209) to obtain

$$Q_T = WC_{\text{ox}} \int_0^L [V_{GS} - V(y) - V_t]dy \tag{2-236}$$

Substituting for dy/dV from Eq. (2-211) into Eq. (2-236), we find

$$Q_T = \frac{W^2 C_{\text{ox}}^2 \mu_n}{I_D} \int_0^{V_{GS}-V_t} (V_{GS} - V - V_t)^2 dV \tag{2-237}$$

where the limit $y = L$ corresponds to $V = (V_{GS} - V_t)$ in saturation. Solution of Eq. (2-237) and use of Eqs. (2-214) and (2-215) gives

$$Q_T = \frac{2}{3} WLC_{\text{ox}}(V_{GS} - V_t) \tag{2-238}$$

and thus

$$C_{gs} = \frac{\partial Q_T}{\partial V_{GS}} = \frac{2}{3} WLC_{\text{ox}} \tag{2-239}$$

In addition, there is a contribution to C_{gs} from the constant parasitic oxide capacitance due to gate overlap of the source region.

The f_T of the MOSFET is given by

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{gs} + C_{gd} + C_{gb}} \tag{2-240}$$

The dependence of MOSFET f_T on device and process parameters can be seen by assuming that the intrinsic device capacitance G_{gs} dominates. Thus from Eq. (2-240) we have

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{gs}} \tag{2-241}$$

Substituting in Eq. (2-241) for g_m from Eq. (2-225) and C_{gs} from Eq. (2-239) we find for a MOSFET

$$f_T = 1.5 \frac{\mu_n}{2\pi L^2} (V_{GS} - V_t) \tag{2-242}$$

It is interesting to compare this with the intrinsic f_T of a bipolar transistor when parasitic depletion-layer capacitance is neglected. From

$$f_T = \frac{1}{2\pi\tau_F} \tag{2-243}$$

and substituting for τ_f and using the Einstein relationship $D_n/\mu_n = kT/q = V_T$, we find for a bipolar transistor

$$f_T = 2 \frac{\mu_n}{2\pi W_B^2} V_T \tag{2-244}$$

The similarity in form between Eqs. (2-242) and (2-244) is striking. In both cases the intrinsic device f_T increases as the inverse square of the critical device dimension across which carriers are in transit. The voltage V_t = 26 mV is fixed for a bipolar transistor, but MOSFET f_T can be increased by operating at high values of $(V_{GS} - V_t)$. Note that the base width W_B in a bipolar transistor is a vertical dimension determined by diffusions or implants and can typically be made much smaller than the channel length L of a MOSFET, which depends on surface geometry and photolithographic processes. Thus bipolar transistors generally have higher f_T than MOSFETs made with comparable processing. Finally, Eq. (2-242) was derived assuming that the MOSFET square law is valid. However, as discussed in Section 2-3-5, submicron MOSFETs depart significantly from square-law characteristics, and we find that for such devices f_T increases as L^{-1} rather than L^{-2}.

2-3-5 Short-Channel Effects in FETs

The evolution of integrated-circuit processing techniques has led to continuing reductions in both the horizontal and vertical dimensions of the active devices (the minimum allowed dimension of passive devices has also increased). This trend is driven primarily by economics in that more devices and circuits can be processed at one time on a given wafer. A second benefit has been that the frequency capability of the active devices continues to increase, as intrinsic f_T values increase with smaller dimensions while parasitic capacitances decrease.

Vertical dimensions such as the base width of a bipolar transistor in production processes may now be on the order of 0.05 μm or less, whereas horizontal dimensions such as bipolar emitter width or FET gate length may be significantly less than 1 μm. Even at these very small dimensions, the large-signal and small-signal models of bipolar transistors given in previous sections remain valid. However, significant short-channel effects become important in FETs of all types at channel lengths on the order of 1 μm or less and require modifications to the FET models given previously. The primary effect is to modify the classical FET square-law transfer characteristic in the saturation region to make the device more closely approach an ideal linear transfer function. Note, however, that even in processes with submicron capability, many of the FETs in a given analog circuit may be deliberately chosen to be larger than the minimum size and may be well approximated by the square-law model.

The most important short-channel effect in FETs is due to velocity saturation of carriers in the channel [17]. At low electric field values, the linear relation between carrier velocity and field implied by Eqs. (2-210) and (2-211) is valid. At high fields, however, the carrier velocities approach the thermal velocities and subsequently the carrier velocities increase more slowly with increasing field. This is illustrated in Figure 2-108, which shows typical measured electron drift velocity v_d versus tangential electric field strength magnitude $\mathscr{E}$ in an NMOS surface channel. Note that at low field values the velocity is proportional to the field, while at high fields the velocity approaches a constant value called the scattering-limited velocity v_{scl}. A first-order analytical approximation to this curve is

$$v_d = \frac{\mu_n \mathscr{E}}{1 + \mathscr{E}/\mathscr{E}_c} \tag{2-245}$$

where $\mathscr{E} \simeq 1.5 \times 10^6$ V/m and $\mu_n \simeq 0.07$ m^2/V · s is the low-field mobility. Equation (2-245) is also plotted in Figure 2-108. From Eq. (2-245), as $\mathscr{E} \to \infty$ we have $v_{\text{scl}} = \mu_n \mathscr{E}_c$. At the critical field value $\mathscr{E}_c$, the carrier velocity is a factor of 2 less than the low-field formula would predict. In a device with a channel length $L = 1$ μm, we need a voltage drop of only 1.5 V along the channel to have an average field equal to $\mathscr{E}_c$, and this condition is readily achieved in small MOSFETs. Similar results are found for PMOS devices.

As an example of the effects of velocity saturation on FET characteristics, we consider the example of the MOSFET. In the analysis of "Transfer Characteristics of MOS Devices" in Section 2-3-3, we now use the more general expression

$$I_D = WQ_I(y)v_d(y) \tag{2-246}$$

Substituting Eq. (2-245) into Eq. (2-246) and using

$$\mathscr{E} = \frac{dV}{dy} \tag{2-247}$$

for the magnitude of the field, we find that

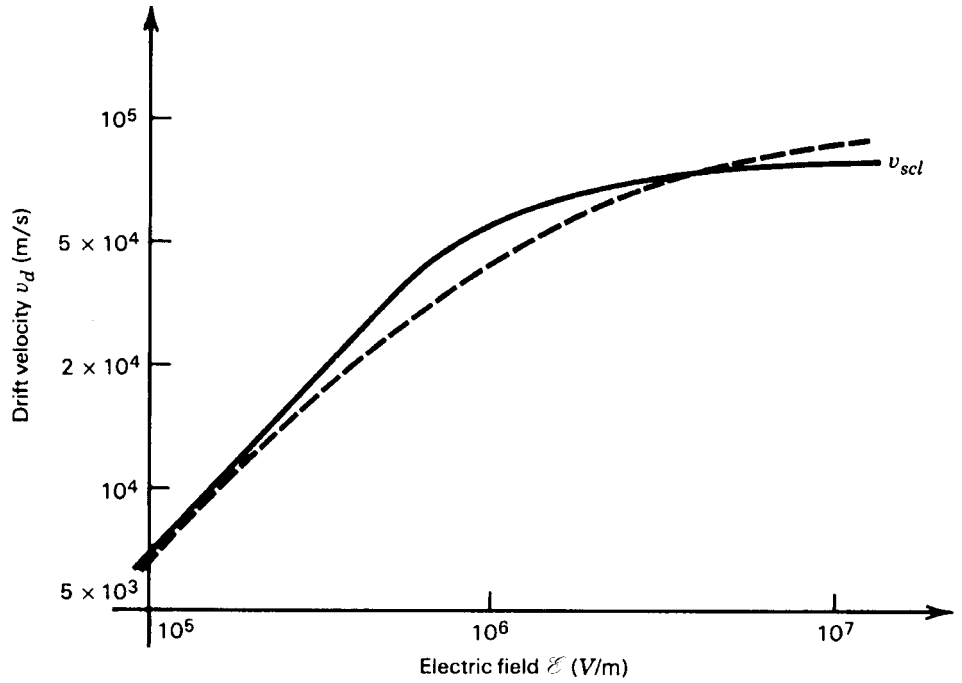


Figure 2-108 Typical measured electron drift velocity v_d versus tangent electric field $\mathscr{E}$ in an MOS surface channel (solid line). Also shown (dotted line) is the analytical approximation of Eq. (2.245) with $\mathscr{E}_c = 1.5 \times 10^6$ V/m and $\mu_n = 0.07$ m^2/V · s.

$$I_D\left(1 + \frac{1}{\mathscr{E}_c}\frac{dV}{dy}\right) = WQ_I(y)\mu_n\frac{dV}{dy} \tag{2-248}$$

Note that as $\mathscr{E}_c \to \infty$ and velocity saturation becomes negligible, Eq. (2-248) approaches the original equation (2-211).

Integrating Eq. (2-248) along the channel we obtain

$$\int_0^L I_D\left(1 + \frac{1}{\mathscr{E}_c}\frac{dV}{dy}\right)dy = \int_0^{V_{DS}} WQ_I(y)\mu_n dV \tag{2-249}$$

and thus

$$I_D = \frac{\mu_n C_{ox}}{2\left(1 + \dfrac{1}{\mathscr{E}_c}\dfrac{V_{DS}}{L}\right)}\frac{W}{L}[2(V_{GS} - V_t)V_{DS} - V_{DS}^2] \tag{2-250}$$

The quantity V_{DS}/L in Eq. (2-250) can be interpreted as the average field in the channel. If this is comparable to $\mathscr{E}_c$, the drain current for a given V_{DS} is less than the simple expression (2-213) would predict.

Equation (2-250) is valid in the triode region. The MOSFET transfer function in saturation can be obtained by using $V_{DS} = (V_{GS} - V_t)$ in Eq. (2-250) to obtain

$$I_D = \frac{k'}{2[1 + \theta(V_{GS} - V_t)]} \frac{W}{L} (V_{GS} - V_t)^2 \tag{2-251}$$

where $\theta = 1/L\mathscr{E}_c$ and has the dimension volts^{-1}. For $L = 1$ μm, a typical value is $\theta \approx 0.7\ \text{V}^{-1}$. Note that in the presence of velocity saturation effects, the device enters the saturation region for $V_{DS} < (V_{GS} - V_t)$. However, Eq. (2-251) still gives a good estimate of the saturation current.

Thus far, we have only considered the effects of the tangential field due to the V_{DS} along the channel when examining velocity saturation effects. However, there exists a normal field originating from the gate voltage that also inhibits channel carrier mobility. Since the normal field depends on the value of V_{GS}, we find that an empirical modification to θ in Eq. (2-251) can adequately model this effect. In practice, θ is determined by a best fit to measured device characteristics.

Returning to Eq. (2-251) we note that for very short channel lengths, θ becomes large and Eq. (2-251) reduces to

$$I_D \propto (V_{GS} - V_t) \tag{2-252}$$

Thus, the FET characteristics tend toward a *linear* transfer function as the channel length becomes very small (less than 1 μm).

Equation (2-251) has a simple circuit representation. Consider the circuit of Figure 2-109, where an ideal square-law MOSFET has a resistance R_{SX} in series with the source of the FET. Assume

$$I_D = \frac{\mu C_{\text{ox}}}{2} \frac{W}{L} (V'_{GS} - V_t)^2 \tag{2-253}$$

Now

$$V_{GS} = V'_{GS} + I_D R_{SX} \tag{2-254}$$

and substituting Eq. (2-254) into Eq. (2-253) we find that

$$I_D = \frac{\mu C_{\text{ox}}}{2} \frac{W}{L} (V_{GS} - I_D R_{SX} - V_t)^2 \tag{2-255}$$

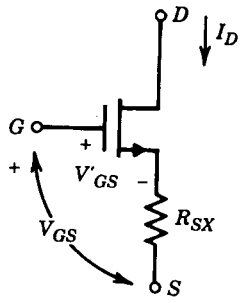


Figure 2-109 Model of velocity saturation of a MOSFET by addition of a series source resistance to an ideal square-law device.

Rearranging Eq. (2-255) we find

$$I_D = \frac{\mu C_{\text{ox}}}{2\left[1 + \mu C_{\text{ox}} \frac{W}{L} R_{SX}(V_{GS} - V_t)\right]} \frac{W}{L} (V_{GS} - V_t)^2 \tag{2-256}$$

This has exactly the same form as Eq. (2-251) if we identify

$$\theta = \mu C_{\text{ox}} \frac{W}{L} R_{SX} \tag{2-257}$$

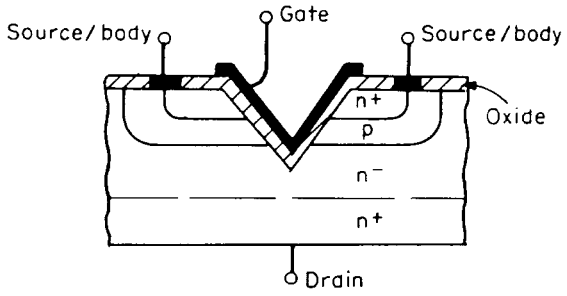
Substituting $\theta = 1/L\mathscr{E}_c$ into Eq. (2-257) we have

$$R_{SX} = \frac{1}{\mathscr{E}_c} \frac{1}{\mu C_{\text{ox}}} \frac{1}{W} \tag{2-258}$$

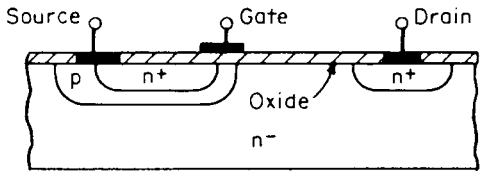
Thus, the influence of velocity saturation on the large-signal characteristics of a FET can be modeled to first order by a resistor R_{SX} in series with the source of an ideal square-law device. Note that R_{SX} varies inversely with W, as does the intrinsic physical series resistance due to source and drain contact regions. Typically, R_{SX} is larger than the physical series resistance. For $W = 2\ \mu\text{m}$, $\mu C_{\text{ox}} = 40\ \mu\text{A}/V^2$, and $\mathscr{E}_c = 1.5 \times 10^6$ V/m, we find $R_{SX} = 8\ \text{k}\Omega$.

The foregoing analysis has developed a modified large-signal mode for the MOSFET including velocity saturation effects. Small-signal MOSFET modeling for small devices can still be done using the equivalent circuit of Figure 2-107 if the values of g_m and g_{mb} are modified to account for the effects of velocity saturation using Eq. (2-251).

Figure 2-110 shows a cross section of VMOS (vertical MOS) and DMOS (double-diffused MOS) transistors. There is also a TMOS device—a Motorola variant of VMOS—and the



(a) VMOS (vertical channel MOS)



(b) DMOS (double-diffused MOS)

Figure 2-110 Cross section of (a) VMOS and (b) DMOS FETs.

now very popular LDMOS (lateral diffused MOS) power transistor (Figures 2-111 and 2-112). LDMOS FETs operate from a single supply (approximately 8 V and up) and are available for output powers surpassing 100 W at 2 GHz. The feedback capacitance is much lower than VMOS, TMOS, and earlier models, significantly improving circuit stability.

Following is a listing of Motorola high-power LDMOS transistors capable of operation up to 2 GHz:

RF High-Power LDMOS Transistors by Motorola
2.0 GHz, V_{DD} = 26 V, Class AB

Device	P_{out} Output Power (watts)	P_{in} Input Power Typical (watts)	G_{ps} (Typ.)/Freq. (dB/MHz)	IMDs (–dBc)	η Efficiency, Typ. %	θ_{JC} (°C/W)	Package/Style
MRF281S	4	0.2	13.6/2000	–30	40	8.75	458/1
MRF281Z	4	0.2	13.6/2000	–30	40	8.75	458A/1
MRF6525-5	5	0.4	11.5/2000	—	40	15	458A/1
MRF6525-10	10	1.0	10/2000	—	40	6.0	458A/1
MRF282S	10	0.5	13/2000	–30	35	2.9	458/1
MRF282Z	10	0.5	13/2000	–30	35	2.9	458A/1
MRF284	30	2.4	11.5/2000	–30	33	2.0	360B/1
MRF284S	30	2.4	11.5/2000	–30	33	2.0	360C/1
MRF286	60	4.75	11.4/2000	–30	33	.73	465/1
MRF286S	60	4.75	11.4/2000	–30	33	.73	465A/1

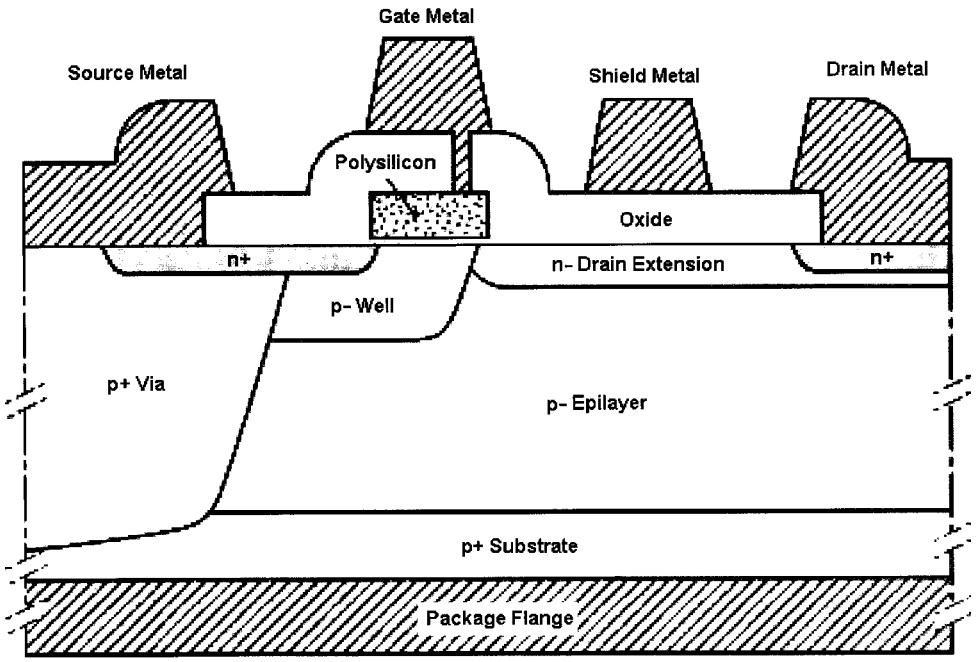


Figure 2-111 Cross section of a Philips LDMOS FET [18].

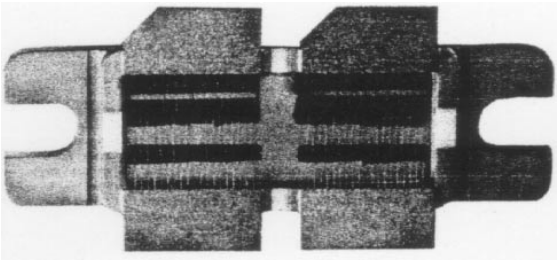


Figure 2-112 Photo of the underside of an LDMOS FET.

2-3-6 Small-Signal Models of MOSFETs

Subthreshold Conduction in MOSFETs. The MOSFET analysis of Section 2-3-3 focused on the normal region of operation where this is a well-defined conducting channel under the gate. Changes in the gate voltage are assumed to cause changes in the channel charge only, and not in the depletion region below. However, for gate voltages less than the extrapolated threshold voltage V_t, the applied gate potential affects both the depletion-region charge and the channel charge, which is then very small but not zero. The device can thus conduct finite (but small) current for $V_{GS} < V_t$, so that Eq. (2-215) is not valid in this region. The electrons in the n^+ source region of an NMOS transistor can overcome the potential barrier to the p-type substrate and enter the channel region. This process is very similar to the turn-on of a bipolar transistor, and in fact the MOSFET characteristics in this subthreshold region (also called weak inversion) are very similar to those of a bipolar transistor. Analysis [19] shows that in the subthreshold region of the MOSFET characteristics can be defined by the equation

$$I_D = k_x \frac{W}{L} e^{V_{GS}/nV_T}(1 - e^{-V_{DS}/V_T}) \tag{2-259}$$

where k_x depends on process parameters and $n \simeq 1.5$. The value of n is not equal to unity in this case (as it is for the bipolar transistor) because the applied voltage V_{GS} appears partially at the silicon surface and partially across the depletion layer.

To illustrate this effect, we show measured NMOS characteristics plotted on three different scales in Figure 2-113. In Figure 2-113a we show the transfer characteristic in the forward-active region plotted on linear scales. For this device W = 20 μm and L = 20 μm and short-channel effects are negligible. The same data are plotted in Figure 2-113b as $\sqrt{I_D}$ versus V_{GS}. The resulting straight line shows that the device characteristic is close to an ideal square law. Plots like this are commonly used to obtain V_t by extrapolation (0.7 V in this case) and also k' from the slope of the curve (54 μA/V^2 in this case). Note that near the threshold voltage, the curve deviates from the straight line representing the square law. This is the subthreshold region. The data are plotted a third time in Figure 2-113c on log–linear scales. The straight line obtained for $V_{GS} < V_t$ fits Eq. (2-259) with n = 1.5. At currents below 10^{-12} A, the effect of leakage currents becomes evident.

The major application of subthreshold operation is in very low power applications at relatively low signal frequencies. The limitation to low signal frequencies occurs because the MOSFET f_T becomes very small. Since the device capacitances at very low bias currents

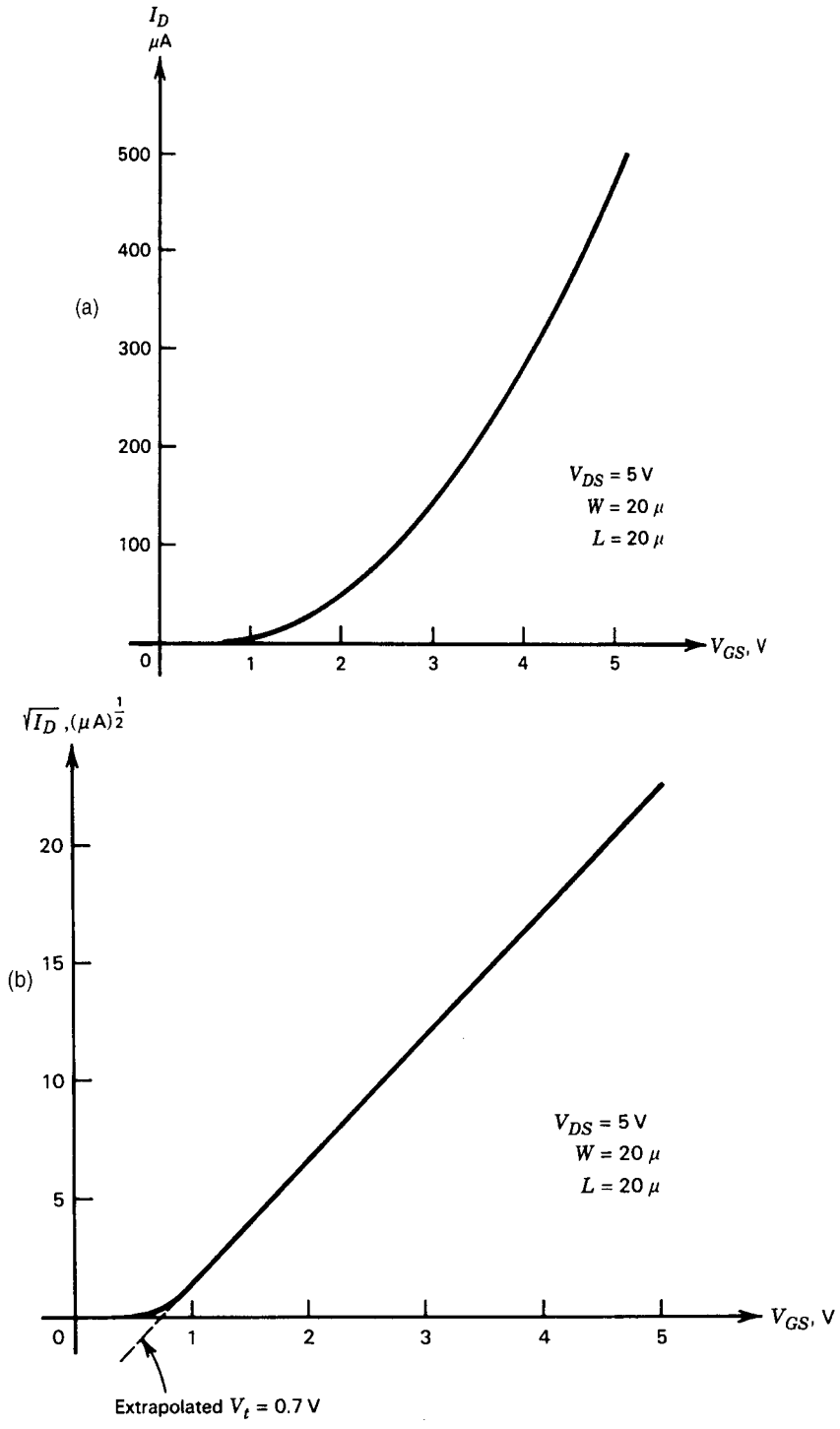


Figure 2-113 (a) Measured NMOS transfer characteristic in the forward-active region. (b) Data from part (a) plotted as $\sqrt{I_D}$ versus V_{GS} showing the square-law characteristic. (c) Data from part (a) plotted on log–linear scales showing the exponential characteristic of the subthreshold region.

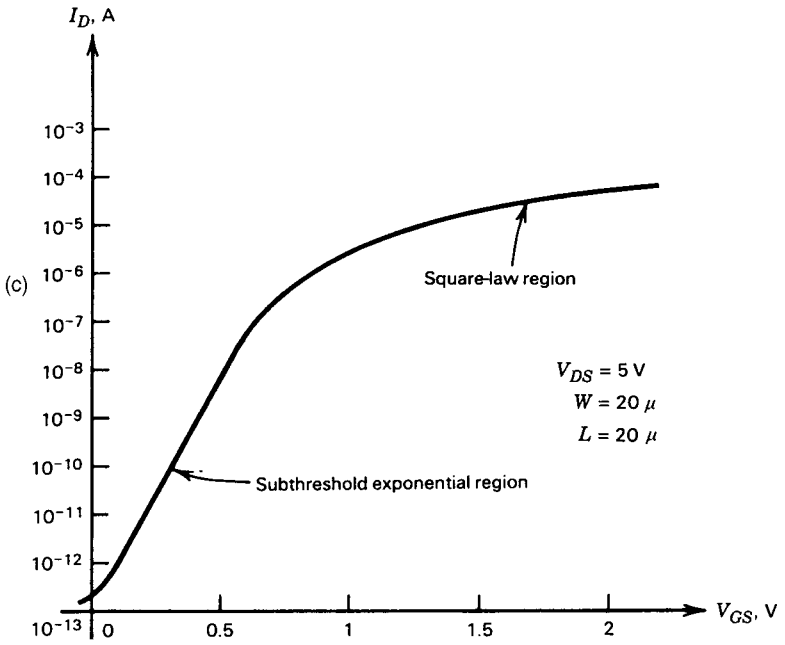


Figure 2-113. (*Continued*)

are essentially fixed, and the small-signal g_m calculated from Eq. (2-259) becomes proportional to I_D, we see that the value of f_T becomes very small at very low values of I_D.

Substrate Flow in MOSFETs. As the reverse-bias voltages on the device are increased, carriers traversing the depletion regions gain sufficient energy to create new hole–electron pairs in lattice collisions by a process known as *impact ionization.* Eventually, at high enough bias voltages, the process results in large avalanche currents. For collector–base bias voltages well below the breakdown value, a small enhanced current flow may occur across the collector–base junction due to this process with little apparent effect on the device characteristics.

Impact ionization also occurs in MOSFETs but has a significantly different effect on the device characteristics. This is because the channel electrons (for the case of NMOS) create hole–electron pairs in lattice collisions in the drain depletion region, and some of the resulting holes then flow to the substrate, creating a substrate current. (The electrons created

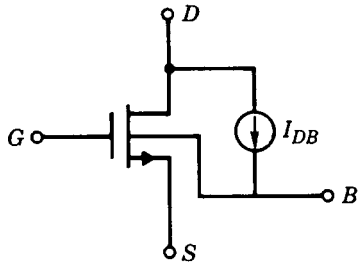


Figure 2-114 Representation of impact ionization in a MOSFET by a drain–substrate current generator.

in the process flow out the drain terminal.) The carriers created by impact ionization are thus not confined within the device as they are in a bipolar transistor. The effect of this phenomenon can be modeled by inclusion of a controlled current generator I_{DB} from drain to substrate, as shown in Figure 2-114 for an NMOS device. The magnitude of this substrate current depends on the voltage across the drain depletion region (which determines the energy of the ionizing channel electrons), and also on the drain current (which is the rate at which channel electrons enter the depletion region). It has been found empirically [20] that the current I_{DB} can be expressed as

$$I_{DB} = K_1(V_{DS} - V_{DS,\text{sat}})I_D e^{-[K_2/(V_{DS}-V_{DS,\text{sat}})]} \tag{2-260}$$

where K_1 and K_2 are process-dependent parameters and $V_{DS,\text{sat}}$ is the value of V_{DS} where the drain characteristics enter the saturation region. Typical values for NMOS devices are $K_1 = 5\ \text{V}^{-1}$ and $K_2 = 30$ V. The effect is generally much less significant in PMOS devices because the holes carrying the charge in the channel are much less efficient in creating hole–electron pairs than are energetic electrons.

The major impact of this phenomenon on circuit performance is that a parasitic resistance from drain to substrate now exists. Since the substrate of an NMOS device in a *p*-substrate process is an ac ground (the common substrate terminal must always be connected to the most negative supply voltage in the circuit), the parasitic resistance shunts the drain to ground and can be a limiting factor in many circuit designs. Differentiating Eq. (2-260) we find for the drain–substrate small-signal conductance

$$g_{db} = \frac{\partial I_{DB}}{\partial V_D}$$

$$= K_2 \frac{I_{DB}}{(V_{DS} - V_{DS,\text{sat}})^2} \tag{2-261}$$

where the gate and source are assumed held at fixed potentials.

In IC applications for RF/UHF circuits, many not-yet-solved problems remain:

1. The flicker corner frequency for MOS devices is somewhere between 100 kHz and 1 MHz.
2. Because of backplane problems in ICs, such as the substrate capacitance, also relative to the source, the performance of an IC-MOS/CMOS is not as good as expected if these effects are not considered.
3. Discrete devices, such as *n*-channel MOS devices, which are also available as dual-gate MOSFETs, can work well up to 1200 MHz. Such performance is not yet shown in silicon-based RFICs.
4. For RF power devices, such as LDMOS and others mentioned, promising results are only shown for discrete devices. There are no equivalent ICs on the market yet.

It will be interesting to see how this technology will work itself upward in frequency. Its main advantage is the BiCMOS process, which allows the mixing of analog and digital circuits nicely. Heterojunction transistors, such as SiGe, are inherently much more linear

and, for the same performance, require less current than MOS technology. On the other hand, we have not seen large-scale production SiGe-based ICs.

In the following five pages, we reproduce the datasheet for the Siemens (now Infineon Technologies) BF999 single-gate MOSFET.* Above 500 MHz, a dual-gate (also referred to as tetrode) MOSFET, such as the BF998, is preferred over a single-gate device. The main reasons for this are:

1. Internally, it is equivalent to two single-gate devices connected in a cascade arrangement, with the output device operating in the grounded-gate configuration, which entirely avoids the Miller effect and therefore remains stable at higher frequencies, operating up to 1.2 GHz.
2. The second gate can be used for AGC. As the Gate 2 voltage changes, the gain varies heavily.
3. By applying a local oscillator signal to Gate 2, this configuration can be used as a mixer.

The major drawbacks of these devices are their wide tolerances and temperature sensitivity. They follow to some degree the CMOS design but seem to have better performance.

The BF995, also called a tetrode, is in reality a dual-gate MOSFET specified to only 300 MHz; however, we are reproducing its datasheet* because the one more applicable for wireless circuits, the BF1005, does not have such a complete set of data.

To demonstrate the capabilities of low-voltage LDMOS transistors, we show the MRF5003 datasheet.† This device is intended for low-voltage applications down to 7.5 V. To the best of our knowledge, this is the lowest-voltage LDMOS transistor on the market. Other members of this family, which produce 5 W (MRF5005) and 15 W (MRF5015), begin to require higher supply voltages. Just to make sure it is not forgotten, the advantage of LDMOS is, besides lower-voltage operation, much-reduced feedback capacitance. LDMOS devices are now available that operate above 1 GHz and produce 60 W. There is now some competition with the recently introduced NEC GaAsFET, which goes up to 100 W in a similar frequency range.

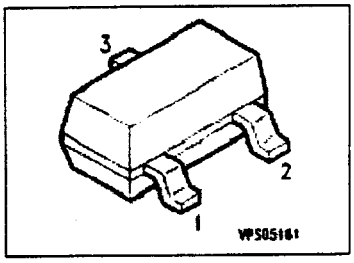
*Reproduced with permission.
†Copyright of Motorola, used by permission.

SIEMENS

Silicon N Channel MOSFET Triode

BF 999

- For high-frequency stages up to 300 MHz, preferably in FM applications



Type	Marking	Ordering Code (tape and reel)	Pin Configuration 1	2	3	Package
BF 999	LB	Q62702-F1132	G	D	S	SOT-23

Maximum Ratings

Parameter	Symbol	Values	Unit
Drain-source voltage	V_{DS}	20	V
Drain current	I_D	30	mA
Gate-source peak current	$\pm I_{GSM}$	10	
Total power dissipation, $T_A \leq 60$ °C	P_{tot}	200	mW
Storage temperature range	T_{stg}	– 55 ... + 150	°C
Channel temperature	T_{ch}	150	

Thermal Resistance

Junction - ambient [1]	$R_{th\,JA}$	≤ 450	K/W

[1] Package mounted on alumina 15 mm × 16.7 mm × 0.7 mm.

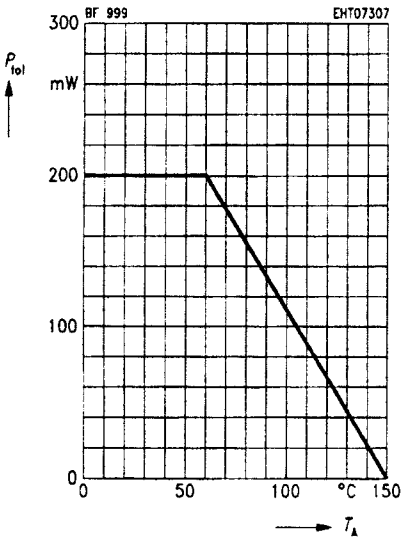
SIEMENS **BF 999**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

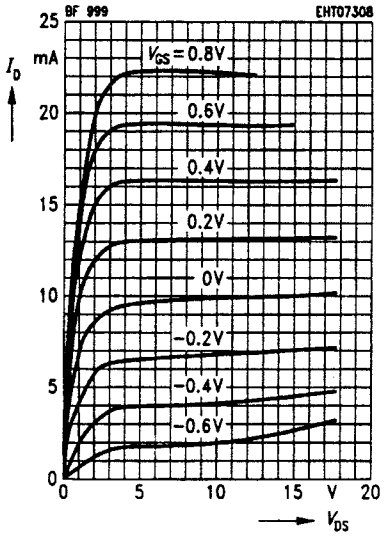
Parameter	Symbol	Values			Unit
		min.	typ.	max.	
DC Characteristics					
Drain-source breakdown voltage I_D = 10 µA, – V_{GS} = 4 V	$V_{(BR)\,DS}$	20	–	–	V
Gate-source breakdown voltage ± I_{GS} = 10 mA, V_{DS} = 0	± $V_{(BR)\,GSS}$	6.5	–	12	
Gate-source leakage current ± V_{GS} = 5 V, V_{DS} = 0	± I_{GSS}	–	–	50	nA
Drain current V_{DS} = 10 V, V_{GS} = 0	I_{DSS}	5	–	18	mA
Gate-source pinch-off voltage V_{DS} = 10 V, I_D = 20 µA	– $V_{GS\,(p)}$	–	–	2.5	V
AC Characteristics					
Forward transconductance V_{DS} = 10 V, I_D = 10 mA, f = 1 kHz	g_{fs}	14	16	–	mS
Gate input capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{gss}	–	2.5	–	pF
Reverse transfer capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{dg}	–	25	–	fF
Output capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{dss}	–	1	–	pF
Power gain (test circuit) V_{DS} = 10 V, I_D = 10 mA, f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS	G_p	–	25	–	dB
Noise figure (test circuit) V_{DS} = 10 V, I_D = 10 mA, f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS	F	–	1	–	

SIEMENS **BF 999**

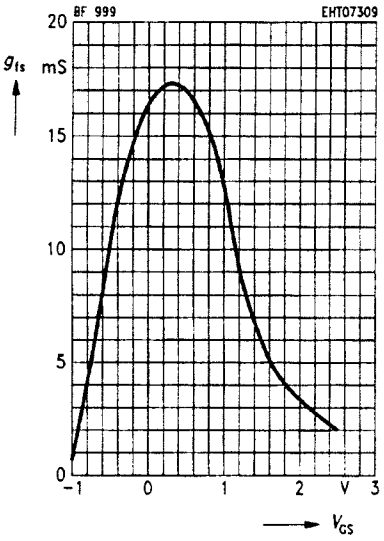
Total power dissipation $P_{tot} = f(T_A)$



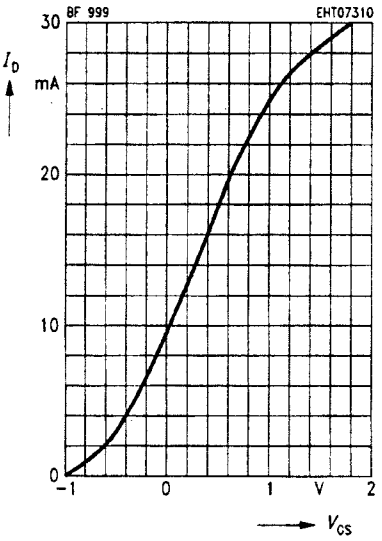
Output characteristics $I_D = f(V_{DS})$



Gate transconductance $g_{fs} = f(V_{GS})$
V_{DS} = 10 V, I_{DSS} = 10 mA, f = 1 kHz

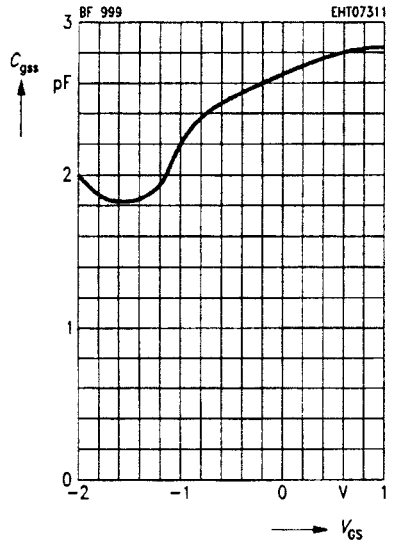


Drain current $I_D = f(V_{GS})$
V_{DS} = 10 V

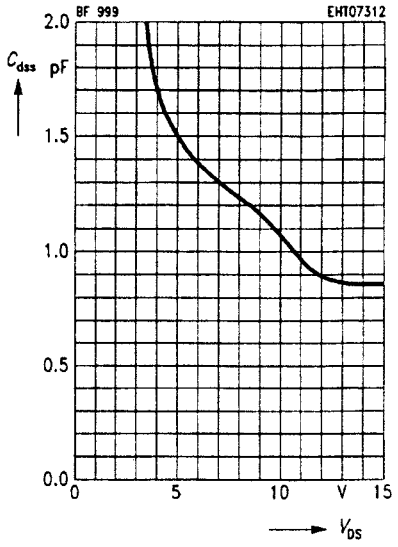


SIEMENS **BF 999**

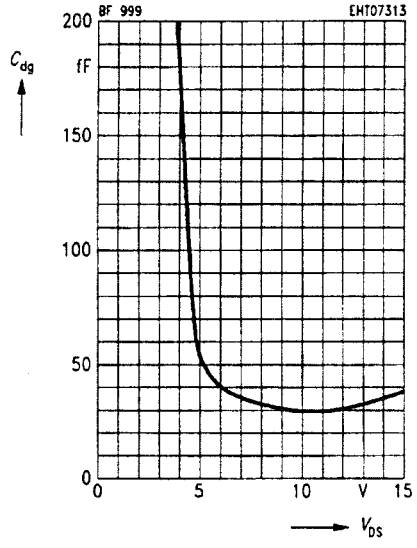
Gate input capacitance $C_{gss} = f(V_{GS})$
V_{DS} = 10 V, I_{DSS} = 10 mA, f = 1 MHz



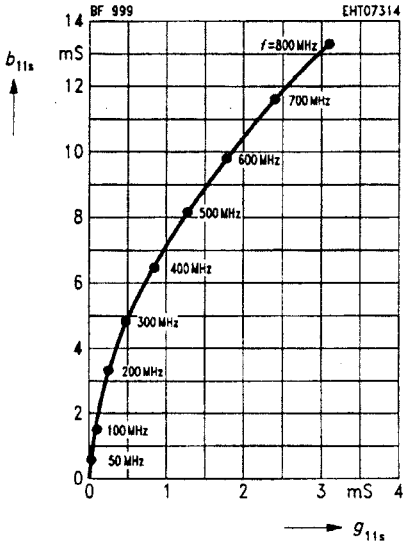
Output capacitance $C_{dss} = f(V_{DS})$
V_{GS} = 0, I_{DSS} = 10 mA, f = 1 MHz



Reverse transfer capacitance
$C_{dg} = f(V_{DS})$
I_{DSS} = 10 mA, f = 1 MHz, V_{GS} = 0

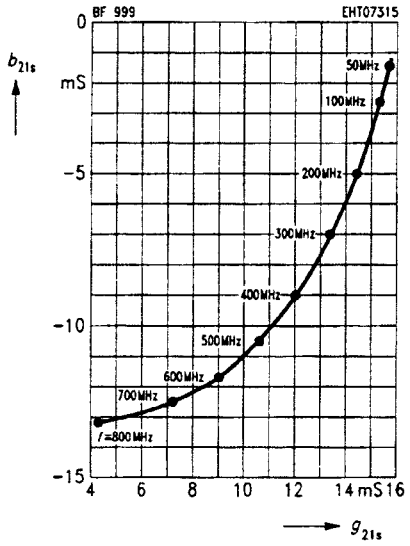


Gate input admittance y_{11s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)

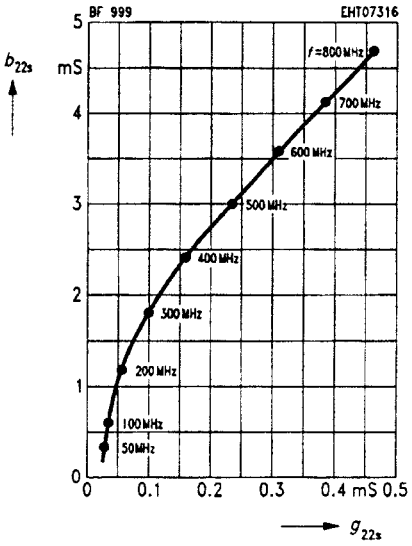


SIEMENS **BF 999**

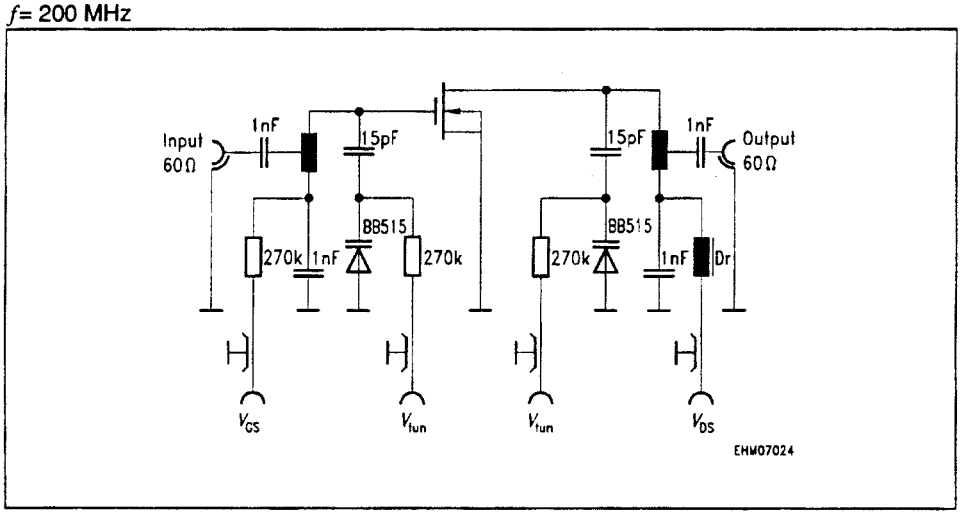
Gate forward transfer admittance y_{21s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)



Output admittance y_{22s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)



Test circuit for power gain and noise figure
f= 200 MHz

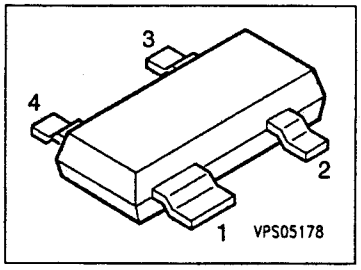


SIEMENS

Silicon N Channel MOSFET Tetrode

BF 995

- For input and mixer stages in FM and VHF TV tuners



Type	Marking	Ordering Code (tape and reel)	Pin Configuration 1	2	3	4	Package
BF 995	MB	Q62702-F936	S	D	G_2	G_1	SOT-143

Maximum Ratings

Parameter	Symbol	Values	Unit
Drain-source voltage	V_{DS}	20	V
Drain current	I_D	30	mA
Gate 1/gate 2 peak source current	$\pm I_{G1/2SM}$	10	
Total power dissipation, T_S < 76 °C	P_{tot}	200	mW
Storage temperature range	T_{stg}	– 55 ... + 150	°C
Channel temperature	T_{ch}	150	

Thermal Resistance

Junction - soldering point	$R_{th\,JS}$	< 370	K/W

SIEMENS **BF 995**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

Parameter	Symbol	Values			Unit
		min.	typ.	max.	
DC Characteristics					
Drain-source breakdown voltage I_D = 10 µA, – V_{G1S} = – V_{G2S} = 4 V	$V_{(BR)\,DS}$	20	–	–	V
Gate 1 source breakdown voltage ± I_{G1S} = 10 mA, V_{G2S} = V_{DS} = 0	± $V_{(BR)\,G1SS}$	8.5	–	14	
Gate 2 source breakdown voltage ± I_{G2S} = 10 mA, V_{G1S} = V_{DS} = 0	± $V_{(BR)\,G2SS}$	8.5	–	14	
Gate 1 source leakage current ± V_{G1S} = 5 V, V_{G2S} = V_{DS} = 0	± I_{G1SS}	–	–	50	nA
Gate 2 source leakage current ± V_{G2S} = 5 V, V_{G1S} = V_{DS} = 0	± I_{G2SS}	–	–	50	
Drain current V_{DS} = 15 V, V_{G1S} = 0, V_{G2S} = 4 V	I_{DSS}	4	–	20	mA
Gate 1 source pinch-off voltage V_{DS} = 15 V, V_{G2S} = 4 V, I_D = 20 µA	– $V_{G1S\,(p)}$	–	–	2.5	V
Gate 2 source pinch-off voltage V_{DS} = 15 V, V_{G1S} = 0, I_D = 20 µA	– $V_{G2S\,(p)}$	–	–	2.0	

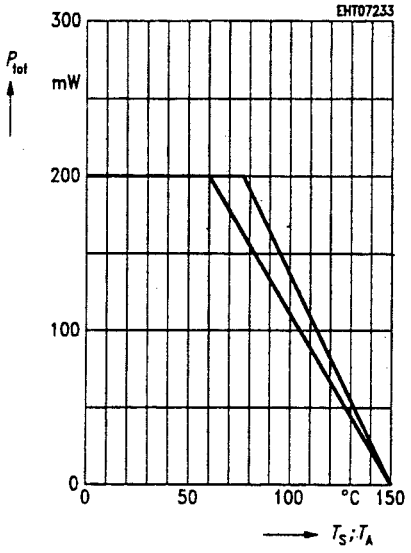
SIEMENS **BF 995**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

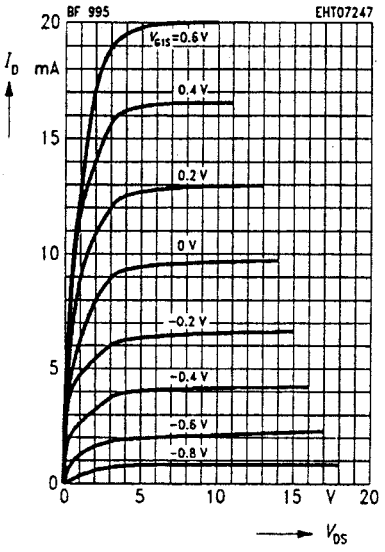
Parameter	Symbol	Values			Unit
		min.	typ.	max.	
AC Characteristics					
Forward transconductance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 kHz	g_{fs}	12	17	–	mS
Gate 1 input capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{g1ss}	–	3.6	–	pF
Gate 2 input capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{g2ss}	–	1.6	–	
Feedback capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{dg1}	–	25	–	fF
Output capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{dss}	–	1.6	–	pF
Power gain V_{DS} = 15 V, I_D = 10 mA f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS 2 Δf = 12 MHz (see test circuit 1)	G_{ps}	–	23	–	dB
Noise figure V_{DS} = 15 V, I_D = 10 mA f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS (see test circuit 1)	F	–	1.1	–	
Gain control range V_{DS} = 15 V, V_{G2S} = 4 ... – 2 V, f = 200 MHz (see test circuit 1)	ΔG_{ps}	–	50	–	
Mixer gain (additive) V_{DS} = 15 V, V_{G2S} = 6 V, R_S = 220 Ω f = 200 MHz, f_F = 36 MHz 2 Δf_F = 5 MHz, V_{osc} = 0.5 V (see test circuit 2)	G_{psc}	–	16	–	
Mixer gain (multiplicative) V_{DS} = 15 V, V_{G1S} = 1.7 V, V_{G2S} = 2.5 V R_S = 220 Ω, f = 200 MHz, f_F = 36 MHz 2 Δf_F = 5 MHz, V_{osc} = 2 V (see test circuit 3)	G_{psc}	–	18	–	

SIEMENS **BF 995**

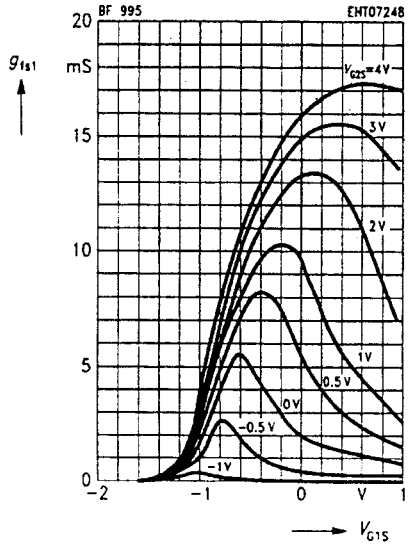
Total power dissipation $P_{tot} = f(T_A)$



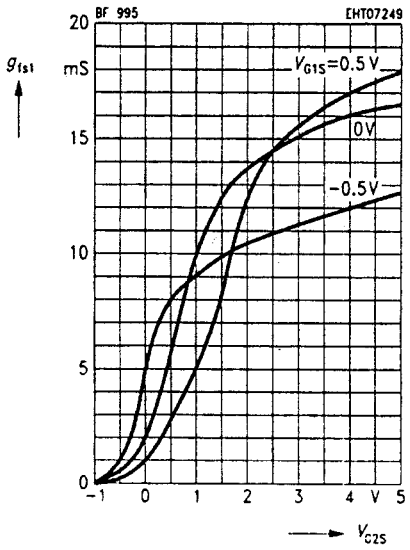
Output characteristics $I_D = f(V_{DS})$
V_{G2S} = 4 V



Gate 1 forward transconductance $g_{fs1} = f(V_{G1S})$
V_{DS} = 15 V, I_{DSS} = 10 mA, f = 1 kHz

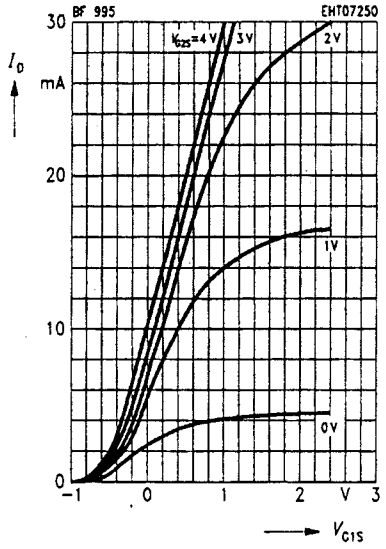


Gate 1 forward transconductance $g_{fs1} = f(V_{G2S})$
V_{DS} = 15 V, I_{DSS} = 10 mA, f = 1 kHz

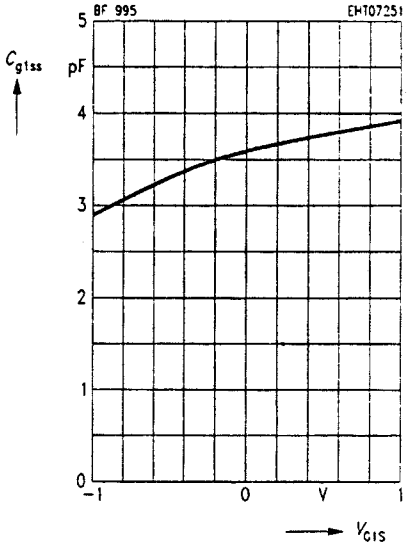


SIEMENS **BF 995**

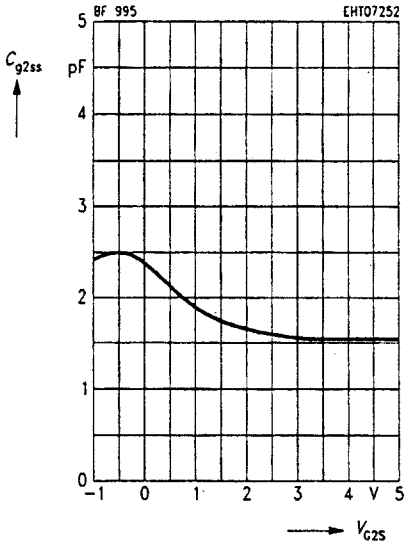
Drain current $I_D = f(V_{G1S})$
V_{DS} = 15 V



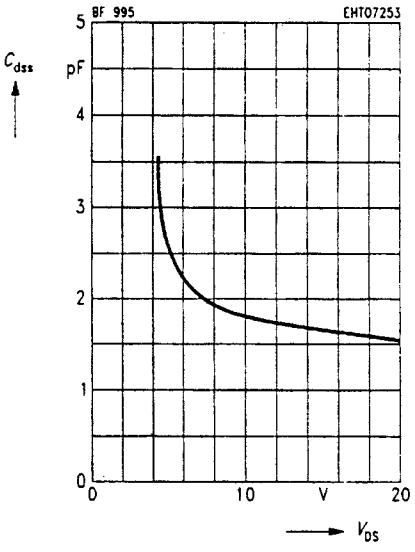
Gate 1 input capacitance $C_{g1ss} = f(V_{G1S})$
V_{G2S} = 4 V, V_{DS} = 15 V
I_{DSS} = 10 mA, f = 1 MHz



Gate 2 input capacitance $C_{g2ss} = f(V_{G2S})$
V_{G1S} = 0 V, V_{DS} = 15 V
I_{DSS} = 10 mA, f = 1 MHz

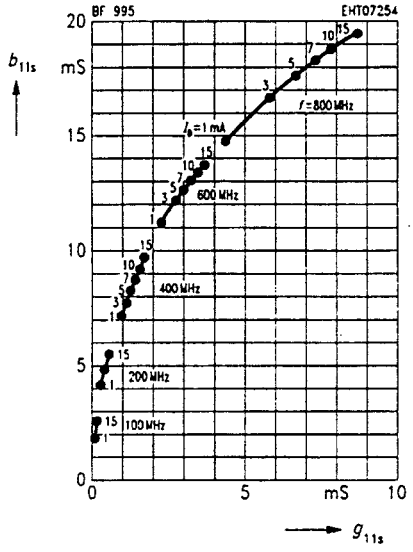


Output capacitance $C_{dss} = f(V_{DS})$
V_{G1S} = 0 V, V_{G2S} = 4 V
I_{DSS} = 10 mA, f = 1 MHz

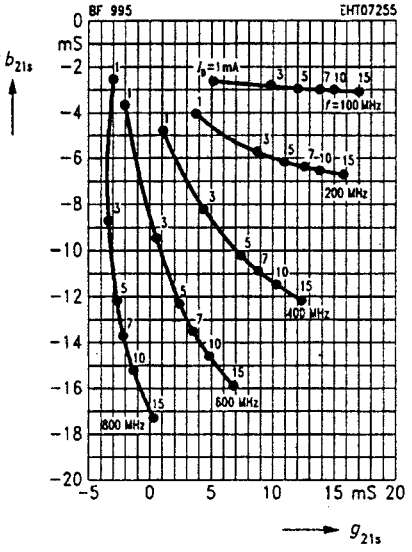


SIEMENS **BF 995**

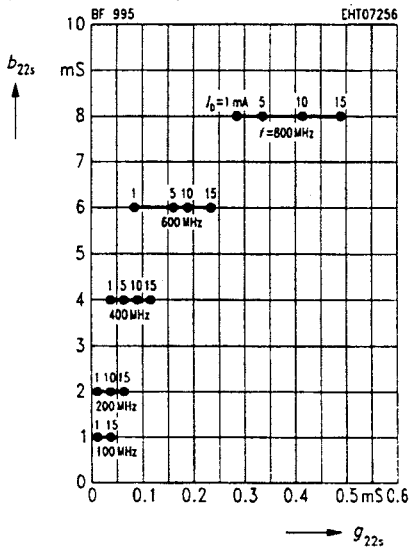
Gate 1 input admittance y_{11s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)



Gate 1 forward transfer admittance y_{21s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)

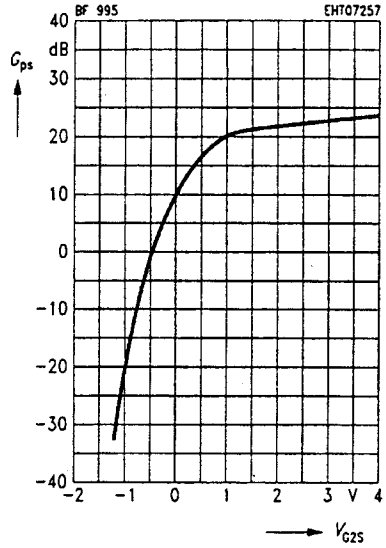


Output admittance y_{22s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)

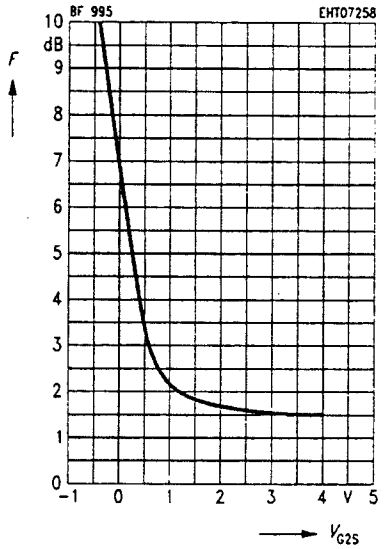


SIEMENS **BF 995**

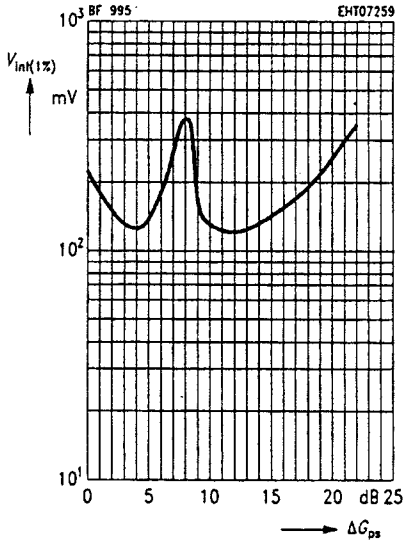
Power gain $G_{ps} = f(V_{G2S})$
V_{DS} = 15 V, V_{G1S} = 0 V, I_{DSS} = 10 mA
f = 200 MHz (see test circuit 1)



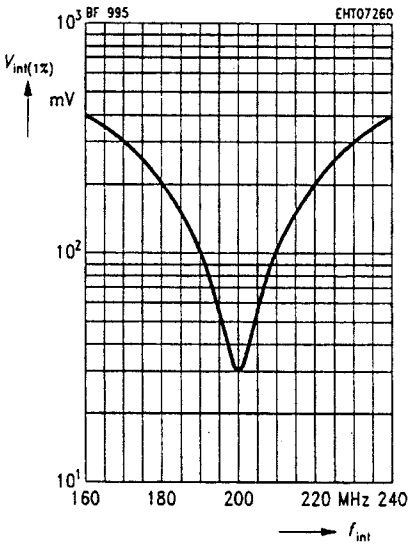
Noise figure $F = f(V_{G2S})$
V_{DS} = 15 V, V_{G1S} = 0 V, I_{DSS} = 10 mA
f = 200 MHz (see test circuit 1)



Interference voltage for 1% cross modulation $V_{int\,(1\%)} = f(\Delta G_{ps})$[1)]
V_{DS} = 15 V, V_{G1S} = 0, f = 200 MHz
f_{int} = 221 MHz (see test circuit 1)



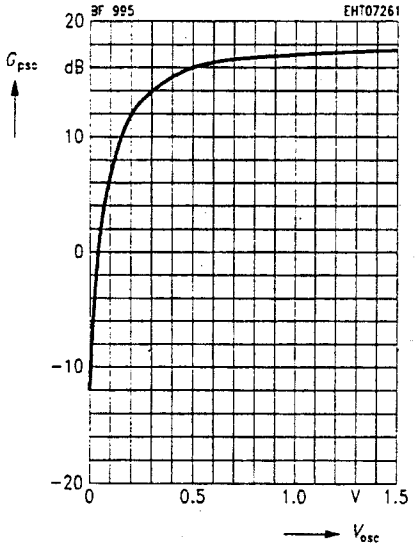
Interference voltage for 1% cross modulation $V_{int\,(1\%)} = f(f_{int})$[1)]
V_{DS} = 15 V, V_{G2S} = 4 V, V_{G1S} = 0
f = 200 MHz (see test circuit 1)



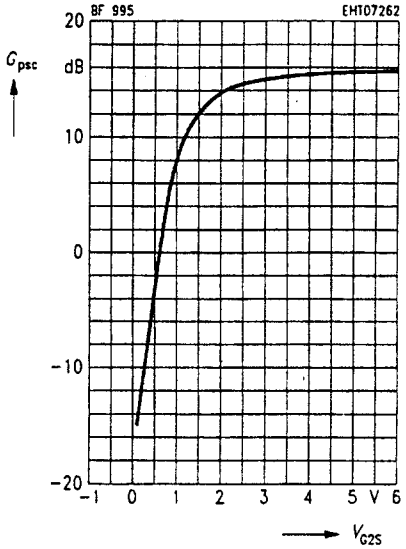
1) For footnote refer to the last page of this data sheet.

SIEMENS **BF 995**

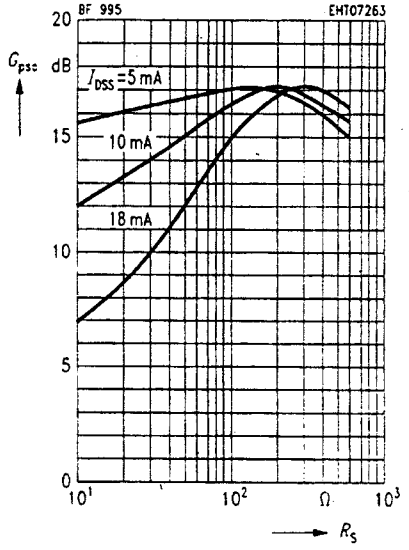
Mixer gain (additive) $G_{psc} = f(V_{osc})$
V_D = 15 V, V_{G1S} = 0, V_{G2S} = 6 V
R_S = 220 Ω, I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)



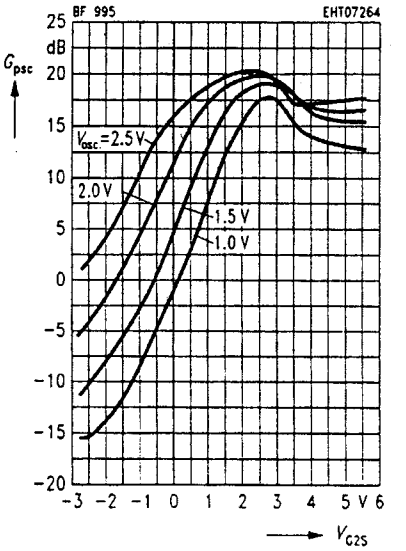
Mixer gain (additive) $G_{psc} = f(V_{G2S})$
V_D = 15 V, V_{G1S} = 0, R_S = 220 Ω
V_{osc} = 0.5 V, I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)



Mixer gain (additive) $G_{psc} = f(R_S)$
V_D = 15 V, V_{G1S} = 0, V_{G2S} = 6 V
V_{osc} = 0.5 V, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)

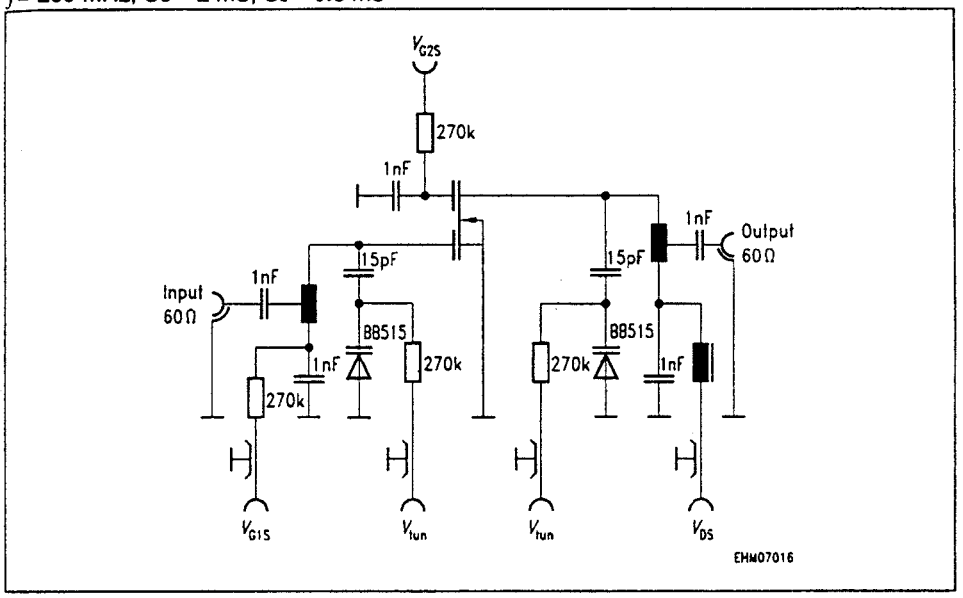


Mixer gain (multiplicative) $G_{psc} = f(V_{G2S})$
V_D = 15 V, V_{G1S} = 1.7 V, R_S = 200 Ω
I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 3)

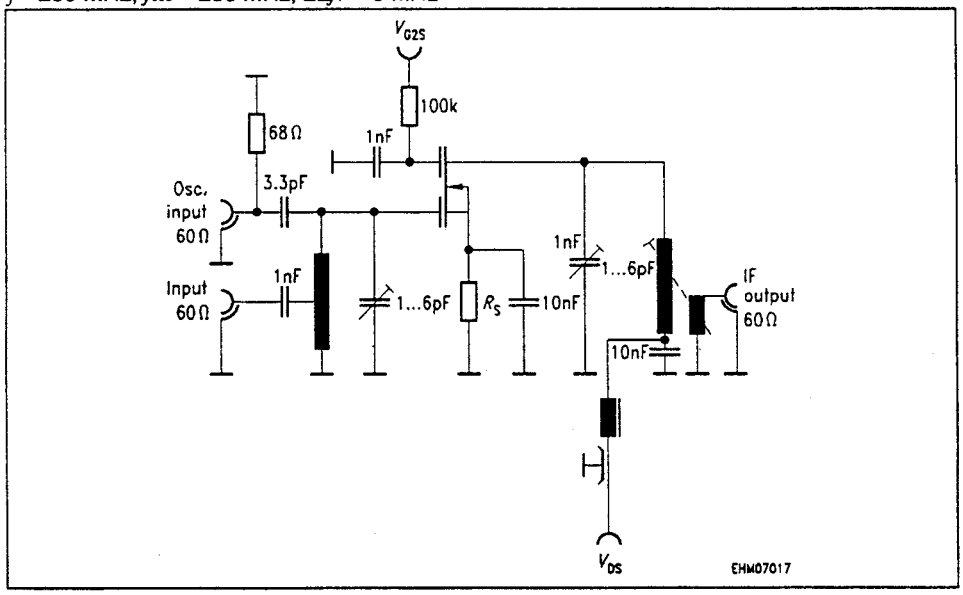


SIEMENS **BF 995**

Test circuit 1 for power gain, noise figure and cross modulation
f= 200 MHz, G_S = 2 mS, G_L = 0.5 mS

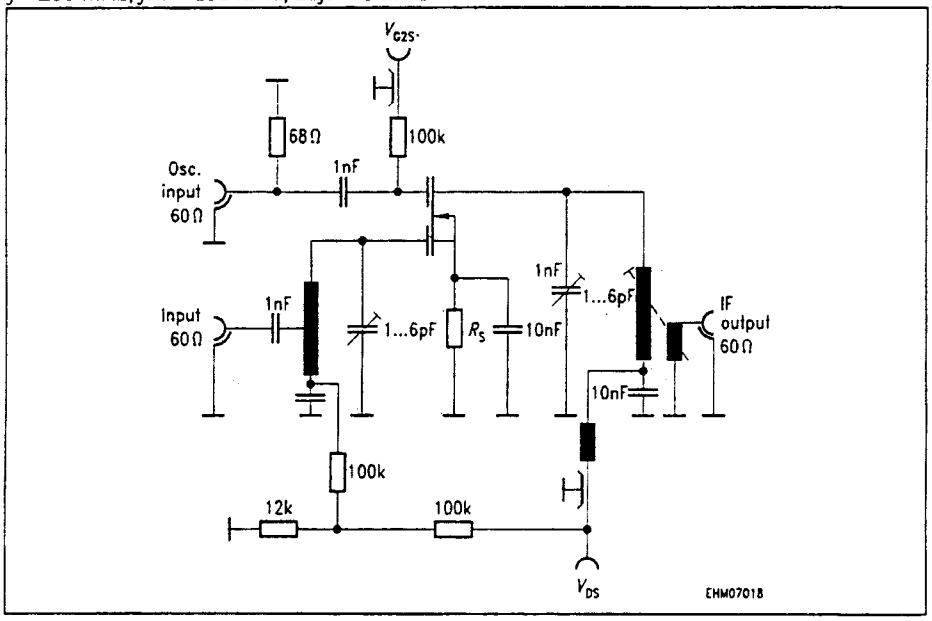


Test circuit 2 for mixer gain (additive)
f= 200 MHz, f_{osc} = 236 MHz, $2\Delta f_F$ = 5 MHz



SIEMENS **BF 995**

Test circuit 3 for mixer gain (multiplicative)
f = 200 MHz, f_{osc} = 236 MHz, $2\Delta f_F$ = 5 MHz



1) $V_{int\,(1\%)}$ is the rms value of half the emf (terminal voltage at matching) of a 100 % sine modulated TV carrier at an internal generator resistance of 60 Ω, causing 1 % amplitude modulation on the active carrier.

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF5003/D

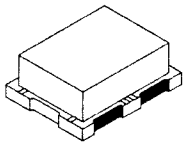
The RF MOSFET Line
RF Power Field Effect Transistor
N–Channel Enhancement–Mode

MRF5003

3.0 W, 7.5 V, 512 MHz
N–CHANNEL
BROADBAND
RF POWER FET

The MRF5003 is designed for broadband commercial and industrial applications at frequencies to 520 MHz. The high gain and broadband performance of this device makes it ideal for large–signal, common source amplifier applications in 7.5 Volt and 12.5 Volt mobile, portable, and base station FM equipment.

- Guaranteed Performance at 512 MHz, 7.5 Volts
  Output Power = 3.0 Watts
  Power Gain = 9.5 dB
  Efficiency = 45%
- Characterized with Series Equivalent Large–Signal Impedance Parameters
- S–Parameter Characterization at High Bias Levels
- Excellent Thermal Stability
- All Gold Metal for Ultra Reliability
- Capable of Handling 20:1 VSWR, @ 15.5 Vdc, 512 MHz, 2.0 dB Overdrive
- Suitable for 12.5 Volt Applications
- True Surface Mount Package
- Available in Tape and Reel by Adding R1 Suffix to Part Number.
  R1 Suffix = 500 Units per 16 mm, 7 inch Reel.
- Circuit board photomaster available upon request by contacting
  RF Tactical Marketing in Phoenix, AZ.



CASE 430–01, STYLE 2

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Drain–Source Voltage	V_{DSS}	36	Vdc
Drain–Gate Voltage (R_{GS} = 1.0 Meg Ohm)	V_{DGR}	36	Vdc
Gate–Source Voltage	V_{GS}	±20	Vdc
Drain Current — Continuous	I_D	1.7	Adc
Total Device Dissipation @ T_C = 25°C Derate above 25°C	P_D	12.5 0.07	Watts W/°C
Storage Temperature Range	T_{stg}	–65 to +150	°C
Operating Junction Temperature	T_J	200	°C

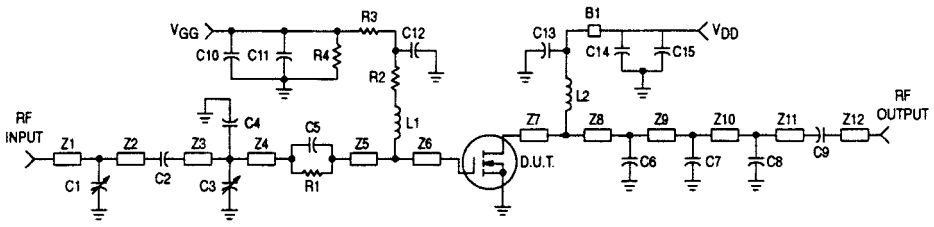
THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	14	°C/W

NOTE – **CAUTION** – MOS devices are susceptible to damage from electrostatic charge. Reasonable precautions in handling and packaging MOS devices should be observed.

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted.)

Characteristic		Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS						
Drain–Source Breakdown Voltage (V_{GS} = 0, I_D = 2.5 mAdc)		$V_{(BR)DSS}$	36	—	—	Vdc
Zero Gate Voltage Drain Current (V_{DS} = 15 Vdc, V_{GS} = 0)		I_{DSS}	—	—	1.0	mAdc
Gate–Source Leakage Current (V_{GS} = 20 Vdc, V_{DS} = 0)		I_{GSS}	—	—	1.0	μAdc
ON CHARACTERISTICS						
Gate Threshold Voltage (V_{DS} = 10 Vdc, I_D = 5.0 mAdc)		$V_{GS(th)}$	1.25	2.25	3.5	Vdc
Drain–Source On–Voltage (V_{GS} = 10 Vdc, I_D = 0.5 Adc)		$V_{DS(on)}$	—	—	0.375	Vdc
Forward Transconductance (V_{DS} = 10 Vdc, I_D = 0.5 Adc)		g_{fs}	0.6	—	—	mho
DYNAMIC CHARACTERISTICS						
Input Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{iss}	—	16.5	—	pF
Output Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{oss}	—	37	—	pF
Reverse Transfer Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{rss}	3.5	4.4	5.4	pF
FUNCTIONAL TESTS (In Motorola Test Fixture)						
Common–Source Amplifier Power Gain (V_{DD} = 7.5 Vdc, P_{out} = 3.0 W, I_{DQ} = 50 mA)	f = 512 MHz f = 175 MHz	G_{ps}	 9.5 —	 10.5 15	 — —	dB
Drain Efficiency (V_{DD} = 7.5 Vdc, P_{out} = 3.0 W, I_{DQ} = 50 mA)	f = 512 MHz f = 175 MHz	h	 45 —	 50 55	 — —	%



C1, C3, C7, C8	0 to 20 pF Johanson	Z1	0.350″ x 0.08″ Microstrip
C2, C9	56 pF, 100 mil Chip	Z2	0.190″ x 0.08″ Microstrip
C4	10 pF, 100 mil Chip	Z3	0.800″ x 0.08″ Microstrip
C5	47 pF, Miniature Clamped Mica Capacitor	Z4	0.380″ x 0.08″ Microstrip
C6	22 pF, 100 mil Chip	Z5	0.150″ x 0.08″ Microstrip
C10, C15	10 µF, 50 V, Electrolytic	Z6	0.285″ x 0.08″ Microstrip
C11, C14	0.1 µF, Capacitor	Z7	0.340″ x 0.08″ Microstrip
C12	1000 pF, 100 mil Chip	Z8	0.070″ x 0.08″ Microstrip
C13	160 pF, 100 mil Chip	Z9	0.280″ x 0.08″ Microstrip
R1	35 Ω, 1/4 W Carbon	Z10	0.840″ x 0.08″ Microstrip
R2	30 Ω, 0.1 W Chip	Z11	0.180″ x 0.08″ Microstrip
R3	1.0 kΩ, 0.1 W Chip	Z12	0.600″ x 0.08″ Microstrip
R4	1.0 MΩ, 1/4 W Carbon	L1	7 Turns, 0.076″ ID, #24 AWG Enamel
B1	Fair Rite Products Short Ferrite Bead (2743021446)	L2	5 Turns, 0.126″ ID, #20 AWG Enamel
Board — Glass Teflon®, 31 mils		Input/Output Connectors — Type N	

Note: Plated ceramic part locators (0.1″ x 0.15″) soldered onto Z6 and Z7.

Figure 1. 512 MHz Narrowband Test Circuit

TYPICAL CHARACTERISTICS

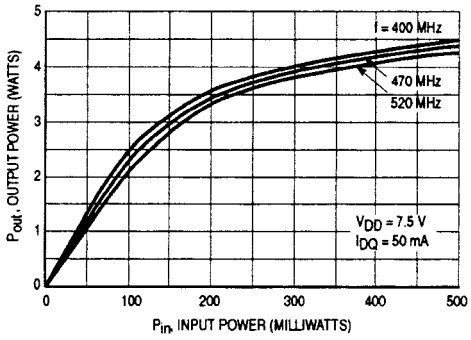


Figure 2. Output Power versus Input Power

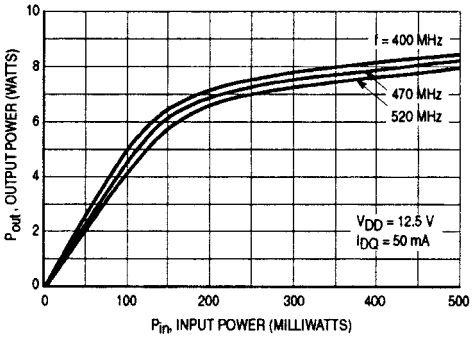


Figure 3. Output Power versus Input Power

TYPICAL CHARACTERISTICS

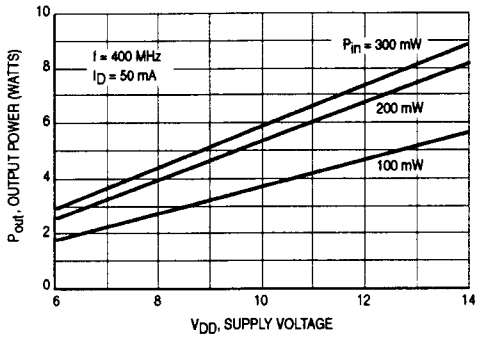


Figure 4. Output Power versus Supply Voltage

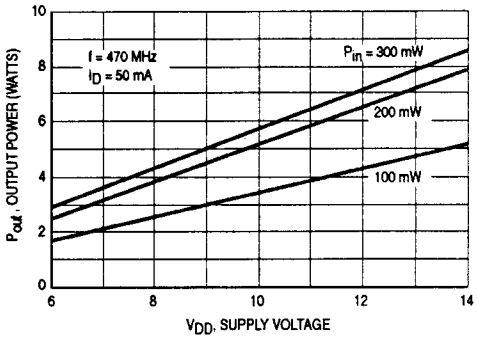


Figure 5. Output Power versus Supply Voltage

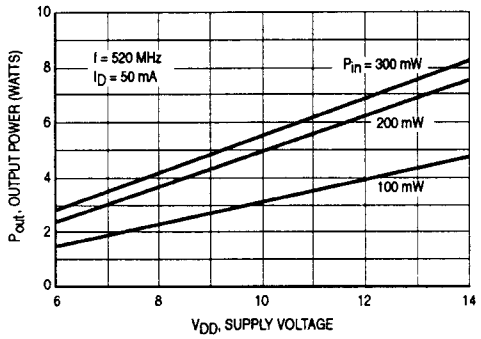


Figure 6. Output Power versus Supply Voltage

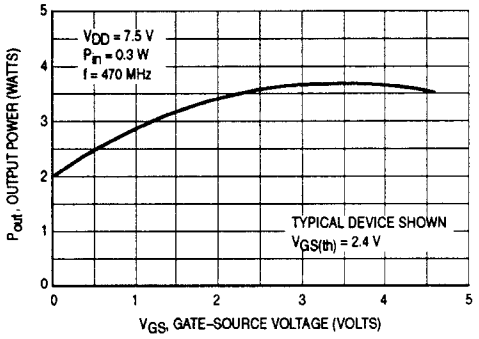


Figure 7. Output Power versus Gate Voltage

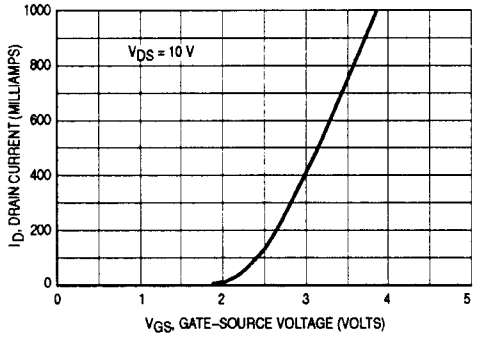


Figure 8. Drain Current versus Gate Voltage (Typical Device Shown)

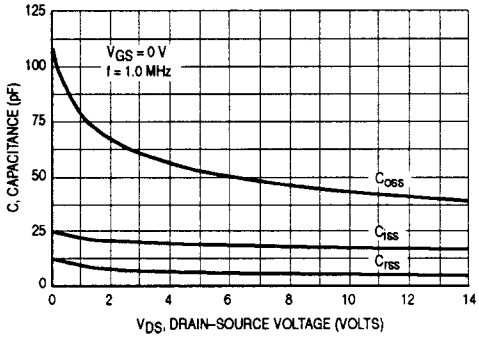


Figure 9. Capacitance versus Voltage

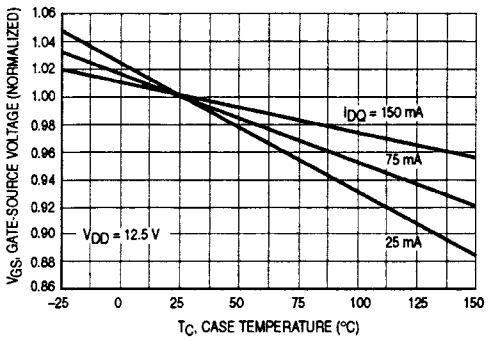


Figure 10. Gate–Source Voltage versus Case Temperature

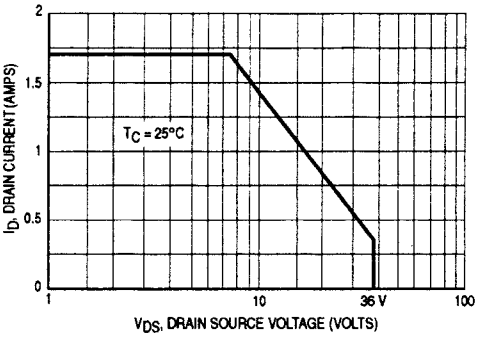
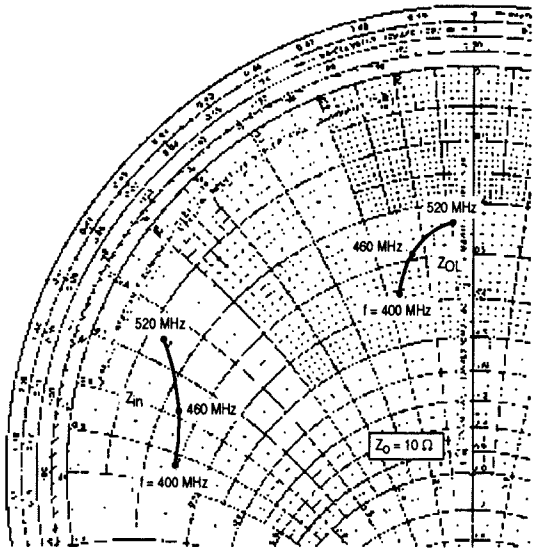


Figure 11. Maximum Rated Forward Biased Safe Operating Area



V_{DD} = 7.5 V, I_{DQ} = 50 mA, P_{out} = 3.0 W

f MHz	Z_{in} Ohms	Z_{OL}^* Ohms
400	2.8 – j9.2	3.6 – j1.7
430	2.7 – j8.5	3.3 – j1.5
460	2.5 – j7.8	2.7 – j1.1
490	2.0 – j7.2	2.5 – j0.8
520	1.3 – j6.5	2.4 – j0.5

Z_{in} = Conjugate of source impedance with parallel 35 Ω resistor and 47 pF capacitor in series with gate.

Z_{OL}^* = Conjugate of the load impedance at given output power, voltage, frequency, and η_D > 50%.

Note: Z_{ol}^* was chosen based on tradeoffs between gain, drain efficiency, and device stability.

Figure 12. Series Equivalent Input and Output Impedance

Table 1. Common Source Scattering Parameters (V_{DS} = 10 V)

I_D = 50 mA

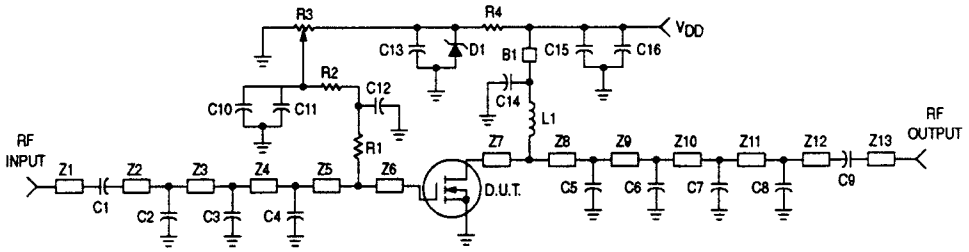
f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.69	−90	10.8	117	0.07	29	0.74	−119
100	0.58	−120	6.0	96	0.08	10	0.78	−146
200	0.58	−139	3.0	75	0.08	−7	0.81	−161
300	0.64	−147	1.9	61	0.07	−16	0.84	−166
400	0.70	−152	1.3	50	0.06	−21	0.86	−169
500	0.75	−157	0.99	41	0.05	−24	0.88	−172
700	0.82	−165	0.61	28	0.03	−15	0.92	−176
850	0.86	−171	0.45	21	0.02	13	0.94	−179
1000	0.89	−176	0.34	16	0.02	47	0.95	178

I_D = 500 mA

f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.76	−124	15.0	109	0.04	23	0.76	−151
100	0.72	−150	7.9	94	0.04	12	0.81	−165
200	0.72	−163	4.0	80	0.04	6	0.83	−172
300	0.73	−168	2.6	71	0.04	5	0.84	−175
400	0.75	−171	1.9	62	0.04	7	0.85	−176
500	0.77	−173	1.5	55	0.03	12	0.86	−178
700	0.81	−177	0.97	42	0.03	29	0.89	−180
850	0.84	−180	0.75	35	0.03	44	0.90	178
1000	0.86	177	0.60	29	0.04	55	0.92	176

I_D = 1.0 A

f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.80	−125	14.6	110	0.04	23	0.75	−155
100	0.76	−150	7.8	95	0.04	10	0.81	−167
200	0.76	−164	3.9	81	0.04	1	0.83	−173
300	0.77	−169	2.6	71	0.04	−3	0.84	−175
400	0.79	−172	1.9	63	0.03	−5	0.85	−176
500	0.80	−174	1.4	56	0.03	−5	0.86	−177
700	0.83	−178	0.95	43	0.03	−1	0.88	−179
850	0.85	179	0.73	35	0.02	9	0.90	179
1000	0.87	177	0.58	28	0.02	22	0.91	178



C1, C9	100 pF 100 mil Chip
C2	16 pF, 100 mil Chip
C3	24 pF, 100 mil Chip
C4	68 pF, 100 mil Chip
C5	51 pF, 100 mil Chip
C6	39 pF, 100 mil Chip
C7	6.2 pF, 100 mil Chip
C8	9.1 pF, 100 mil Chip
C10, C15	39000 pF, 100 mil Chip
C11, C16	10 µF, 50 V Electrolytic
C12	10000 pF, 100 mil Chip
B1	Fair Rite Products Short Ferrite Bead (2743021446)
C13	0.1 µF, 100 mil Chip
C14	160 pF, 100 mil Chip
R1	43 Ω, 0.1 W Chip Resistor
R2	1000 Ω, 0.1 W Chip Resistor
R3	10 kΩ Potentiometer
R4	3000 Ω, 0.1 W Chip Resistor
L1	5 Turns, 0.126″ ID, #20 AWG Enamel
Z1 to Z13	See Photomaster
D1	1N4734 Motorola Zener
Board — G10, 1/32″	
Input/Output Connectors — SMA	

Figure 13. Schematic of Broadband Demonstration Amplifier

PERFORMANCE CHARACTERISTICS OF BROADBAND DEMONSTRATION AMPLIFIER

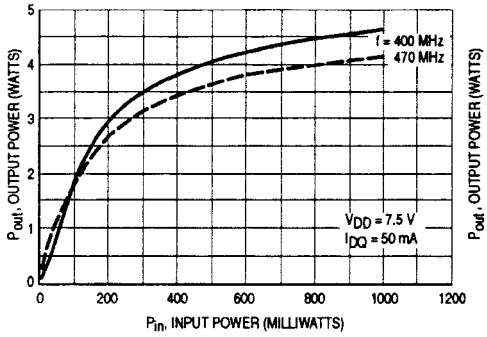


Figure 14. Output Power versus Input Power

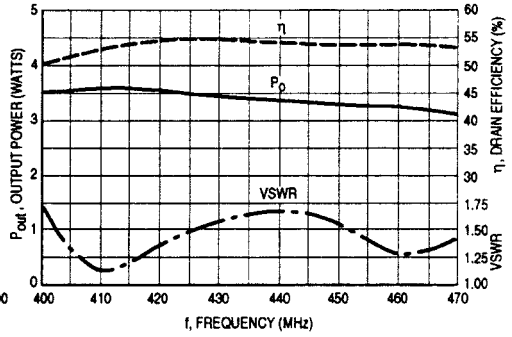


Figure 15. Output Power, Drain Efficiency and VSWR versus Frequency

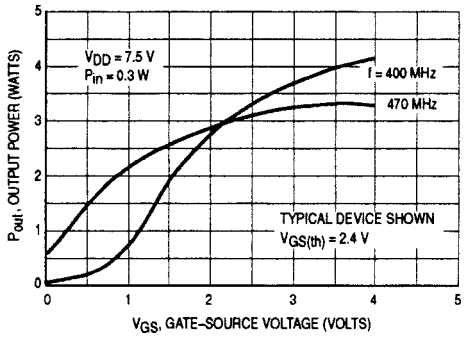


Figure 16. Output Power versus Gate Voltage

DESIGN CONSIDERATIONS

The MRF5003 is a common–source, RF power, N–Channel enhancement mode, Metal–Oxide Semiconductor Field–Effect Transistor (MOSFET). Motorola RF MOSFETs feature a vertical structure with a planar design. Motorola Application Note AN211A, "FETs in Theory and Practice", is suggested reading for those not familiar with the construction and characteristics of FETs.

This surface mount packaged device was designed primarily for VHF and UHF power amplifier applications. Manufacturability is improved by utilizing the tape and reel capability for fully automated pick and placement of parts.

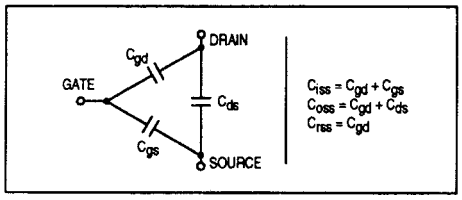
The major advantages of RF power MOSFETs include high gain, simple bias systems, relative immunity from thermal runaway, and the ability to withstand severely mismatched loads without suffering damage.

MOSFET CAPACITANCES

The physical structure of a MOSFET results in capacitors between all three terminals. The metal oxide gate structure determines the capacitors from gate–to–drain (C_{gd}), and gate–to–source (C_{gs}). The PN junction formed during fabrication of the RF MOSFET results in a junction capacitance from drain–to–source (C_{ds}). These capacitances are characterized as input (C_{iss}), output (C_{oss}) and reverse transfer (C_{rss}) capacitances on data sheets. The relationships between the inter–terminal capacitances and those given on data sheets are shown below. The C_{iss} can be specified in two ways:

1. Drain shorted to source and positive voltage at the gate.
2. Positive voltage of the drain in respect to source and zero volts at the gate.

In the latter case, the numbers are lower. However, neither method represents the actual operating conditions in RF applications.



DRAIN CHARACTERISTICS

One critical figure of merit for a FET is its static resistance in the full–on condition. This on–resistance, $R_{DS(on)}$, occurs in the linear region of the output characteristic and is specified at a specific gate–source voltage and drain current. The drain–source voltage under these conditions is termed $V_{DS(on)}$. For MOSFETs, $V_{DS(on)}$ has a positive temperature coefficient at high temperatures because it contributes to the power dissipation within the device.

GATE CHARACTERISTICS

The gate of the RF MOSFET is a polysilicon material, and is electrically isolated from the source by a layer of oxide. The input resistance is very high — on the order of $10^9\ \Omega$ — resulting in a leakage current of a few nanoamperes.

Gate control is achieved by applying a positive voltage to the gate greater than the gate–to–source threshold voltage, $V_{GS(th)}$.

Gate Voltage Rating — Never exceed the gate voltage rating. Exceeding the rated V_{GS} can result in permanent damage to the oxide layer in the gate region.

Gate Termination — The gates of these devices are essentially capacitors. Circuits that leave the gate open–circuited or floating should be avoided. These conditions can result in turn–on of the devices due to voltage build–up on the input capacitor due to leakage currents or pickup.

Gate Protection — These devices do not have an internal monolithic zener diode from gate–to–source. If gate protection is required, an external zener diode is recommended with appropriate RF decoupling.

Using a resistor to keep the gate–to–source impedance low also helps dampen transients and serves another important function. Voltage transients on the drain can be coupled to the gate through the parasitic gate–drain capacitance. If the gate–to–source impedance and the rate of voltage change on the drain are both high, then the signal coupled to the gate may be large enough to exceed the gate–threshold voltage and turn the device on.

DC BIAS

Since the MRF5003 is an enhancement mode FET, drain current flows only when the gate is at a higher potential than the source. See Figure 8 for a typical plot of drain current versus gate voltage. RF power FETs operate optimally with a quiescent drain current (I_{DQ}), whose value is application dependent. The MRF5003 was characterized at I_{DQ} = 50 mA, which is the suggested value of bias current for typical applications. For special applications such as linear amplification, I_{DQ} may have to be selected to optimize the critical parameters.

The gate is a dc open circuit and draws no current. Therefore, the gate bias circuit may generally be just a simple resistive divider network. Some special applications may require a more elaborate bias system.

GAIN CONTROL

Power output of the MRF5003 may be controlled from its rated value down to zero (negative gain) with a low power dc control signal, thus facilitating applications such as manual gain control, ALC/AGC and modulation systems. Figure 16 is an example of output power variation with gate–source bias voltage. This characteristic is very dependent on frequency and load line.

MOUNTING

The specified maximum thermal resistance of 14°C/W assumes a majority of the 0.100″ x 0.200″ source contact on the back side of the package is in good contact with an appropriate heat sink. In the test fixture shown in Figure 1, the device is clamped directly to a copper pedestal. In the demonstration amplifier, the device was mounted on top of the G10 circuit board and heat removal was accomplished through several solder filled plated through holes. As with all RF power devices, the goal of the thermal design should be to minimize the temperature at the back side of the package.

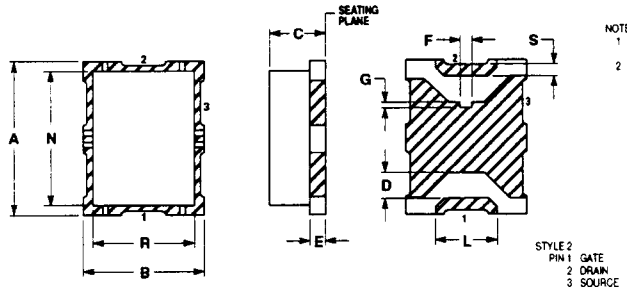
AMPLIFIER DESIGN

Impedance matching networks similar to those used with bipolar transistors are suitable for the MRF5003. For examples see Motorola Application Note AN721, "Impedance Matching Networks Applied to RF Power Transistors." Both small–signal S–parameters and large–signal impedances are provided. While the S–parameters will not produce an exact design solution for high power operation, they do yield a good first approximation. This is an additional advantage of RF power MOSFETs.

Since RF power MOSFETs are triode devices, they are not unilateral. This coupled with the very high gain of the MRF5003 yield a device capable of self oscillation. Stability may be achieved by techniques such as drain loading, input shunt resistive loading, or output to input feedback. Different stabilizing techniques were applied to the test fixture and demonstration amplifiers. The RF test fixture implements a parallel resistor and capacitor in series with the gate while the demonstration amplifier utilizes a 43 Ω shunt resistor from gate to ground. Both circuits have a load line selected for a higher efficiency, lower gain, and more stable operating region.

Two port stability analysis with the MRF5003 S–parameters provides a useful tool for selection of loading or feedback circuitry to assure stable operation. See Motorola Application Note AN215A, "RF Small–Signal Design Using Two–Port Parameters," for a discussion of two port network theory and stability.

PACKAGE DIMENSIONS



NOTES
1 DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2 CONTROLLING DIMENSION: INCH.

	INCHES		MILLIMETERS	
DIM	MIN	MAX	MIN	MAX
A	0.260	0.270	6.60	6.86
B	0.200	0.210	5.08	5.33
C	0.090	0.104	2.29	2.64
D	0.040	0.050	1.02	1.27
E	0.022	0.028	0.56	0.71
F	0.015	0.025	0.38	0.64
G	0.005	0.015	0.13	0.38
L	0.100	0.110	2.54	2.79
N	0.226	0.236	5.74	5.99
R	0.166	0.176	4.22	4.47
S	0.025	0.035	0.64	0.89

CASE 430–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters can and do vary in different applications. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and Ⓜ are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

2-3-7 GaAs MESFETs

Introduction. More than twenty years ago, for RF applications, the JFET was the dominant FET. As mentioned earlier, its major disadvantage is the fact that the f_T cutoff frequency, defined as

$$f_T = g_m / 2\pi C_{GS} \tag{2-262}$$

limits the operating range of this semiconductor device to 1 GHz at most. Also, the JFET's temperature sensitivity and loose tolerances in transconductance and other RF-related values made it difficult to use. Finally, no RFICs made use of this technology. While at the moment silicon-based MOS technology probably is the most cost-effective and is produced in the largest quantities, it does not rival the metal-semiconductor FET designed on GaAs technology. Figure 2-115 shows the structure of the silicon MOSFET, silicon JFET, and the GaAs MESFET.

Large-Signal Behavior of GaAs MESFETs. The mathematics for the large-signal behavior of the GaAsFET is basically quite similar to that for the JFET; however, the computations for JFET channel current, diode currents, and capacitance are much simpler than for the GaAsFET. Temperature effects are embedded in all the equations for all the transistors mentioned so far. In our opinion, the modified Materka model used by Ansoft and a large number of users is the most complete one; however, the following models are supported by most CAD tools:

- Chalmers (Angelov)
- Curtice–Ettenberg cubic
- Curtice quadratic
- IAF (Berroth)
- ITT PFET and TFET
- Modified Materka–Kacprzak
- Physics-based MESFET
- Raytheon (Statz)
- TriQuint (TOM-1, TOM-2, and TOM-3)
- With some restrictions, metal-insulator-semiconductor FETs (MISFETs), modulation-doped FETs (MODFETs), and high-electron-mobility transistors (HEMTs)

The main advantages of GaAsFET technology derive from the fact that it uses a metal-semiconductor junction with a barrier voltage of 0.8 V, its input capacitance is typically less than 0.2 pF, and the reverse feedback capacitance is less than 0.02 pF, or roughly 10% of the input capacitance. As a result of this, $f_{\max}$ is approximately five times higher than f_T. Table 2-16 shows a comparison of silicon BJT and GaAsFET technologies. While $f_{\max}$ for the bipolar transistor is

$$f_{\max} \simeq \sqrt{\frac{f_t}{8\pi r'_{bb} C_C}} \tag{2-263}$$

the $f_{\max}$ determination for the GaAsFET is given by

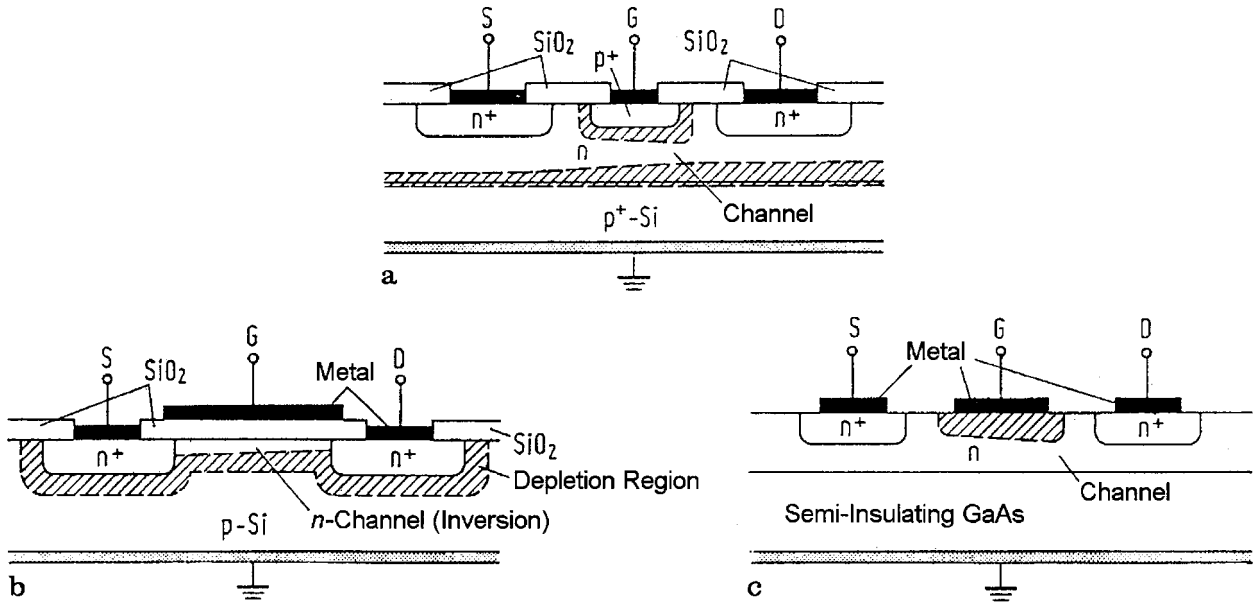


Figure 2-115 This picture is provided to differentiate between the structure of the three technologies. (a) The JFET and (b) the MOSFET are based on silicon; (c) the GaAsFET is based on semi-insulating GaAs. While the MOS transistor has silicon dioxide (SiO_2) as gate insulation, both the JFET and GaAs MESFET have a diode from gate to source, with a barrier voltage of 0.65 V for silicon and 0.8 V for GaAs.

Table 2-16 Comparison of Si BJT, SiGe HBT, and GaAsFET technologies

	Si Bipolar	SiGe HBT	GaAs FET
f_T	25 GHz	100 GHz	22 GHz
$f_{\max}$	40 GHz	200 GHz	110 GHz
Features	Low cost, low 1/*f* noise (5 kHz)	Low 1/*f* noise, very low distortion, high cost, IC only	Highest flexibility, lowest NF_0, well established

$$f_{\max} = \frac{f_T}{2}\sqrt{\frac{R_0}{R_i + R_s + R_g}} \tag{2-264}$$

As a sample calculation,

$$f_{\max}\ \frac{21.9\text{ GHz}}{2}\sqrt{\frac{450}{1 + 1.5 + 2}} = 110\text{ GHz} \tag{2-265}$$

The three major drawbacks of the GaAsFET are:

1. Much higher flicker corner frequency (somewhere between 10 and 100 MHz), probably due to a lack of surface passivation.
2. Much higher output conductance. This tends to load down any circuit connected to the drain. On the other hand, since the transconductance is quite high for even low currents, these devices have very high gains at low frequencies, which can make them quite unstable. In the saturated mode it is not uncommon to find a drain–source resistance of 100–500 Ω, while BJTs and JFETs offer values of several kilohms and higher.
3. Because of the very high flicker corner frequency (from 10 to 100 MHz), MESFETs are really not useful for low-noise mixers and oscillators, and unless there are no devices available in the frequency range above 30 GHz, they should be avoided for these applications.

As to the MESFET's construction and dc properties, Figure 2-116 shows a MESFET's cross section and dc *I*–*V* characteristics.

It was outlined in the beginning that while the GaAsFET is a close relative to the JFET and MOSFET, its actual behavior was found to be described best by a set of analytic equations. The first such model was the one by Curtice in the form of quadratic and cubic models, but it does not have enough derivatives to give enough insight into subtle things such as third- and higher-order intermodulation distortion and accurate harmonic generation. Other researchers have addressed various areas, but we still find that the Materka model has the best approximation, especially using the two-transistor approach outlined later [20, 21]. The large-signal topology for all FETs consists of an intrinsic model with some extrinsic parameters, further complicated by the package, as shown by Figures 2-117 and 2-118. Table 2-17 lists their keywords. The actual intrinsic model and its parameter definition depend on the particular model and since designs using GaAsFET will always be done using CAD tools, we will not go into any details of the equations but will list them. They are not dissimilar from the JFET and MOSFET equations.

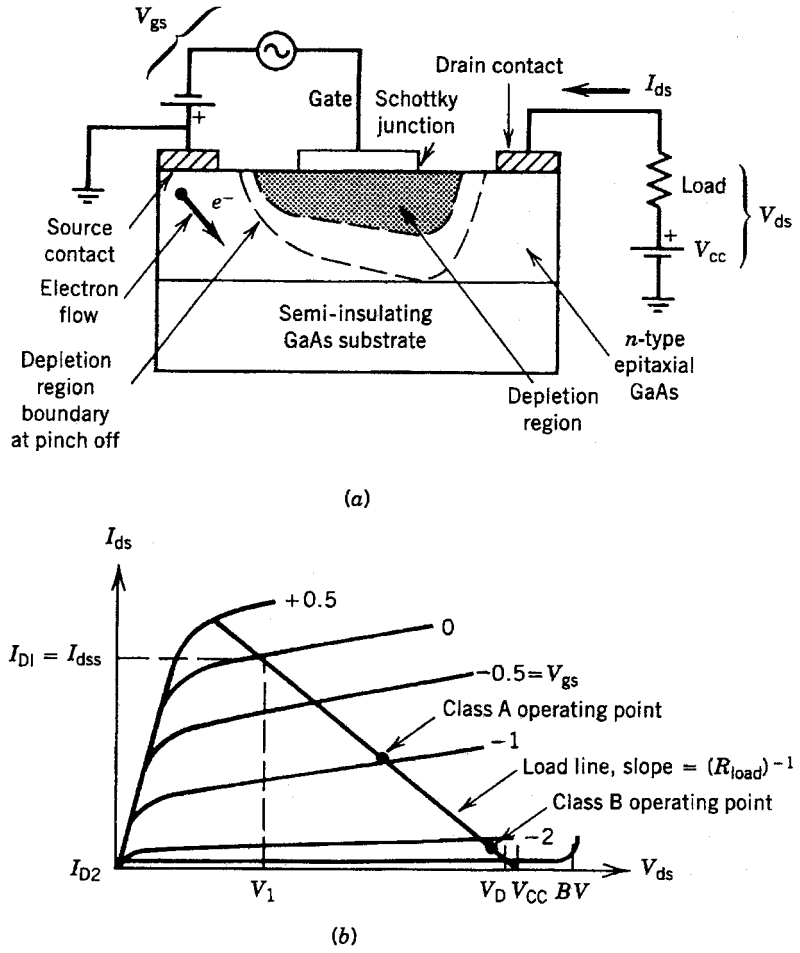


Figure 2-116 (a) Cross section and bias circuit and (b) dc *I–V* curve, including ac load line, for a MESFET.

The Modified Materka-Kacprzak Model Figure 2-119 shows the intrinsic model of the Materka FET. Table 2-18 lists its keywords.

Large-Signal Equations

Device Equations

V_{gsi} = Intrinsic gate–source voltage
V_{dsi} = Intrinsic drain–source voltage
V_1 = Voltage across C_{GS} and R_i
V_{gdi} = Intrinsic gate–drain voltage
$V_T = kT_J/q$ (thermal voltage)
k = Boltzmann's constant
q = Electron charge
T_J = Analysis temperature (kelvins)

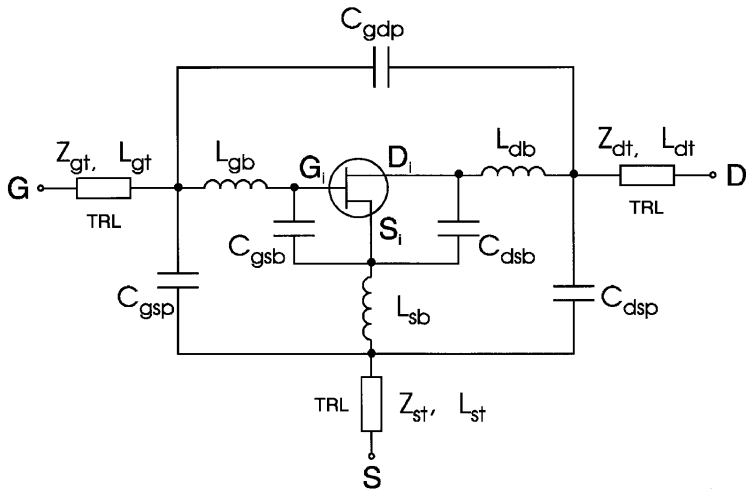


Figure 2-117 MESFET extrinsic model.

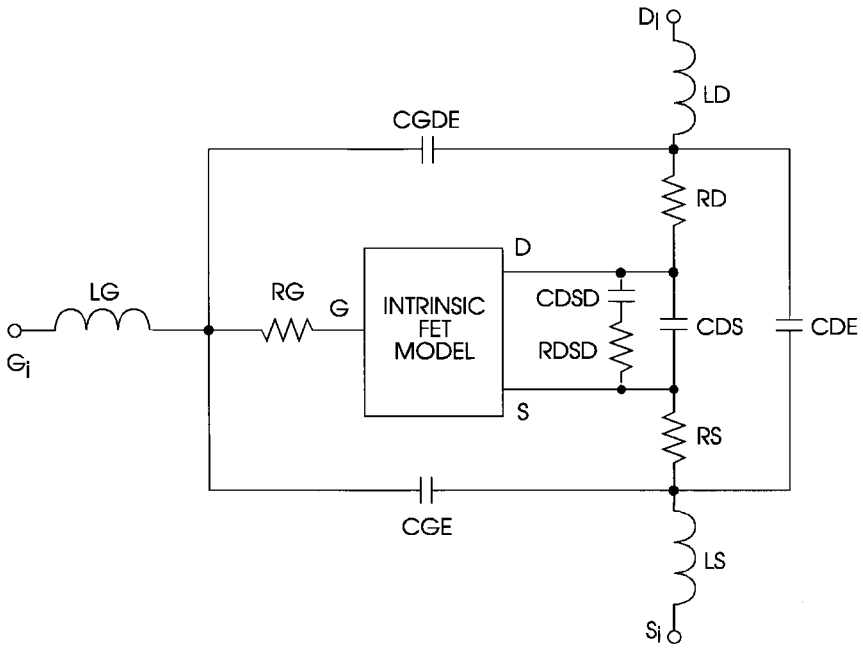


Figure 2-118 MESFET package model.

Table 2-17 Large-signal MESFET model keywords

Keyword	Description	Unit	Default
	Extrinsic Parameters		
RG	Gate bulk and ohmic resistance	ohm	0.0
RD	Drain bulk and ohmic resistance	ohm	0.0
RS	Source bulk and ohmic resistance	ohm	0.0
LG	Gate-lead inductance (metallization)	henry	0.0
LD	Drain-lead inductance (metallization)	henry	0.0
LS	Source-lead inductance (via)	henry	0.0
CDS	Drain–source capacitance	farad	0.0
CDSD	Low-frequency trapping capacitor	farad	0.0
RDSD	Channel trapping resistance	ohm	∞
CGE	Gate–source electrode capacitance	farad	0.0
CDE	Drain–source electrode capacitance	farad	0.0
CGDE	Gate–drain electrode capacitance	farad	0.0
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate-to-source package capacitance	farad	0.0
CDSP	Drain-to-source package capacitance	farad	0.0
CGDP	Gate-to-drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for $\varepsilon_r = 1$	meter	0.0

Channel Current

$$I_{ds} = \mathrm{IDSS}\left(1 + \mathrm{SS}\,\frac{V_{dsi}}{\mathrm{IDSS}}\right)\left(1 - \frac{V_{gsi}(t-T)}{\mathrm{VP0} + \mathrm{GAMA}V_{dsi}}\right)^{(\mathrm{E}+\mathrm{KE}V_{gsi}(t-T))}$$

$$\times \tanh\left(\frac{\mathrm{SL}V_{dsi}}{\mathrm{IDSS}(1 - \mathrm{KG}V_{gsi}(t-T))}\right)$$

Diode

$$I_{gd} = I_{gdc} - \begin{cases} \mathrm{IB0}\,\exp[-\mathrm{AFAB}(V_{gdi} + \mathrm{VBC})] \\ \dfrac{\mathrm{GMAX}}{4}\,\{\tanh[\mathrm{K1D}(V_{gsi} - \mathrm{K2D})] - 1\}\,[V_{gdi} + \mathrm{VBC} - \sqrt{(V_{gdi} + \mathrm{VBC})^2 + \mathrm{K3D}}] \end{cases}$$

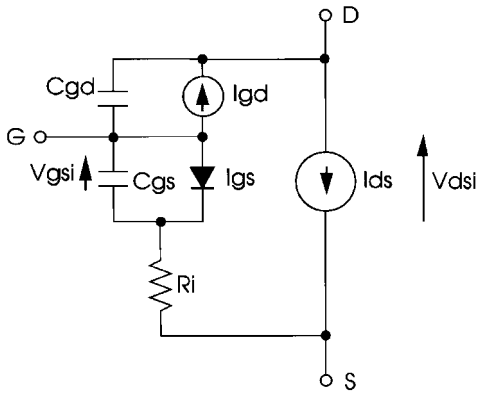


Figure 2-119 Intrinsic model of the modified Materka–Kacprzak MESFET.

where

$$I_{gdc} = \text{IG0}[\exp(\text{AFAB}V_{gdi}) - 1]$$

Channel Resistance

$$R_i = \begin{cases} \text{R10}(1 - \text{KR}V_{gsi}), & \text{KR}V_{gsi} < 1.0 \\ 0, & \text{KR}V_{gsi} \geq 1.0 \end{cases}$$

Capacitance Model

$$C_{gs} = \text{CGS0}\frac{F1F2}{\sqrt{1 - \dfrac{V_{\text{new}}}{\text{VBI}}}} + \text{CGD0}\ F3$$

$$C_{gd} = \text{CGS0}\frac{F1F3}{\sqrt{1 - \dfrac{V_{\text{new}}}{\text{VBI}}}} + \text{CGD0}\ F2$$

where

$$F1 = \frac{1}{2}\left(1 + \frac{V_{\text{eff}} - V_T}{\sqrt{(V_{\text{eff}} - V_T)^2 + \delta^2}}\right)$$

$$F2 = \frac{1}{2}\left(1 + \frac{V_{gsi} - V_{gdi}}{\sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}}\right)$$

Table 2-18 Materka–Kacprzak MESFET model keywords

Keyword	Description	Unit	Default
	Area, Noise, and Name		
AREA	Area multiplier		1.0
KFN	Flicker noise coefficient (Materka model only)[a]		0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
	Channel Current Model		
IDSS	Drain saturation current for $V_{GS} = 0$	ampere	0.1
VP0	Pinchoff voltage for $V_{DS} = 0$	volt	–2.0
GAMA	Voltage slope parameter of pinchoff voltage	/volt	0.0
E	Constant part of power-law parameter		2.0
KE	Dependence of power law on V_{GS}	/volt	0.0
SL	Slope of the $V_{GS} = 0$ drain characteristic in the linear region	ampere/volt	0.15
KG	Drain dependence on V_{GS} in the linear region	/volt	0.0
SS	Slope of the drain characteristic in the saturated region	ampere/volt	0.0
T	Channel transit-time delay	second	0.0
IG0	Diode saturation current	ampere	0
AFAG	Slope factor of forward diode current	/volt	38.696
IB0	Breakdown saturation current	ampere	0
AFAB	Slope factor of breakdown current	/volt	0
VBC	Breakdown voltage	volt	∞
GMAX	Breakdown conductance	ampere/volt	0
K1D	Fitting parameter	/volt	0
K2D	Fitting parameter	volt	0
K3D	Fitting parameter	$volt^2$	0
R10	Intrinsic channel resistance for $V_{GS} = 0$	ohm	0.0
KR	Slope factor of intrinsic channel resistance	/volt	0.0
	Materka Capacitance Model		
C10	Gate–source Schottky barrier capacitance for $V_{GS} = 0$	farad	0.0
K1	Slope parameter of gate–source capacitance	/volt	1.25
MGS	Gate–source grading coefficient		0.5
C1S	Constant parasitic component of gate–source capacitance	farad	0.0
CF0	Gate–drain feedback capacitance for $V_{GD} = 0$	farad	0.0
KF	Slope parameter of gate–drain feedback capacitance	/volt	1.25
MGD	Gate–drain grading coefficient		0.5
FCC	Forward-bias depletion capacitance coefficient		0.8

[a]The flicker noise parameter of the Materka model is KFN so as not to conflict with the KF parameter in the capacitance model.

$$F3 = \frac{1}{2}\left(1 - \frac{V_{gsi} - V_{gdi}}{\sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}}\right)$$

$$V_{\text{new}} = \begin{cases} A1, & A1 < \text{VMAX} \\ \text{VMAX}, & A1 \geq \text{VMAX} \end{cases}$$

$$A1 = \frac{1}{2}\left(V_{\text{eff}} + V_T + \sqrt{(V_{\text{eff}} - V_T)^2 + \delta^2}\right)$$

$$V_{\text{eff}} = \frac{1}{2}\left(V_{gsi} + V_{gdi} + \sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}\right)$$

$$V_T = \text{VP0} + \text{GAMA}\, V_{dsi}$$

$$\delta = 0.2$$

Some of the modifications to the Materka model have been done by Ansoft under various Department of Defense contracts, and by Raytheon and Texas Instruments under similar contracts, with Ansoft being a subcontractor.

The most relevant equation is really the channel current. Its derivatives are largely responsible for the accuracy of the intermodulation distortion, power-added efficiency, and, of course, its dc *I–V* curves.

Enhancement/Depletion FETs To make the designer's life more difficult, it turns out that there are two types of GaAsFETs:

1. *Depletion FETs (DFETs).* Most similar to the JFET; here V_G must be negative to control the device. They are the most commonly produced and are the FET type most referred to in this book.

On one hand:

- They require a negative gate voltage with respect to the source.
- Self-bias allows operation from a single supply voltage.

On the other hand:

- For low-voltage operation, a negative voltage generator may be required.
- Supply voltage must be doubled to accommodate full-swing operation.

2. *Enhancement FETs (EFETs).* Most similar to the MOSFET; here V_G must be positive to bring life to the device. Practically speaking, EFETs are used mostly in integrated circuits; they're typically not available in discrete, packaged form.

On one hand:

- They need only positive supply for biasing.
- They provide higher g_m/mA (for same device width)—5.1 mS versus 3.9 mS at 8 mA.

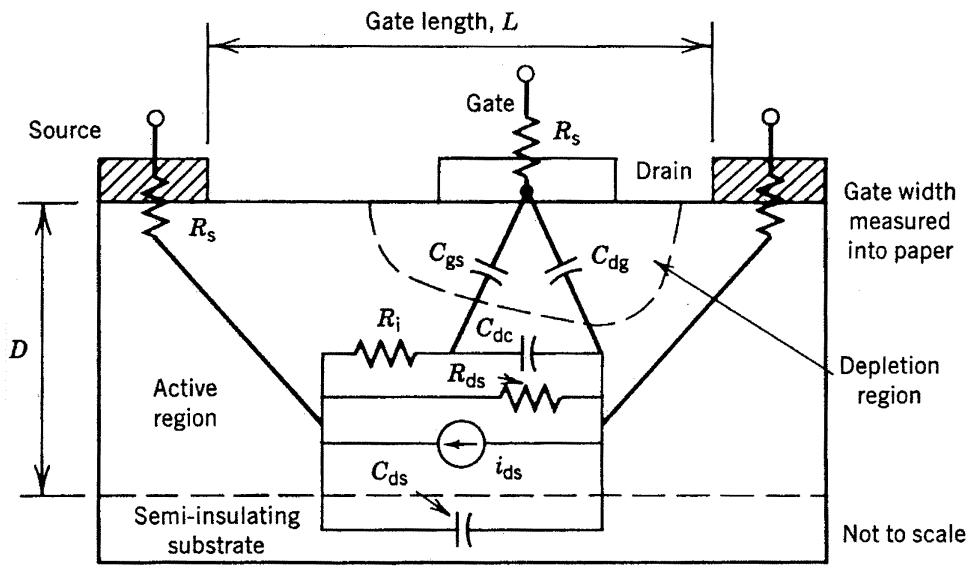


Figure 2-120 Location of lumped-element components for a MESFET.

- They are good for low-power LNAs, giving slightly better NF that DFETs, an NF of 1 dB at 1 GHz, and $I_{dd} < 10$ mA.

On the other hand:

- They have a very limited gate bias range (V_{GS} between 0.15 and 0.7 V).
- The gate conduction degrades NF and input impedance.
- The gate capacitance is higher than that of DFETs.
- The linearity is not as good as that of DFETs.

With today's technologies, all GaAs devices are *n*-channel; we have not seen any *p*-channels yet. More information about biasing will be given in the next chapter.

Figure 2-120 is a lumped-element, two-part equivalent circuit of a MESFET showing the location of lumped-element components.

2-3-8 Small-Signal GaAs MESFET Model

Figure 2-121 shows the applicable linear equivalent circuit for a MESFET and Table 2-19 lists its keywords. As with the MOS transistors, there is a gallium-arsenide dual-gate MOSFET available that is used mostly in special circuits, such as preamplifiers and mixers, whose IF has to be significantly higher than the flicker corner frequency, for example, higher than 200 MHz.

The following two datasheets* show typical GaAsFETs, both low-noise and medium-power. Such devices are, of course, available from a large variety of manufacturers—to name at least two, Siemens and the Nippon Electric Company, which is represented in the United States by California Eastern Laboratories.

*Reproduced with permission.

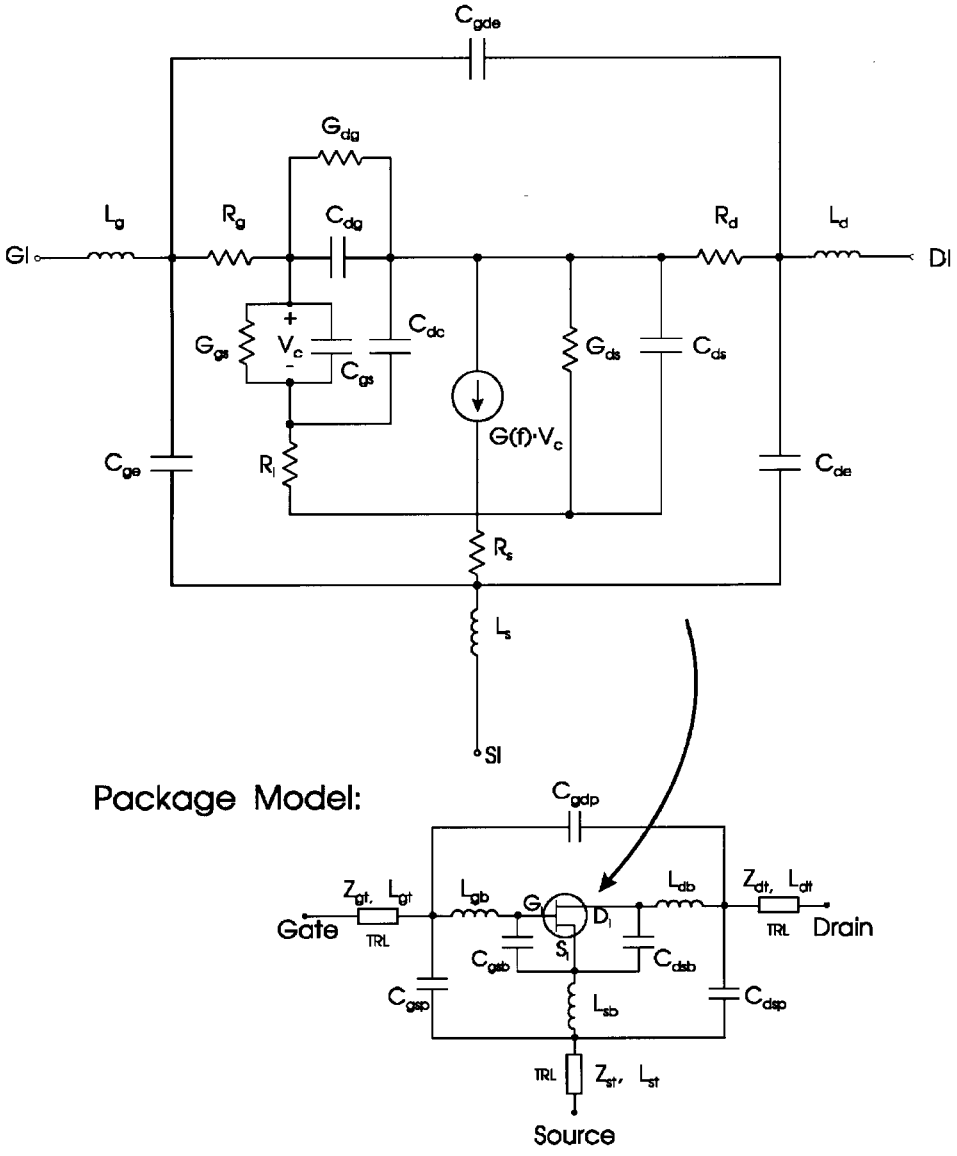


Figure 2-121 Small-signal model of a MESFET.

Table 2-19 Small-signal MESFET model keywords

Keyword	Description	Unit	Default
G	Transconductance at dc, G_0[a]	/ohm	
CGS	Gate–source capacitance	farad	
F	3-dB Roll-off frequency	hertz	∞
T	Time delay	second	0.0
TDS	Drain–source time delay	second	0.0
GGS	Gate–source conductance	/ohm	0.0
CDG	Drain–gate capacitance	farad	0.0
CDC	Dipole-layer capacitance	farad	0.0
CDS	Drain–source capacitance	farad	0.0
GDS	Drain–source conductance	/ohm	0.0
RI	Channel resistance	ohm	0.0
RG	Gate resistance	ohm	0.0
RD	Drain resistance	ohm	0.0
RS	Source resistance	ohm	0.0
CGE	External gate capacitance	farad	0.0
CDE	External drain capacitance	farad	0.0
LG	Gate-lead inductance	henry	0.0
LD	Drain-lead inductance	henry	0.0
LS	Source-lead inductance	henry	0.0
CGDE	External gate–drain capacitance	farad	0.0
GDG	Gate–drain conductance	/ohm	0.0
TJ	Chip temperature	kelvin	298
	Package Parasitics		
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate-to-source package capacitance	farad	0.0
CDSP	Drain-to-source package capacitance	farad	0.0
CGDP	Gate-to-drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for at $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for at $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for at $\varepsilon_r = 1$	meter	0.0
FC	Corner frequency of flicker (1/*f*) noise[b]	hertz	10 MHz
FCP	Shape factor of the 1/*f* noise response	1.0	
label	User-defined term that refers to temperature coefficient		

[a]The transconductance of this model may approximately be described by $g_m = G\dfrac{e^{-j\omega T}}{1 = j(f/F)}$ where $\omega = 2\pi f$ and f is frequency.

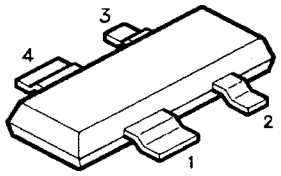
[b]The flicker noise frequency dependence is given by $1/(f/F_c)^{\mathrm{FCP}}$

SIEMENS AlGaAs / InGaAs HEMT CFY 77

Datasheet

Features

* Very low noise
* Very high gain
* For low noise front end amplifiers up to 20 GHz
* For DBS down converters



ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering code (taped)	Package 1)
CFY77-08	HG	Q62702-F1549	MW-4
CFY77-10	HH	Q62702-F1559	MW-4

Maximum ratings	Symbol		Unit
Drain-source voltage	V_{DS}	3.5	V
Drain-gate voltage	V_{DG}	4.5	V
Gate-source voltage	V_{GS}	-3.0	V
Drain current	I_D	60	mA
Channel temperature	T_{Ch}	150	°C
Storage temperature range	T_{stg}	-65...+150	°C
Total power dissipation ($T_s \leq 51$°C) 2)	P_{tot}	180	mW
Thermal resistance			
Channel-soldering point source	R_{thChS}	550	K/W

1) Dimensions see chapter Package Outlines
2) T_S: Temperature measured at soldering point

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Electrical characteristics at T_A = 25°C

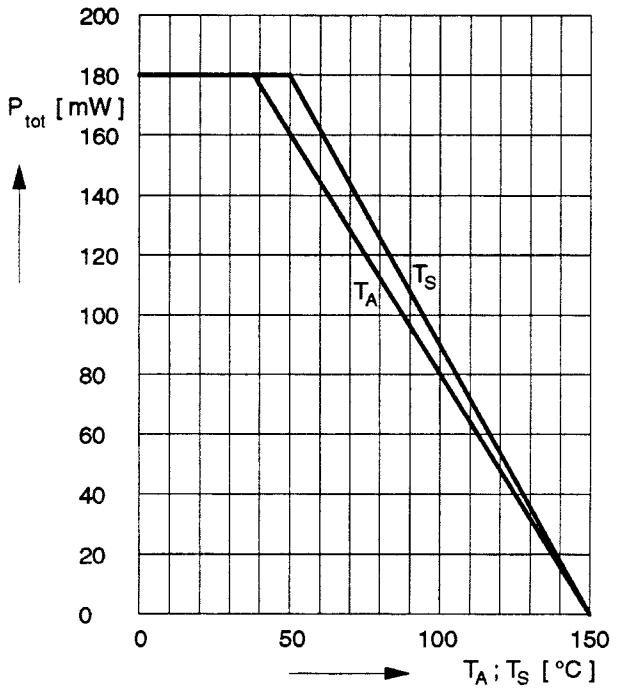
unless otherwise specified

Characteristics	Symbol	min	typ	max	Unit
Drain-source saturation current V_{DS} = 2 V V_{GS} = 0 V	I_{DSS}	15	30	60	mA
Pinch-off voltage V_{DS} = 2 V I_D = 1 mA	$V_{GS(P)}$	-2	-0.7	-0.2	V
Gate leakage current V_{DS} = 2 V I_D = 15 mA	I_G	-	0.05	2	µA
Transconductance V_{DS} = 2 V I_D = 15 mA	g_m	50	65	-	mS
Noise figure V_{DS} = 2 V I_D = 15 mA f = 12 GHz	F				dB
CFY77-08 CFY77-10		- -	0.7 0.9	0.8 1	
Associated gain V_{DS} = 2 V I_D = 15 mA f = 12 GHz	G_a				dB
CFY77-08 CFY77-10		10 9.5	10.5 10	- -	

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Total Power Dissipation $P_{tot} = f(T_s;T_A)$

Package mounted on alumina



Typical Common Source Noise Parameters

I_D = 15 mA U_{DS} = 2.0 V Z_0 = 50 Ω

f	F_{min}	G_a	Γ_{opt}		R_n	r_n	N	$F_{50\Omega}$
GHz	dB	dB	MAG	ANG	Ω			dB
2	0.36	19.4	0.79	27	13.7	0.274	0.03	1.2
4	0.44	15.9	0.72	60	10.1	0.202	0.04	1.1
6	0.51	13.9	0.63	92	5.85	0.117	0.05	1.05
8	0.58	12.4	0.56	134	2.35	0.047	0.06	1.0
10	0.65	11.2	0.52	180	1.1	0.022	0.07	1.0
12	0.72	10.4	0.54	-135	2.9	0.058	0.08	1.1
14	0.80	9.7	0.59	-108	7.15	0.143	0.10	1.5

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Typical Common Source S-Parameters

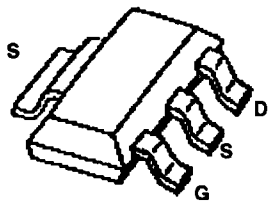
I_D = 15 mA U_D = 2.0 V Z_o = 50 Ω

	S11		S21		S12		S22	
GHz	Mag	Ang	Mag	Ang	Mag	Ang	Mag	Ang
1	0.98	-22.8	5.55	159.6	0.030	87.5	0.633	-16.3
2	0.94	-46.1	5.40	139.3	0.053	57.8	0.60	-32.5
3	0.88	-68.4	5.09	120.1	0.074	44.9	0.54	-48.0
4	0.82	-90.6	4.77	101.2	0.089	30.7	0.48	-63.3
5	0.77	-110.8	4.45	84.0	0.101	18.1	0.42	-77.5
6	0.72	-131.4	4.16	67.3	0.112	7.9	0.35	-92.6
7	0.66	-153.6	3.88	50.2	0.119	-3.3	0.28	-110.8
8	0.63	-175.2	3.58	34.5	0.122	-12.7	0.22	-132.0
9	0.62	164.4	3.29	18.9	0.120	-22.0	0.16	-157.3
10	0.62	145.0	3.01	4.0	0.119	-29.5	0.14	177.3
11	0.64	128.3	2.76	-10.3	0.119	-37.4	0.15	136.2
12	0.64	113.1	2.51	-23.5	0.114	-44.0	0.18	115.4
13	0.66	101.3	2.32	-35.7	0.114	-47.3	0.23	100.9
14	0.67	89.4	2.18	-48.2	0.116	-53.1	0.25	91.0
15	0.69	73.6	2.06	-62.4	0.116	-58.6	0.28	75.4
16	0.73	59.2	1.85	-75.9	0.115	-65.8	0.36	57.1
17	0.76	51.7	1.65	-86.5	0.112	-69.4	0.39	53.1
18	0.78	45.4	1.56	-96.7	0.115	-72.3	0.42	43.8
19	0.77	36.2	1.51	-108.6	0.121	-76.7	0.44	38.8

SIEMENS GaAs FET **CLY 15**

Datasheet

* Power amplifier for mobile phones
* For frequencies from 400 MHz to 2.5 GHz
* Operating voltage range: 2.7 to 6 V
* P_{out} at V_D=3V, f=1.8 GHz typ. 31.5 dBm
* Efficiency better 50%



ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering code (taped)	Package 1)
CLY 15	CLY 15	Q62702-L99	SOT 223

Maximum ratings	Symbol		Unit
Drain-source voltage	V_{DS}	9	V
Drain-gate voltage	V_{DG}	12	V
Gate-source voltage	V_{GS}	-6	V
Drain current	I_D	5	A
Channel temperature	T_{Ch}	150	°C
Storage temperature	T_{stg}	-55...+150	°C
Total power dissipation (Ts ≤ 80°C) Ts: Temperature at soldering point	P_{tot}	4.7	W
Thermal resistance Channel-soldering point (GND)	R_{thChS}	< 15	K/W

1) Dimensions see chapter Package Outlines

SIEMENS GaAs FET **CLY 15**

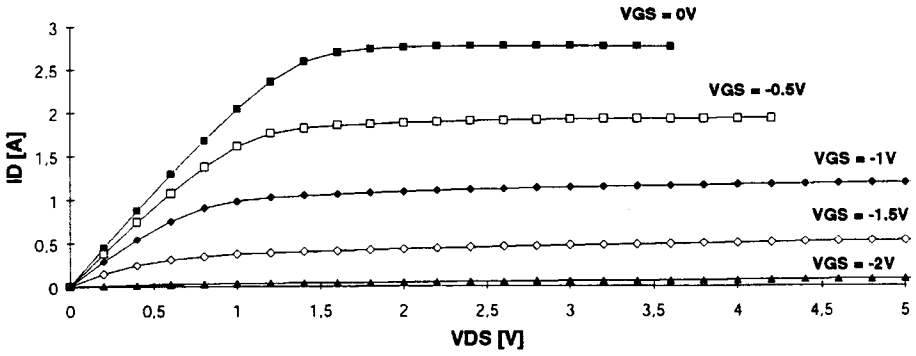
Electrical characteristics (T_A = 25°C , unless otherwise specified)

Characteristics	Symbol	min	typ	max	Unit
Drain-source saturation current *) V_{DS} = 3V V_{GS} = 0V	I_{DSS}	2.4	3.2	4.8	A
Cut-off current V_{DS} = 3V V_{GS} = -3.8V	I_D	-	-	400	µA
Gate cut-off current V_{DS} = 3V V_{GS} = -3.8V	I_G	-	20	70	µA
Pinch-off Voltage VDS=3V ID=400µA	$V_{GS(p)}$	-3.8	-2.8	-1.8	V
Small Signal Gain *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 5dBm	G	-	6	-	dB
Output Power *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 29dBm	P_o	32	32.5	-	dBm
Output Power *) V_{DS} = 5V I_D = 1.4A f = 1.8GHz P_{in} = 30 dBm	P_o	34.5	35	-	dBm
1dB-Compression Point *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz	P_{1dB}	-	31.5	-	dBm
1dB-Compression Point *) V_{DS} = 5V I_D = 1.4A f = 1.8GHz	P_{1dB}	-	34.5	-	dBm
Power Added Efficiency *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 29dBm	η_D	45	50	-	%

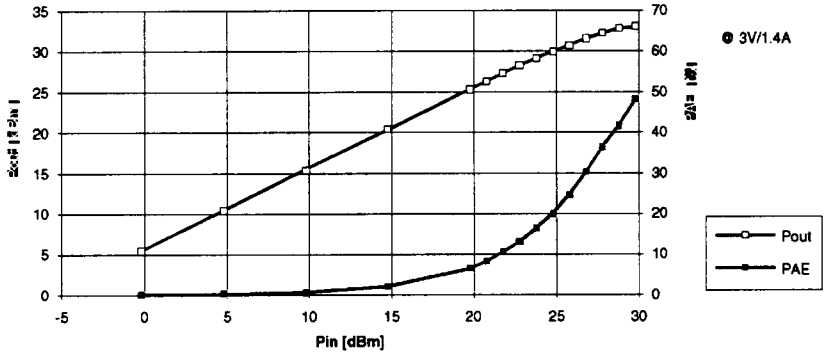
*) pulsed measurement; duty cycle 1:10; t_{on} = 1ms, power matching conditions.

SIEMENS GaAs FET **CLY 15**

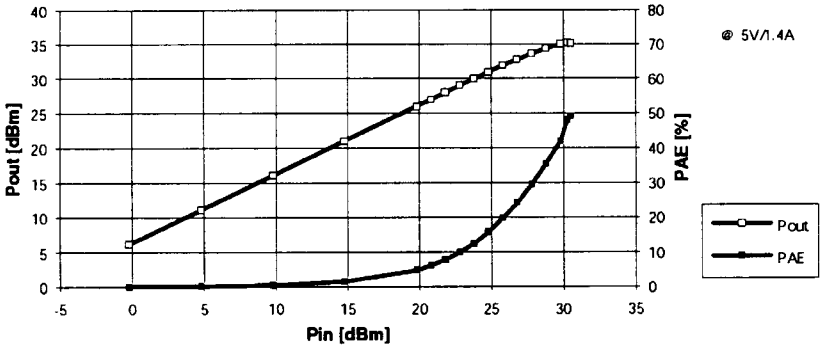
Output Characteristics



Power Characteristics



Power Characteristics



SIEMENS GaAs FET **CLY 15**

typ. Common Source S-Parameter

V_{DS} = 3V I_D=1.4A Z_O=50Ω

f	S11		S21		S12		S22	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
0.200	0.91	-150.9	5.69	99.7	0.01	48.5	0.90	176.0
0.250	0.91	-160.6	4.63	93.5	0.02	45.9	0.87	173.8
0.300	0.90	-167.9	3.89	88.7	0.02	46.8	0.88	171.8
0.350	0.90	-173.7	3.34	84.6	0.02	47.4	0.87	170.8
0.400	0.90	-178.7	2.92	80.9	0.02	47.5	0.87	168.8
0.450	0.90	176.9	2.60	77.5	0.02	47.6	0.87	167.3
0.500	0.90	173.0	2.34	74.6	0.02	48.1	0.87	165.8
0.550	0.90	169.5	2.12	71.4	0.02	47.7	0.87	164.2
0.600	0.90	166.1	1.95	68.7	0.03	47.0	0.87	162.8
0.650	0.90	163.1	1.79	66.1	0.03	47.1	0.87	161.2
0.700	0.90	160.0	1.66	63.5	0.03	46.6	0.87	159.7
0.750	0.90	157.2	1.54	60.9	0.03	45.6	0.87	158.3
0.800	0.90	154.6	1.45	58.6	0.03	45.0	0.87	156.9
0.850	0.90	152.0	1.36	56.1	0.03	43.9	0.87	155.6
0.900	0.90	149.3	1.28	53.8	0.04	43.0	0.87	154.0
0.950	0.90	146.9	1.21	51.5	0.04	41.9	0.87	152.6
1.000	0.90	144.5	1.15	49.0	0.04	41.0	0.87	151.3
1.200	0.91	135.2	0.95	40.3	0.05	36.1	0.87	145.8
1.400	0.91	126.7	0.81	31.8	0.05	31.9	0.88	140.1
1.600	0.92	118.5	0.70	23.8	0.06	26.1	0.88	134.7
1.800	0.92	110.6	0.61	16.3	0.06	20.8	0.88	129.7
2.000	0.93	103.2	0.55	8.7	0.06	15.6	0.89	124.3
2.200	0.93	96.3	0.49	2.1	0.07	10.4	0.88	119.1
2.400	0.93	89.3	0.44	-4.1	0.07	5.2	0.90	114.4
2.600	0.94	82.8	0.40	-10.0	0.07	0.2	0.90	109.3
2.800	0.94	77.0	0.37	-14.9	0.07	-4.2	0.90	104.5
3.000	0.94	71.3	0.34	-19.6	0.08	-9.7	0.91	99.8
3.200	0.93	66.0	0.32	-23.4	0.08	-15.0	0.92	95.1
3.400	0.92	61.6	0.31	-26.8	0.08	-19.4	0.93	90.8
3.600	0.91	57.3	0.30	-29.7	0.07	-23.7	0.92	87.0
3.800	0.90	53.1	0.31	-33.1	0.07	-28.1	0.93	83.1
4.000	0.89	49.2	0.32	-38.1	0.07	-31.9	0.93	79.8
4.200	0.86	46.4	0.34	-44.9	0.07	-35.4	0.92	76.4
4.400	0.83	44.7	0.36	-55.4	0.07	-37.5	0.92	73.4
4.600	0.89	44.2	0.07	-36.2	0.07	-38.1	0.92	71.0
4.800	0.83	43.7	0.34	-80.6	0.07	-39.4	0.92	68.2
5.000	0.85	42.2	0.30	-92.1	0.07	-40.3	0.92	65.2
5.200	0.88	39.4	0.27	-100.8	0.07	-42.5	0.92	62.2
5.400	0.89	36.5	0.24	-107.8	0.07	-45.0	0.92	58.7
5.600	0.90	33.1	0.22	-113.6	0.07	-48.6	0.92	55.7
5.800	0.91	29.6	0.19	-118.9	0.07	-51.4	0.92	52.1
6.000	0.92	26.4	0.18	-124.4	0.07	-54.4	0.92	48.0

SIEMENS GaAs FET **CLY 15**

typ. Common Source S-Parameter
V_{DS} = 5V I_D=1.4A Z_O=50Ω

f	S11		S21		S12		S22	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
0.200	0.90	-151.1	7.61	98.8	0.01	46.5	0.84	176.7
0.250	0.89	-160.6	6.18	92.4	0.02	43.4	0.82	174.7
0.300	0.89	-167.8	5.19	87.5	0.02	46.5	0.82	172.9
0.350	0.89	-173.7	4.45	83.3	0.02	46.0	0.82	171.8
0.400	0.88	-178.7	3.90	79.4	0.02	46.5	0.82	169.7
0.450	0.89	177.0	3.47	75.9	0.02	47.3	0.82	168.3
0.500	0.88	173.2	3.11	72.8	0.02	47.9	0.82	166.7
0.550	0.88	169.6	2.82	69.5	0.02	47.8	0.82	165.5
0.600	0.89	166.4	2.59	66.6	0.03	47.4	0.82	163.9
0.650	0.88	163.1	2.38	63.9	0.03	47.4	0.82	162.6
0.700	0.89	160.3	2.20	61.1	0.03	46.5	0.82	161.0
0.750	0.89	157.5	2.05	58.4	0.03	45.6	0.82	159.6
0.800	0.89	154.9	1.91	55.9	0.03	45.3	0.82	158.0
0.850	0.89	152.1	1.79	53.2	0.03	44.8	0.82	156.8
0.900	0.89	149.7	1.69	50.7	0.03	43.9	0.82	155.4
0.950	0.89	147.1	1.59	48.4	0.04	42.7	0.82	154.1
1.000	0.89	144.7	1.51	45.7	0.04	42.0	0.82	152.8
1.200	0.89	135.5	1.24	36.2	0.04	37.8	0.83	147.3
1.400	0.90	127.1	1.04	27.1	0.05	32.2	0.83	141.9
1.600	0.91	119.1	0.90	18.3	0.05	27.4	0.84	136.8
1.800	0.92	111.1	0.78	10.1	0.06	22.5	0.84	131.6
2.000	0.92	103.7	0.68	2.1	0.06	18.2	0.85	126.3
2.200	0.93	96.6	0.61	-5.1	0.06	12.5	0.86	121.3
2.400	0.93	89.8	0.54	-12.0	0.07	7.3	0.86	116.1
2.600	0.93	83.2	0.48	-18.6	0.07	2.6	0.87	111.4
2.800	0.93	77.3	0.43	-24.0	0.07	-2.6	0.88	106.3
3.000	0.93	71.8	0.39	-29.3	0.07	-7.2	0.89	101.8
3.200	0.92	66.6	0.37	-33.5	0.07	-12.1	0.90	97.2
3.400	0.92	61.8	0.34	-37.2	0.07	-16.8	0.91	92.5
3.600	0.91	57.9	0.32	-40.5	0.07	-21.1	0.91	88.8
3.800	0.90	54.1	0.32	-43.9	0.07	-24.9	0.92	85.1
4.000	0.88	50.5	0.31	-48.3	0.07	-27.6	0.92	81.4
4.200	0.87	47.8	0.32	-53.8	0.07	-31.5	0.92	78.1
4.400	0.86	45.7	0.32	-60.9	0.07	-33.4	0.92	74.9
4.600	0.85	43.4	0.32	-68.9	0.07	-35.4	0.92	72.2
4.800	0.85	42.3	0.31	-77.5	0.07	-37.2	0.92	69.3
5.000	0.86	40.3	0.30	-86.7	0.07	-39.3	0.92	66.2
5.200	0.87	37.7	0.28	-94.5	0.07	-41.6	0.92	63.1
5.400	0.88	35.2	0.26	-101.8	0.07	-44.0	0.92	59.6
5.600	0.89	32.1	0.24	-108.4	0.08	-48.5	0.92	56.6
5.800	0.90	29.0	0.22	-114.5	0.08	-50.3	0.92	53.0
6.000	0.90	26.0	0.21	-121.1	0.08	-54.0	0.93	49.0

2-4 PARAMETER EXTRACTION OF ACTIVE DEVICES

2-4-1 Introduction

We have already seen that all linear models that are used on the market are really derived from large-signal models, and it is our point that the quality of the model depends mostly on the quality of the parameters used to describe the model.

As far as modeling is concerned there are two options. The first option is a physics-based model, which can be applied to items such as diodes and bipolar transistors, including its most advanced versions, such as HBTs, SiGe versions, and other future derivatives. The bipolar transistor is essentially a combination of two diodes whereby the base–emitter junction generates the electrons or holes (depending on how the reader is accustomed to viewing the process) and then transfers those to the collector. The emitter, as its name states, emits the charged particles and the collector collects them, minus some current losses, which are expressed in the current gain of the device. A current gain (β) of 100 means that the collector receives 1% less than the emitter emits.

The second modeling option involves the construction of an analytical mathematic equivalent for the device modeled. Earlier in this chapter, we went through a detailed physics-based derivation about the inner workings of these devices, explaining how, for a given dc bias point, one can simplify them to a small-signal equivalent circuit. This is possible since we assume that the RF current and voltages are less than 1% of the dc equivalent currents and voltages. Reducing a device's complex nonlinear equivalent circuit to a quasistatic linear equivalent circuit allows its manufacturer to generate and publish sets of *S* parameters for it. These parameters are bias, voltage, and temperature dependent. At higher frequencies, parasitics play an enormous role in determining a device's RF performance.

A device operating at signal levels that are more than 1% of its dc equivalent currents and voltages must be evaluated in terms of large-signal performance. Most of the large-signal-equivalent circuits for nonbipolar devices, such as FETs (MOS, silicon JFETs, and members of the GaAsFET family), are analytic approximations that essentially bear no physical insight into the inner workings of the transistors. This fact is particularly painful because to look at high-order intermodulation distortion, one needs to have third and fourth derivatives of mathematically continuous equations. The SPICE approach cannot easily be translated into modern harmonic-balance simulators because such cop-outs as IF THEN ELSE statements in programming provide everything else but a continuous model. In the case of our own work, we used complicated curve-fitting equations to obtain analytic equations and their derivatives to accurately provide a large-signal model that is valid over a wide range of dc and RF.

2-4-2 Typical SPICE Parameters

Since we are about to evaluate bipolar microwave transistors, junction FETs, MOSFETs (model level 3), and GaAsFETs, Tables 2-20, 2-21, 2-22, and 2-23 list typical parameters for the devices we have used. These parameters can be obtained by the Scout program with the appropriate measurements. The BSIM model for MOSFETs, applicable for submicron technology transistors, requires an enormous level of parameter extraction and has not fully been validated for the LDMOS-type transistors currently favored for RF and microwave applications.

Table 2-20 BFR193W BJT

IS = 2.738E – 16	BF = 125	NF = 0.95341
VAF = 24	IKF = 0.26949	ISE = 1.0627E – 14
NE = 1.935	BR = 14.267	NR = 1.4289
VAR = 3.8742	IKR = 0.037925	ISC = 3.7409E – 17
NC = 0.94371	RB = 15	IRB = 0.00091763
RBM = 1.8368	RE1 = 0.76534	RC2 = 0.11938
CJE = 1.1824E – 15	VJE = 0.70276	MJE = 0.48654
TF = 1.8828E – 11	XTF = 0.69477	VTF = 0.8
ITF = 0.00096893	PTF = 0	CJC = 9.3503E – 13
VJC = 1.1828	MJC = 0.30002	XCJC = 0.053563
TR = 1.0037E – 09	VJS = 0.75	MJS = 0
XTB = 0	EG = 1.11	XTI = 3
FCC = 0.72063	LB = 0.57E – 9	LC = 0.00E – 9
LE = 0.43E – 9	CBCP = 0.101E – 12	CCEP = 0.175E – 12
CBEP = 0.061E – 12	VCMX = 10 V	

The meaning and significance of the various parameters is best explored in a book on SPICE or semiconductor physics [22–24].

2-4-3 Noise Modeling

Diode Noise Model. The noise model for the diodes (Figure 2-122 and Table 2-24) consists of two contributions: the shot noise and the flicker noise. The shot noise is computed automatically and does not require any parameters. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications).
2. Using bias-dependent flicker noise coefficients (specifying KF and AF at multiple bias points).

The noise generators in the diode noise model are the series parasitic resistance, R_S, and the intrinsic junction. Figure 2-122 illustrates the intrinsic junction noise generator. Let Δf be the bandwidth (usually normalized to 1 Hz). The intrinsic noise generator has a mean-square value of

$$\langle i_{Dn}^2 \rangle = 2qI_D\,\Delta f + \mathrm{KF}\,\frac{I_D^{\mathrm{AF}}}{f^{\mathrm{FCP}}}\,\Delta f \tag{2-266}$$

Notes on the Diode Noise Model:

1. Shot noise is always present unless the SN parameter is set to zero. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.

Table 2-21 2SK125 JFET[a]

IDSS	= 0.5250E – 01	VP0	= –0.3111E + 01	GAMA	= –0.1867E – 01	E	= 0.1520E + 01
KE	= –0.3856E – 03	SL	= 0.2818E – 01	KG	= –0.2398E + 00	T	= 0.0000E + 00
SS	= 0.7448E – 04	IG0	= 0.2000E – 14	AFAG	= 0.3846E + 02	IB0	= 0.1000E – 04
AFAB	= 0.3800E + 02	VBC	= 0.3000E + 02	R10	= 0.1711E + 02	KR	= 0.0000E + 00
C10	= 0.6609E – 11	K1	= 0.1675E + 01	C1S	= 0.6818E – 33	CF0	= 0.7261E – 11
KF	= 0.1156E + 01	RG	= 0.5000E + 00	RD	= 0.1542E + 01	RS	= 0.1333E + 01
LG	= 0.6098E – 09	LD	= 0.5159E – 08	LS	= 0.1482E – 08	CDS	= 0.4813E – 16
CGE	= 0.1590E – 11	CDE	= 0.3394E – 26	CGSP	= 0.8282E – 13	CDSP	= 0.4832E – 12
ZGT	= 0.5000E + 02	LGT	= 0.4712E – 01	ZDT	= 0.5000E + 02	LDT	= 0.3998E – 01
CGDP	= 0.3653E – 12	ZST	= 0.5000E + 02	LST	= 0.1495E – 01	CGDE	= 0.3831E – 12
CGSB	= 0.3120E – 13	CDSB	= 0.5896E – 12	VDMX	= 10		

[a]The currently implemented model for the junction FET is too primitive for serious RF applications. We have therefore taken the approach (liberty) of using the Materka parameter extraction approach for silicon junction FETs. This has resulted in unparalleled high-quality parameters; in particular, the knee voltage behavior has significantly improved, as well as the overall frequency response.

Table 2.22 GaAs MESFET

IDSS = 0.1077E+00	VP0 = –0.1800E+01	GAMA = –0.5741E-01	E = 0.1290E+01
KE = –0.1155E-01	SL = 0.1652E+00	KG = –0.1782E+00	T = 0.0000E+00
SS = –0.1208E+01	IG0 = 0.2130E-11	AFAG = 0.2740E+02	IB0 = 0.5680E-09
AFAB = 0.1826E+01	VBC = 0.9000E+01	R10 = 0.8382E+01	KR = 0.6359E+00
C10 = 0.5964E-12	K1 = 0.1296E+01	C1S = 0.0000E+00	CF0 = 0.6110E-13
KF = 0.9775E+00	RG = 0.1996E+01	RD = 0.1296E+01	RS = 0.1234E+01
CDS = 0.7852E-13	CDSD = 0.1000E-07	RDSD = 0.1581E+03	CGE = 0.1609E-12
CDE = 0.8674E-13	VDMX = 8		

Table 2.23 The 1 μm × 750 μm level 3 LDMOS FET

CBD = 0.863E-12	CGD0 = 166E-12	CGS0 = 246E-12	GAMA = 0.211
IS = 6.53E-16	KAPA = 0.809	MJ = 0.536	NSUB = 1E+15
PB = 0.71	PBSW = 0.71	PHI = 0.579	RD = 39
RS = 0.1	THET = 0.588	TOX = 4E-8	U0 = 835
VMAX = 3.38E5	VT0 = 2.78	XQC = 0.41	

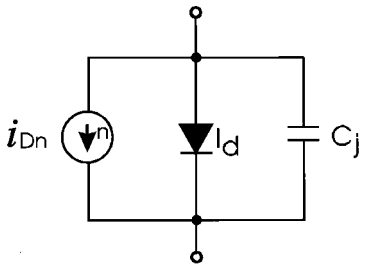


Figure 2-122 Equivalent noise circuit for a diode chip.

Table 2.24 Diode noise model keywords

Keyword	Description	Unit	Default
ID	Required bias current for the data point	Ampere	
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner fequency	hertz	

2. If the value of KF is specified as zero, then the flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.
3. The corner frequency noise model uses the system noise floor to internally compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the diode parameters and *kT*.
4. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

BJT Noise Model. The noise model for the Gummel–Poon BJT model consists of two contributions: shot noise and flicker noise. The shot noise is computed automatically and does not require any parameters. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications) (see Table 2-25).
2. Using bias-dependent flicker noise coefficients (specifying KF and AF at multiple bias points).

Option 1: *Specifying the Bias-Independent Flicker Noise Coefficient*. This option involves the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model.

1. Shot noise is always present unless it is turned off. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying The Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows a bias-dependent flicker noise coefficient (i.e., KF and AF vary with the bias point).

Notes on the BJT Noise Model:

Table 2.25 BJT noise model keywords

Keyword	Description	Unit	Default
IB	Required base bias current for the data point	ampere	
VCE	Required collector–emitter voltage for the data point	volt	
VBS	Base–substrate voltage required for LPNP type when four nodes are used	volt	
VCS	Collector–substrate voltage required for NPN or PNP type when four nodes are used	volt	
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

1. KF, AF, and FC can be specified as bias dependent. If only one set of noise data is specified, the corresponding bias point is not meaningful because all parameters are considered constant over all bias values. However, the bias point is needed for the program to identify the data as bipolar noise data.
2. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
3. This noise model of course considers the actual operating temperature, which must be supplied to the model.

Figure 2-123 shows the BJT noise model. Let Δf be the bandwidth (usually normalized to a 1-Hz bandwidth). The noise generators introduced in the intrinsic device are shown below and have mean-square values of

$$\langle i_{bn}^2 \rangle = 2qI_B \,\Delta f + \text{KF}\,\frac{I_B^{\text{AF}}}{f^{\text{FCP}}}\,\Delta f \tag{2-267}$$

$$\langle i_{cn}^2 \rangle = 2qI_C \Delta f \tag{2-268}$$

$$\langle i_{R_{bb}}^2 \rangle = \frac{4kT}{R_{bb}}\,\Delta f \tag{2-269}$$

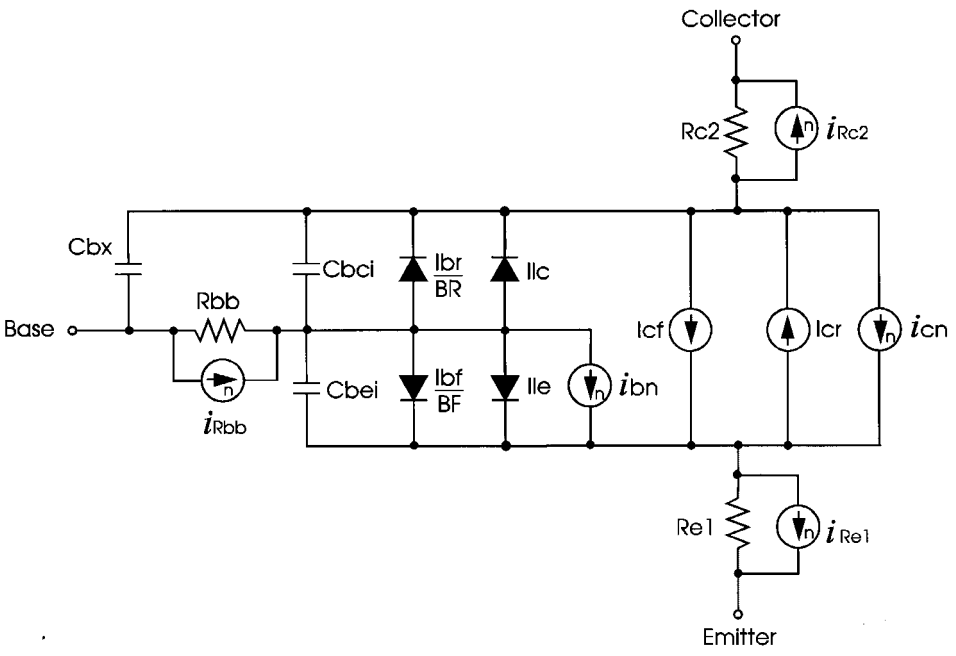


Figure 2-123 BJT noise model (not showing extrinsic parasitics). Current sources with *n* are noise sources.

$$\langle i_{R_{e1}}^2 \rangle = \frac{4kT}{R_{e1}} \Delta f \tag{2-270}$$

$$\langle i_{R_{c2}}^2 \rangle = \frac{4kT}{R_{c2}} \Delta f \tag{2-271}$$

$$I_B = \frac{I_{bf}}{BF} + I_{le} \tag{2-272}$$

$$I_C = I_{cf} - I_{cr} \tag{2-273}$$

JFET and MESFET Noise Models. The noise model for the FETs consists of two contributions: the shot noise and the flicker noise. There are two options to specify noise in the FET model:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (Table 2-26) to determine the flicker noise (this option is usually sufficient for most applications). The shot noise will be automatically computed using the SPICE equation.
2. Using bias-dependent flicker noise coefficients through a reference in the DATA block (specifying KF and AF at multiple bias points) and specifying the four noise parameters [$F_{\min}$, Γ_{opt} (magnitude and phase), and R_n] at multiple bias points.

Option 1: *Specifying the Enhanced SPICE Noise Model.* Option 1 is the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model. The drain noise model has the form

$$\langle |I_{dn}|^2 \rangle = 4K_B T \frac{2g_m}{3} \Delta f + \text{KF} \frac{|I_D|^{\text{AF}}}{f^{\text{FCP}}} \Delta f \tag{2-274}$$

Table 2.26 FET noise model keywords

Keyword	Description	Unit	Default
FN	Noise data measurement frequency	hertz	1.0 GHz
VGS	Required gate–source voltage for the data point	volt	
VDS	Required drain–source voltage for the data point	volt	
FMIN	Required minimum noise figure in dB at FN		
MGO	Required magnitude of optimum noise reflection coefficient at FN		
PGO	Required phase of optimum noise reflection coefficient at FN		
RN	Required normalized noise resistance at FN		
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

where the shot noise is derived from g_m and the flicker noise is proportional to KF and the drain channel current, I_D, and inversely proportional to frequency. The AF and FCP parameters tailor the flicker noise dependence on bias and frequency, respectively.

1. Shot noise is always present unless it is turned off. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, then flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying the Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows the specification of the complex bias-dependent nature of the shot noise and flicker noise. At high frequencies, the equivalent noise sources are correlated (the SPICE noise model does not account for this correlation). The complete evaluation of the shot noise sources can be determined from the four noise parameters. Since these are functions of bias, they can be specified over the (V_{GS}, V_{DS}) bias plane.

Additionally, a bias-dependent flicker noise coefficient (i.e., KF and AF vary with current) can be specified.

The MESFET noise model uses the four measured noise data [$F_{\min}$, Γ_{opt} (magnitude and phase), and R_n] at one frequency and multiple arbitrary bias points. The program uses these data and the FET model parameters to de-embed the noise data to an intrinsic noise model. The intrinsic model is accurate at all frequencies and, therefore, can predict the noise performance at all frequencies given data at just one frequency point. Built-in bias-dependent characteristics are used if multi-bias noise data are not provided.

Notes on the FET Noise Model

1. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
2. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

Noise in a MESFET is produced by sources intrinsic to the device. The same approach, but with different flicker corner frequencies, is highly applicable to JFETs and MOSFETs. For more detail as to simulation, see the Element library book for the active-device portion of Ansoft's Serenade Design Environment product. The equivalent noisy circuit of an intrinsic FET is represented in Figure 2-124.

The intrinsic FET is internally represented as a noiseless nonlinear two-port with one equivalent noise current connected across the gate–source terminal, and one across the drain–source terminal. The correlations of the gate and drain noise current sources are

$$\langle |I_{gn}|^2 \rangle = 4K_B T \, \Delta f \frac{\omega^2 C_{gs}^2}{g_m} R \tag{2-275}$$

$$\langle |I_{gn}|^2 \rangle = 4K_B T \, \Delta f \, g_m P \tag{2-276}$$

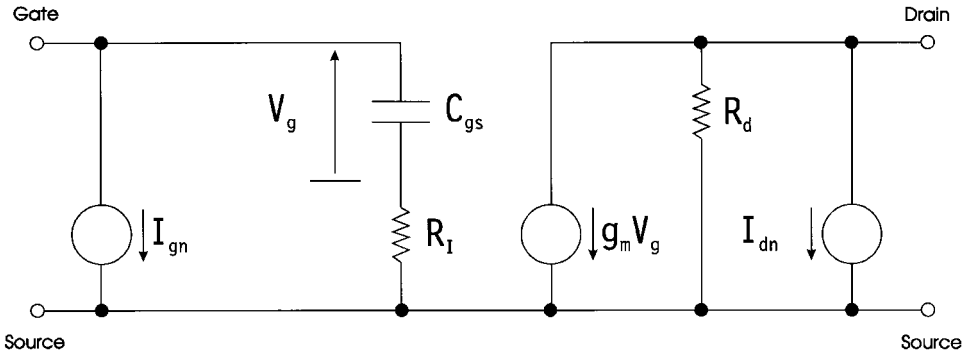


Figure 2-124 Equivalent noise circuit of an intrinsic FET device.

$$\langle I_{gn} I_{dn}^{*} \rangle = 4K_B T\, \Delta f\, j\omega C_{gs} \sqrt{PR}\, C \tag{2-277}$$

The correlation matrix of the noise current sources is

$$C_{dc}(\omega) = \frac{2}{\pi} K_B T\, d\omega \begin{bmatrix} \dfrac{\omega^2 C_{gs}^2}{g_m} R & -j\omega C_{gs}\sqrt{PR}\, C \\ j\omega C_{gs}\sqrt{PR}\, C & g_m P \end{bmatrix} \tag{2-278}$$

The gate and drain noise parameters R and P and the correlation coefficient C are related to the physical noise sources acting in the channel and are functions of the device structure and bias noise parameters. By defining measured noise parameters, $F_{\min}$, R_n, and Γ_{opt}, and using a noise de-embedding procedure, the parameters R, P, and C and the intrinsic noise correlation matrix of a FET device as functions of device bias are determined by the program.

In addition to the noise sources shown above, the flicker (1/f) noise can also be modeled by means of a noise current source connected in parallel with the intrinsic drain port. The flicker noise component in a narrow band, Δf, is expressed in the form

$$\langle |I_f|^2 \rangle = \mathrm{Q}\, \Delta f \frac{|I_D|^{\mathrm{AF}}}{f^{\mathrm{FCP}}} \tag{2-279}$$

where I_D is the instantaneous value of the channel current, and Q, AF, and FCP are empirical parameters. In most practical cases, AF and FCP are directly obtained from measurements (typically, AF = 2 and FCP = 1), while Q is not. In Ansoft's Serenade Design Environment, Q is either provided directly using KF or is computed by providing the flicker corner frequency (FC). FC is the frequency at which the flicker noise equals the shot/diffusion noise. The corner frequency is defined by the equation

$$\mathrm{Q} \frac{|I_D|^{\mathrm{AF}}}{f_{\mathrm{C}}^{\mathrm{FCP}}} = g_m P \tag{2-280}$$

Given the corner frequency FC and the measurement bias point V_{gs} and V_{ds}, the program automatically computes I_D, g_m, and P, and finally Q.

More information on FET noise modeling can be found in the literature [25–31].

MOSFET Noise Model. The MOSFET noise model (Figure 2-125 and Table 2-27) consists of two contributions: the shot noise and the flicker noise. The shot noise is computed automatically and does not require any parameters. It can be turned off by specifying SN = 0. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications).

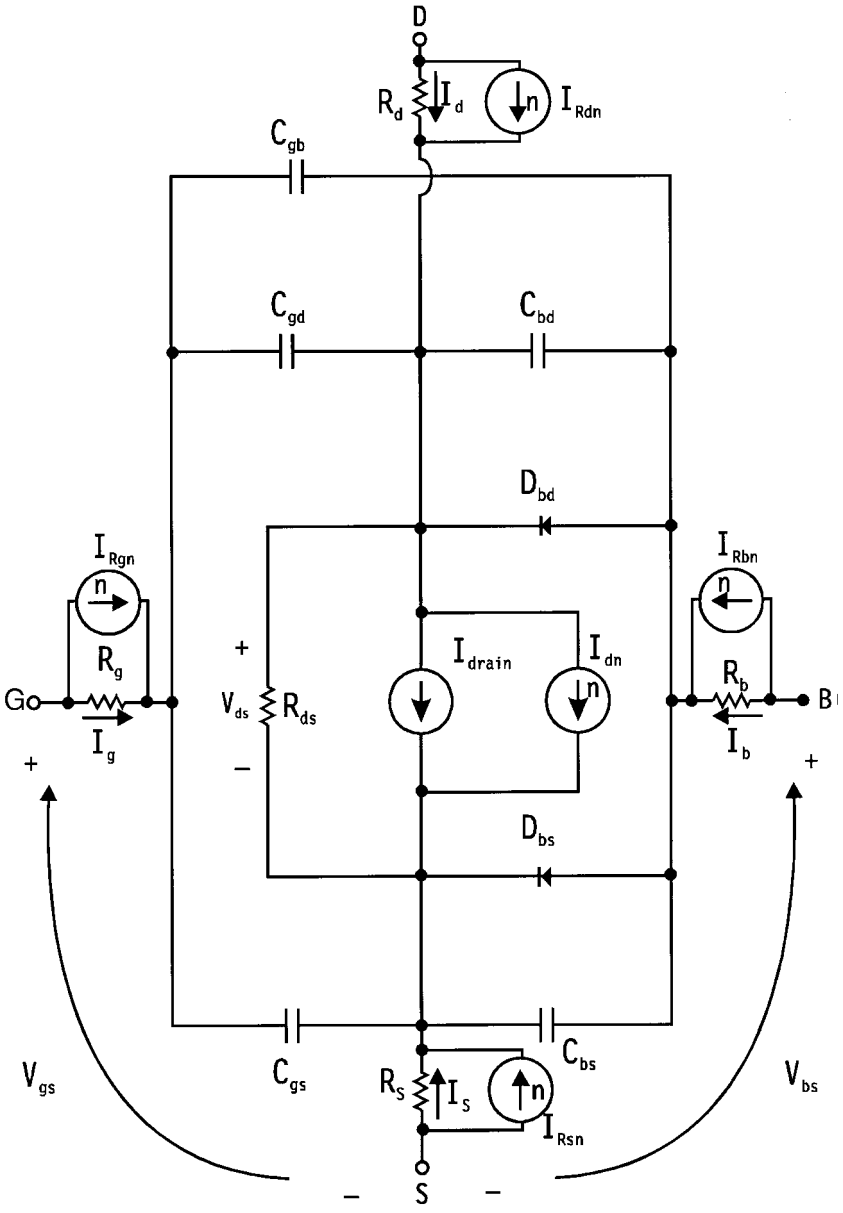


Figure 2-125 Equivalent noise circuit of an intrinsic MOSFET device.

Table 2.27 MOSFET noise model keywords

Keyword	Description	Unit	Default
VGS	Required gate–source bias for the data point	volt	
VDS	Required drain–source for the data point	volt	
VBS	Required drain–bulk for the data point	volt	
KF	Flicker noise coefficient		1.0E–13
AFT	Bias exponent of the flicker noise model		2.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

2. Using bias-dependent flicker noise coefficients through a reference (specifying KF and AF at multiple bias points).

Option 1: *Specifying the Enhanced SPICE Noise Model.* This option is the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model (the flicker noise is considered bias dependent).

1. Shot noise is always present unless the SN parameter is set to zero. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, then the flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying the Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows a bias-dependent flicker noise coefficient (i.e., KF and AF vary with drain current). The MOSFET noise model data are given and referenced by a model parameter.

Notes on the MOSFET Noise Model

1. KF, AF, and FC can be specified as bias dependent. If only one set of noise data is specified, the corresponding bias point is not meaningful because all parameters are considered constant over all bias values. However, the bias point is needed for the program to identify the data as MOSFET noise data.
2. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
3. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

Let Δf be the bandwidth (normalized to 1 Hz). The noise generators introduced in the intrinsic device are shown below and have mean-square values of

$$\langle i_{dn}^2 \rangle = \frac{8kTg_m}{3}\Delta f + \mathrm{KF}\frac{I_D^{\mathrm{AF}}}{f^{\mathrm{FCP}}}\Delta f$$

$$\langle i_{Rgn}^2 \rangle = 4\frac{kT}{R_g}\Delta f$$

$$\langle i_{Rdn}^2 \rangle = 4\frac{kT}{R_d}\Delta f$$

$$\langle i_{Rsn}^2 \rangle = 4\frac{kT}{R_s}\Delta f$$

$$\langle i_{Rbn}^2 \rangle = 4\frac{kT}{R_b}\Delta f$$

We include this MOSFET noise model (used for quite awhile) for completeness. At the moment, we do not know which MOSFET noise model the industry will settle on in the future.

Modern CAD tools, such as Ansoft's Serenade product, use these models to generate quite accurate noise data based on a good linear equivalent model. Internally Serenade uses the noise-correlation-matrix method (first introduced by Russer). The same noise-correlation techniques apply to BJTs, all FETs except MOS, and HBTs (such as SiGe devices).

2-4-4 Scalable Device Models

Since diodes and transistors are scalable, here are guidelines for how to scale them:

Microwave Diode

$$I_D = Area \times I_D$$
$$C_j = Area \times C_j$$
$$R_D = R_D/\text{Area}$$

PIN Diode

$$I_D = \text{Area} \times I_D$$
$$C_j = \text{Area} \times C_j$$
$$R_S = R_S/\text{Area}$$
$$R_{\max} = R_{\max}/\text{Area}$$

Bipolar

$$I_{bf} = \text{Area} \times I_{bf} \qquad I_{br} = \text{Area} \times I_{br}$$
$$I_{le} = \text{Area} \times I_{le} \qquad I_{lc} = \text{Area} \times I_{lc}$$
$$I_{cf} = \text{Area} \times I_{cf} \qquad I_{cr} = \text{Area} \times I_{cr}$$
$$C_{bc} = \text{Area} \times C_{bc} \qquad C_{be} = \text{Area} \times C_{be}$$
$$C_{bx} = \text{Area} \times C_{bx} \qquad R_{bb} = R_{bb}/\text{Area}$$
$$\text{RB2} = \text{RB2/Area} \qquad \text{RC2} = \text{RC2/Area}$$
$$\text{RE1} = \text{RE1/Area} \qquad I_{jss} = I_{jss} \times \text{Area}$$
$$C_{jss} = C_{jss} \times \text{Area}$$

Materka FET

$$\begin{aligned}
\text{IGSS} &= \text{IGSS} \times \text{Area} \\
\text{CGS0} &= \text{CGS0} \times \text{Area} \\
\text{CGS1} &= \text{CGS1} \times \text{Area} \\
\text{CDVC} &= \text{CDVC} \times \text{Area} \\
\text{CDVS} &= \text{CDVS} \times \text{Area} \\
R_G &= R_G \times \text{Area}/(\text{Number of fingers} \times 2) \\
R_D &= R_D/\text{Area} \\
R_S &= R_S/\text{Area}
\end{aligned}$$

Generating a Databank for Parameter Extraction. Given the fact that one can measure with an automated system (Rohde & Schwarz, Hewlett-Packard) the bias-dependent *S* parameters of any active device, bipolar and FET, and furthermore, that Hewlett-Packard makes a computer-controllable dc *I–V* curve tracer, we can generate a huge databank that allows us to have corresponding values for both RF and dc operating points. To validate this, one needs to run a large-signal simulator, such as Ansoft's Serenade Design Environment, or its equivalent from other manufacturers, enter the large-signal parameters, and generate a set of *S*-parameters that correspond to the dc values for which we have measurements. The agreement between measured and predicted points for both RF and dc values is a good measure to evaluate the accuracy of the parameter extraction and the simulator's performance. We cannot stress often enough, however, the fact that while most simulator models are quite good, the weakest point of the link is the parameter extraction program.

Currently, the best parameter extraction for bipolar transistors and FETs, excluding the BSIM model, is implemented in the Scout program by Ansoft, based on modified techniques found in the literature. These techniques are based on the fact that present-day nonlinear modeling of microwave devices, especially FETs, is not exactly adequate to describe all effects found within these devices. In the case of FETs, the most troublesome parameter is R_{DS}, which is really more bias dependent than practically all existing models take into consideration. Techniques used in published extraction methods extract nonlinear models independently for each section of a device model, and then assume that the proper response will be generated when all parts are put together. For instance, many extraction techniques independently fit the *I–V* data of the section describing the *I–V* curves and then fit the *C–V* data to the equivalent small-signal models. This results in an approximation to the bias-dependent response of the complete model.

In the technique used in the Scout program, model parameters are extracted simultaneously so the effects between model sections are accounted for and the device is treated as a whole. The most critical area happens to be in the linear region, typically around the pinchoff area, and for very small supply voltages. The Scout techniques use measured data taken at a wide variety of bias ranges and frequencies to define the model response. The possible solution region for the complete model becomes well identified and uniqueness of solution is vastly improved. The Scout program is interactive in nature and displays simultaneously measured versus predicted data. This allows for fine-tuning of the parameters.

Over the years, the industry has adapted a large number of models. Those supported by the Scout parameter extraction program include:

- *npn* Gummel–Poon BJT
- *pnp* Gummel–Poon BJT

- *npn* heterojunction BJT
- *pnp* heterojunction BJT
- General-purpose *n*-channel JFET
- General-purpose *p*-channel JFET
- Chalmers (Angelov) MESFET/HEMT
- Curtice–Ettenberg cubic MESFET/HEMT
- Curtice quadratic MESFET/HEMT
- IAF (Berroth) MESFET/HEMT
- Modified Materka–Kacprzak MESFET/HEMT
- Raytheon (Statz) MESFET/HEMT
- TriQuint (TOM-1 and TOM-2) MESFET/HEMT
- Ansoft physics-based MESFET

The Ansoft physics-based MESFET model is unique and politically sensitive because its users are essentially required to enter all device fabrication parameters to get the appropriate answer, and most, if not all, companies will stay away from handing out such company proprietary recipes. The validation of this model was done using manufacturing data for transistors that were developed under government contract and therefore accessible for the team members. Compact Software/Ansoft over the last ten years has been involved in a large number of government-funded research and development programs, and practically all modern modeling implementations and their validation were done in conjunction with the largest semiconductor manufacturers.

As far as further validation is concerned, besides our looking at the matching of dc *I–V* curves and *S* parameters, important parameters such as 1-dB compression point, power-added efficiency, and harmonic content are important criteria to evaluate the quality of such an undertaking. As far as hardware is concerned, one needs to have a network analyzer, such as the HP8510 or its equivalent. Its selection is determined by the highest cutoff frequency where measurements are necessary. A dc feed capability, up to several amperes, through the measuring ports, and appropriate power supplies are essential, as is a dc *I–V* curve tracer (HP4145). The built-in various optimizers of the program establish the best possible match between the selected parameters and the measured results.

The area of proper modeling has always been fascinating because the active-device model can make or break first-pass success. Figure 2-126 shows the doubler gain comparison of a model supplied by NEC (the device manufacturer, parameter extraction done with Scout, and parameter extraction supplied from the *Microwave Engineering Europe* (*MEE*) magazine in its CAD review of May 1994). It is somewhat unclear where the *MEE* model came from. The actual task given by *MEE* was to simulate a frequency doubler against measured data. This type of simulation required two areas of high precision, one being the model and the other being the electromagnetic modeling of the discontinuities of the circuit—in this case, a stub element. The circuit itself was very simple. The manufacturer-supplied data gave the poorest results (we believe the parameter extraction came from a third party not quite up to speed in this area). The results that came from the Scout parameter extraction and the third one that *MEE* supplied were fairly close, yet the Scout solution, including the electromagnetic simulation part, gave the best answer. More simulators have since appeared on the market, and it would be interesting to revisit this topic. Figure 2-127 shows a physical layout of the doubler circuit.

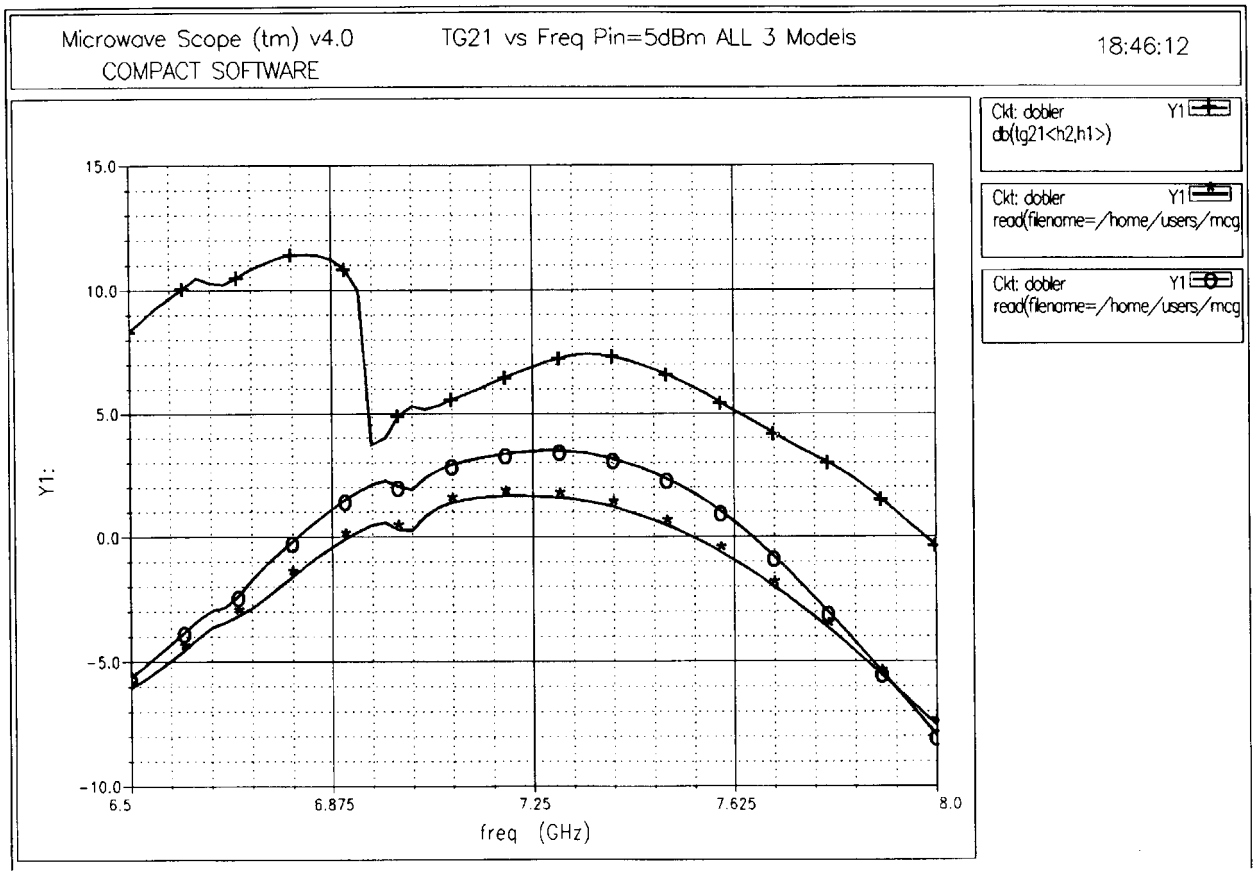


Figure 2-126 Gain comparison of a MESFET doubler based on the three device models used: NEC (+), Ansoft (O), and *Microwave Engineering Europe* (*).

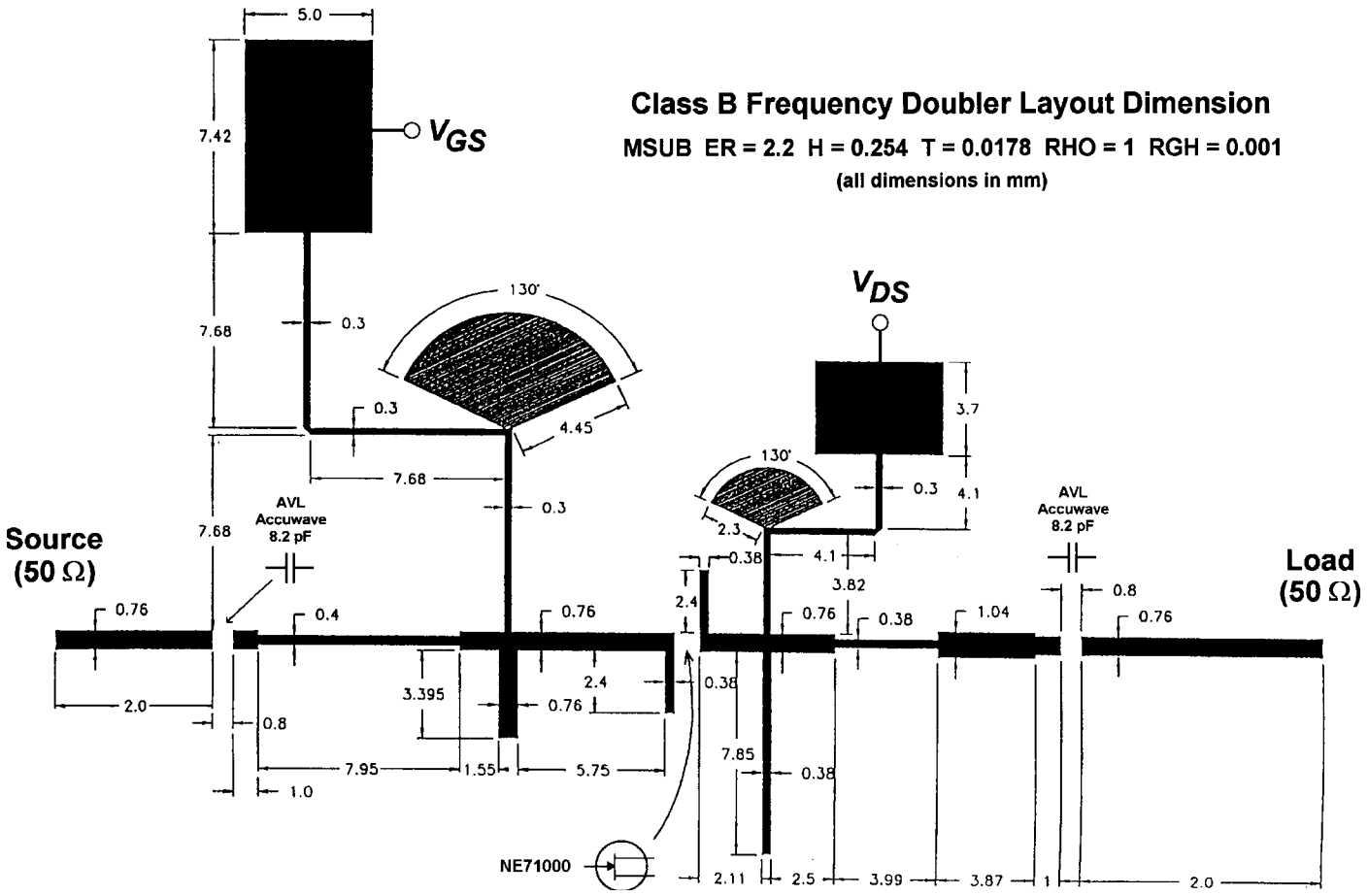


Figure 2-127 Layout of the MESFET doubler.

Test Setup. The measuring equipment required for the actual extraction includes:

- Semiconductor parameter analyzer (HP4145)
- Vector network analyzer (ZVR, HP8510, W360)
- Bias supplies and meters
- Power RF source, power meters, spectrum analyzer, and so on

The required software tools include:

- GPIB controller: Compact NETCOM (network analyzer control program)
- Parameter extraction performed by Scout
- Small-signal and harmonic-balance validation: Serenade nonlinear analysis (Microwave Harmonica) or its equivalent

After we have assembled all these tools, the device extraction process can begin. It will follow the flow as shown in Figure 2-128. Depending on the numbers for the dc/ac measurement, this process can take up to two hours. The Scout program, with its built-in optimizer, can take another two hours. It is therefore reasonable to assume a thorough parameter extraction of a device will take about one day.

Parameter Extraction. As we have outlined, the parameter extraction program Scout, which is optimized for high-frequency and millimeter-wave application, is a Windows-based parameter extraction and optimization program. It requires a network-communication program to acquire dc and *S* parameter measurement data. It supports interactive device fitting and optimization and supports the various models listed earlier. For ease of use, it has an attractive graphical user interface with multiple windows, as can be seen from Figure 2-129. Hewlett-Packard offers more general RF parameter extraction programs, but we did not have access to those. We also expect that some other companies will develop similar tools. An interesting program named HarPE, developed by Professor John Bandler in Canada, may have been withdrawn from the market after his company merger with HP. The program was acclaimed to provide quite good results using measured harmonics. One of the unfortunate deficiencies of most BJT parameter extraction programs known to us is the fact that (as one

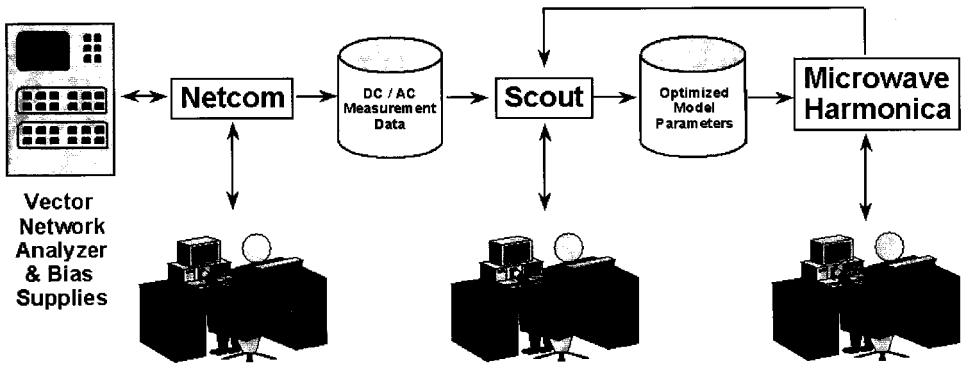


Figure 2-128 Flowchart of the parameter extraction process.

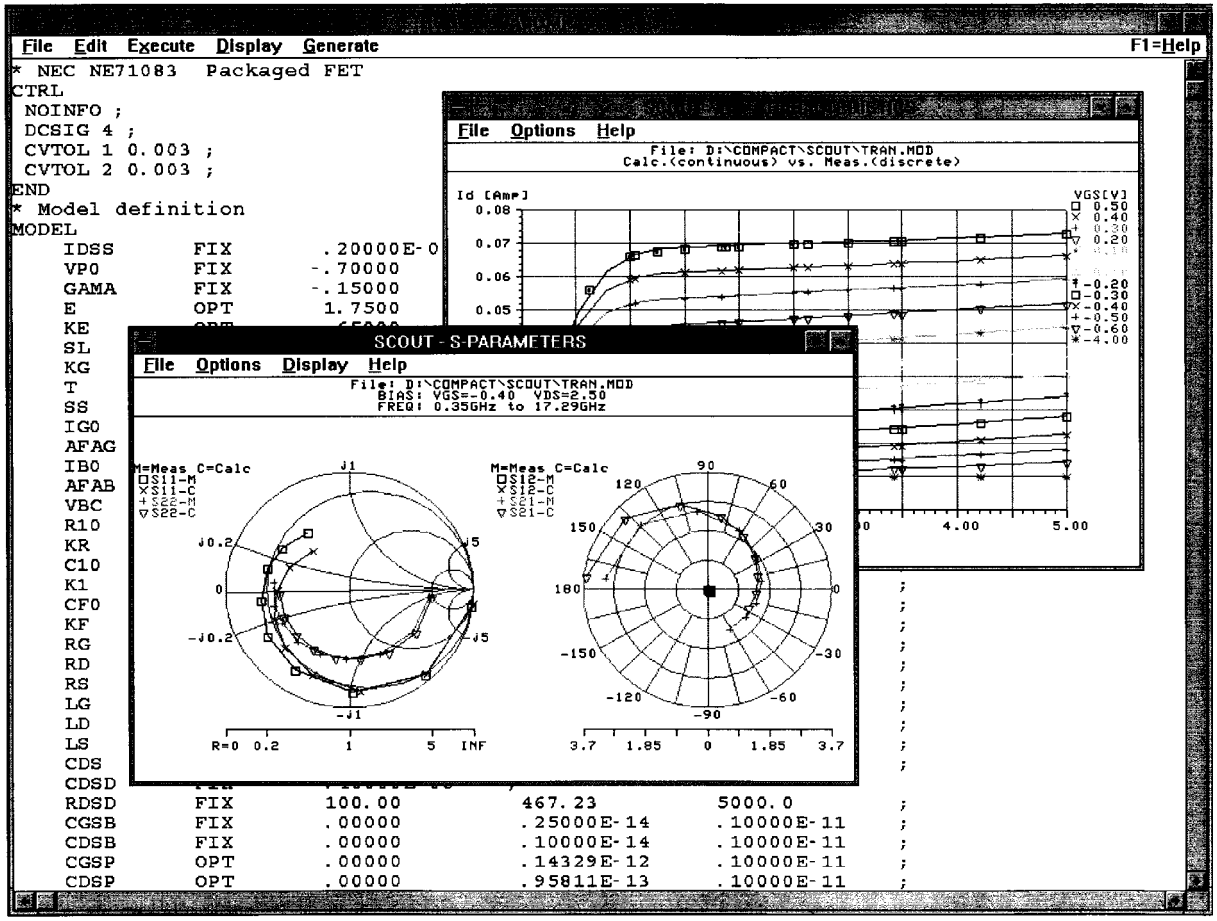


Figure 2-129 The Scout user interface.

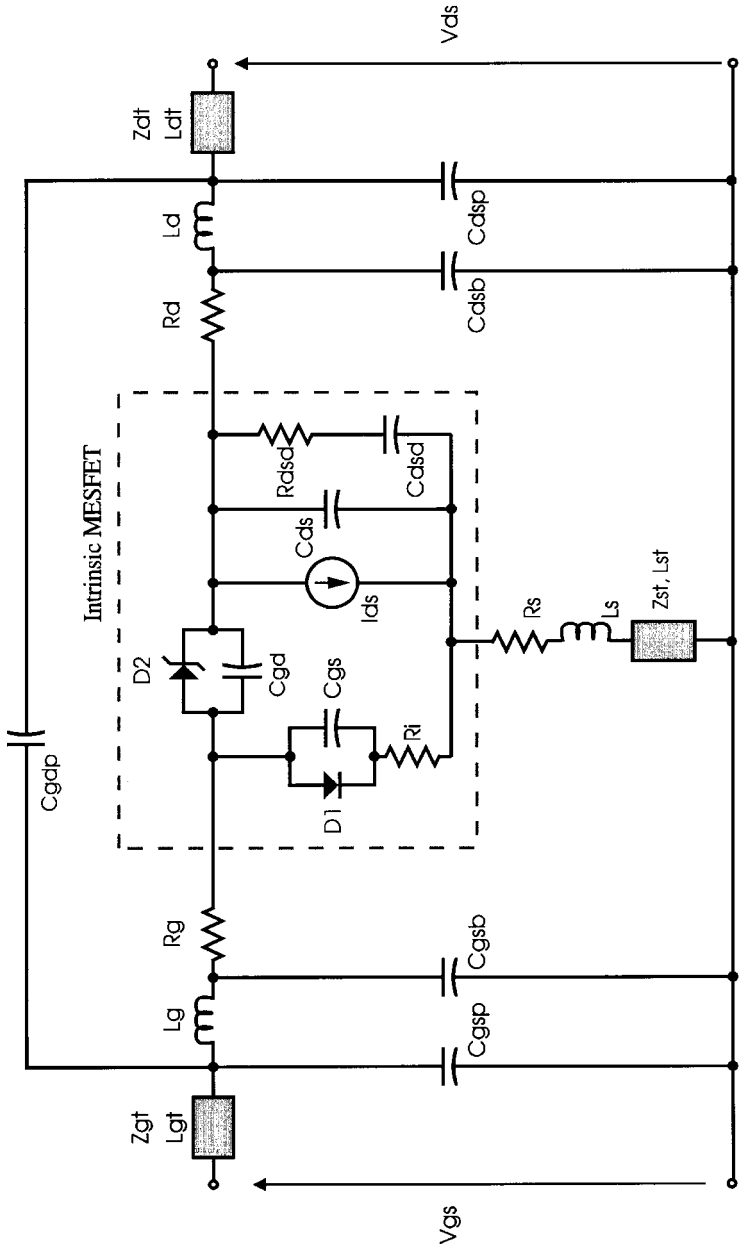


Figure 2-130 Intrinsic model and complete chip/package model.

example) the sensitivity to determine the actual base-spreading resistor for bipolar transistors is poor. The only useful workaround is considering the noise figure of the device, given the fact that the midrange noise factor of any bipolar transistor can be calculated from

$$F = 1 + \frac{R'_{bb}}{R_g} + \frac{0.5(26\ \text{mV}/I_C)}{R_g}$$

(approximate equation at medium frequencies) (2-281)

By solving the equation for R'_{bb}, since we know that R_g, the generator impedance, is 50 Ω, and the emitter diffusion resistance (26 mV/I_C) can be determined from the bias point, we can already eliminate the need to extract R'_{bb} from the HP test equipment. By using HP's program, errors of a factor up to 10 are not uncommon because of the low sensitivity for this parameter. Another headache in the case of the BJT is the determination of the excess phase (PTF) and forward and reverse transit time (TF and TR, respectively). The lack of quality modeling can best be found by examining the measured versus predicted values of S_{22}. Again, practically all the parameter extraction programs we have seen so far suffer from this deficiency.

MESFETs. In the case of a member of the MESFET family, a model as shown in Figure 2-130 is used. It consists of an intrinsic model and a complete chip/package parasitic model. In the case of the MESFET, to obtain first-class results, the following measurements are necessary:

- Dc *I*–*V* measurements
- Dc diode measurements
- Cold-FET and forward-biased-gate *S* parameters
- *S*-parameter measurements over bias
- Harmonic power measurements
- *Z*-Parameter Equations for Figure 2-131

In order to simplify the modeling, it is useful to do cold-FET [Figure 2-131 and Eqs. (2-282), (2-283), and (2-284)] and pinched-FET [Figure 2-132 and Eqs. (2-285), (2-286), and (2-287)] measurements to derive some of the intrinsic or extrinsic values.

$$Z_{11} = R_s + R_g + R_{gs} + \alpha_g R_{ch} + j\omega(L_g + L_s) \tag{2-282}$$

$$Z_{12} = Z_{21} = R_s + \alpha R_{ch} + j\omega L_s \tag{2-283}$$

$$Z_{22} = R_s + R_d + 2\alpha R_{ch} + j\omega(L_s + L_d) \tag{2-284}$$

Capacitance Equations for Figure 2-132

$$C_b = -\frac{\text{Im}(Y_{12})}{\omega} \tag{2-285}$$

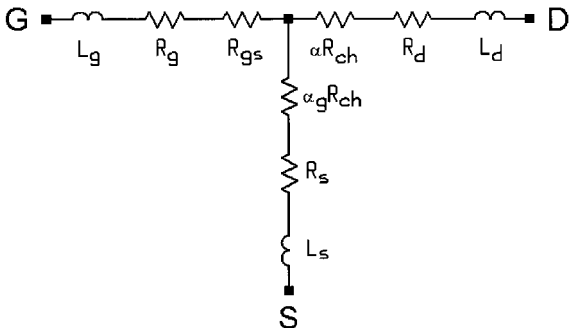


Figure 2-131 Forward-biased gate FET model for cold-FET (V_{DS} = 0, V_{GS} 0) measurements. This approach allows direct extraction of parasitic resistances and inductances.

$$C_{pg} = \frac{\mathrm{Im}(Y_{11})}{\omega} - 2C_b \tag{2-286}$$

$$C_{pd} + C_{ds} = \frac{\mathrm{Im}(Y_{22})}{\omega} - C_b \tag{2-287}$$

For the highest accuracy, the following extraction strategies are recommended:

- Pre-extraction using basic dc and *S* parameters—for example, C_{GS} from S_{11}
- Six extraction steps to determine parameters:
  Basic dc model
  Advanced dc model
  Basic ac model
  Advanced ac model
  Package determination
  Diode characterization
- Large-signal validation
  Power compression and bias shift
  Harmonic power comparison
  Intermodulation characterization

As an example, we are going to use the popular and well-characterized NE71000, which is not a particularly modern device, but has been used for various research projects. Figure

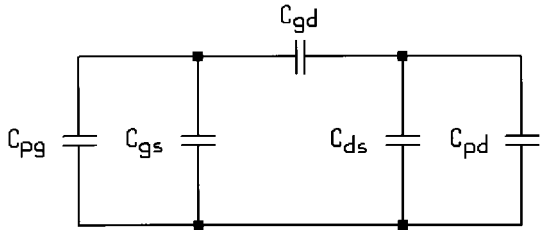


Figure 2-132 Pinched-FET (V_{DS} = 0, V_{GS} *VP*) model. This approach allows direct extraction of parasitic gate and drain capacitances.

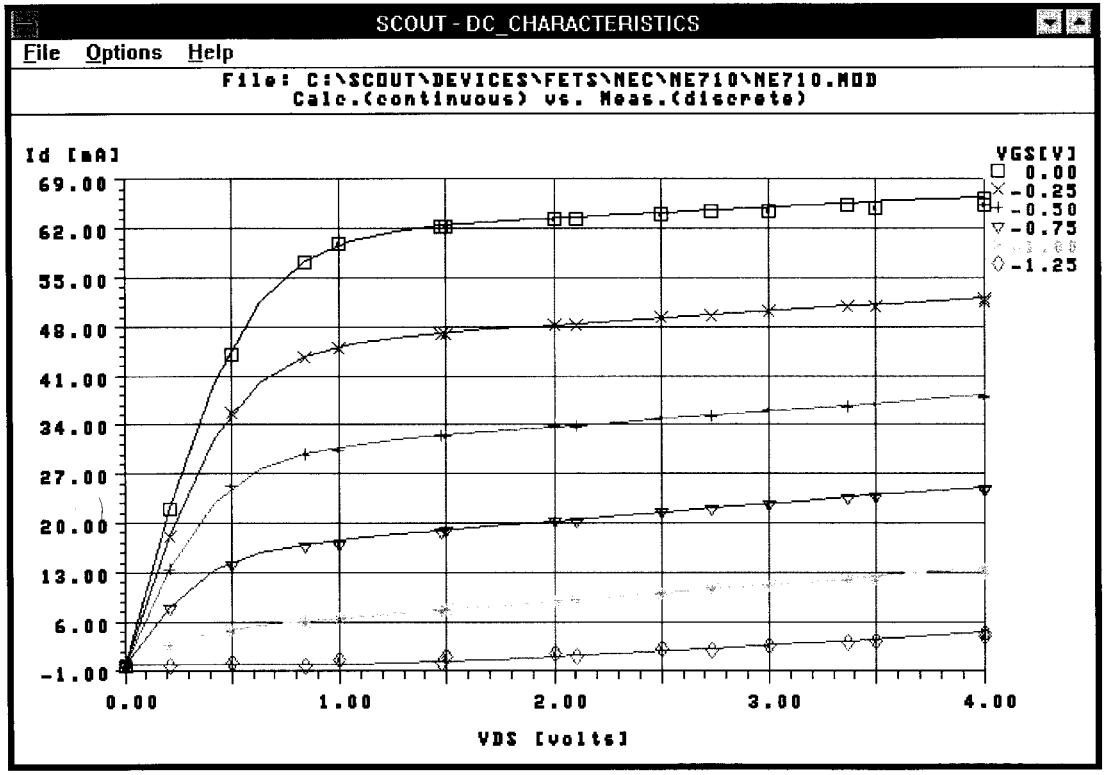


Figure 2-133 Measured NE71000 dc *I–V* curves versus those simulated using the Ansoft Materka model.

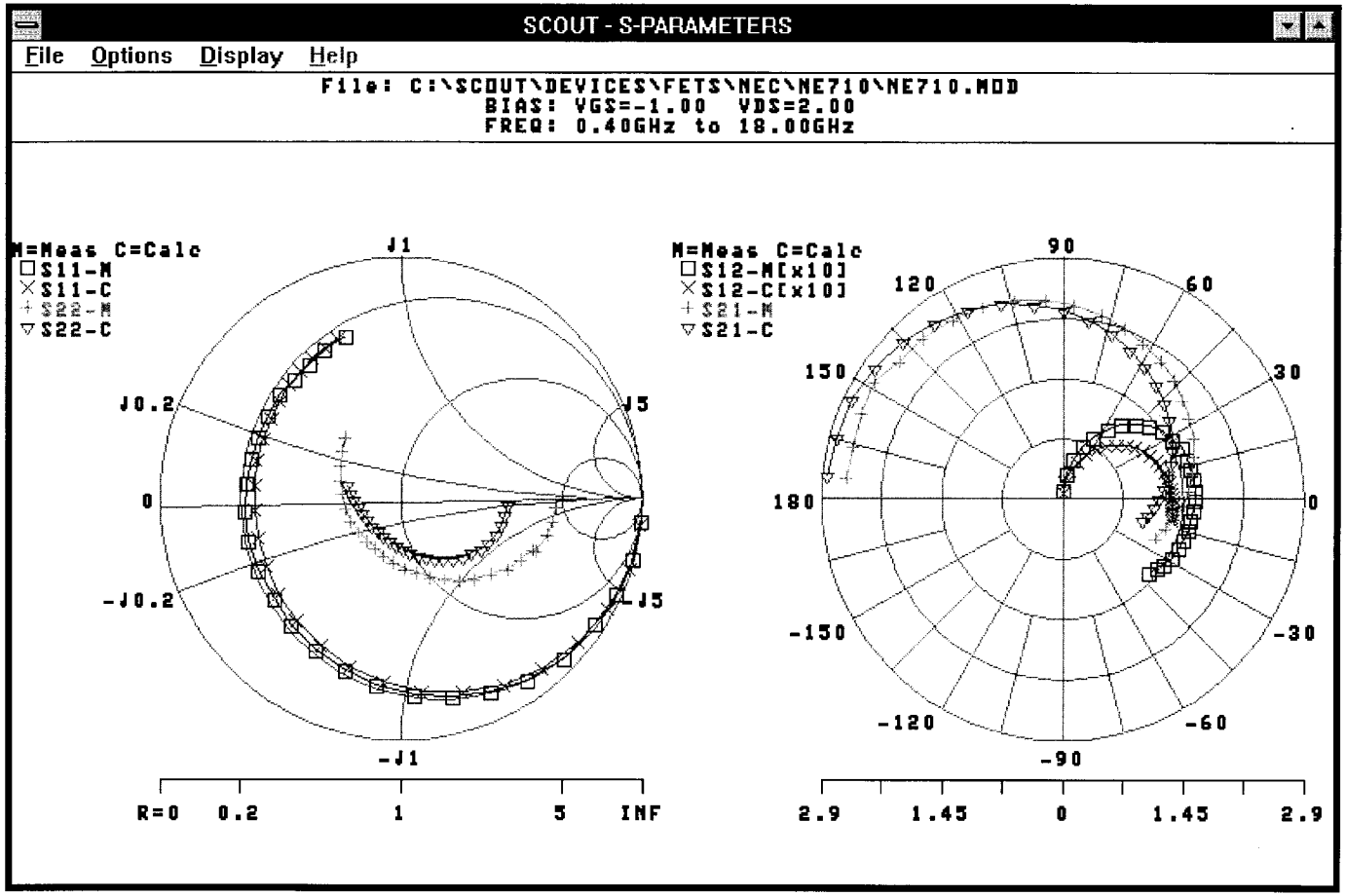


Figure 2-134 Measured versus calculated NE71000 *S* parameters for $V_{GS} = -1.00$ and $V_{DS} = 2.00$.

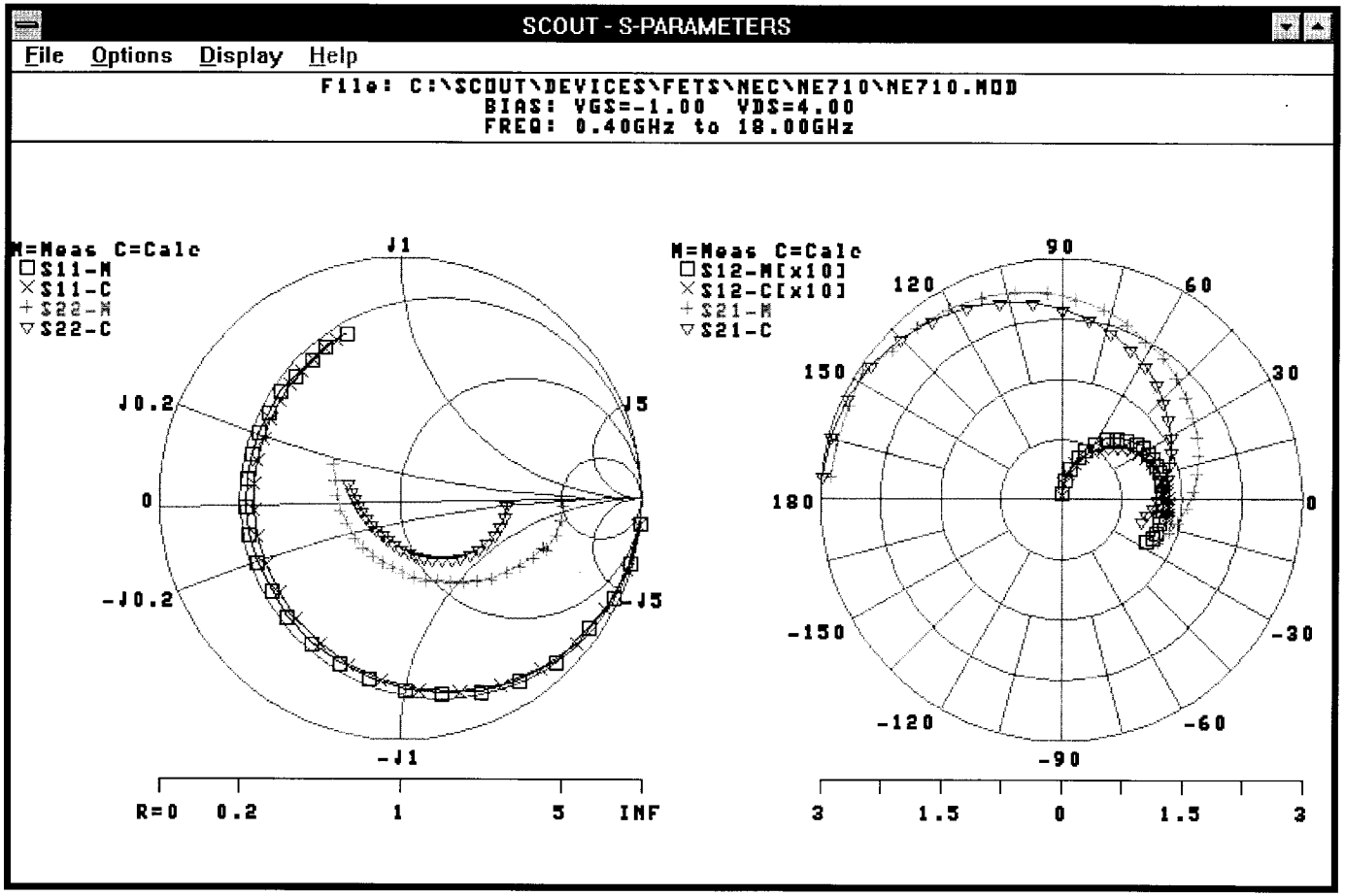


Figure 2-135 Measured versus calculated NE71000 *S* parameters for V_{GS} = –1.00 and V_{DS} = 4.00.

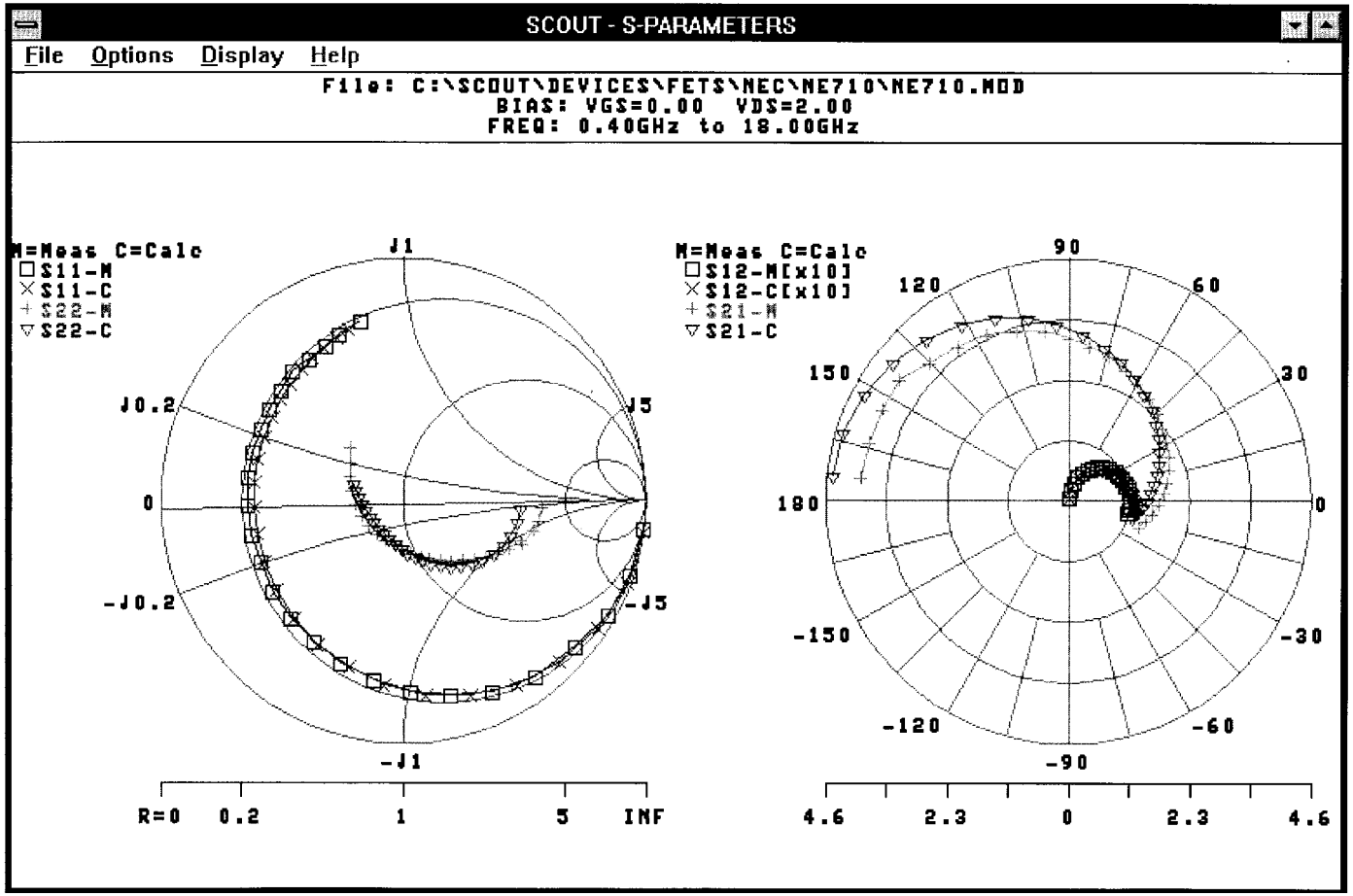


Figure 2-136 Measured versus calculated NE71000 *S* parameters for V_{GS} = 0.00 and V_{DS} = 2.00.

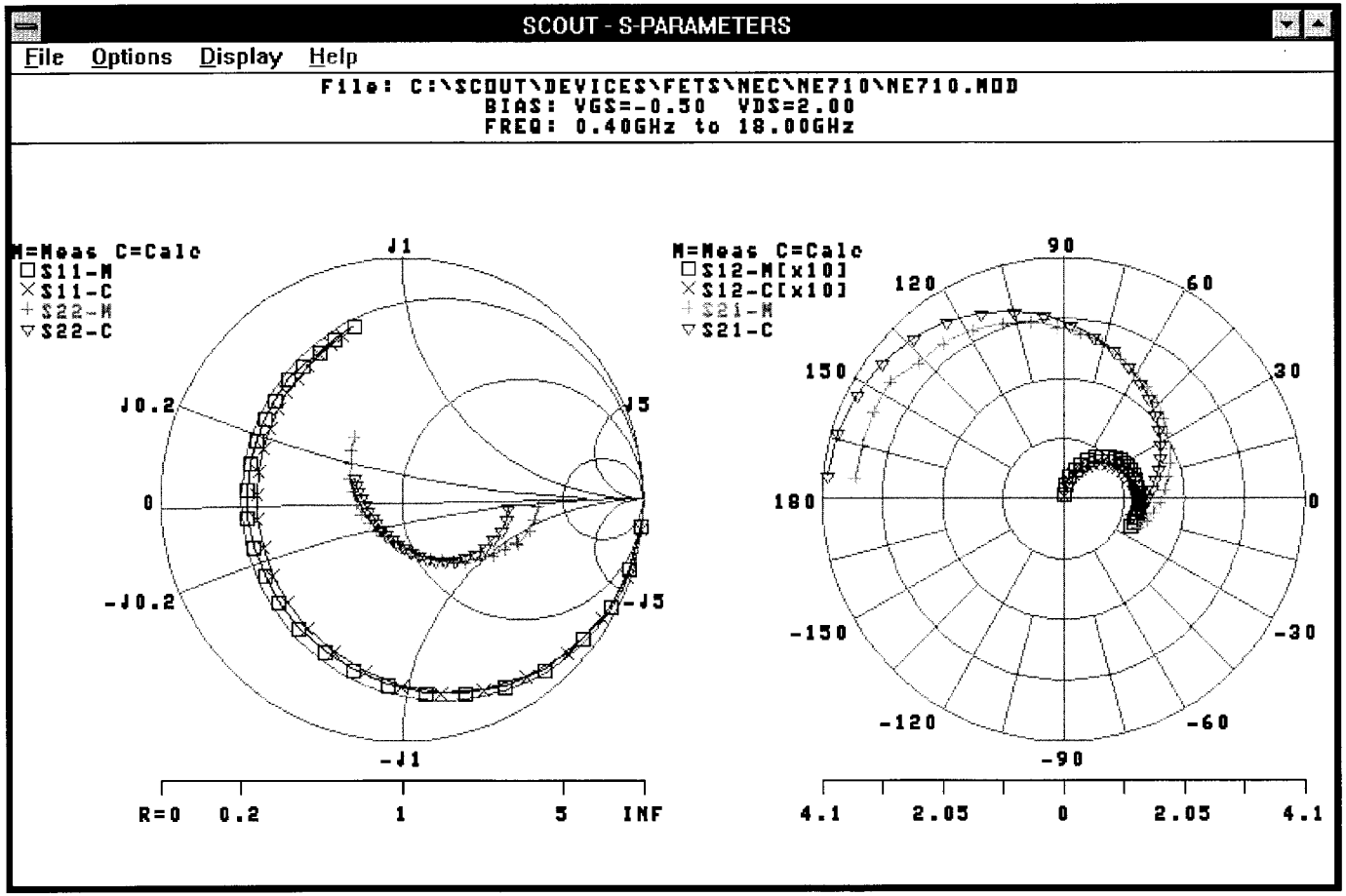


Figure 2-137 Measured versus calculated NE71000 *S* parameters for V_{GS} = –0.50 and V_{DS} = 2.00.

2-133 shows measured versus simulated dc I–V curves for the NE71000. The agreement between measured and predicted dc I–V curves is quite good. In a similar fashion, we obtain the S-parameter fits over bias. They are shown in Figures 2-134 through 2-137.

As mentioned above, 1-dB compression, power saturation, and drain current are functions of the drive. This particular model point is not particularly impressive. We believe this has to do with the original Materka model.

Having brought up the subject of models, Figure 2-138 compares measured data with that obtained with the Angelov, cubic, quadratic, TriQuint, Raytheon, and Materka models. The figure shows that there is no perfect model yet. To put this in perspective, in Figure 2-139 we show the dc I–V error bars for these six models. This graph shows that the least errors for the dc measurement versus modeling are obtained with the Curtice cubic model, followed by the modified Materka model (implemented in Serenade) and TriQuint models. The difference in S parameter matching for RF application is significantly more relevant for the designer.

Things become much more interesting if we compare the S-parameter errors between the Materka shown in Figure 2-140, the Curtice cubic model (Figure 2-141), and the TOM model (Figure 2-142). Improved modeling regarding the R_{DS} is being developed, and we will discuss this at the end of this chapter.

Since the trend is to go to low-voltage power RF operation, several issues must be battled:

- Models are most inaccurate at low V_{DS} and near pinchoff.
- Modeling of low-voltage operation is adversely impacted.
- Thermal effects must be considered, but older models do not account for this.

A Case Study. As a case study, we will consider the Oki KGF1608 power MESFET (0.5 μm × 28 mm). Figures 2-143 and 2-144 show the dc I–V curves and S parameters, respectively. The resolution of the S parameters, based on the scale used, is not particularly impressive.

Figure 2-145 shows the dc I–V curves, with load line, for the KGF1608; the self-heating effect of the transistor can be seen in the droop of the top two curves. Figures 2-146 and 2-147 show drain current and output power versus drive, respectively. The test conditions for these three figures were f = 850 MHz, V_{DS} = 3.4 V, and V_{DS} = 2.025 V.

2-4-5 Conclusions

From all plots of dc and ac error contours, dc I–V curves, and dc error bar charts we can draw the following conclusions.

Dc errors

- *In the Saturation Region*. All the models considered work very well, with a relative error of less than 15%. This means that all models are suitable for typical applications where the device operates in the saturation region.
- *In the Linear Region.* Though the errors in the saturation region are similar for different models, the errors in the linear region are quite different. The Materka, Curtice cubic, and TriQuint models behave better than other models in the linear region. The relative errors of these three models are less than 10%, while other models have more than 60% relative errors at some points.

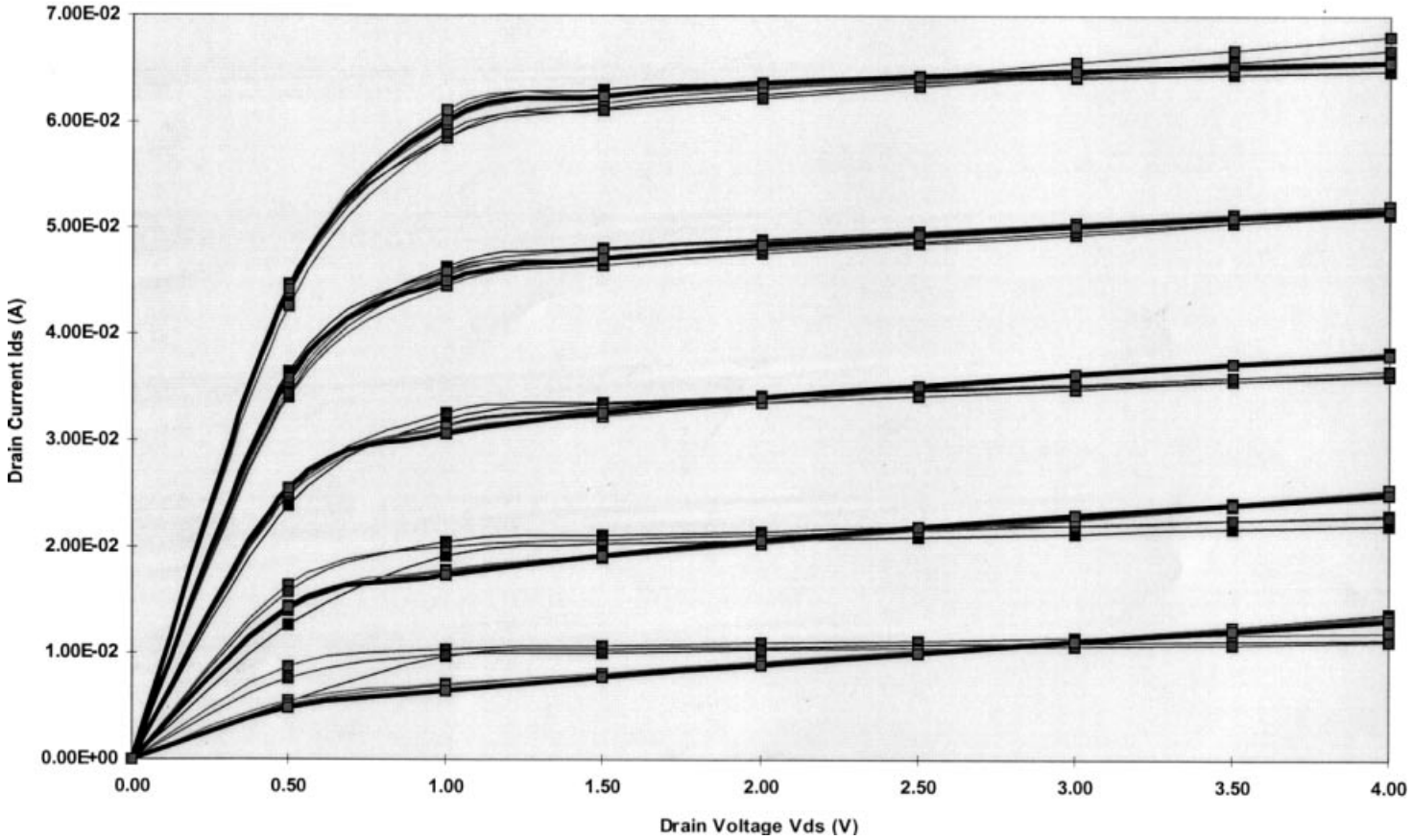


Figure 2-138 Comparison of dc models: Chalmers (Angelov), Curtice cubic, Curtice quadratic, Materka, Raytheon, and TOM (TriQuint).

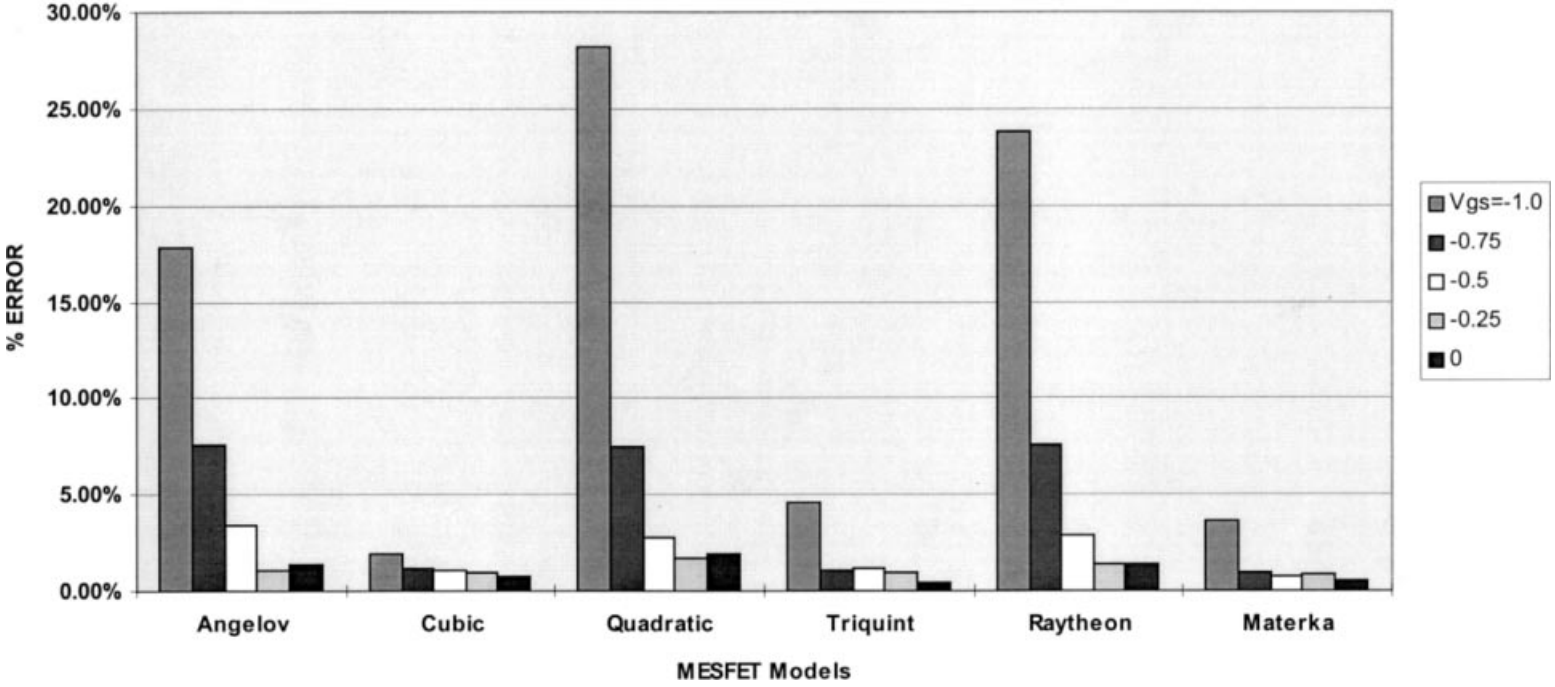


Figure 2-139 Dc *I–V* error bars: Chalmers (Angelov), Curtice cubic, Curtice quadratic, Materka, Raytheon, and TOM.

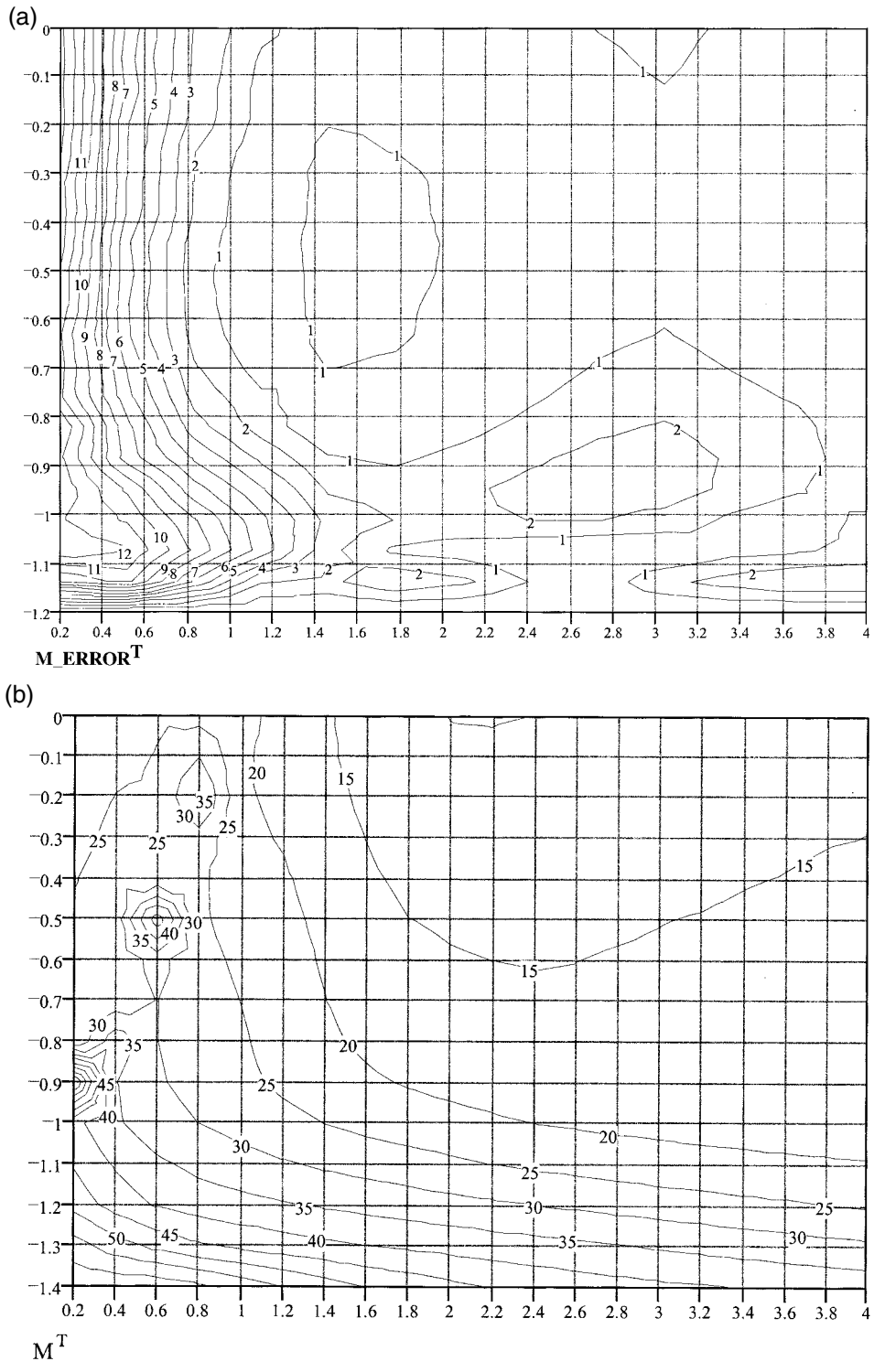


Figure 2-140 NE71000, Materka model, % error contours for (a) dc *I–V* and (b) *S* parameters.

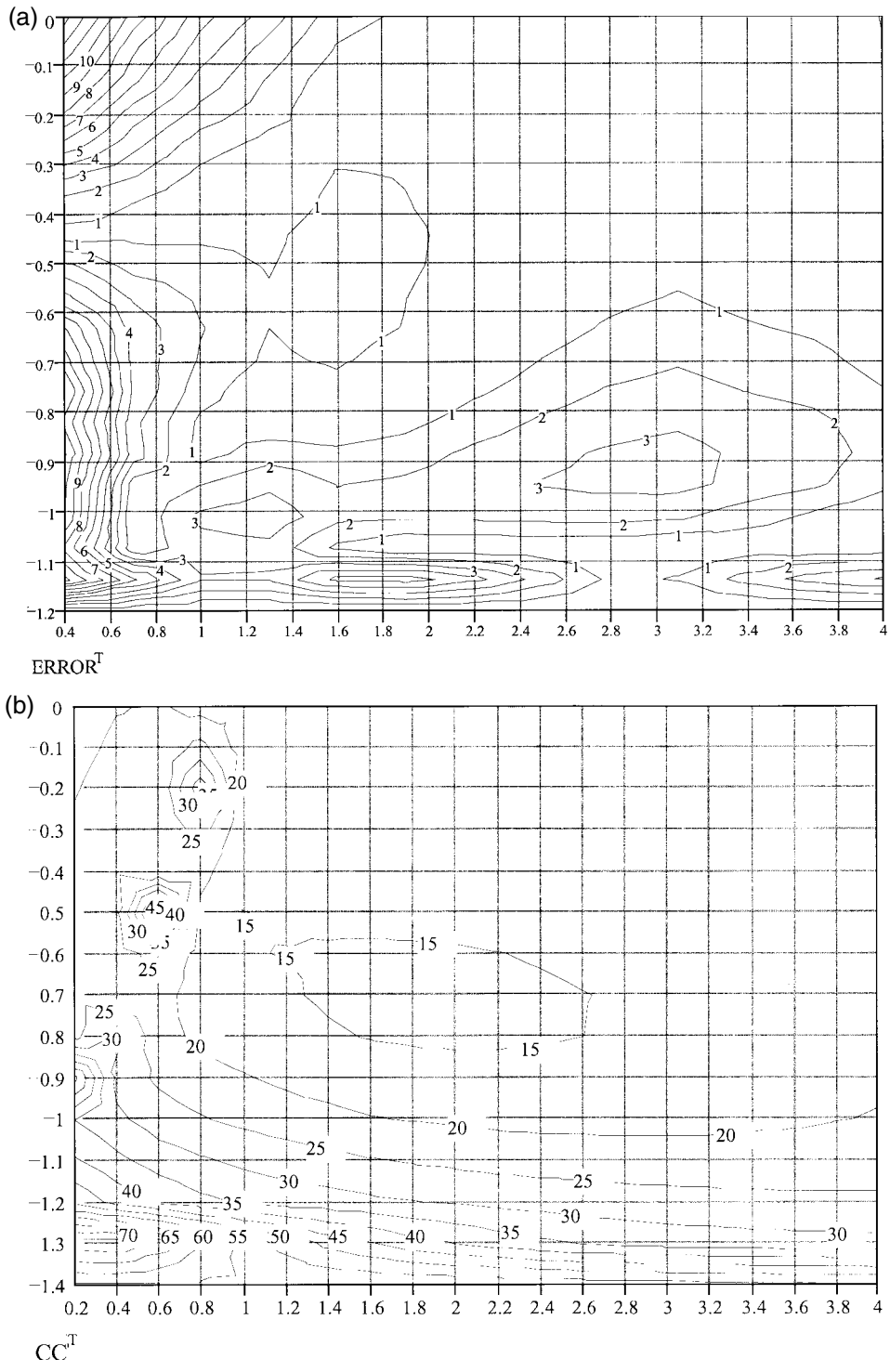


Figure 2-141 NE71000, Curtice cubic model, % error contours for (a) dc *I–V* and (b) *S* parameters.

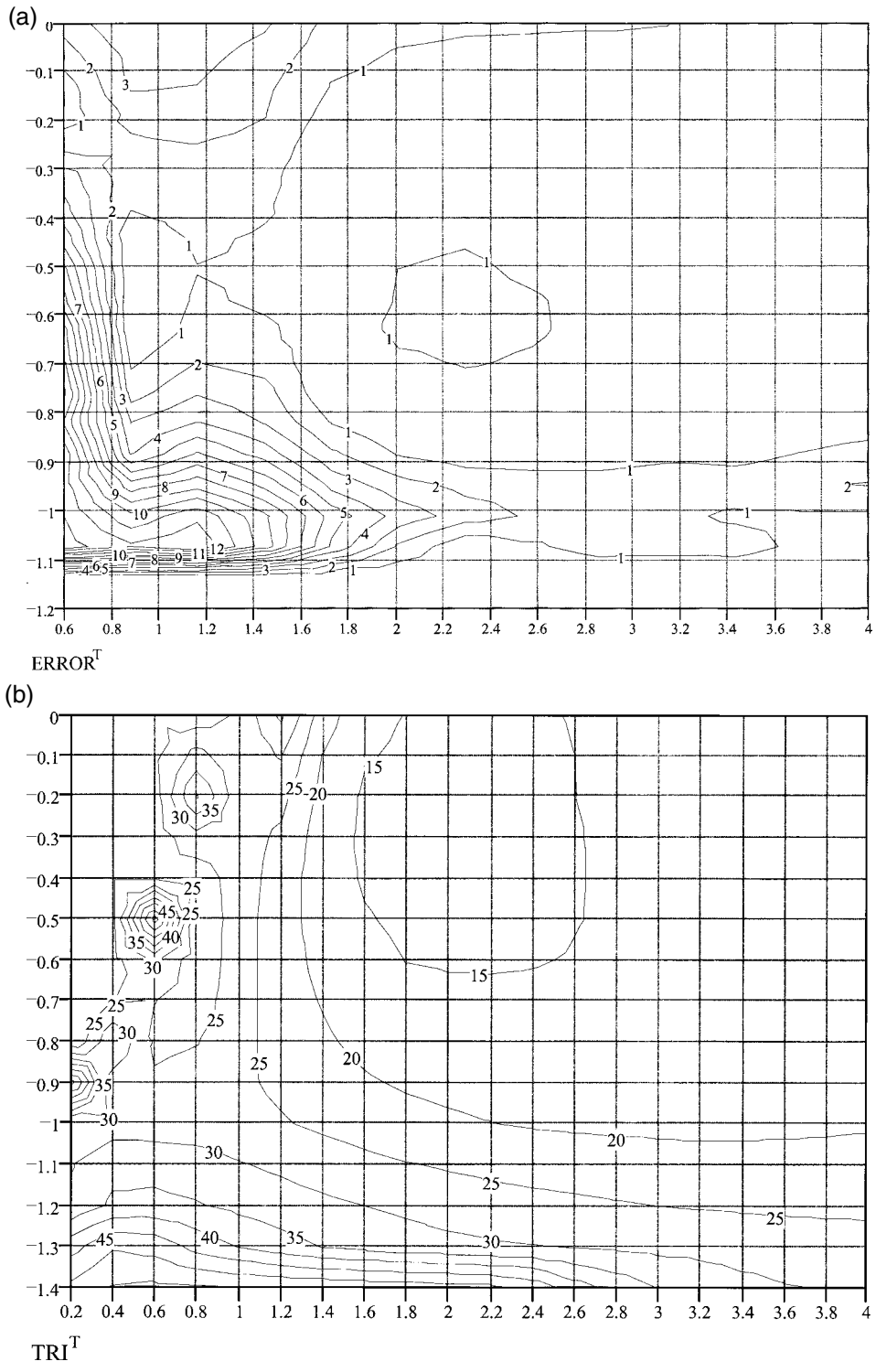


Figure 2-142 NE71000, TOM model, % error contours for (a) dc *I–V* and (b) *S* parameters.

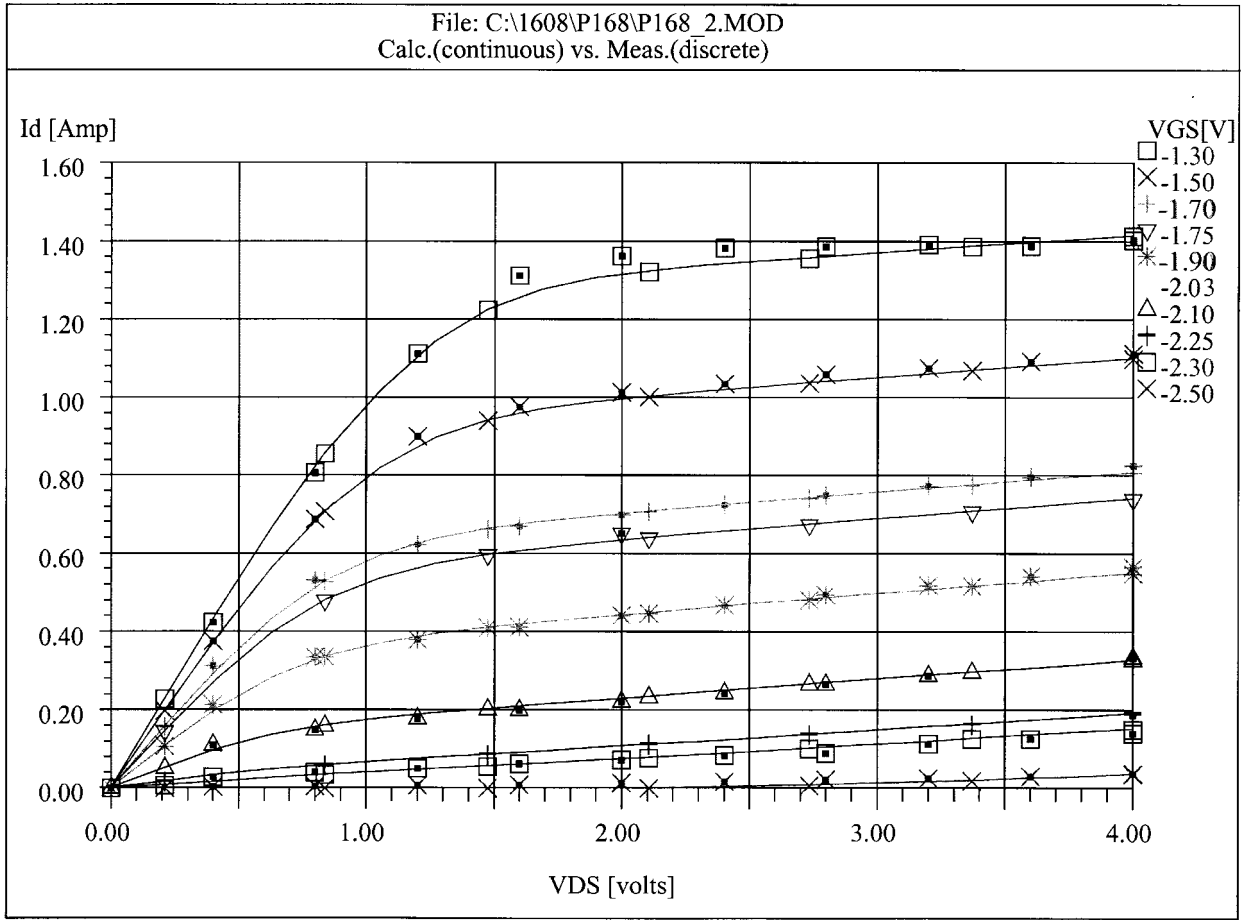


Figure 2-143 Calculated versus measured dc *I–V* curves for the Oki KGF1608 package MESFET.

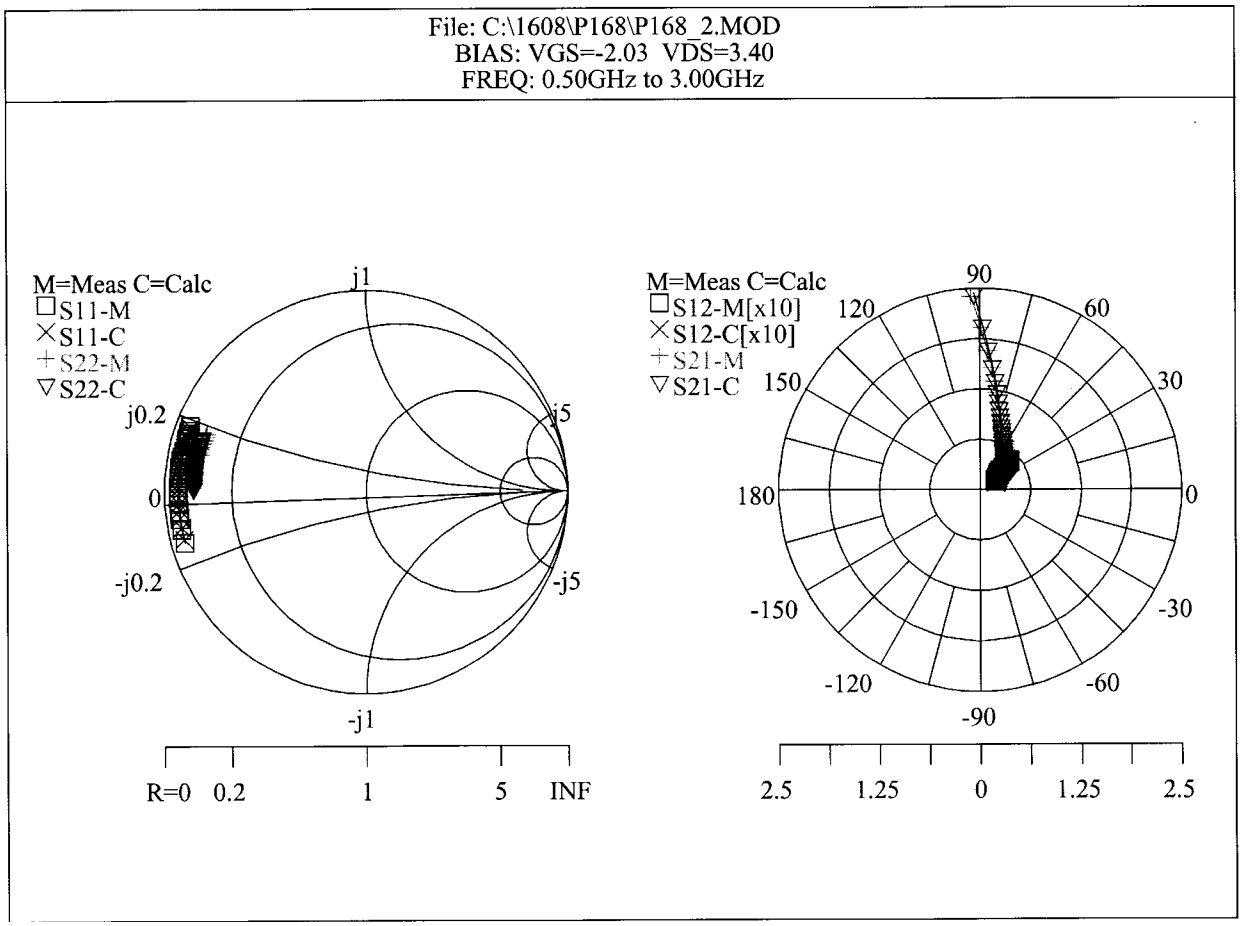


Figure 2-144 Calculated versus measured *S* parameters for the Oki KGF1608 package MESFET.

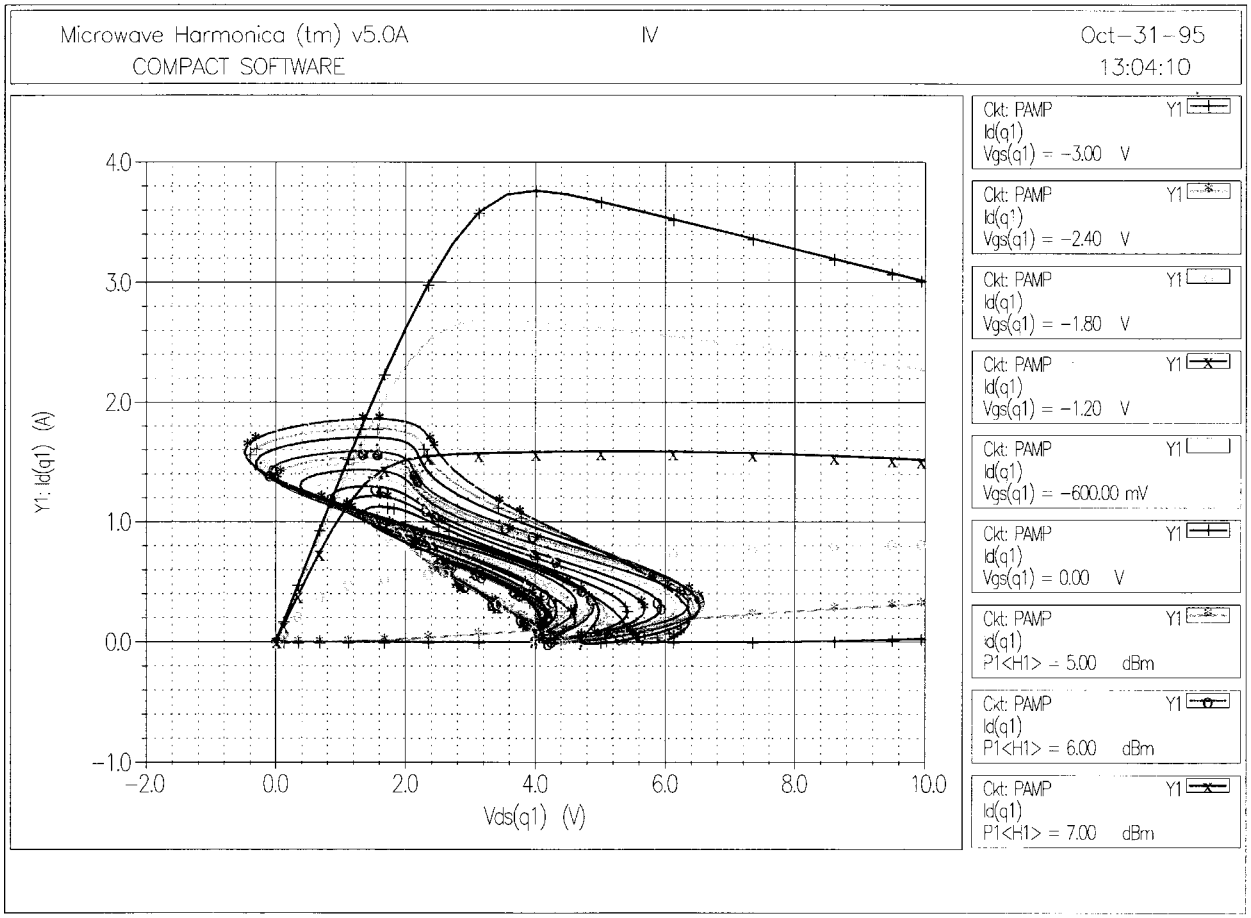


Figure 2-145 Load line and dc *I–V* curves for the KGF1608 MESFET.

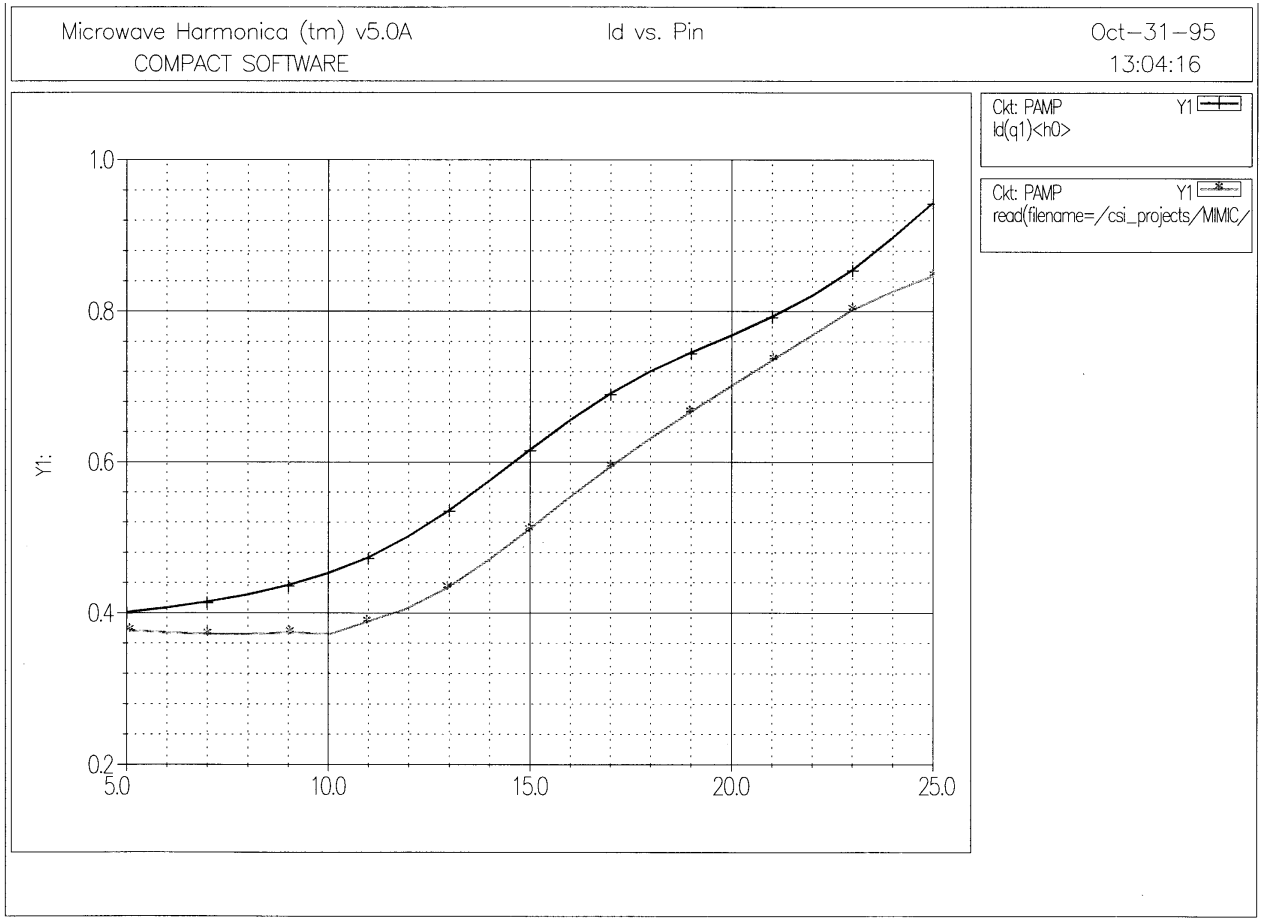


Figure 2-146 Measured versus calculated drain current for the KGF1608 MESFET.

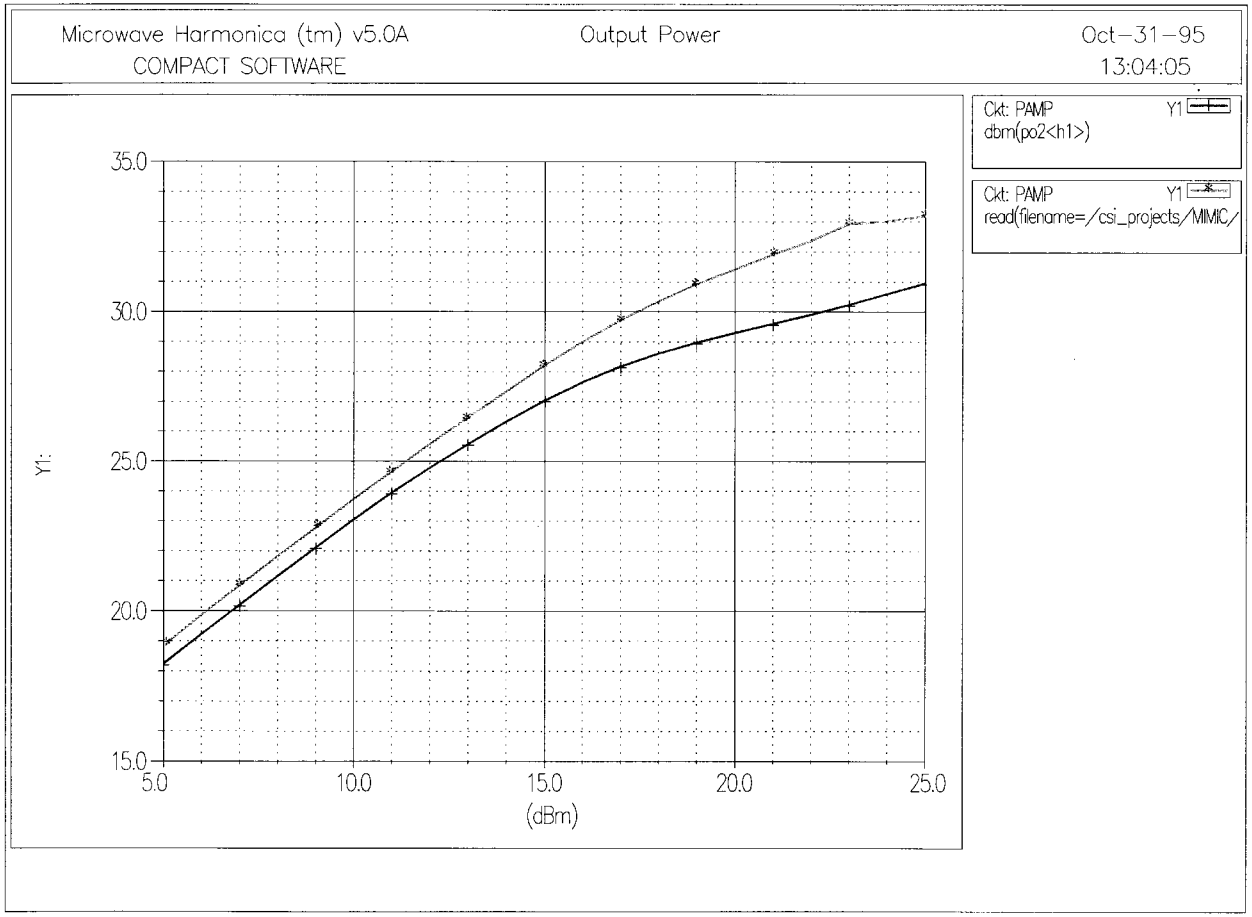


Figure 2-147 Measured versus calculated output power for the KGF1608 MESFET.

ac Errors. From the ac error contour plots we can see that the Chalmers (Angelov) model emerges as the most accurate for ac small-signal operation. The comparison also shows that all models considered are good for ac operations biased in the saturation region with a typical relative error of less than 15%. However, the ac relative errors are much higher in the linear region, particularly for the Curtice quadratic model, which exhibits more than 80% error at some points.

Notes. The above conclusions are based on the matching between the simulated responses of different models and the measured data from a typical NE71000 device; the models may behave differently for a different device. It is not possible to draw any broad-based conclusions as to what model is the best. In order to select the most suitable model for your application, you must consider all the factors, such as bias range, frequency range, and small-signal or large-signal operation, according to your experiences in device modeling.

2-4-6 Device Libraries

All these measurements lead to a nonlinear device library. Figures 2-148 and 2-149 shows a typical datasheet for a device in the library with measured and modeled data. The previous high-power example does not always lead to acceptable results as the models are far from being fully developed. We also show the simulation results using the physics-based model for FETs.

Physics-Based MESFET Modeling. We mentioned a physics-based model earlier. Figure 2-150 shows the measured versus predicted dc *I–V*, I/O power, and *S* parameters for an Alpha Industries MBE MESFET developed under a government contract. The simulation does not handle the trapped energy, which explains the negative slope on the third trace of the output dc *I–V* curve.

The process for bipolar transistors is quite similar and less tricky than the ones just described. Good examples can be found in the user manual for Compact Scout.

2-4-7 A Novel Approach for Simulation at Low Voltage and Near Pinchoff Voltage

As mentioned earlier, the most critical zone is the area of operating the FET both close to the pinchoff voltage and at low operating voltages. A crucial reason for the inadequate results is the modeling of the drain–source resistance as a function of bias point. This is particularly important in calculating the power-added efficiency. When doing your modeling, there is always a temptation to look at both high currents and operating voltages. Figure 2-151 shows a reasonable agreement between measured and predicted dc *I–V* performance. However, the very moment the transistor is modeled close to the linear region (Figure 2-152), things change drastically and the results from 0 to 6 mA have little resemblance to those from 0 to 50 mA. The modified Materka model (Figures 2-153 and 2-154) gives a significantly better match than the Angelov model.

In order to overcome the modeling inaccuracy for the FETs, Thomas Meier of Siemens AG in Munich has come up with the concept of combining two transistors in parallel, whereby the second one is not really connected for dc (Figure 2-155). Figure 2-156 compares measured and simulated R_{DS} using this new approach.

NE67300

Manufacturer : **NEC**
Model Keyword : **NE - 67300**
Device Type : **Mesfet**
Model Type : **Materka**
Package Type : **Chip**

Electrical Characteristics:
Idss : 40mA
Vp : 1.1V
Gain : 8.5dB @ 18GHz
P1db : N/A
Fmin : N/A

Model Data Range:
Vgs : -1.2V to 0.0V
Vds : 0.0V to 4.0V
Freq: 0.4GHz to 18GHz

Verification :
Type : Power sweep, 50Ω
Vgs : -0.6V Vds : 3.0V Ids : 30mA Freq: 6.0GHz

Comments : The S-parameters that appear in the figure were measured and simulated with wire bonds present. The length of each wire bond is approximately 30 mils. To obtain the S-parameters for the manufacture's data sheets, the wire bond inductances would be removed from the equivalent circuit.

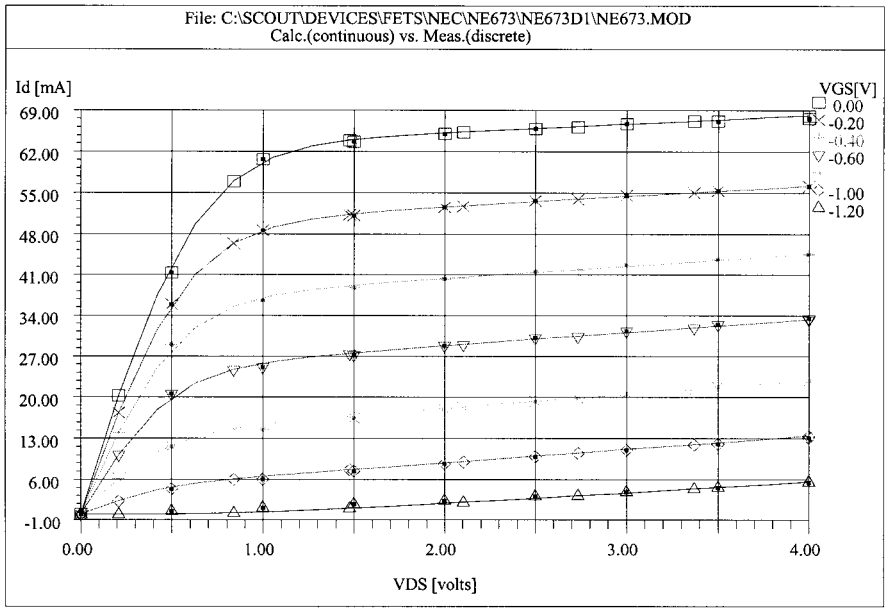
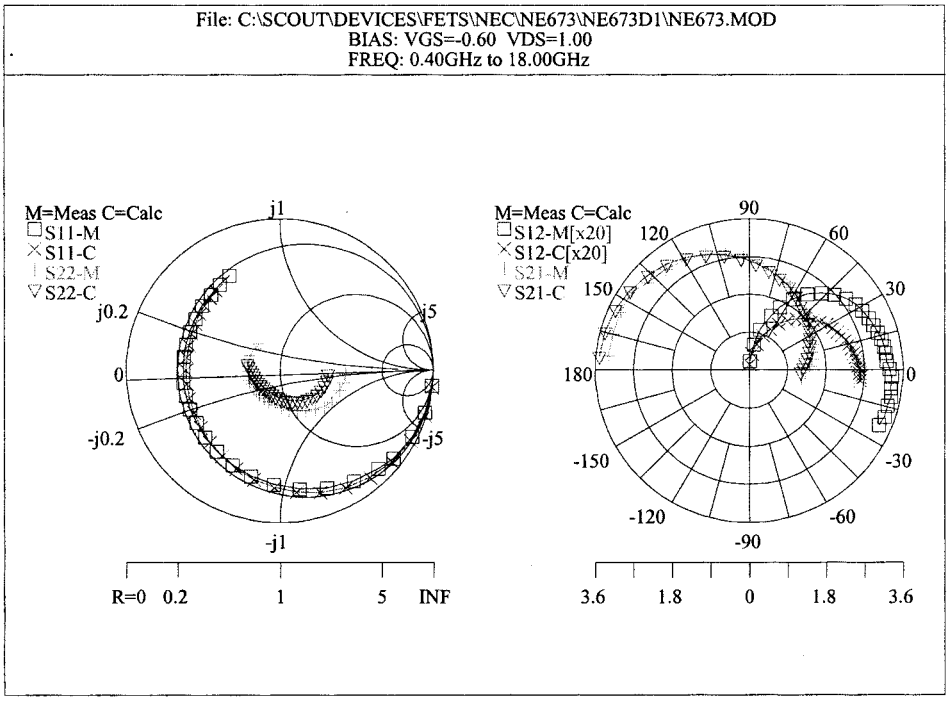


Figure 2-148 Page 1 of a nonlinear device library datasheet for the NE67300.



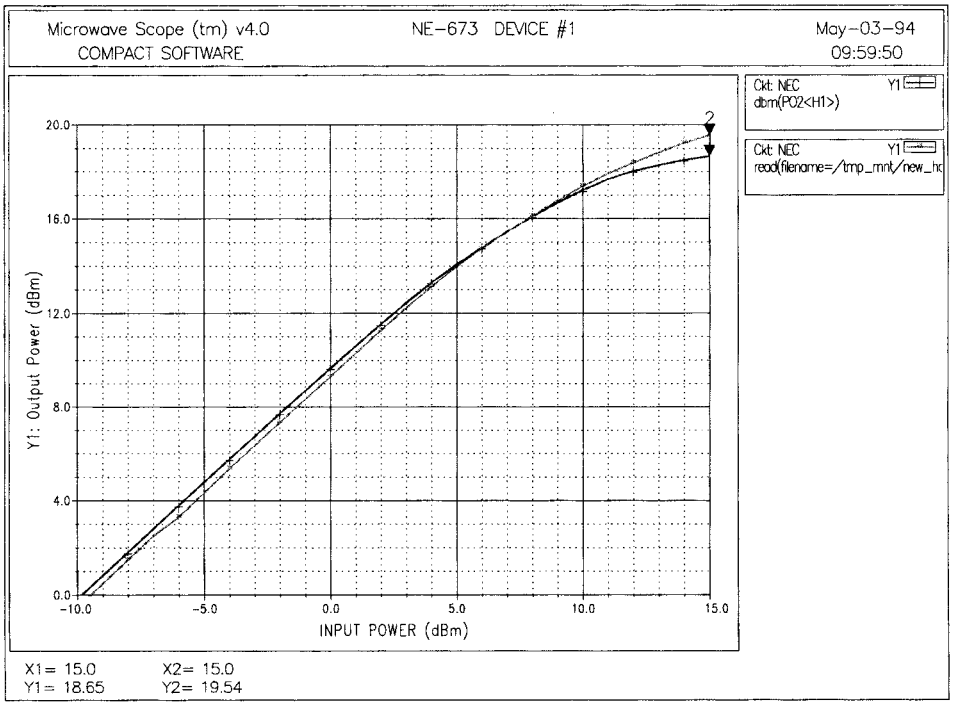


Figure 2-149 Page 2 of a nonlinear device library datasheet for the NE67300.

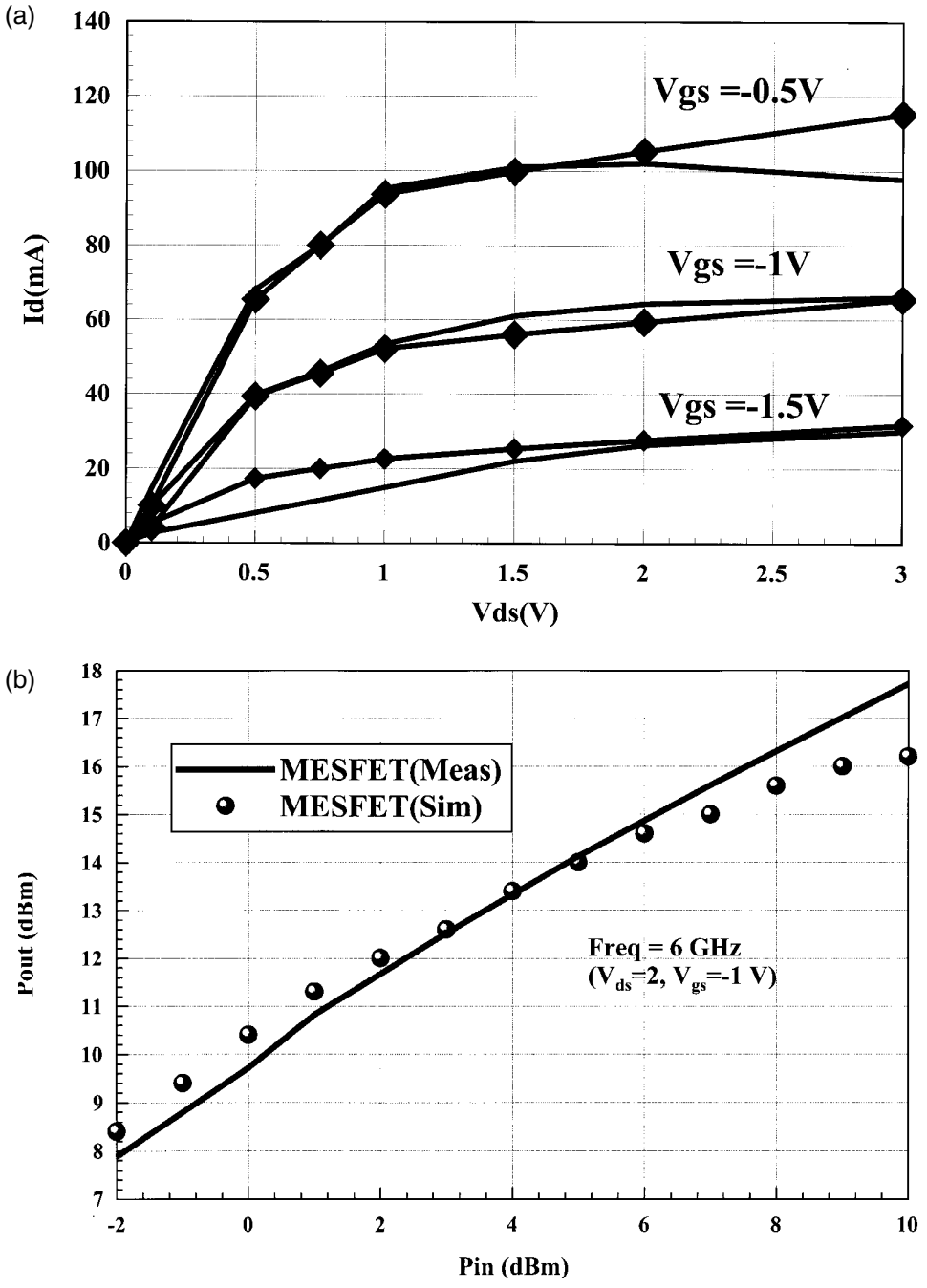


Figure 2-150 Comparison of measured data with PHYSFET and TOM models for an Alpha Industries MBE MESFET (0.25 μm × 400 μm) at $V_{GS} = -1$, $V_{DS} = 2$.

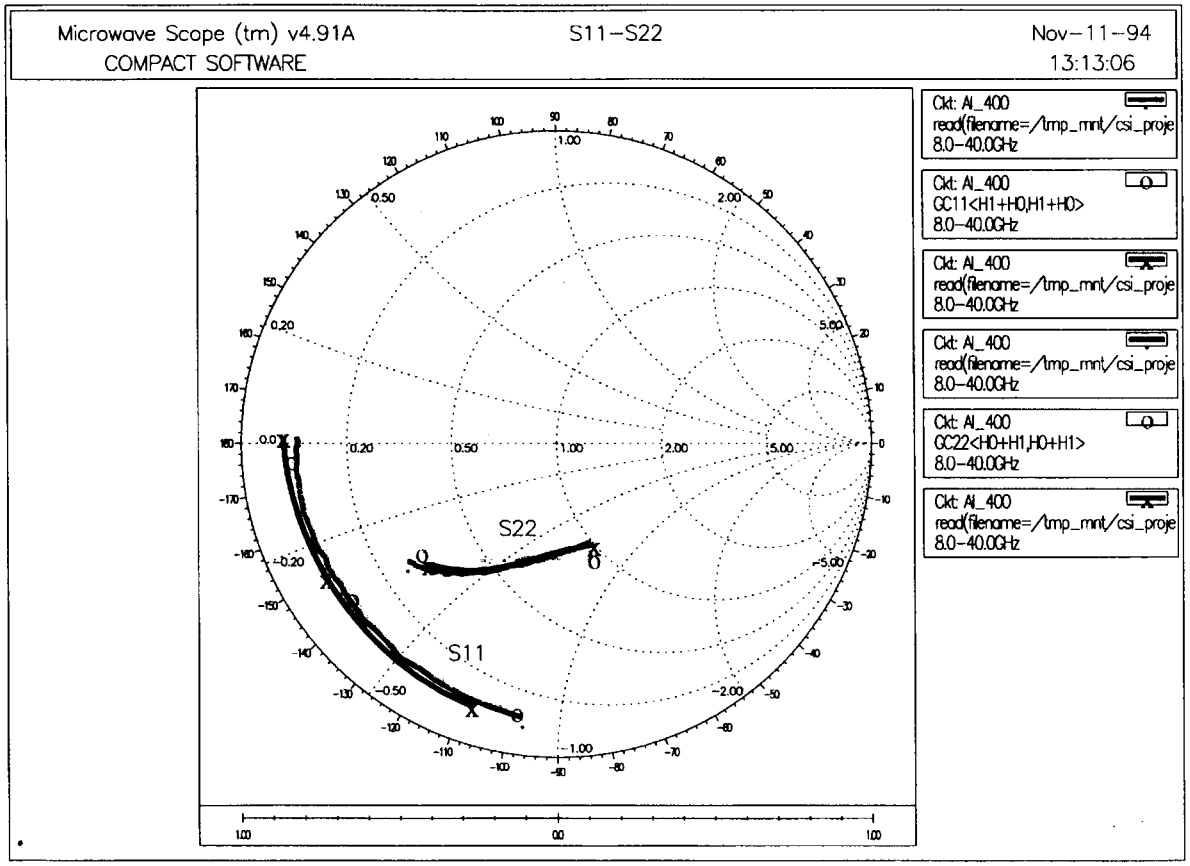


Figure 2-150 (*Continued*)

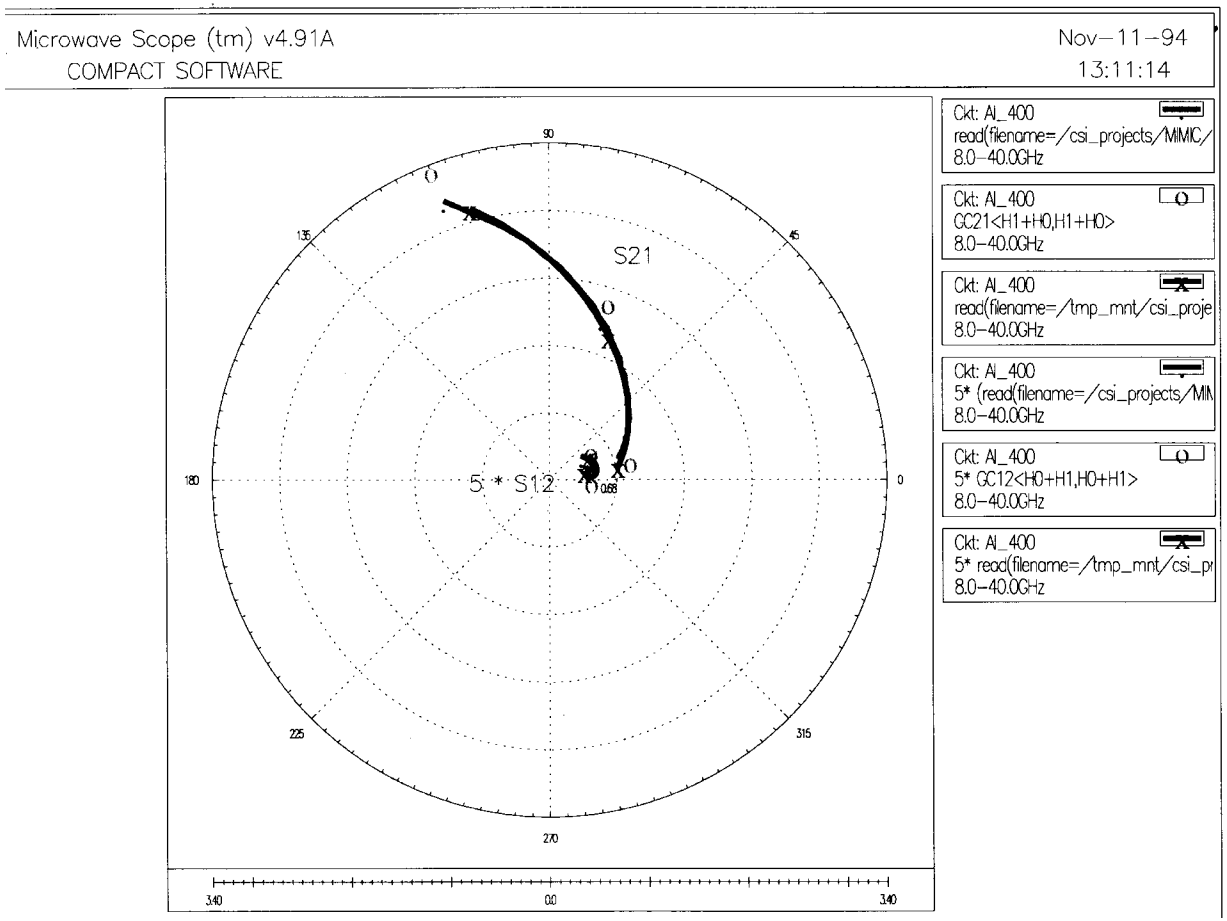


Figure 2-150 (*Continued*)

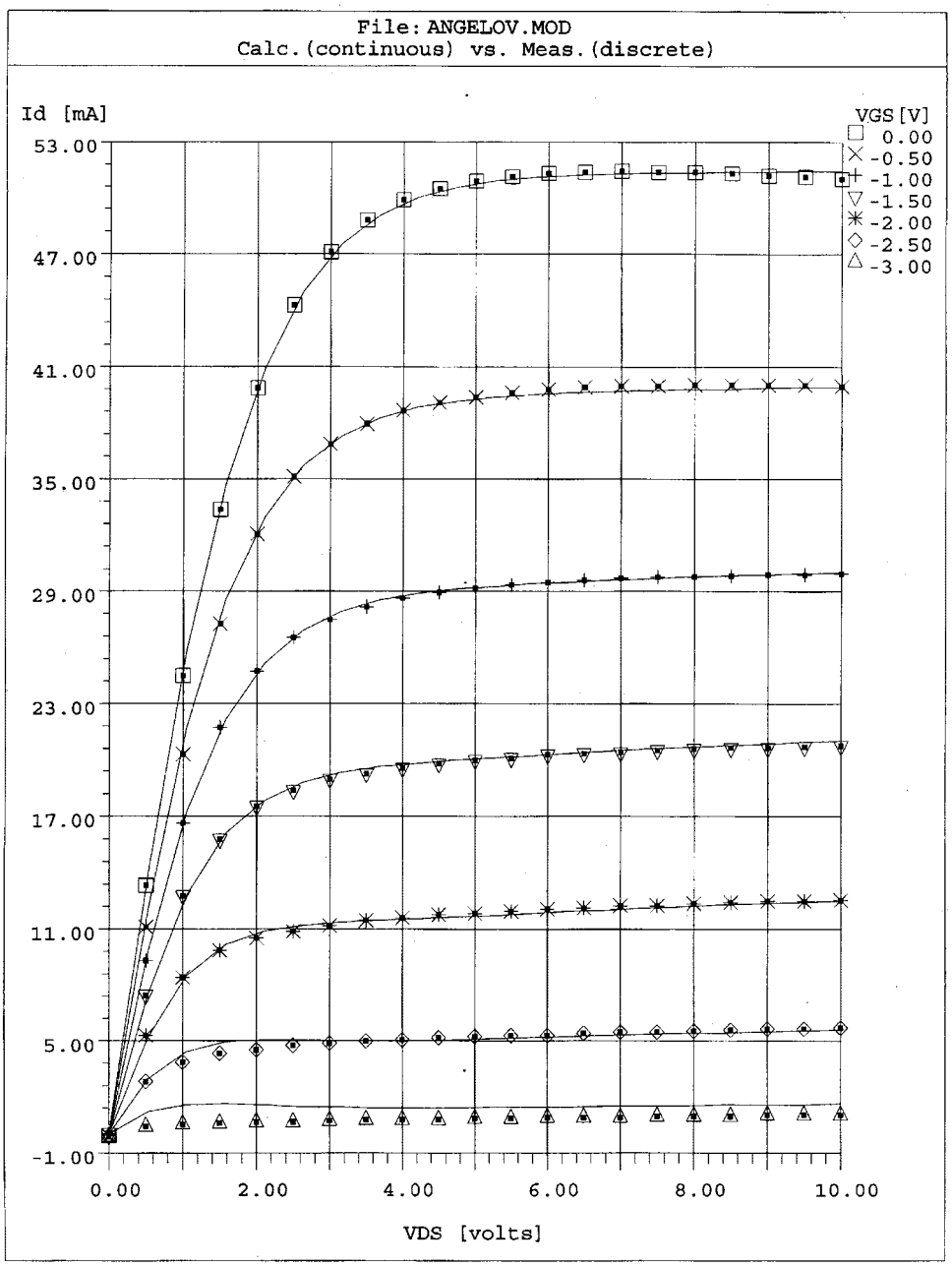


Figure 2-151 High-I_D dc I–V curves for the Angelov FET model.

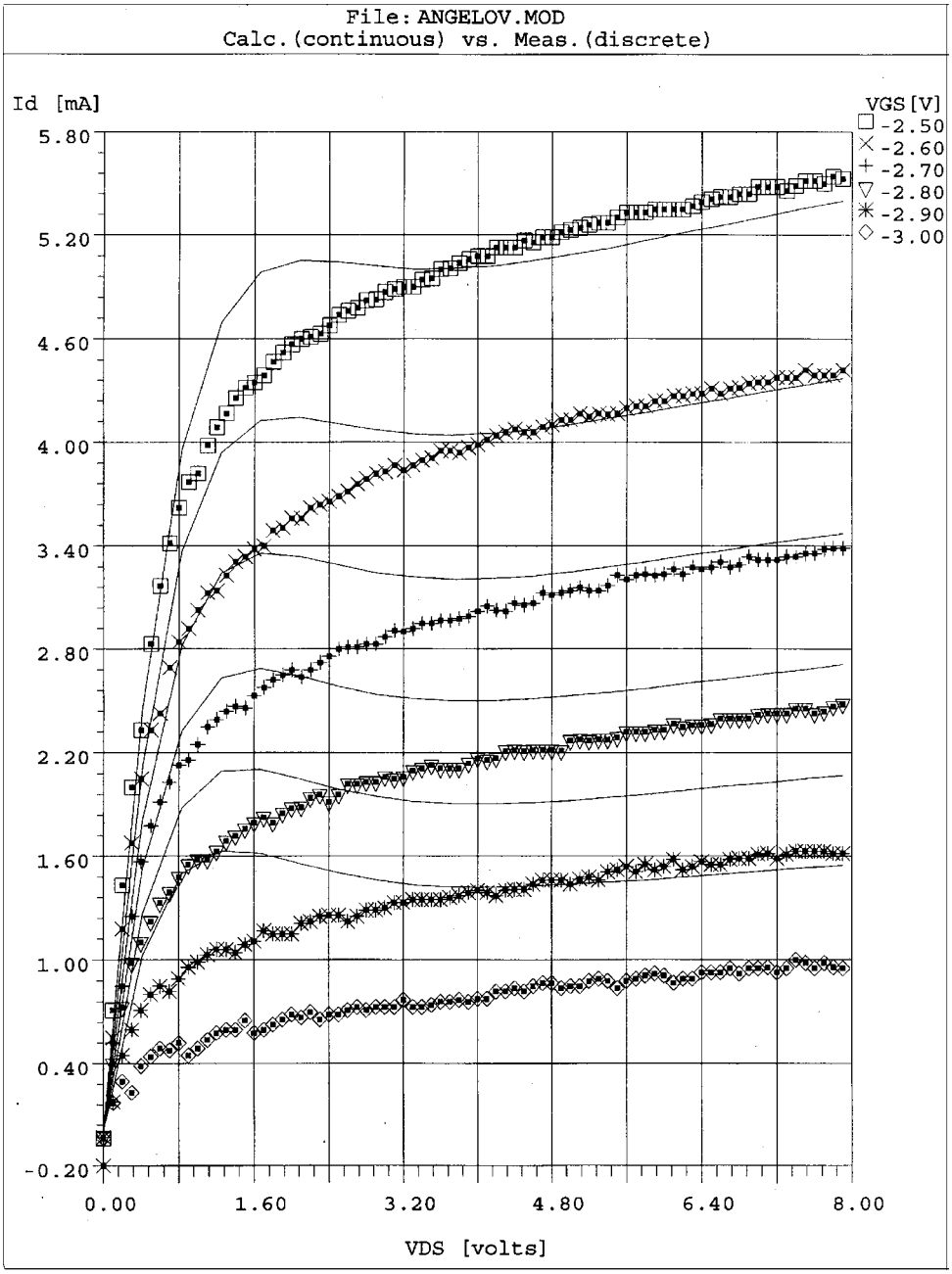


Figure 2-152 Low-I_D dc I–V curves for the Angelov FET model.

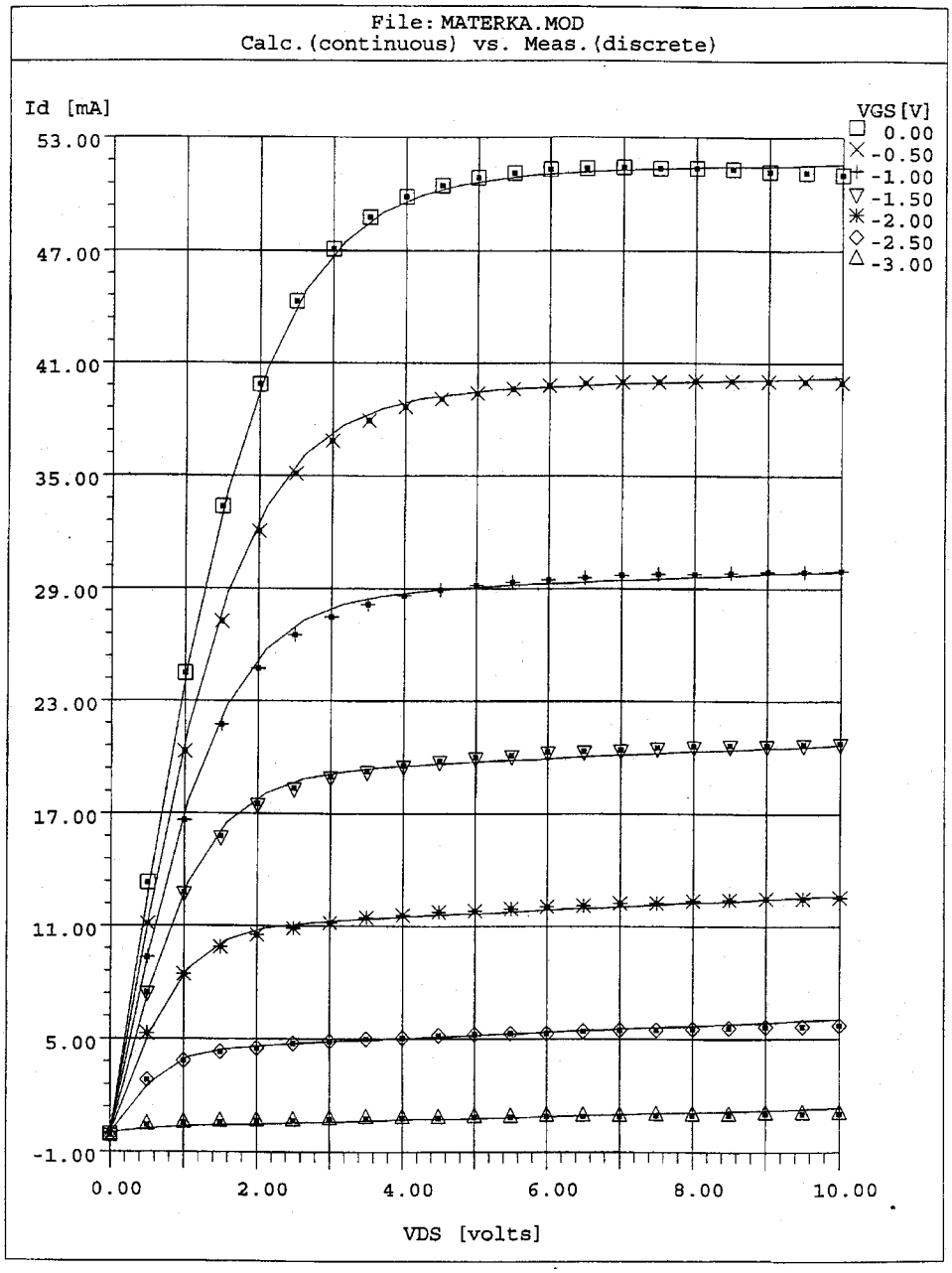


Figure 2-153 High-I_D dc I–V curves for the modified Materka FET model.

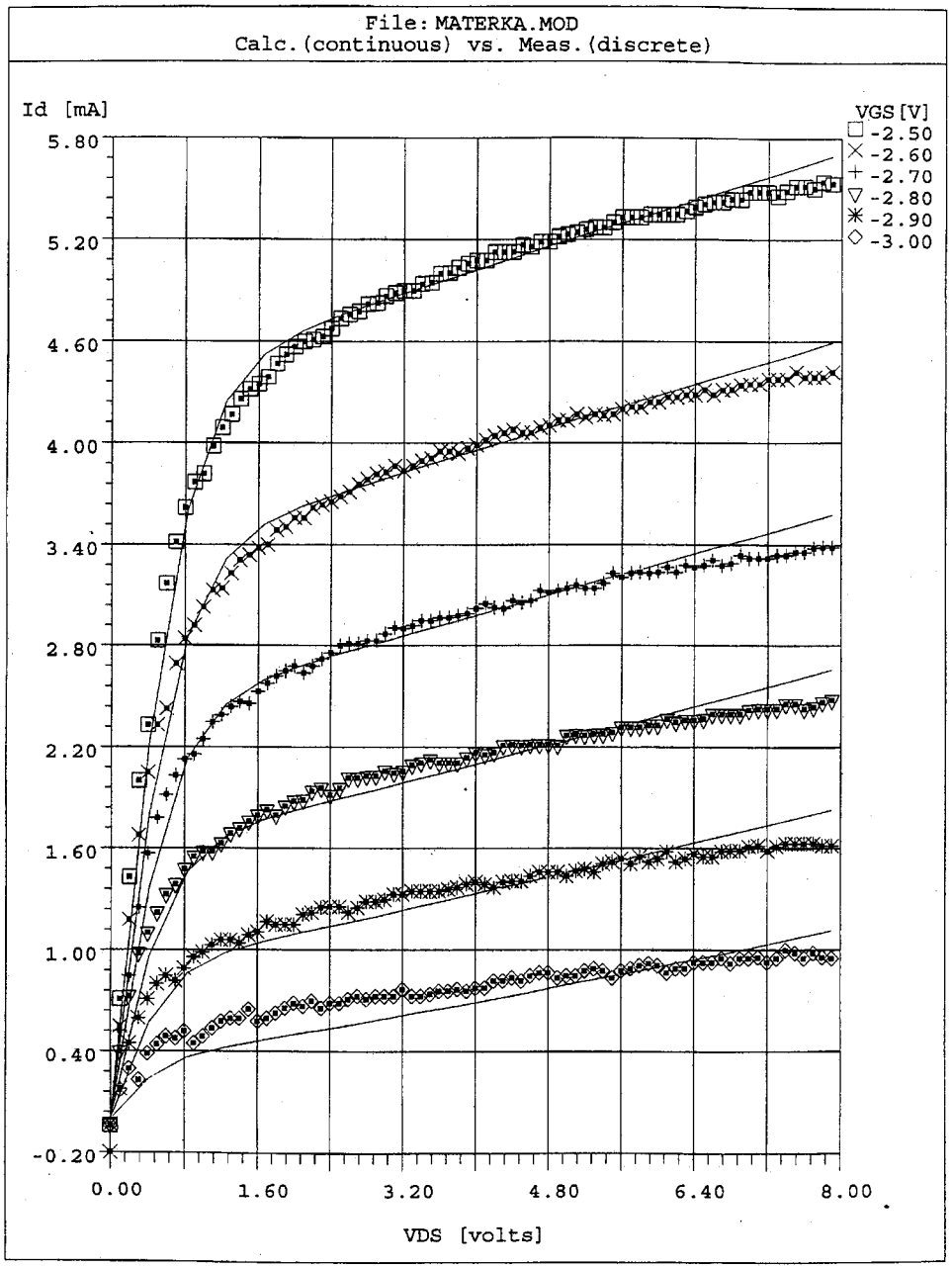


Figure 2-154 Low-I_D dc I–V curves for the modified Materka FET model.

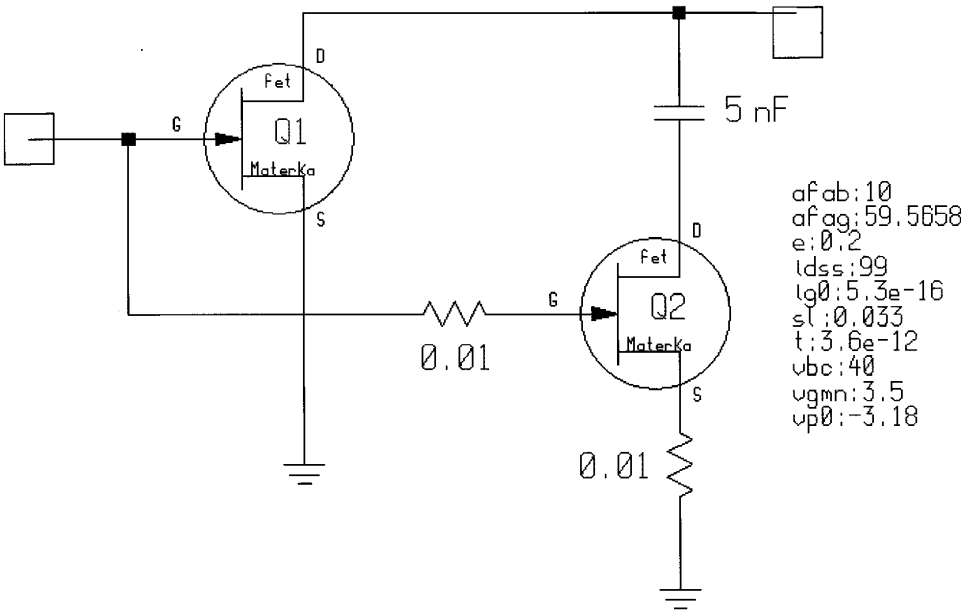


Figure 2-155 By putting two transistors in parallel, with the second one having no dc connection at the drain, one can simulate a dynamic R_{DS} quite well. Regardless of parameter values used for Q1, the parameters listed are necessary to get an accurate R_{DS} result.

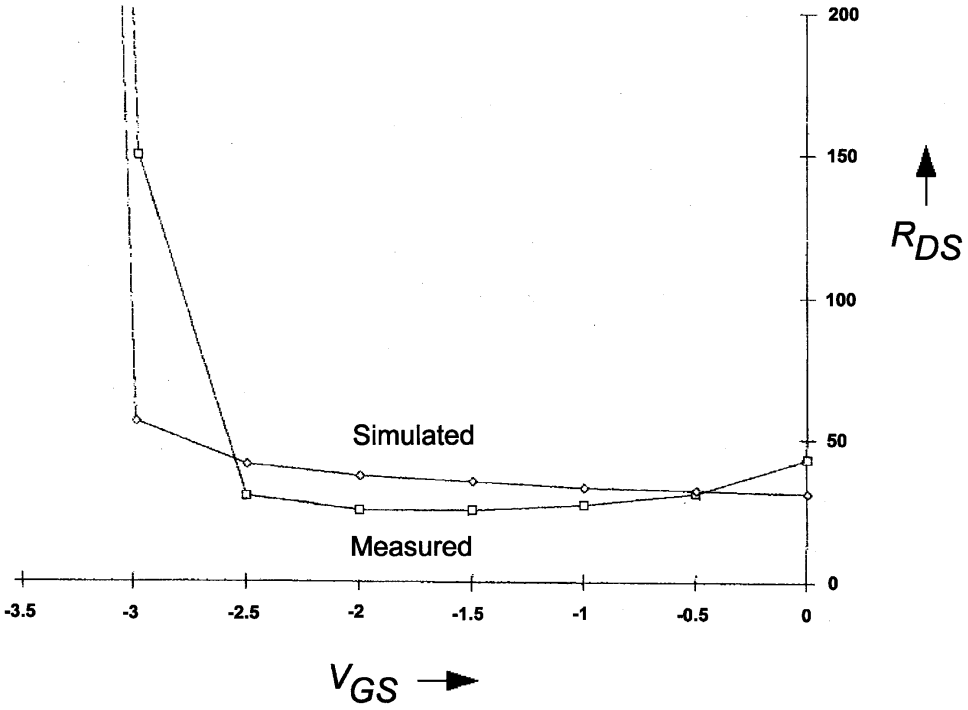


Figure 2-156 Measured versus simulated R_{DS} for the paralleled MESFET approach.

One might ask the question: "Why not fix the existing models?" It turns out that in order to get the best possible convergence and speed, one has to use analytic derivatives for voltage and current; at least the first and second derivatives, if not the third, are required. The programming of such a model becomes so difficult, however, and the calculation of its derivative is so highly likely to have subtle errors, that it turns out that the parallel configuration shown in Figure 2-155 is the best possible compromise for fast computation and accurate results.

A similar approach for bipolar transistors has been discussed by Philips with their MEXTRAM model. However, we have not seen any parameter extraction for this model and also we have not seen any industry acceptance for it. While, in principle, it will cover all derivatives of all members of the BJT family, its topology consists out of three discrete transistors connected and we doubt very much that a reliable parameter extraction program/mathematically continuous model for nonlinear simulation is supplied by the university that developed this jointly with Philips. At the time of writing this book, Philips has not yet provided us with the code for both, and until this happens, we are not in a position to qualify this any further.

2-4-8 Example: Improving the BFR193W Model

We have already mentioned the difficulty encountering the proper base-spreading resistor values and others by using standard parameter extraction programs. Table 2-28 is a copy from the Siemens semiconductor book showing Gummel–Poon Berkeley SPICE syntax parameters for the BFR193W BJT.

The first thing that will happen is that the observant reader will notice that the minimum possible base-spreading resistor, RBM (valid at large current densities), is larger than the actual base-spreading resistor, RB, "claimed to be measured" at 1 Ω (see Table 2-28).

Another headache is caused by the measured noise performance of the same device at an operating point of 10 mA. Using the above-mentioned simplified equation

Table 2.28. BFR193W transistor chip data: SPICE parameters (Gummel–Poon model, Berkeley–SPICE 2G.6 syntax)

IS	=	0.2738	fA	BF	=	125	—	NF	=	0.95341	—
VAF	=	24	V	IKF	=	0.26949	A	ISE	=	10.627	fA
NE	=	1.935	—	BR	=	14.267	—	NR	=	1.4289	—
VAR	=	3.8742	V	IKR	=	0.037925	A	ISC	=	0.037409	fA
NC	=	0.94371	—	RB	=	1	Ω	IRB	=	0.91763	mA
RBM	=	1.8368	Ω	RE	=	0.76534	Ω	RC	=	0.11938	Ω
CJE	=	1.1824	fF	VJE	=	0.70276	V	MJE	=	0.48654	—
TF	=	18.828	ps	XTF	=	0.69477	—	VTF	=	0.8	V
ITF	=	0.96893	mA	PTF	=	0	deg	CJC	=	935.03	fF
VJC	=	1.1828	V	MJC	=	0.30002	—	XCJC	=	0.053563	—
TR	=	1.0037	ns	CJS	=	1	fF	VJS	=	0.75	V
MJS	=	0	—	XTB	=	0	—	EG	=	1.1	eV
XTI	=	3	—	FC	=	0.72063	—	TNOM	=	300	K

All parameters are ready to use, no scaling is necessary. Extracted on behalf of SIEMENS Small Signal Semiconductors by Institut für Mobil-und Satellitenfunktechnik (IMST).

$$F = 1 + \frac{R'_{bb}}{R_g} + \frac{0.5(26\ \text{mV}/I_C)}{R_g}$$

(approximate equation at medium frequencies) (2-288)

we obtain the following results:

1. Inserting 1 Ω (value for RB provided by IMST) and R_B = 15 Ω (measured by the authors), here are the interesting results:

$$F = 1 + \frac{1}{50} + \frac{2.6}{100} = 1.046 \tag{2-289}$$

as noise factor or, as noise figure, F (dB) = 0.2 dB (IMST).

2. Or using the measured 15 Ω, one obtains

$$f = 1 + \frac{15}{50} + \frac{2.6}{100} = 1.326 \tag{2-290}$$

as noise factor or, as noise figure, F (dB) = 1.22.

The Siemens datasheet claims a noise figure of 1.3 dB, and therefore the measured versus calculated data differ by 0.08 dB. Since our noise figure was calculated well below f_T, which is seen from the fact that there are no frequency-dependent terms in the noise figure calculation, it is not unusual that the 900-MHz noise figure is 1.3 dB or slightly higher. This simple example shows how easy it is to provide measurement data that are inconsistent with the laws of physics. It is therefore highly recommended to add the noise figure as an additional parameter in the parameter extraction to reduce the number of variables—specifically, those for which the standard extraction has little sensitivity.

The collector–emitter junction capacitance is another important parameter that frequently gets measured incorrectly. The way around this is to remember the definition

$$f_{\max} = \sqrt{\frac{f_T}{8\pi R'_{bb} C_c}} \tag{2-291}$$

f_T can be measured from the 3-dB point of the emitter current gain (β) and R'_{bb} was just computed above, so the equation can be solved for C_c, as there is a relationship between measured $g_{\max}$ and $f_{\max}$. There are several correlations between the parameters of the equivalent circuit as published in different places, and therefore one can restrict the degree of freedom within the optimization process and obtain much better results. This approach is applicable to all types of transistors, bipolar and FET, and can be used to improve the accuracy of parameters that not only affect a transistor's RF performance but also its noise performance.

REFERENCES

1. H. A. Watson, *Microwave Semiconductor Devices and Their Circuit Applications* . New York: McGraw-Hill, 1969, Chap. 2.
2. A. S. Grove, *op. cit.,* Chap. 4.
3. A. S. Grove, *op. cit.,* Chap. 7.
4. P. E. Gray, D. DeWitt, A. R. Boothroyd, and J. G. Gibbons, *Physical Electronics and Circuit Models of Transistors* . New York: Wiley, 1964, p. 10.
5. P. E. Gray et al., *op. cit.,* p. 129.
6. P. E. Gray et al., *op. cit.,* p. 180.
7. Charles A. Liechti, "Microwave Field-Effect Transistors—1976," *IEEE Trans. Microwave Theory Tech.* , **MTT-24**, 279–300, June 1976.
8. Edward A. Wolff and Roger Kaul, *Microwave Engineering and Systems Applications* . New York: Wiley, 1988, p. 395.
9. Arthur D. Evans, ed., *Designing with Field-Effect Transistors* . New York: McGraw-Hill, 1981.
10. R. D. Middlebrook, "A Simple Derivation of Field-Effect Transistor Characteristics," *Proc. IEEE*, **50**, 1146–1147, 1963.
11. Tzu-Hwa and Craig P. Snapp, "Low Noise Microwave Bipolar Transistor with Sub-Half Micrometer Emitter Width," *IEEE Trans. Electron Devices* , **ED-25**, 723–730, June 1978.
12. Ulrich L. Rohde, "Designing a Match Low Noise Amplifier Using CAD Tools," *Microwave J.* , **20**(10) 154–160, October 1986.
13. R. A. Pucel, W. Strubel, R. Hallgren, and U. L. Rohde, "A General Noise De-embedding Procedure for Packaged Two-Port Linear Active Devices," *IEEE Trans. Microwave Theory Tech.* , **40**(11), 2013–2024, November 1992.
14. R. A. Pucel and U. L. Rohde, "An Exact Expression for Noise Resistance R_n of a Bipolar Transistor for the Hawkins Noise Model," *IEEE Microwave and Guided Wave Letters* , **3**(2), 35–37, February 1993.
15. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits.* New York: Wiley, 1986, p. 386.
16. D. Frohman-Bentchkowsky and A. S. Grove, "Conductance of MOS Transistors in Saturation," *IEEE Trans. Electron Devices* , **ED-16**, 108–113, January 1969.
17. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits* . New York: Wiley, 1986, p. 480.
18. "New LDMOST Technology Superb for RF Applications," *Philips Semiconductors World News,* Vol. 7, Issue 3, Article No. 29, http://www-us2.semiconductors.philips.com/publications/content/file_344.html.
19. Y. P. Tsividis, *Operation and Modeling of the MOS Transistor* . New York: McGraw-Hill, 1987, p. 136.
20. A. Materka and T. Kacprzak, "Computer Calculation of Large-Signal GaAs FET Amplifier Characteristics," *IEEE Trans. Microwave Theory Tech.* , **MTT-33**(2), 129–135, February 1985.
21. R. E. Anholt and S. E. Swirhun, "Experimental Investigation of the Temperature Dependence of GaAs FET Equivalent Circuits," *IEEE Trans. Electron Devices* , **ED-39**(9), 2029–2036, September 1992.
22. Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits* . New York: Wiley, 1986.
23. Paul W. Tuinenga, *SPICE: A Guide to Circuit Simulation and Analysis Using PSpice ®*. Englewood Cliffs, NJ: Prentice Hall, 1992.
24. Andrei Vladimirescu, *The SPICE Book* . New York: Wiley, 1994.

25. H. Statz, H. Haus, and R. A. Pucel, "Noise Characteristics of Gallium Arsenide Field-Effect Transistors," *IEEE Trans. Electron Devices,* **ED-21**, 549–562, September 1974.

26. V. Rizzoli, F. Mastri, and C. Cecchetti, "Computer-Aided Noise Analysis of MESFET and HEMT Mixers," *IEEE Trans. Microwave Theory Tech.* , **MTT-37**, 1401–1410, September 1989.

27. V. Rizzoli, F. Mastri, and D. Masotti, "General-Purpose Noise Analysis of Forced Nonlinear Microwave Circuits," *Military Microwave*, 1992.

28. Ulrich L. Rohde and Chao-Ren Chang, "Parameter Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits," *Microwave J.,* **36**(5), 222–239, May 1993.

29. Ulrich L. Rohde, "Parameters Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits Like Mixers and Oscillators," *23rd European Microwave Conference,* Madrid, Spain, September 6–9, 1993.

30. Ulrich L. Rohde, Chao-Ren Chang, and J. Gerber, "Design and Optimization of Low-Noise Oscillators Using Non-linear CAD Tools," presented at the *Frequency Control Symposium,* Boston, June 1–2, 1994.

31. Ulrich L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design* . New York: Wiley, 1997, Appendix B-3, "Bias Independent Noise Model."

FURTHER READING

George Anderson, "Moto Aims FETs at UHF," *Electron. Eng. Times*, 96, September 2, 1996.

Jean–Jacques Bouny, "Advantages of LDMOS in High Power Linear Amplification," *Microwave Eng. Eur.*, 37–40, April 1996.

Noël Boutin, "Develop SPICE High Frequency Small Signal Transistor Models from Data Sheets," *Appl. Microwave & Wireless*, 30–36, April 1999.

V. Danelon, F. Aniel, J. Benchimol, J. Mba, M. Riet, P. Crozat, G. Vernet, and R. Adde, "Parameters of InP-Based Double-Heterojunction Base Collector Self-Aligned Bipolar Transistors," *IEEE Microwave Guided Wave Lett.*

Suman Datta, Shen Shi, Kenneth P. Roenker, Marc M. Cahay, and William Stanchina, "Simulationand Design of InAlAs/InGaAs *pnp* Heterojunction Bipolar Transistors," *IEEE Trans. Electron Devices*, **ED-45**(8), 1634–1643, August 1998.

Russell J. Elias, Gordon C. Ma, and Laurence S. Golonka, "Statistical Optimization High Frequency LDMOS Devices Via Hyper-Fractionated Experimental Designs," *1996 IEEE/CPMT International Electronics Manufacturing Technology Symposium,* October 14–16, 1996, pp. 462–466.

Al Franceschino, "LDMOS Devices Enhance Cellular Amplifier Design," *Microwaves & RF*, **35**(10), 67–68, 70, 72, 74, October 1996.

GaAs IC Symposium Technical Digest 1995 , IEEE Gallium Arsenide Integrated Circuit Symposium, San Diego, CA, October 29 to November 1, 1995.

Hewlett-Packard, *Communications Components Designer's Catalog* , Hewlett-Packard, Jan Jose, CA, 1997.

ITT Capacitance Diodes, Tuner Diodes, Diode Switch, PIN Diodes, Basics and Applications , International Telephone and Telegraph Corporation, 1976.

Aron Kain, *Final Report for Bias Dependence Noise Modeling of Heterojunction Bipolar Transistors* , USAF SBIR Phase II (PIIN), F33615-95-C-1707, November 1997. Issued by USAF/AFMC/ASC, Wright Laboratory WL/AAKE BLD 7, 2530 C ST, Wright-Patterson AFB, OH 45433-7607.

X. McDade and X. Schiavone, "Switching Time Performance of Microwave PIN Diodes," *Microwave J.*, **17**, September 1974.

Danny Molezion, "A 1 watt Power Amplifier Design for 800–900 MHz Applications Using Discrete LDMOS FETs," *Wireless Technology '96 Conference & Exposition Conference Proceedings,* October 7–11, 1996, pp. 260–267.

Nadir Rahman, "A Survey of Existing Literature on Reverse Modeling of Field Effect Transistors," http://www-mtl.mit.edu/CAPAM/memos/95-3/95-3.html.

Jin-Ho Shin, Jiyoung Kim, Yuji Chugn, Joonwoo Lee, Youngsuk Suh, Kyu Hwan Ahn, and Bumman Kim, "Low-Frequency Noise Characterization of Self-Aligned AlGaAs/GaAs Heterojunction Bipolar Transistors with a Noise Corner Frequency Below 3 kHz," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, pp. 1604–1613, November 1998.

K. Y. Toh, P. K. Ko, and R. G. Meyer, "An Engineering Model for Short-Channel MOS Devices," *IEEE J. Solid-State Circuits*, **23**, 950–958, August 1988.

Robert J. Zavrel, Jr., "Silicon–Germanium and the World of RF Design," *RF Des.*, 46–50, June 1997.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

2

MODELS FOR ACTIVE DEVICES

Because we are dealing with amplifiers, both small-signal and power, oscillators, and active mixers, active devices will play a major role in circuit design. At our frequency range of interest (about 1000 MHz and above), distributed elements predominate, and even in device modeling we must consider some distributed effects such as packaging.

As we dive into providing functional explanations of the active devices used in practical circuits, there is some temptation not to provide enough information on semiconductor behavior. Because manufacturers of microwave products will use the services of a foundry or specify semiconductors from their suppliers, we will cover the device operation in depth. That said, for the beginner looking into designing circuits from high frequencies to low microwave frequencies, we will provide examples of a variety of devices unique to this frequency range. Some treatment of device physics will therefore be useful in providing a basic understanding of the nonlinear behavior of active devices and modeling in preparing for practical circuits using a computer-aided-design (CAD) tool. For CAD applications, we assume the user will have access to nonlinear CAD software, such as a high-performance SPICE program or a harmonic-balance program, such as Ansoft Serenade 8.0 or its derivatives. It is interesting that most semiconductor houses today still only offer very limited nonlinear data sets, while in the linear area a large body of useful data is made available, mostly on CDs. What CAD does for us—and this applies to SPICE and harmonic-balance programs alike—is predict the dc bias point of a device in the context of its surrounding circuitry. Configuring the simulation to drive the device into large-signal operation provides us with insight into the various forms of distortion and compression.

In evaluating active devices for RF circuit design, we have a few different technologies to consider. First, we will examine the semiconductor junction diode—a basic *pn* junction—and venture from there to the bipolar transistor. An investigation of metal-oxide semiconductor (MOS) devices, another important set of semiconductors for medium frequencies, will follow. We will also examine the gallium arsenide metal-gate Schottky field-effect transistor (GaAs MESFET) and members of its family. Since critical large-signal parameters for some devices may not be available from their manufacturers, we will examine

parameter-extraction techniques and problems. For various device types, we will provide insight into the noise properties of the device.

2-1 DIODES*

The diode model contains a nonlinear current source that follows the Shockley equation:

$$\text{Current} = \text{IS}\left(e^{V_j/\text{N}V_t} - 1\right) \tag{2-1}$$

where V_j = voltage across the junction
V_t = thermal voltage (= kT/q)

These values, with the model parameters IS and N, are used to model the current–voltage effects of the semiconductor junction. This does not include the nonideal operation of real diodes. For example, at low currents (less than 1 nA), other semiconductor processes that increase the flow of currents become noticeable.

By setting IS to different values, we can obtain the characteristics of other devices, such as a Schottky barrier diode, or a silicon diffused-junction diode. High-current effects are modeled, grossly, by including a series resistance that is intended to combine the effects of bulk resistance (the material on each side of the junction) and high-level injection. At high currents, the observed diode current stops following the Shockley form

$$I_{\text{forward}} = \text{IS} \times e^{V_j/\text{N}V_t} \tag{2-2}$$

and approaches a modified form

$$I_{\text{forward}} = \text{IS} \times e^{V_j/2\text{N}V_t} \tag{2-3}$$

2-1-1 Large-Signal Diode Model

Three diode models are used in the industry:

- Microwave diode model
- PIN diode model
- Enhanced SPICE diode model

Figure 2-1 shows the large-signal microwave diode model. Its keywords appear in Table 2-1. This model can also be used to simulate varactor diodes.

Table 2-2 lists SPICE parameters for a selection of Schottky mixer diodes by Alpha.

In most cases, the diode capacitance is modeled by a voltage-dependent capacitor, which is connected in parallel with the nonlinear current generator described previously, to represent the charge-storage effects of the junction. There are two components to this charge

- The reverse-voltage capacitive effect of the depletion region
- The forward-voltage charge represented by mobile carriers in the diode junction

*Portions of this chapter's diode coverage are based on material in Alpha Industries Semiconductor Application Reports 80800, 80200, and 80500. Used with permission.

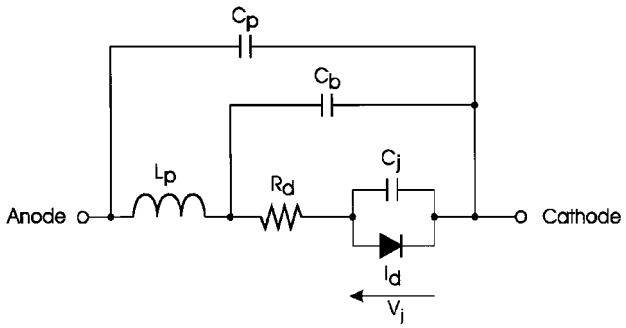


Figure 2-1 The large-signal microwave diode model. This model is temperature dependent.

Reverse-voltage capacitance follows the simple approximation that the depletion region (the area of the junction that is depleted of carriers) serves as the gap between the "plates" of a capacitor. This region varies in thickness, and therefore the capacitance varies with applied voltage. For a step (abrupt) junction, or linearly graded junction, the capacitance approximation is

Table 2-1 Large-signal microwave diode model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
JS	Saturation current	ampere	0
ALFA	Slope factor of conduction current	/volt	38.696
JB	Breakdown saturation current	ampere	10 mA
VB	Breakdown voltage	volt	$-\infty$
E	Power-law parameter of breakdown current	—	10.0
CT0	Zero-bias depletion capacitance	farad	0
FI	Built-in barrier potential	volt	0.8
GAMA	Capacitance power-law parameter	0.5	
GC1	Varactor capacitance polynomial coefficient 1	/volt	0.0
GC2	Varactor capacitance polynomial coefficient 2	/volt^2	0.0
GC3	Varactor capacitance polynomial coefficient 3	/volt^3	0.0
CD0	Zero-bias diffusion capacitance (*pn* diodes)	farad	0
AFAC	Slope factor of diffusion capacitance	/volt	38.696
R0	Bias-dependent part of series resistance in forward-bias condition	ohm	0
T	Intrinsic time constant of depletion layer for abrupt-junction diodes	second	0
KF	Flicker noise coefficient	—	0.0
AF	Flicker noise exponent	—	1.0
FCP	Flicker noise frequency shape factor	—	1.0
AREA	Area multiplier	—	1.0
	Extrinsic Model		
CP	Package parasitic capacitance	farad	0.0
CB	Beam-lead parasitic capacitance	farad	0.0
LP	Package parasitic inductance	henry	0.0

Table 2-2 Diode SPICE parameters

Parameter	Unit	SMS1546	SMS3922	SMS3923	SMS3924	SMS3926	SMS3927	SMS3928	SMS7621	SMS7630
I_S	A	3E–7	3E–8	5E–9	2E–11	2.5E–07	1.3E–09	9E–13	4E–8	5E–06
R_S	Ω	4	9	11	11	4	4	4	12	30
n	—	1.04	1.08	1.05	1.08	1.04	1.04	1.04	1.05	1.05
T_d	S	1E–11	8E–11	8E–11	8E–11	1E–11	1E–11	1E–11	1E–11	1E–11
C_{J0}	PF	0.38	0.9	0.93	1.6	0.42	0.39	0.39	0.1	0.14
m	—	0.36	0.26	0.24	0.4	0.32	0.37	0.42	0.35	0.4
E_G	EV	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
X_{ti}	—	2	2	2	2	2	2	2	2	2
F_C	—	0.5	0.5	0.5	0.5	0.5	0.5	0.5	5	0.5
B_J	V	3	20	46	100	2	3	4	3	2
I_{BV}	A	1E–5	1E–5	1E–5	1E–5	1.00E–05	1.00E–05	1.00E–05	1E–5	0.0001
V_J	—	0.51	0.65	0.15	0.84	0.495	0.595	0.800	0.51	0.34

$$\text{Capacitance} = \frac{C_{J0}}{\left(1 - V_j/\phi\right)^M} \tag{2-4}$$

where C_{J0} is the zero-bias value, ϕ (phi) is the junction barrier potential, and M is the grading coefficient that varies ($\frac{1}{2}$ is used for step junctions and $\frac{1}{3}$ is used for linearly graded junctions, and most junctions are somewhere in between).

There is often confusion about the barrier potential, which appears in the capacitance equation. From capacitance measurements, ϕ (model parameter VJ, and not to be confused with V_j in the equations) takes on a value of nearly 0.7 volt for regular (silicon) junction diodes and a range of 0.58–0.85 volt for various Schottky barrier diodes. This value is sometimes confused with the forward-current voltage drop of the diode or the energy gap of the material; it is *neither* of these.

Varying M generates a variety of reverse-bias capacitance characteristics. Inspection of the capacitance formula reveals that it predicts infinite capacitance for a forward bias, which is not the case for a real junction. Several depletion-capacitance formulas have been proposed that more correctly fit observed operation; however, SPICE uses a simple approach: For forward biases beyond some fraction (set by the parameter FC) of the value for ϕ, the capacitance is calculated as the linear extrapolation of the capacitance at the departure. This provides a continuous numerical result and does not affect circuit operation significantly because, for forward bias, the device capacitance is normally dominated by diffusion capacitance.

The diffusion charge (and therefore the capacitance) varies with forward current and is simply modeled as a transit time (model parameter TT) for the carriers to cross the diffusion region of the junction. The total charge is

$$\text{Diffusion charge} = \text{Device current} \times \text{Transit time} \tag{2-5}$$

and capacitance is the derivative, with respect to bias, of this:

$$\text{Diffusion capacitance} = \text{TT}\,\frac{\text{IS}}{\text{N}V_t}\,e^{V_j/\text{N}V_t} \tag{2-6}$$

Diffusion charge manifests itself as the *storage time* of a switching diode, which is the time required to discharge the diffusion charge in the junction, which must happen before the junction can be reverse-biased (switched off). Storage time is normally specified as the time to discharge the junction so that it is supporting only a fraction (typically 10%) of the initial reverse current. First, a forward current is supplied to the device to charge the junction. Then, as quickly as possible, a reverse current is supplied to the device. Internally, the junction is still forward-biased to a voltage nearly the same as before the switch in current; the junction is still conducting at the forward-current rate. This internal current adds to the external current as the total current discharging the junction. As the junction voltage decreases, the internal current falls off exponentially (according to the Shockley equation). This system is a relatively simple differential equation that can be solved to an explicit equation for the TT parameter (assuming complete discharge) as follows:

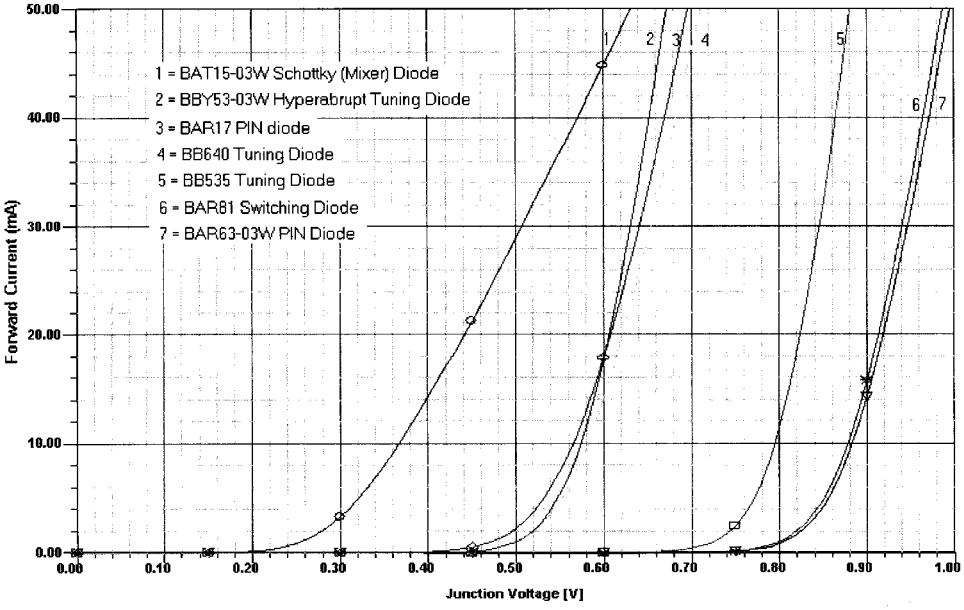


Figure 2-2 The dc *I–V* curves for seven diodes, showing the various barrier voltages that result from different doping profiles.

$$\text{Transit time} = \frac{\text{Storage time}}{\ln\left(\dfrac{I_F - I_R}{-I_R}\right)} \tag{2-7}$$

The diffusion charge dominates the reverse recovery characteristic of the diode. During the last part of the recovery, as the junction becomes reverse-biased, the depletion capacitance dominates. This causes the small tail at the end of the discharge cycle. Total capacitance is taken to be the sum of these capacitances: the depletion approximation dominates for reverse bias as the device current is small, and the diffusion proximity dominates for forward bias as the device current is large.

A special case for diode application is the switching diode, and its description and application will be part of a later chapter. Figure 2-2 shows the dc *I–V* curves, which indicate the different voltage potentials that are a result of the different doping profiles.

2-1-2 Mixer and Detector Diodes

Electrical Characteristics and Physics of Schottky Barriers. Schottky barrier diodes differ from junction diodes in that current flow involves only one type of carrier instead of both types. That is, in *n*-type Schottky barrier diodes the forward current consists of holes flowing from the *n*-type material into the metal.

Diode action results from a contact potential set up between the metal and the semiconductor, similar to the voltage between the two metals in a thermocouple. When metal is brought into contact with an *n*-type semiconductor (during fabrication of the chip), electronics diffuse out of the semiconductor into the metal, leaving a region under the contact that

has no free electrons ("depletion layer"). This region contains donor atoms that are positively charged (because each lost its excess electron), and this charge makes the semiconductor positive with respect to the metal. Diffusion continues until the semiconductor is so positive with respect to the metal that no more electrons can go into the metal. The internal voltage difference between the metal and the semiconductor is called the contact potential and is usually in the range of 0.3–0.8 volt for typical Schottky diodes. A cross section is shown in Figure 2-3.

When a positive voltage is applied to the metal, the internal voltage is reduced, and electrons can flow into the metal. The process is similar to thermionic emission of electrons from the hot cathode of a vacuum tube, except that the electrons are "escaping" into a metal instead of into a vacuum. Unlike the vacuum tube case, room temperature is "hot" enough for this to happen if enough voltage is applied. However, only those electronics whose thermal energy happens to be many times the average can escape, and these "hot electrons" account for all the forward current from the semiconductor into the metal.

One important thing to note is that there is no flow of minority carriers from the metal into the semiconductor and thus no neutral plasma of holes and electrons is formed. Therefore, if the forward voltage is removed, current stops "instantly," and reverse voltage can be established in a few picoseconds. There is no delay effect to charge storage as in junction diodes. This accounts for the exclusive use of barrier diodes in microwave mixers, where the diode must switch conductance states at microwave oscillator rates.

The current–voltage relationship for a barrier diode is described by the Richardson equation (which also applies to thermionic emission from a cathode). The derivation is given in many textbooks (e.g., Sze).

$$I = AA_{\mathrm{RC}}T^2 \exp\left(-\frac{q\phi_B}{kT}\right)\left[\exp\left(\frac{qV_J}{kT}\right) - M\right] \tag{2-8}$$

where A = area (cm^2)

A_{RC} = modified Richardson constant ($\mathrm{A}/K^2 \cdot \mathrm{cm}^2$)

k = Boltzmann's constant

T = absolute temperature (K)

ϕ_{B} = barrier height in volts

V_J = external voltage across the depletion layer (positive for external voltage) = $V - IR_S$

R_S = series resistance

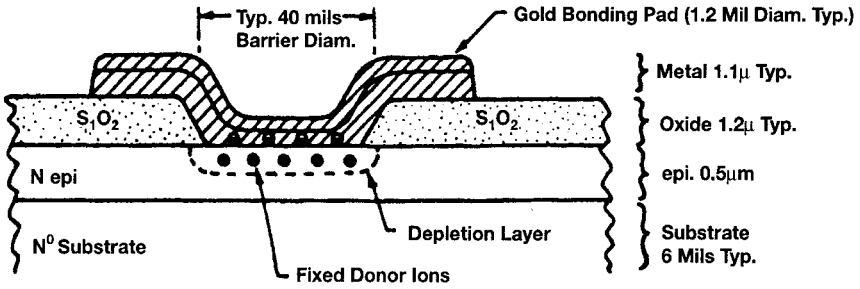


Figure 2-3 Schottky diode chip cross section.

M = avalanche multiplication factor
I = diode current in amperes (positive forward current)

The barrier height, ϕ_B, is typically a few tenths of a volt higher than the contact potential, ϕ_C (about 0.15 V higher than ϕ_C for silicon). This equation agrees well with experimental data for diodes without surface leakage, but it is difficult to use because A_{RC}, ϕ_B, and M are all dependent on applied voltage.

The major cause for variation in ϕ_B with voltage is the "image effect," in which the barrier height is lowered as the electric field near the metal is increased, especially at the edges. A better equation for circuit designers to use is one in which all parameters are independent of voltage and current. The simplest one that agrees reasonably well with Richardson's equation is

$$I = I_S\left[\exp\left(\frac{V_J}{0.028}\right) - 1 + \frac{K}{\left(1 - V_B/V\right)}\right] \tag{2-9}$$

where I_S = saturation current (a temperature-dependent quantity)
0.028 = nkT/q at room temperature (n = 1.08)
n = forward slope factor (derived from the variation of ϕ_B with forward voltage)
K = reverse slope factor (expressing the variation of ϕ_B with reverse voltage)
V_B = breakdown voltage (the voltage at which M = 1)

As before, V and I are considered positive for forward bias and negative for reverse bias.

Typical ranges for these parameters for microwave Schottky and point-contact mixer diodes are:

I_S = 10^{-12} to 10^{-5} A
n = 1.04–1.10
R_S = 2–20 Ω
K = 8–100
V_B = 2–20 V

The quantities I_S and 0.028 are strongly temperature dependent, while both R_S and V_B increase with temperature to a slight degree. R_S increases with current at high current levels (due to carrier velocity saturation) but is essentially independent of current at 10 mA and below for mixer diodes. Thus, for normal mixer and detector operation, R_S can be considered constant.

Agreement between Eqs. (2-8) and (2-9) is not perfect but Eq. (2-9) is much easier to use and is preferred by most circuit designers. A comparison of the two equations near zero bias gives the following relationship between zero-bias barrier height, ϕ_0, and saturation current:

$$I_S = AA_{RC}T^2\exp - \frac{q\phi_0}{kT}$$

$$\cong \left(\frac{10^7 A}{\text{cm}^2}\right)A\exp\left(-\frac{\phi_0}{0.026}\right) \quad \text{(for } n \text{ silicon at room temperature)} \tag{2-10}$$

Small-Signal Parameters. By combining Eqs. (2-9) and (2-10), the values of the parameters in Eq. (2-9) can be derived from a few simple measurements. Many specific equations can be derived, but the following are commonly used for production measurements:

$$R_S = \frac{V_{F10} - V_{F1} - 0.065}{0.009} \quad (\text{for } n = 1.08) \tag{2-11}$$

$$\phi_0 = \frac{V_{F1} - 0.001R_S + 0.28 + 0.12\log_{10} D}{1.08} \tag{2-12}$$

$$n = \frac{V_{F1} - V_{F0.1} - 0.0009R_S}{0.060} \tag{2-13}$$

$$K = \left(\frac{I_{R1}}{I_S} - 1\right)V_B \tag{2-14}$$

$$I_S = \exp\left(-\frac{V_{F1}}{0.028} + \frac{R_S}{28}\right) \quad (\text{in mA}) \tag{2-15}$$

where $V_{F0.1}$, V_{F1}, and V_{F10} are the forward voltages at 0.1, 1, and 10 mA, respectively, and I_{R1} is the reverse current at 1 volt. (The derivation of these equations requires that I_S be small compared to 0.1 mA.) The quantity D is the diameter of the metal–silicon contact in mils.

Measuring V_{F1} at 1 mA and 10 mA instead of some other current levels leads to the best accuracy for typical mixer diodes.

The total dynamic resistance for a forward-biased diode is given by

$$R_T = \frac{dV}{dI} = R_S + \frac{nkT}{q(I + I_S)} = R_S + R_B \tag{2-16}$$

and

$$R_B = \frac{28}{I + I_S} \quad \text{at room temperature (with } I \text{ and } I_S \text{ in mA, } n = 1.08) \tag{2-17}$$

This equation is also good at zero bias (unless K is very large or there is significant surface leakage). That is,

$$R_0 = R_S + \frac{28}{I_S} \tag{2-18}$$

For reverse voltages of a few volts, the dynamic resistance is dominated by the K term:

$$R_R = \text{Reverse resistance} = \frac{dV}{dI} \cong \frac{V_B}{KI_S} \tag{2-19}$$

For typical values of I_S, R_0 is larger than 5000 Ω and R_R is larger than 100 kΩ. For some zero-bias Schottky applications, it is desirable for R_0 to be made smaller than this.

The factors that determine R_S are (1) the thickness of the epitaxial layer, (2) the epitaxial layer doping level (N_D), (3) the barrier diameter, (4) the substrate resistivity ("spreading resistance"), (5) the contact resistances of the metals used for the barrier and the substrate contact, and (6) the resistance associated with the bonding wire or whisker. The barrier height is about 0.15 volt higher than the contact potential between the barrier metal and the semiconductor and is influenced by the method used to apply the metal, conditions at the edge of the junction, and the doping level. Saturation current depends on barrier height, junction area, and temperature; and the slope factors, n and K, depend on doping level, punchthrough voltage, and edge conditions.

Junction Capacitance. The capacitance of a Schottky barrier chip results mainly from two sources: the depletion layer under the metal–semiconductor contact and the capacitance of the oxide layer under the bonding pad (the so-called overlay capacitance). The bonding pad is required because the typical Schottky barrier diameter is so small that it is impractical to bond directly to the metal on the junction. If the semiconductor epitaxial layer is uniformly doped, the capacitance–voltage characteristic is similar to that of a textbook "abrupt-junction" diode.

$$C_J = \frac{\varepsilon_S \varepsilon_0 A'}{X_D} + C_0 \tag{2-20}$$

$$X_D = \sqrt{\frac{2\varepsilon_S \varepsilon_0 (\phi_C - V)}{qN}} \tag{2-21}$$

where ϕ = contact potential
C_0 = overlay (bonding pad) capacitance
ε_S = dielectric constant of the semiconductor ($\cong$ 12 for Si or GaAs)
N = doping level for the epitaxial layer
A' = effective contact area, including fringing corrections

In practical terms, the capacitance can be related to the 0-V barrier capacitance defined by

$$C_{B0} = \frac{\varepsilon_S \varepsilon_0 A'}{X_{D0}} \tag{2-22}$$

where

$$X_{D0} = \sqrt{\left(\frac{1.3 \times 10^{15}}{N_D}\right)\phi_C} \quad \text{(in } \mu\text{m)} \tag{2-23}$$

The resulting C–V relationship can be written

$$C_J = \frac{C_{B0}}{\sqrt{1 - V/\phi_C}} + C_0 \tag{2-24}$$

The contact potential, ϕ_C, is related to the barrier height as follows:

$$\phi_C = \phi_B - 0.026\left[1 + L_n\left(\frac{N_C}{N}\right)\right] \tag{2-25}$$

The theoretical meaning of these terms can be clarified by looking at Figure 2-4.

Parameter Trade-Offs

Barrier Height. The barrier height of a Schottky diode is important because it directly determines the forward voltage. In order to get a good noise figure, the LO drive voltage, V_L, must be large compared to V_T, which is essentially V_{F1}. Normally, it is best to have a low

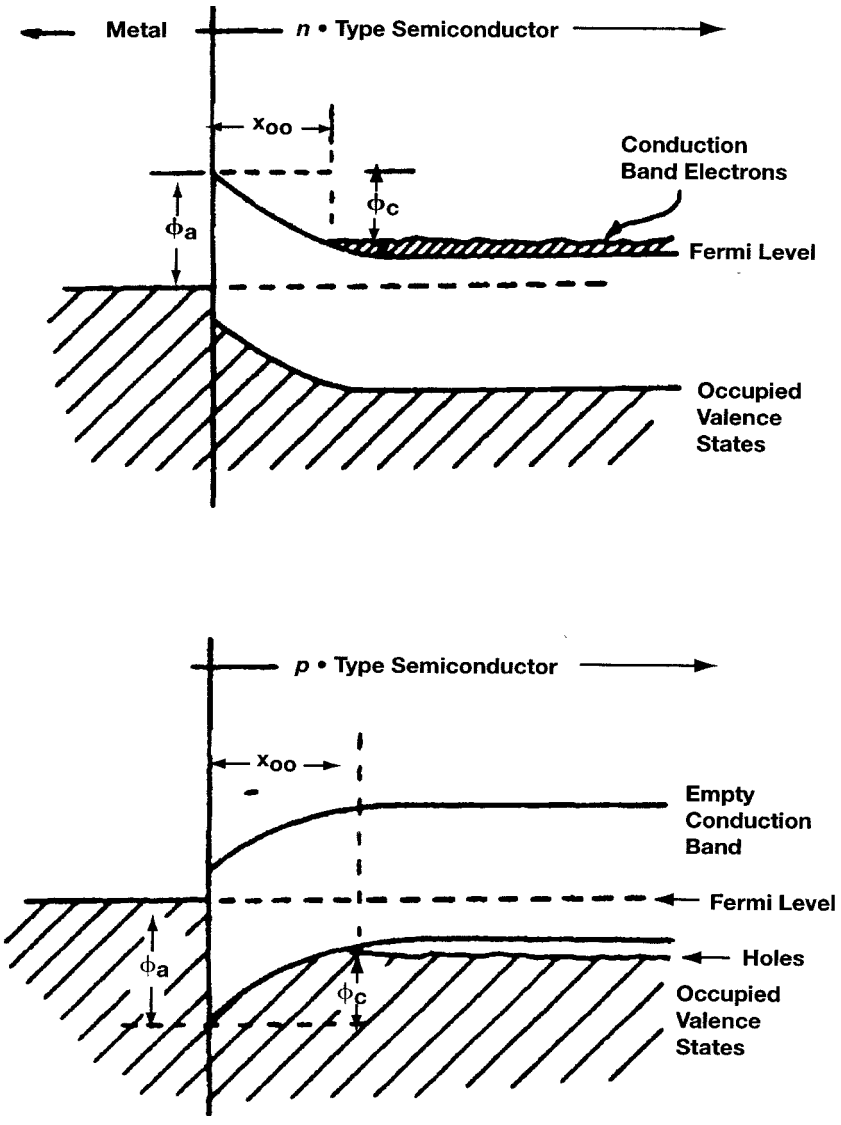


Figure 2-4 Schottky diode band diagrams.

Table 2-3 LO power and V_{F1} for various applications

Type	Typical V_{F1}	LO Power	Application
"Zero bias"	0.10–0.25	< 0.1 mW	Mainly for detectors
Low barrier	0.25–0.35	0.2–2 mW	Low-drive mixers
Medium barrier	0.35–0.50	0.5–10 mW	General purpose
High barrier	0.50–0.80	> 10 mW	High dynamic range

forward voltage (low V_{F1}), or low drive) diode, to reduce the amount of LO power needed. However, if high dynamic range is important, high LO power is needed, and the diode can have a higher V_F and should also have a high V_B (see Table 2-3).

Noise Figure Versus LO Power. At low LO drive levels, the noise figure is poor because of poor conversion loss, due to a too-low conduction angle. At very high LO drive levels, the noise figure again increases due to diode heating, excess noise, and reverse conduction.

If a high LO drive level is needed, for example, to get higher dynamic range, a high V_B (> 5 V) should be specified. However, Nature requires that you pay for this with higher R_S (lower f_C), so the noise figure will be degraded compared to what could be obtained with diodes designed for lower LO drive. Forward voltage and breakdown are basically independent parameters, but high breakdown is not needed or desirable unless high LO power is used.

Such a high-breakdown diode will have low reverse current (which is important only if the diode has to run hot).

Silicon Versus GaAs. Typical silicon Schottky diodes have cutoff frequencies in the 800–200-GHz range, which is good enough through the Ku band.

At the Ku band and above, or for image-enhanced mixers, higher f_C may be needed, which calls for the use of GaAs diodes. These have lower R_S due to higher mobility, which translates to cutoff frequencies in the 400–1000-GHz range.

However, if your IF is low, be careful; GaAs diodes have high 1/*f* noise. They also have high V_{F1}, so more LO power is required.

C_J Versus Frequency. There is quite a lot of latitude in choosing C_J. However, in general, the capacitive reactance should be a little lower than the transformed line impedance (Z_0). If Z_0 is not known, a good way to start is to use $X_C = 100\ \Omega$. Experience has shown that most practical mixers use an X_C near this value (a little higher in waveguide, and lower in 50-Ω systems). This translates to the following rule of thumb for choosing the junction capacitance of a diode for operation at frequency *f* (in GHz):

$$C_{J0} \approx \frac{100}{\omega}$$

$$\approx \frac{1.6}{f} \quad \text{(in pF)} \tag{2-26}$$

In order to evaluate possible tolerances, we show the range of forward current as a function of diode voltage (Figure 2-5), the junction capacitance as a function of the bias voltage

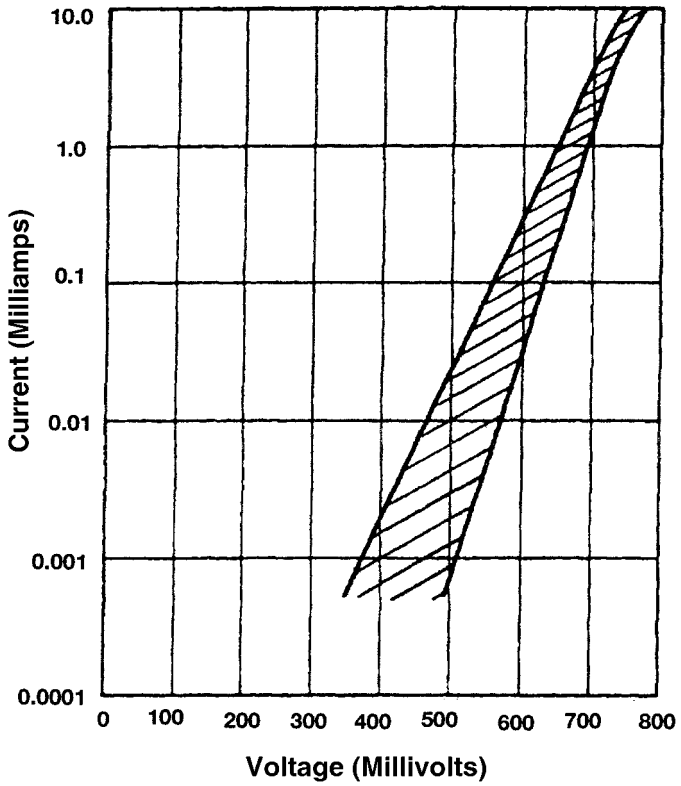


Figure 2-5 Forward dc characteristic curve range—current versus voltage.

(Figure 2-6), and finally some important RF parameters, such as noise figure and IF impedance, as a function of the local-oscillator drive (Figure 2-7).

Mixer Diodes. As an example of some of the parameters for mixer diodes, Table 2-4 gives data on some of the X-band mixer diodes. NF is measured at 9.375 GHz.

Linear Diode Model. Figure 2-8 shows the linear diode model. Its keywords appear in Table 2-5.

The circuit simulators, such as those supplied by Ansoft and Hewlett-Packard, provide a model library that has SPICE-type parameters for diodes (regular diodes, varactor diodes, and PIN diodes) as well as bipolar transistors and FETs, which will be discussed later.

2-1-3 PIN Diodes

Introduction. The PIN diode, in comparison with other microwave semiconductor devices, is fairly easy to understand. This makes possible the reduction of its complex behavior to relatively simple terms.

We do not attempt to describe the many possible microwave circuits in which PIN diodes are used. Rather, we attempt to explain the behavior of the diode in all aspects, giving the facts and some of the theory behind the facts. We offer the circuit designer the opportunity

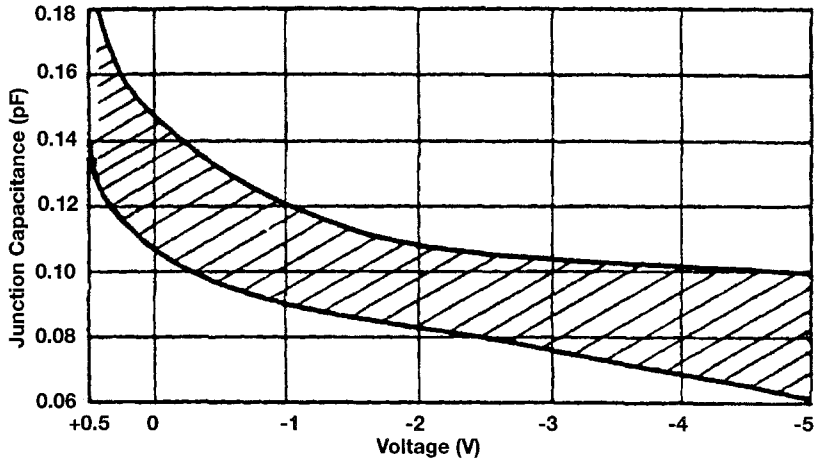


Figure 2-6 Junction capacitance range versus voltage.

to understand the PIN, so that he/she can understand its behavior in circuits. We assume the reader knows the circuit equations; to that knowledge we hope to add diode equations.

Most of the material presented consists of generalized data and explanations of the behavior of PIN diodes; we conclude with a brief description of circuit performance, test methods, and some hints on proper PIN specification writing.

The user can then evaluate the trade-offs involved in diode design and performance and be able to select the most nearly optimum diode from the wide range of diodes offered.

Large-Signal PIN Diode Model. Figure 2-9 shows the large-signal model for a PIN diode. Table 2-6 lists its keywords.

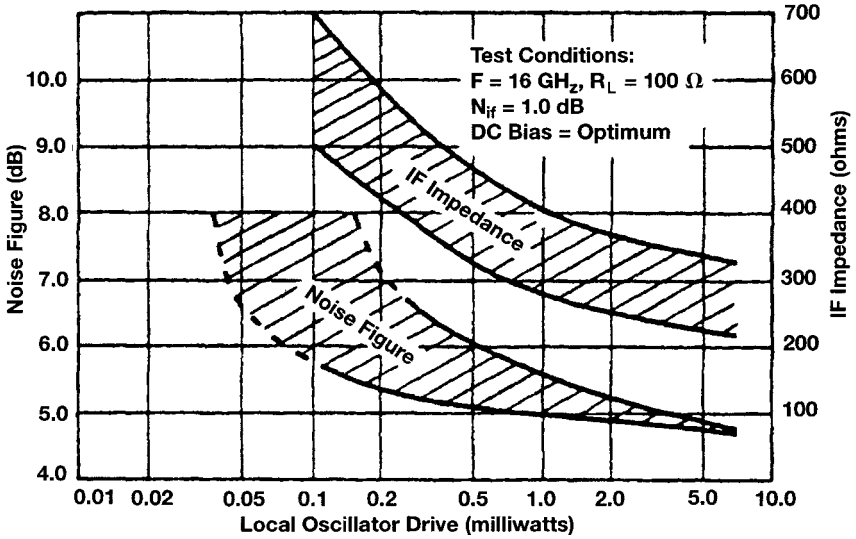


Figure 2-7 RF parameters versus local-oscillator drive level.

Table 2-4 X-band mixer diode data

Material	Barrier	Typical V_F (@1 mA)	Typical F_{C0} (GHz)	Typical R_S (Ω)	Typical C_{J0} (pF)	Maximum NF (dB)
n GaAs	High	0.70	1000	—	0.15	5.0[a]
n GaAs (BL)	High	0.70	500	—	0.15	6.0[a]
n GaAs (chip)	High	0.70	1000	—	0.15	5.3[a]
n Silicon (BL)	Low	0.28	150	6	0.20	6.5
n Silicon (quad)	Low	0.28	150	6	0.20	6.5
p Silicon (BL)	Low	0.28	150	12	0.20	6.5
n Silicon (BL)	High	0.60	100	8	0.20	6.5
n Silicon (quad)	High	0.60	100	8	0.20	6.5
n Silicon	Low	0.28	200	6	0.15	5.5
p Silicon	Low	0.28	200	18	0.14	6.0
p Silicon	Medium	0.40	150	150	0.12	6.5
n Silicon	Low	0.28	150	8	0.18	6.5
p Silicon	Low	0.28	150	12	0.18	6.5

[a]Specified for N_F = 1.0 dB.

Basic Theory: Variable Resistance. Intrinsic or "pure" silicon, as it can be grown in a laboratory, is an almost lossless dielectric. Some of is physical properties include:

Dielectric constant (relative)	12
Dielectric strength	400 V/mil (approximate)
Specific density	2.3
Specific heat	0.72 J/g · °C
Thermal conductivity	1.5 W/cm · °C
Resistivity	300,000 Ω/cm

Since a PIN diode is valuable essentially because it is a variable resistor, let us concentrate initially on the resistivity. Consider a volume comparable to a typical PIN diode chip, say, 20 mils in diameter and 2 mils thick. This chip has a dc resistance of about 0.75 MΩ. High

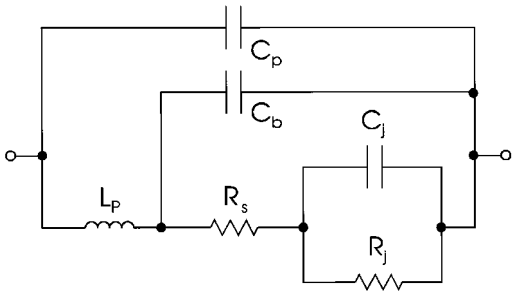


Figure 2-8 The linear diode model. This model is temperature dependent.

Table 2-5 Linear diode model keywords

Keyword	Description	Unit	Default
LP	Package inductance	henry	0.0
CB	Beam-lead capacitance	farad	0.0
CP	Package capacitance	farad	0.0
RS	Contact resistance	ohm	0.0
RJ	Junction resistance	ohm	
CJ	Junction capacitance	farad	

resistivity in any material indicates that most of the likely carriers of electric charge, electrons and holes, are tightly held in the crystal lattice and cannot "conduct."

In real life there are impurities, typically boron, that cannot be segregated out of the crystal. Such impurities contribute carriers, holes or electrons, that are not very tightly bound to the lattice and therefore lower the resistivity of the silicon.

Through various techniques we can adjust the level of impurities, called *dopants*, to produce resistivities ranging from 10 kΩ/cm (for good PIN diodes) to 0.001 Ω/cm (for substrates).

If the impurity adds "electrons" to the crystal, it is called a *donor*; if it adds holes it is called an *acceptor*. Boron adds holes, hence it is an acceptor, and the silicon plus boron combination is called *p*-type, or *positive*, because it has an excess of positive carriers. Phosphorus, on the other hand, is a *donor*, adding electrons, and the corresponding mix is *n*-type, or *negative*.

Notes on the PIN Diode Model

1. The PIN diode model is used to model a bias-dependent RF resistance for use in PIN diode circuits such as attenuators and switches. The resistance varies from RMAX to RS using the R function above. A typical R versus *I* characteristic is shown in Figure 2-10 with parameters IS = 5.96 nA, RS = 2.016, RMAX = 6500, K1 = 0.1272, K2 = 1.0, N = 2.077.
2. The transit-time parameter, TT, can also be used to approximately model a switching PIN diode's reverse-recovery time—a value often provided by diode manufacturers.
3. Diode breakdown can be modeled by specifying IBV and BV parameters.

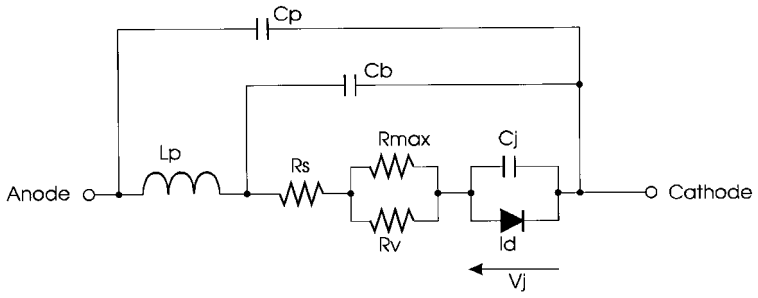


Figure 2-9 The large-signal PIN diode model. This model is temperature dependent.

Table 2-6 PIN diode model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
IS	Saturation current	ampere	1.0E–14
N	Emission coefficient		1.0
IBV	Magnitude of current at the reverse breakdown voltage	ampere	1.0E–10
BV	Magnitude of the reverse breakdown voltage	volt	∞
FC	Coefficient for forward-bias depletion capacitance		0.5
CJ0	Zero-bias *pn* junction capacitance	farad	0.0
VJ	Built-in junction potential	volt	1.0
M	*pn* Junction grading coefficient		0.5
GC1	Varactor capacitance polynomial coefficient 1	/volt	0.0
GC2	Varactor capacitance polynomial coefficient 2	/volt^2	0.0
GC3	Varactor capacitance polynomial coefficient 3	/volt^3	0.0
TT	Transit time	second	0.0
K1	Variable resistance coefficient	volt	0.0
K2	Variable resistance current exponent		1.0
RMAX	Maximum resistance of PIN intrinsic region	ohm	0.0
KF	Flicker noise coefficient		0.0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
AREA	Area multiplier		1.0
	Extrinsic Model		
RS	Series resistance (minimum resistance of PIN diode)	ohm	0.0
CP	Package parasitic capacitance	farad	0.0
CB	Beam-lead parasitic capacitance	farad	0.0
LP	Package parasitic inductance	henry	0.0

4. The reverse-bias capacitance characteristics can be more accurately modeled than the common expression derived from *pn* junction theory. The capacitance grading coefficient exponent can be expressed as a polynomial function of voltage by specifying values for GC1, GC2, and GC3.
5. The PIN diode model was derived from J. Walston, "SPICE Circuit Yields Recipe for PIN Diode," *Microwaves & RF*, November 1992.
6. Following Simon M. Sze, *Physics of Semiconductor Devices* (New York: John Wiley, 1982), the variable resistance may be modeled by setting

$$\text{K1} = \frac{3}{8} V_T \frac{W^2}{D_a \tau_a} \qquad \text{K2} = 1.0$$

where W is the width of the intrinsic region, D_a is the ambipolar diffusion coefficient, τ_a is the ambipolar lifetime, and V_T is the thermal voltage.

There are many concepts important to the physicist, but not to the diode user, that elaborate on the impact of impurities on the behavior of silicon. These can be studied in Watson [1]. The more carriers added, the lower the resistivity.

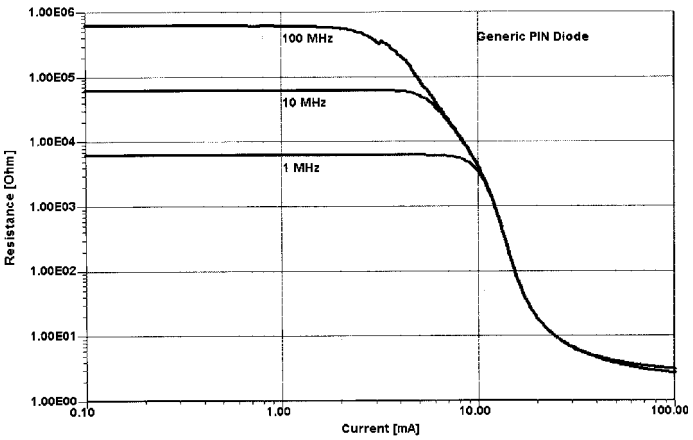


Figure 2-10 Simulated PIN diode resistance as a function of dc at 1, 10, and 100 MHz.

If one wished to vary the resistance of a given diode, in principle he/she could bring it into a semiconductor laboratory, add or subtract carriers as desired, and perhaps even make the process reversible. However, this is a slow, expensive, and impractical way to make a variable resistor; one would be better advised to take a wrench and a soldering iron and replace a component.

The PIN diode derives its value from the fact that the free-charge-carrier concentration in silicon, and hence its resistance, can be varied electronically by means of current from a simple bias supply. This can be done rapidly (in nanoseconds in some cases), reversibly, repeatably, and accurately. The thing that makes this possible is called a *junction*, the interface between the relatively pure silicon in the middle of the PIN diode (the I stands for *intrinsic*) and the heavily doped layers on either end, p^+ and n^+. The p^+ region is rich in holes; the n^+ region is rich in electrons. Both of these regions have low resistance. The I region is the variable resistive element in the diode (see Figure 2-11). In the absence of any external bias, internal effects within the crystal keep the charges fixed; the resistance of the I region is high.

When the p^+ region (anode) is biased positively with respect to the n^+ region (cathode), the interface potential "barrier" is overcome, and direct current flows in the form of holes

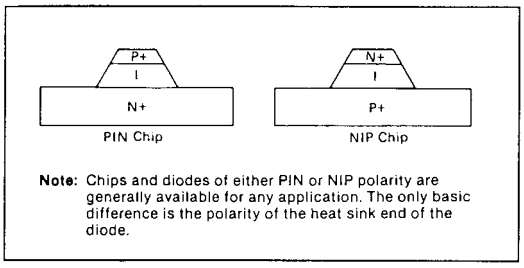


Figure 2-11 General outline of PIN diode construction.

streaming from p^+ toward n^+, with electrons moving in the opposite direction; we say that free carriers have been injected into the I region. The resistance of the I region becomes low.

The number of free carriers within the I region determines the resistivity of the region and thus the resistance of the diode.

Consider "one hole" and "one electron" drifting in opposite directions in the I region under the impetus of the applied field. Under certain conditions, imperfections in the silicon may cause these carriers to recombine. They are no longer available to constitute current or to lower the resistivity of the I region.

It can be shown that the amount of "recombination" between holes and electrons that continuously takes place in a semiconductor is governed by a property of the lattice called *lifetime*. In fact, lifetime is defined as the reciprocal of recombination rate.

Thus, $Q_S = Q_0 \exp(-t/T_L)$, where Q_S is the total amount of free charge "stored" in the I region, and T_L is the lifetime, or the mean time between recombination events. In the steady-state condition, the bias supply must deliver current to maintain a constant Q_S. The required current is

$$I_{\text{dc}} = \frac{Q_S d}{dt} = \frac{-Q_S}{T_L} \tag{2-27}$$

or $Q_S = I_{\text{dc}} T_L$, dropping the minus sign.

Ignoring some details that are not crucial to this section, we can now calculate the resistance of a given diode, of area A and thickness W. (W stands for *base width*, the width or thickness of the intrinsic layer.) The p^+ and n^+ regions have essentially zero resistance, as they are very heavily doped.

The resistivity of a given material is inversely proportional to the number of free carriers, N, and the mobility (not quite the same as velocity) of the carriers. Thus,

$$\rho = \frac{1}{q(\mu_n N + \mu_p P)} \tag{2-28}$$

where μ_n and μ_p are mobilities of electrons and holes, and N and P are numbers of electrons and holes. Simplifying, we find

$$\rho = \frac{C}{Q_S d} \tag{2-29}$$

where C is a collection of constants, and $Q_S d$ is the stored charge density (numbers per unit volume). For our piece of silicon, volume is WA, and

$$Q_S d = \frac{Q_S}{WA} \tag{2-30}$$

The resistivity is

$$\rho = \frac{CWA}{Q_S} \tag{2-31}$$

and the resistance is

$$R_S = \rho \frac{W}{A} = \frac{CW^2}{I_{\text{dc}} T_L} \tag{2-32}$$

This is a fundamental equation in PIN diode theory and design.

Rigorous analysis shows that

$$R = \frac{2kT/q}{I_f} \sinh\left(\frac{W}{2\sqrt{DT_L}} \tan^{-1}\left[\sinh\frac{W}{2\sqrt{DT_L}}\right]\right) \tag{2-33}$$

where k = Boltzmann's constant
T = temperature (in kelvins)
D = diffusion coefficient = $\mu kt/q$

For most PIN diodes, W/DT_L is less than unity, and the equation simplifies to the simple equation above.

Typical data on R_S as a function of bias current are shown in Figure 2-12. Many design choices are available, as the data indicate. Many combinations of W and T_L have been developed to satisfy the full range of applications.

Breakdown Voltage, Capacitance, Q Factor. The previous section on R_S explained how a PIN can become a low resistance, or a "short." This paragraph will describe the other state: a high impedance, or an "open." Clearly, the better PIN diode is the one that has the better on/off ratio at the frequency and power level of interest.

If we return to the undoped or intrinsic I region, we note that it is an almost lossless dielectric. As such, it has a dielectric strength of about 400 V/mil, and all PIN diodes have

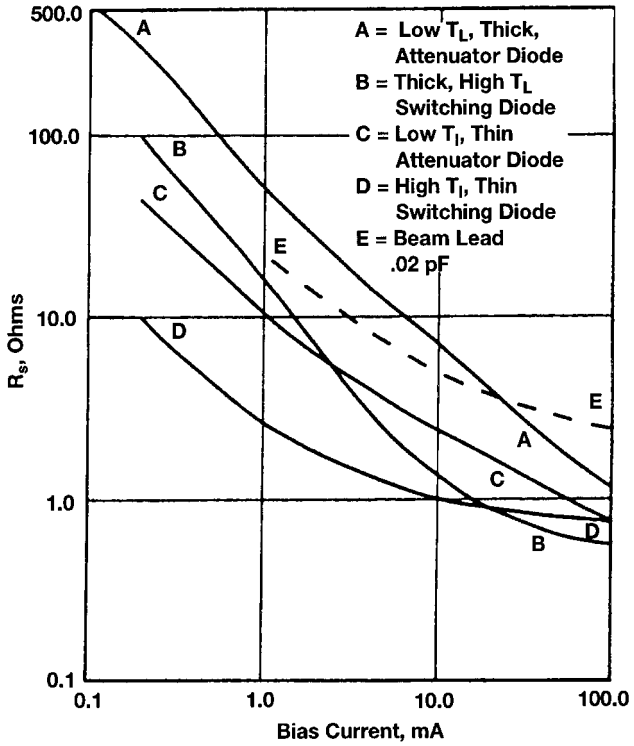


Figure 2-12 Typical series resistance as a function of bias (1 GHz).

a parameter called V_b, breakdown voltage, which is a direct measure of the width of the I region. Voltage in excess of this parameter results in a rapid increase in current flow (called avalanche current); see Figure 2-13. When the negative bias voltage is below the bulk breakdown of the I region, a few nanoamperes will be drawn. As V_b is approached, the leakage current increases often gradually, as is exaggerated in the curve. This current is primarily caused by less-than-perfect diode fabrication, although there is some contribution from temperature. Typically, the leakage current occurs at the periphery of the I region. For this reason, various *passivation* materials (silicon dioxide, silicon nitride, hard glass) are grown or deposited to protect and stabilize this surface and minimize leakage. These techniques have been well advanced over the years, and PIN diode reliability has improved as a result.

Most diodes are specified in terms of minimum V_b for a nominal leakage, usually 10 μA.

It will be noted later that RF voltage swings in excess of the rated V_b are permitted, for the mechanisms causing leakage current do not always respond at radiofrequencies. However, bulk breakdown is effectively instantaneous, and that voltage should never be exceeded.

The next characteristic of our "open" circuit is the capacitance. In simplest form, the capacitance of a PIN diode is determined by the area and width of the I region and the dielectric constant of silicon; however, we have discussed the fact that intrinsic material does contain some carriers and therefore has some conductivity. An *E* field could not exist unless all these carriers were swept out, or depleted.

Application of a reverse bias accomplishes this. At zero bias, the excess carriers on either side of the junction are separated, held apart, by "built-in" fields. This is the contact potential (about 0.5 V for silicon). If there are only a few excess carriers in the I region, this "potential" can separate the charges more easily. The junction "widens" in the sense that, starting at the p^+ and I interface, there is a region of no free carriers called the *depletion zone*. Beyond this depletion zone, the I region still contains the free charges with which it started. With the application of reverse bias, the depletion zone widens.

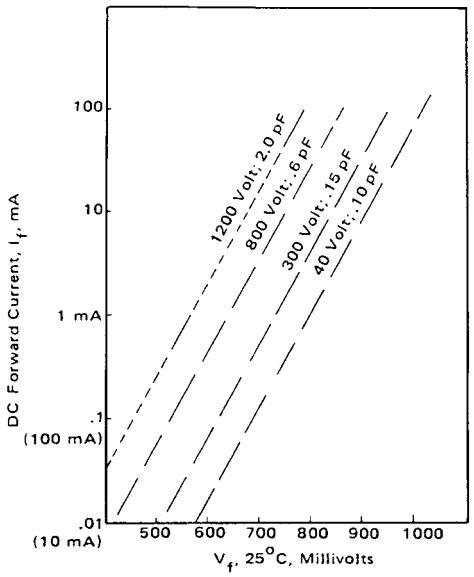


Figure 2-13 Current versus voltage for various PIN diodes.

Eventually, at a bias equal to a so-called *punchthrough* voltage (V_{PT}), the depletion zone fills the entire I region. At this voltage, the 1-MHz capacitance bottoms out and the diode Q reaches its maximum. Figure 2-14 illustrates the equivalent circuit of the I region before punchthrough.

Some very interesting facts can be derived from this model. Consider the undepleted region; this is a lossy dielectric consisting of a volume (area A, length ℓ) of silicon of permittivity 12 and resistivity ε. The capacitance is

$$12\frac{\varepsilon_0 A}{\ell} \tag{2-34}$$

and the admittance is

$$2\pi\frac{12\varepsilon_0 A}{\ell} \tag{2-35}$$

The resistance is $\varepsilon\ell/A$ and the conductance is $A/\varepsilon\ell$.

At very low frequencies, the undepleted zone looks like a pure resistor. At very high frequencies it looks like a lossy capacitor. The "crossover" frequency depends on the resistivity of the I-region material. For ε of 160 Ω/cm, the frequency is 1 GHz. Higher resistivity is generally used for PINs—say, 1000 Ω/cm, and the crossover frequency is 160 MHz.

Diode manufacturers measure junction capacitance at 1 MHz; clearly, what is measured is the depletion-zone capacitance.

If the I region thickness is W and the depletion is X_d, the undepleted region is $(W - X_d)$.

The capacitance of the depleted zone is, proportionally,

$$\frac{1}{X_d} \tag{2-36}$$

and of the undepleted zone,

$$\frac{1}{W - X_d} \tag{2-37}$$

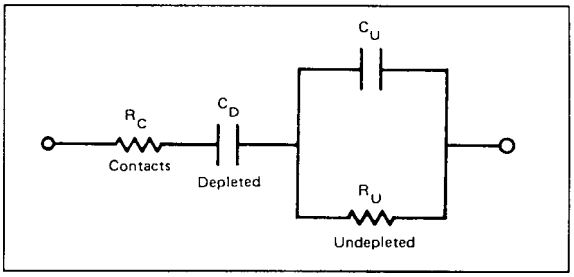


Figure 2-14 Equivalent circuit of I region before punchthrough.

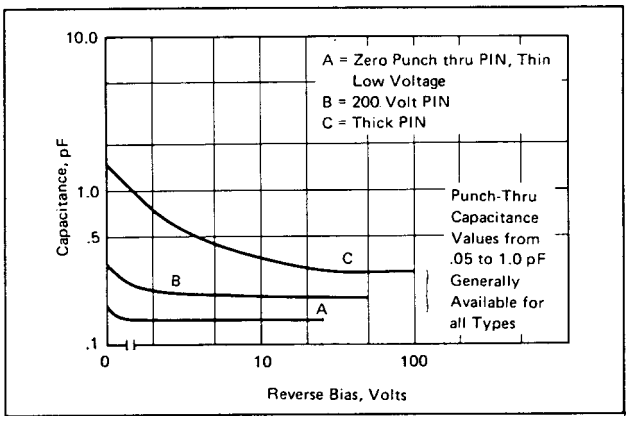


Figure 2-15 Typical 1-MHz capacitance.

The 1-MHz capacitance as a function of reverse bias is seen in Figure 2-15. The 1-MHz capacitance decreases with bias until punchthrough, where $X_d = W$. However, at microwave frequencies well above the crossover, the junction looks like two capacitors in series:

$$C_T = \frac{C_d C_u}{C_d + C_u} \propto \frac{1}{W} \tag{2-38}$$

That is, the microwave capacitance tends to be constant, independent of X_d and bias voltage.

However, since the undepleted zone is lossy, an increase in bias up to the punchthrough voltage reduces the loss.

At any given frequency, the equivalent network can now be drawn as Figure 2-16. R_v is now the equivalent series resistance of the undepleted region. Typical R_v data are shown in Figure 2-17. An alternate equivalent network is shown in Figure 2-18, and typical R shunt data are shown in Figure 2-19.

A good way to understand the effects of series resistance is to observe the insertion loss of a PIN chip shunt mounted in a 50-Ω line, as shown in Figure 2-20. An accepted way to include reverse loss in the figure of merit of a PIN diode is to write

$$Q = \frac{1}{2\pi + \sqrt{R_s R_v C}} \tag{2-39}$$

where R_s and R_v are measured under the expected forward- and reverse-bias conditions at the frequency of interest.

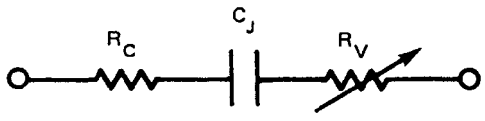


Figure 2-16 Simplified equivalent series circuit.

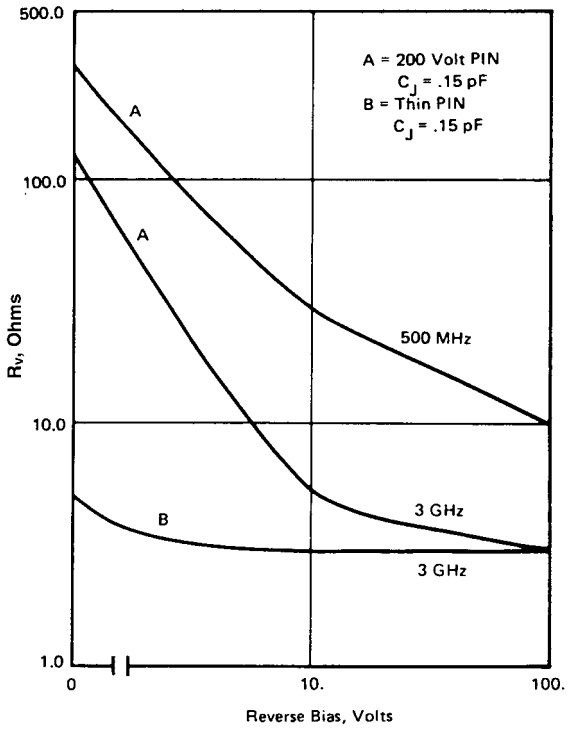


Figure 2-17 Reverse series resistance R_V.

The punchthrough voltage is a function of the resistivity and thickness of the I region. It is advisable to measure loss as a function of bias voltage and RF voltage to determine if the correct diode has been selected for your application. (*Note*: At frequencies below crossover, and diodes with thin I regions, the effective junction capacitance can increase substantially at low forward bias, on the order of 1–200 μA.)

Incidentally, if you are working with PIN or NIP chips that do not have an opaque covering, note that PIN diodes are photosensitive. Incident light causes photogeneration of carriers in the I region, increasing the chip's insertion loss.

PIN Diode Applications. If the intrinsic zone is thick (10–100 μm), we then have a high-reverse-voltage rectifier with a low forward-voltage drop at high current or, in other

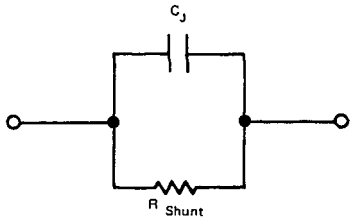


Figure 2-18 Simplified equivalent circuit, shunt.

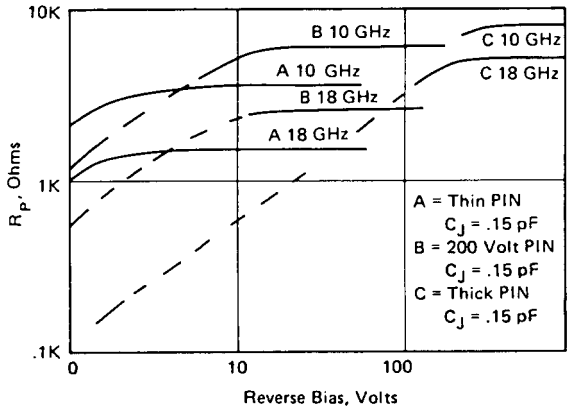


Figure 2-19 Reverse shunt resistance, R_p.

words, a highly efficient rectifier. The low forward voltage results from the fact that the conductivity of the I zone can be modulated by large amounts of charge carriers injected from the *p* and the *n* zones.

Another application of PIN diodes is the high-frequency (HF) field. Here, the fact is exploited that, due to the long carrier lifetime at frequencies beyond approximately 10 MHz, a rectifying effect will no longer occur and the PIN diode rather behaves like a real resistance, the magnitude of which depends on the forward direct current passed by the device and

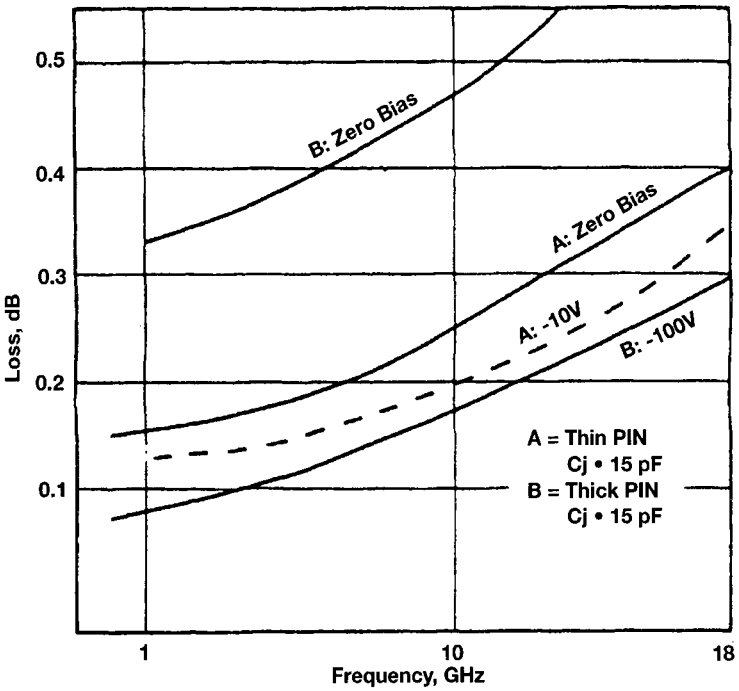


Figure 2-20 Insertion loss versus frequency.

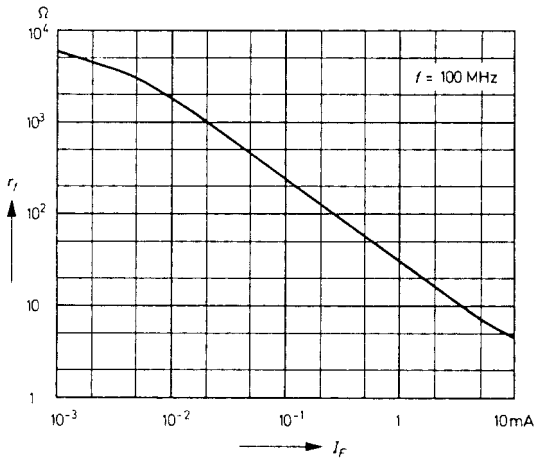


Figure 2-21 Forward resistance versus forward current.

produces an equal effect on both half waves of the HF signal. In view of this behavior, the PIN diode can be used as a switch or a variable resistor for HF signals. Thus it becomes possible, for example, to subject an HF signal to amplitude modulation by means of an AF-controlled PIN diode.

An important application of PIN diodes that has found favor in recent times is for dc-operated attenuators in TV tuners and antenna distribution amplifiers. Figure 2-21 shows the real HF forward resistance r_f as a function of the forward current I_f, measured at 100 MHz. Figures 2-22 and 2-23 show second-order IMD and cross-modulation for PIN diodes.

Applying the PIN in Amplitude Control of High-Frequency Signals. In conventional transistorized TV tuners, automatic gain control (AGC) is usually achieved by varying the emitter current of the input transistor. This method exploits the fact that the input transistor exhibits maximum gain at a certain level of emitter current and that this gain decreases when the emitter current is either increased or decreased relative to this point. Since modulation capacity and cross-modulation resistance grow with the emitter current, "upward" gain control has lately become the preferred solution, and according to it, gain reduction is achieved by an increase of the input transistor's emitter current. An input stage based on this design, therefore, is least resistant to cross-modulation when receiving a signal from a weak transmitter because it is operating at or below the point at which AGC action begins. This effect is particularly disadvantageous when the signal of a weak transmitter is to be received in the presence of a strong local transmitter. The unsatisfactory cross-modulation properties of such an input stage manifest themselves more and more frequently in the form of image perturbations as the TV reception band is crowded by an increasing number of transmitters. Another disadvantage is the variation in the transistor parameters that results from the control action. The variations affect antenna matching at the input and the response of the RF filter at the output.

To obviate these problems, methods being adopted today provide for the input stage to be equipped with cross-modulation-resistant transistors with fixed bias, gain control being provided by a variable attenuator preceding the transistor(s). The input-filter termination

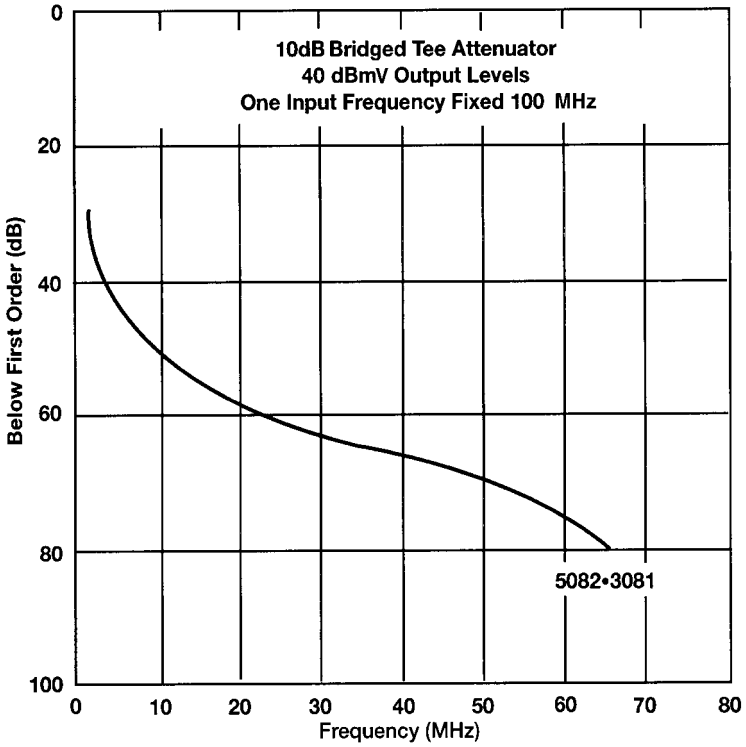


Figure 2-22 Second-order IMD in a PIN diode. Equation (1-48) can be used to determine the diode's IP_2 from the absolute values of its fundamental and IM_2 outputs.

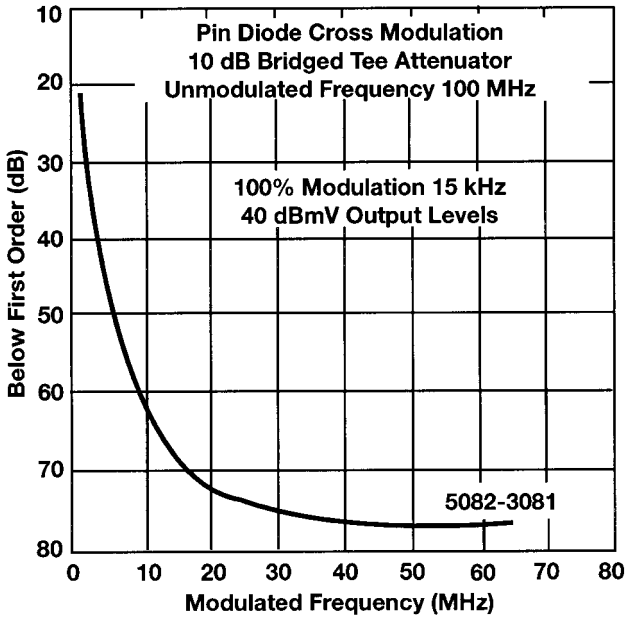


Figure 2-23 Cross-modulation in a PIN diode. Equation (2-61) relates cross-modulation level to IMD.

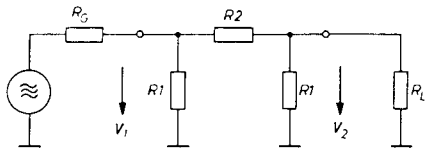


Figure 2-24 Basic circuit of a π-mode filter attenuator.

provided by this attenuator must be independent of the degree of attenuation to preserve the filter's response characteristics.

The first two requirements are met by the properties of the PIN diode and the latter requirement can be met by designing a PIN diode control in the form of a π network, a schematic of which is shown in Figure 2-24. The attenuation of a matched π network is expressed by the following equations:

$$\frac{a}{\text{dB}} = 20 \log \frac{Z + R_2}{R_2 - Z} \tag{2-40}$$

$$\frac{a}{\text{dB}} = 20 \log \left(\frac{R_1}{Z} + \sqrt{\frac{R_1^2}{Z^2} + 1} \right) \tag{2-41}$$

Figure 2-25 plots the relationship between the resistance levels of R_1 and R_2 and the input and output voltages of the π network. Figure 2-26 shows the circuit diagram of an implementation of the π network composed of PIN diodes. The resistance characteristic required according to Figure 2-25 is achieved approximately by varying the control current I_{co}.

If the control current I_{co} in Figure 2-26 equals zero, then—due to the fact that $R_1 = R_2$ and $R_3 = R_4$—the auxiliary voltage V_h causes forward currents of equal magnitude to flow through diodes D_1 and D_3. The voltage drops across R_1 and R_2 are of equal magnitude, so that no current passes through diode D_2. Therefore, the latter presents a high resistance, whereas D_1 and D_3 present a low resistance. Under these conditions, the network produces maximum attenuation. A control current I_{co} reduces the forward currents through D_1 and D_3 and allows a forward current to be passed by diode D_2. When it has attained its maximum value, diodes

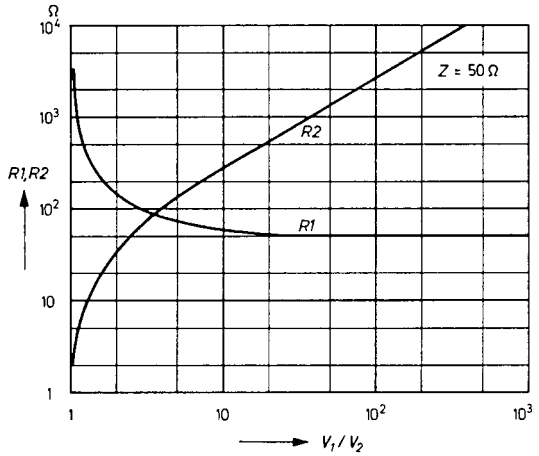


Figure 2-25 Values of R_1 and R_2 in Figure 2-24 versus relation of input and output voltage.

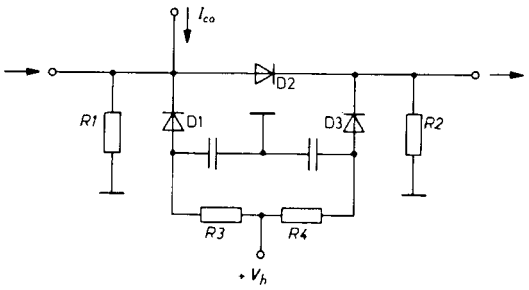


Figure 2-26 PIN diode π-mode attenuator.

D_1 and D_3 are blocked—that is, they are high-ohmic—and D_2 is low-ohmic. Under these conditions, the network produces its minimum attenuation.

If the maximum attenuation of the π network is to be increased, it is possible to make the two shunt diodes D_1 and D_3 more low-ohmic than the impedance Z, although this would reduce the reflection coefficient of attenuation. For example, if a reflection coefficient of about 7 dB is accepted over the entire control range, then a maximum attenuation of 25–30 dB could be achieved in the VHF range. The series inductances of the transverse diodes and the shunt capacitance of the series diode are responsible for this relatively unsatisfactory value. Improved attenuation is brought about with the integrated π network described in the next section.

Example: A PIN Diode π Network for TV Tuners. The since-discontinued TDA1053 is an example of an integrated PIN π network attenuator intended for TV-tuner use. It comprises three silicon planar PIN diodes connected to form a π network (Figure 2-27) and serves for the electronic amplitude control of the input signals of TV tuners and antenna distribution amplifiers in the 40–1000-MHz range. Its input and output impedances remain constant over the entire control range. This can also be achieved with discrete diodes.

The TDA1053 was normally supplied with vertical leads. The characteristics stated below apply to devices of this configuration.

MAXIMUM RATINGS OF INDIVIDUAL DIODES

Reverse voltage	V_R	30	V
Forward current at $T_{amb} = 25$ °C	I_F	50	mA
Junction temperature	T_j	125	°C
Storage temperature range	T_s	−55 to +125	°C

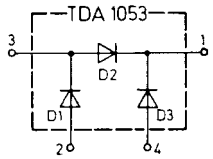


Figure 2-27 Internal circuitry of the TDA1053.

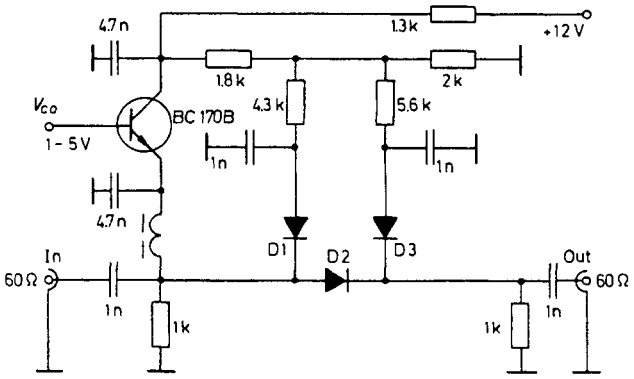


Figure 2-28 Test and application circuit.

MAXIMUM RATINGS OF THE π NETWORK

Ambient operating temperature range when operating according to Figure 2-28	T_{amb}	100	°C

CHARACTERISTICS OF INDIVIDUAL DIODES AT T_{amb} = 25 °C

Forward voltage at I_F = 50 mA	V_F	< 1.2	V
Forward current at V_R = 15 V	I_R	< 500	nA
Differential forward resistance:			
At I_F = 10 mA, f = 100 MHz	r_f	5	Ω
At I_F = 10 µA, f = 100 MHz	r_f	1.4	kΩ

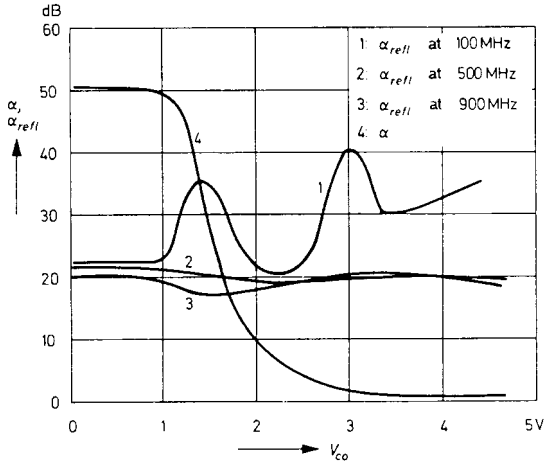


Figure 2-29 Attenuation and reflection attenuation versus control voltage.

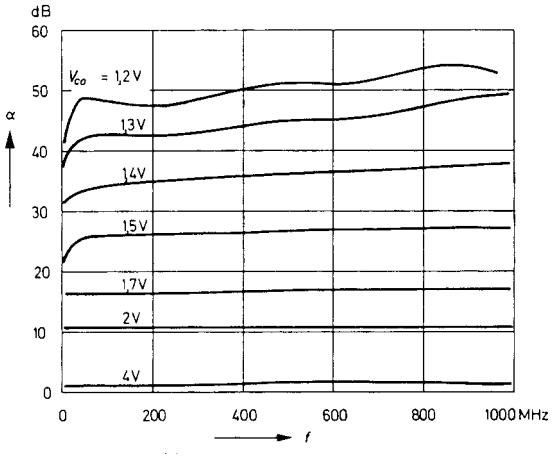


Figure 2-30 Attenuation versus frequency.

CHARACTERISTICS IN THE TEST CIRCUIT (FIGURE 2-28) AT T_{amb} = 25 °C

Voltage for 1% cross-modulation	V_{cr}	1	V
Attenuation in the 40–1000-MHz range:			
At V_{co} = 1.5 V (1–2 V)	a_{max}	45 (> 36)	dB
At V_{co} = 5 V (4–5 V)	a_{min}	1.5 (< 2)	dB
Reflection coefficient in the 40–1000-MHz range over the entire control range, depending on circuit design	a_{refl}	20 (> 16)	dB

These data reveal that the compact design of the three PIN diodes in a common 50B4 plastic package guarantees favorable values for minimum and maximum attenuation, as well as reflection attenuation. The test and application circuit shown in Figure 2-28 also comprises the transistor control-signal amplifier. The typical characteristic of the attenuation and the reflection attenuation for this circuit are shown in Figure 2-29, as a function of the control voltage V_{co}. Figure 2-30 shows the attenuation at different control voltages, as a function of frequency.

2-1-4 Tuning Diodes

Introduction. In recent years, continuous development of tuning diodes—also known as *varactors* or *varicaps*—together with increased commercial and military use has led to substantial improvement in Q, reproducibility, and reliability. Concurrently, new techniques for producing and controlling a hyperabrupt dopant profile in the semiconductor permit the capacitance–voltage law to be much faster than the classical square root or cube root behavior.

Current tuning diode materials include silicon and gallium arsenide; silicon is favored for low cost and lower Q applications from HF through microwave frequencies. Hyperabrupt varactors, also of silicon, are finding many applications in commercial television tuners, where their high tuning ratios, linear tuning, and low cost are needed. New developments include low-capacitance hyperabrupts for microwave and wireless applications.

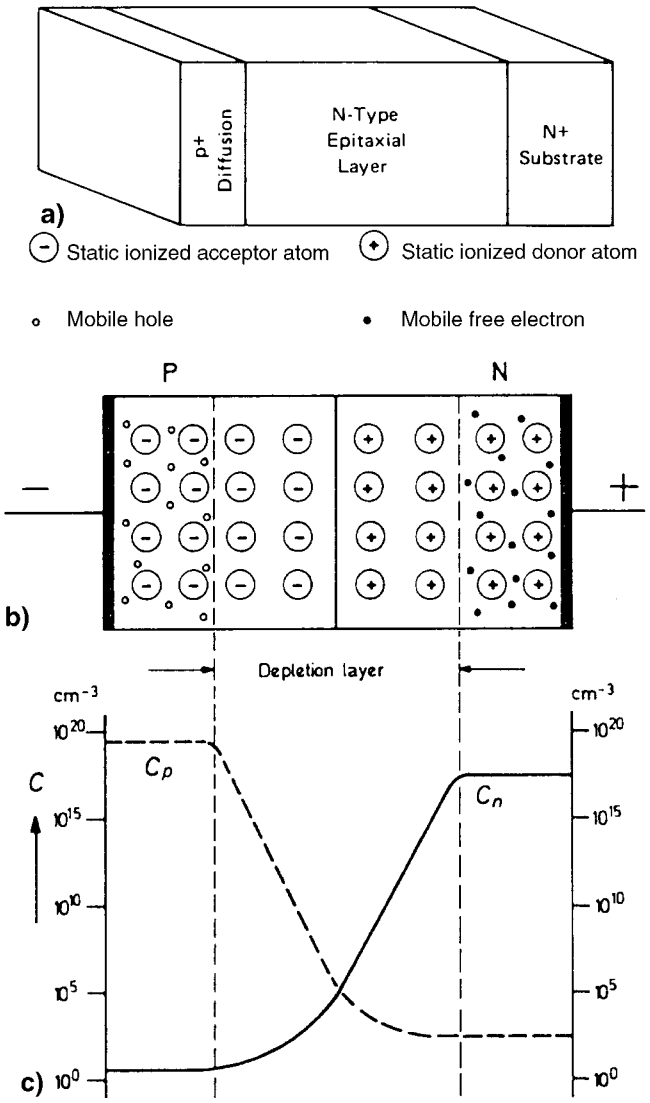


Figure 2-31 (a) Basic PIN structure; (b) cross section of a reverse-biased *pn* junction; and (c) density distribution of free charge carriers.

Gallium arsenide is used with high operating frequency and dictates the highest Q possible, as in parametric amplifiers and millimeter multipliers.

This section will acquaint the reader with tuning diodes: how they work, and what they can or cannot be expected to do in an electronic circuit. The basic properties of a tuning diode will be described in terms of the parameters that manufacturers use in characterizing them. The following topics will also be addressed:

- Capacitance ratio with respect to voltage and voltage breakdown
- Q as a function of design and operating conditions

- Stability—leakage current, temperature coefficient, and post-tuning drift
- Distortion products
- Packaging parasitics
- Applications—suggestions on how to specify a varactor

Tuning Diode Physics

Introduction. All junction diodes are made up of the same physical parts: a *pn* junction, a carefully controlled epitaxial layer, and a very-low-resistance substrate. These parts are shown in Figure 2-31.

No matter what type of junction device we are discussing—a tuning diode, a step-recovery diode, or a PIN diode—these parts are all present; the main difference between these devices is the resistivity and thickness of the epitaxial layer. Tuning diodes and multiplier diodes need epitaxial layers where both the resistivity and thickness are carefully controlled.

Abrupt Junction. An abrupt-junction diode is one in which the p^+ diffused region of the diode is much more highly doped than the epitaxial layer. Also, the high doping drops to the doping level of the epitaxial layer in a distance that is short compared to the epitaxial layer thickness, and the doping level of the epitaxial layer is constant over its thickness. This is shown in Figure 2-32, with the corresponding C–V curve shown in Figure 2-33. When these requirements are satisfied, the diode capacity, diode area, epitaxial layer doping level, and diode voltage are related by

$$\frac{C(V)}{A} = K\left(\frac{N}{V+\varphi}\right)^n \tag{2-42}$$

where $C(V)$ = capacitance of the diode at voltage V
A = area of the diode
N = doping level of the epitaxial layer

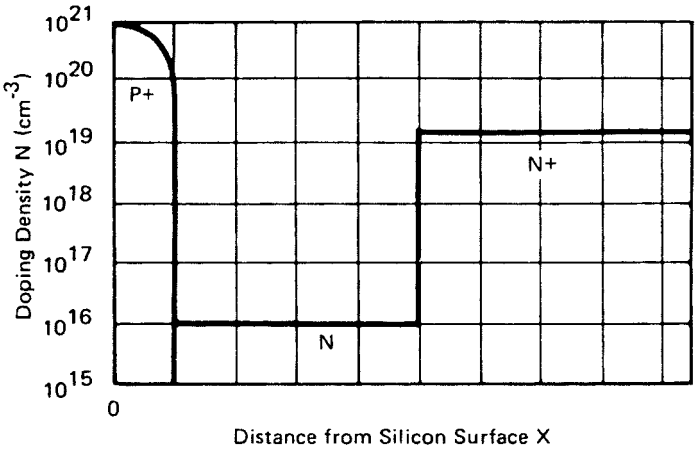


Figure 2-32 *N–X* abrupt junction.

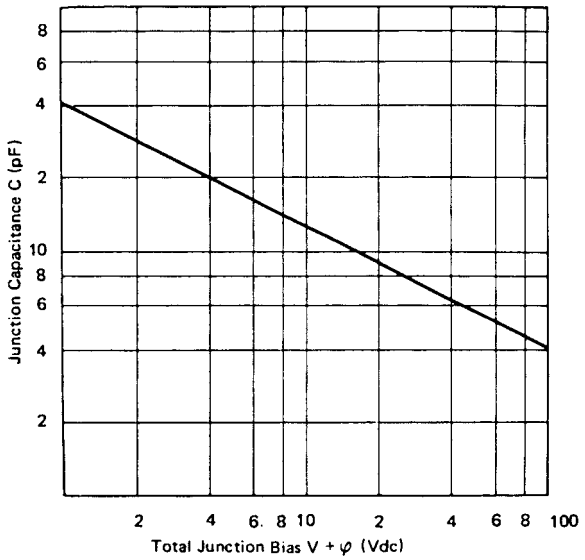


Figure 2-33 Capacitance versus total junction bias for abrupt-junction diode.

V = voltage applied to the diode
φ = built-in potential of the diode (0.6–0.8 V)
n = slope of diode C–V curve; $n \approx 0.5$ for an abrupt-junction diode
K = constant

As a consequence of the physical properties of a *pn* junction, a depletion layer is formed between the *p* and *n* regions whose width depends on the voltage applied to the diode. The capacitance of the diode is inversely proportional to the width of the depletion layer; that is, $C \sim 1/\ell_0$. In addition, the series resistance of the diode is proportional to the width of the undepleted epitaxial layer. Thus, as diode reverse bias is increased, the depletion layer increases, causing a decrease in capacitance and a decrease in series resistance. As the diode reverse bias is increased further, a point is reached where the electric field caused by the reverse bias reaches a critical level, and current through the diode increases rapidly; this is

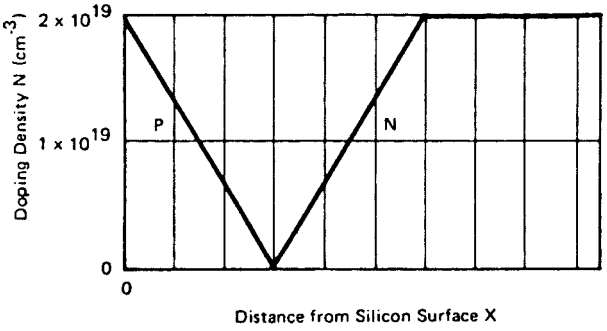


Figure 2-34 *N–X* linearly graded junction.

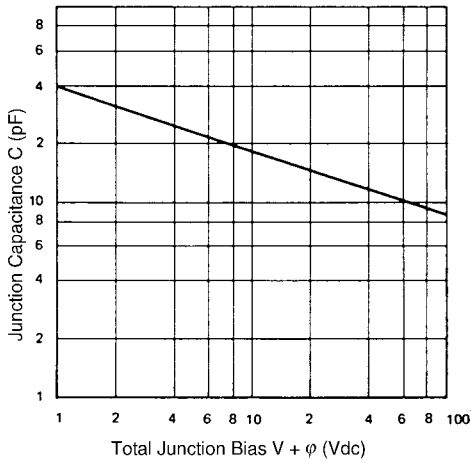


Figure 2-35 Capacitance versus total junction bias for linearly graded junction.

breakdown voltage of the diode. If, at the breakdown voltage, the epitaxial layer is not completely depleted, the diode will have excessive series resistance. Conversely, if the epitaxial layer is depleted before the breakdown voltage is reached, no further capacitance decrease occurs after the total depletion, and a condition called *punchthrough* occurs.

While, in the ideal case, voltage breakdown will occur just as the epitaxial layer is totally depleted, this seldom occurs in practice, and we generally have a condition of either punchthrough or excess series resistance.

Linearly Graded Junction. If, instead of the junction profile shown in Figure 2-32, we have a p^+ region and an n region whose doping levels increase linearly with distance from the pn junction as shown in Figure 2-34, with its corresponding C–V curve in Figure 2-35, then we have what is called a linearly graded junction diode. This diode follows Eq. (2-42) with the exception that the exponent n is equal to $\frac{1}{3}$. This means that, for a given voltage change, the linearly graded junction will have a smaller capacitance change than an abrupt-junction diode. Since, in most cases, the designer is looking for the maximum

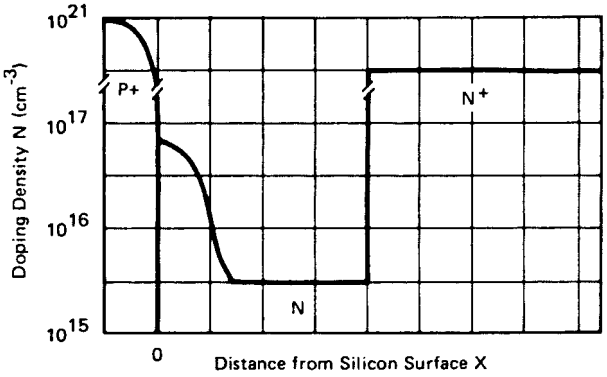


Figure 2-36 *N–X* hyperabrupt-junction diode.

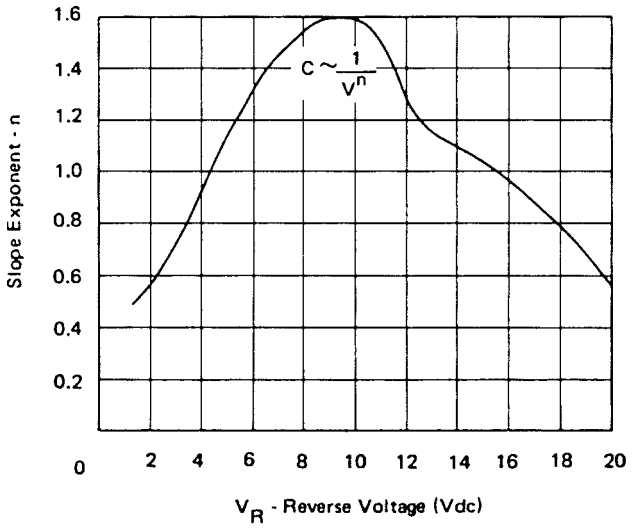


Figure 2-37 Typical *n* versus reverse voltage characteristic for hyperabrupt-junction diode.

capacitance change obtainable, the linearly graded junction is not used as a tuning diode. This structure found its greatest use several years ago as a "cube law" multiplier, but even this use has decreased as new structures have been developed.

Hyperabrupt Junction. The hyperabrupt-junction diode provides a greater capacitance change than the abrupt-junction diode for a given voltage change, as well as a linear frequency versus voltage characteristic over a limited voltage range. The structure of the hyperabrupt-junction diode is shown in Figure 2-36 and can be seen to be an abrupt-junction diode with an additional, increased doping level at the *pn* junction. This diode also follows Eq. (2-42) with the exception that n is not a function of voltage and is generally in the range of 0.5–2. A typical curve of n versus voltage is shown in Figure 2-37.

The C–V curve in a hyperabrupt-junction diode is shown in Figure 2-38 and is seen to start at a high value of capacitance per unit area at low bias (high epitaxial doping) and change to a lower value of capacitance per unit area (low epitaxial doping) at high bias. The details of the curve depend on details of the shape of the more highly doped region near the *pn* junction.

Unfortunately, with a hyperabrupt-junction diode, you must settle for a lower Q than an abrupt-junction diode with the same breakdown voltage and same capacitance at 4 V.

It should be noted that any diode that has an n value that exceeds 0.5 at any bias voltage is, by definition, a hyperabrupt-junction diode. Thus, the hyperabrupt-junction diode family can have an infinite number of different C–V curves. Since the abrupt-junction diode has a well-defined C–V curve, the capacitance value at one voltage is sufficient to define the C–V capacitance at any other voltage. This is not the case for the hyperabrupt-junction diode. In order to adequately define the C–V characteristics of a hyperabrupt-junction diode, two and sometimes three points on the curve must be specified.

Silicon Versus Gallium Arsenide. Everything mentioned so far applies to both silicon and gallium arsenide (GaAs) diodes. The main difference between silicon and GaAs from a

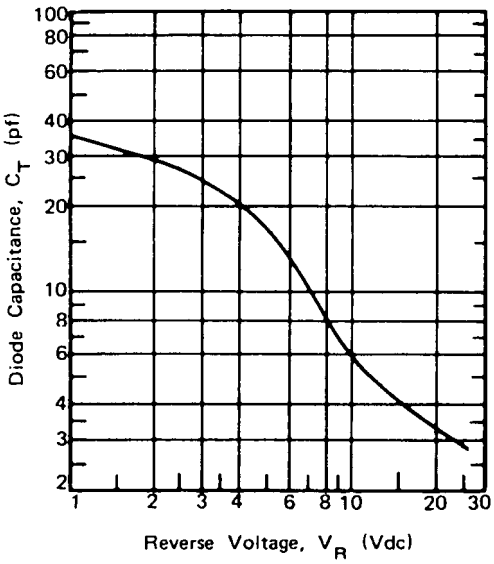


Figure 2-38 Capacitance versus junction bias for hyperabrupt-junction diode.

user's point of view is that higher Q can be obtained from GaAs devices. This is due to the lower resistivity of GaAs from a given doping level N. The resistivity of the epitaxial layer, or substrate, of a diode is given by

$$\rho = \frac{1}{Ne\mu} \tag{2-43}$$

where ρ = the resistivity
N = the doping level of the layer
e = the charge on an electron
μ = is the mobility of the charge carriers in the layer

Gallium arsenide has a mobility about four times that of silicon and, thus, a lower resistivity and higher Q for a given doping level N. Since diode capacitance is proportional to $\sqrt{N}$, independent of resistivity, a silicon diode and a GaAs diode of equal area and doping will have a capacitance difference proportional to the square root of the dielectric constant ratio. This gives the GaAs diode a 5% higher capacitance and is thus of little practical significance. The penalty paid for using GaAs is an unpassivated diode and a more expensive diode due to higher material and processing costs. If the higher Q of the GaAs device is not really needed, a substantial price saving will be obtained by using a silicon device.

Planar Versus Mesa Construction. The two basic construction techniques used to manufacture tuning diodes are planar and mesa; cross sections for both of these devices are shown in Figure 2-39. The planar process, which is the backbone of the integrated circuit industry, lends itself to large-volume production techniques and is the one used for the 1N series of tuning diodes. Mesa processing, on the other hand, requires more

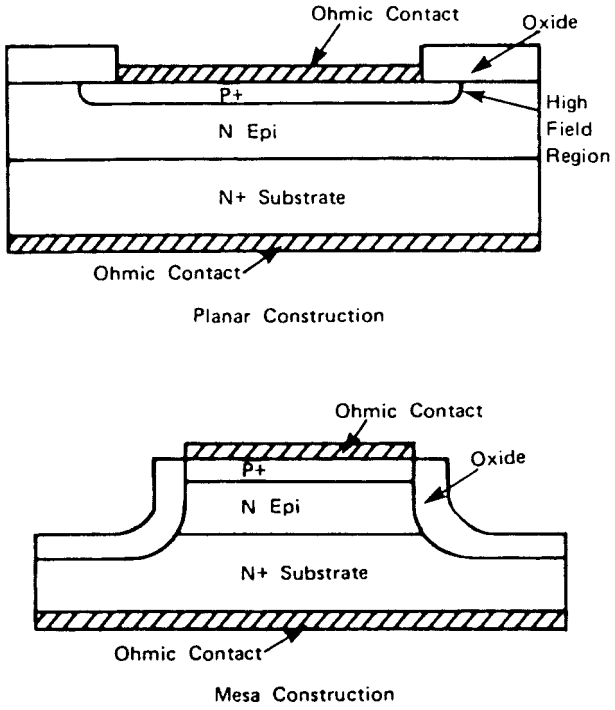


Figure 2-39 Cross sections of planar and mesa devices.

processing steps and is generally done on a wafer-by-wafer basis. This results in a more costly process and thus a more expensive diode. All microwave tuning diodes are of mesa design because of greatly higher Q. Due to the relatively small radius of curvature at the junction edge of a planar diode, the electric field in this area is greater than the electric field in the center (flat) portions of the junction. As a result, the breakdown voltage of the diode is determined by both the epitaxial resistivity and the radius of curvature of the junction edge. Thus, for a given breakdown voltage, a planar diode must use higher resistivity epitaxial material than a mesa diode, which has a completely flat junction. The end result is that the planar diode has a greater series resistance than a mesa diode for the same capacitance and breakdown voltage, and thus lower Q.

Capacitance

Capacitance Ratio. From the user's point of view, the capacitance ratio is simply the capacitance change available in the circuit. Thus, a user tuning from, say, –4 to –45 volts defines capacitance ratio as

$$R = \frac{C_T(-4)}{C_T(-45)} \tag{2-44}$$

where C_T includes C_J plus C_P plus C_F. The manufacturer, however, defines C_T as $C_J + C_P$.

To explore the significance of this difference, let's take two examples, one with a large C_J and another with a small C_J, in chip, package, and "typical" fringe situations. Both are 45-V tuning varactors.

Device	C_{J0}	C_{J45}	Ratio
A	0.6 pF	0.1 pF	6.0
B	15.0 pF	2.5 pF	6.0

Put both devices in a standard 023 package with C_P (strap and ceramic) of 0.18 pF:

Device	C_{T0}	C_{T45}	Ratio
C	0.78 pF	0.28 pF	2.75
D	15.18 pF	2.68 pF	5.67

Notice the drop in capacitance ratio, especially for the low-C_J diode. If we now add a typical 0.04 pF for external fringe capacitance, we get the following results:

Device	C_{T0}	C_{T45}	Ratio
C	0.82 pF	0.32 pF	2.56
D	15.22 pF	2.72 pF	5.6

The reduction in capacitance ratio, and thus circuit tuning capability, by the fringing fields is quite obvious and amounts to 7% in this example.

Because of the often stringent specifications on tuning ratio, it is mandatory that the manufacturer and customer clearly agree on the exact design of the holder used to measure the varactor in question.

Having described how to measure capacitance, it is relatively easy to describe the results. The section on diode physics described the various types of "laws," or C–V curves, and we will not repeat them here. Nonetheless, several important points must be covered.

The first is "available capacitance swing." The laws indicate a steadily decreasing capacitance with voltage, which indicates that the epitaxial region is widening and the electric field is increasing. (For an abrupt junction, since $C \propto 1/\sqrt{V}$, the depletion zone with W is increasing as $\sqrt{V}$, and the electric field V/W increases as $\sqrt{V}$.)

Two things can happen:

1. The junction width widens so that the entire intrinsic region is depleted. The capacitance bottoms out, resulting in voltage punchthrough.
2. The electric field exceeds the dielectric strength of silicon (or GaAs), and "solid-state discharge" or "avalanche" current is drawn.

The diode impedance drops, the varactor no longer "varacts," and circuit operation ceases. Moreover, if more than a few milliamperes of current is drawn, localized overheating may destroy the diode, resulting in breakdown voltage. All varactors are characterized for breakdown voltage—for example, 45 V minimum.

The theoretical tuning varactor is designed so that the punchthrough occurs at a voltage equal to the voltage breakdown of the diode. Logically, this means that in order to obtain

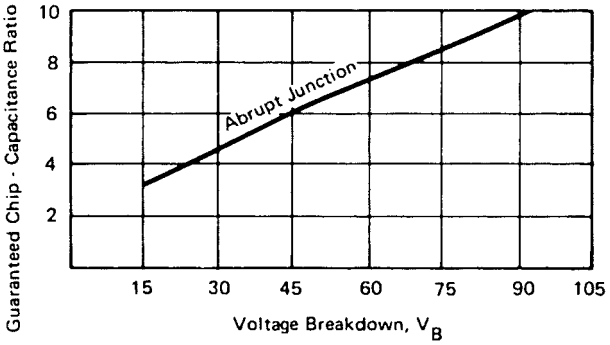


Figure 2-40 Capacitance ratio versus breakdown voltage.

greater tuning ratios, it is necessary to be able to increase the depletion-layer width without reaching punchthrough or breakdown. You must have a thicker epitaxial region to make this possible.

Figure 2-40 shows catalog ratio values, from zero bias to breakdown, as a function of breakdown voltage necessary. In the next section, on Q, we will discuss other elements in your choice of V_B.

To complete this section, we should mention that semiconductor processing control has been refined so well that capacitance tracking to within ±1% over the full range from zero to breakdown is now readily obtainable in production quantities.

Temperature Coefficient of Capacitance (T_{CC}). Unfortunately, since most datasheets give the value of T_{CC} at 4 V, it is sometimes assumed that this value applies at all bias voltages. This is not the case. Consider Eq. (2-45), a rewritten form of Eq. (2-42):

$$C(V) = \frac{C(O)}{(V + \varphi)^n} \tag{2-45}$$

Taking the derivative of this with respect to temperature T we have

$$\frac{dC(V)}{dT} = \frac{+nC(O)}{(v + \varphi)(v + \phi)^n} \frac{d\varphi}{dT} \tag{2-45a}$$

or, after substituting Eq. (2-45):

$$T_{CC} = \frac{1}{C(V)} \frac{dC(V)}{dT} = \frac{-n}{(V + \varphi)} \frac{d\varphi}{dT} \tag{2-46}$$

As a first approximation, we can say that $d\varphi/dT = -2.3$ mV/°C over the temperature range of interest.

From Eq. (2-46) we can draw the following conclusions:

1. The temperature coefficient is inversely proportional to the applied voltage.
2. The temperature coefficient is directly proportional to the diode slope, n.

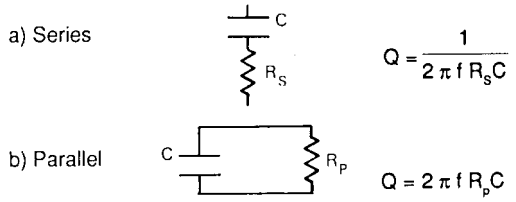
For an abrupt-junction diode that has a constant value of n (0.5), the temperature coefficient has a smooth curve in the form $K/(v + \varphi)$. However, in the case of hyperabrupt-junction diodes, n is a function of voltage, and the shape of the T_{CC} curve depends on the details of the $n(V)$ curve. A typical T_{CC} curve for an abrupt-junction diode is shown in Figure 2-41 and for an Alpha DKV6520 series hyperabrupt-junction diode in Figure 2-42. The inflection in the hyperabrupt T_{CC} is due to the fact that in this voltage range $n(V)$ is increasing faster than $1/V$, giving an increase in T_{CC}. It should also be noted, however, that over the range of the T_{CC} minimum the temperature coefficient is relatively constant, and operation in this area may be advantageous in some applications where a restricted tuning range can be used.

Q Factor or Diode Loss

Definitions. The classical definition of the Q of any device or circuit is

$$Q = \frac{2\pi \times \text{Energy stored}}{\text{Energy dissipated per cycle}} \tag{2-47}$$

For a capacitor, two formulations are possible:



Clearly, the two definitions must be equal at any frequency, which establishes

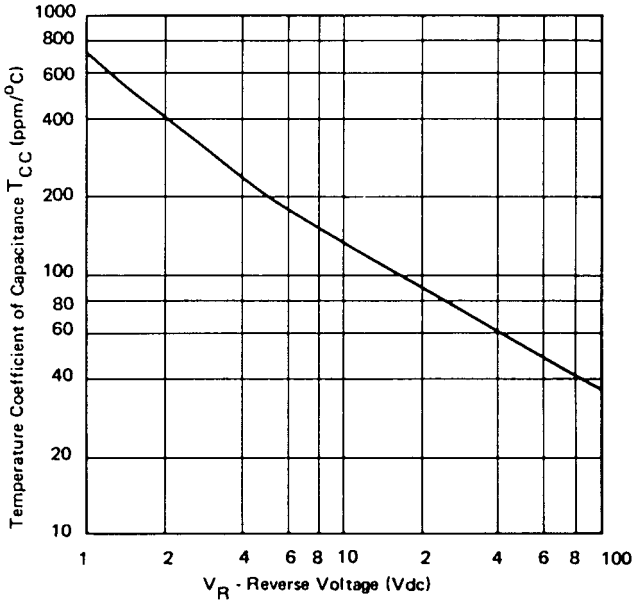


Figure 2-41 Temperature coefficient of capacitance versus tuning voltage for an abrupt-junction diode.

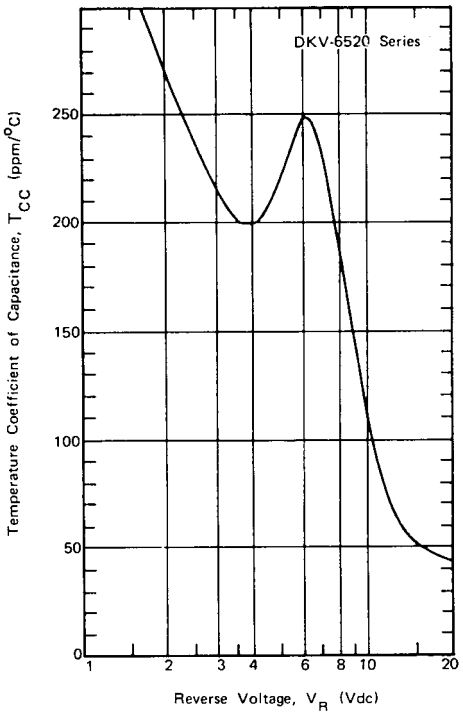


Figure 2-42 Temperature coefficient of capacitance versus tuning voltage ($T_A = 25$ °C) hyperabrupt-junction diode.

$$R_P = \frac{1}{(2\pi f)^2 C^2 R_S} \tag{2-48}$$

In the case of a high-Q tuning diode, the proper physical model is the series configuration, because the depleted region is almost perfectly pure capacitance, and the undepleted region, due to its relatively low resistivity, is almost a pure resistor in series with the capacitance. Furthermore, the contact resistances are also clearly in series. Q, then, for a tuning diode, is given by

$$Q_{(-v)} = \frac{1}{2\pi f_o R_{(-v)} C_{(-v)}} \tag{2-49}$$

where f_o is the operating frequency, $C_{(-v)}$ is the junction capacitance, and $R_{(-v)} = R(\text{epi}) + R_{C'}$, the sum of the resistance of the undepleted epitaxial layer and the fixed contact resistance.

Cutoff frequency, f_c, is defined as that frequency at which Q equals unity. Thus,

$$f_{c_{(-v)}} = \frac{1}{2\pi R_{(-v)} C_{(-v)}} \tag{2-50}$$

Historically, the tuning diode business has specified Q at 50 MHz, despite the fact that Q values of microwave diodes are so high that it is almost impossible to measure them at 50

Table 2-7 Parameters for 0.6-pF diode

Breakdown Voltage V_B	Resistivity ($\Omega \cdot$ cm)	Junction Diameter (mills)	Epitaxial Region (μm)	Depleted Epitaxial $V = -4$ V (μm)	Undepleted Epitaxial $V = -4$ V (μm)	R Undepleted (Ω)	R_{sp} (Ω)	Q_4
30	0.31	2.4	1.37	0.54	0.83	0.81	0.18	5200
45	0.52	2.8	2.25	0.73	1.52	1.86	0.15	2600
60	0.74	3.2	3.20	0.90	2.30	3.24	0.13	1500
90	1.25	3.7	5.27	1.21	4.06	7.17	0.10	700

MHz. Instead, as discussed below, Q is measured at microwave frequencies (e.g., 1–3 GHz) and related to 50 MHz by the relationship

$$Q_{(f1)} = Q_{(f2)} \frac{f_2}{f_1} \tag{2-51}$$

which derives quickly from the assumption that f_c is independent of the measuring frequency.

Since both junction capacitance and epitaxial layer resistance are functions of the applied bias, it is not possible to calculate Q as a function of bias from a measurement of capacitance alone. Catalog specifications typically show Q at –4 V, together with the capacitance at two or more voltages.

Relative to $Q_{(-4)}$, Q increases faster than the reduction in capacitance for bias greater than 4 V and, conversely, decreases faster for bias less than 4 V.

In the following section, we will discuss the diode design parameters that determine Q. Following this, we will describe some elementary Q measurement techniques.

Causes. In the discussion of device physics, the resistivity of the epitaxial region was discussed, together with its impact on punchthrough and breakdown. For example, Table 2-7 supplies typical resistivity and relative parameters of 0.6 pF ($C_{J(4)}$) diodes of different breakdowns. If we remember that if at any bias lower than breakdown the epitaxial region is not completely depleted, it follows that the undepleted portion presents a resistance in series with the pure capacitance of the depleted zone. The magnitude of this "undepleted" resistance is also shown in the table.

The entry R_{sp} ($R_{spreading}$) is the series resistance between the epitaxial region and the low-resistivity substrate. The calculations are for idealized cylindrical epitaxial regions of uniform resistivity, low-resistivity contact on the anode (top), and low-resistivity substrate on the cathode. This resistance is constant, independent of bias; also shown are epitaxial thickness region and width of the depletion zone at –4 V bias.

Note the substantial reduction in Q for high-voltage diodes caused by the increased epitaxial layer resistance; this is true for any value of capacitance or any type of junction. For greater voltage breakdown, the epitaxial region thickness must be increased, which requires an increase in epitaxial layer resistivity; the higher the resistivity of the undepleted zone, multiplied by the fact that it is much wider for high-voltage diodes, means the resistance increases substantially.

Consequently, a rule of thumb emerges: For maximum Q, never choose a tuning diode with a voltage breakdown in excess of what is needed for the necessary tuning range. If the

Table 2-8 ***Q*** **versus bias for 0.6-pF diode**

Breakdown Voltage V_B	C_{J0}	$Q(0)$	$C_J(-4)$	$Q(-4)$	$C_J(-10)$	Q (−10)	$C_J(-30)$	$Q(-30)$	$C_J(-45)$	$Q(-45)$	$C_J(-60)$	$Q(-60)$	$C_J(-90)$	Q(−90)
30	1.43	1700	0.6	5200	.4	10k	0.23	64k						
45	1.43	850	0.6	2600	.4	5k	0.23	20k	0.19	90k				
60	1.43	550	0.6	1500	.4	3k	0.23	9k	0.19	23k	0.17	170k		
90	1.43	270	0.6	700	.4	1.3k	0.23	3.5	0.19	6k	0.17	10k	0.14	220k

Table 2-9 Capacitance ratios (C_{J0}/C_JV_B)

Breakdown Voltage V_B	Optimum Ratio	Minimum Guaranteed Ratio	$Q_{(-4)}$ Typical
30	6.2	4.5	3000
45	7.5	6.0	2500
60	8.4	7.5	1400
90	10.2	8.7	650

required tuning range is an octave, requiring a 4:1 ratio, the selection of a 30-V diode will result in diode losses half those of a 60-V diode.

Table 2-8 lists capacitance and Q for each of these chips as a function of bias. Please remember that Q is calculated at 50 MHz.

Table 2-9 rewrites the data of Table 2-8 to show available capacitance ratios between zero bias and breakdown. The first column is the theoretical optimum, as tabulated. The second column is the typical catalog specification. The reduction in tuning ratio below theoretical optima is caused by nonideal junction fabrication. The junctions are never perfectly abrupt.

Although the tables and numbers given refer to abrupt-junction silicon devices, the principle applies without exception to all types of tuning diodes. For comparison, Table 2-10 lists available ratios and Q values for a number of different varactors. The high Q values for GaAs and the low values for hyperabrupts are apparent.

One last point: The Q values and series resistance refer to chips only. The effects of package parasitics will be discussed later, but it is important to consider circuit contact losses here. For low-capacitance diodes—for example, $C_{J4} = 0.6$ pF—the epitaxial region contributes a high value of resistance and dominates Q except at punchthrough. Diode contact losses are less significant.

Post-Tuning Drift. Post-tuning drift (PTD) is the change in oscillator frequency with time after the tuning voltage has stabilized. The minimization of PTD has assumed greater importance with the design of more sophisticated electronic countermeasure systems, where rapid, accurate frequency changes are required.

Table 2-10 Comparative Tuning Diodes

Type	C_{J0}	Breakdown Voltage V_B	$Q_{(-4)}$[a]	Ratio C_{J0}/C_JV_B[a]
Silicon abrupt	1.0	30	5,000	4.5
Silicon abrupt	2.5	30	4,600	4.5
Silicon abrupt	5.0	30	3,800	4.5
Gallium arsenide abrupt	1.0	25	10,000	3.6
Gallium arsenide abrupt	0.5	10	17,000	2.5
Silicon hyperabrupt	50.0	22	300	17.0
Silicon hyperabrupt	2.5	22	500	14

[a] Minimum guaranteed.

Post-tuning drift can be characterized as short-term and long-term. Short-term PTD occurs in the time range of tens of nanoseconds to a few seconds, while long-term PTD occurs in the time range of seconds to minutes, hours, or days.

Short-term PTD is mainly dependent on the thermal properties of the diode and is improved by high Q (low power loss) and flip chip construction. Long-term PTD depends on oxide stability and freedom of mobile charge in the oxide. It should be noted that actual oscillation frequency change may occur even with a perfect tuning diode because of variation with frequency in the power dissipated by the diode, changes in the diode heat-sink temperature, and frequency changes due to other circuit elements. Less than 0.01% short- and long-term PTD can be obtained.

Distortion Products. Inasmuch as nonlinear components generate harmonics and other distortion products, an understanding of this mechanism is of prime interest to the circuit designer. In some instances, the distortion products are the desired end result of the circuit design, as in frequency multipliers, where harmonics of the input signal frequency are the required output signal. For other applications, such as tuning-diode-tuned linear circuits, distortion products are extremely undesirable, and in some instances the end product specification may set a maximum limit on the distortion products allowed.

Cross-Modulation. Cross-modulation is the transfer of the modulation on one signal to another signal and is caused by third-order and higher odd-order nonlinearities in the behavior of the device. Rewriting Eq. (2-42) we have

$$C(V) = \frac{C_0}{(1 + V/\varphi)^n} \tag{2-52}$$

where C_0 = capacitance
V = applied voltage = $V_0 + v$
V_0 = dc applied voltage
v = ac applied voltage

Then, for a desired signal of

$$S_1 = v_1 \sin \omega_1 t \tag{2-53}$$

and a second, amplitude-modulated signal of

$$S_2 = v_2(1 + m \cos \omega_m t)\sin \omega_2 t \tag{2-54}$$

it can be shown that the cross-modulation, γ, defined by

$$\text{Output signal} \sim v_1 \sin \omega_1 t + \gamma \sin (\omega_1 \pm \omega_m)t \tag{2-55}$$

is found to be

$$\gamma = \frac{n(n+1)mv_2^2}{4(V_0+\varphi)^2} \tag{2-56}$$

From this equation it can be seen that cross-modulation is:

- Proportional to the square of the interfering signal
- Directly proportional to the interfering signal's modulation index, m
- Independent of the strength of the desired signal
- Independent of the frequencies of the desired and interfering signals (assuming that the nonlinearity to which both signals are subjected is sufficiently frequency-indiscriminate so this is the case)
- Present for all values of n; that is, no value of n gives zero cross-modulation

Solving Eq. (2-56) for the signal level v_2 required to produce cross-modulation of value γ, we have

$$v_2 = \frac{2(V_0+\phi)\gamma}{n(n+1)^m} \tag{2-57}$$

The interfering signal levels required to produce 1% cross-modulation from a 30%-modulated interfering signal applied to an abrupt-junction diode and a hyperabrupt-junction diode are shown in Figure 2-43.

From this figure, it can be seen that the hyperabrupt-junction diode is more susceptible to cross-modulation than the abrupt-junction diode in the region of maximum slope of the

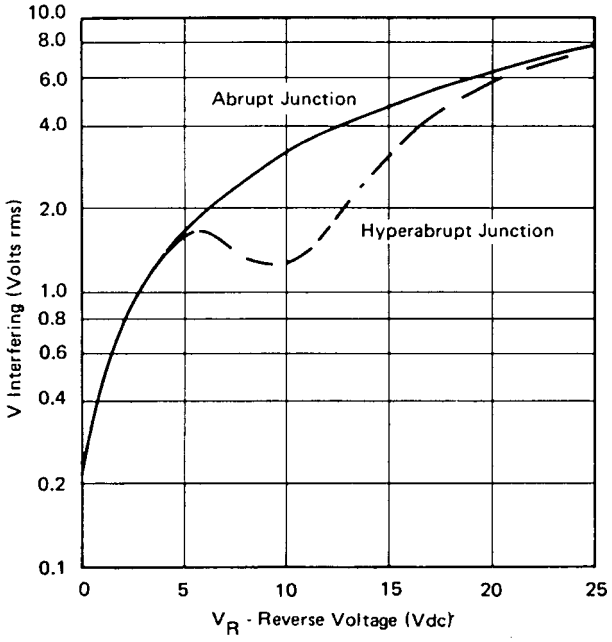


Figure 2-43 Interfering signal (30% amplitude modulated) level versus bias for 1% cross-modulation for abrupt-junction and hyperabrupt-junction diodes.

hyperabrupt-junction diode. For many applications, however, distortion products will be generated by other devices, such as transistors, at signal levels considerably below those given in Figure 2-43.

Intermodulation. Intermodulation is the production of undesired frequencies in the form

$$\sin(2\omega_1 t - \omega_2 t) \quad \text{and} \quad \sin(\omega_1 t - 2\omega_2 t) \tag{2-58}$$

from an input signal in the form of

$$v(\cos \omega_1 t + \cos \omega_2 t) \tag{2-59}$$

From an analysis similar to that done for cross-modulation, it can be shown that

$$\text{Intermodulation} = \frac{n(n+1)v^2}{8(V_0 + \varphi)^2} \tag{2-60}$$

or

$$\text{Cross-modulation} = (2m) \times \text{Intermodulation} \tag{2-61}$$

Harmonic Distortion. Harmonic distortion products are integral multiples of the signal frequencies and decrease in amplitude as the harmonic number decreases. Due to passband considerations and amplitude decrease with harmonic number, the second harmonic is the one of prime concern. Again, it can be shown that the second harmonic, v_2, of a signal of amplitude v_1 is

$$v_2 = \frac{n}{3(V_0 + \phi)} v_1^2 \tag{2-62}$$

Figure 2-44 shows the signal level required to produce 10% second-harmonic distortion in an abrupt-junction and a hyperabrupt-junction diode. Again, as in the case of cross-modulation, the hyperabrupt-junction diode is slightly worse than the abrupt-junction diode in the region of maximum slope of the hyperabrupt-junction diode.

Reduction of Distortion Products. In some cases, the signal levels applied to the diode generate distortion products larger than desirable for the circuit application. In this case, significant reduction in the distortion products can be achieved by using two diodes in a back-to-back configuration, as shown in Figure 2-45. Analysis shows that the fundamental signal components through the diodes are in phase and add, while some distortion products are out of phase and cancel, thus improving distortion performance.

Since the gradient of the electrical field produced in the depletion layer by a reverse bias applied to the device is proportional to the space charge density, the following equations can be written for the junction width W as a function of the reverse bias V_R:

For an Abrupt *pn* Junction (Alloyed Diodes)

$$W = {}_2\sqrt{2\frac{\varepsilon_r \varepsilon_0}{q}\left(\frac{1}{C_p} + \frac{1}{C_n}\right)(V_R + V_D)} \tag{2-63}$$

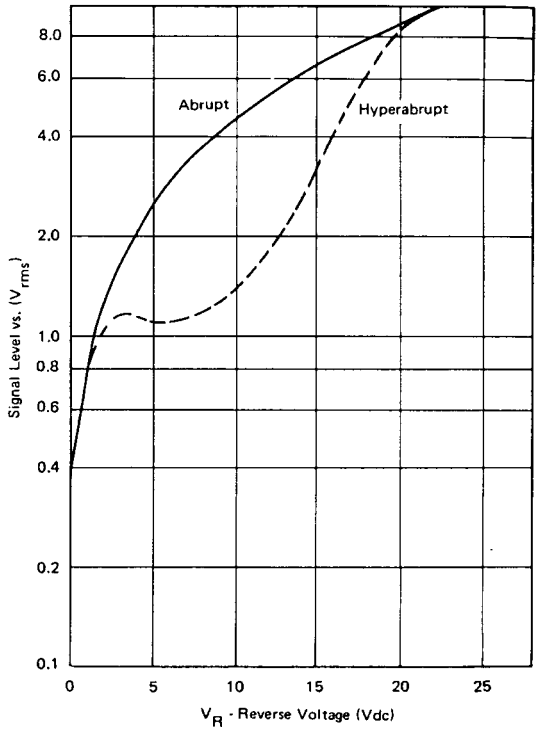


Figure 2-44 Signal level versus reverse voltage for 10% harmonic distortion.

For Linear *pn* Junctions (Single-Diffused Diodes, such as BA110–BA112)

$$W = \sqrt[3]{12\frac{\varepsilon_r \varepsilon_0}{aq}(V_R + V_D)} \tag{2-64}$$

where a is the impurity gradient within the depletion layer, ε_0 is the absolute dielectric constant ($8.85 \times 10^{-14}\ A_s/V_{cm}$, and $\varepsilon_r \approx 12$, the relative dielectric constant of silicon.

The junction capacitance, which is inversely proportional to the junction width, therefore varies in alloyed diodes with the square root, and in single-diffused diodes with the cube root of the externally applied reverse bias, and can be calculated from the general equation

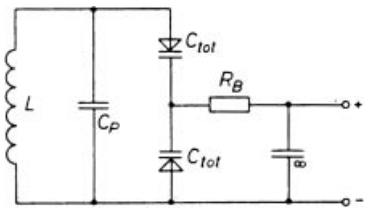


Figure 2-45 Back-to-back diodes. C_P is a fixed parallel capacitance, R_B is a bias decoupling resistor, and the capacitor marked ∞ is a low-impedance bypass.

$$C = \frac{\varepsilon_r \varepsilon_0 S}{W} \tag{2-65}$$

where S is the surface area of the *pn* junction. By way of approximation, we can also use the equation

$$C = \frac{K}{(V_R + V_D)^n} \tag{2-66}$$

where all constants and all parameters determined by the manufacturing process are contained in K. The exponent n is a measure of the slope of the capacitance/voltage characteristics and is 0.5 for alloyed diodes, 0.33 for single-diffused diodes, and (on average) 0.75 for tuner diodes with a hyperabrupt *pn* junction. Figure 2-46 shows the capacitance/voltage characteristics of an alloyed, a diffused, and a tuner diode.

Recently, an equation has been indicated which, although purely formal, describes the practical characteristics better than Eq. (2-66):

$$C = C_0 \left(\frac{A}{A + V_R} \right)^m \tag{2-67}$$

where C_0 is the capacitance at $V_R = 0$, and A is a constant whose dimension is a voltage. The exponent m is much less dependent on voltage than the exponent n in Eq. (2-66).

Equations (2-63) through (2-67) express the pure junction capacitance of the capacitance diode, but to this must still be added a constant capacitance, determined by structure parameters, in order to obtain the diode capacitance C_{tot}, which interests the user. With high inverse voltages—that is, low junction capacitance—a difference will therefore arise between the theoretical capacitance/voltage characteristic according to Eq. (2-66) and the practical characteristic, as shown in Figure 2-47.

The operating range of a capacitance diode or its useful capacitance ratio

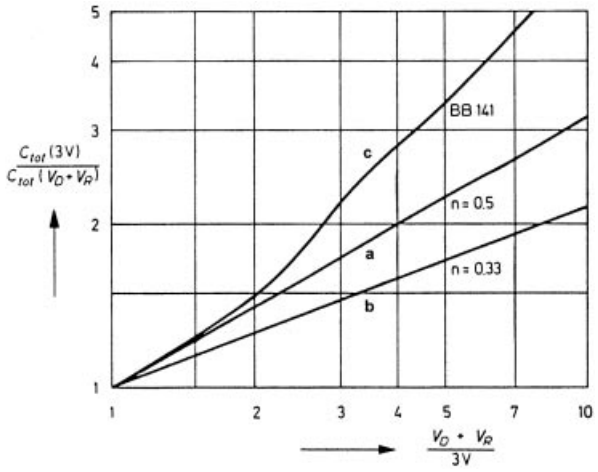


Figure 2-46 Capacitance/voltage characteristic for (a) an alloyed capacitance diode, n = 0.5; (b) a diffused capacitance diode, n = 0.33; and (c) a wide-range tuner diode (BB141).

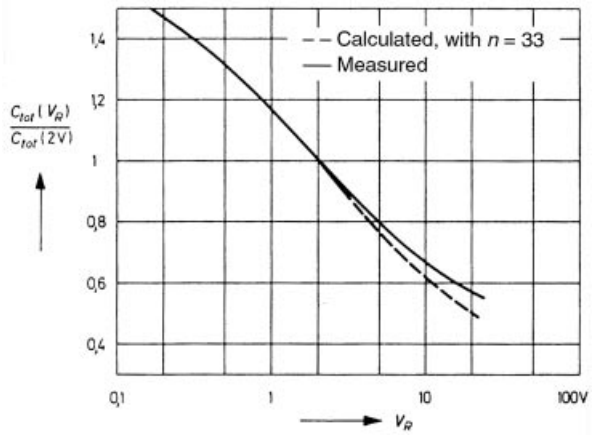


Figure 2-47 Capacitance/voltage characteristic of the BA110 diode.

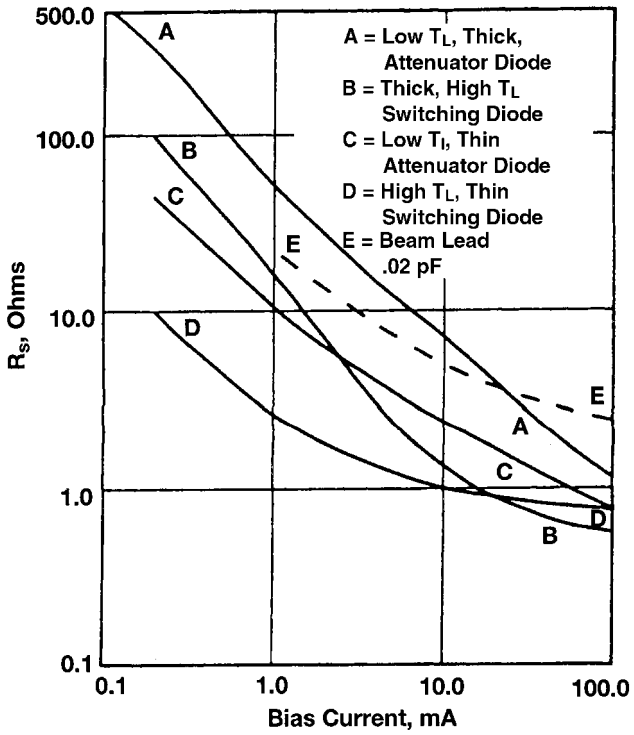


Figure 2-48 Basic current/voltage and capacitance/voltage characteristics.

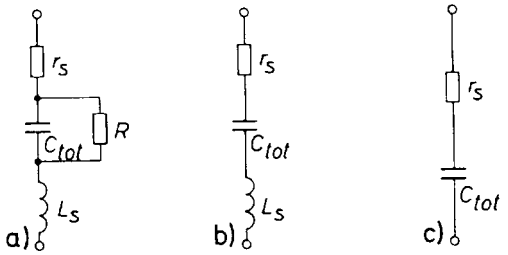


Figure 2-49 Equivalent circuits for capacitance diodes: (a) complete circuit, (b) simplified circuit, and (c) further simplification for low frequencies.

$$\frac{C_{\max}}{C_{\min}} = \frac{C_{\text{tot}}(V_{R\min})}{C_{\text{tot}}(V_{R\max})} \tag{2-68}$$

is limited by the fact that the diode must not be driven by the alternating voltage superimposed on the tuning voltage either into the forward mode or the breakdown mode. Otherwise, rectification, which would shift the diode's bias and considerably affect its figure of merit, would take place. Figure 2-48 plots the capacitance/voltage characteristics of a capacitance diode to clarify the relationship. The useful operating range lies between the voltages

$$V_{\min} > \hat{V} - V_F \quad \text{and} \quad V_{\max} < V_{(\text{BR})R} - \hat{V} \tag{2-69}$$

As has already been indicated, the exponent n for large-capacitance-ratio or tuning diodes used today for TV tuners is not constant but voltage-dependent and subject to manufacturing tolerances. This means that the capacitance/voltage characteristic of these diodes is likewise subject to manufacturing tolerances. Since, in a TV tuner, it is necessary for two or three circuits to be tuned uniformly, tuner diodes must be selected empirically for identical characteristics and supplied in equipment lots.

Electrical Properties of Tuning Diodes. In this section, the electrical properties of capacitance diodes are described, with reference to data published on the tuner diodes BB141 and BB142.

Equivalent Circuit. Since a capacitance diode is not an ideal capacitor, it is useful to introduce an equivalent circuit that can serve as a basis for discussing the diode's electrical properties. Figure 2-49 shows various types of equivalent circuits.

The complete equivalent circuit (Figure 2-49a), which conforms closely to physical conditions, comprises, in addition to the diode capacitance C_{tot}, the series resistance r_S and the series inductance L_S, and also the reverse resistance $R = dV_R/dI_R$. At higher frequencies, this resistance can usually be disregarded, so that the equivalent circuit is simplified in terms of Figure 2-49b—the configuration usually employed. In many cases, the series inductance can also be disregarded, in which case we obtain Figure 2-49c.

Capacitance. Figure 2-50 shows the capacitance/voltage characteristic borrowed from the datasheets of diodes BB141 and BB142. As this characteristic, like most characteristics published in datasheets, represents merely a typical curve, it would not convey sufficient

information for the user to dimension a tuner. Therefore, the datasheets contain additional data on capacitance and useful capacitance ratio:

Capacitance		BB141A (pF)	BB141B (pF)	BB142 (pF)
At $V_R = 1$ V	C_{tot}	16	19	17
At $V_R = 3$ V	C_{tot}	11	13	12
At $V_R = 25$ V	C_{tot}	2–2.35	2.25–2.65	2–3

$$\text{Useful capacitance ratio} = \frac{C_{tot}(3\text{ V})}{C_{tot}(25\text{ V})} = 4 \text{ to } 6$$

On the basis of the spread of the diode capacitance C_{tot} guaranteed for $V_R = 25$ V, it is possible to calculate the tuner with the aid of the spread guaranteed for the capacitance ratio.

Three further capacitance graphs are shown in Figures 2-51 through 2-53. Figure 2-51 shows a (typical) curve representing the normalized slope as a function of the reverse voltage, Figure 2-52 the normalized capacitance as a function of the junction temperature, and Figure 2-53 the temperature coefficient of the capacitance as a function of the reverse voltage.

The variation of diode capacitance with ambient temperature, as shown in Figures 2-52 and 2-53, is really only a function of the temperature dependence of the diffusion voltage [see Eq. (2-66)]. The diffusion voltage drops with rising temperatures by about 2 mV/°C , which means that the diode capacitance rises with temperature. The influence of the diffusion voltage, and thus the temperature coefficient of the capacitance, decreases as the reverse voltage rises. It is therefore advisable to run a capacitance diode with as high a reverse voltage as the required capacitance ratio permits. Compensation for the dependence of capacitance on temperature will be discussed later.

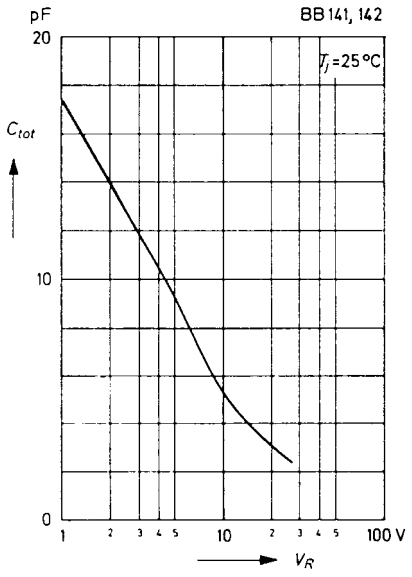


Figure 2-50 Capacitance/voltage characteristic of diodes BB141 and BB142.

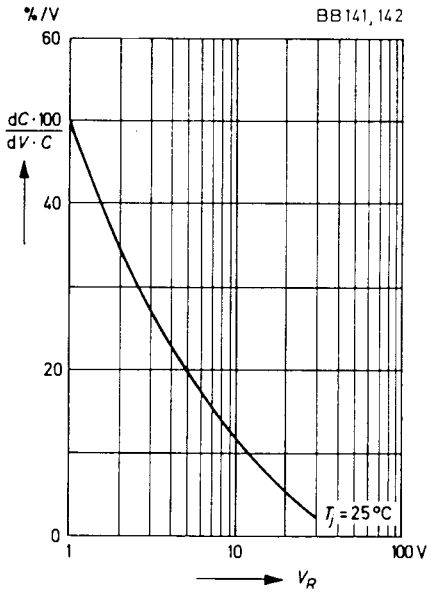


Figure 2-51 Slope (normalized) as a function of the reverse voltage.

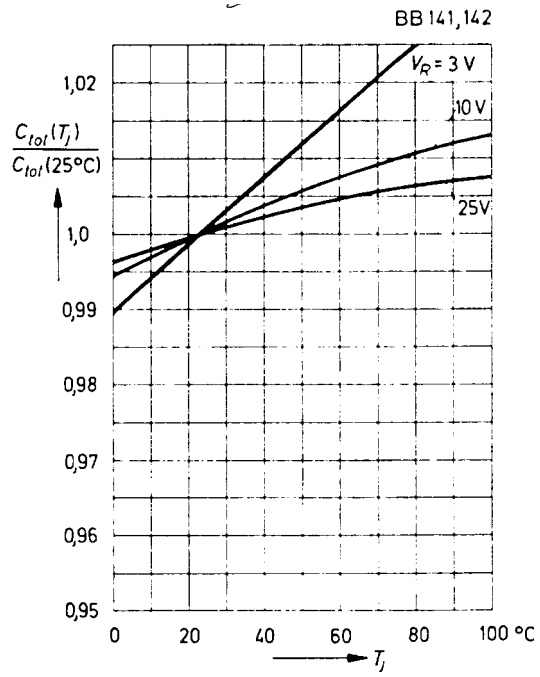


Figure 2-52 Capacitance (normalized) as a function of the junction temperature.

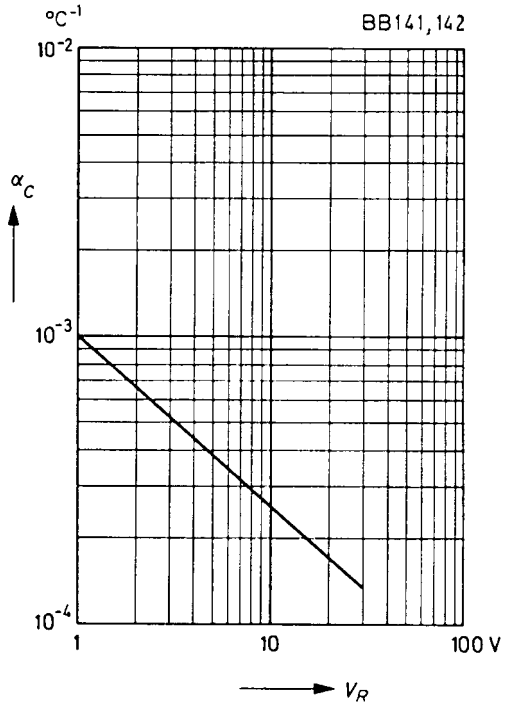


Figure 2-53 Temperature coefficient of capacitance as a function of the reverse voltage.

Series Resistance, Figure of Merit (Q). Since capacitance diodes are intended for employment in resonant circuits, indications are required as to the circuit attenuation that they produce. Usually, the series resistance or the figure of merit that can be calculated therefrom is indicated. With reference to the equivalent circuit shown in Figure 2-49a, the figure of merit (or Q factor) is calculated:

$$Q = \frac{1}{\omega C_{\text{tot}} r_S + 1/\omega C_{\text{tot}} R} \tag{2-70}$$

Figure 2-54 shows the theoretical (normalized) Q as a function of frequency. With the frequency range of 10–1000 MHz, which is essential for tuner diodes, the parallel resistance R caused by the reverse current (see Figure 2-49b) can be disregarded, and Eq. (2-70) can be simplified to read as follows:

$$Q = \frac{1}{\omega C_{\text{tot}} r_S} \tag{2-71}$$

In Figure 2-55, Q is plotted as a function of frequency for the tuner diodes BB141 and BB142. The datasheet indicates, in addition to this curve, the value of the series resistance at $f = 470$ MHz and $C_{\text{tot}} = 9$ pF:

	BB141A, BB141B	BB142
Series resistance (r_S)	0.6 (< 0.8)	0.9 (< 1.2)

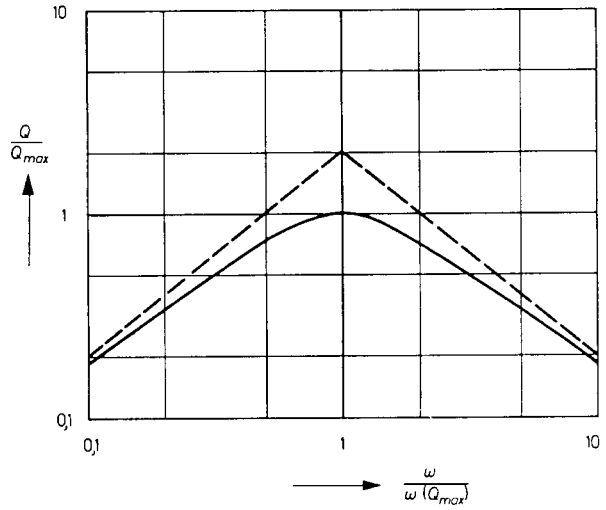


Figure 2-54 *Q* (normalized) as a function of frequency.

Since the diode capacitance and the series resistance decrease as the reverse voltage rises, Q is likewise dependent on the reverse voltage. It is smallest at a low reverse voltage, that is, at a low frequency. The series and parallel capacitances inevitable in a diode-tuned resonant circuit influence the effective Q of the tuning capacitance, as will be described later.

Series Inductance, Series Resonant Frequency, Cutoff Frequency for Q = 1. These three parameters define the behavior of the diode at high frequencies when the influence of

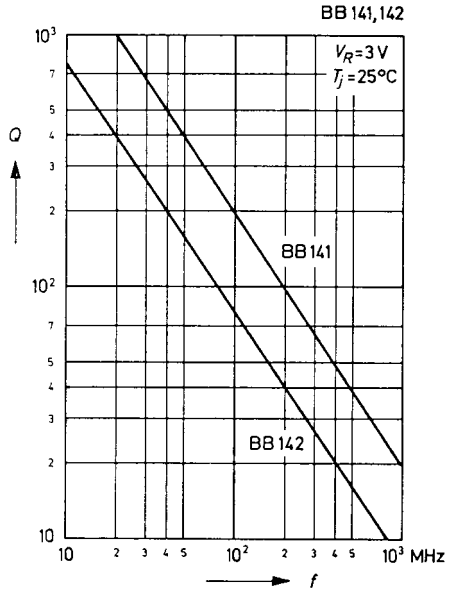


Figure 2-55 *Q* as a function of frequency.

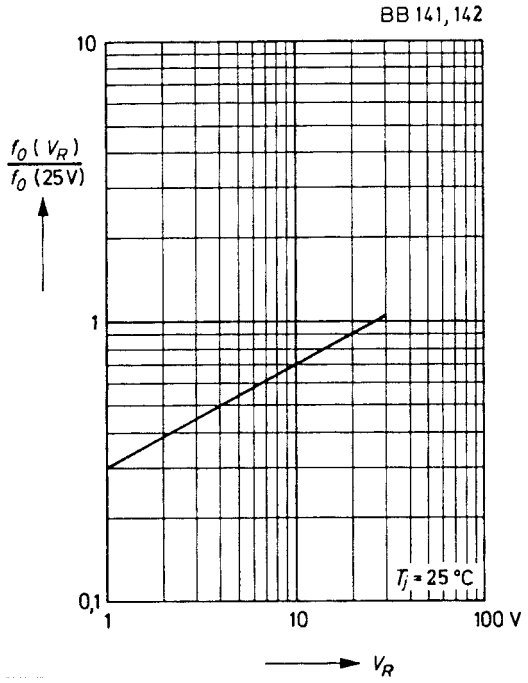


Figure 2-56 Series-resonant frequency (normalized) as a function of the reverse voltage.

the series inductance can no longer be neglected. The parameter on which the calculation is based is the series inductance L_S, which is determined by the connection leads. As indicated in Figure 2-49a, it is in series with the capacitance and, given a suitable frequency, forms a series-resonant circuit therewith. In that case

$$\omega L_S = \frac{1}{\omega C_{\text{tot}}} \tag{2-72}$$

and the series-resonant frequency is

$$f_0 = \frac{1}{2\pi\sqrt{C_{\text{tot}} L_S}} \tag{2-73}$$

Above this frequency, the impedance of the diode is inductive, but it depends on the reverse voltage. Figure 2-56 shows how the (normalized) series-resonant frequency depends on the reverse voltage for the tuner diodes BB141 and BB142.

The cutoff frequency for $Q = 1$ is defined as the frequency at which Q assumes the value of unity. In a rough calculation, the reactive impedance of the diode capacitance can be neglected and the frequency f_{Q1} obtained for the condition in which the inductive reactance equals the series resistance:

$$\omega L_S = r_S \tag{2-74}$$

$$f_{Q1} = \frac{r_S}{2\pi L_S} \tag{2-75}$$

For the tuner diodes BB141 and BB142, Figure 2-57 shows how the (normalized) cutoff frequency depends on the reverse voltage. The datasheets contain the following numerical information:

		BB141	BB142	
Series inductance measured 1.5 mm from the case	L_S	2.5	2.5	nH
Series-resonant frequency at $V_R = 25$ V	f_0	2	1.8	GHz
Cutoff frequency for $Q = 1$ at $V_R = 3$ V	f_{Q1}	24	16	GHz

Depending on the intended application, either the series-resonant frequency or the cutoff frequency for $Q = 1$ determines the maximum useful frequency range.

Leakage Current, Breakdown Voltage. These two parameters mark the dc performance of the diode when a reverse bias is applied to it. In the databook for the tuner diodes BB141 and BB142, the breakdown voltage is quoted as follows:

$$\text{Reverse breakdown voltage at } I_R = 100\ \mu\text{A}: \quad V_{(BR)R} > 30\ \text{V}$$

This means that the maximum reverse voltage that may be applied to the diode as a tuning voltage is 30 V.

The leakage current for the above-mentioned tuner diodes is guaranteed as follows:

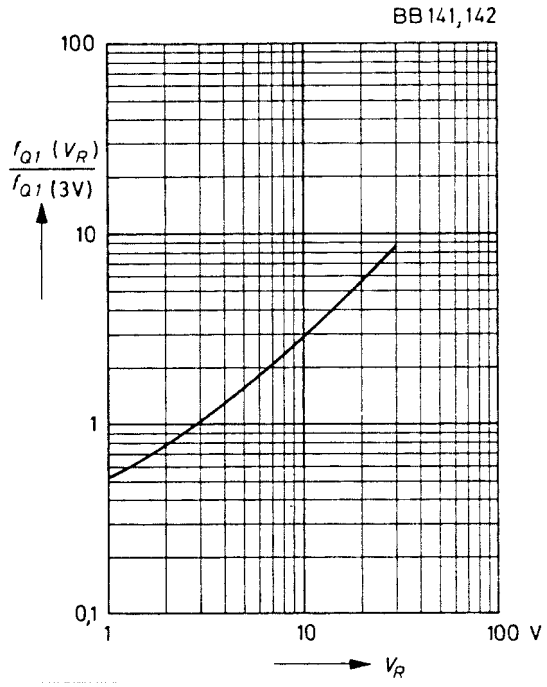


Figure 2-57 Cutoff frequency for $Q = 1$ (normalized) as a function of the reverse voltage.

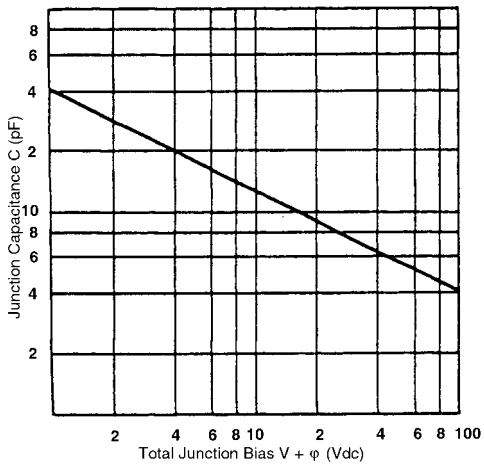


Figure 2-58 Leakage current as a function of reverse voltage.

$$\text{Leakage current at } V_R = 28 \text{ V:} \quad I_R < 50 \text{ nA}$$

Moreover, Figure 2-58 shows a graph that illustrates the leakage current as a function of the reverse voltage. Since this current is temperature dependent (as with every silicon diode)—it doubles with each temperature rise of about 10 °C—care should be taken when dimensioning the tuner circuit that the leakage current does not cause inadmissible voltage variations at increased ambient temperatures.

Matching of Tuner Diodes, Uniform Parameters. As has already been mentioned, the capacitance versus voltage characteristics of modern tuner diodes are subject to a certain scatter, so that it becomes necessary to test these diodes to obtain equipment lots empirically.

Diode-Tuned Resonant Circuits

Tuner Diode in the Parallel-Resonant Circuit. Figures 2-59, 2-60, and 2-61 illustrate three basic circuits for the tuning of parallel-resonant circuits by means of capacitance diodes.

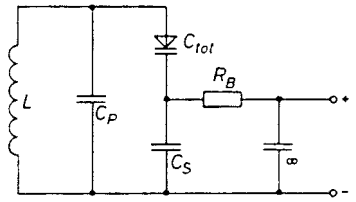


Figure 2-59 Parallel-resonant circuit with tuner diode, and bias resistor parallel to the series capacitor.

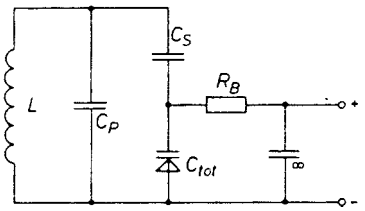


Figure 2-60 Parallel-resonant circuit with tuner diode, and bias resistor parallel to the diode.

In the circuit diagram of Figure 2-59, the tuning voltage is applied to the tuner diode via the input coil and the bias resistor R_B. Series-connected to the tuner diode is the series capacitor C_S, which completes the circuit for ac but isolates the cathode of the tuner diode from the coil and thus from the negative terminal of the tuning voltage. Moreover, a fixed parallel capacitance C_P is provided. The decoupling capacitor preceding the bias resistor is large enough for its value to be disregarded in the following discussion. Since for high-frequency purposes the biasing resistor is connected in parallel with the series capacitor, it is transformed into the circuit as an additional equivalent shunt resistance R_C. We have the equation

$$R_C = R_B\left(1 + \frac{C_S}{C_{\text{tot}}}\right)^2 \tag{2-76}$$

If in this equation the diode capacitance is substituted by the resonant circuit frequency ω, we obtain

$$R_C = R_B\left(\frac{\omega^2 LC_S}{1 - \omega^2 LC_P}\right)^2 \tag{2-77}$$

The resistive loss R_C caused by the bias resistor R_B is seen to be highly frequency dependent, and this may result in the bandwidth of the tuned circuit being independent of frequency if the capacitance of the series capacitor C_S is not chosen sufficiently high.

Figure 2-60 shows that the tuning voltage can also be applied directly and in parallel to the tuner diode. For the parallel loss resistance transformed into the circuit, we have the expression

$$R_C = R_B\left(\frac{1 + C_{\text{tot}}}{C_S}\right)^2 \tag{2-78}$$

and

$$R_C = R_B\left(\frac{\omega^2 LC_S}{\omega^2 L(C_S + C_P) - 1}\right)^2 \tag{2-79}$$

The influence of the bias resistor R_B in this case is larger than in the circuit of Figure 2-60, provided that

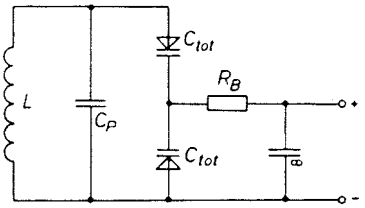


Figure 2-61 Parallel-resonant circuit with two tuner diodes.

$$C_S^2 > C_S(C_{\text{tot}} + C_P) + C_{\text{tot}}C_P \tag{2-80}$$

This is usually the case because the largest possible capacitance will be preferred for the series capacitor C_S, and the smallest for the shunt capacitance C_P. The circuit of Figure 2-59 is therefore normally preferred to that of Figure 2-60. An exception would be the case in which the resonant circuit is meant to be additionally damped by means of the bias resistor at higher frequencies.

In the circuit of Figure 2-61, the resonant circuit is tuned by two tuner diodes that are connected in parallel via the coil for tuning purpose, but series-connected in opposition for high-frequency signals. This arrangement has the advantage that the capacitance shift caused by the ac modulation (see section entitled "Modifying the Diode Capacitance by the Applied ac Voltage") takes effect in opposite directions in these diodes and therefore cancels itself. The bias resistor R_B, which applies the tuning voltage to the tuner diodes, is transformed into the circuit at a constant ratio throughout the whole tuning range. Given two identical, loss-free tuner diodes, we obtain the expression

$$R_C = 4R_B \tag{2-81}$$

Capacitances Connected in Parallel or Series with the Tuner Diode. Figures 2-59 and 2-60 show that a capacitor is usually in series with the tuner diode, in order to close the circuit for alternating current and, at the same time, to isolate one terminal of the tuner diode from the rest of the circuit with respect to direct current, so as to enable the tuning voltage to be applied to the diode. As far as possible, the value of the series capacitor C_S will be chosen such that the effective capacitance variation is not restricted. However, in some cases, as, for example, in the oscillator circuit of receivers whose intermediate frequency is of the order of magnitude of the reception frequency, this is not possible and the influence of the series capacitance will then have to be taken into account. By connecting the capacitor C_S, assumed to be lossless, in series with the diode capacitance C_{tot}, the tuning capacitance is reduced to the value

$$C^* = C_{\text{tot}} \frac{1}{1 + C_{\text{tot}}/C_S} \tag{2-82}$$

The Q of the effective tuning capacitance, taking into account the Q of the tuner diode, increases to

$$Q^* = Q\left(1 + \frac{C_{\text{tot}}}{C_S}\right) \tag{2-83}$$

The useful capacitance ratio is reduced to the value

$$\frac{C^*_{\max}}{C^*_{\min}} = \frac{C_{\max}}{C_{\min}} \frac{1 + C_{\min}/C_S}{1 + C_{\max}/C_S} \tag{2-84}$$

where $C_{\max}$ and $C_{\min}$ are the maximum and minimum capacitances of the tuner diode.

On the other hand, the advantage is gained that, due to capacitive potential division, the amplitude of the alternating voltage applied to the tuning diode is reduced to

$$\hat{v}^* = \hat{v} \frac{1}{1 + C_{\text{tot}}/C} \tag{2-85}$$

so that the lower value of the tuning voltage can be smaller, and this results in a higher maximum capacitance $C_{\max}$ of the tuner diode and a higher useful capacitance ratio. The influence exerted by the series capacitor, then, can actually be kept lower than Eq. (2-83) would suggest.

The parallel capacitance C_P that appears in Figures 2-59 to 2-61 is always present, since wiring capacitances are inevitable and every coil has its self-capacitance. By treating the capacitance C_P, assumed to be lossless, as a shunt capacitance, the total tuning capacitance rises value and, if C_S is assumed to be large enough to be disregarded, we obtain

$$C^* = C_{\text{tot}} \left(1 + \frac{C_P}{C_{\text{tot}}} \right) \tag{2-86}$$

The Q of the effective tuning capacitance, as derived from the Q of the tuner diode, is

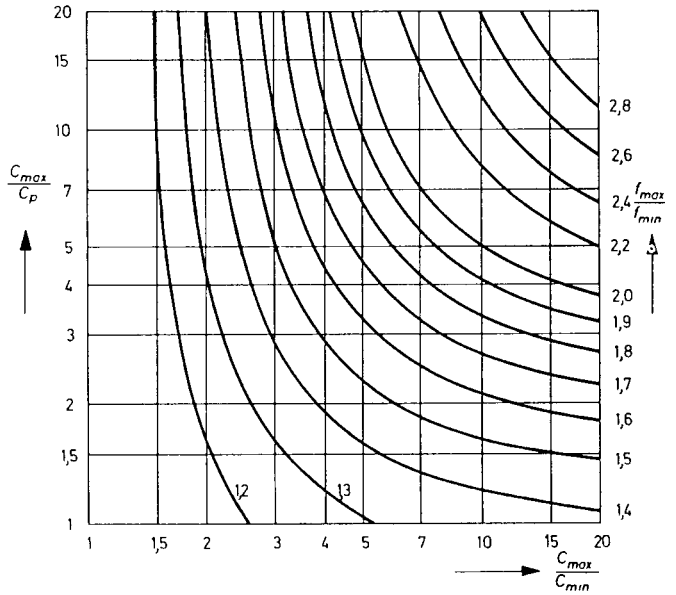


Figure 2-62 Diagram for determining the capacitance ratio and minimum capacitance.

$$Q^* = Q\left(1 + \frac{C_P}{C_{\text{tot}}}\right) \tag{2-87}$$

or, in other words, it rises with the magnitude of the parallel capacitance. The useful capacitance ratio is reduced:

$$\frac{C^*_{\text{max}}}{C^*_{\text{min}}} = \frac{C_{\text{max}}}{C_{\text{min}}} \frac{1 + C_P/C_{\text{max}}}{1 + C_P/C_{\text{max}}} \tag{2-88}$$

In view of the fact that even a comparatively small shunt capacitance reduces the capacitance ratio considerably, it is necessary to ensure low wiring and coil capacitances in the circuit-design stage.

TUNING RANGE. The frequency range over which a parallel-resonant circuit according to Figure 2-59 can be tuned by means of the tuner diode depends on the useful capacitance ratio of the diode and on the parallel and series capacitances present in the circuit.

The ratio can be found from

$$\frac{f_{\text{max}}}{f_{\text{min}}} = \sqrt{1 + \frac{\dfrac{C_{\text{max}}}{C_P\,(1 + C_{\text{max}}/C_S)}}{1 + \dfrac{C_{\text{max}}}{C_P\,(C_{\text{max}}/C_{\text{min}} + C_{\text{max}}/C_S)}}} \tag{2-89}$$

In many cases, the series capacitor can be chosen large enough for its effect to be negligible. In that case, Eq. (2-89) is simplified as follows:

$$\frac{f_{\text{max}}}{f_{\text{min}}} = \sqrt{\frac{1 + C_{\text{max}}/C_P}{1 + C_{\text{min}}/C_P}} \tag{2-89a}$$

From this equation, the diagram shown in Figure 2-62 is computed. With the aid of this diagram, the tuner diode parameters required for tuning a resonant circuit over a stipulated frequency range—that is, the maximum capacitance and the capacitance ratio—can be determined.

Whenever the series capacitance C_S cannot be disregarded, the effective capacitance ratio is reduced according to Eq. (2-84).

TRACKING. Some applications require the maintenance of a fixed frequency relationship between two or more tuned circuits as their tuning is simultaneously adjusted. Referred to as *tracking*, this technique requires narrow tolerances of capacitance versus tuning voltage. Minimizing tracking error requires special care if the tracking circuits must cover the same frequency span beginning at different start and end frequencies, as is necessary when simultaneously tuning oscillator and mixer/RF circuitry in a superheterodyne receiver. Then, tracking error must be minimized by means of series and shunt capacitances in accordance with methods known from variable capacitors. The frequency deviations that must be anticipated are summarized in the equation

$$\frac{df}{f} = -\frac{1}{2}\frac{dC_0}{C_0} - \frac{1}{2}\frac{d(L-L_0)}{L} - \frac{1}{2}\frac{dL_0}{L_0} + \frac{n}{2}\frac{dV_R}{V_R + V_D} \tag{2-90}$$

The spread of parameters dC_0/C_0 and dL_0/L_0 can only be compensated by varying the circuit inductances $d(L–L_0)/L$ or by varying the bias $dV_R/(V_R + V_D)$.

Modulating the Diode Capacitance by the Applied ac Voltage. In normal operation, the sum of the tuning voltage and the alternating signal voltage of the resonant circuits is applied to the tuner diode. The bias, and thus the capacitance, of the tuner diode therefore varies at the rhythm of the alternating voltage. Due the nonlinear character of the capacitance versus voltage curve, voltage distortions and capacitance shifts are inevitable, and these must be kept within adequate limits. This is done by maintaining the ac applied to the diode(s) at sufficiently low ac amplitude and by choosing an adequate minimum value for the tuning voltage. In the resonant circuit, a tuner diode is modulated predominantly by a current free from harmonics, according to the equation

$$i = \hat{i}\cos \omega t \tag{2-91}$$

The alternating voltage across the diode is

$$v = (V_R + V_D)\left[\left(1 + \frac{\hat{i}(1-n)}{\omega C_{\text{tot}} V_R}\sin \omega t\right)^{1/(1-n)} - 1\right] \tag{2-92}$$

An evaluation of this equation shows that especially the first harmonic makes its appearance. The capacitance shift caused by the alternating voltage superimposed on the tuning voltage is shown in Figure 2-63. However, the voltage distortion, and thus the capacitance shift, can be largely avoided if two tuner diodes are used, as in Figure 2-61.

COMPENSATING TEMPERATURE DEPENDENCE. As has already been mentioned, variations are caused in the capacitance of the diode mainly by the dependence of the diffusion voltage on

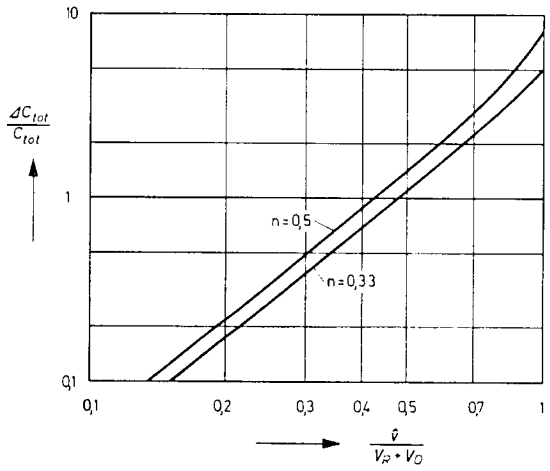


Figure 2-63 Capacitance increase as a function of the ac voltage drop across the tuner diode.

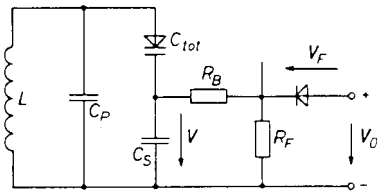


Figure 2-64 Temperature-compensation circuit with diode.

temperature. In a diode-tuned resonant circuit, the temperature coefficient of the resonant frequency, therefore, depends on the tuning voltage and thus on the resonant frequency. It is therefore impossible to compensate the temperature dependence of the resonant frequency by means of temperature-dependent capacitors—the method usually adopted for mechanically tuned circuits.

To achieve satisfactory compensation, the tuning voltage should be increased by an amount equal to that by which the diffusion voltage of the diode is reduced with rising temperature; that is, by approximately 2 mV/°C. This can be done easily by connecting a forward-biased silicon diode in series with the source of the tuning voltage, as shown in Figure 2-64. Since the forward voltage of the diode varies by –2 mV/°C with ambient temperature, the voltage

$$V = V_0 + V_D - V_F$$

which determines the bias of the tuner diode, and therefore also the diode capacitance, is almost temperature independent. In practice, it is advisable to feed this diode with additional current via the resistor R_F to keep its differential forward resistance sufficiently low and thus to prevent capacitive noise voltages from leaking into the tuning circuit.

To prevent the load placed by the resistor R_F on the source of the tuning voltage from fluctuating in a manner depending on the tuning voltage, a circuit such as shown in Figure 2-65 may be used in which the emitter diode of the transistor effects the compensation, and the only load placed on the source of the tuning voltage is the base current of this transistor.

Dynamic Stability. The resonant frequency of a diode-tuned RF circuit follows the equation

$$f = f_0 \sqrt{\left(\frac{V_R + V_D}{V_{R0} + V_D}\right)^n} \tag{2-93}$$

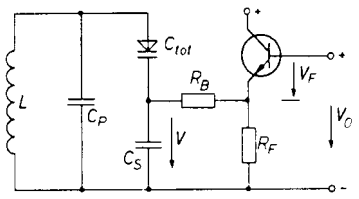


Figure 2-65 Temperature-compensation circuit with transistor.

The frequency f_0 is set with the aid of the circuit inductance when the tuning voltage equals V_{R0}. Since the signal amplitude in the circuit depends on the field strength of the received transmitter and since the resonant frequency of the circuit depends on the amplitude, undesirable feedback between the tuning voltage and the received alternating voltage may occur at the tuner diode. A disturbance that has a similar effect is the microphonic effect, which can sometimes be observed with mechanically variable capacitors. In an investigation of the feedback in diode-tuned resonant circuits, the following assumptions were made:

- The impedance of the antenna (or of the signal generator) is negligible.
- The full alternating voltage is applied to the diode (whereas, in practice, the series capacitor brings about a potential division).
- The capacitance versus voltage characteristic is an exponential function.
- The Q of this circuit is independent of the amplitude of the alternating voltage.

As a result of a high alternating voltage at the resonant circuit, three undesirable effects—harmonic generation, frequency shift, and cross-modulation—are encountered.

HARMONIC GENERATION. The nonlinearity of the capacitance versus voltage characteristic produces harmonics. However, the selectivity of the receiver reduces the effects of these harmonics to noise level.

FREQUENCY SHIFT. When a sinusoidal voltage v is applied to the tuner diode of the RF circuit, the variation of the diode capacitance does not follow the sine law, but the equation

$$\frac{C_\omega}{C} = \left(1 + \frac{v}{V_R + V_D} \sin \omega t\right)^{-n} \tag{2-94}$$

This results in a change of the resonant frequency of the circuit. The inductive slope of the resonant curve steepens and may even turn back. This could lead to bistable behavior of the resonant circuit. The change of frequency takes place as if the tuning voltage, and thus the resonant frequency, has decreased. The order of magnitude of this change of resonant frequency was determined with the aid of a computer from Fourier analysis of

$$\frac{\Delta f}{f} = 1 + \frac{1}{2\pi} \int_0^{2\pi} \left(1 + \frac{v}{V_R + V_D} \sin \omega t\right)^{n/2} d(\omega t) \tag{2-95}$$

Figure 2-66 shows the admissible alternating voltage at the diode, computed for a 2% frequency shift, as a function of the tuning voltage. The measured values differ by not more than 10% from the computed values. Since the Q of the circuits is about 50 for the VHF as well as the UHF range, a 2% frequency shift is permissible.

CROSS-MODULATION. Cross-modulation is a disturbance caused by nonlinearities of the characteristics of the components concerned. The modulation of an undesired amplitude-modulated signal is here transferred to the carrier of the amplitude-modulated intelligence signal. Cross-modulation is virtually independent of the amplitude of the intelligence signal.

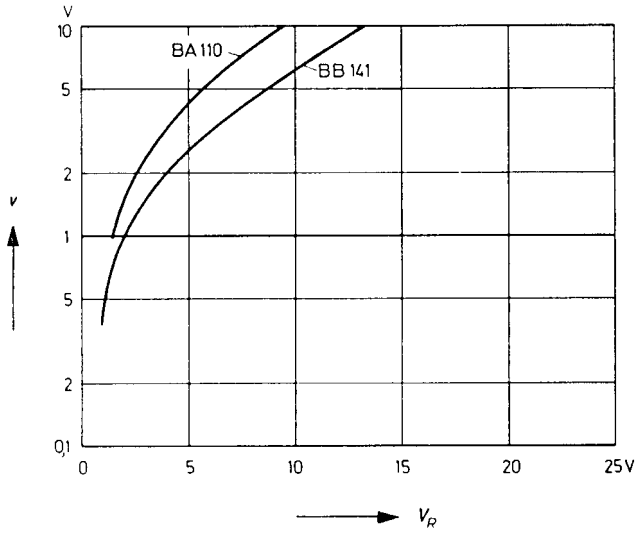


Figure 2-66 Admissible alternating voltage for a 2% frequency shift as a function of tuning voltage.

The cross-modulation factor for tuner diodes is

$$k = \frac{v^2}{(V_R + V_D)^2} m_v \frac{n(n+1)}{2} \tag{2-96}$$

Resolving this equation in terms of v, we obtain

$$v = \frac{V_R + V_D}{n} \sqrt{\frac{2kn}{m_v (n+1)}} \tag{2-97}$$

The usual definition of the cross-modulation percentage for measurement purposes is as follows: A 100% modulated noise signal ($m_v = 1$) of amplitude v is received simultaneously with the intelligence signal in the diode-tuned RF circuit. If a 1% modulation of the signal carrier by the noise signal takes place, this amounts to a cross-modulation factor $k = 0.01$. We thus obtain for tuner diodes

$$V = 0.14 \frac{V_R + V_D}{n} \sqrt{\frac{n}{n+1}} \tag{2-98}$$

Figure 2-67 gives the graphs for this equation. In practical operation, FM and VHF/UHF tuners equipped with the tuner diodes BB121 or BB141 exhibit merely the cross-modulation caused by the BJTs or FETs employed. For bipolar transistors, $v \approx 15$–100 mV; for FETs, 100–300 mV.

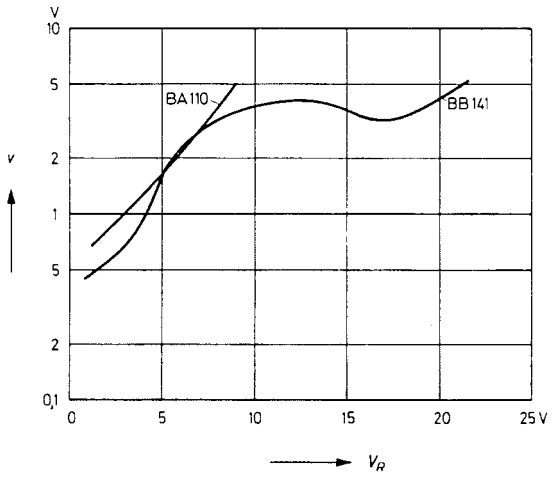


Figure 2-67 Graphical representation of Eq. (2-98).

Static Stability. To ensure trouble-free operation, the static parameters of the tuner diodes, especially the leakage current, must be taken into account. When differentiating the expression for the tuning voltage

$$V_R = V_B - I_R R_B \tag{2-99}$$

we obtain for a known leakage current I_R the tuning voltage variation

$$dV_R = dV_B - R_B I_R \left(\frac{dR_B}{R_B} + \frac{dI_R}{I_R} + dV_B \right) \tag{2-100}$$

This leads to a change in the resonant frequency

$$\frac{df}{f} = \frac{n}{2(V_R + V_D)} \left[dV_D - R_B I_R \left(\frac{dR_B}{R_B} + \frac{dI_R}{I_R} \right) + dV_B \right] \tag{2-101}$$

The resonant frequency is more stable the smaller the exponent n and the higher the tuning voltage V_R. The influence of the temperature dependence of the diffusion voltage V_D and its compensation have already been discussed. If the series bias resistance R_B is not too high, the effect of its variation ($dR_B/R_{B)}$ and the effect of the temperature dependence of the leakage current—doubling for every 10 °C temperature rise—may be neglected.

Generating the Tuning Voltage. The supply voltage V_B in Eq. (2-101) exerts an influence in many respects. The temperature coefficient of the voltage source—usually, for economic reasons, only a Zener diode is provided for stabilization—affects conditions not only with ambient temperature variations, but also when the Zener diode heats up after the set is switched on, or when its operating current varies. Because external temperature compensation of the operating voltage of the Zener diode cannot cover all detrimental effects and, moreover, increases the series bias resistance R_B, as well as any variations the latter may

undergo, the use of a temperature-compensated Zener diode, such as the 1N4065A (nominal Zener voltage, 33 V; temperature coefficient, 0.002 °C/W), is recommended.

Diode Switches. The diode switches described here differ somewhat from the switching diodes used in computer and pulse technology. Whereas in the case of switching diodes the signal itself triggers the switching operation—current does or does not pass through the switching diode depending on the signal level—diode switches allow an alternating current to be switched on or off by means of a direct voltage or a direct current. Figure 2-68 shows an example of the use of diode switches.

Diode Switch Technology. Special manufacturing techniques are needed to attain the low junction capacitance, low differential forward resistance r_F, and the low-inductance structure required for the special applications envisaged. Two conflicting requirements have to be met: On the one hand, a high-resistivity basic material should be used and the diode area should be kept as small as possible to minimize the junction capacitance for a given reverse voltage; on the other hand, the bulk resistance should be kept low. The size of the junction area, therefore, can represent only a compromise between these two requirements. To obtain a low bulk resistance R_B despite the use of high-resistivity silicon around the junction, the diodes are manufactured by the epitaxial process, characterized by a thin high-resistivity layer on a low-resistivity substrate. If the forward current I_F is sufficiently high, then the specific resistance of the epitaxial layer does not affect the bulk resistance, because under these conditions the resistivity of this layer is considerably reduced by being flooded with carriers.

The junction impedance Z_j is somewhat more critical (Figure 2-69). It is composed of the junction capacitance C_S, which does not depend on current, the current-dependent diffusion capacitance

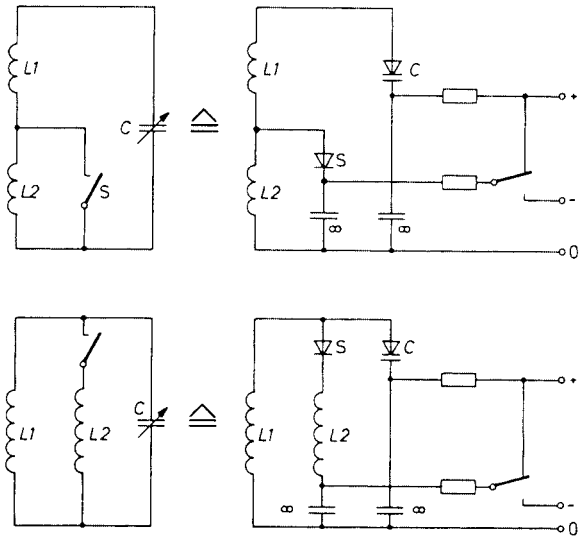


Figure 2-68 Comparison of mechanically and electronically tuned and switched resonant circuits.

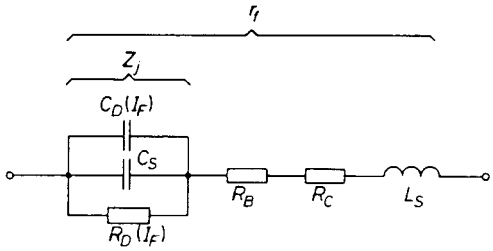


Figure 2-69 Equivalent circuit of diode switches in the frequency range from 1 MHz to 1 GHz.

$$C_D = \frac{\mathfrak{t}_p I_F}{V_T} \tag{2-102}$$

and the diffusion resistance

$$R_D = \frac{V_T}{I_F} \tag{2-103}$$

where $\mathfrak{t}_p$ is the lifetime of the minority charge carriers and

$$V_T = \frac{kT}{q} = 26 \text{ mV} \tag{2-104}$$

is the voltage equivalent of thermal energy. The resistive component of the impedance Z_j is calculated as follows:

$$\text{Re}(Z_j) = \frac{R_D}{1 + \omega^2 R_D^2 (C_S + C_D)^2} \tag{2-105}$$

As can be seen from the following approximations, this resistance becomes rather dependent on current at higher frequencies:

$$\text{For } I_F \to 0 \quad \text{Re}(Z_j) \approx \frac{I_F}{\omega^2 C_S^2 V_T} \tag{2-106}$$

$$\text{For } I_F \to \infty \quad \text{Re}(Z_j) \approx \frac{I_F}{\omega t_p I_F} \tag{2-107}$$

In Figure 2-70 the differential forward resistance is plotted as a function of the forward current. The only parameter by which the differential forward resistance can be reduced for a given current and a given frequency according to Eq. (2-107) is the carrier lifetime. Steps will have to be taken, therefore, to make the carrier lifetime as long as possible.

Because diode switches are designed for applications up to the gigahertz range, a geometry resulting in a low inherent capacitance and inductance had to be used. Furthermore, the device had to be assembled in such a way that virtually no additional contact resistance R_C was introduced. For this reason, a double-plug construction (hard-glass pressure contacts)

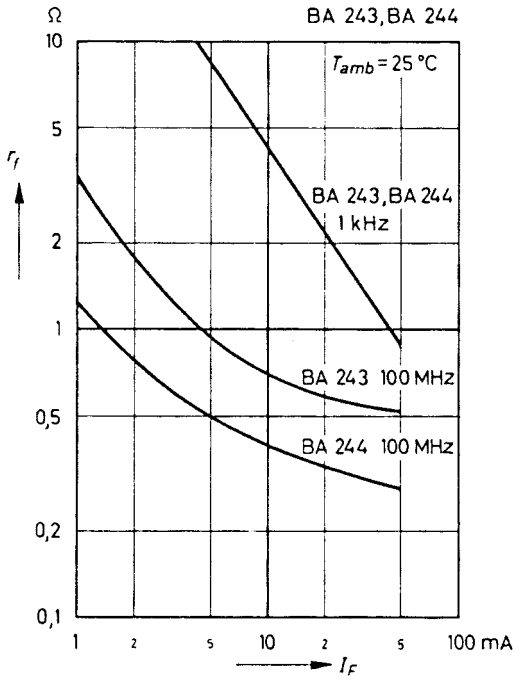


Figure 2-70 Differential forward resistance as a function of forward current.

was chosen to avoid the S- or U-bends customary with other encapsulations. The system is pressure-clamped between two stud-shaped enlargements of the connecting leads, the necessary high pressure being produced as the glass shrinks during cooling. This construction results in an encapsulation of short length (only 4 mm) and low series inductance, the advantages of which can be fully exploited because soldered connections to the leads can be made directly at the glass body.

Diode Switch Data. The silicon planar diode switches BA243 and BA244 are designed for electronic band selection in tuners in the frequency range from 10 to 1000 MHz. The differential forward resistance remains constant and is very low over a wide frequency and current range. The diode capacitance is likewise small and independent of voltage through a wide range. The devices are characterized by low-inductance geometry. The BA243 is intended for the VHF range, and the BA244 for the UHF range. See Table 2-11.

Resonant Circuits Incorporating Diode Switches. Figure 2-71 illustrates the function of the diode switch with reference to a parallel-resonant circuit whose resonant frequency can be switched by short-circuiting part of the circuit inductance. Such a switching procedure may be necessary for a VHF tuner, for example, to switch from band I to band III, because the frequencies of these bands are too far apart.

Capacitor C_K represents the capacitance of the resonant circuit. The symbol ∞ at capacitor C is meant to indicate that the value of this capacitor should be chosen to be much higher than the circuit capacitance. C completes the ac circuit when the inductor L_2 is shorted and

Table 2-11 BA243/244 specifications

Maximum Ratings			
Reverse voltage	V_R	20	V
Forward current at $T_A = 25$ °C	I_F	100	mA
Junction temperature	T_j	150	°C
Storage temperature range	T_s	–55 to +150	°C
Characteristics at $T_A = 25$ °C			
Forward voltage at $I_F = 100$ mA	V_F	<1	V
Leakage current at $V_R = 15$ V	I_R	< 100	nA
Dynamic forward resistance at $f = 50$–1000 MHz, $I_F = 10$ mA			
BA243	r_F	0.7 (< 1)	Ω
BA244	r_F	0.4 (< 0.5)	Ω
Relative variation of dynamic forward resistance with the variation of forward current in the range of $I_F + 2$–40 mA	$\frac{\Delta r_F 100}{r_F \Delta I_F}$	5	%/mA
Capacitance at $V_R = 15$ V, $f = 1$ MHz	C_{tot}	1.3 (< 2)	pF
Relative variation of capacitance with the variation of reverse voltage in the range of $V_R = 7$–20 V, $f = 100$ MHz	$\frac{\Delta C_{tot} 100}{C_{tot} \Delta V_R}$	1	%/V
Series inductance across case	L_S	2.5	nH

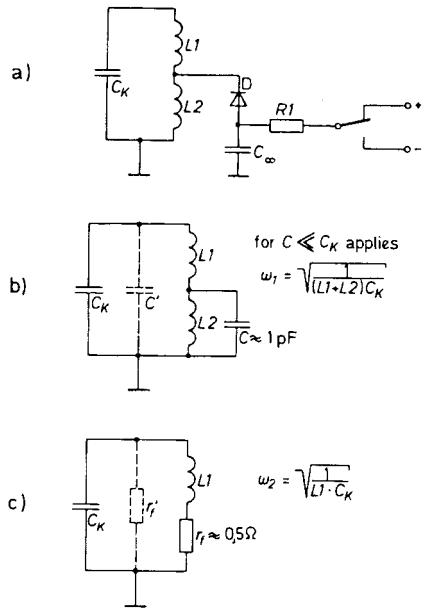


Figure 2-71 (a) Diode switch as bandswitch, shorting a partial inductance; (b) equivalent circuit with blocked diode switch; (c) equivalent circuit with conducting diode switch.

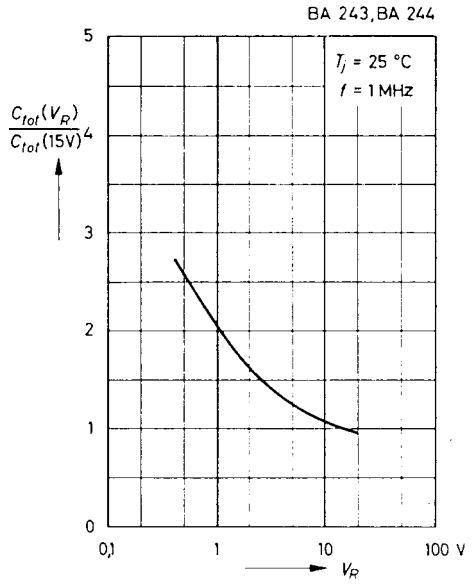


Figure 2-72 Capacitance as a function of the reverse voltage.

enables the switching or reverse voltage to be applied at a point of virtually zero RF potential. The resistor R_1 limits the diode current in the forward direction.

At a reverse voltage of $V_R = 15$ V, for example, the diode switch represents a 1-pF capacitor, as shown in Figure 2-71b. This represents parasitic capacitance C' in the circuit, which reduces the effective capacitance variation of the resonant circuit in the case of capacitive tuning. The capacitance C' is small enough to be negligible in most cases relative to the wiring and winding capacitances. Figure 2-72 shows the relative capacitance of the blocked-diode switches, as a function of the reverse voltage.

At a forward current of, for example, 10 mA, the diode has a differential resistance of approximately 0.5 Ω (see Figure 2-70). The effect of this resistance r_F on the resonant circuit is illustrated in Figure 2-71c. Conversion of a series loss resistance r_F into a parallel loss resistance r'_F is based on the following considerations: The equivalent parallel resistance r'_F affects the tuned circuit in the same way as the series resistance r_F; that is, the Q of the circuit should be the same in both cases. If one assumes that $\omega L >> r_F$, the equivalent of parallel resistance can be expressed as follows:

$$Q = \frac{\omega L_1}{r_F} = \frac{r'_F}{\omega L_1} \tag{2-108}$$

$$r'_f = \frac{(\omega L_1)^2}{r_F} \tag{2-109}$$

An evaluation of this equation shows that a differential resistance $r_F \approx 0.5$ Ω can be tolerated in most applications as regards amplification and selectivity.

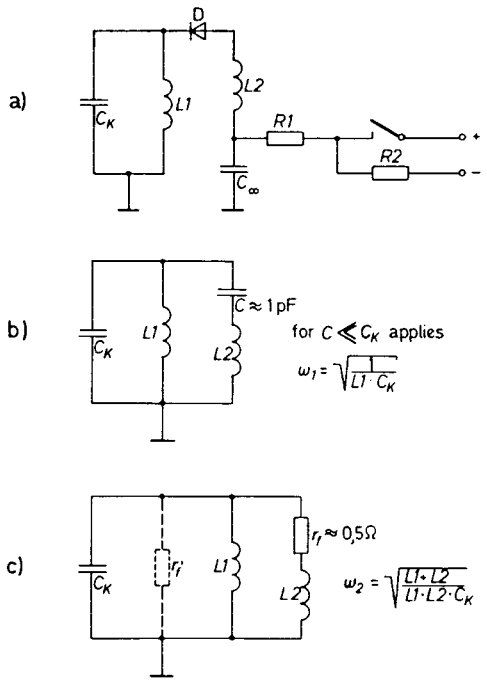


Figure 2-73 (a) Diode switch as a bandswitch, with an inductance in parallel; (b) equivalent circuit with blocked-diode switch; (c) equivalent circuit with conducting diode switch.

Instead of altering the resonant frequency of a tuned circuit by short-circuiting part of its inductance, the same result can be obtained by connecting a second inductance in parallel. The use of the diode switch for this purpose is illustrated in Figure 2-73. This differs from the one shown in Figure 2-71 in that the circuit inductance consists of either L_1 or of L_1 and L_2 in parallel. By applying a permanent negative bias to the diode via a high-value resistor R_2, it is possible to dispense with the use of one switch contact. Figure 2-73b shows the equivalent circuit for the diode under reverse-bias conditions (corresponding to a low resonant frequency), and Figure 2-73c shows the equivalent circuit under forward-bias conditions (corresponding to a higher resonant frequency).

In radio and television receivers, it usually is possible to derive the necessary reverse bias for the diode by rectification of the local oscillator signal, since the power required for blocking the diodes is extremely small. Note that under reverse-bias conditions the inductance L_2 and the capacitance C form a series-tuned acceptor circuit. However, this capacitance being extremely small, any undesirable effects due to its presence can normally be avoided. The same diode switching arrangements described for single-tuned circuits can also be employed in multisection bandpass and broadband filters.

In the following paragraph, a practical example for the use of diode switches will be given, and it will be seen that it is possible, in principle, to replace all of the mechanical RF switch contacts normally employed in receivers by diode switches. In cases where application of the diode dc potentials to a "cold" point in the circuit is not feasible, it may be necessary to apply them via *LC* or *RC* isolating networks.

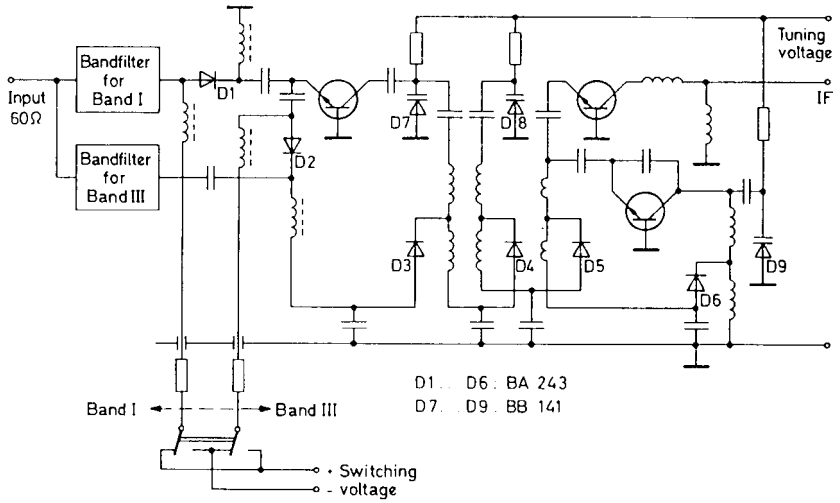


Figure 2-74 Schematic of a VHF television tuner with electronic band selection.

Use of the Diode Switch in a Television Receiver. Figure 2-74 shows the schematic of an electronically tuned and switched band I (48–62 MHz)/band III (175–224 MHz) TV tuner that requires switching of the input filter, bandpass filter, and oscillator circuit. These switching functions are performed by BA243 diode switches. As is customary in most European countries, in this example a broadband input filter between the antenna and the first RF transistor is used rather than a selective front-end circuit. In view of the need to correct noise and power matching over a relatively large frequency range, a switchable filter is employed in most tuners, switching being effected in this case by means of diodes D_1 and D_2.

Diodes D_3–D_6 are used for band-filter switching, for switching the coupling to the mixer transistor, and for switching the oscillator frequency. To minimize the bias-current requirements (each diode requires a bias current of approximately 10 mA), all diodes that are to be activated at the same time are connected in series. This arrangement requires only a few additional RF chokes, which are easily provided because of the high frequencies involved. The direct current required for the diode switches could be obtained by simple rectification, for example, from the heater circuit of the television receiver CRT, and the diode reverse bias could be obtained by rectification of the local oscillator signal, in which case the minute additional power requirement would have to be allowed for in the oscillator design.

A technically sophisticated design for a television tuner takes the form of a combined VHF/UHF unit incorporating electronically switched circuits in which several elements would be common. Such an all-band tuner can be assembled from only three transistors and a combined VHF/UHF band filter in $\lambda/2$ or $\lambda/4$ technology for UHF reception. The VHF circuits are short-circuited when the UHF position is selected.

Multistandard TV receivers require a large number of switch contacts in the tuner, the IF amplifier, the video amplifier, and the time base. The use of diode switches is particularly advantageous in this situation because it renders the physical layout of the equipment virtually independent of the location of the controls and obviates the need for long RF cable runs, which may be prone to interference.

2-2 BIPOLAR TRANSISTORS

2-2-1 Transistor Structure Types

All current RF silicon transistors are of the bipolar *npn* planar epitaxial type. Briefly, the significance of each of these terms is as follows.

Bipolar. In its broadest sense, it is the basic structure shown schematically in Figure 2-75, that is, the familiar three-semiconductor-region structure. *Bipolar* specifically means that the charge carriers of both negative (electron) and positive (hole) polarities are involved in the transistor action. In way of contrast, unipolar types include the junction-gate and insulated-gate field-effect transistors, which are basically one- or two-semiconductor-region structures in which carriers of a single polarity dominate.

npn. This abbreviation is for *negative–positive–negative* and identifies the regions of the structure as to polarity of the dominant or majority carrier in each region. The other polarity type is *pnp*. (See Figure 2-75.)

Silicon. Silicon (Si) is one of two elements from the fourth column of the periodic table that are in widespread use for transistor fabrication [the other is germanium (Ge)]. Other materials used include the compound gallium arsenide (GaAs). Although silicon is in predominant use because it results in the most favorable compromise among high-frequency, high-temperature, high-reliability, and ease-of-use attributes of the usable semiconductor material, silicon–germanium (SiGe) technology is increasingly important because it affords high-frequency performance unachievable with silicon alone.

Planar. The term planar denotes that both emitter–base and base–collector junctions of the transistor intersect the device surface in a common plane (hence, a better term might be *coplanar*). However, the real significance of the so-called planar structure is that the technique of diffusing dopants through an oxide mask, used in fabricating such a structure, results in junctions being formed beneath a protective oxide layer. These protected junctions are less prone to the surface problems sometimes associated with other types of structures, such as the mesa.

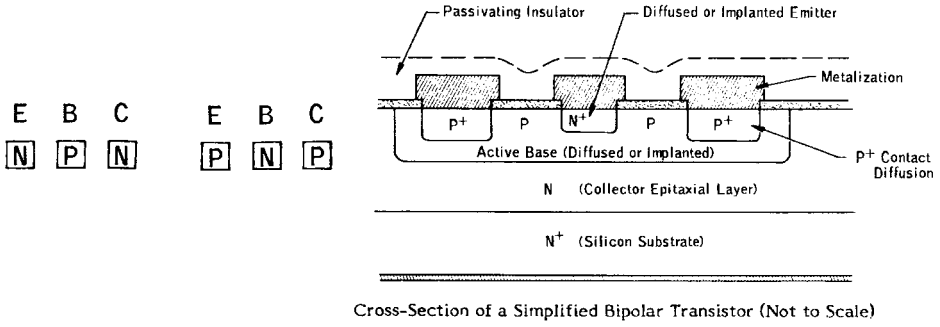


Figure 2-75 Transistor structure schematic.

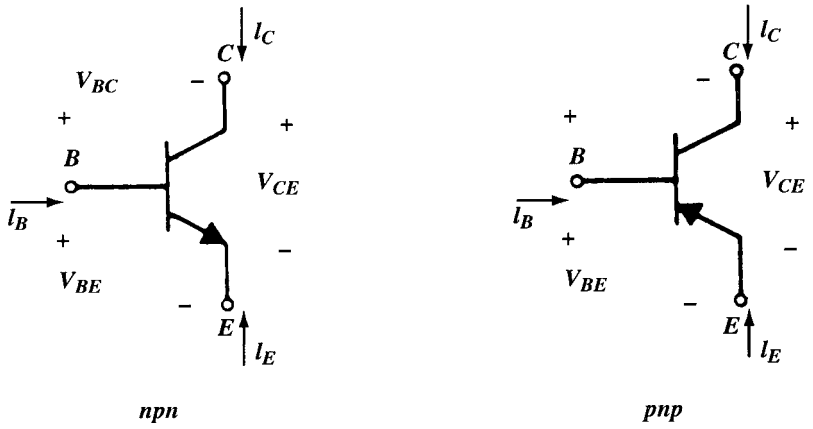


Figure 2-76 Bipolar transistor sign convention.

Epitaxial. This term, as it is commonly used, is actually a shortening of the term *epitaxial-collector*. That is, the collector region of the transistor is formed by the epitaxial technique, rather than by diffusion, which is commonly used to form the base and emitter regions. The epitaxial layer is formed by condensing a single-crystal film of semiconductor material on a wafer or substrate that is usually of the same material. Thus, an epitaxial (collector) transistor is one in which the collector region is formed on a low-resistivity substrate. Subsequently, the base and emitter regions are diffused into the "epi" layer. The epitaxial technique lends itself to precise tailoring of collector-region thickness and resistivity with consequent improved device performance and uniformity.

2-2-2 Large-Signal Behavior of Bipolar Transistors*

In this section, the large-signal or dc behavior of bipolar transistors is considered. Large-signal models are developed for the calculation of total currents and voltages in transistor circuits, and such effects as breakdown-voltage limitations, which are usually not included in models, are also considered. Second-order effects, such as current-gain variation with collector current and Early voltage, can be important in many circuits and are treated in detail.

The sign conventions used for bipolar transistor currents and voltages are shown in Figure 2-76. All bias currents for both *npn* and *pnp* transistors are assumed positive going into the device.

Electrical and Performance Characteristics and Specifications. Electrical characteristics may be described as uniquely defined, measurable electrical transistor properties that are not a function of the measuring circuit or apparatus (except insofar as standard terminations and measurement accuracy are concerned). Performance or operating characteristics are also electrical properties but they are, in general, not unique because their values depend on the measuring circuit (in particular, source and load impedance, which may be arbitrary). As might be expected, the terms are often used somewhat loosely (and sometimes

* Portions of this section, Sections 2-2-3 through 2-2-5, and Sections 2-3-3 through 2-3-6 are based on Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed., ©1993 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

interchangeably), especially in cases where only subtle differences are involved. The terms are generally used on transistor datasheets to segregate (for emphasis) under performance or operating characteristics those properties most directly applicable to the primary intended application.

Performance (Operating) Characteristics. Of the numerous performance characteristics that can be specified for high-frequency transistors, perhaps the most fundamental and pertinent characteristics are:

- Power gain and noise figure, for small-signal applications
- Power gain, power output, and efficiency, for large-signal applications

All of these characteristics are, of course, functions of frequency, bias temperature, and so on, and to completely characterize a transistor over its full frequency, bias, and temperature ranges would be prohibitively costly. Consequently, characterization data are given only for restricted ranges of these variables. These data should portray sufficiently the capabilities of a particular device for its intended applications. In the case of maximum ratings, some applications may require additional characterization by the user or through applications assistance from the manufacturer.

POWER GAIN

G_{max}. Of the various definitions for the measurement of power flow in an active two-port device, such as a transistor, two are unique enough to allow specification without recourse to specifying the complete measuring circuit in detail. One of these definitions is termed *maximum available gain* (G_{max}) and is the power gain obtained when the input and output ports are simultaneously conjugately matched to source and load impedances, respectively. Implicit in this definition is the assumption that the two-port is unconditionally stable; that is, that no combination of input/output tuning can result in increasing gain to the point of oscillation.

$|S_{21}|^2$. The other unique power gain is the gain realized when the transistor is inserted between a source and load with identical impedances (in practice, usually $50 + j0\ \Omega$). This particular insertion of transducer gain happens to coincide with the usual definition of the two-port forward scattering parameter, S_{21}. More precisely, it is equal to the magnitude squared of this parameter and is therefore often identified by the symbol $|S_{21}|^2$. For wideband applications, $|S_{21}|^2$ is important because wideband terminations "not too different from 50 Ω" are more easily realized than are wideband transforming networks that provide the matching required for G_{max}.

NOISE FIGURE. A common measure of the noise generated by an active two-port device—noise that sets a lower limit on amplifier sensitivity—is noise factor (F). This is defined as

$$F = \frac{\text{Input signal-to-noise ratio}}{\text{Output signal-to-noise ratio}} \tag{2-110}$$

or more generally as

$$F = \frac{\text{Total output noise power}}{\text{Output power due to source resistance}} \tag{2-111}$$

At high frequencies, spot noise factor or noise factor for a small fractional bandwidth (say, 1%) is used and is usually expressed as *noise figure* (*NF*), in decibels:

$$NF = 10 \log F \tag{2-112}$$

As already discussed, noise figure is a function of source impedance (as well as functions of frequency, bias, etc.) and hence, there is an infinity of noise figures associated with a given device corresponding to the infinity of possible impedances that may be presented to the device input. The only unique one, in the sense that it does not involve arbitrary source impedances, is $NF_{\min}$, the minimum noise figure obtained (at a given bias and frequency) when the input is tuned to optimize this parameter. This is the noise figure value commonly given on datasheets.

In practical amplifiers that involve more than one stage, *overall noise factor* (F_O) is given by

$$F_O = F_1 + \frac{F_2 - 1}{G_1} + \cdots + \frac{F_n - 1}{G_{(n-1)}} \tag{2-113}$$

where n = number of stages
G_n = gain of the nth stage
F_n = noise factor of the nth stage

This expression emphasizes the important fact that for low-noise amplifiers, the first stage must be designed for the lowest noise figure and highest gain possible. [Note that the noise contribution of the second stage is divided (reduced) by the gain of the first stage.] Since the optimum source impedance and bias currents for optimum gain and noise figure do not often coincide, very careful circuit design is required to minimize overall noise figure. For more accuracy, we recommend the noise correlation approach.

POWER OUTPUT. This characteristic is important for both amplifier and oscillator transistors. In both cases, it is extremely circuit sensitive. For amplifiers, maximum useful output is often limited to that power output level at which gain has compressed 1 dB, an indicator of the upper limit of the device's linearity range. For oscillators, it is merely a quantitative measure of RF power for given dc input power.

EFFICIENCY. In the most general sense, this characteristic expresses as a percentage the ratio of RF power output to the total circuit input power, both dc and RF. That expression is efficiency, η, defined as

$$\eta = \frac{P_O}{P_i + P_{\text{dc}}} \times 100 \tag{2-114}$$

where P_O = RF output power
P_i = RF input power
P_{dc} = total dc power input

Since for oscillator transistors there is no RF power input, and for amplifier transistors the maximum input RF power is calculable from the power gain and power output specifications,

the inclusion of P_i in efficiency is redundant. Moreover, since the major portion of the dc power is dissipated by the transistor collector, a more restricted definition of efficiency is pertinent. This parameter, termed *collector efficiency* (η_C), is given by

$$\eta_C = \frac{P_O}{P_{CC}} \times 100 \tag{2-115}$$

where $P_{CC} = V_{CC} \times I_{CC}$
V_{CC} = collector supply voltage
I_{CC} = collector supply current

Electrical Characteristics. Electrical characteristics may be conveniently classified into two main types, dc and ac.

dc CHARACTERISTICS. The importance of dc characteristics of high-frequency transistors lies primarily in biasing and reliability considerations. However, certain dc characteristics are also directly related to high-frequency performance. For example, high-frequency noise figure is affected by dc current gain. The dc characteristics discussed here are those usually found on high-frequency transistor datasheets.

$V_{(BR)CBO}$, I_{CBO}. These two parameters serve to characterize the reverse-biased collector–base *pn* junction and are defined as follows (with the aid of Figure 2-77a). The collector–base breakdown voltage, $V_{(BR)CBO}$, identifies the voltage at which collector current tends to increase without limit, usually due to the high electric field developed across the junction. This voltage sets a limit on the maximum transistor operating voltage and, as mentioned before under maximum ratings, usually is the basis for the collector–base maximum voltage rating. $V_{(BR)CBO}$ should be specified at a value of $I_C = I_{C1}$ in the figure, which is within the avalanche (or high slope) region of the device's reverse characteristic. Typical values of I_{C1} are in the 1–10-μA region for high-frequency transistors.

To further define the quality of the reverse *I–V* characteristic, a specification is usually placed on the collector cutoff current, I_{CBO}, which is measured at some value of collector–base voltage less than $V_{(BR)CBO}$. For a good-quality silicon junction ("sharp" instead of soft; see Figure 2-77a), I_{CBO} is in the nanoampere range.

$V_{(BR)EBO}$, I_{EBO}. These two parameters characterize the reverse-biased emitter–base *pn* junction in an analogous manner to the collector–base junction parameters $V_{(BR)CBO}$ and I_{CBO}, given above, and are shown in Figure 2-77b. No further discussion of these parameters will be given here.

$V_{(BR)CEO}$, I_{CEO}. The collector–emitter breakdown voltage and cutoff current are somewhat more complex in nature than the collector–base and emitter–base parameters. In the latter two, only a single *pn* diode is involved. In the collector–emitter case, two diodes are involved. Moreover, each is influenced by the other through transistor action, since the reverse current of the collector–base diode flows through the emitter–base junction as forward current. Thus, the collector–base reverse current is amplified by the dc current gain of the transistor, resulting in:

- I_{CEO} being greater than I_{CBO} (for a given voltage)
- Typically, the familiar negative-resistance region in the *I–V* characteristic as shown in Figure 2-77c.

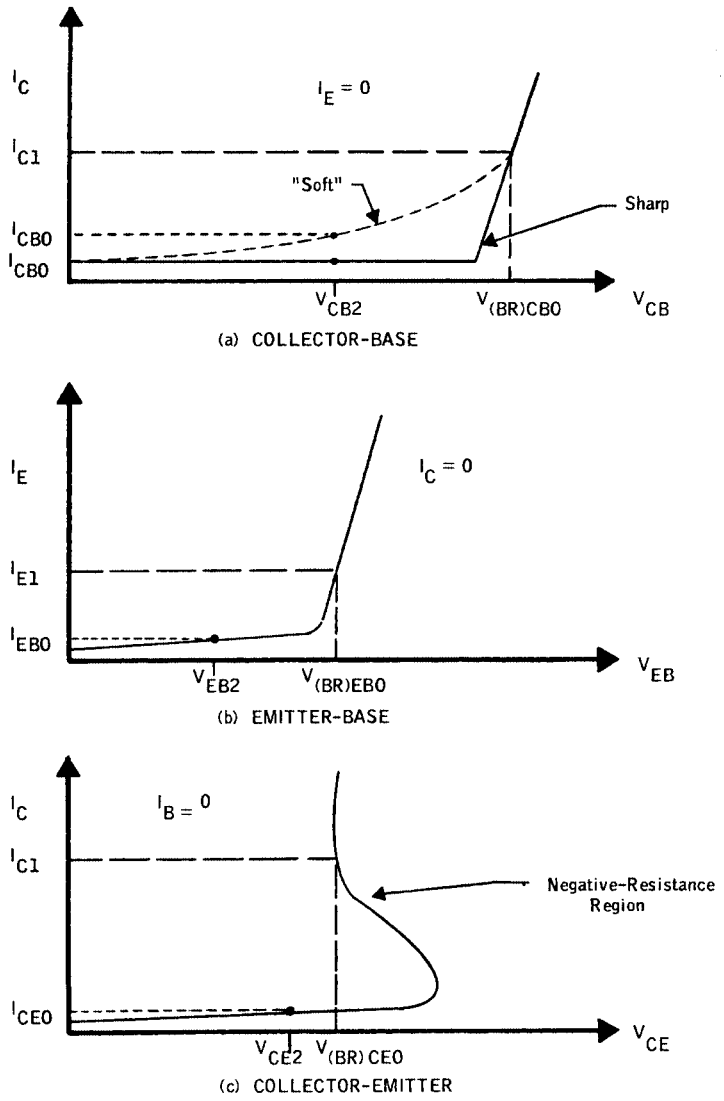


Figure 2-77 Transistor reverse *I–V* characteristics.

Consequently, $V_{(BR)CEO}$ is typically specified at collector currents one to three orders of magnitude higher than in the case of $V_{(BR)CBO}$ and $V_{(BR)EBO}$ to establish the minimum value of this characteristic.

h_{fe}. This parameter is simply the dc common-emitter current gain; that is, the ratio of collector current to base current at some specified collector voltage and current.

ac CHARACTERISTICS. Of the numerous ac characteristics defined for transistors, only relatively few are commonly used in characterizing high-frequency transistors. Some of the more pertinent parameters are briefly covered here.

S Parameters. By far the most useful and conveniently measured set of two-port parameters for transistor high-frequency (roughly 100 MHz and above) characterization is

the *S*-parameter or scattering-matrix set. These parameters completely and uniquely define the small-signal gain and input/output immittance properties of any linear "black box." (By definition, a transistor or any active device is linear under small-signal conditions.) However, these parameters reveal nothing (except possibly indirectly and approximately) about large-signal or noise behavior. Simply interpreted (more general definitions and other interpretations abound in the technical literature), the *S* parameters are merely insertion gains, forward and reverse, and reflection coefficients, input and output, with driven and nondriven ports both terminated in equal impedances, usually $50 + j0\ \Omega$. Such an interpretation tends to make *S* parameters very attractive, once some familiarity is gained, at high (especially microwave) frequencies, since the power flow or gain and reflection-coefficient concepts are more intuitively meaningful than voltage and current conceptual schemes. It should also be mentioned that *S*-parameters can be converted through straightforward matrix transformations to other two-port parameter sets, such as *h*-, *y*-, and *z*-parameters.

Proceeding with more specific definitions, the *S*-parameters are defined analytically by

$$b_1 = S_{11}a_1 + S_{12}a_2 \tag{2-116}$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \tag{2-117}$$

or, in matrix form,

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{2-118}$$

where, referring to Figure 2-78

a_1 = (incoming power at Port 1)$^{1/2}$
b_1 = (outgoing power at Port 1)$^{1/2}$
a_2 = (incoming power at Port 2)$^{1/2}$
b_2 = (outgoing power at Port 2)$^{1/2}$
E_1, E_2 = electrical stimuli at Port 1, Port 2

From the figure and defining linear equations, for $E_2 = 0$, then $a_2 = 0$, and (skipping through numerous rigorous steps)

$$S_{11} = \frac{b_1}{a_1}$$

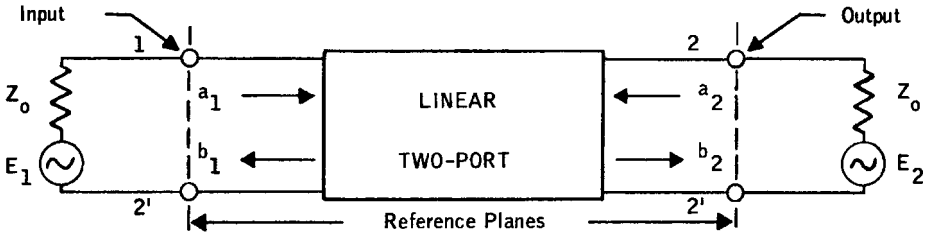


Figure 2-78 Two-port *S*-parameter definition.

$$= \left(\frac{\text{Outgoing input power}}{\text{incoming input power}}\right)^{1/2}$$

$$= \frac{\text{Reflected voltage}}{\text{Incident voltage}}$$

$$= \text{Input reflection coefficient} \tag{2-119}$$

$$S_{21} = \frac{b_2}{a_1}$$

$$= \left(\frac{\text{Outgoing output power}}{\text{incoming input power}}\right)^{1/2}$$

$$= \left(\frac{\text{Output power}}{\text{Available input power}}\right)^{1/2}$$

$$= (\text{Forward transducer gain})^{1/2} \tag{2-120}$$

or, more precisely in the case of S_{21}:

$$\text{Forward transducer gain} = G_{\text{TF}} = |S_{21}|^2 \tag{2-121}$$

$$Z_i = Z_o \tag{2-122}$$

Similarly, at Port 2 for $E_1 = 0$, then $a_1 = 0$, and

$$S_{22} = \frac{b_2}{a_2} = \text{Output reflection coefficient} \tag{2-123}$$

$$S_{12} = \frac{b_1}{a_2} = (\text{Reverse transducer gain})^{1/2} \tag{2-124}$$

$$G_{\text{TR}} = |S_{12}|^2 \tag{2-125}$$

Since many measurement systems display *S*-parameter magnitudes in decibels, the following relationships are particularly useful:

$$|S_{11}|_{\text{dB}} = 10\log|S_{11}|^2$$

$$= 20\log|S_{11}| \tag{2-126}$$

$$|S_{22}|_{\text{dB}} = 20\log|S_{22}| \tag{2-127}$$

$$|S_{21}|_{\text{dB}} = 10\log|S_{21}|^2$$

$$= 20 \log|S_{21}|$$

$$= 10 \log|G_{TF}| = |G_{TF}|_{dB} \tag{2-128}$$

$$|S_{12}|_{dB} = 10 \log|S_{12}|^2$$

$$= 20 \log|S_{12}|$$

$$= 10 \log|G_{TR}| = |G_{TR}|_{dB} \tag{2-129}$$

Figure 2-79 shows an example of the use of *S* parameters in transistor stage design.

Transition Frequency. One of the better known, but perhaps least understood, figures of merit for high-frequency transistors is the so-called transition frequency, f_T. Part of the misunderstanding that appears to exist is due to the use of the misleading (but common, for

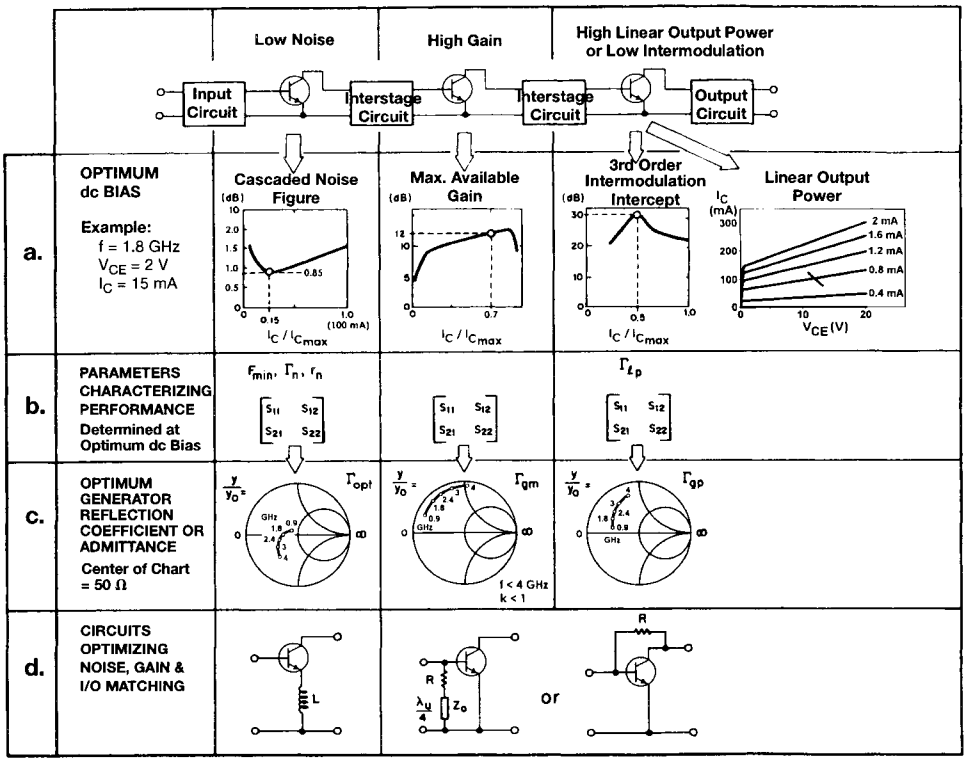


Figure 2-79 Key parameters in applying a BJT in low-noise front-end, high-gain, and linear-power stages. (a) The BJT characteristics that lead to optimum dc bias for a low-noise front end; (b) parameters characterizing device performance; (c) Γ_{opt} for the first stage, and S_{11} and S_{22} for the output stage as simulated by CAD software. The example is based on the Siemens BJTs BFP420 (low-noise stage), BFP450 (high-gain stage), and BFG235 (output stage). Circuits to optimize stage noise, gain, and I/O matching are shown in (d). (Presentation based on Figure 14 in Charles A. Liechti, "Microwave Field-Effect Transistors—1976," *IEEE Trans. Microwave Theory Tech.,* vol. MTT-24, pp. 279–300, June 1976.)

historical reasons) term *short-circuit gain–bandwidth product* for this parameter. By definition, f_T is that characteristic frequency described by the equation

$$f_T = h_{fe} \times f_{\text{meas}} \tag{2-130}$$

where h_{fe} is the magnitude of the small-signal common-emitter short-circuit current gain, and f_{meas} is the frequency of measurement, chosen such that

$$2 \le h_{fe} \le \frac{h_{fe0}}{2} \tag{2-131}$$

where h_{fe0} is the low-frequency value of h_{fe}.

To varying degrees of approximation, depending on the transistor type, f_T is the frequency at which h_{fe} approximates unity. It is not, in general, the frequency at which h_{fe} is precisely equal to unity. To clarify these points further, consider the plot of h_{fe} against frequency sketched in Figure 2-80.

At low frequencies, $f << f_B$, h_{fe} is constant and equal to h_{fe0}.

At f_B, h_{fe} has decreased to 0.707 h_{fe0}; that is, f_B is the 3-dB cutoff frequency for common-emitter short-circuit current gain, h_{fe}.

For frequencies such that

$$2f_B < f < f_T \tag{2-132}$$

h_{fe} varies inversely proportional to frequency. That is, the f_T defining relationship holds:

$$h_{fe} \times f_{\text{meas}} = \text{constant} = f_T \tag{2-133}$$

At frequencies approaching f_T, other parameters, especially package parasitics, can cause $|h_{fe}|$ to depart significantly from this $1/f$ variation. Therefore, the frequency, f_1, at which $|h_{fe}|$ actually equals unity can be somewhat different from f_T.

Applying these frequency-gain characteristics to the common-emitter wideband, low-pass amplifiers gives rise to the terminology of f_T being a "gain–bandwidth product." However,

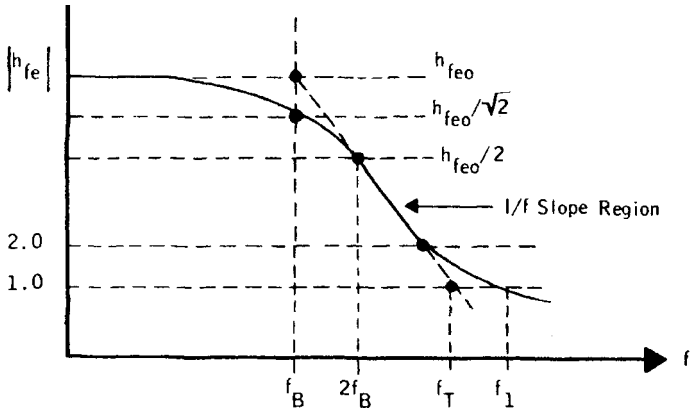


Figure 2-80 $|h_{fe}|$ frequency characteristics.

this is an optimistic approximation at best, since the product of low-frequency circuit gain and the 3-dB cutoff frequency is reduced from f_T by an amount that depends on circuit impedances.

The real significance of f_T lies in the fact that it is a measure of certain internal transistor parameters that do, in fact, affect high-frequency performance: for example, gain (though not in the convenient quantitative manner implied by the term "gain–bandwidth product"). In particular, good high-frequency noise performance requires that f_T be high. Thus, f_T is included on transistor datasheets primarily as a figure of merit, not as a parameter to be used directly in design.

Collector–Base Time Constant, $r_b'C_c$. This is an internal device parameter that relates only indirectly to high-frequency performance. It is primarily a measure of internal feedback within the transistor. It also relates to transistor high-frequency impedance. As the name says (in symbols), it is a measure of transistor base resistance and collector capacitance in combination; however, except for certain low-frequency transistors, it cannot be considered the simple two-element lumped *RC* time constant implied by the terminology. (In high-frequency transistors, base resistance and collector impedance must be considered to be distributed when considered in detail). As a figure of merit, it is included on transistor datasheets to indicate how well base resistance and collector impedance have been minimized. It also allows the estimation of certain gain properties of the transistor (see the $f_{\max}$ parameter, which we'll describe shortly).

Collector–Base Capacitance, C_{cb}. This parameter is simply the total collector–base *pn* junction capacitance measured at a low frequency (typically, 1 MHz), where it can be considered a single lumped element. For high-frequency transistors it is, of course, desirable that C_{cb} be small from bandwidth and stability considerations as well as from gain considerations alone.

Maximum Frequency of Oscillation, $f_{\max}$. This is another figure-of-merit parameter, as opposed to measurable parameters directly usable in the application of transistors. Its importance stems from the following approximate relationships (which will not be derived here):

$$f_{\max} \approx \left(\frac{f_T}{8\pi r_b' C_C} \right)^{1/2} \tag{2-134}$$

$$G_{\max} \approx \left(\frac{f_{\max}}{f_{\text{oper}}} \right)^2 \tag{2-135}$$

These expressions illustrate in a quantitative way the importance and the interrelationship between high f_T and low $r_b'C_C$ insofar as high-frequency gain is concerned. However, since they are approximations and since their derivation involves several assumptions that are not always valid, they must be interpreted with caution. For example, the expression for $G_{\max}$ is obviously not applicable at low frequencies, since as $f \to 0$, $G_{\max} \to \infty$, according to this expression. As a rule of thumb, the $G_{\max}$ expression is a reasonable approximation for frequencies such that

$$5 > \frac{f_{\max}}{f_{\text{oper}}} > 1 \tag{2-136}$$

For accurate analysis of transistor gain and stability, a complete set of two-port parameters must be employed in exact expressions, such as those from which the approximations shown above were derived.

Datasheet Transistor Specifications. The following is a reprint of a Siemens microwave transistor datasheet. This 25-GHz f_T transistor will soon be replaced by a 40-GHz family, based on a silicon–germanium design. Key parameters describing the transistor are:

- The dc parameters, such as maximum current and voltage and dissipation
- Noise figure as a function of generator impedance, bias point, and frequency
- Frequency-dependent behavior, typically expressed in *S* parameters, and also as a function of bias point and frequency
- Large-signal performance to be calculated based on a nonlinear model, such as the Gummel–Poon

The transistor description typically gives details about packaging, as the Gummel–Poon model is intrinsic (without parasitics). Not only does the package determine the parasitics, it also affects the device's ability to transfer heat to its heat sink. The following figures, reproduced with permission from the datasheet for the BFP420 transistor by Siemens (now Infineon Technologies), are a typical representation of the dc-I_C curves, *S* parameters, and noise parameters of a normal, packaged microwave transistor. Besides its dc parameters and capacitance, there are other critical parameters. A BJT's noise figure is a strong function of its dc current gain, and one can determine a 3-dB cutoff frequency referred to as f_T, which shows the gain roll-off as a function of collector current and frequency. Another important measure or figure of merit is $f_{\max}$, the maximum frequency of oscillation:

$$f_{\max} = \sqrt{\frac{|Y_{21}|}{16\pi R_{bb} C_{be} C_{bc}}} \tag{2-137}$$

Figure 2-81 shows a photograph of the die for the BFP420.

Since the intercept point is an important figure of merit, Figure 2-82 shows intercept point as a function of bias point for the BFR93AW BJT. We wish more manufacturers would supply this information.

2-2-3 Large-Signal Transistors in the Forward-Active Region

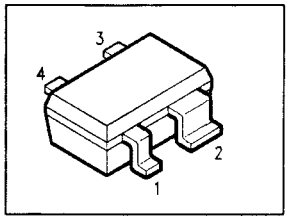
We saw in the diode case that its description is physics-based and it is, therefore, logical to apply the same modeling for the transistors. The first complete modeling for transistors was described in the Gummel–Poon model, but many engineers use the somewhat more simpler Ebers–Moll model. By combining two diodes, the parasitic elements, and the base separating resistor, we can show the "intrinsic" nonlinear model for the bipolar transistor.

Many equations that have been published contain some approximations, but for exact nonlinear analysis, we have to resort to the somewhat more complex or complete equations as shown in Appendix A for the bipolar model.

Recently, some other models have been proposed by researchers, such as the MEXTRAM models (Philips), but the trouble with these models is that while they may give quite good results for RF applications, it is virtually impossible to obtain parameters for them. So far the only industry-standard model is the Gummel–Poon, and to the authors' knowledge there

SIEMENS **SIEGET®25 BFP420**

NPN Silicon RF Transistor



- For High Gain Low Noise Amplifiers
- For Oscillators up to 10 GHz
- Noise Figure F = 1.05 dB at 1.8 GHz
  Outstanding G_{ms} = 20 dB at 1.8 GHz
- Transition Frequency f_T = 25 GHz
- Gold metallization for high reliability
- **SIEGET®25-Line**
  Siemens **G**rounded **E**mitter **T**ransistor-
  25 GHz f_T-**Line**

ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering Code (8-mm taped)	Pin Configuration 1	2	3	4	Package [1]
BFP420	AMs	Q62702-F1591	B	E	C	E	SOT343

Maximum Ratings

Parameter	Symbol		Unit
Collector-emitter voltage	V_{CEO}	4.5	V
Collector-base voltage	V_{CBO}	15	V
Emitter-base voltage	V_{EBO}	1.5	V
Collector current	I_C	35	mA
Base current	I_B	3	mA
Total power dissipation, Ts ≤ 107°C [2)3)]	P_{tot}	160	mW
Junction temperature	T_j	150	°C
Ambient temperature range	T_A	-65...+150°C	°C
Storage temperature range	T_{stg}	-65...+150°C	°C

Thermal Resistance

Junction-soldering point [2)]	$R_{th\ JS}$	270	K/W

1) For detailed information see chapter Package

2) T_S is measured on the emitter lead at the soldering point to the pcb.

3) P_{tot} due to Maximum Ratings.
At typical Ts ≤ 80°C: P_{tot} = 250 mW due to thermical characteristics.

SIEMENS **SIEGET®25 BFP420**

Electrical Characteristics

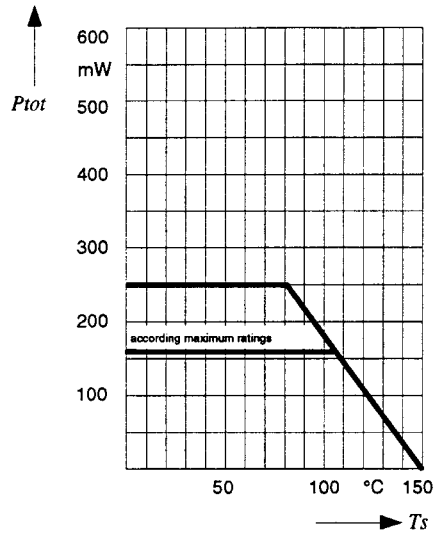
at T_A = 25 °C, unless otherwise specified.

Parameter	Symbol	Value			Unit
		min.	typ.	max.	
DC Characteristics					
Collector-emitter breakdown voltage I_C = 1 mA	$V_{(BR)CEO}$	4.5	5	6.5	V
Collector-cutoff current V_{CB} = 5 V, I_E = 0	I_{CBO}	-	-	200	nA
Emitter base cuttoff current V_{EB} = 1.5 V, I_C = 0	I_{EBO}	-	-	35	µA
DC current gain I_C = 20 mA, V_{CE} = 4 V	h_{FE}	50	80	150	
AC Characteristics					
Transition frequency I_C = 30 mA, V_{CE} = 3 V, f = 2 GHz	f_T	20	25	-	GHz
Collector-base capacitance V_{CB} = 2 V, V_{BE} = v_{be} = 0, f = 1 MHz	C_{cb}	-	0.15	0.24	pF
Collector-emitter capacitance V_{CE} = 2 V, V_{BE} = v_{be} = 0, f = 1 MHz	C_{ce}	-	0.41	-	pF
Emitter-base capacitance V_{EB} = 0.5 V, V_{CB} = v_{cb} = 0, f = 1 MHz	C_{eb}	-	0.55	-	pF
Noise figure I_C = 5 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}	F	-	1.05	1.4	dB
Power gain I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt} , Z_L = Z_{Lopt}	G_{ms} [1]	-	20	-	dB
Insertion power gain I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_L = 50Ω	$\lvert S_{21}\rvert^2$	14	17	-	dB
Third order intercept point at output I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}, Z_L = Z_{Lopt}	IP_3	-	22	-	dBm
1dB Compression point I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}, Z_L = Z_{Lopt}	P_{-1dB}	-	12	-	dBm

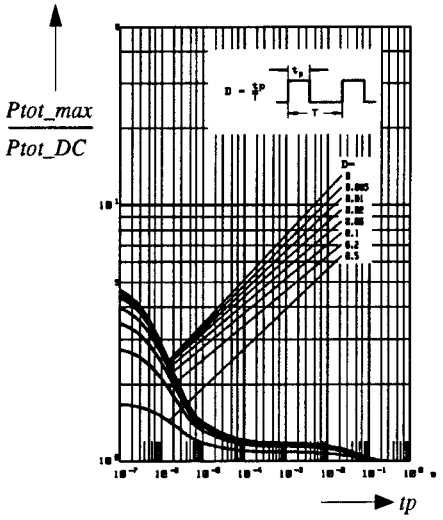
1) $G_{ms} = \left|\frac{S21}{S12}\right|$

SIEMENS **SIEGET®25 BFP420**

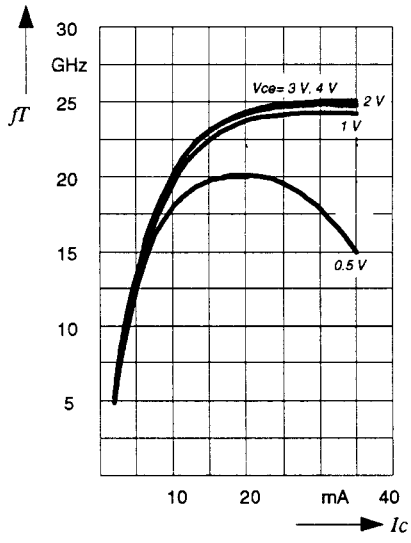
Total Power Dissipation
versus Soldering Point Temperature



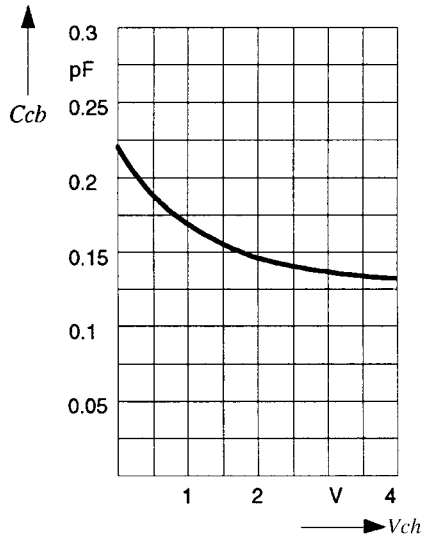
Permissible Pulse Power Dissipation
versus On-Time (V_{CE0max} = 4.5 V)



Transition Frequency
versus Collector Current
f = 2 GHz



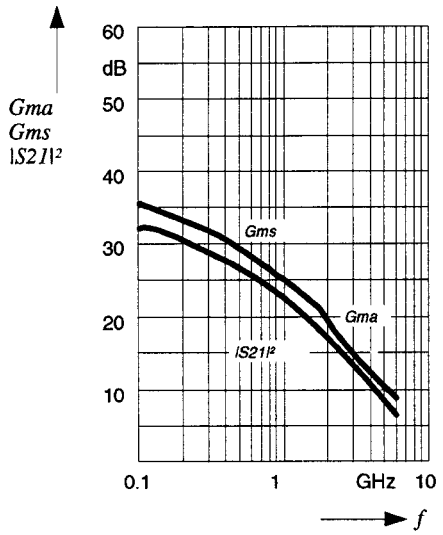
Collector-base Capacitance
versus Collector-base Voltage
V_{BE} = 0 V, f = 1MHz



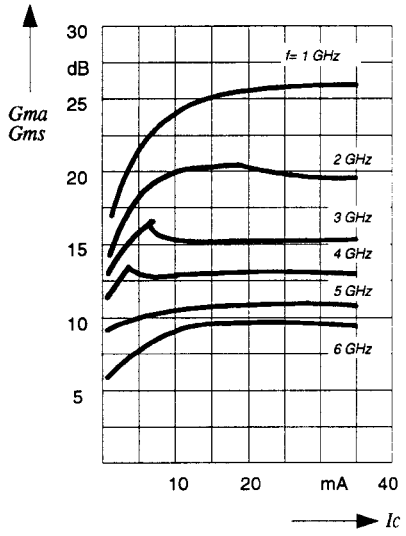
SIEMENS

SIEGET®25 BFP420

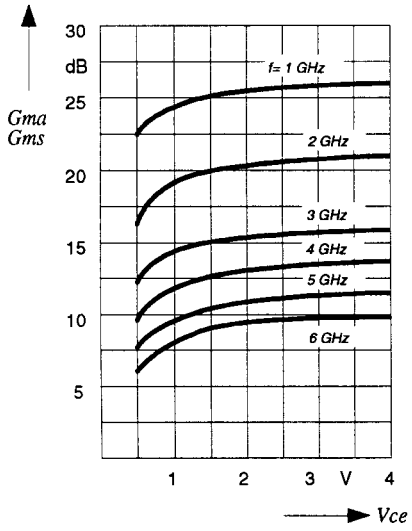
Power Gain
versus Frequency
V_{CE} = 2 V , I_C = 20 mA



Power Gain
versus Collector Current
V_{CE} = 2 V

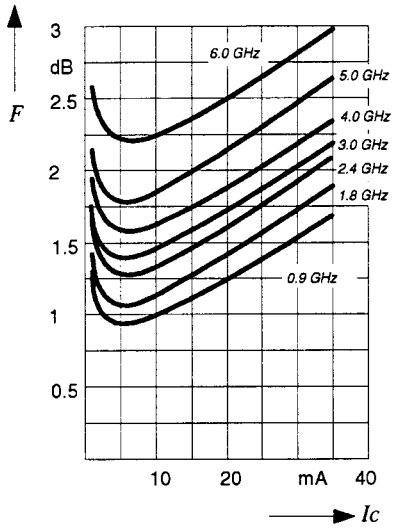


Power Gain
versus Collector Voltage
I_C = 20 mA

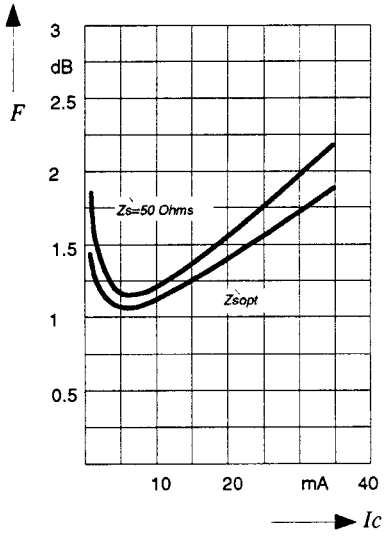


SIEMENS SIEGET®25 BFP420

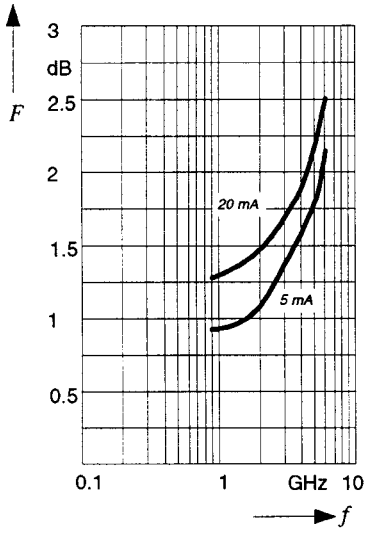
Noise Figure
versus Collector Current
V_{CE} = 2 V, Z_S = Z_{Sopt}



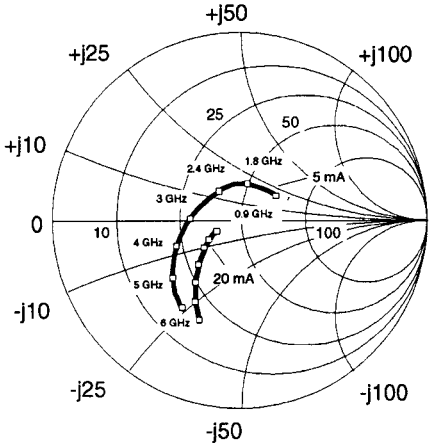
Noise Figure
versus Collector Current
VCE = 2 V, f = 1.8 GHz



Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA,
Z_S = Z_{Sopt}



Source Impedance for min.
Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA



SIEMENS **SIEGET®25 BFP420**

Common Emitter S-Parameters

f	S_{11}		S_{21}		S_{12}		S_{22}	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
V_{CE} = 2 V, I_C = 20 mA								
0.01	0.452	-2.3	37.62	178.3	0.0011	94.4	0.956	-0.6
0.1	0.447	-25.1	36.30	164.7	0.0068	82.5	0.941	-12.4
0.5	0.386	-101.1	23.41	121.0	0.0262	61.7	0.632	-47.2
1.0	0.378	-146.2	13.99	96.0	0.0395	57.8	0.395	-63.9
2.0	0.405	173.5	7.18	70.8	0.0664	54.0	0.222	-87.3
3.0	0.446	149.4	4.77	52.6	0.0949	47.1	0.133	-111.3
4.0	0.501	130.0	3.52	36.8	0.1206	38.5	0.133	-158.5
6.0	0.599	104.8	2.27	8.2	0.1646	18.9	0.196	142.0
8.0	0.700	78.5	1.51	-20.8	0.1800	-2.4	0.289	99.3
9.0	0.758	67.6	1.25	-34.4	0.1820	-13.0	0.379	84.1
10.0	0.800	62.0	1.04	-43.5	0.1800	-19.3	0.465	76.6
V_{CE} = 2 V, I_C = 5 mA								
0.01	0.790	-1.0	15.14	179.2	0.0012	83.4	0.988	-0.7
0.1	0.786	-11.6	14.98	171.8	0.0092	84.1	0.982	-6.5
0.5	0.702	-55.7	12.86	140.1	0.0398	62.8	0.857	-29.8
1.0	0.589	-99.1	9.63	112.6	0.0603	46.5	0.647	-48.6
2.0	0.507	-156.0	5.60	79.4	0.0798	34.6	0.401	-70.3
3.0	0.511	168.5	3.84	57.1	0.0957	29.8	0.267	-84.2
4.0	0.549	142.0	2.87	38.5	0.1121	25.1	0.207	-110.5
5.0	0.604	123.9	2.26	22.1	0.1285	19.4	0.150	-137.3
6.0	0.633	110.0	1.86	6.7	0.1442	13.1	0.173	-169.8

Common Emitter Noise Parameters

f	F_{min} 1)	G_a 1)	Γ_{opt}		R_N	r_n	$F_{50\Omega}$ 2)	$\lvert S_{21}\rvert^2$ 2)
GHz	dB	dB	MAG	ANG	Ω	-	dB	dB
V_{CE} = 2 V, I_C = 5 mA								
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	-175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	-157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	-142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	-123.0	15.0	0.30	3.30	5.3

1) Input matched for minimum noise figure, output for maximum gain 2) Z_S=Z_L=50 Ω

SIEMENS **SIEGET®25 BFP420**

SPICE Parameters:

Transistor Chip Data T502 (Berkeley-SPICE 2G.6 Syntax):

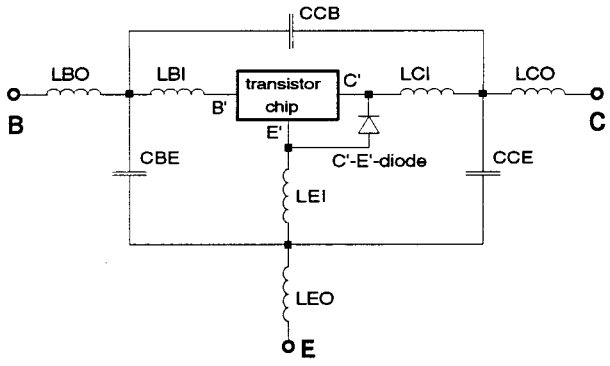
IS =	0.20045	fA	BF =	72.534	-	NF =	1.2432	-
VAF =	28.383	V	IKF =	0.48731	A	ISE =	19.049	fA
NE =	2.0518	-	BR =	7.8287	-	NR =	1.3325	-
VAR =	19.705	V	IKR =	0.69141	A	ISC =	0.019237	fA
NC =	1.1724	-	RB =	3.4849	OHM	IRB =	0.72983	mA
RBM =	8.5757	OHM	RE =	0.31111	OHM	RC =	0.10105	OHM
CJE =	1.8063	fF	VJE =	0.8051	V	MJE =	0.46576	-
TF =	6.7661	ps	XTF =	0.42199	-	VTF =	0.23794	V
ITF =	1.0	mA	PTF =	0	deg	CJC =	234.53	fF
VJC =	0.81969	V	MJC =	0.30232	-	XCJC =	0.3	-
TR =	2.3249	ns	CJS =	0	fF	VJS =	0.75	V
MJS =	0	-	XTB =	0	-	EG =	1.11	eV
XTI =	3.0	-	FC =	0.73234	-	Tnom =	300	K

C'-E'- Diode Data (Berkeley-SPICE 2G.6 Syntax):

IS =	3.5	fA	N =	1.02	-	RS =	10	Ω

All parameters are ready to use, no scaling is necessary. Caution: In these extracted parameters, RBM > the operating RB.

Package Equivalent Circuit:



SOT343-3:

LBI = 0.47 nH
LBO = 0.53 nH
LEI = 0.23 nH
LEO = 0.05 nH
LCI = 0.56 nH
LCO = 0.58 nH
CBE = 136 fF
CCE = 134 fF
CCB = 6.9 fF

valid up to 6 GHz

The SOT343 package has two emitter leads. To avoid high complexity of the package equivalent circuit, both leads are combined in one electrical connection.

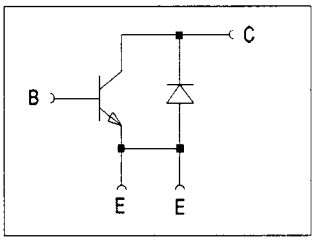
Extracted on behalf of SIEMENS Small Signal Semiconductors by:
Institut für Mobil- und Satellitenfunktechnik (IMST)
© 1996 SIEMENS AG

For non-linear simulation:

- Use transistor chip parameters in Berkeley SPICE 2G.6 syntax for all simulators.
- If you need simulation of the reverse characteristics, add the diode with the C'-E'-diode data between collector and emitter.
- Simulation of the package is not necessary for frequencies < 100 MHz.
  For higher frequencies add the wiring of the package equivalent circuit around the non-linear transistor and diode model.

Note:

- This transistor is constructed in a common emitter configuration. This feature causes an additional, reverse biased diode between emitter and collector, which does not effect normal operation.



Transistor Schematic Diagram

The common emitter configuration shows the following advantages:

- Higher gain because of lower emitter inductance.
- Power is dissipated via the grounded emitter leads, because the chip is mounted on the copper emitter leadframe.

Please note, that the broadest lead is the emitter lead.

- The AC-Characteristics are verified by random sampling.

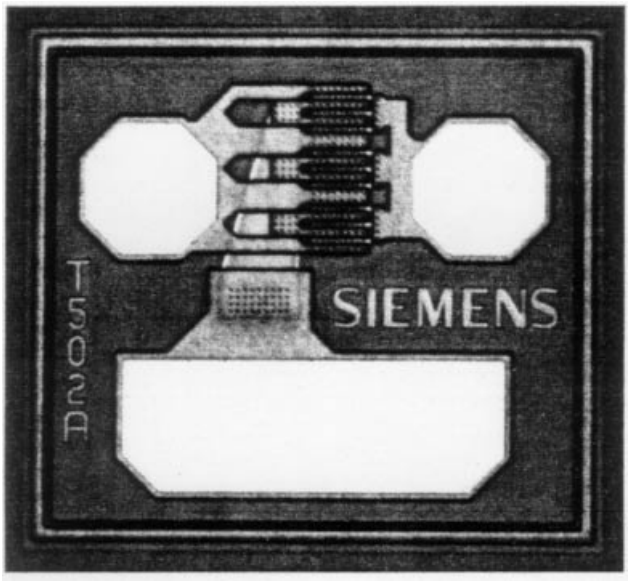


Figure 2-81 Photograph of the BFP420 die (310 × 290 μm). The bonding pad for the base is at the upper right; for the collector at the upper left. The large, lower pad connects to the emitter.

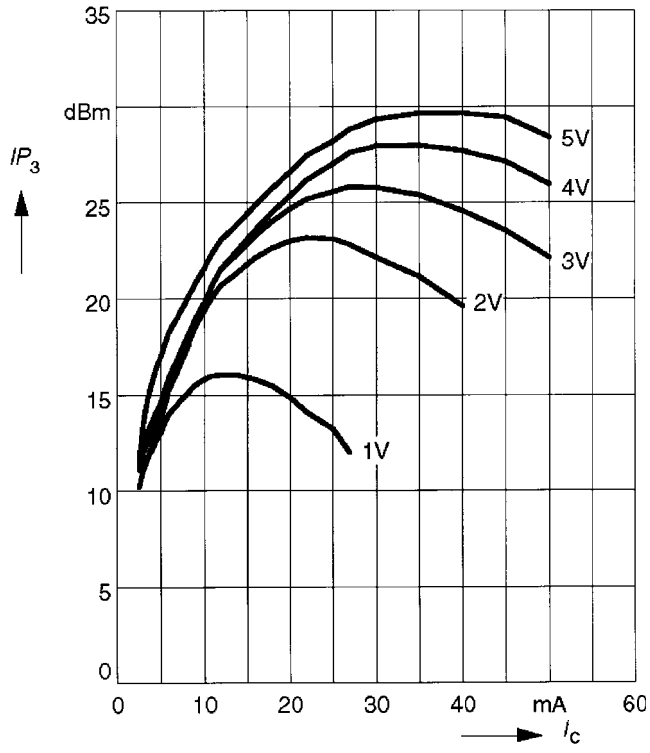


Figure 2-82 $IP_{3,out}$ versus I_C at five values of V_{CE} for the BRF93AW BJT.

are no accepted or affordable parameter extractor programs available, not even to consider the cost of hardware for such extractions. Hewlett-Packard recently published a system with a hardware cost in the vicinity of $1 million.

A typical *npn* planar bipolar transistor structure is shown in Figure 2-83a, where collector, base, and emitter are labeled *C*, *B*, and *E*, respectively. The impurity doping density in the base and the emitter of such a transistor is not constant but varies with distance from the top surface. However, many of the characteristics of such a device can be predicted by analyzing

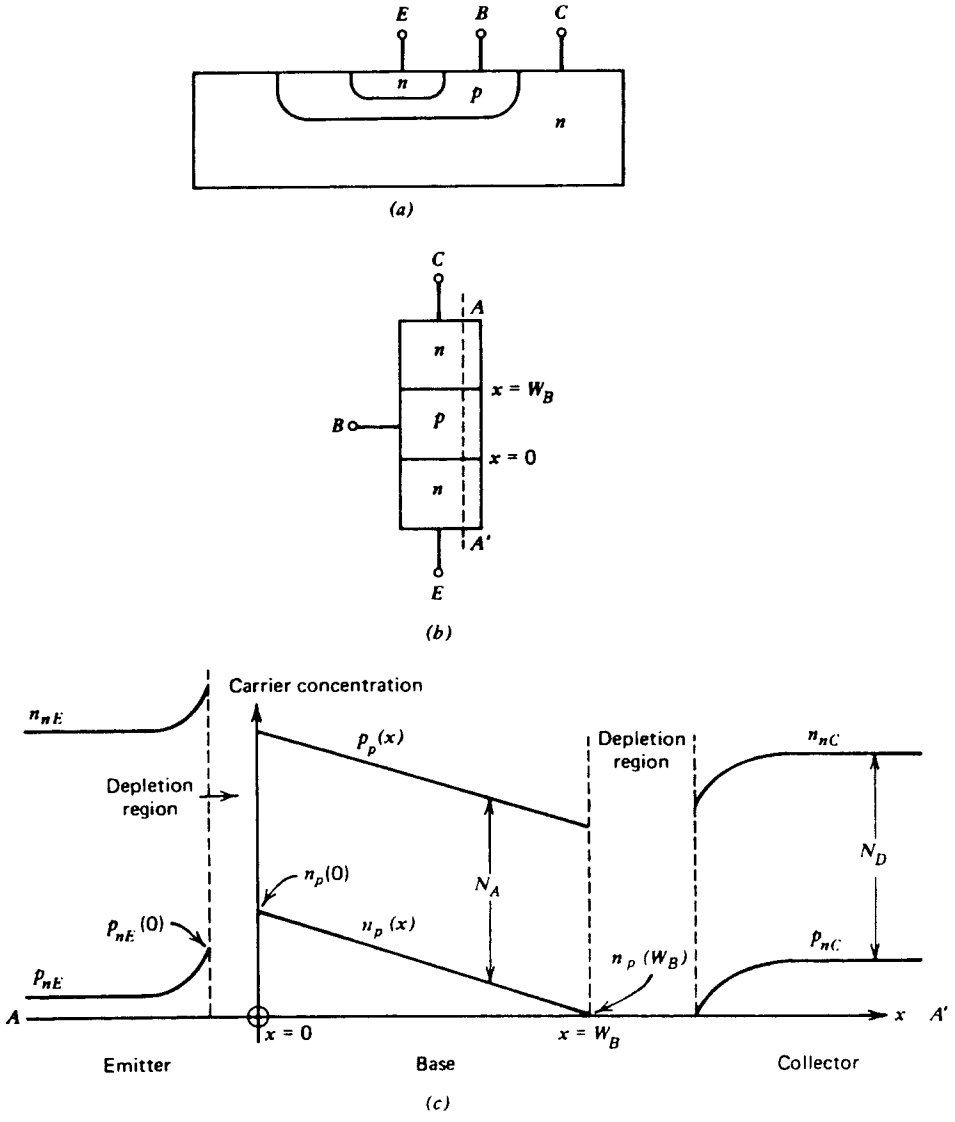


Figure 2-83 (a) Cross section of a typical *npn* planar bipolar transistor structure; (b) idealized transistor structure; (*c*) carrier concentrations along the cross section *AA*′ of the transistor in (b). Uniform doping densities are assumed. (Not to scale)

the idealized transistor structure shown in Figure 2-83b. In this structure, the base and emitter doping densities are assumed constant, and this is sometimes called a "uniform-base" transistor. Where possible in the following analyses, the equations for the uniform-base analysis are expressed in a form that applies also to nonuniform-base transistors.

A cross section AA' is taken through the device of Figure 2-83b and carrier concentrations along this section are plotted in Figure 2-83c. Hole concentrations are denoted by p and electron concentrations by n with subscripts p or n representing p-type or n-type regions. The n-type emitter and collector regions are distinguished by subscripts E and C, respectively. The carrier concentrations shown in Figure 2-83c apply to a device in the *forward-active region*. That is, the base–emitter junction is forward-biased and the base–collector junction is reverse-biased. The minority-carrier concentrations in the base at the edges of the depletion regions can be calculated from a Boltzmann approximation to the Fermi–Dirac distribution function to give [2]

$$n_p(0) = n_{po} \exp \frac{V_{BE}}{V_T} \tag{2-138}$$

$$n_p(W_B) = n_{po} \exp \frac{V_{BC}}{V_T} \simeq 0 \tag{2-139}$$

where W_B is the width of the base from the base–emitter junction depletion layer edge of the base–collector depletion layer edge and n_{po} is the equilibrium concentration of electrons in the base. Note that V_{BC} is negative for an *npn* transistor in the forward-active region and thus $n_p(W_B)$ is very small. Low-level injection conditions are assumed in the derivation of Eqs. (2-138) and (2-139). This means that the minority-carrier concentrations are always assumed much smaller than the majority-carrier concentrations.

If *recombination* of holes and electrons in the base is small, it can be shown that the minority-carrier concentration $n_p(x)$ in the base varies *linearly* with distance [3]. Thus, a straight line can be drawn joining the concentrations at $x = 0$ and $x = W_B$ in Figure 2-83c.

For charge neutrality in the base, it is necessary that

$$N_A + n_p(x) = p_p(x) \tag{2-140}$$

and thus

$$p_p(x) - n_p(x) = N_A \tag{2-141}$$

where $p_p(x)$ is the hole concentration in the base and N_A is the base doping density that is assumed constant. Equation (2-141) indicates that the hole and electron concentrations are separated by a constant amount and thus $p_p(x)$ also varies linearly with distance.

Collector current is produced by minority-carrier electrons in the base diffusing in the direction of the concentration gradient and being swept across the collector–base depletion region by the field existing there. The diffusion current density due to electrons in the base is

$$J_n = qD_n \frac{dn_p(x)}{dx} \tag{2-142}$$

where D_n is the diffusion constant for electrons. From Figure 2-83c,

$$J_n = -qD_n \frac{n_p(0)}{W_B} \tag{2-143}$$

If I_C is the collector current and is taken as positive flowing *into* the collector, it follows from Eq. (2-143) that

$$I_C = qA D_n \frac{n_p(0)}{W_B} \tag{2-144}$$

where A is the cross-sectional area of the emitter. Substitution of Eq. (2-138) into Eq. (2-144) gives

$$I_C = \frac{qAD_n n_{po}}{W_B} \exp \frac{V_{BE}}{V_T} \tag{2-145}$$

$$= I_S \exp \frac{V_{BE}}{V_T} \tag{2-146}$$

where

$$I_S = \frac{qA D_n n_{po}}{W_B} \tag{2-147}$$

and I_S is a constant used to describe the transfer characteristic of the transistor in the forward-active region. Equation (2-147) can be expressed in terms of the base doping density by noting that [4]

$$n_{po} = \frac{n_i^2}{N_A} \tag{2-148}$$

and substitution of Eq. (2-148) in Eq. (2-147) gives

$$I_S = \frac{qAD_n n_i^2}{W_B N_A} = \frac{qA\overline{D}_n n_i^2}{Q_B} \tag{2-149}$$

where $Q_B = W_B N_A$ is the number of doping atoms in the base per unit area of the emitter and n_i is the intrinsic carrier concentration in silicon. In this form Eq. (2-149) applies to both uniform- and nonuniform-base transistors and D_n has been replaced by $\overline{D}_n$, which is an average effective value of the electron diffusion constant in the base. This is necessary for nonuniform-base devices because the diffusion constant is a function of impurity concentration. Typical values of I_S as given by Eq. (2-149) are 10^{-14} to 10^{-16} A.

Equation (2-146) gives the collector current as a function of base–emitter voltage. The base current I_B is also an important parameter and, at moderate current levels, consists of two major components. One of these (I_{B1}) represents recombination of holes and electrons in the base and is proportional to the minority-carrier charge Q_e in the base. From Figure 2-83c, the minority-carrier charge in the base is

$$Q_e = \frac{1}{2} n_p(0) W_B q A \tag{2-150}$$

and we have

$$I_{B1} = \frac{Q_e}{\tau_b} = \frac{1}{2} \frac{n_p(0) W_B q A}{\tau_b} \tag{2-151}$$

where τ_b is the minority-carrier lifetime in the base. I_{B1} represents a flow of majority holes from the base lead into the base region. Substitution of Eq. (2-138) into Eq. (2-151) gives

$$I_{B1} = \frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} \exp \frac{V_{BE}}{V_T} \tag{2-152}$$

The second major component of base current (usually the dominant one in the integrated-circuit *npn* devices) is due to injection of holes from the base into the emitter. This current component depends on the gradient of minority-carrier holes into the emitter and is [5]

$$I_{B2} = \frac{q A D_p}{L_p} p_{nE}(0) \tag{2-153}$$

where D_p is the diffusion constant for holes and L_p is the diffusion length (assumed small) for holes in the emitter. $p_{nE}(0)$ is the concentration of holes in the emitter at the edge of the depletion region and is

$$P_{nE}(0) = p_{nEo} \exp \frac{V_{BE}}{V_T} \tag{2-154}$$

If N_D is the donor atom concentration in the emitter (assumed constant) then

$$p_{nEo} \simeq \frac{n_i^2}{N_D} \tag{2-155}$$

The emitter is deliberately doped much more heavily than the base, making N_D large and p_{nEo} small, so that the base-current component, I_{B2}, is minimized.

Substitution of Eqs. (2-154) and (2-155) into Eq. (2-153) gives

$$I_{B2} = \frac{q A D_p}{L_p} \frac{n_i^2}{N_D} \exp \frac{V_{BE}}{V_T} \tag{2-156}$$

The total base current, I_B, is the sum of I_{B1} and I_{B2}:

$$I_B = I_{B1} + I_{B2} = \left(\frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} + \frac{q A D_p}{L_p} \frac{n_i^2}{N_D} \right) \exp \frac{V_{BE}}{V_T} \tag{2-157}$$

Although this equation was derived assuming uniform base and emitter doping, it gives the correct functional dependence of I_B on device parameters for practical double-diffused

nonuniform-base devices. Second-order components of I_B, which are important at low current levels, are considered later.

Since I_C in Eq. (2-146) and I_B into Eq. (2-147) are both proportional to $\exp(V_{BE}/V_T)$ in this analysis, the base current can be expressed in terms of collector current as

$$I_B = \frac{I_C}{\beta_F} \tag{2-158}$$

where β_F is the forward current gain. An expression for β_F can be calculated by substituting Eqs. (2-145) and (2-157) into Eq. (2-158) to give

$$\beta_F = \frac{qAD_n n_{po}/W_B}{\frac{1}{2}\frac{n_{po}W_B qA}{\tau_b} + \frac{qAD_p n_i^2}{L_p N_D}} = \frac{1}{\frac{W_B^2}{2\tau_b D_n} + \frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \tag{2-159}$$

where Eq. (2-148) has been substituted for n_{po}. Equation (2-159) shows that β_F is maximized by minimizing the base width W_B and maximizing the ratio of emitter to base doping densities N_D/N_A. Typical values of β_F for *npn* transistors in integrated circuits are 50–500, whereas lateral *pnp* transistors have values of 10–100. Finally, the emitter current is

$$I_E = -(I_C + I_B) = -\left(I_C + \frac{I_C}{\beta_F}\right) = -\frac{I_C}{\alpha_F} \tag{2-160}$$

where

$$\alpha_F = \frac{\beta_F}{1+\beta_F} \tag{2-161}$$

The value of α_F can be expressed in terms of device parameters by substituting Eq. (2-159) into Eq. (2-161) to obtain

$$\alpha_F = \frac{1}{1+\frac{1}{\beta_F}} = \frac{1}{1+\frac{W_B^2}{2\tau_b D_n} + \frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \simeq \alpha_T \gamma \tag{2-162}$$

where

$$\alpha_T = \frac{1}{1+\frac{W_B^2}{2\tau_b D_n}} \tag{2-163}$$

$$\gamma = \frac{1}{1+\frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \tag{2-164}$$

The validity of Eq. (2-162) depends on $W_B^2/2\tau_b D_n << 1$ and $(D_p/D_n)(W_B/L_p)(N_A/N_D) << 1$, and this is always true if β_F is large [see Eq. (2-159)]. The term γ in Eq. (2-162) is

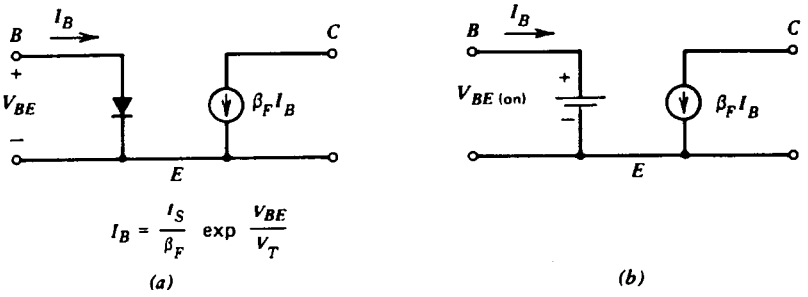


Figure 2-84 Large-signal models of *npn* transistors for use in bias calculations. (a) Circuit incorporating an input diode. (b) Simplified circuit with an input voltage source.

called the *emitter injection efficiency* and is equal to the ratio of the electron current (*npn* transistor) injected into the base from the emitter to the total hole and electron current crossing the base–emitter junction. Ideally $\gamma \to 1$, and this is achieved by making N_D/N_A large and W_B small. In that case very little reverse injection occurs from base to emitter.

The term α_T in Eq. (2-162) is called the *base transport factor* and represents the fraction of carriers injected into the base (from the emitter) that reach the collector. Ideally $\alpha_T \to 1$, and this is achieved by making W_B small. It is evident from the above development that fabrication changes that cause α_T and γ to approach unity also maximize the value of β_F of the transistor.

The results derived above allow formulation of a large-signal model of the transistor suitable for bias-circuit calculations with devices in the forward-active region. One such circuit is shown in Figure 2-84 and consists of a base–emitter diode to model Eq. (2-157) and a controlled collector-current generator to model Eq. (2-158). Note that the collector voltage ideally has no influence on the collector current and the collector node acts as a high-impedance current source. A simpler version of this equivalent circuit, which is often useful, is shown in Figure 2-84b, where the input diode has been replaced by a battery with a value $V_{BE(\text{on})}$, which is usually 0.6–0.7 V. This represents the fact that in the forward-active region the base–emitter voltage varies very little because of the steep slope of the exponential characteristic. In some circuits the temperature coefficient of $V_{BE(\text{on})}$ is important and a typical value for this is –2 mV/°C. The equivalent circuits of Figure 2-84 apply for *npn* transistors. For *pnp* devices the corresponding equivalent circuits are shown in Figure 2-85.

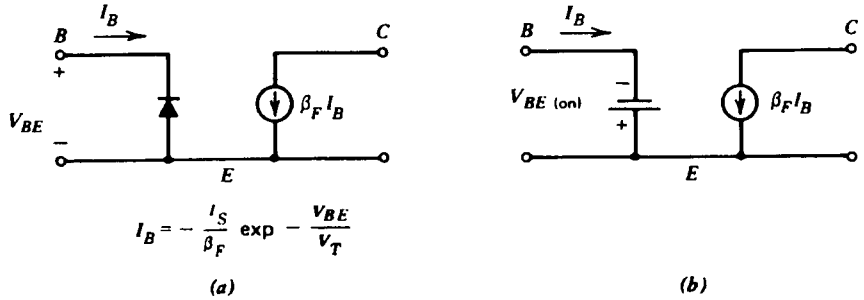


Figure 2-85 Large-signal models of *pnp* transistors corresponding to the circuits of Figure 2-84.

2-2-4 Effects of Collector Voltage on Large-Signal Characteristics in the Forward-Active Region

In the analysis of the previous section, the collector–base junction was assumed reverse-biased and ideally had no effect on the collector currents. This is a useful approximation for first-order calculations, but it is not strictly true in practice. There are occasions where the influence of collector voltage on collector current is important, and this will now be investigated.

The collector voltage has a dramatic effect on the collector current in two regions of device operation. These are the saturation (V_{CE} approaches zero) and breakdown (V_{CE} very large) regions that will be considered later. For values of collector–emitter voltage V_{CE} between these extremes, the collector current increases slowly as V_{CE} increases. The reason for this can be seen from Figure 2-86, which is a sketch of the minority-carrier concentration in the base of the transistor. Consider the effect of changes in V_{CE} on the carrier concentration for constant V_{BE}. Since V_{BE} is constant, the change in V_{CB} equals the change in V_{CE}, and this causes an increase in the collector–base depletion-layer width as shown. The change in the base width of the transistor, ΔW_B, equals the change in depletion-layer width and causes an increase ΔI_C in the collector current.

From Eqs. (2-146) and (2-149) we have

$$I_C = \frac{qA\bar{D}_n n_i^2}{Q_B} \exp \frac{V_{BE}}{V_T} \tag{2-165}$$

Differentiation of Eq. (2-165) yields

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{qA\bar{D}_n n_i^2}{Q_B^2} \left(\exp \frac{V_{BE}}{V_T} \right) \frac{dQ_B}{dV_{CE}} \tag{2-166}$$

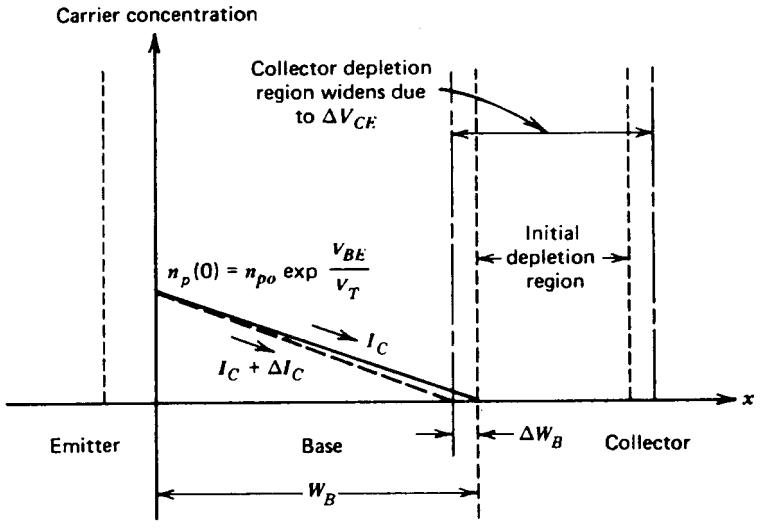


Figure 2-86 Effect of increases in V_{CE} on the collector depletion region and base width of a bipolar transistor.

and substitution of Eq. (2-165) in Eq. (2-166) gives

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{I_C}{Q_B}\frac{dQ_B}{dV_{CE}} \tag{2-167}$$

For a uniform-base transistor $Q_B = W_B N_A$ and Eq. (2-167) becomes

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{I_C}{W_B}\frac{dW_B}{dV_{CE}} \tag{2-168}$$

Note that since the base width *decreases* as V_{CE} increases, dW_B/dV_{CE} in Eq. (2-168) is negative and thus $\partial I_C/\partial V_{CE}$ is positive. dW_B/dV_{CE} is a function of the bias value of V_{CE}, but the variation is typically small for a reverse-biased junction and dW_B/dV_{CE} is often assumed constant. The resulting predictions agree adequately with experimental results.

Equation (2-168) shows that $\partial I_C/\partial V_{CE}$ is proportional to the collector-bias current and inversely proportional to the transistor base width. Thus narrow-base transistors show a greater dependence of I_C on V_{CE} in the forward-active region. The dependence of $\partial I_C/\partial V_{CE}$ on I_C results in typical transistor output characteristics as shown in Figure 2-87. In accordance with the assumptions made in the foregoing analysis, these characteristics are shown for constant values of V_{BE}. However, in most integrated-circuit transistors the base current is dependent only on V_{BE} and not on V_{CE}, and thus constant-base-current characteristics can often be used in the following calculation. The reason for this is that the base current is usually dominated by the I_{B2} component of Eq. (2-156), which has no dependence on V_{CE}. Extrapolation of the characteristics of Figure 2-87 back to the V_{CE} axis gives an intercept V_A called the Early voltage, where

$$V_A = \frac{I_C}{\dfrac{\partial I_C}{\partial V_{CE}}} \tag{2-169}$$

Substitution of Eq. (2-168) in Eq. (2-169) gives

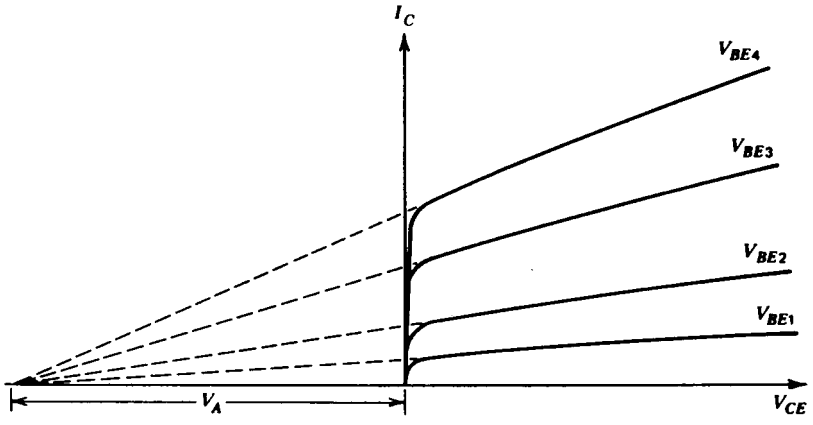


Figure 2-87 Bipolar transistor output characteristics showing the Early voltage, V_A.

$$V_A = -W_B \frac{dV_{CE}}{dW_B} \tag{2-170}$$

which is a constant, independent of I_C. Thus all the characteristics extrapolate to the same point on the V_{CE} axis. The variation of I_C with V_{CE} is called the Early effect and V_A is a common model parameter for circuit-analysis computer programs. Typical values of V_A for integrated-circuit transistors are 15–100 V. The inclusion of Early effect in dc bias calculations is usually limited to computer analysis because of the complexity introduced into the calculation. However, the influence of the Early effect is often dominant in small-signal calculations for high-gain circuits and this point will be considered later.

Finally, the influence of the Early effect on the transistor large-signal characteristics in the forward-active region can be represented approximately by modifying Eq. (2-146) to

$$I_C = I_S \left(1 + \frac{V_{BE}}{V_A}\right) \exp \frac{V_{BE}}{V_T} \tag{2-171}$$

This is a common means of representing the device output characteristics for computer simulation.

2-2-5 Saturation and Inverse Active Regions

Saturation is a region of device operation that is usually avoided in analog circuits because the transistor gain is very low in this region. Saturation is much more commonly encountered in digital circuits, where it provides a well-specified output voltage that represents a logic state. In saturation, both emitter–base and collector–base junctions are forward biased. Consequently, the collector–emitter voltage V_{CE} is quite small and usually in the range 0.05–0.3 V. The carrier concentrations in a saturated *npn* transistor with uniform base doping are shown in Figure 2-88. The minority-carrier concentration in the base of the edge of the depletion region is again given by Eq. (2-139) as

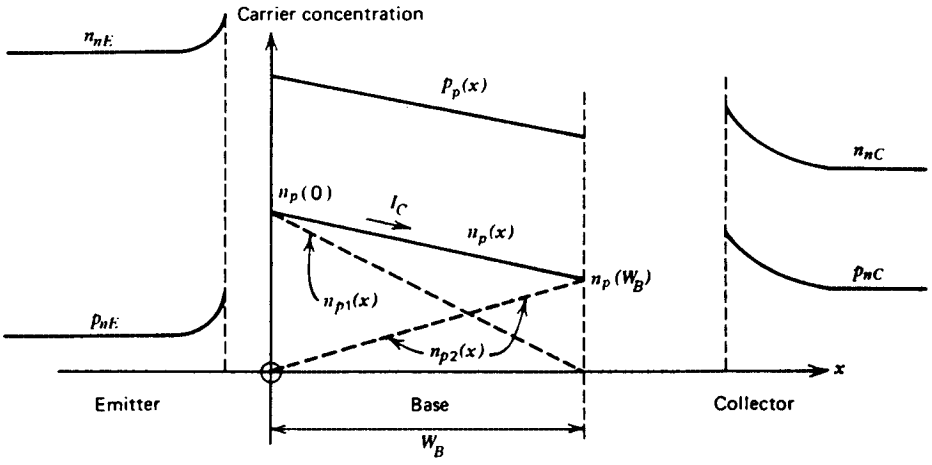


Figure 2-88 Carrier concentrations in a saturated *npn* transistor. (Not to scale)

$$n_p(W_B) = n_{po} \exp \frac{V_{BC}}{V_T} \tag{2-172}$$

but since V_{BC} is now positive, the value of $n_p(W_B)$ is no longer negligible. Consequently, changes in V_{CE} with V_{BE} held constant (which cause equal changes in V_{BC}) directly affect $n_p(W_B)$. Since the collector current is proportional to the slope of the minority-carrier concentration in the base [see Eq. (2-142)], it is also proportional to $[n_p(0) - (n_p(W_B)]$ from Figure 2-88. Thus, changes in $n_p(W_B)$ directly affect the collector current, and the collector node of the transistor appears to have a *low impedance*. As V_{CE} is decreased in saturation with V_{BE} held constant, V_{BC} increases as does $n_p(W_B)$ from Eq. (2-172). Thus, from Figure 2-88 the collector current decreases because the slope of the carrier concentration decreases. This gives rise to the saturation region of the $I_C - V_{CE}$ characteristic shown in Figure 2-89. The slope of the $I_C - V_{CE}$ characteristic in this region is largely determined by the resistance in series with the collector lead due to the finite resistivity of the *n*-type collector material. A useful model for the transistor in this region is shown in Figure 2-90

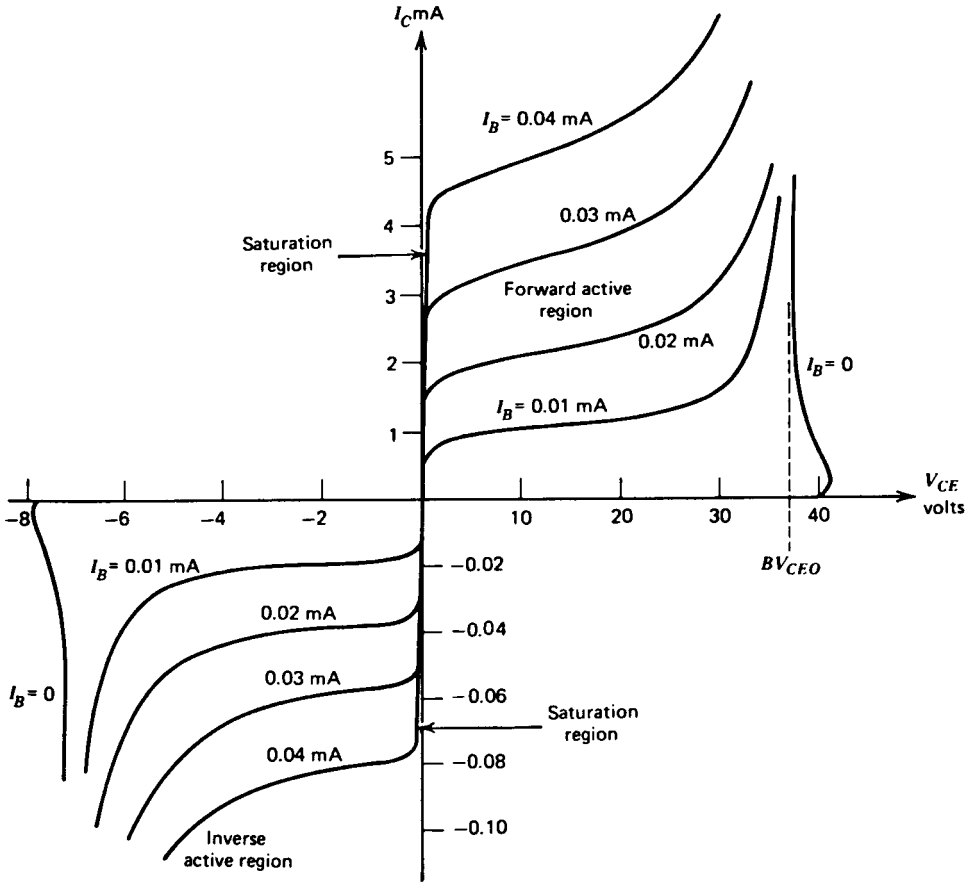


Figure 2-89 Typical $I_C - V_{CE}$ characteristics for an *npn* bipolar transistor. Note the different scales for positive and negative currents and voltages.

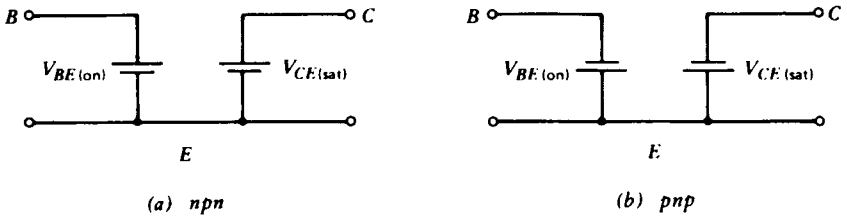


Figure 2-90 Large-signal models for bipolar transistors in the saturation region.

and consists of a fixed voltage source to represent $V_{BE(on)}$, and a fixed voltage source to represent the collector–emitter voltage $V_{CE(sat)}$. A more accurate but more complex model includes a resistor in series with the collector. This resistor can have a value ranging from 20 to 500 Ω, depending on the device structure.

An additional aspect of transistor behavior in the saturation region is apparent from Figure 2-88. For a given collector current, there is now a much larger amount of stored charge in the base than there is in the forward-active region. Thus the base-current contribution represented by Eq. (2-152) will be larger in saturation. In addition, since the collector–base junction is now forward biased, there is a new base-current component due to injection of carriers from the base to the collector. These two effects result in a base current I_B in saturation, which is larger than in the forward-active region for a given collector current I_C. Ratio I_C/I_B in saturation is often referred to as the *forced* β and is always less than β_F. As the forced β is made lower with respect to β_F, the device is said to be more *heavily saturated*.

The minority-carrier concentration in saturation, shown in Figure 2-88, is a straight line joining the two end points, assuming that recombination is small. This can be represented by a linear superposition of the two dotted distributions as shown. The justification for this is that the terminal currents depend *linearly* on the concentrations $n_p(0)$ and $n_p(W_B)$. This picture of device-carrier concentrations can be used to derive some general equations describing transistor behavior. Each of the distributions in Figure 2-88 is considered separately and the two contributions are combined. The *emitter* current that would result from $n_{p1}(x)$ above is given by the classical diode equation

$$I_{EF} = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) \tag{2-173}$$

where I_{ES} is a constant that is often referred to as the "saturation current" of the junction (no connection with the transistor saturation described above). Equation (2-173) predicts that the junction current is given by $I_{EF} \simeq I_{ES}$ with a reverse-bias voltage applied. However, in practice Eq. (2-173) is applicable only in the forward-bias region, since second-order effects dominate under reverse-bias conditions and typically result in a junction current several orders of magnitude larger than I_{ES}. The junction current that flows under reverse-bias conditions is often called the "leakage current" of the junction.

Returning to Figure 2-88, we can describe the *collector* current resulting from $n_{p2}(x)$ alone as

$$I_{CR} = -I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-174}$$

where I_{CS} is a constant. The total collector current I_C is given by I_{CR} plus the fraction of I_{EF} that reaches the collector (allowing for recombination and reverse emitter injection). Thus,

$$I_C = \alpha_F I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) - I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-175}$$

where α_F has been defined previously by Eq. (2-162). Similarly, the total emitter current is composed of I_{EF} plus the fraction of I_{CR} that reaches the emitter with the transistor acting in an inverted mode. Thus,

$$I_E = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + \alpha_R I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-176}$$

where α_R is the ratio of emitter to collector current with the transistor operating *inverted* (i.e., with the collector–base junction forward biased and emitting carriers into the base and the emitter–base junction reverse biased and collecting carriers). Typical values of α_R are 0.5–0.8. An inverse current gain β_R is also defined,

$$\beta_R = \frac{\alpha_R}{1 - \alpha_R} \tag{2-177}$$

and has typical values 1–5. This is the current gain of the transistor when operated inverted and is much lower than β_F because the device geometry and doping densities are designed to maximize β_F. The inverse-active region of device operation occurs for V_{CE} negative in an *npn* transistor and is shown in Figure 2-89. In order to display these characteristics adequately in the same figure as the forward-active region, the negative voltage and current scales have been expanded. The inverse-active mode of operation is rarely encountered in analog circuits.

Equations (2-175) and (2-176) describe *npn* transistor operation in the saturation region when V_{BE} and V_{BC} are both positive, and also in the forward-active and inverse-active regions. These equations are from the *Ebers–Moll* equations. In the forward-active region, they degenerate into a form similar to that of Eqs. (2-146), (2-158), and (2-160) derived earlier. This can be shown by putting V_{BE} positive and V_{BC} negative in Eqs. (2-175) and (2-176) to obtain

$$I_C = \alpha_F I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + I_{CS} \tag{2-178}$$

$$I_E = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) - \alpha_R I_{CS} \tag{2-179}$$

Equation (2-178) is similar in form to Eq. (2-146) except that leakage currents that were previously neglected have now been included. This minor difference is significant only at high temperatures or very low operating currents. Comparison of Eq. (2-178) with Eq. (2-146) allows us to identify $I_S = \alpha_F I_{ES}$ and it can be shown [6] in general that

$$\alpha_F I_{ES} = \alpha_R I_{CS} = I_S \tag{2-180}$$

where this expression represents a reciprocity condition. Using Eq. (2-180) in Eqs. (2-175) and (2-176) allows the Ebers–Moll equations to be expressed in the general form

$$I_C = I_S\left(\exp\frac{V_{BE}}{V_T} - 1\right) - \frac{I_S}{\alpha_R}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-181}$$

$$I_E = -\frac{I_S}{\alpha_F}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + I_S\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-182}$$

This form is often used for computer representation of transistor large-signal behavior.

The effect of leakage currents mentioned above can be further illustrated as follows. In the forward-active region, we have, from Eq. (2-179),

$$I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) = -I_E - \alpha_R I_{CS} \tag{2-183}$$

Substitution of Eq. (2-183) in Eq. (2-178) gives

$$I_C = -\alpha_F I_E + I_{CO} \tag{2-184}$$

where

$$I_{CO} = I_{CS}(1 - \alpha_R\alpha_F) \tag{2-185}$$

and I_{CO} is the collector–base leakage current with the emitter open. Although I_{CO} is given theoretically by Eq. (2-185), in practice, surface leakage effects dominate when the collector–base junction is reverse biased and I_{CO} is typically several orders of magnitude larger than the value given by Eq. (2-185). However, Eq. (2-184) is still valid if the appropriate measured value for I_{CO} is used. Typical values of I_{CO} are 10^{-10} to 10^{-12} A at 25 °C, and the magnitude doubles about every 8 °C. As a consequence, these leakage terms can become very significant at high temperatures. For example, consider the base current I_B. This is

$$I_B = -(I_C + I_E) \tag{2-186}$$

If I_E is calculated from Eq. (2-184) and substituted in Eq. (2-186), the result is

$$I_B = \frac{1 - \alpha_F}{\alpha_F} I_C - \frac{I_{CO}}{\alpha_F} \tag{2-187}$$

But from Eq. (2-161)

$$\beta_f = \frac{\alpha_F}{1 - \alpha_F} \tag{2-188}$$

and use of Eq. (2-188) in Eq. (2-187) gives

$$I_B = \frac{I_C}{\beta_F} - \frac{I_{CO}}{\alpha_F} \tag{2-189}$$

Since the two terms in Eq. (2-189) have opposite signs, the effect of I_{CO} is to *decrease* the magnitude of the external base current at a given value of collector current.

2-2-6 Small-Signal Models of Bipolar Transistors

Figure 2-91 shows the small-signal equivalent BJT model. It consists of an intrinsic model and a package model. Table 2-12 lists its keywords.

In order to better understand the difference between the various technologies, we show Table 2-13, which allows the reader to compare their performances.

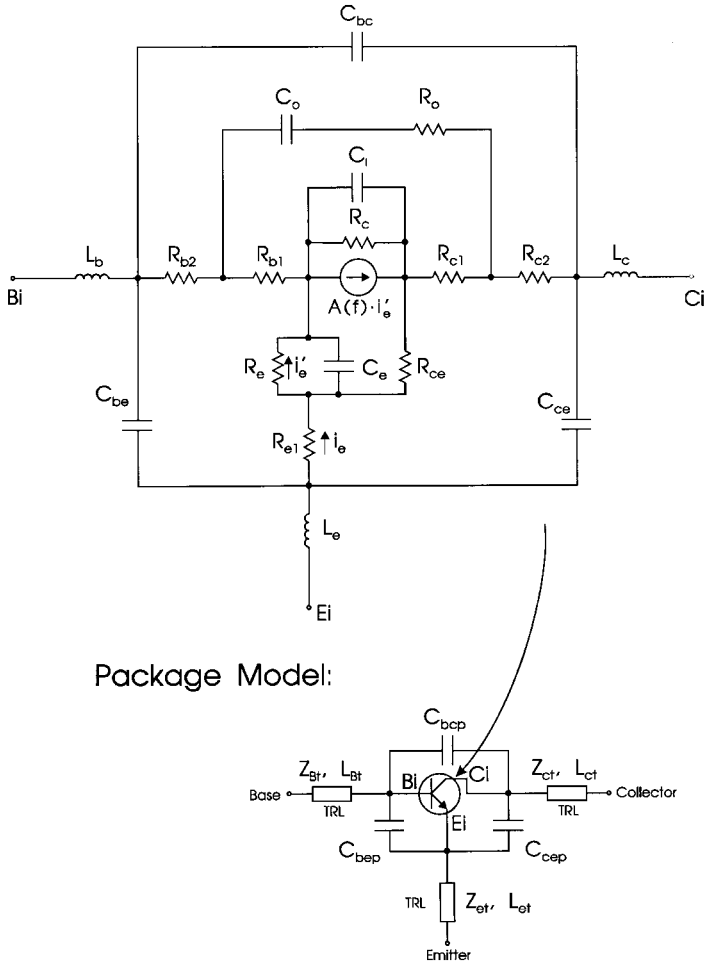


Figure 2-91 Linear BJT model.

Table 2-12 Small-signal equivalent BJT model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
A	Ratio of I_C to I_E at dc		0.0
RE	Emitter resistance	ohm	0.0
F	Current generator roll-off frequency	hertz	∞
T	Time delay	second	0.0
CE	Emitter capacitance	farad	0.0
CI	Collector capacitance	farad	0.0
RCE	Collector–emitter resistance	ohm	∞
RC	Collector resistance	ohm	∞
RO	Extrinsic base–collector resistance	ohm	0.0
CO	Extrinsic base–collector capacitance	farad	0.0
RB1	Intrinsic base resistance (R_{bb})	ohm	0.0
RC1	Parasitic collector resistance	ohm	0.0
RE1	Parasitic emitter resistance	ohm	0.0
RB2	Parasitic base resistance	ohm	0.0
RC2	Parasitic collector resistance	ohm	0.0
CBE	Base-to-emitter package capacitance	farad	0.0
CBC	Base-to-collector package capacitance	farad	0.0
CCE	Collector-to-emitter package capacitance	farad	0.0
LB	Base-lead inductance	henry	0.0
LC	Collector-lead inductance	henry	0.0
LE	Emitter-lead inductance	henry	0.0
TJ	Chip temperature	Kelvin	298
NFAC	Noise factor proportional to drive		1.0
FC	Flicker noise (1/*f* noise) corner frequency	hertz	50
	Package Model		
CBCP	Base-to-collector package capacitance	farad	0.0
CBEP	Base-to-emitter package capacitance	farad	0.0
CCEP	Collector-to-emitter package capacitance	farad	0.0
ZBT	Base transmission line impedance	ohm	50
ZCT	Collector transmission line impedance	ohm	50
ZET	Emitter transmission line impedance	ohm	50
LBT	Base transmission line length at $\varepsilon_r = 1$	meter	0.0
LCT	Collector transmission line length at $\varepsilon_r = 1$	meter	0.0
LET	Emitter transmission line length at $\varepsilon_r = 1$	meter	0.0

Notes:

1. $A \equiv \alpha = I_c/I_e$; β = dc current gain = $\alpha(1 - \alpha)$.
2. The bipolar current gain in this model is described by

$$\mathrm{A} = \mathrm{A}(0) = \frac{e^{-j\omega\mathrm{T}}}{1 + j\,(f/F)}$$

where $\omega = 2\,\pi f$ and f is frequency.

3. The current source is controlled by the current through R_e. The current generator has a cutoff frequency with respect to the total emitter current, I_E:

$$\mathrm{F} = g_m/2\pi C_e$$

where $g_m = 1/R_e$. This frequency becomes infinite for the default value for C_e (0.0). The parammeter F specifies the frequency roll-off for the current generator with respect to the current through R_e. Effectively, this frequency parameter may be used to model additional delays in the device.

Table 2-13 Overview of current BJT technologies including BiCMOS

Parameter	MOSAIC 3	MOSAIC 5	MOSAIC 5SE	RFBICMOS1	TFSOI
Emitter Size	$0.95 \times 3.2\ \mu m^2$	$0.3 \times 0.9\ \mu m^2$	$0.3 \times 0.9\ \mu m^2$	$0.4 \times 1.3\ \mu m^2$	$0.1 \times 1.2\ \mu m^2$
β	150	150	90	80	40–65
F_t (GHz)	14@ $V_{ce} = 1$ V	14@ $V_{ce} = 1$ V	27@ $V_{ce} = 1$ V	16@ $V_{ce} = 1$ V	8@ $V_{ce} = 2$ V
C_{je} (fF)	19.9	4.3	3.8	5.1	1.5
C_{jc} (fF)	25.1	2.8	2.8	3.7	0.34
C_{js} (fF)	49.6	4.3	3.6	7.4	0.95
R_e (Ω)	22	36	50	34	72
R_b (Ω)	462	578	583	462	n/a
BV_{ceo} (V)	6.5	7.0	4.8	6.0	>3.5
BV_{cbo} (V)	20	20	15	18	>10
BV_{ebo} (V)	5.2	4.2	3.8	6.0	4.3
V_A (V)	30	48	20	>25	25–40
Minimum T_{pd}	75 ps @ 1.5 mA	35 ps @ 800 μA	36 ps @ 800 μA	43 ps @ 300 μA	52 ps @ 25 μA

Parameter	TFSOI Ring Device	RFBICMOS3 Analog	RFBICMOS3 Digital	BiGCMOS3 Analog	BiGCMOS3 Digital
Emitter Size	$0.1 \times 7.2\ \mu m^2$	$0.25 \times 0.75\ \mu m^2$	$0.25 \times 0.75\ \mu m^2$	$0.2 \times 0.6\ \mu m^2$	$0.2 \times 0.6\ \mu m^2$
β	40–65	100	150	100	100
F_t (GHz)	8 @ $V_{ce} = 2$ V	15 @ $V_{ce} = 2$ V	25 @ $V_{ce} = 2$ V	15 @ $V_{ce} = 1$ V	35 @ $V_{ce} = 1$ V
C_{je} (fF)	3.4	1.8	1.8	1.7	1.7
C_{jc} (fF)	2.8	1.1	1.1	0.2	0.3
C_{js} (fF)	2.4	12.0	12.0	4	4
R_e (Ω)	26	140	140	113	113
R_b (Ω)	790	n/a	n/a	10 kΩ/sq.	15 kΩ/sq.
BV_{ceo} (V)	>3.5	6.5	4	6.5	4.8
BV_{cbo} (V)	>10	15	15	20	15
BV_{ebo} (V)	4.3	4	4	4	4
V_A (V)	30–40	40	25	40	20
Minimum T_{pd}	75 ps @ 100 μA	n/a	n/a	35 ps @ 50 μA	22 ps @ 200 μA

Parameter	AT&T BEST2[a]	Fujitsu POSET	HP HP-25	HP ISOSAT II[b]	Hitachi SMI
Emitter Size	$1.0 \times 2.0\ \mu m^2$	$0.1 \times 1.7\ \mu m^2$	$0.4 \times 20\ \mu m^2$	$0.6 \times ?\ \mu m^2$	$0.2 \times 2.0\ \mu m^2$
β	100	110	75	100	500
F_t (GHz)	23 @ $V_{ce} = 3$ V	38 @ $V_{ce} = 3$ V	28.5 @ $V_{ce} = 3$ V	14/F_{max} > 20 GHz	82@ $V_{ce} = 2$ V
C_{je} (fF)	5.0	3.5	42	n/a	2.9
C_{jc} (fF)	5.0	2.2	37	n/a	4.8
C_{js} (fF)	8.0	n/a	123	n/a	4.1
R_e (Ω)	25	25	n/a	n/a	30
R_b (Ω)	10 kΩ/sq.	960	100	n/a	240
BV_{ceo} (V)	5.5	5.4	5.2	15	2.5
BV_{cbo} (V)	11	18.2	15	25	6.4
BV_{ebo} (V)	8.0	4.5	>5	2.0	6.2
V_A (V)	24	n/a	n/a	25	8

(continued)

Table 2-13 (Continued)

Parameter	AT&T BEST2[a]	Fujitsu POSET	HP HP-25	HP ISOSAT II[b]	Hitachi SMI
Minimum T_{pd}	37 ps @ 2.1 mA 48 ps @ 600 µA	21.5 ps @ 320 µA	n/a	n/a	16 ps @ $I_g = 1.3$

Parameter	IBM SiGe[c]	Maxim GST-1[d]	Maxim GST-2[e]	Mitsubishi[f]	National ABIC V
Emitter Size	0.5×1.0 µm^2	1.2×2.6 µm^2	0.7×1.5 µm^2	0.5×5.7 µm^2	0.5×0.5 µm^2
β	100	75	190	90	85
F_t (GHz)	41 @ $V_{ce} = 1$ V	12.5	26 @ $V_{ce} = 4$ V	20	13 @ $V_{ce} = 2$ V
C_{je} (fF)	5.2	9	4	n/a	1.9
C_{jc} (fF)	3.4	15	7	9	3.7
C_{js} (fF)	3.6	11	9	n/a	9.6
R_e (Ω)	n/a	n/a	80	n/a	n/a
R_b (Ω)	300	600	710	155	383
BV_{ceo} (V)	3.3	>5	>4	6.2	8.5
BV_{cbo} (V)	9.5	>12	>14	n/a	20
BV_{ebo} (V)	n/a	>4	>4	n/a	7.5
V_A (V)	n/a	20	15	n/a	n/a
Minimum T_{pd}	n/a	n/a	n/a	n/a	38 ps @ 1.8 mA

Parameter	NEC BSA	NEC SiGe[g]	NEC NESAT III	Philips QBiC2	Philips PRET[g]
Emitter	0.4×7.6 µm^2	0.2×1.6 µm^2	0.6×40 µm^2	0.5×1.2 µm^2	0.2×0.9 µm^2
β	n/a	120	80	100	50
F_t (GHz)	43	51	20 @ $V_{ce} = 3$ V	17	14
C_{je} (fF)	20	n/a	120	2.4	1.5
C_{jc} (fF)	16	1.7	140	3.5	1.0
C_{js} (fF)	n/a	n/a	160	6.5	5.0
R_e (Ω)	n/a	80	n/a	120	120
R_b (Ω)	43	220	55	700 Ω/sq.	1 kΩ/sq.
BV_{ceo} (V)	3.3	2.7	6	4.5	7
BV_{cbo} (V)	n/a	n/a	17	13	12.5
BV_{ebo} (V)	2	2.2	4.5	5.5	5.5
V_A (V)	n/a	n/a	n/a	15	20
Minimum T_{pd}	29 ps @ 300 µA	19 ps @ 300 µA	n/a	n/a	n/a

Parameter	Philips SiGe[g]	Siemens B6HF	Siemens B6HF Plus	Siemens SiGe[g]	SONY[h]
Emitter Size	0.4×1.4 µm^2	0.8×3.3 (drawn) µm^2	0.25×2.7 µm^2	0.27×2.5 µm^2	$0.2 \times ?$ µm^2
β	100	>60	115	130	100
F_t (GHz)	45 @ $V_{ce} = 1.5$ V	25 @ $V_{cb} = 0$ V	50 @ $V_{cb} = 2$ V	61 @ $V_{cb} = 1$ V	30
C_{je} (fF)	9.1	11	8.5	7.5	n/a
C_{jc} (fF)	4.6	6.5	8.0	5.5	n/a
C_{js} (fF)	10.8	31	26.5	13	n/a

(continued)

Table 2-13 Overview of current BJT technologies including BiCMOS (Continued)

Parameter	Phillips SiGe[g]	Siemens B6HF	Siemens B6HF Plus	Siemens SiGe[g]	SONY[h]
R_e (Ω)	n/a	n/a	n/a	n/a	n/a
R_b (Ω)	n/a	165	14 kΩ/sq.	8 kΩ/sq.	n/a
BV_{ceo} (V)	3.2	>3.5	7	2.5	4.0
BV_{cbo} (V)	11	n/a	10.8	9.9	18
BV_{ebo} (V)	3.5	n/a	2.0	3.0	3.5
V_A (V)	70–100	n/a	20	130	10
Minimum T_{pd}	14 ps @ 6 mA	25 ps @ 1.2 mA (CML)	16 ps @ 2.0 mA (CML)	11 ps @ 1.9 mA (CML)	21 ps @ I_g = 3.5 mA (CML)

Parameter	STM HSB[i]	STM HSB2A	iSTM HSB3[g,j]	Toshiba	Toshiba Epi Base[g]
Emitter Size	$0.35 \times ? \ \mu m^2$	$0.35 \times ? \ \mu m^2$	$0.2 \times ? \ \mu m^2$	$0.4 \times 5.0 \ \mu m^2$	$0.6 \times 5.0 \ \mu m^2$
β	>50	>50	>50	60	105
F_t (GHz)	n/a	20	~50	22 @ V_{ce} = 2 V	52 @ V_{ce} = 3 V
C_{je} (fF)	n/a	n/a	n/a	9.0	16.1
C_{jc} (fF)	n/a	n/a	n/a	3.6	14.1
C_{js} (fF)	n/a	n/a	n/a	11	n/a
R_e (Ω)	n/a	n/a	n/a	14	22.7
R_b (Ω)	10 kΩ/sq.	n/a	n/a	90	205
BV_{ceo} (V)	4.5	n/a	n/a	6.2	4.1
BV_{cbo} (V)	n/a	n/a	n/a	14.8	9.5
BV_{ebo} (V)	n/a	n/a	n/a	4.2	n/a
V_A (V)	n/a	>30	n/a	n/a	26.5
Minimum T_{pd}	n/a	30 ps	n/a	19 ps @ 1.2 mA	n/a

[a] 1.0 µm Design Rule; NF = 3.25 dB, G_A = 12 dB at I_C = 2 mA, freq. = 2 GHz for $16 \times 1 \times 2 \ \mu m^2$; Resistors: 150 Ω/sq. (mono), 300 Ω/sq. (mono), 500 Ω/sq. (poly); Layers of metal, Schottky diodes, linear capacitor (silicided n^+ poly/n^+Si); Gate length = 0.68 µm (*n*-ch)/0.8 µm (*p*-ch), Gate Oxide = 12.5 nm (3 V)/15 nm (5 V), V_{th} = 0.65 V (*n*-ch)/–0.9 V (*p*-ch), Gm (µS/µm) = 204 (*n*-ch)/65 (*p*-ch).

[b] 0.6 µm Design Rule; Gold metallization, ILD: polyimide; Inductors: $Q > 10$ for $L \leq 10$ nH; Capacitors: $Q > 25$ for $C \leq 20$ pF; Schottky diodes (f_c = 100 GHz), Zener diodes, MIS capacitor, thin film resistors, and LPNP, 10 mask process.

[c] Analog devices using this process.

[d] LPNP F_t = 80 MHz; three-layer metallization; Schottky diodes (R_sC_j ~ 4 ps); Implanted resistors: 200 Ω/sq.; Poly resistors: 400/2000 Ω/sq.; NiCr thin film resistors (50 Ω/sq.) with laser trim option; MOS capacitor.

[e] LPNP F_t = 85 MHz; two-layer metallization; Schottky diodes (R_sC_j ~ 0.8 ps); Implanted resistors: 170/600/1200 Ω/sq.; NiCr thin film resistors (50 Ω/sq.) with laser trim option; MOS capacitor.

[f] L_n = 0.8 µm, T_{ox} = 18 nm, V_{th} = 0.76 V (nmos), V_{th} = –0.78 V (pmos), 2poly-MIM capacitor: 2.7 fF/µm^2, Schottkys, inductors.

[g] Not known if this process is in production.

[h] 0.8 µm Design Rules; 2.5 fF/µm^2 MIS capacitor; 600 MHz PJFET; 12L; F_t = 4 GHz VPNP; poly resistors 80/400/2000 Ω/sq.

[i] Targeted for 900-MHz GSM radio front-end blocks, a CT2 cordless single-chip radio, and DECT (digital enhanced cordless telephones).

[j] Targeted for multimode 900-MHz GSM, 1.8-GHz PCN and 1.9-GHz PCS based systems. Company is also evaluating integrating SiGe in HSB3.

2-3 FIELD-EFFECT TRANSISTORS

Both diodes and *pnp/npn* transistors work on the injection principle, which means the base–emitter junction generates either free electrons or other carriers, which are being "collected" by the collector connection and the base connection controls it. While this is a fairly complex process, its physics is well understood, and its latest spin-offs, heterojunction BJTs and silicon–germanium transistors, seem to offer a bright future.

In opposition to this, the field-effect transistor (FET) controls the electronic conduction in a solid by an electric field, and this concept actually predates the invention of the bipolar transistor. J. E. Lilienfeld filed for a patent on such a device in 1925 (U.S. Patent No. 1,745,175), and W. Shockley presented a comprehensive theory of the FET in 1952. It took until 1960 before the first commercially available FETs came on the market.

Several types of semiconductor materials have been used for making FETs: silicon, germanium, gallium arsenide, and others have been used. By far the most widely used is silicon, followed now by gallium arsenide.

The FET is a class of electronic semiconductor device in which the conduction of a "channel" between source (*S*) and drain (*D*) terminals is controlled by an electric field impressed on the channel via a gate (*G*) terminal. The conduction channel may utilize *n*-type carriers (electrons) or *p*-type carriers (holes). The electric field that controls the channel conduction may be introduced via a *pn* junction [for a "junction" FET (JFET)], a metal plate separated from the semiconductor channel by an oxide dielectric [for a metal-oxide semiconductor (MOS) FET], or a combination of the two methods. The polarity of the controlling electrical field is a function of the type of carriers in the channel. A FET "family tree" is shown in Figure 2-92.

For today's microwave and RF integrated circuits, the JFET with its high-frequency problems (low cutoff frequency) and large variation in parameters, such as transconductance and pinchoff voltage, has fallen out of favor and is mainly used in discrete circuits such as two-way radios or other RF applications that tolerate these wide variations and temperature dependence. The most-used FET types are MOS for small-signal applications, LDMOS (lateral diffused MOS) for power applications, and GaAs MESFET for lower-power applications (low-voltage applications) as well as high-power applications (recent advertisements by NEC have shown output power up to 100 W). Additional information on the FET family is shown in Figure 2-93.

All members of the FET family are high-impedance-input devices. Figure 2-94 shows the key operating parameters for MOSFETs and MESFETs in low-noise stages, high-gain stages, and linearized power stages. The linearization also improves the matching. This figure is similar to its counterpart in Section 2-2-2.

The following several pages present a typical example of a datasheet* for a JFET (U310) and the chip (NZA) on which it is based, followed by an example of a device with a much higher dynamic range (the CP640 family). The CP640 family and devices similar to them are being replaced by low-power LDMOS, which are not yet well-documented. At least one company, InterFET, is devoted entirely to producing discrete JFETs and JFET-related hybrid and small-scale-integration IC products.

* Reproduced with permission of Vishay-Siliconix.

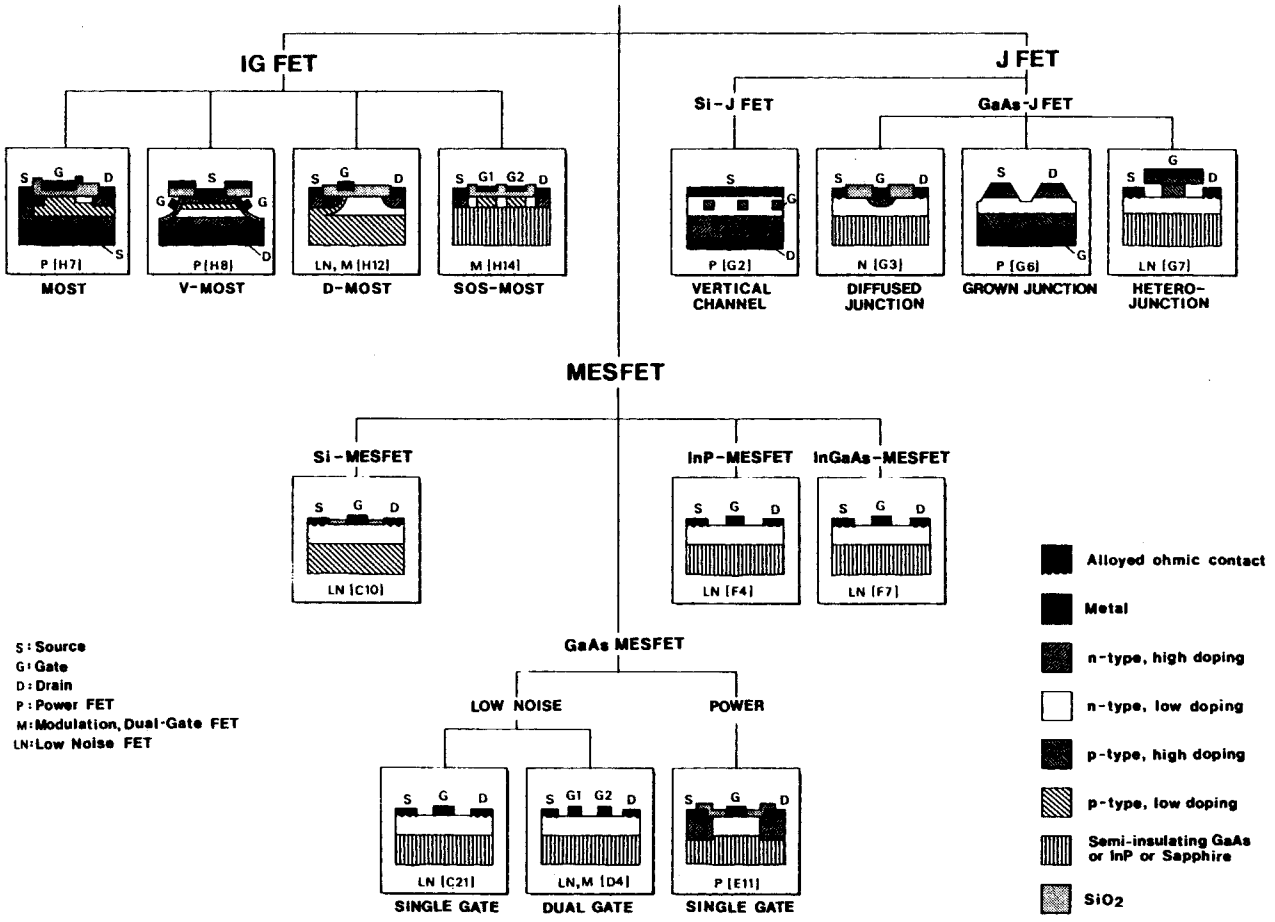


Figure 2-92 "Family tree" of microwave FET types showing cross sections of their structures [7]. Copyright ©1976 IEEE.

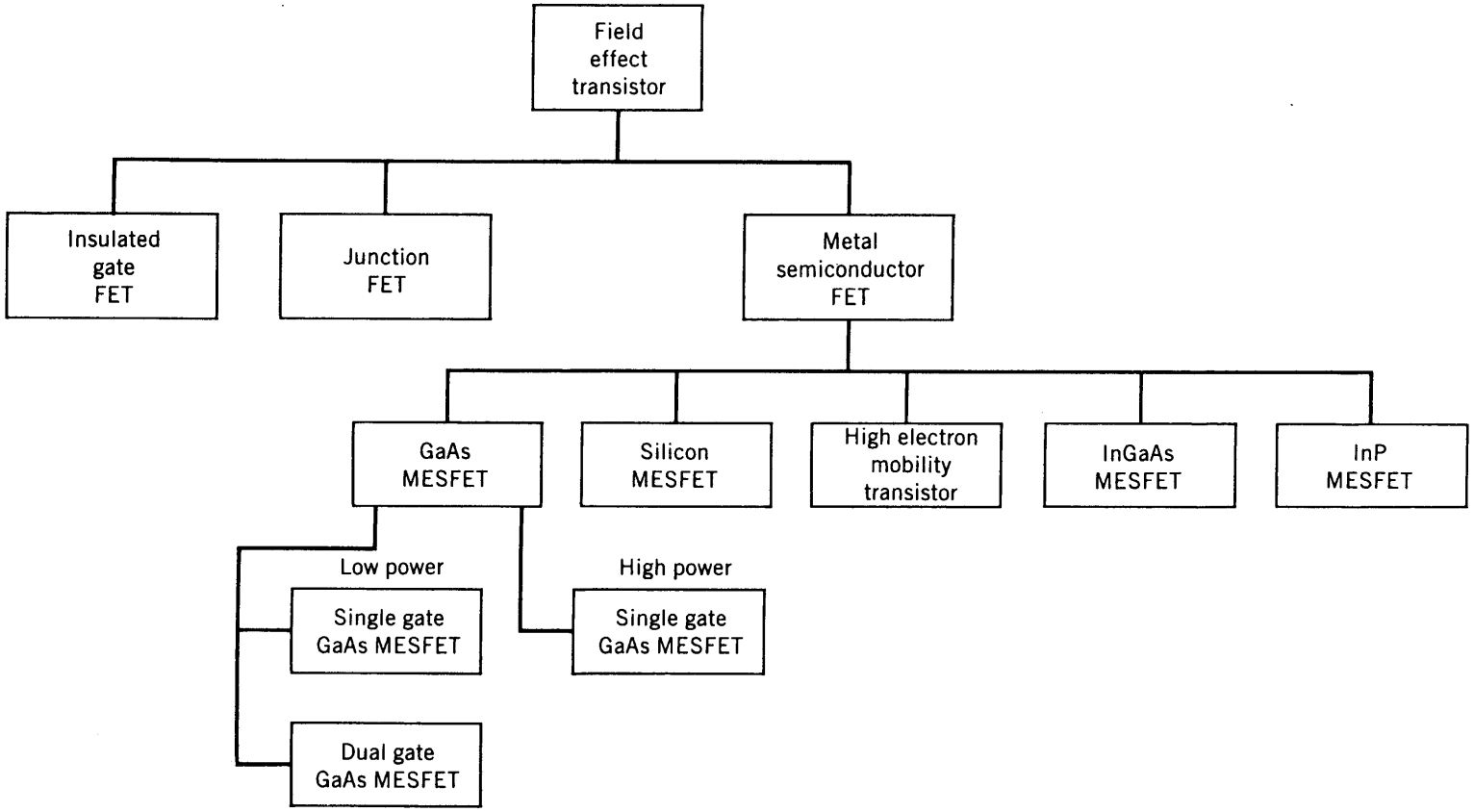


Figure 2-93 Another FET family tree [8].

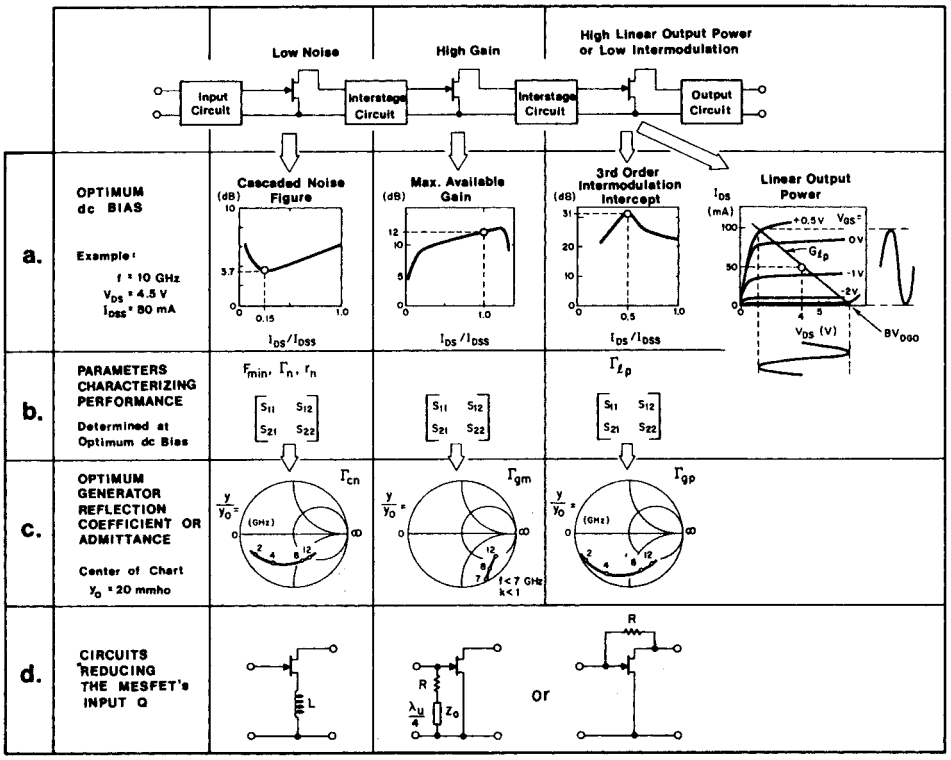


Figure 2-94 Key parameters in operating a MESFET in a low-noise front-end stage, a high-gain intermediate stage, and a linear power stage, showing (a) the device characteristics that lead to optimum dc bias for each service, (b) the parameters that characterize the device performance, (c) the optimum generator reflection coefficient that must be synthesized by each circuit, and (d) circuit arrangements that reduce the transistor's input *Q* or stabilize it at low frequencies. Although this example is based on an HP MESFET with a gate length of 1 μm and a gate width of 500 μm, this principle applies to all members of the FET family [7]. Copyright ©1976 IEEE.

n-channel JFETs *designed for . . .*

Siliconix

Performance Curves NZA See Section 4

- **VHF Amplifiers**
- **Front End High Sensitivity Amplifiers**
- **Oscillators**
- **Mixers**

BENEFITS

- Industry Standard
- High Power Gain
  16 dB at 105 MHz, Common-Gate
  11 dB at 450 MHz, Common-Gate
- Low Noise
  2.7 dB Noise Figure at 450 MHz
- Wide Dynamic Range
  Greater than 100 dB
- 75 Ω Input Match Common Gate

ABSOLUTE MAXIMUM RATINGS (25°C)

Gate-Drain or Gate-Source Voltage . . . . . -25 V
Gate Current . . . . . 20 mA
Total Power Dissipation at T_A = 25°C . . . . . 500 mW
Power Derating to 150°C . . . . . 4.0 mW/°C
Storage Temperature Range . . . . . -65 to +200°C
Lead Temperature
(1/16" from case for 10 seconds) . . . . . 300°C

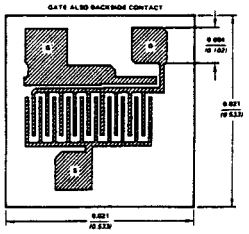
TO-52
See Section 6

ELECTRICAL CHARACTERISTICS (25°C unless otherwise noted)

		Characteristic		U308 Min	U308 Typ	U308 Max	U309 Min	U309 Typ	U309 Max	U310 Min	U310 Typ	U310 Max	Unit	Test Conditions	
1	STATIC	I_{GSS}	Gate Reverse Current			-150			-150			-150	pA	V_{GS} = -15 V,	
2						-150			-150			-150	nA	V_{GS} = 0	T_A = 125°C
3		BV_{GSS}	Gate-Source Breakdown Voltage	-25			-25			-25			V	I_G = -1 μA, V_{DS} = 0	
4		$V_{GS(off)}$	Gate-Source Cutoff Voltage	-1.0		-6.0	-1.0		-4.0	-2.5		-6.0		V_{DS} = 10 V, I_D = 1 nA	
5		I_{DSS}	Saturation Drain Current (Note 1)	12		60	12		30	24		60	mA	V_{DS} = 10 V, V_{GS} = 0	
6		$V_{GS(f)}$	Gate-Source Forward Voltage			1.0			1.0			1.0	V	I_G = 10 mA, V_{DS} = 0	
7	DYNAMIC	g_{fg}	Common-Gate Forward Transconductance (Note 1)	10	17		10	17		10	17		mmho	V_{DS} = 10 V, I_D = 10 mA	f = 1 kHz
8		g_{og}	Common-Gate Output Conductance			250			250			250	μmho		
9		C_{gd}	Drain-Gate Capacitance			2.5			2.5			2.5	pF	V_{GS} = -10 V, V_{DS} = 10 V	f = 1 MHz
10		C_{gs}	Gate-Source Capacitance			5.0			5.0			5.0			
11		$\overline{e_n}$	Equivalent Short Circuit Input Noise Voltage		10			10			10		$\frac{nV}{\sqrt{Hz}}$	V_{DS} = 10 V, I_D = 10 mA	f = 100 Hz
12	HI FREQ	g_{fg}	Common-Gate Forward Transconductance		15			15			15		mmho	V_{DS} = 10 V, I_D = 10 mA	f = 105 MHz
13					14			14			14				f = 450 MHz
14		g_{og}	Common-Gate Output Conductance		0.18			0.18			0.18				f = 105 MHz
15					0.32			0.32			0.32				f = 450 MHz
16		G_{pg}	Common-Gate Power Gain (Note 2)	14	16		14	16		14	16		dB		f = 105 MHz
17				10	11		10	11		10	11				f = 450 MHz
18		NF	Noise Figure		1.5	2.0		1.5	2.0		1.5	2.0			f = 105 MHz
19					2.7	3.5		2.7	3.5		2.7	3.5			f = 450 MHz

NOTES:
1. Pulse test duration = 2 ms.
2. Gain (G_{pg}) measured at optimum input noise match.

NZA



n-channel JFET
designed for . . .

- VHF/UHF Amplifiers
- Front End High Sensitivity Amplifiers
- Oscillators
- Mixers

BENEFITS

- Industry Standard
- High Power Gain
  16 dB at 100 MHz, Common Gate
  11 dB at 450 MHz, Common Gate

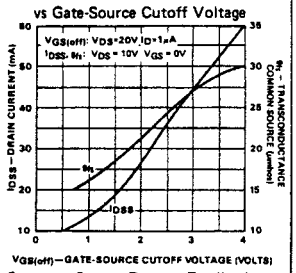
Siliconix

BENEFITS

- Low Noise
  3 dB Noise Figure at 450 MHz
- Wide Dynamic Range
  Greater Than 100 dB
- 75 Ohm Input Match Common Gate

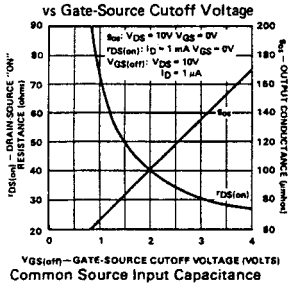
TYPE	PACKAGE	PRINCIPAL DEVICES
Single	TO-52	U308-10
Single	TO-72	U311
Single	TO-92	J308-10
Dual	TO-99	U430-1
Single	Chip	J308CHP-10CHP, U308CHP-10CHP, U311CHP
Dual	Chip	U430CHP-1CHP

Drain Current & Transconductance vs Gate-Source Cutoff Voltage

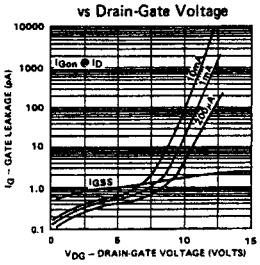


PERFORMANCE CURVES (25°C unless otherwise noted)

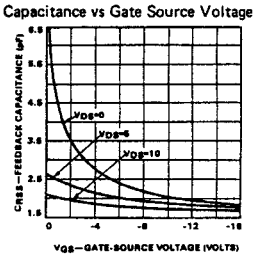
On Resistance & Output Conductance vs Gate-Source Cutoff Voltage



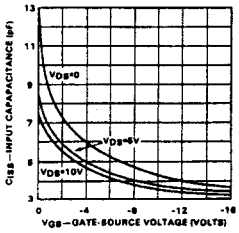
Gate Operating Current vs Drain-Gate Voltage



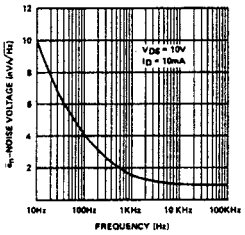
Common Source Reverse Feedback Capacitance vs Gate Source Voltage



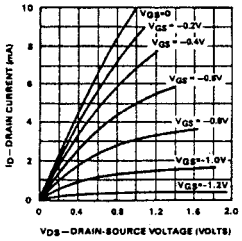
Common Source Input Capacitance vs Gate-Source Voltage



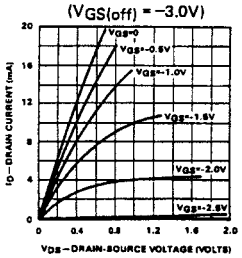
Noise Voltage vs Frequency



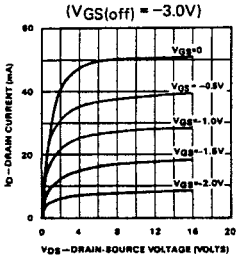
Output Characteristic ($V_{GS(off)} = -1.7V$)



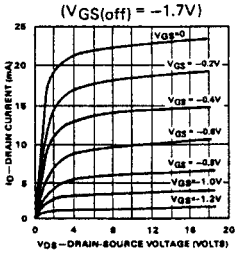
Output Characteristic ($V_{GS(off)} = -3.0V$)



Output Characteristic ($V_{GS(off)} = -3.0V$)



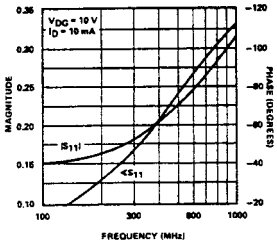
Output Characteristic ($V_{GS(off)} = -1.7V$)



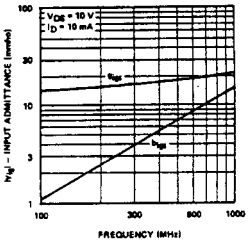
NZA

PERFORMANCE CURVES (Con't) (25°C unless otherwise noted)

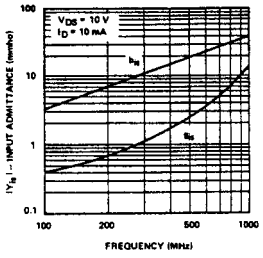
Forward Reflection Coefficient Common Gate



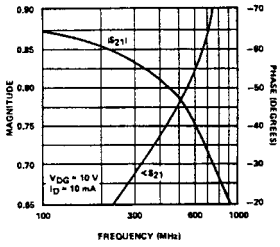
Input Admittance Common Gate



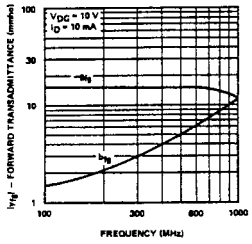
Input Admittance Common Source



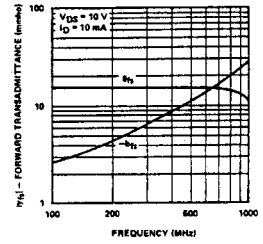
Forward Transmission Coefficient Common Gate



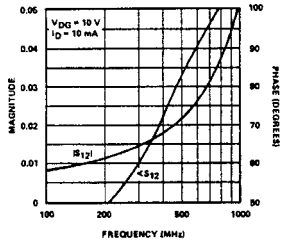
Forward Transfer Admittance Common Gate



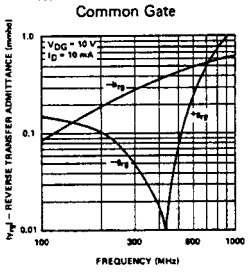
Forward Transfer Admittance Common Source



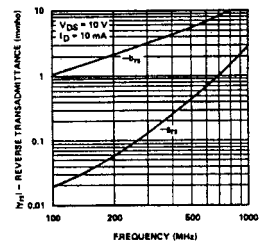
Reverse Transmission Coefficient Common Gate



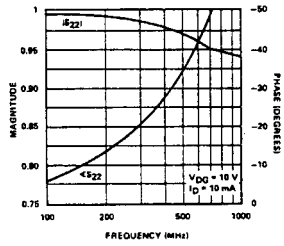
Reverse Transfer Admittance Common Gate



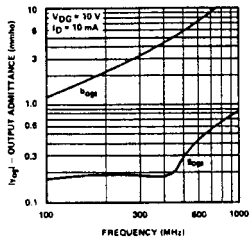
Reverse Transfer Admittance Common Source



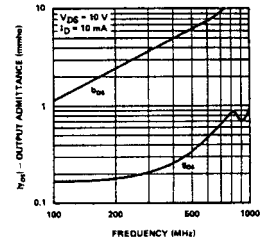
Reverse Reflection Coefficient Common Gate



Output Admittance Common Gate



Output Admittance Common Source



BROADBAND RF FET
SILICON EPITAXIAL JUNCTION N-CHANNEL FIELD EFFECT TRANSISTOR

HIGH DYNAMIC RANGE HF AND VHF AMPLIFIER
FOR USE IN COMMON GATE CONFIGURATION

- USABLE TO OVER 300 MHz
- 50 Ohm VSWR < 1.5:1 0.5-50 MHz (FIG. 1)
- LOW NOISE FIGURE — 2.2 dB TYPICAL @ 50 MHz
- INPUT Z CONSTANT 0.5-50 MHz
- HIGH IM INTERCEPT POINT — > + 40 dBm
- HIGH TRANSCONDUCTANCE — 100,000 μmhos (TYP.)
- 1 dB COMPRESSION POINT > + 20 dBm
- DYNAMIC RANGE > 140 dB (TO 1 dB COMPRESSION)
- HIGH VOLTAGE—TO 50 V.

CP640
CP664
CP665
CP666

ELECTRICAL DATA ABSOLUTE MAXIMUM RATINGS

PARAMETER	SYMBOL	CP 640	CP 664	CP 665	CP 666	UNITS
Drain to Source Voltage	BVDSO	20	30	40	50	Volts
Drain to Gate Voltage	BVDGO	20	30	40	50	Volts
Gate to Source Voltage	BVGSO	−15	−20	−20	−20	Volts
Peak Drain Current	ID	1.2	1.2	1.2	1.2	Amps
Power Dissipation 25 °C CASE	PD	8.0	8.0	8.0	8.0	Watts
Derating Factor (slope)	DF	22	22	22	22	°C/W
Junction Temp.(Oper. & Store)	TJ	−55 °C to +200 °C				

TYPICAL TWO TONE 3rd ORDER IM PRODUCTS — CIRCUIT FIGURE 1

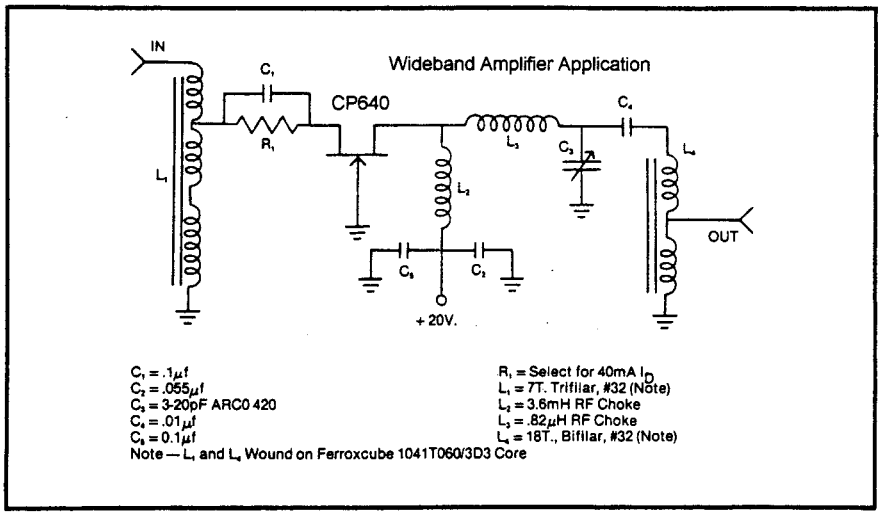
Tones at 3MHz/5MHz

Signal Level EMF	3rd Order Product
1 Volt	-44 dB
0.3 Volt	-75 dB
0.25 Volt (0dBM)	-80 dB

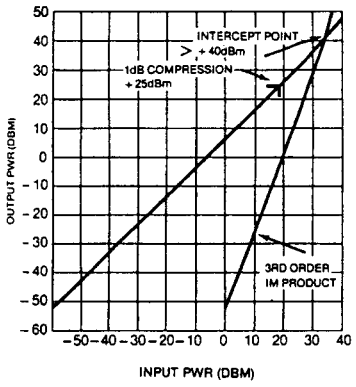
ELECTRICAL CHARACTERISTICS: T_{CASE} = 25 °C (UNLESS OTHERWISE STATED)

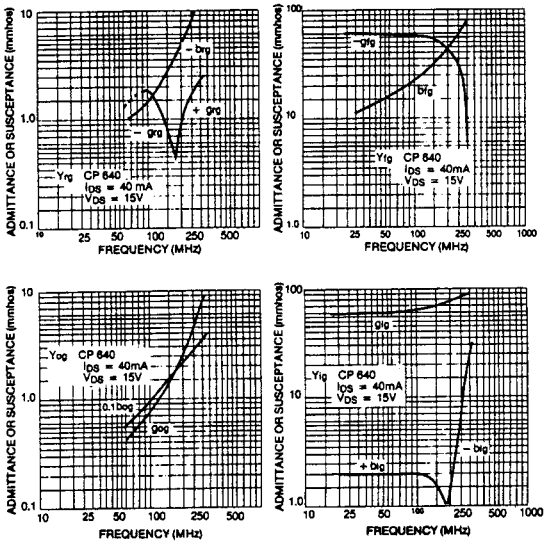
PARAMETERS	CONDITIONS		SYMBOL	Min.	Typ.	Max.	UNITS
Gate Leakage Current	V_{GS} = 15V, V_{DS} = 0	25 °C	I_{GSS}		5	100	nA
		150 °C	I_{GSS}			10	μA
Operating Transconductance	V_{DS} = 15V, I_{DS} = 40 mA		g_{fg}	40	60	80	mmho
Zero Bias Transconductance	V_{DS} = 15V, V_{GS} = 0(1)		g_{fg}	75	100	200	mmho
Gate-Source Cut-Off Voltage	V_{DS} = 5V, I_{DS} = 1.0 mA		$V_{GS(off)}$	2	5	10	Volts
Zero Bias Drain Current	V_{DS} = 15V, V_{GS} = 0 (1)		I_{DSS}	100	400	800	mA
Gate to Source Cap.	V_{GS} = −20V		C_{GS}		15	20	pf
Gate to Drain Cap.	V_{GD} = −20V		C_{GD}		15	20	pf
Power Gain	I_{DS} = 40mA, f = 50MHz, Fig. 1		Gpg	8	8.5	9.5	dB
Noise Figure	I_{DS} = 40mA, f = 30MHz, Fig. 1		N.F.		2.2	3.0	dB
Voltage Standing Wave Ratio	f = 0.5-50MHz, 50 Ω Source, Fig. 1		VSWR			1.5:1	
Common Gate Input Conductance	f = 0.5-50MHz, V_{DS} = 15, I_D = 40mA		g_{igs}		60		mmho
Common Gate Output Conductance	f = −50MHz, V_{DS} = 15, I_D = 40mA		g_{ogs}		0.4		mmho

(1)Pulse Measurement 1% Duty Cycle 10 mS Max.



TYPICAL INTERCEPT AND COMPRESSION POINT





2-3-1 Large-Signal Behavior of JFETs

A section view of a junction field-effect transistor (JFET) is shown in Figure 2-95. This structure contains, between the source and the drain contacts, an *n*-type "channel" embedded in a *p*-type silicon substrate. If it is assumed that the *pn* junction forms a barrier to current flow, then it can be seen that channel conduction is a function of the channel width, length, and thickness and of the density and mobility of the carriers. In this structure, current can flow equally well in either direction through the channel; that is, the drain can be positive or negative with respect to the source.

Since we have found that the SPICE models for JFETs are really incomplete and give poor answers, we will only briefly touch here on the JFET's performance. The modified Materka model, one of the best GaAsFET models, turns out to be best suited for JFETs because both transistors act similarly in the dc area. Both have a gate-to-source diode that becomes conductive above 0.7 V at the input and they are also quite similar in other respects. Their equivalent circuits are also quite similar, as parameter extractions for both approaches have shown.

In general, in a three-terminal device, the drain current I_D is a function of two variables: V_{DS} and V_{GS}. This function is best represented by families of characteristic curves, as shown in Figure 2-96. These curves are for the "common source" configuration, with the drain as the output and the gate as the input. They reveal that for this device, if V_{DS} is greater than about 2 V (but less than the drain breakdown), I_D is primarily determined by the gate voltage V_{GS}. Under these conditions it is valid for small signals to characterize the FET by a single transfer characteristic curve, commonly called the "forward transconductance curve," such as the upper curve shown in Figure 2-96b. Some of the relationships between the forward transfer curve and the output characteristic curves of Figure 2-96a will be examined. The value of V_{GS} that reduces I_D to approximately zero is the gate–source cutoff voltage, $V_{GS(\text{off})}$. With reference to the output curve, Figure 2-96a, note that the drain current at $V_{GS} = 0$ tends to become saturated at a drain voltage approximately equal in magnitude to $-V_{GS(\text{off})}$. This drain voltage is often referred to as the pinchoff voltage V_P; however, in this section pinchoff voltage is used interchangeably with gate–source cutoff voltage. V_P will have the same meaning as $V_{GS(\text{off})}$. The symbol I_{DSS} is commonly used to indicate the value of saturated drain current at $V_{GS} = 0$.

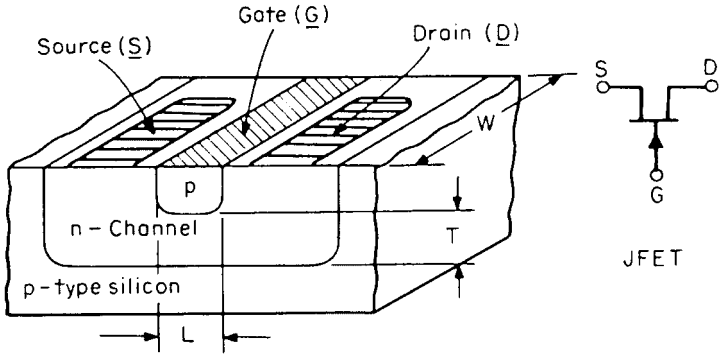


Figure 2-95 Junction field-effect transistor [9].

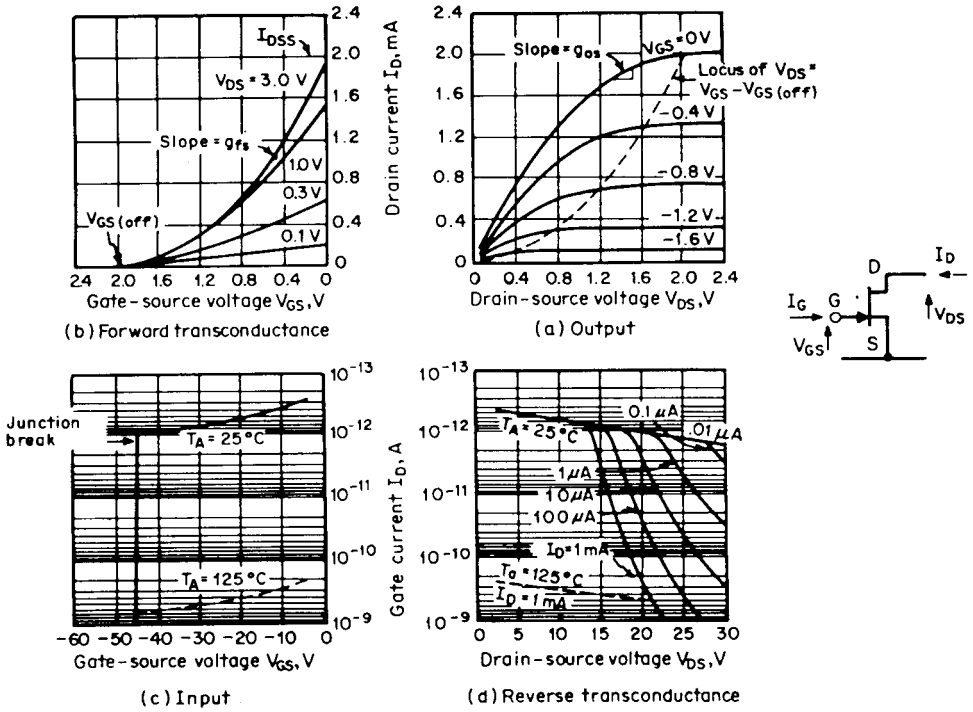


Figure 2-96 Static characteristics of an *n*-channel JFET.

The forward transconductance characteristic of Figure 2-96b can be approximated by a power law relation expressed as

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right)^n \tag{2-190}$$

if $V_{DS} \geq -V_{GS(\text{off})}$. By differentiation the small-signal transconductance g_{fs} is given by

$$g_{fs} = \frac{dI_D}{dV_{GS}} = -n\frac{I_{DSS}}{V_{GS(\text{off})}}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right)^{n-1} \tag{2-191}$$

Some texts indicate a value of $\frac{3}{2}$ for n; however, experimental measurements on a number of n-channel JFET geometries indicate that the exponent n is close to 2, which is the value derived in an approximate treatment by R. D. Middlebrook [10].

A useful relationship between g_{fso}, I_{DSS}, and $V_{GS(\text{off})}$ is derived from the ratio of Eqs. (2-190) and (2-191):

$$\frac{g_{gs}}{I_D} = n(V_{GS} - V_{GS(\text{off})})^{-1} \tag{2-192}$$

At $V_{GS} = 0$, $I_D = I_{DSS}$ and $g_{fs} = g_{fso}$. Using 2 as the value of the constant n leads to

$$g_{fso} = -2\frac{I_{DSS}}{V_{GS(\text{off})}} \tag{2-193}$$

For n-channel FETs I_{DSS} is positive and $V_{GS(\text{off})}$ is negative; for p-channel FETs I_{DSS} is negative and $V_{GS(\text{off})}$ is positive; thus g_{fs} is a positive quantity for both p- and n-channel FETs.

Equation (2-190) indicates that for $V_{GS} = V_{GS(\text{off})}$, $I_D = 0$. In a real device this does not happen. Starting from zero, as the gate voltage is made more negative, the drain current decreases until it reaches a very low value equal to the drain-leakage current. At this value, the source current will consist of source–gate leakage. Any further increase in the magnitude of the negative gate voltage will result in an increase in I_D leakage. For the small-signal device illustrated, this minimum I_D is on the order of 2×10^{-13} A.

For some types of applications of the FET, it is helpful to understand the characteristics at very low values of V_{DS}, such as those shown in Figure 2-97. A "very low value" is one that is small compared to the magnitude of $V_{GS} - V_{GS(\text{off})}$. In this region, V_{DS} is small enough to have little effect on channel thickness, so that the I_D/V_{DS} slope is nearly linear. Since the slope is a function of V_{GS}, the FET can be utilized as a voltage-controlled resistor. The conductance slope ($\Delta I_D/V_{DS}$) at $V_{DS} = 0$ is approximately a linear function of $V_{GS} - V_{GS(\text{off})}$.

If g_{ds} at $V_{GS} = 0$ is given the term g_{dso}, then

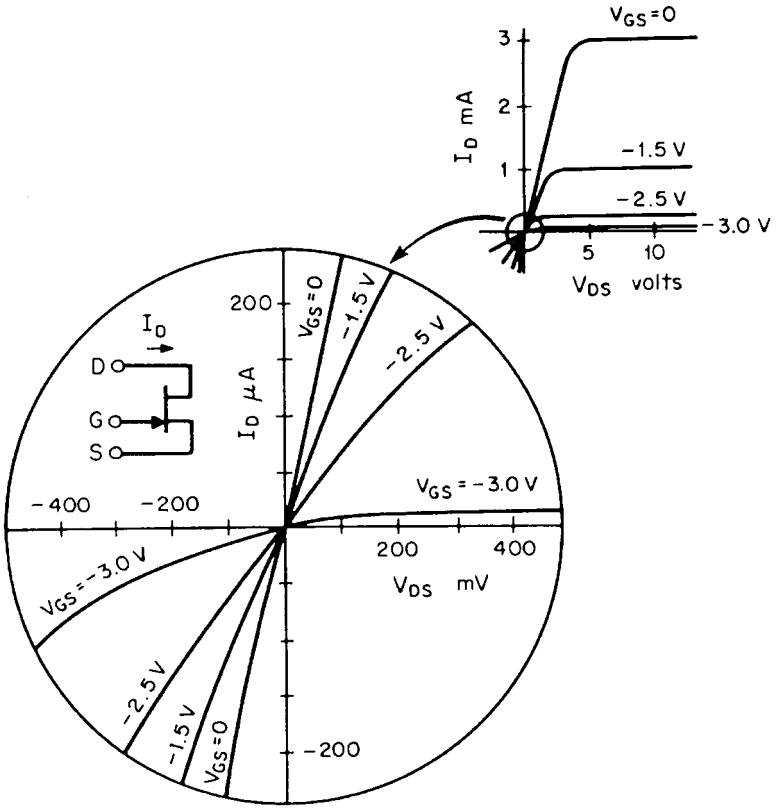


Figure 2-97 Enlargement of n-channel output characteristic around $V_{DS} = 0$.

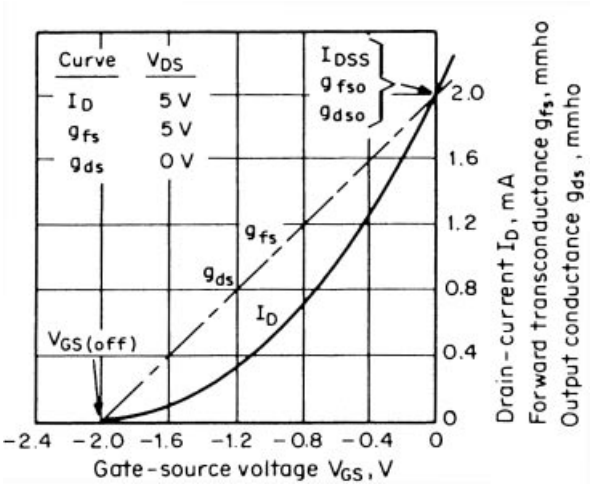


Figure 2-98 JFET I_D, g_{fs}, and g_{ds}.

$$g_{ds} = g_{dso}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right) \tag{2-194}$$

with $V_{DS} = 0$. A plot of this characteristic is shown in Figure 2-98, along with g_{fs} and I_D characteristics.

The relationship between g_{dso}, $V_{GS(\text{off})}$, and I_{DSS} is given by the equation

$$g_{dso} = -2\frac{I_{DSS}}{V_{GS(\text{off})}} \tag{2-195}$$

where I_{DSS} and $V_{GS(\text{off})}$ are as indicated in Figure 2-96. It is important to note that Eqs. (2-194) and (2-195) and the g_{ds} curve of Figure 2-98 are valid only for the case where V_{DS} is very small compared to V_P. Drain–source conductance at higher values of V_{DS} will be discussed later. FETs designed to be used as voltage-controlled resistors typically have a high $V_{GS(\text{off})}$ because the $V_{DS}/(V_{GS} - V_{GS(\text{off})})$ ratio should be low to keep distortion low. The actual large-signal JFET model as used in most simulators is shown in Figure 2-99, including a list of its intrinsic keywords (Table 2-14). Again, we had very little luck using this model as published by several SPICE CAD software manufacturers as offered. While we cannot judge on the quality of the parameters that came with the software, it was not possible to closely match even optimized sets of parameters to this large-signal model and obtain acceptable results. However, since the modified Materka model worked very well for this, we recommend that this popular JFET model not be used at frequencies above 10 MHz.

2-3-2 Small-Signal Behavior of JFETs

Figure 2-100 shows the small-signal equivalent transistor model we would recommend. It consists of an empirical FET model and a package model. Table 2-15 lists its keywords. Its advantage is that its built-in noise model is very accurate, even for JFETs and metal semiconductor FETs.

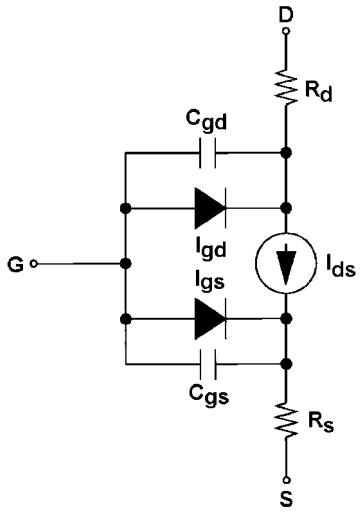


Figure 2-99 The nonlinear JFET model. Table 2-14 lists its intrinsic and extrinsic parameters. In our opinion, this model is limited to frequencies below 10 MHz; above this frequency, the modified Materka model should replace it.

Table 2-14 Large-signal JFET model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
NJF/PJF	Channel-type selection (NJF = *n*-channel, PJF = *p*-channel)		NJF
VT0	Threshold voltage	volt	–2.0
BETA	Transconductance coefficient	ampere/volt^2	1.0E–4
LAMB	Channel-length modulation	/volt	0.0
IS	Gate-junction saturation current	ampere	1.0E–14
PB	Gate-junction potential	volt	1.0
FC	Forward-bias depletion capacitance coefficient		0.5
CGS	Zero-bias gate–source junction capacitance	farad	0.0
CGD	Zero-bias gate–drain junction capacitance	farad	0.0
T	Channel transit-time delay	second	0.0
KF	Flicker noise coefficient		0.0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
SN	Switch to turn device shot noise on (1) or off (0)		1
AREA	Area multiplier		1.0
NOIS	Reference label to a set of noise data		
NAME	Required user-specified name up to 8 characters		
	Extrinsic Model		
RD	Drain ohmic resistance	ohm	0.0
RS	Source ohmic resistance	ohm	0.0

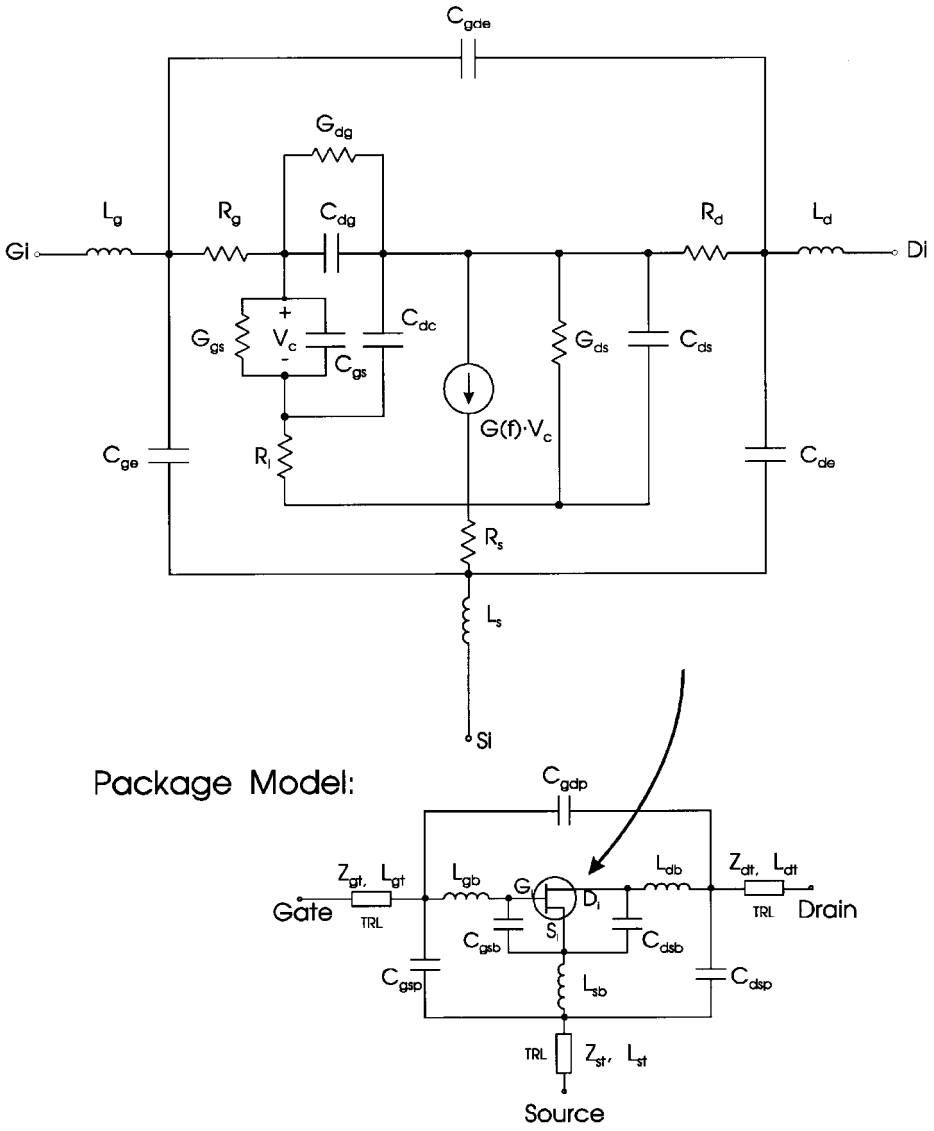


Figure 2-100 Linear FET model.

Electrical noise generated within a JFET is usually represented by equivalent noise sources, $\overline{e}_n$ and $\overline{i}_n$. Both noise voltage $\overline{e}_n$ and noise current $\overline{i}_n$ are frequency dependent and have the characteristics shown in Figure 2-101.

An equivalent noise circuit is shown in Figure 2-102. Above the frequency f_1, $\overline{e}_n$ is approximately given by

$$\overline{e}_n \cong \left(4KTB\,\frac{0.67}{g_{fs}}\right)^{1/2} \tag{2-196}$$

Table 2-15 Small-signal FET model keywords

Keyword	Description	Unit	Default
	Intrinsic model		
G	Transconductance at dc, G_0[a]	/ohm	
CGS	Gate–source capacitance	farad	
F	3-dB roll-off frequency	hertz	∞
T	Time delay	second	0.0
TDS	Drain–source time delay	second	0.0
GGS	Gate–source conductance	/ohm	0.0
CDG	Drain–gate capacitance	farad	0.0
CDC	Dipole layer capacitance	farad	0.0
CDS	Drain–source capacitance	farad	0.0
GDS	Drain–source conductance	/ohm	0.0
RI	Channel resistance	ohm	0.0
RG	Gate resistance	ohm	0.0
RD	Drain resistance	ohm	0.0
RS	Source resistance	ohm	0.0
CGE	External gate capacitance	farad	0.0
CDE	External drain capacitance	farad	0.0
LG	Gate-lead inductance	henry	0.0
LD	Drain-lead inductance	henry	0.0
LS	Source-lead inductance	henry	0.0
CGDE	External gate–drain capacitance	farad	0.0
GDG	Gate–drain conductance	/ohm	0.0
TJ	Chip temperature	kelvin	298
	Package Parasitics		
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate to source package capacitance	farad	0.0
CDSP	Drain to source package capacitance	farad	0.0
CGDP	Gate to drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for at $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for $\varepsilon_r = 1$	meter	0.0
FC	Corner frequency of flicker (1/*f*) noise[b]	hertz	10 MHz
FCP	Shape factor of the 1/*f* noise response		1.0
label	User-defined term that refers to temperature coefficient		

[a]The transconductance of this model may be approximately described by $g_m = G\,\frac{e^{-j\omega T}}{1 + j(f/F)}$ where $\omega = 2\pi f$ and f is frequency.
[b]The flicker noise frequency dependence is given by $1/(f/F_c)^{\mathrm{FCP}}$.

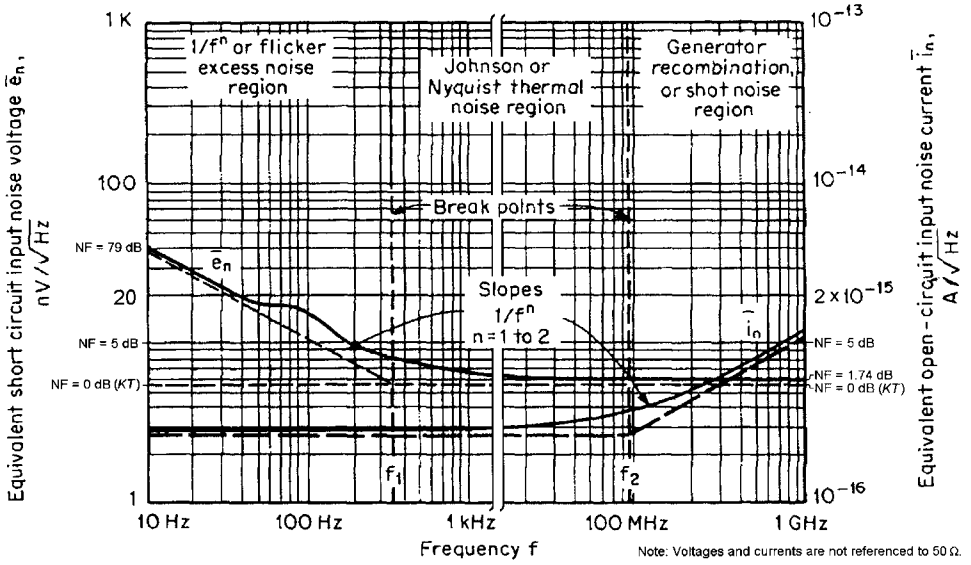


Figure 2-101 JFET noise characteristics.

where $K = 1.374 \times 10^{-23}$ J/K

T = absolute temperature in kelvins (273 K = 0 °C)

B = frequency range in hertz

g_{fs} = transconductance of FET

With the input short-circuited, the noise voltage across the load R_L resulting from the FET is

$$\text{Output noise voltage} = \bar{e}_n A_V \tag{2-197}$$

Below the frequency f_1, $\bar{e}_n$ increases proportional to $1/f^n$ and is expressed as

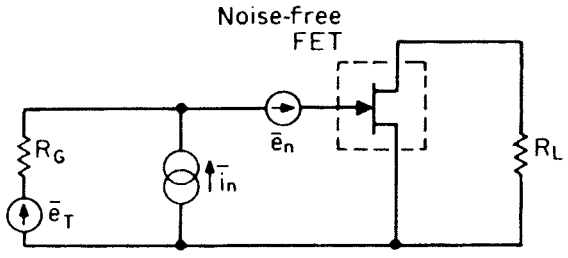


Figure 2-102 Equivalent noise circuit of the FET.

$$\bar{e}_n = \left[4KTB \left(\frac{0.67}{g_{fs}} \right) \left(1 + \frac{f_1}{f^n} \right)^{1/2} \right] \tag{2-198}$$

The low-frequency corner frequency f_1 for JFETs is typically in the 100-Hz to 1-kHz range, and the exponent n is usually between 1 and 2. As indicated by the equations, $\bar{e}_n$ is inversely proportional to the square root of g_{fs}.

The equivalent input noise current $\bar{i}_n$ is caused by the current in the gate-to-channel junction. Its approximate value below f_2 is

$$\bar{i}_n = (2qI_G B)^{1/2} \tag{2-199}$$

where $q = 1.602 \times 10^{-19}$
B = frequency range in hertz
I_G = dc gate current

This expression is fairly accurate when I_G is the result of the active device conductance. Typically, $\bar{i}_n$ will be lower than the calculated value because part of I_G is due to conductance across the device package.

At higher frequencies (above f_2),

$$\bar{i}_n = \left(\frac{4KTB}{R_P} \right)^{1/2} \tag{2-200}$$

where R_P is the real part of gate input impedance. R_P, in terms of Y_{11}, can range from several tens of megohms (at audiofrequencies) to 1 kΩ or less (at VHF/UHF frequencies). The high-frequency corner f_2 is typically in the range of 5–50 kHz and more than 100 MHz for high-frequency devices.

Another form of noise is known as "popcorn" or burst noise, the causes of which have not been completely identified. It shows up as a random short-duration step-function change in drain current, equivalent to an input gate–source voltage change of a few tenths of a microvolt. It is not unlike big bubbles on the surface of boiling water.

As far as the noise is concerned, we can also state that the Materka built-in proprietary noise model is significantly more accurate than the one mentioned above [11–14]. This model is implemented, with minimal documentation of its inner workings because of its proprietary nature, in the linear portion of the circuit simulator in Ansoft's Serenade Design Environment. In many cases, the manufacturer supplies only limited data, so it is most convenient that after SPICE-type parameter extraction is done by using the Scout program, one even gets good noise prediction for JFETs using the modified Materka model.

2-3-3 Large-Signal Behavior of MOSFETs

Metal-oxide-semiconductor field-effect transistors (MOSFETs) are important components in contemporary analog integrated circuits. Whereas initial applications were centered on all-MOS processes, combined bipolar and MOS processes give the designer the best of both worlds. A major advantage of MOS processes for realizing analog functions is that complex, dense digital functions can be realized on the same chip.

Transfer Characteristics of MOS Devices. A cross section of a typical enhancement-mode *n*-channel MOS transistor (NMOS) is shown in Figure 2-103a. Heavily doped *n*-type source and drain regions are fabricated in a *p*-type substrate (often called the body). A thick layer of silicon dioxide is grown over the substrate material and a conductive gate material (metal or polycrystalline silicon) covers this oxide between source and drain. The operation of the device is very similar to that of a JFET in that the gate–source voltage is used to modify the conductance of the region under the gate. This allows the gate voltage to control the current flowing between source and drain, giving gain in analog circuits and switching characteristics in digital circuits.

The enhancement-mode NMOS device of Figure 2-103a shows significant conduction between source and drain only when an *n*-type channel exists under the gate, and this is the origin of the "*n*-channel" designation. The term "enhancement mode" refers to the fact that no conduction occurs for $V_{GS} = 0$, and thus the channel must be "enhanced" to cause conduction. MOS devices can be made equally well by using an *n*-type substrate with a *p*-type conducting channel. Such devices are called enhancement-mode *p*-channel MOS transistors (PMOS). In technologies employing one or the other of these device types, the circuit symbol of Figure 2-103b is commonly used for either one. In complementary MOS technology (CMOS) both device types are present and the circuit symbols of Figure 2-103c will be used to distinguish them. In PMOS and NMOS technologies, the substrate is common to all devices, invariably connected to a dc power supply voltage, and usually not shown on the circuit diagram. In CMOS technology, however, devices of one type or another are fabricated in individual, separate isolation regions, which may or may not be connected to a power supply voltage. If these isolation regions *are* connected to the appropriate power supply, the symbols of Figure 2-103c will be used and the substrate connection will not be shown. If the individual isolation regions are connected elsewhere, however, the devices will be represented by the symbols of Figure 2-103d, where the substrate is labeled *B*. Finally, in NMOS technology an additional device type called a "depletion-mode" device is usually available. This is a conducting channel implanted between the source and drain so that conduction occurs for $V_{GS} = 0$. This device has characteristics that are almost identical to that of a JFET and will be represented by the symbol of Figure 2-103e.

The derivation of the transfer characteristics of the enhancement-mode NMOS device of Figure 2-103a begins by noting that with $V_{GS} = 0$, the source and drain regions are separated by back-to-back *pn* junctions. These junctions are formed between the *n*-type source and drain regions and the *p*-type substrate, resulting in an extremely high resistance (about 10^{12} Ω) between drain and source when the device is off.

Now consider substrate, source, and drain grounded and a positive voltage V_{GS} applied to the gate as shown in Figure 2-104. The gate and substrate then form the plates of a capacitor with the SiO_2 as a dielectric. Positive charge accumulates on the gate and the negative charge in the substrate. Initially, the negative charge in the *p*-type substrate is manifested by creation of a *depletion region* and resulting exclusion of holes under the gate. This is shown in Figure 2-104. The depletion-layer width *X* under the oxide is

$$X = \left(\frac{2\varepsilon\phi}{qN_A} \right)^{1/2} \tag{2-201}$$

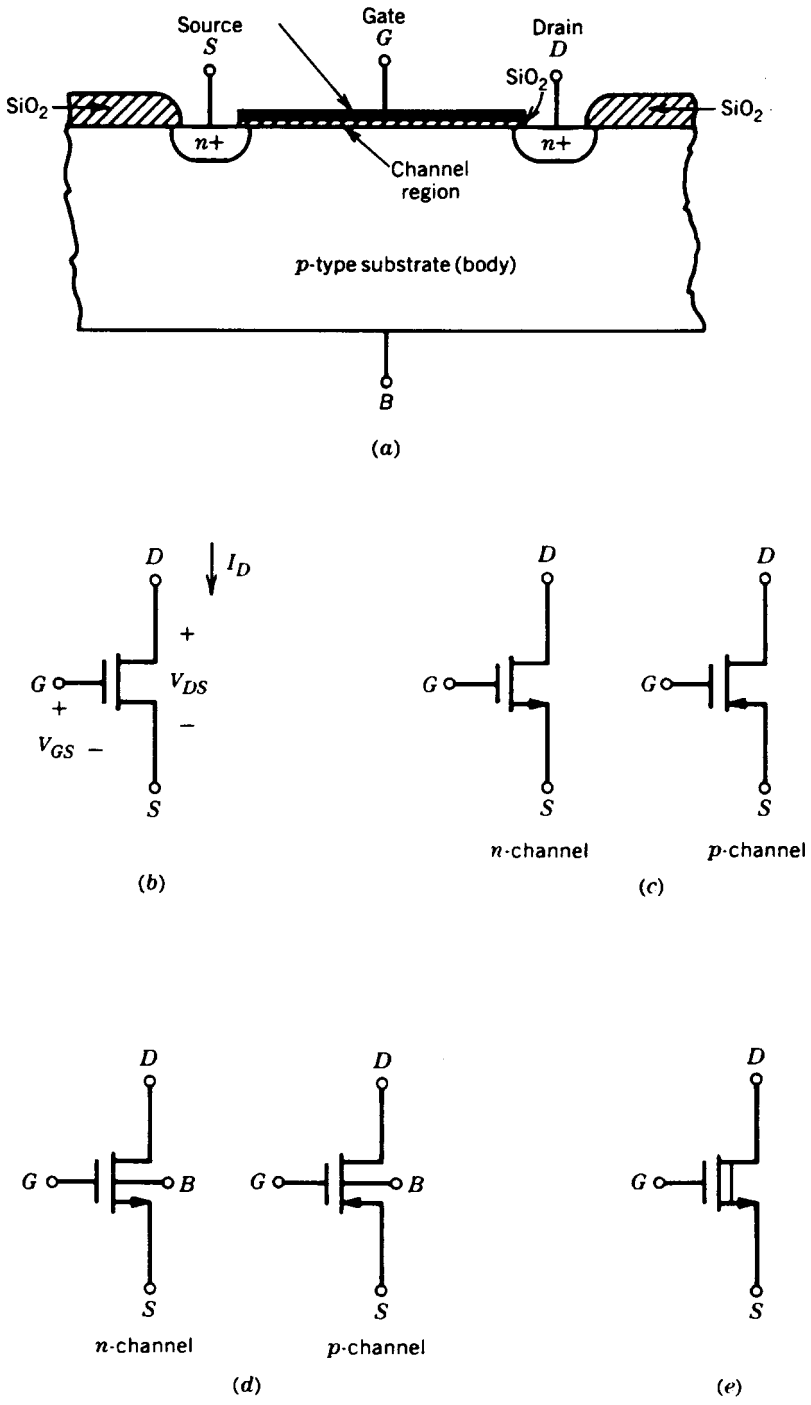


Figure 2-103 (a) Typical enhancement-mode NMOS structure. (b) Enhancement-mode NMOS or PMOS circuit symbol when one device type only is present. (c) NMOS and PMOS symbols used in CMOS circuits. (d) NMOS and PMOS symbols used when the substrate connection is nonstandard. (e) Depletion MOS device symbol.

where ϕ is the potential in the depletion layer at the oxide–silicon interface, N_A atoms/cm^3 is the doping density (assumed constant) of the p-type substrate, and ε is the permittivity of the silicon. The charge per unit area in this depletion region is

$$Q = qN_A X = \sqrt{2qN_A \varepsilon \phi} \tag{2-202}$$

When the potential in the silicon reaches a critical value equal to twice the Fermi level, $\phi_f \simeq 0.3$ V, a phenomenon known as "inversion" occurs [15]. Further increases in gate voltage produce no further changes in the depletion-layer width but instead a thin layer of electronics is induced in the depletion layer directly under the oxide. This produces a continuous n-type region with the source and drain regions and is the conducting channel between source and drain. This channel can then be modulated by increases or decreases in the gate voltage. In the presence of an inversion layer, and with no substrate bias, the depletion region contains a fixed charge

$$Q_{b0} = \sqrt{2qN_A \varepsilon 2\phi_f} \tag{2-203}$$

If a substrate bias voltage V_{SB} (source positive for n-channel devices) is applied between source and substrate, the potential required to produce inversion becomes $(2\phi_f + V_{SB})$ and the charge stored in the depletion region in general is

$$Q_b = \sqrt{2qN_A \varepsilon (2\phi_f + V_{SB})} \tag{2-204}$$

The gate voltage V_{GS}, required to produce an inversion layer, is called the threshold voltage V_t and can now be calculated. This voltage consists of several components. First, a voltage $[2\phi_f + (Q_b/C_{ox})]$ is required to sustain the depletion-layer charge Q_b, where C_{ox} is the gate oxide capacitance per unit area. Second, a work-function difference ϕ_{ms} exists between the

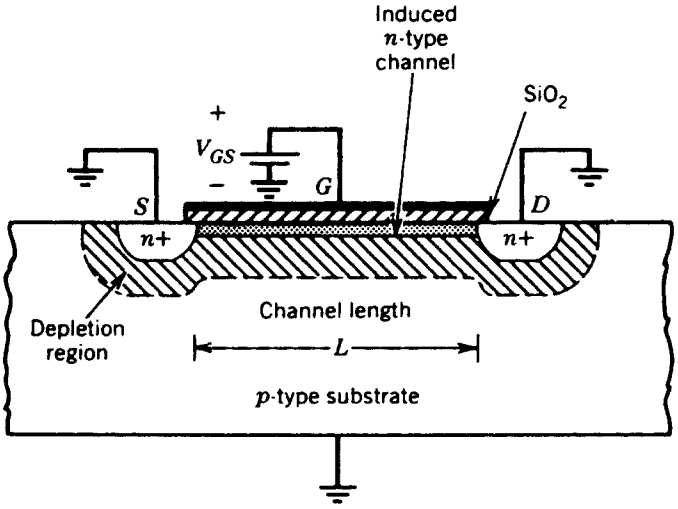


Figure 2-104 Idealized NMOS device cross section with positive V_{GS} applied, showing depletion regions and the induced channel.

gate metal and the silicon. Third, there is always charge density Q_{ss} (positive) in the oxide at the silicon interface. This is caused by crystal discontinuities at the Si–SiO_2 interface and must be compensated by a gate voltage contribution of $-Q_{ss}/C_{ox}$. Thus we have a threshold voltage

$$V_t = \phi_{ms} + 2\phi_f + \frac{Q_b}{C_{ox}} - \frac{Q_{ss}}{C_{ox}}$$

$$= \phi_{ms} + 2\phi_f + \frac{Q_{b0}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} + \frac{Q_b - Q_{b0}}{C_{ox}} \tag{2-205}$$

$$= V_{t0} + \gamma\left(\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f}\right) \tag{2-206}$$

where Eqs. (2-203) and (2-204) have been used and V_{t0} is the threshold voltage with $V_{SB} = 0$. The parameter γ is defined as

$$\gamma = \frac{1}{C_{ox}}\sqrt{2q\varepsilon N_A} \tag{2-207}$$

and

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \tag{2-208}$$

where ε_{ox} and t_{ox} are the permittivity and thickness of the oxide, respectively. A typical value of γ is 0.5 $V^{1/2}$ and $C_{ox} = 3.5 \times 10^{-4}$ pF/μ^2 for $t_{ox} = 0.1$ μm.

In practice, the value of V_{t0} is usually adjusted in processing by implanting additional impurities into the channel region. Extra *p*-type impurities are implanted in the channel to make $V_{t0} \simeq 0.5$–1.5 V for *n*-channel enhancement devices. By implanting *n*-type impurities in the channel region, a conducting channel can be formed, even for $V_{GS} = 0$. This MOS transistor is called a depletion device, with typical values of V_{t0} in the range –1 to –4 V. If Q_i is the charge density per unit area due to the implant, then the threshold voltage given by Eq. (2-205) is shifted by approximately Q_i/C_{ox}.

The preceding equations can now be used to calculate the large-signal characteristics of an NMOS transistor. For purposes of analysis the source is assumed grounded and bias voltages V_{GS}, V_{DS}, and V_{SB} are applied as shown in Figure 2-105. If V_{GS} is greater than V_t, a conducting channel exists and V_{DS} causes a larger reverse bias from drain to substrate than exists from source to substrate, and thus a wider depletion region exists at the drain. However, for simplicity, we assume that the voltage drop along the channel itself is small so that the depletion layer is constant along the channel.

At a distance *y* along the channel the voltage with respect to the source is $V(y)$ and the gate-to-channel voltage at that point is $V_{GS} - V(y)$. We assume this voltage exceeds the threshold voltage V_t, and thus the induced electron charge per unit area in the channel is

$$Q_I(y) = C_{ox}[V_{GS} - V(y) - V_t] \tag{2-209}$$

The resistance *dR* of a length *dy* of the channel is

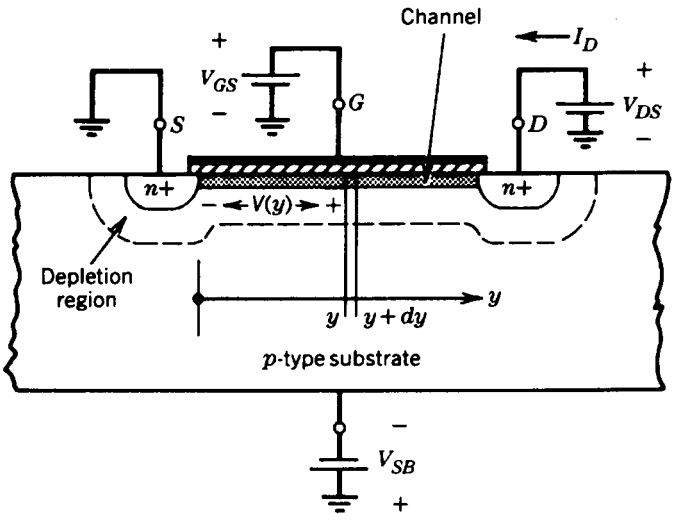


Figure 2-105 NMOS device with bias voltages applied.

$$dR = \frac{dy}{W\mu_n Q_I(y)} \tag{2-210}$$

where W is the width of the device, perpendicular to the plane of Figure 2-105, and μ_n is the average electron mobility of the channel.

The voltage drop dV along the length of the channel dy is

$$dV = I_D\, dR = \frac{I_D}{W\mu_n Q_I(y)}\, dy \tag{2-211}$$

If L is the total channel length, then substitution of Eq. (2-209) into Eq. (2-211) and integration gives

$$\int_0^L I_D\, dy = \int_0^{V_{DS}} W\mu_n C_{\text{ox}}(V_{GS} - V - V_t)\, dV \tag{2-212}$$

This results in

$$I_D = \frac{k'}{2}\frac{W}{L}[2(V_{GS} - V_t)V_{DS} - V_{DS}^2] \tag{2-213}$$

where

$$k' = \mu_n C_{\text{ox}} = \frac{\mu_n \varepsilon_{\text{ox}}}{t_{\text{ox}}} \tag{2-214}$$

Equation (2-213) is important and describes the I–V characteristics of an MOS transistor, assuming a continuous induced channel. A typical value of k' for $t_{\text{ox}} = 0.1$ μm is about 20 μA/V^2 for an n-channel device.

As the value of V_{DS} is increased, the induced conducting channel narrows at the drain end and Eq. (2-209) indicates that Q_I at the drain end approaches zero as V_{DS} approaches $(V_{GS} - V_t)$. This results in the same pinchoff phenomenon as occurs in a JFET, and further increases in V_{DS} produce little change in I_D. Equation (2-213) is thus no longer valid if V_{DS} is greater than $(V_{GS} - V_t)$. The value of I_D in this region is obtained by substituting $V_{DS} = (V_{GS} - V_t)$ in Eq. (2-213), giving

$$I_D = \frac{k'}{2}\frac{W}{L}(V_{GS} - V_t)^2 \tag{2-215}$$

for an MOS transistor in the pinchoff region. As in the case of a JFET, the drain current in the pinchoff region varies slightly as the drain voltage is varied. This is due to the presence of a depletion region between the physical pinchoff point in the channel at the drain end and the drain region itself. If this depletion-layer width is X_d, then the *effective* channel length is given by

$$L_{\text{eff}} = L - X_d \tag{2-216}$$

If L_{eff} is used in place of L in Eq. (2-215), we obtain a more accurate formula for the pinchoff region:

$$I_D = \frac{k'}{2}\frac{W}{L_{\text{eff}}}(V_{GS} - V_t)^2 \tag{2-217}$$

The fact that X_d (and thus L_{eff}) is a function of the drain–source voltage results in a variation of I_D and V_{DS} in the pinchoff region. Using Eqs. (2-216) and (2-217), we obtain

$$\frac{\partial I_D}{\partial V_{DS}} = -\frac{k'}{2}\frac{W}{L_{\text{eff}}^2}(V_{GS} - V_t)^2\frac{dL_{\text{eff}}}{dV_{DS}} \tag{2-218}$$

and thus

$$\frac{\partial I_D}{\partial V_{DS}} = \frac{I_D}{L_{\text{eff}}}\frac{dX_d}{dV_{DS}} \tag{2-219}$$

This equation is analogous to Eq. (2-168) for bipolar transistors. Following a similar procedure, a voltage analogous to the Early voltage can be defined as

$$V_A = \frac{I_D}{\partial I_D/\partial V_{DS}} \tag{2-220}$$

and thus

$$V_A = L_{\text{eff}}\left(\frac{dX_d}{dV_{DS}}\right)^{-1} \tag{2-221}$$

For MOS transistors, the most widely used parameter for the characterization of output resistance is

$$\lambda = \frac{1}{V_A} \tag{2-222}$$

As in the bipolar case, the large-signal properties of the transistor can be approximated by assuming that λ and V_A are constants, independent of bias conditions. Thus we can formulate a better approximation to the I–V characteristics as

$$I_D = \frac{k'}{2}\frac{W}{L}(V_{GS} - V_t)^2 (1 + \lambda V_{DS}) \tag{2-223}$$

In practical MOS transistors variation of X_d with voltage is complicated by the fact that the field distribution in the drain depletion region is not one dimensional. A vertical component in the field distribution is introduced by the potential difference between the gate and channel and the gate and drain. As a result, the calculation of λ from the device structure is quite different [16], and it is usually necessary to develop effective values of λ from experimental data. The parameter λ is a linear function of effective channel length and is an increasing function of the doping level in the channel. Typical values of λ are in the range 0.05–0.005 V^{-1}. Note that, in the case of a JFET, the pinchoff region for MOS devices is often called the *saturation* region.

A plot of I_D versus V_{DS} for an NMOS transistor is shown in Figure 2-106. Below pinchoff, the device behaves as a nonlinear voltage-controlled resistor, which is often called the *ohmic* or *triode* region. Above pinchoff, the device approximates a voltage-controlled current source. Note that for depletion MOS devices, V_t is negative and I_D is finite, even for $V_{GS} = 0$. For PMOS devices, all polarities of voltage and current are reversed.

The results as previously derived can be used to form a large-signal model of an MOS transistor. The model topology in the pinchoff region is the same as that for the JFET, but using Eq. (2-215) for the controlled-current generator.

MOS Device Voltage Limitations. The voltage limitations of MOS transistors depend on the gate length L. For small values of L (less than about 10 μm), the drain-depletion region

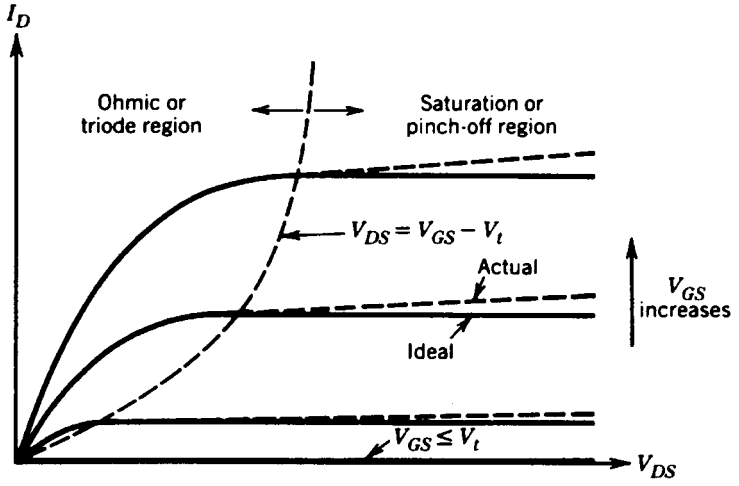


Figure 2-106 NMOS device characteristics.

exerts an appreciable influence on the channel. This causes I_D to rise with increasing V_{DS} in a similar fashion to the bipolar transistor curves of Figure 2-89.

For long channel lengths, the drain-depletion region has little effect on the channel and the I_D-versus-V_{DS} curves follow closely the ideal curves of Figure 2-106. Eventually the drain–substrate *pn*-junction breakdown voltage is exceeded and a sharp breakdown characteristic is obtained, which is similar to the JFET characteristic.

In addition to V_{DS} limitations, MOS devices must also be protected against excessive gate voltages. Typical gate oxides break down with about 25–50 V applied from gate to channel, and this process is destructive to the transistor.

2-3-4 Small-Signal Model of the MOS Transistor in Saturation

The preceding large-signal equations can now be used to derive the small-signal model of the MOS transistor in the saturation or pinchoff region. The important equations are (2-220) and (2-206). Note from Eq. (2-206) that the source–substrate voltage V_{BS} affects V_t, and thus I_D. This is due to the influence of the substrate acting as a second gate and is called *body effect*. As a consequence, I_D is a function of both V_{GS} and V_{BS}, and we require *two* transconductance generators in the small-signal model as shown in Figure 2-107. Variations in the voltage v_{bs} from source to body cause current $g_{mb}v_{bs}$ to flow from drain to source. Note that the body (or substrate) of an NMOS integrated circuit is usually connected to the most negative supply voltage and is thus an ac ground. However, the source connection can have a significant ac voltage impressed on it. Parasitic resistances due to the channel contact regions should be included in series with the source and drain of the model but are usually neglected in hand calculations. These resistances have an inverse dependence on channel width W and have typical values of 50–100 Ω for devices with W of about 1 μm.

The parameters of Figure 2-107 can be determined from Eq. (2-223) by differentiating.

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = k' \frac{W}{L} (V_{GS} - V_t)(1 + \lambda V_{DS}) \tag{2-224}$$

If $\lambda V_{DS} << 1$, this is often approximated as

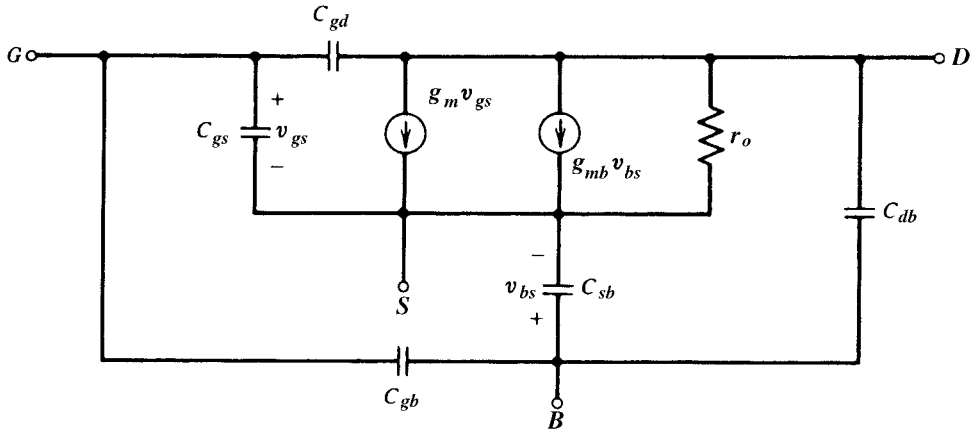


Figure 2-107 Small-signal MOS transistor equivalent circuit.

$$g_m = k' \frac{W}{L} (V_{GS} - V_t) = \sqrt{2k' \frac{W}{L} I_D} \tag{2-225}$$

Like the JFET and unlike the bipolar transistor, the transconductance of the MOSFET depends on both bias current and the *W*/*L* ratio (also on the oxide thickness via *k*′). Similarly,

$$g_{mb} = \frac{\partial I_D}{\partial V_{BS}} = -k' \frac{W}{L} (V_{GS} - V_t) (1 + \lambda V_{DS}) \frac{\partial V_t}{\partial V_{BS}} \tag{2-226}$$

From Eq. (2-206)

$$\frac{\partial V_t}{\partial V_{BS}} = -\frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} = -\chi \tag{2-227}$$

where this equation defines a factor χ. This parameter is equal to the rate of change of threshold voltage with body bias voltage. One obtains

$$\chi = \frac{C_{js}}{C_{\text{ox}}} \tag{2-228}$$

where C_{js} is the capacitance per unit area of the depletion region under the channel, assuming a one-sided step junction with a built-in potential $\psi_0 = 2\phi_f$.

Substitution of Eq. (2-227) into Eq. (2-226) gives

$$g_{mb} = \frac{\gamma k' (W/L) (V_{GS} - V_t) (1 + \lambda V_{DS})}{2\sqrt{2\phi_f + V_{SB}}} \tag{2-229}$$

Again, if $\lambda V_{DS} << 1$, we have

$$g_{mb} = \frac{\gamma \sqrt{k' (W/L) I_D}}{\sqrt{2(2\phi_f + V_{SB})}} \tag{2-230}$$

An important quantity is the ratio g_{mb}/g_m, and from Eqs. (2-224) and (2-229) we find

$$\frac{g_{mb}}{g_m} = \frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} = \chi \tag{2-231}$$

The factor χ is typically in the range 0.1–0.3.

Finally, the small-signal output resistance can be obtained directly from Eq. (2-218):

$$r_o = \left(\frac{\partial I_D}{\partial V_{DS}} \right)^{-1} = \frac{L_{\text{eff}}}{I_D} \left(\frac{dX_d}{dV_{DS}} \right)^{-1} \tag{2-232}$$

and using Eqs. (2-221) and (2-222), we find

$$r_o = \frac{1}{\lambda I_D} = \frac{V_A}{I_D} \tag{2-233}$$

Small-signal model capacitances are also shown in Figure 2-107. Of these, only the gate–source capacitance C_{gs} is intrinsic to the device operation in the saturation region. Capacitances C_{sb} and C_{db} are parasitic depletion-region capacitances between the substrate and the source and drain regions, respectively. These can be expressed as follows:

$$C_{sb} = \frac{C_{sb0}}{(1 + V_{SB}/\psi_0)^{1/2}} \tag{2-234}$$

$$C_{db} = \frac{C_{db0}}{(1 + V_{DB}/\psi_0)^{1/2}} \tag{2-235}$$

These capacitances are proportional to the gate and source region areas (including sidewalls), and C_{sb} also includes depletion-region capacitance from the induced channel in the body.

Capacitance C_{gb} between gate and substrate models parasitic oxide capacitance between the gate contact material and the substrate outside the active-device area. This is a constant capacitance and models coupling between polysilicon and metal interconnects and the underlying substrate. In fact, parasitic capacitance of this type underlies all polysilicon and metal traces on the chip and should be taken into account when simulating and calculating high-frequency circuit and device performance. Typical values depend on oxide thicknesses and range from about 0.04 to 0.15 fF per square micron of interconnect, with fringing effects becoming important for narrow lines (several microns or less in width).

Capacitances C_{gs} and C_{gd} exist from gate to source and drain, respectively. If C_{ox} is the oxide capacitance per unit area from gate to channel then the total capacitance under the gate is $C_{\text{ox}}WL$. This capacitance is intrinsic to the device operation and models the gate control of the channel conductance. In the ohmic region of device operation, this capacitance is split equally between source and drain so that $C_{gs} = C_{gd} = 1/2C_{\text{ox}}WL$. However, in the saturation region, the channel is very narrow at the drain end and the drain voltage exerts little influence on either the channel or the gate charge. As a consequence, the intrinsic portion of C_{gd} is essentially zero in the saturation region and C_{gd} then consists of a constant parasitic oxide-capacitance contribution due to gate overlap of the drain region. This is on the order of 1–10 fF for small devices.

In order to calculate the corresponding value of C_{gs} in the saturation region, we must calculate the total charge Q_T stored in the channel. This can be obtained by integrating Eq. (2-209) to obtain

$$Q_T = WC_{\text{ox}} \int_0^L [V_{GS} - V(y) - V_t]dy \tag{2-236}$$

Substituting for dy/dV from Eq. (2-211) into Eq. (2-236), we find

$$Q_T = \frac{W^2 C_{\text{ox}}^2 \mu_n}{I_D} \int_0^{V_{GS}-V_t} (V_{GS} - V - V_t)^2 dV \tag{2-237}$$

where the limit $y = L$ corresponds to $V = (V_{GS} - V_t)$ in saturation. Solution of Eq. (2-237) and use of Eqs. (2-214) and (2-215) gives

$$Q_T = \frac{2}{3} WLC_{\text{ox}}(V_{GS} - V_t) \tag{2-238}$$

and thus

$$C_{gs} = \frac{\partial Q_T}{\partial V_{GS}} = \frac{2}{3} WLC_{\text{ox}} \tag{2-239}$$

In addition, there is a contribution to C_{gs} from the constant parasitic oxide capacitance due to gate overlap of the source region.

The f_T of the MOSFET is given by

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{gs} + C_{gd} + C_{gb}} \tag{2-240}$$

The dependence of MOSFET f_T on device and process parameters can be seen by assuming that the intrinsic device capacitance G_{gs} dominates. Thus from Eq. (2-240) we have

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{gs}} \tag{2-241}$$

Substituting in Eq. (2-241) for g_m from Eq. (2-225) and C_{gs} from Eq. (2-239) we find for a MOSFET

$$f_T = 1.5 \frac{\mu_n}{2\pi L^2} (V_{GS} - V_t) \tag{2-242}$$

It is interesting to compare this with the intrinsic f_T of a bipolar transistor when parasitic depletion-layer capacitance is neglected. From

$$f_T = \frac{1}{2\pi\tau_F} \tag{2-243}$$

and substituting for τ_f and using the Einstein relationship $D_n/\mu_n = kT/q = V_T$, we find for a bipolar transistor

$$f_T = 2 \frac{\mu_n}{2\pi W_B^2} V_T \tag{2-244}$$

The similarity in form between Eqs. (2-242) and (2-244) is striking. In both cases the intrinsic device f_T increases as the inverse square of the critical device dimension across which carriers are in transit. The voltage V_t = 26 mV is fixed for a bipolar transistor, but MOSFET f_T can be increased by operating at high values of $(V_{GS} - V_t)$. Note that the base width W_B in a bipolar transistor is a vertical dimension determined by diffusions or implants and can typically be made much smaller than the channel length L of a MOSFET, which depends on surface geometry and photolithographic processes. Thus bipolar transistors generally have higher f_T than MOSFETs made with comparable processing. Finally, Eq. (2-242) was derived assuming that the MOSFET square law is valid. However, as discussed in Section 2-3-5, submicron MOSFETs depart significantly from square-law characteristics, and we find that for such devices f_T increases as L^{-1} rather than L^{-2}.

2-3-5 Short-Channel Effects in FETs

The evolution of integrated-circuit processing techniques has led to continuing reductions in both the horizontal and vertical dimensions of the active devices (the minimum allowed dimension of passive devices has also increased). This trend is driven primarily by economics in that more devices and circuits can be processed at one time on a given wafer. A second benefit has been that the frequency capability of the active devices continues to increase, as intrinsic f_T values increase with smaller dimensions while parasitic capacitances decrease.

Vertical dimensions such as the base width of a bipolar transistor in production processes may now be on the order of 0.05 μm or less, whereas horizontal dimensions such as bipolar emitter width or FET gate length may be significantly less than 1 μm. Even at these very small dimensions, the large-signal and small-signal models of bipolar transistors given in previous sections remain valid. However, significant short-channel effects become important in FETs of all types at channel lengths on the order of 1 μm or less and require modifications to the FET models given previously. The primary effect is to modify the classical FET square-law transfer characteristic in the saturation region to make the device more closely approach an ideal linear transfer function. Note, however, that even in processes with submicron capability, many of the FETs in a given analog circuit may be deliberately chosen to be larger than the minimum size and may be well approximated by the square-law model.

The most important short-channel effect in FETs is due to velocity saturation of carriers in the channel [17]. At low electric field values, the linear relation between carrier velocity and field implied by Eqs. (2-210) and (2-211) is valid. At high fields, however, the carrier velocities approach the thermal velocities and subsequently the carrier velocities increase more slowly with increasing field. This is illustrated in Figure 2-108, which shows typical measured electron drift velocity v_d versus tangential electric field strength magnitude $\mathscr{E}$ in an NMOS surface channel. Note that at low field values the velocity is proportional to the field, while at high fields the velocity approaches a constant value called the scattering-limited velocity v_{scl}. A first-order analytical approximation to this curve is

$$v_d = \frac{\mu_n \mathscr{E}}{1 + \mathscr{E}/\mathscr{E}_c} \tag{2-245}$$

where $\mathscr{E} \simeq 1.5 \times 10^6$ V/m and $\mu_n \simeq 0.07$ m^2/V · s is the low-field mobility. Equation (2-245) is also plotted in Figure 2-108. From Eq. (2-245), as $\mathscr{E} \to \infty$ we have $v_{\text{scl}} = \mu_n \mathscr{E}_c$. At the critical field value $\mathscr{E}_c$, the carrier velocity is a factor of 2 less than the low-field formula would predict. In a device with a channel length $L = 1$ μm, we need a voltage drop of only 1.5 V along the channel to have an average field equal to $\mathscr{E}_c$, and this condition is readily achieved in small MOSFETs. Similar results are found for PMOS devices.

As an example of the effects of velocity saturation on FET characteristics, we consider the example of the MOSFET. In the analysis of "Transfer Characteristics of MOS Devices" in Section 2-3-3, we now use the more general expression

$$I_D = WQ_I(y)v_d(y) \tag{2-246}$$

Substituting Eq. (2-245) into Eq. (2-246) and using

$$\mathscr{E} = \frac{dV}{dy} \tag{2-247}$$

for the magnitude of the field, we find that

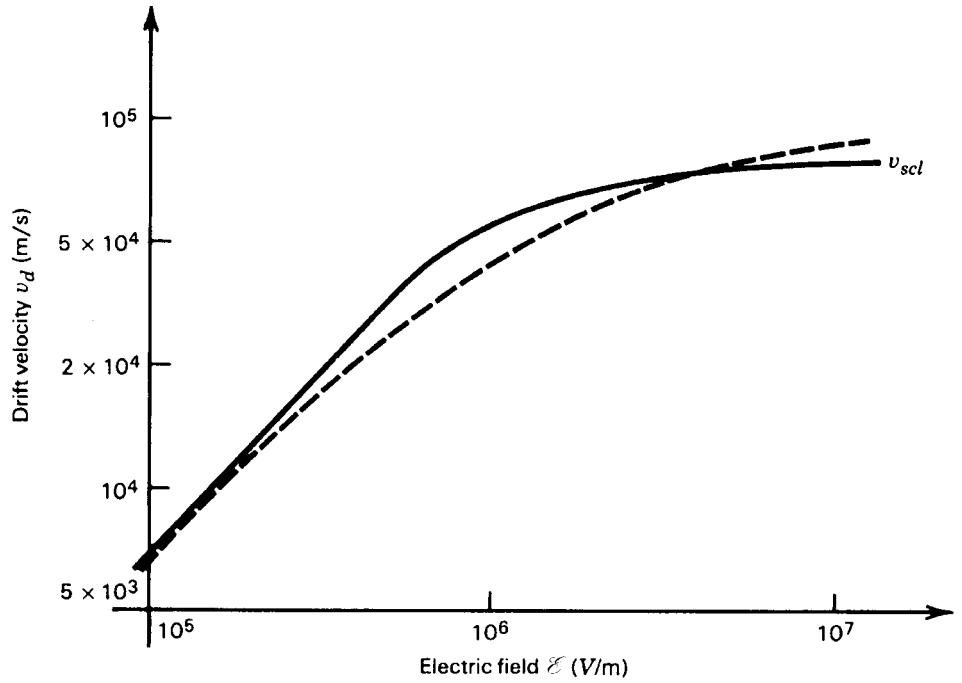


Figure 2-108 Typical measured electron drift velocity v_d versus tangent electric field $\mathscr{E}$ in an MOS surface channel (solid line). Also shown (dotted line) is the analytical approximation of Eq. (2.245) with $\mathscr{E}_c = 1.5 \times 10^6$ V/m and $\mu_n = 0.07$ m^2/V · s.

$$I_D\left(1 + \frac{1}{\mathscr{E}_c}\frac{dV}{dy}\right) = WQ_I(y)\mu_n\frac{dV}{dy} \tag{2-248}$$

Note that as $\mathscr{E}_c \to \infty$ and velocity saturation becomes negligible, Eq. (2-248) approaches the original equation (2-211).

Integrating Eq. (2-248) along the channel we obtain

$$\int_0^L I_D\left(1 + \frac{1}{\mathscr{E}_c}\frac{dV}{dy}\right)dy = \int_0^{V_{DS}} WQ_I(y)\mu_n dV \tag{2-249}$$

and thus

$$I_D = \frac{\mu_n C_{ox}}{2\left(1 + \dfrac{1}{\mathscr{E}_c}\dfrac{V_{DS}}{L}\right)}\frac{W}{L}[2(V_{GS} - V_t)V_{DS} - V_{DS}^2] \tag{2-250}$$

The quantity V_{DS}/L in Eq. (2-250) can be interpreted as the average field in the channel. If this is comparable to $\mathscr{E}_c$, the drain current for a given V_{DS} is less than the simple expression (2-213) would predict.

Equation (2-250) is valid in the triode region. The MOSFET transfer function in saturation can be obtained by using $V_{DS} = (V_{GS} - V_t)$ in Eq. (2-250) to obtain

$$I_D = \frac{k'}{2[1 + \theta(V_{GS} - V_t)]} \frac{W}{L} (V_{GS} - V_t)^2 \tag{2-251}$$

where $\theta = 1/L\mathscr{E}_c$ and has the dimension volts^{-1}. For $L = 1$ μm, a typical value is $\theta \approx 0.7\ \text{V}^{-1}$. Note that in the presence of velocity saturation effects, the device enters the saturation region for $V_{DS} < (V_{GS} - V_t)$. However, Eq. (2-251) still gives a good estimate of the saturation current.

Thus far, we have only considered the effects of the tangential field due to the V_{DS} along the channel when examining velocity saturation effects. However, there exists a normal field originating from the gate voltage that also inhibits channel carrier mobility. Since the normal field depends on the value of V_{GS}, we find that an empirical modification to θ in Eq. (2-251) can adequately model this effect. In practice, θ is determined by a best fit to measured device characteristics.

Returning to Eq. (2-251) we note that for very short channel lengths, θ becomes large and Eq. (2-251) reduces to

$$I_D \propto (V_{GS} - V_t) \tag{2-252}$$

Thus, the FET characteristics tend toward a *linear* transfer function as the channel length becomes very small (less than 1 μm).

Equation (2-251) has a simple circuit representation. Consider the circuit of Figure 2-109, where an ideal square-law MOSFET has a resistance R_{SX} in series with the source of the FET. Assume

$$I_D = \frac{\mu C_{\text{ox}}}{2} \frac{W}{L} (V'_{GS} - V_t)^2 \tag{2-253}$$

Now

$$V_{GS} = V'_{GS} + I_D R_{SX} \tag{2-254}$$

and substituting Eq. (2-254) into Eq. (2-253) we find that

$$I_D = \frac{\mu C_{\text{ox}}}{2} \frac{W}{L} (V_{GS} - I_D R_{SX} - V_t)^2 \tag{2-255}$$

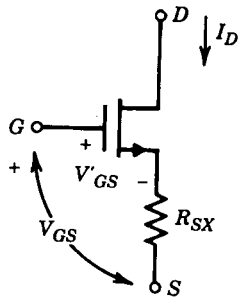


Figure 2-109 Model of velocity saturation of a MOSFET by addition of a series source resistance to an ideal square-law device.

Rearranging Eq. (2-255) we find

$$I_D = \frac{\mu C_{\text{ox}}}{2\left[1 + \mu C_{\text{ox}} \frac{W}{L} R_{SX}(V_{GS} - V_t)\right]} \frac{W}{L} (V_{GS} - V_t)^2 \tag{2-256}$$

This has exactly the same form as Eq. (2-251) if we identify

$$\theta = \mu C_{\text{ox}} \frac{W}{L} R_{SX} \tag{2-257}$$

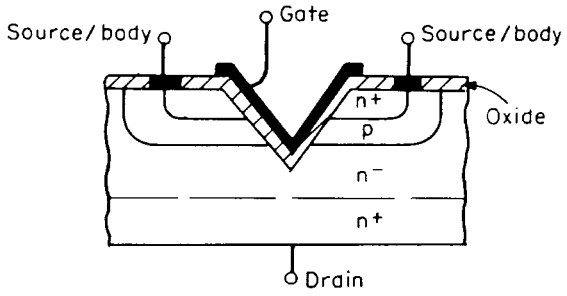
Substituting $\theta = 1/L\mathscr{E}_c$ into Eq. (2-257) we have

$$R_{SX} = \frac{1}{\mathscr{E}_c} \frac{1}{\mu C_{\text{ox}}} \frac{1}{W} \tag{2-258}$$

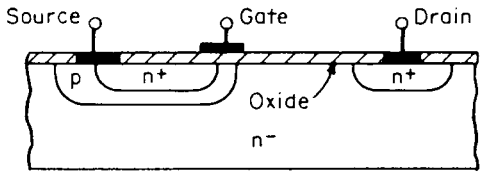
Thus, the influence of velocity saturation on the large-signal characteristics of a FET can be modeled to first order by a resistor R_{SX} in series with the source of an ideal square-law device. Note that R_{SX} varies inversely with W, as does the intrinsic physical series resistance due to source and drain contact regions. Typically, R_{SX} is larger than the physical series resistance. For $W = 2\ \mu\text{m}$, $\mu C_{\text{ox}} = 40\ \mu\text{A}/V^2$, and $\mathscr{E}_c = 1.5 \times 10^6$ V/m, we find $R_{SX} = 8\ \text{k}\Omega$.

The foregoing analysis has developed a modified large-signal mode for the MOSFET including velocity saturation effects. Small-signal MOSFET modeling for small devices can still be done using the equivalent circuit of Figure 2-107 if the values of g_m and g_{mb} are modified to account for the effects of velocity saturation using Eq. (2-251).

Figure 2-110 shows a cross section of VMOS (vertical MOS) and DMOS (double-diffused MOS) transistors. There is also a TMOS device—a Motorola variant of VMOS—and the



(a) VMOS (vertical channel MOS)



(b) DMOS (double-diffused MOS)

Figure 2-110 Cross section of (a) VMOS and (b) DMOS FETs.

now very popular LDMOS (lateral diffused MOS) power transistor (Figures 2-111 and 2-112). LDMOS FETs operate from a single supply (approximately 8 V and up) and are available for output powers surpassing 100 W at 2 GHz. The feedback capacitance is much lower than VMOS, TMOS, and earlier models, significantly improving circuit stability.

Following is a listing of Motorola high-power LDMOS transistors capable of operation up to 2 GHz:

RF High-Power LDMOS Transistors by Motorola
2.0 GHz, V_{DD} = 26 V, Class AB

Device	P_{out} Output Power (watts)	P_{in} Input Power Typical (watts)	G_{ps} (Typ.)/Freq. (dB/MHz)	IMDs (–dBc)	η Efficiency, Typ. %	θ_{JC} (°C/W)	Package/Style
MRF281S	4	0.2	13.6/2000	–30	40	8.75	458/1
MRF281Z	4	0.2	13.6/2000	–30	40	8.75	458A/1
MRF6525-5	5	0.4	11.5/2000	—	40	15	458A/1
MRF6525-10	10	1.0	10/2000	—	40	6.0	458A/1
MRF282S	10	0.5	13/2000	–30	35	2.9	458/1
MRF282Z	10	0.5	13/2000	–30	35	2.9	458A/1
MRF284	30	2.4	11.5/2000	–30	33	2.0	360B/1
MRF284S	30	2.4	11.5/2000	–30	33	2.0	360C/1
MRF286	60	4.75	11.4/2000	–30	33	.73	465/1
MRF286S	60	4.75	11.4/2000	–30	33	.73	465A/1

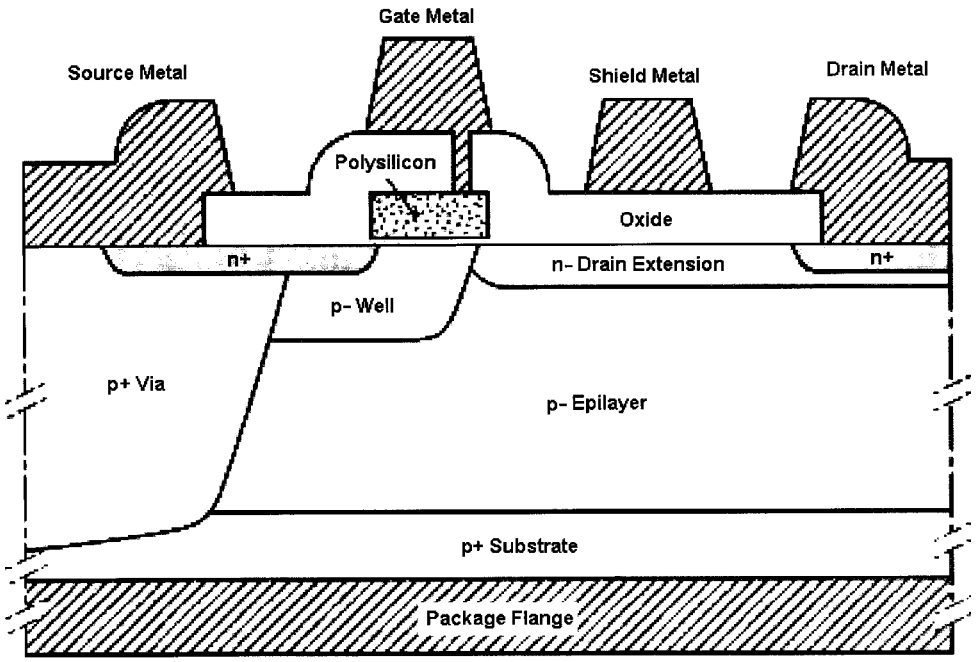


Figure 2-111 Cross section of a Philips LDMOS FET [18].

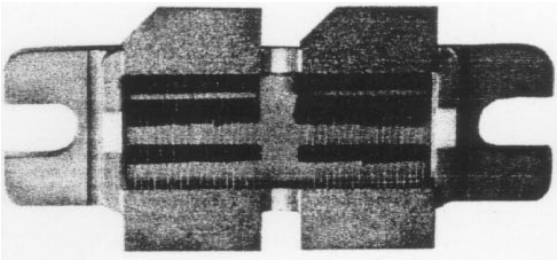


Figure 2-112 Photo of the underside of an LDMOS FET.

2-3-6 Small-Signal Models of MOSFETs

Subthreshold Conduction in MOSFETs. The MOSFET analysis of Section 2-3-3 focused on the normal region of operation where this is a well-defined conducting channel under the gate. Changes in the gate voltage are assumed to cause changes in the channel charge only, and not in the depletion region below. However, for gate voltages less than the extrapolated threshold voltage V_t, the applied gate potential affects both the depletion-region charge and the channel charge, which is then very small but not zero. The device can thus conduct finite (but small) current for $V_{GS} < V_t$, so that Eq. (2-215) is not valid in this region. The electrons in the n^+ source region of an NMOS transistor can overcome the potential barrier to the p-type substrate and enter the channel region. This process is very similar to the turn-on of a bipolar transistor, and in fact the MOSFET characteristics in this subthreshold region (also called weak inversion) are very similar to those of a bipolar transistor. Analysis [19] shows that in the subthreshold region of the MOSFET characteristics can be defined by the equation

$$I_D = k_x \frac{W}{L} e^{V_{GS}/nV_T}(1 - e^{-V_{DS}/V_T}) \tag{2-259}$$

where k_x depends on process parameters and $n \simeq 1.5$. The value of n is not equal to unity in this case (as it is for the bipolar transistor) because the applied voltage V_{GS} appears partially at the silicon surface and partially across the depletion layer.

To illustrate this effect, we show measured NMOS characteristics plotted on three different scales in Figure 2-113. In Figure 2-113a we show the transfer characteristic in the forward-active region plotted on linear scales. For this device W = 20 μm and L = 20 μm and short-channel effects are negligible. The same data are plotted in Figure 2-113b as $\sqrt{I_D}$ versus V_{GS}. The resulting straight line shows that the device characteristic is close to an ideal square law. Plots like this are commonly used to obtain V_t by extrapolation (0.7 V in this case) and also k' from the slope of the curve (54 μA/V^2 in this case). Note that near the threshold voltage, the curve deviates from the straight line representing the square law. This is the subthreshold region. The data are plotted a third time in Figure 2-113c on log–linear scales. The straight line obtained for $V_{GS} < V_t$ fits Eq. (2-259) with n = 1.5. At currents below 10^{-12} A, the effect of leakage currents becomes evident.

The major application of subthreshold operation is in very low power applications at relatively low signal frequencies. The limitation to low signal frequencies occurs because the MOSFET f_T becomes very small. Since the device capacitances at very low bias currents

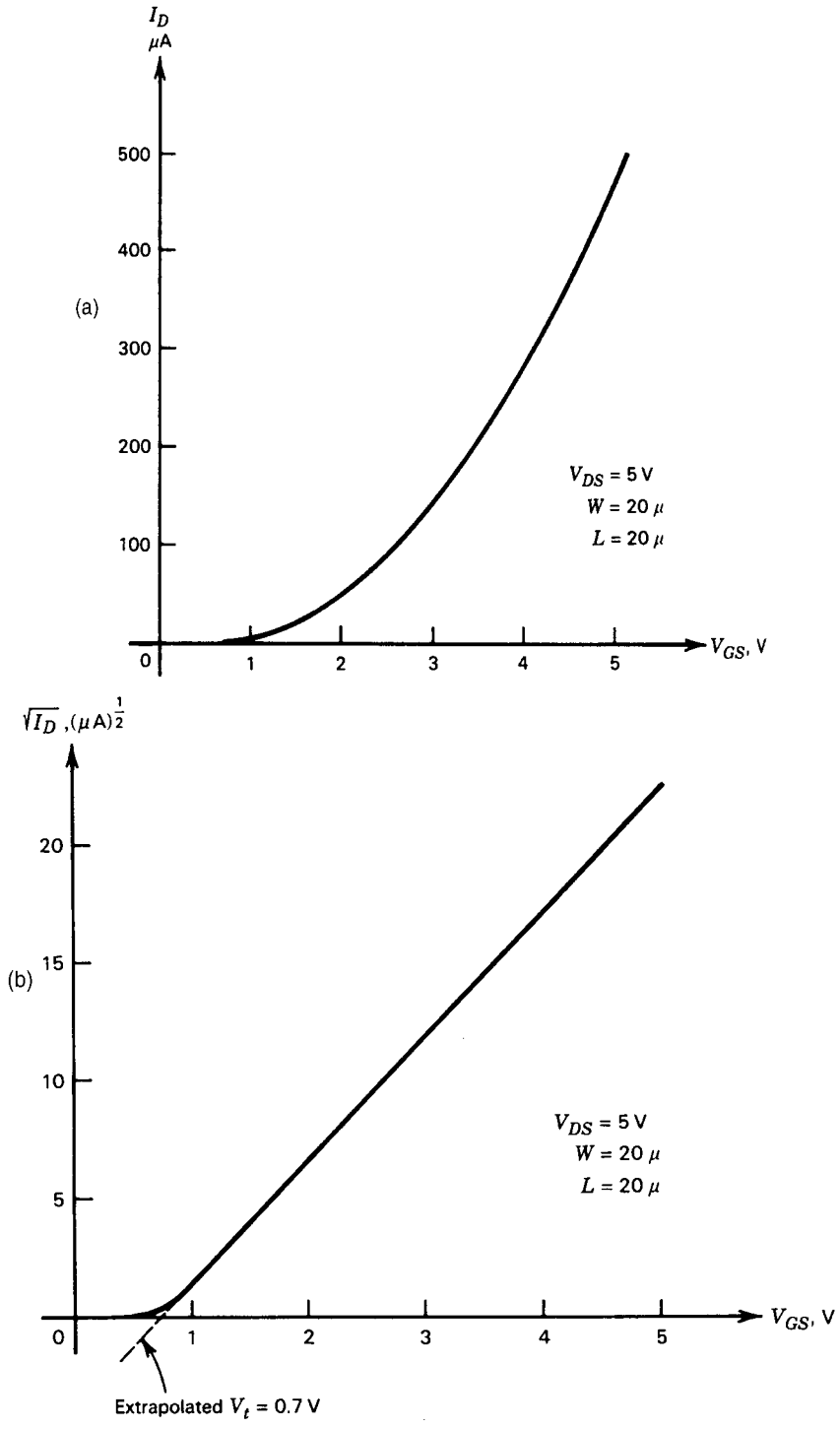


Figure 2-113 (a) Measured NMOS transfer characteristic in the forward-active region. (b) Data from part (a) plotted as $\sqrt{I_D}$ versus V_{GS} showing the square-law characteristic. (c) Data from part (a) plotted on log–linear scales showing the exponential characteristic of the subthreshold region.

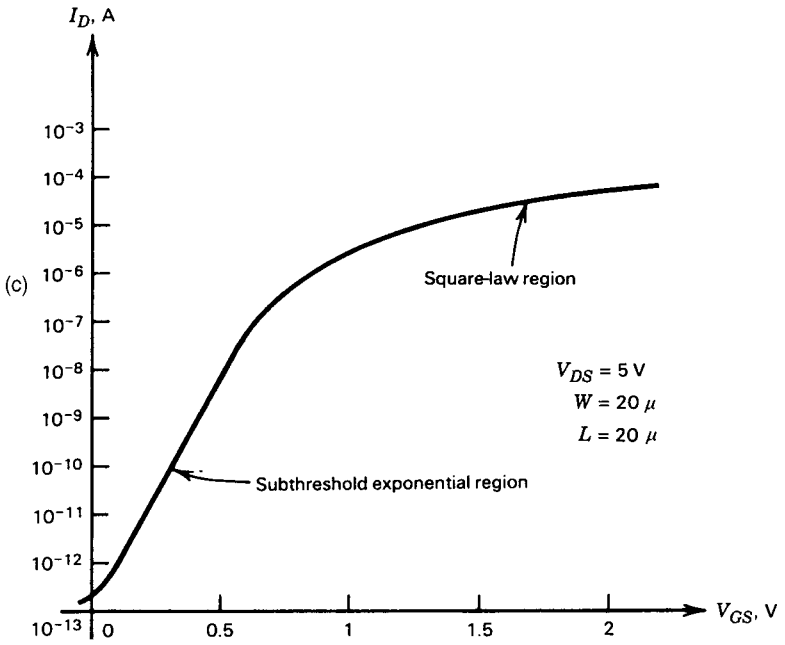


Figure 2-113. (*Continued*)

are essentially fixed, and the small-signal g_m calculated from Eq. (2-259) becomes proportional to I_D, we see that the value of f_T becomes very small at very low values of I_D.

Substrate Flow in MOSFETs. As the reverse-bias voltages on the device are increased, carriers traversing the depletion regions gain sufficient energy to create new hole–electron pairs in lattice collisions by a process known as *impact ionization.* Eventually, at high enough bias voltages, the process results in large avalanche currents. For collector–base bias voltages well below the breakdown value, a small enhanced current flow may occur across the collector–base junction due to this process with little apparent effect on the device characteristics.

Impact ionization also occurs in MOSFETs but has a significantly different effect on the device characteristics. This is because the channel electrons (for the case of NMOS) create hole–electron pairs in lattice collisions in the drain depletion region, and some of the resulting holes then flow to the substrate, creating a substrate current. (The electrons created

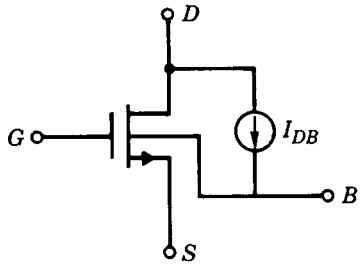


Figure 2-114 Representation of impact ionization in a MOSFET by a drain–substrate current generator.

in the process flow out the drain terminal.) The carriers created by impact ionization are thus not confined within the device as they are in a bipolar transistor. The effect of this phenomenon can be modeled by inclusion of a controlled current generator I_{DB} from drain to substrate, as shown in Figure 2-114 for an NMOS device. The magnitude of this substrate current depends on the voltage across the drain depletion region (which determines the energy of the ionizing channel electrons), and also on the drain current (which is the rate at which channel electrons enter the depletion region). It has been found empirically [20] that the current I_{DB} can be expressed as

$$I_{DB} = K_1(V_{DS} - V_{DS,\text{sat}})I_D e^{-[K_2/(V_{DS}-V_{DS,\text{sat}})]} \tag{2-260}$$

where K_1 and K_2 are process-dependent parameters and $V_{DS,\text{sat}}$ is the value of V_{DS} where the drain characteristics enter the saturation region. Typical values for NMOS devices are $K_1 = 5\ \text{V}^{-1}$ and $K_2 = 30$ V. The effect is generally much less significant in PMOS devices because the holes carrying the charge in the channel are much less efficient in creating hole–electron pairs than are energetic electrons.

The major impact of this phenomenon on circuit performance is that a parasitic resistance from drain to substrate now exists. Since the substrate of an NMOS device in a *p*-substrate process is an ac ground (the common substrate terminal must always be connected to the most negative supply voltage in the circuit), the parasitic resistance shunts the drain to ground and can be a limiting factor in many circuit designs. Differentiating Eq. (2-260) we find for the drain–substrate small-signal conductance

$$g_{db} = \frac{\partial I_{DB}}{\partial V_D}$$

$$= K_2 \frac{I_{DB}}{(V_{DS} - V_{DS,\text{sat}})^2} \tag{2-261}$$

where the gate and source are assumed held at fixed potentials.

In IC applications for RF/UHF circuits, many not-yet-solved problems remain:

1. The flicker corner frequency for MOS devices is somewhere between 100 kHz and 1 MHz.
2. Because of backplane problems in ICs, such as the substrate capacitance, also relative to the source, the performance of an IC-MOS/CMOS is not as good as expected if these effects are not considered.
3. Discrete devices, such as *n*-channel MOS devices, which are also available as dual-gate MOSFETs, can work well up to 1200 MHz. Such performance is not yet shown in silicon-based RFICs.
4. For RF power devices, such as LDMOS and others mentioned, promising results are only shown for discrete devices. There are no equivalent ICs on the market yet.

It will be interesting to see how this technology will work itself upward in frequency. Its main advantage is the BiCMOS process, which allows the mixing of analog and digital circuits nicely. Heterojunction transistors, such as SiGe, are inherently much more linear

and, for the same performance, require less current than MOS technology. On the other hand, we have not seen large-scale production SiGe-based ICs.

In the following five pages, we reproduce the datasheet for the Siemens (now Infineon Technologies) BF999 single-gate MOSFET.* Above 500 MHz, a dual-gate (also referred to as tetrode) MOSFET, such as the BF998, is preferred over a single-gate device. The main reasons for this are:

1. Internally, it is equivalent to two single-gate devices connected in a cascade arrangement, with the output device operating in the grounded-gate configuration, which entirely avoids the Miller effect and therefore remains stable at higher frequencies, operating up to 1.2 GHz.
2. The second gate can be used for AGC. As the Gate 2 voltage changes, the gain varies heavily.
3. By applying a local oscillator signal to Gate 2, this configuration can be used as a mixer.

The major drawbacks of these devices are their wide tolerances and temperature sensitivity. They follow to some degree the CMOS design but seem to have better performance.

The BF995, also called a tetrode, is in reality a dual-gate MOSFET specified to only 300 MHz; however, we are reproducing its datasheet* because the one more applicable for wireless circuits, the BF1005, does not have such a complete set of data.

To demonstrate the capabilities of low-voltage LDMOS transistors, we show the MRF5003 datasheet.† This device is intended for low-voltage applications down to 7.5 V. To the best of our knowledge, this is the lowest-voltage LDMOS transistor on the market. Other members of this family, which produce 5 W (MRF5005) and 15 W (MRF5015), begin to require higher supply voltages. Just to make sure it is not forgotten, the advantage of LDMOS is, besides lower-voltage operation, much-reduced feedback capacitance. LDMOS devices are now available that operate above 1 GHz and produce 60 W. There is now some competition with the recently introduced NEC GaAsFET, which goes up to 100 W in a similar frequency range.

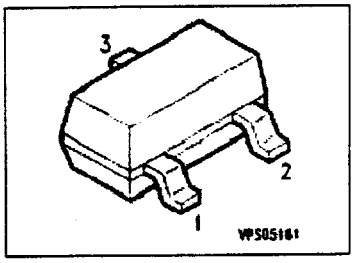
*Reproduced with permission.
†Copyright of Motorola, used by permission.

SIEMENS

Silicon N Channel MOSFET Triode

BF 999

- For high-frequency stages up to 300 MHz, preferably in FM applications



Type	Marking	Ordering Code (tape and reel)	Pin Configuration 1	2	3	Package
BF 999	LB	Q62702-F1132	G	D	S	SOT-23

Maximum Ratings

Parameter	Symbol	Values	Unit
Drain-source voltage	V_{DS}	20	V
Drain current	I_D	30	mA
Gate-source peak current	$\pm I_{GSM}$	10	
Total power dissipation, $T_A \leq 60$ °C	P_{tot}	200	mW
Storage temperature range	T_{stg}	– 55 ... + 150	°C
Channel temperature	T_{ch}	150	

Thermal Resistance

Junction - ambient [1]	$R_{th\,JA}$	≤ 450	K/W

[1] Package mounted on alumina 15 mm × 16.7 mm × 0.7 mm.

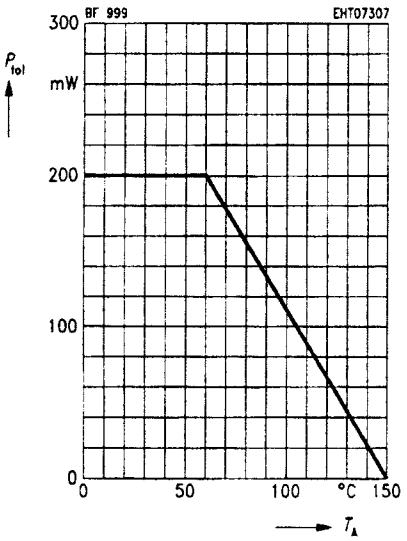
SIEMENS **BF 999**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

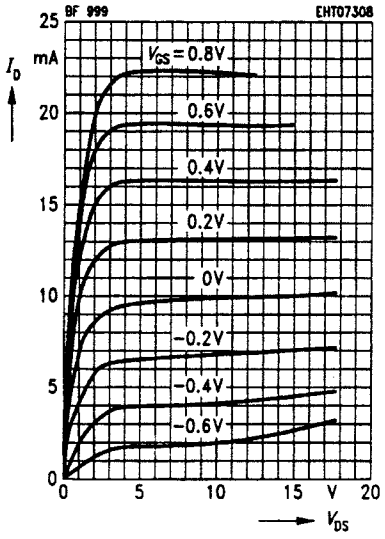
Parameter	Symbol	Values			Unit
		min.	typ.	max.	
DC Characteristics					
Drain-source breakdown voltage I_D = 10 µA, – V_{GS} = 4 V	$V_{(BR)\,DS}$	20	–	–	V
Gate-source breakdown voltage ± I_{GS} = 10 mA, V_{DS} = 0	± $V_{(BR)\,GSS}$	6.5	–	12	
Gate-source leakage current ± V_{GS} = 5 V, V_{DS} = 0	± I_{GSS}	–	–	50	nA
Drain current V_{DS} = 10 V, V_{GS} = 0	I_{DSS}	5	–	18	mA
Gate-source pinch-off voltage V_{DS} = 10 V, I_D = 20 µA	– $V_{GS\,(p)}$	–	–	2.5	V
AC Characteristics					
Forward transconductance V_{DS} = 10 V, I_D = 10 mA, f = 1 kHz	g_{fs}	14	16	–	mS
Gate input capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{gss}	–	2.5	–	pF
Reverse transfer capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{dg}	–	25	–	fF
Output capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{dss}	–	1	–	pF
Power gain (test circuit) V_{DS} = 10 V, I_D = 10 mA, f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS	G_p	–	25	–	dB
Noise figure (test circuit) V_{DS} = 10 V, I_D = 10 mA, f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS	F	–	1	–	

SIEMENS **BF 999**

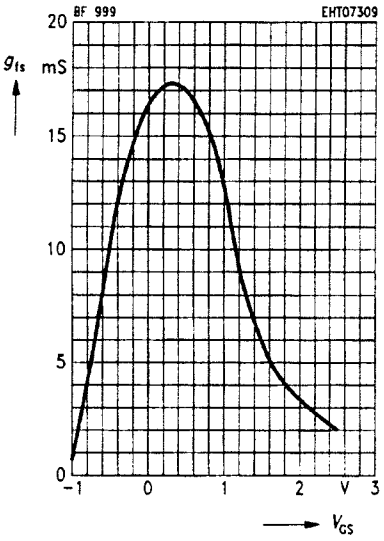
Total power dissipation $P_{tot} = f(T_A)$



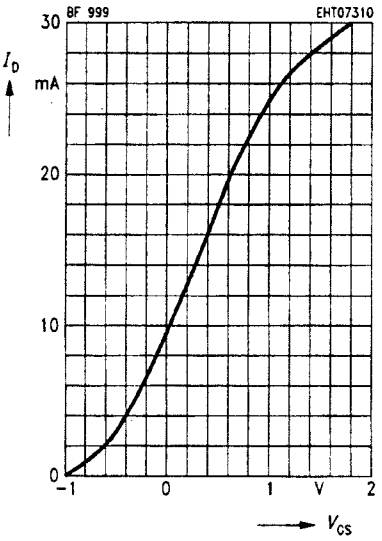
Output characteristics $I_D = f(V_{DS})$



Gate transconductance $g_{fs} = f(V_{GS})$
V_{DS} = 10 V, I_{DSS} = 10 mA, f = 1 kHz

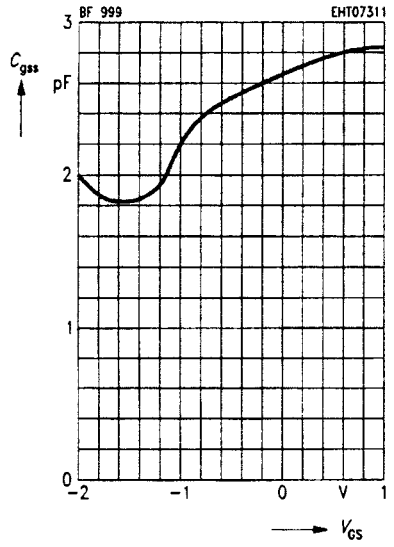


Drain current $I_D = f(V_{GS})$
V_{DS} = 10 V

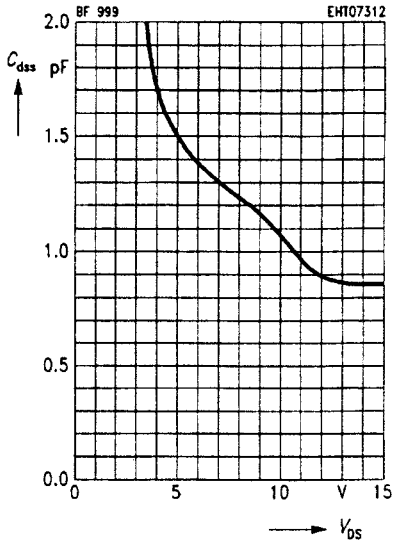


SIEMENS **BF 999**

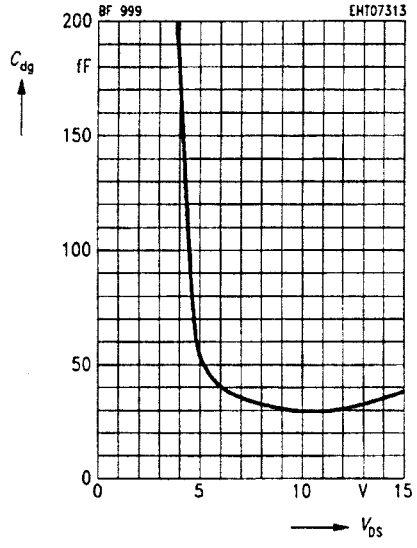
Gate input capacitance $C_{gss} = f(V_{GS})$
V_{DS} = 10 V, I_{DSS} = 10 mA, f = 1 MHz



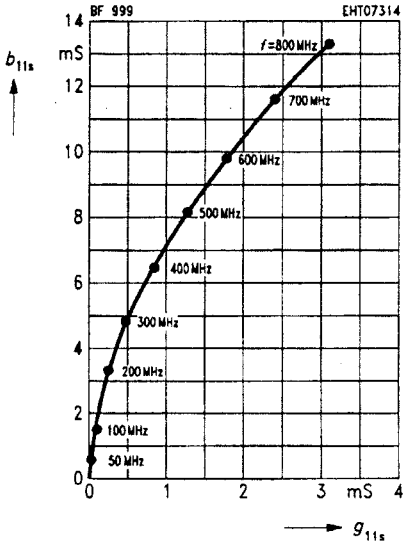
Output capacitance $C_{dss} = f(V_{DS})$
V_{GS} = 0, I_{DSS} = 10 mA, f = 1 MHz



Reverse transfer capacitance
$C_{dg} = f(V_{DS})$
I_{DSS} = 10 mA, f = 1 MHz, V_{GS} = 0

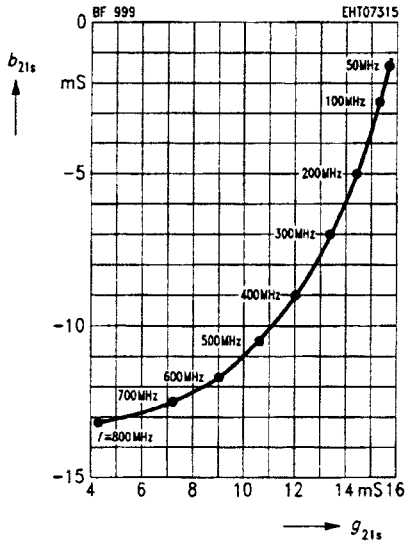


Gate input admittance y_{11s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)

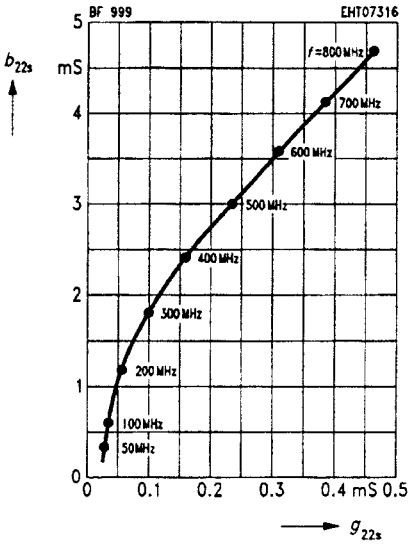


SIEMENS **BF 999**

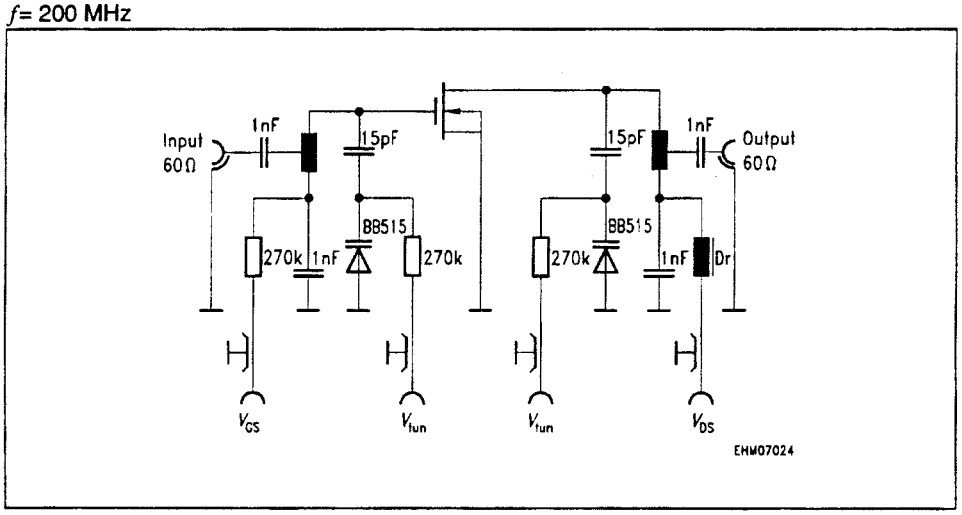
Gate forward transfer admittance y_{21s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)



Output admittance y_{22s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)



Test circuit for power gain and noise figure
f= 200 MHz

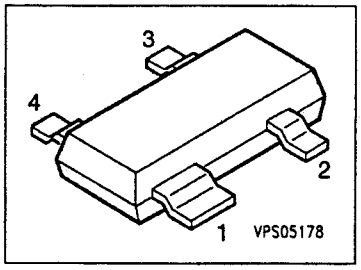


SIEMENS

Silicon N Channel MOSFET Tetrode

BF 995

- For input and mixer stages in FM and VHF TV tuners



Type	Marking	Ordering Code (tape and reel)	Pin Configuration 1	2	3	4	Package
BF 995	MB	Q62702-F936	S	D	G_2	G_1	SOT-143

Maximum Ratings

Parameter	Symbol	Values	Unit
Drain-source voltage	V_{DS}	20	V
Drain current	I_D	30	mA
Gate 1/gate 2 peak source current	$\pm I_{G1/2SM}$	10	
Total power dissipation, T_S < 76 °C	P_{tot}	200	mW
Storage temperature range	T_{stg}	– 55 ... + 150	°C
Channel temperature	T_{ch}	150	

Thermal Resistance

Junction - soldering point	$R_{th\,JS}$	< 370	K/W

SIEMENS **BF 995**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

Parameter	Symbol	Values			Unit
		min.	typ.	max.	
DC Characteristics					
Drain-source breakdown voltage I_D = 10 µA, – V_{G1S} = – V_{G2S} = 4 V	$V_{(BR)\,DS}$	20	–	–	V
Gate 1 source breakdown voltage ± I_{G1S} = 10 mA, V_{G2S} = V_{DS} = 0	± $V_{(BR)\,G1SS}$	8.5	–	14	
Gate 2 source breakdown voltage ± I_{G2S} = 10 mA, V_{G1S} = V_{DS} = 0	± $V_{(BR)\,G2SS}$	8.5	–	14	
Gate 1 source leakage current ± V_{G1S} = 5 V, V_{G2S} = V_{DS} = 0	± I_{G1SS}	–	–	50	nA
Gate 2 source leakage current ± V_{G2S} = 5 V, V_{G1S} = V_{DS} = 0	± I_{G2SS}	–	–	50	
Drain current V_{DS} = 15 V, V_{G1S} = 0, V_{G2S} = 4 V	I_{DSS}	4	–	20	mA
Gate 1 source pinch-off voltage V_{DS} = 15 V, V_{G2S} = 4 V, I_D = 20 µA	– $V_{G1S\,(p)}$	–	–	2.5	V
Gate 2 source pinch-off voltage V_{DS} = 15 V, V_{G1S} = 0, I_D = 20 µA	– $V_{G2S\,(p)}$	–	–	2.0	

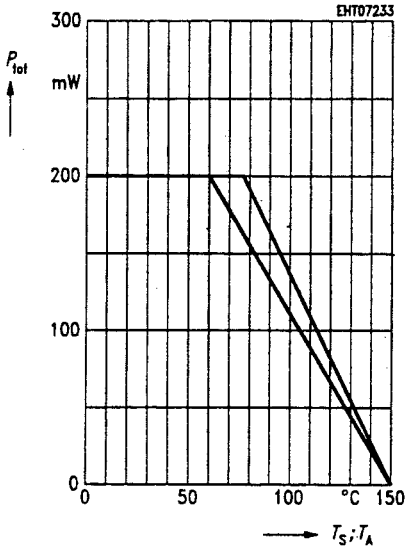
SIEMENS **BF 995**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

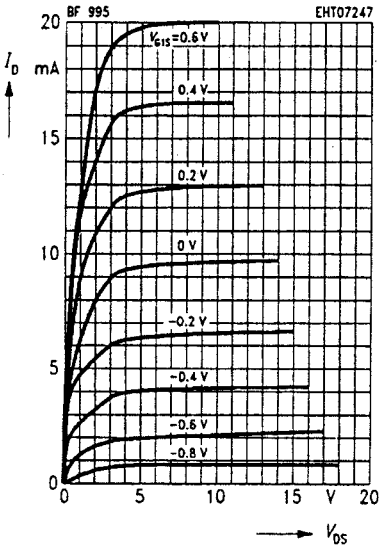
Parameter	Symbol	Values			Unit
		min.	typ.	max.	
AC Characteristics					
Forward transconductance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 kHz	g_{fs}	12	17	–	mS
Gate 1 input capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{g1ss}	–	3.6	–	pF
Gate 2 input capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{g2ss}	–	1.6	–	
Feedback capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{dg1}	–	25	–	fF
Output capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{dss}	–	1.6	–	pF
Power gain V_{DS} = 15 V, I_D = 10 mA f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS 2 Δf = 12 MHz (see test circuit 1)	G_{ps}	–	23	–	dB
Noise figure V_{DS} = 15 V, I_D = 10 mA f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS (see test circuit 1)	F	–	1.1	–	
Gain control range V_{DS} = 15 V, V_{G2S} = 4 ... – 2 V, f = 200 MHz (see test circuit 1)	ΔG_{ps}	–	50	–	
Mixer gain (additive) V_{DS} = 15 V, V_{G2S} = 6 V, R_S = 220 Ω f = 200 MHz, f_F = 36 MHz 2 Δf_F = 5 MHz, V_{osc} = 0.5 V (see test circuit 2)	G_{psc}	–	16	–	
Mixer gain (multiplicative) V_{DS} = 15 V, V_{G1S} = 1.7 V, V_{G2S} = 2.5 V R_S = 220 Ω, f = 200 MHz, f_F = 36 MHz 2 Δf_F = 5 MHz, V_{osc} = 2 V (see test circuit 3)	G_{psc}	–	18	–	

SIEMENS **BF 995**

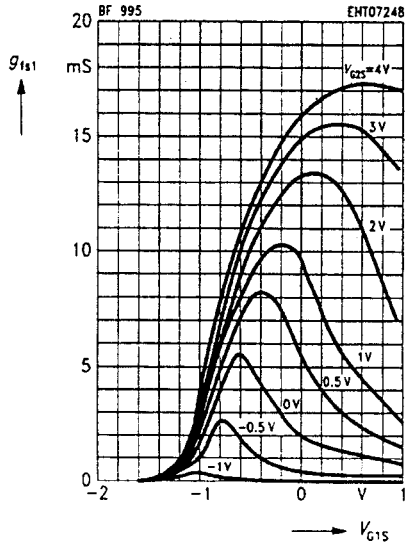
Total power dissipation $P_{tot} = f(T_A)$



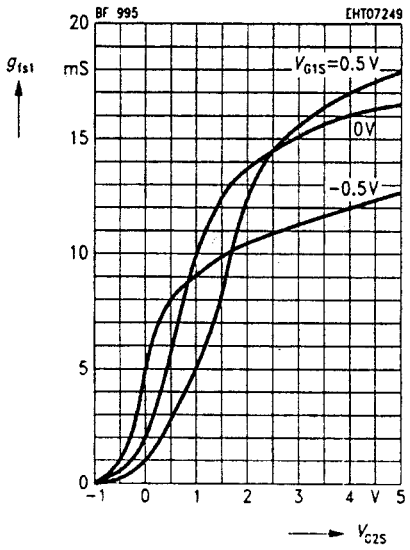
Output characteristics $I_D = f(V_{DS})$
V_{G2S} = 4 V



Gate 1 forward transconductance $g_{fs1} = f(V_{G1S})$
V_{DS} = 15 V, I_{DSS} = 10 mA, f = 1 kHz

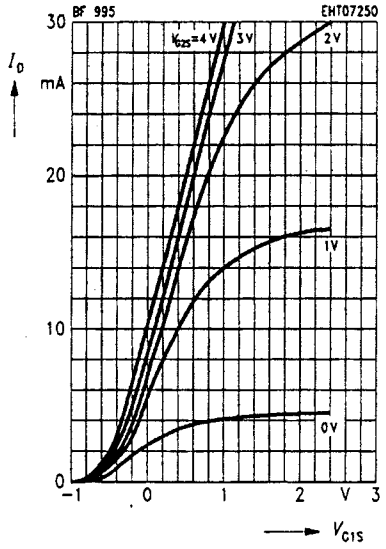


Gate 1 forward transconductance $g_{fs1} = f(V_{G2S})$
V_{DS} = 15 V, I_{DSS} = 10 mA, f = 1 kHz

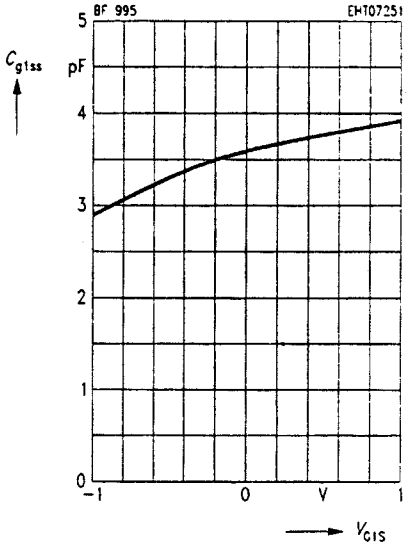


SIEMENS **BF 995**

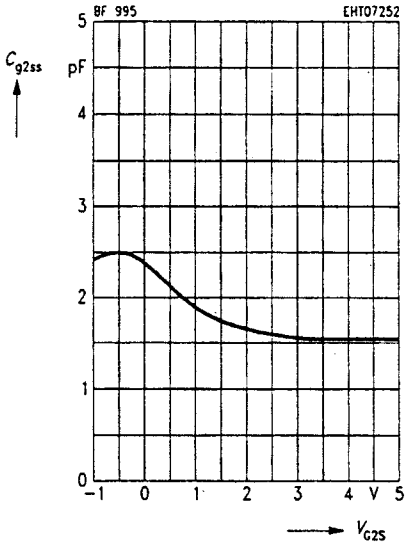
Drain current $I_D = f(V_{G1S})$
V_{DS} = 15 V



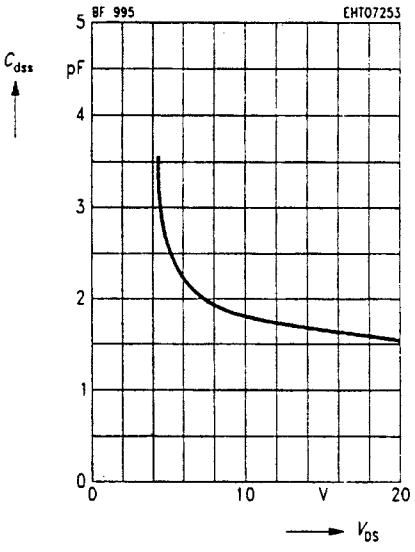
Gate 1 input capacitance $C_{g1ss} = f(V_{G1S})$
V_{G2S} = 4 V, V_{DS} = 15 V
I_{DSS} = 10 mA, f = 1 MHz



Gate 2 input capacitance $C_{g2ss} = f(V_{G2S})$
V_{G1S} = 0 V, V_{DS} = 15 V
I_{DSS} = 10 mA, f = 1 MHz

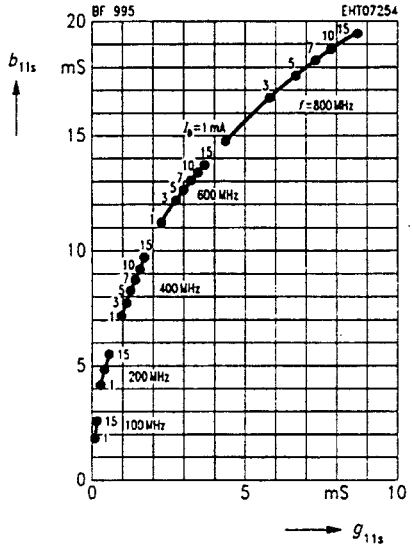


Output capacitance $C_{dss} = f(V_{DS})$
V_{G1S} = 0 V, V_{G2S} = 4 V
I_{DSS} = 10 mA, f = 1 MHz

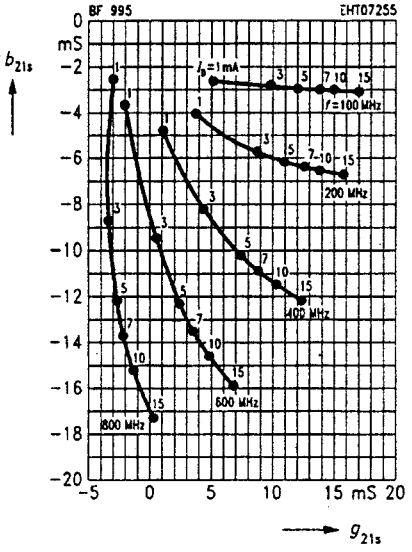


SIEMENS **BF 995**

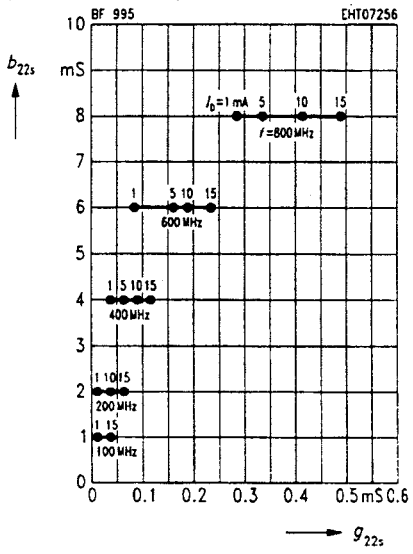
Gate 1 input admittance y_{11s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)



Gate 1 forward transfer admittance y_{21s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)

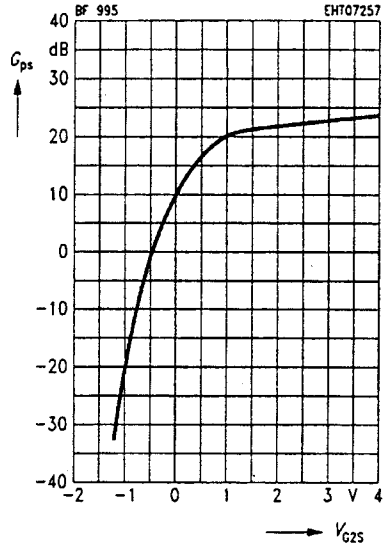


Output admittance y_{22s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)

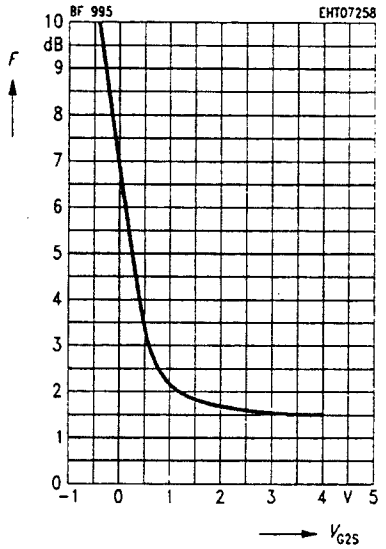


SIEMENS **BF 995**

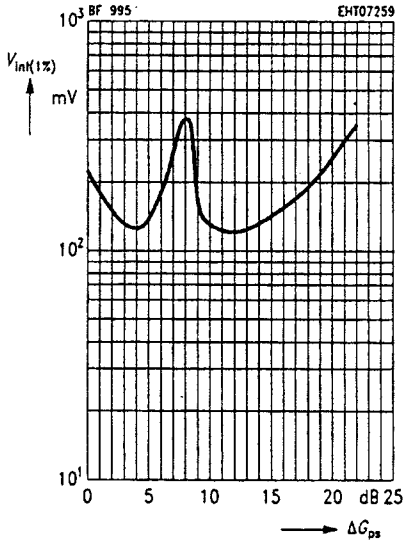
Power gain $G_{ps} = f(V_{G2S})$
V_{DS} = 15 V, V_{G1S} = 0 V, I_{DSS} = 10 mA
f = 200 MHz (see test circuit 1)



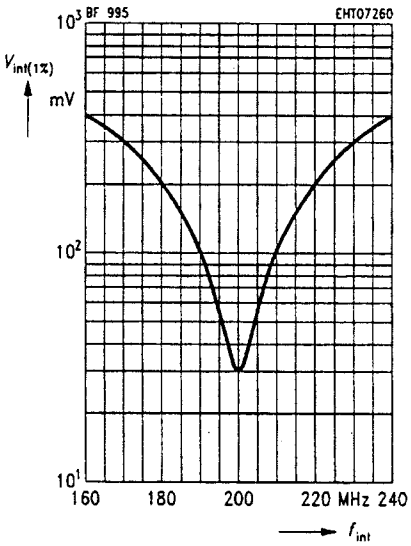
Noise figure $F = f(V_{G2S})$
V_{DS} = 15 V, V_{G1S} = 0 V, I_{DSS} = 10 mA
f = 200 MHz (see test circuit 1)



Interference voltage for 1% cross modulation $V_{int\,(1\%)} = f(\Delta G_{ps})$[1)]
V_{DS} = 15 V, V_{G1S} = 0, f = 200 MHz
f_{int} = 221 MHz (see test circuit 1)



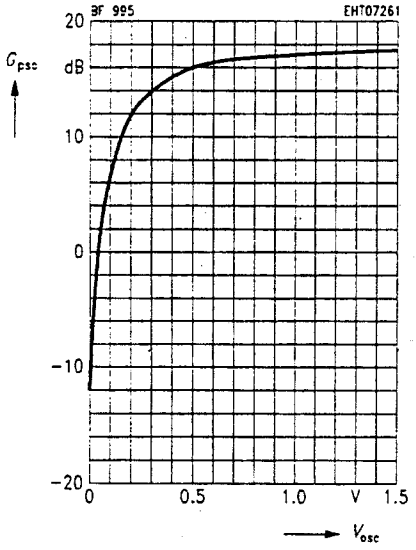
Interference voltage for 1% cross modulation $V_{int\,(1\%)} = f(f_{int})$[1)]
V_{DS} = 15 V, V_{G2S} = 4 V, V_{G1S} = 0
f = 200 MHz (see test circuit 1)



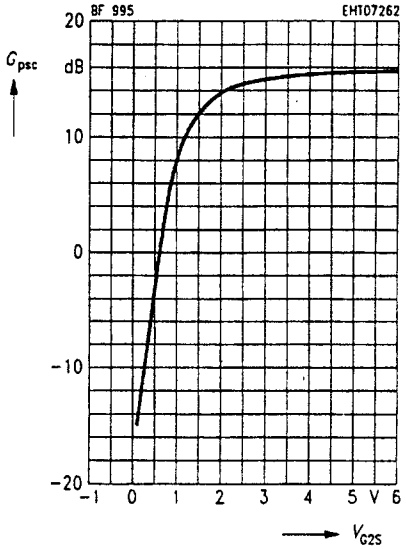
1) For footnote refer to the last page of this data sheet.

SIEMENS **BF 995**

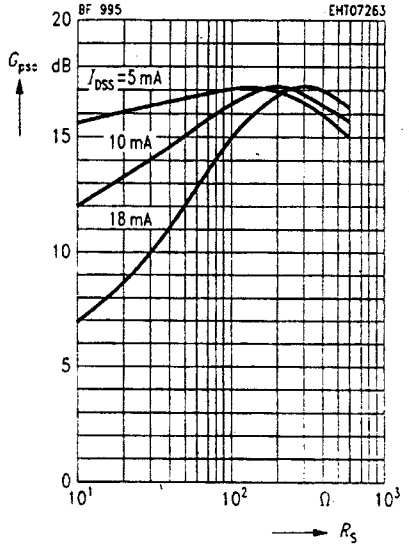
Mixer gain (additive) $G_{psc} = f(V_{osc})$
V_D = 15 V, V_{G1S} = 0, V_{G2S} = 6 V
R_S = 220 Ω, I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)



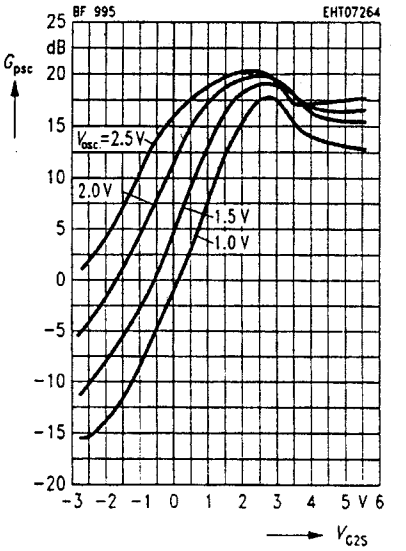
Mixer gain (additive) $G_{psc} = f(V_{G2S})$
V_D = 15 V, V_{G1S} = 0, R_S = 220 Ω
V_{osc} = 0.5 V, I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)



Mixer gain (additive) $G_{psc} = f(R_S)$
V_D = 15 V, V_{G1S} = 0, V_{G2S} = 6 V
V_{osc} = 0.5 V, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)

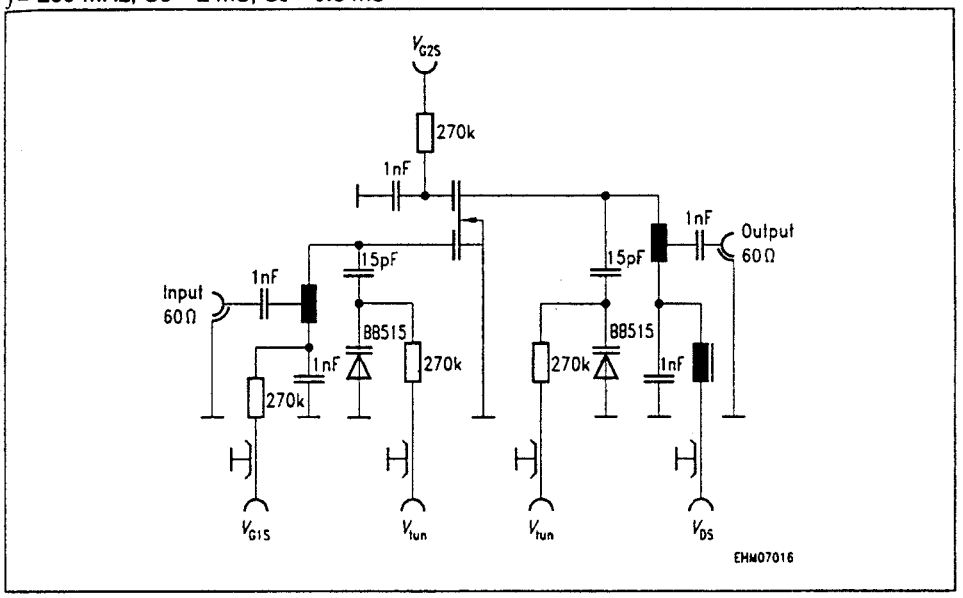


Mixer gain (multiplicative) $G_{psc} = f(V_{G2S})$
V_D = 15 V, V_{G1S} = 1.7 V, R_S = 200 Ω
I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 3)

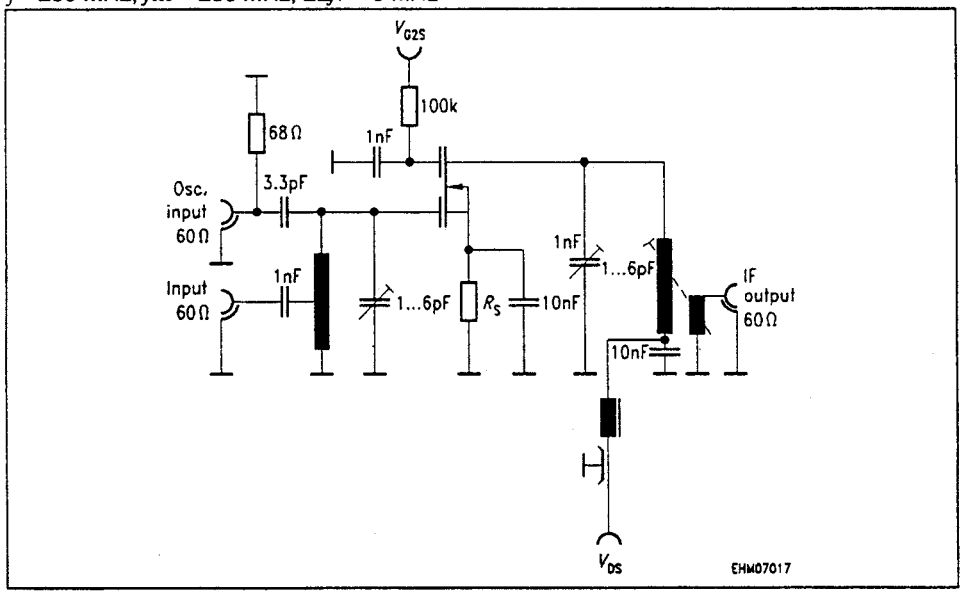


SIEMENS **BF 995**

Test circuit 1 for power gain, noise figure and cross modulation
f= 200 MHz, G_S = 2 mS, G_L = 0.5 mS

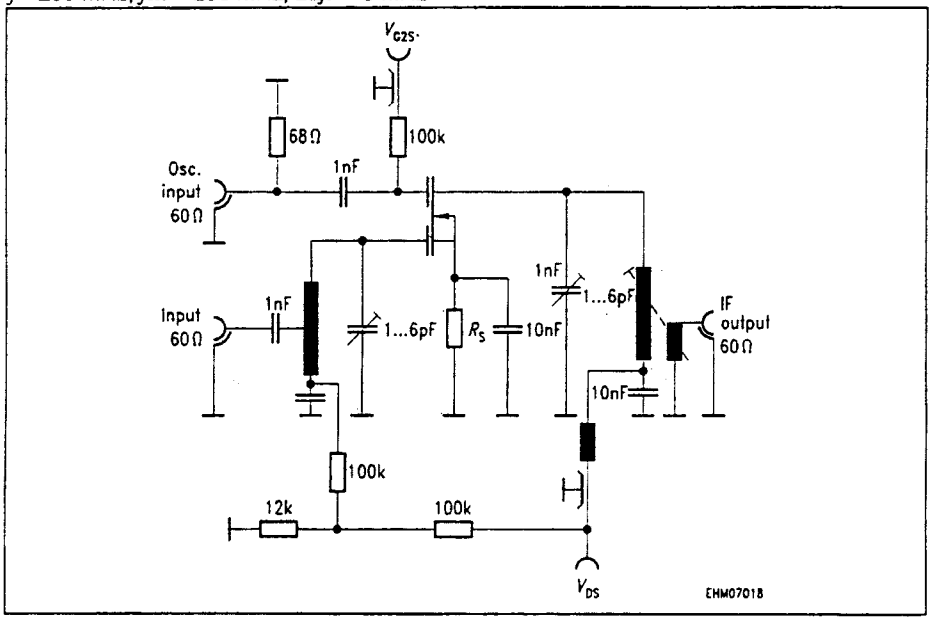


Test circuit 2 for mixer gain (additive)
f= 200 MHz, f_{osc} = 236 MHz, $2\Delta f_F$ = 5 MHz



SIEMENS **BF 995**

Test circuit 3 for mixer gain (multiplicative)
f = 200 MHz, f_{osc} = 236 MHz, $2\Delta f_F$ = 5 MHz



1) $V_{int\,(1\%)}$ is the rms value of half the emf (terminal voltage at matching) of a 100 % sine modulated TV carrier at an internal generator resistance of 60 Ω, causing 1 % amplitude modulation on the active carrier.

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF5003/D

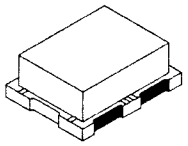
The RF MOSFET Line
RF Power Field Effect Transistor
N–Channel Enhancement–Mode

MRF5003

3.0 W, 7.5 V, 512 MHz
N–CHANNEL
BROADBAND
RF POWER FET

The MRF5003 is designed for broadband commercial and industrial applications at frequencies to 520 MHz. The high gain and broadband performance of this device makes it ideal for large–signal, common source amplifier applications in 7.5 Volt and 12.5 Volt mobile, portable, and base station FM equipment.

- Guaranteed Performance at 512 MHz, 7.5 Volts
  Output Power = 3.0 Watts
  Power Gain = 9.5 dB
  Efficiency = 45%
- Characterized with Series Equivalent Large–Signal Impedance Parameters
- S–Parameter Characterization at High Bias Levels
- Excellent Thermal Stability
- All Gold Metal for Ultra Reliability
- Capable of Handling 20:1 VSWR, @ 15.5 Vdc, 512 MHz, 2.0 dB Overdrive
- Suitable for 12.5 Volt Applications
- True Surface Mount Package
- Available in Tape and Reel by Adding R1 Suffix to Part Number.
  R1 Suffix = 500 Units per 16 mm, 7 inch Reel.
- Circuit board photomaster available upon request by contacting
  RF Tactical Marketing in Phoenix, AZ.



CASE 430–01, STYLE 2

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Drain–Source Voltage	V_{DSS}	36	Vdc
Drain–Gate Voltage (R_{GS} = 1.0 Meg Ohm)	V_{DGR}	36	Vdc
Gate–Source Voltage	V_{GS}	±20	Vdc
Drain Current — Continuous	I_D	1.7	Adc
Total Device Dissipation @ T_C = 25°C Derate above 25°C	P_D	12.5 0.07	Watts W/°C
Storage Temperature Range	T_{stg}	–65 to +150	°C
Operating Junction Temperature	T_J	200	°C

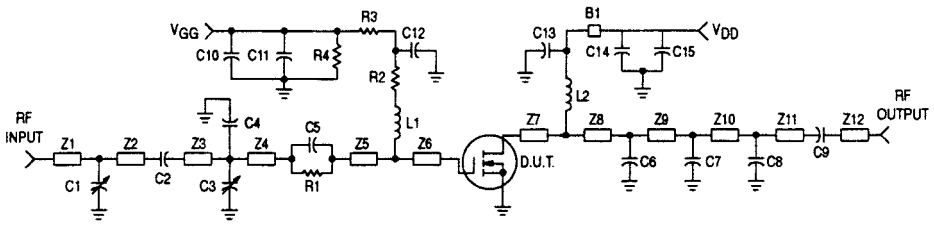
THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	14	°C/W

NOTE – **CAUTION** – MOS devices are susceptible to damage from electrostatic charge. Reasonable precautions in handling and packaging MOS devices should be observed.

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted.)

Characteristic		Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS						
Drain–Source Breakdown Voltage (V_{GS} = 0, I_D = 2.5 mAdc)		$V_{(BR)DSS}$	36	—	—	Vdc
Zero Gate Voltage Drain Current (V_{DS} = 15 Vdc, V_{GS} = 0)		I_{DSS}	—	—	1.0	mAdc
Gate–Source Leakage Current (V_{GS} = 20 Vdc, V_{DS} = 0)		I_{GSS}	—	—	1.0	μAdc
ON CHARACTERISTICS						
Gate Threshold Voltage (V_{DS} = 10 Vdc, I_D = 5.0 mAdc)		$V_{GS(th)}$	1.25	2.25	3.5	Vdc
Drain–Source On–Voltage (V_{GS} = 10 Vdc, I_D = 0.5 Adc)		$V_{DS(on)}$	—	—	0.375	Vdc
Forward Transconductance (V_{DS} = 10 Vdc, I_D = 0.5 Adc)		g_{fs}	0.6	—	—	mho
DYNAMIC CHARACTERISTICS						
Input Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{iss}	—	16.5	—	pF
Output Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{oss}	—	37	—	pF
Reverse Transfer Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{rss}	3.5	4.4	5.4	pF
FUNCTIONAL TESTS (In Motorola Test Fixture)						
Common–Source Amplifier Power Gain (V_{DD} = 7.5 Vdc, P_{out} = 3.0 W, I_{DQ} = 50 mA)	f = 512 MHz f = 175 MHz	G_{ps}	 9.5 —	 10.5 15	 — —	dB
Drain Efficiency (V_{DD} = 7.5 Vdc, P_{out} = 3.0 W, I_{DQ} = 50 mA)	f = 512 MHz f = 175 MHz	h	 45 —	 50 55	 — —	%



C1, C3, C7, C8	0 to 20 pF Johanson	Z1	0.350″ x 0.08″ Microstrip
C2, C9	56 pF, 100 mil Chip	Z2	0.190″ x 0.08″ Microstrip
C4	10 pF, 100 mil Chip	Z3	0.800″ x 0.08″ Microstrip
C5	47 pF, Miniature Clamped Mica Capacitor	Z4	0.380″ x 0.08″ Microstrip
C6	22 pF, 100 mil Chip	Z5	0.150″ x 0.08″ Microstrip
C10, C15	10 µF, 50 V, Electrolytic	Z6	0.285″ x 0.08″ Microstrip
C11, C14	0.1 µF, Capacitor	Z7	0.340″ x 0.08″ Microstrip
C12	1000 pF, 100 mil Chip	Z8	0.070″ x 0.08″ Microstrip
C13	160 pF, 100 mil Chip	Z9	0.280″ x 0.08″ Microstrip
R1	35 Ω, 1/4 W Carbon	Z10	0.840″ x 0.08″ Microstrip
R2	30 Ω, 0.1 W Chip	Z11	0.180″ x 0.08″ Microstrip
R3	1.0 kΩ, 0.1 W Chip	Z12	0.600″ x 0.08″ Microstrip
R4	1.0 MΩ, 1/4 W Carbon	L1	7 Turns, 0.076″ ID, #24 AWG Enamel
B1	Fair Rite Products Short Ferrite Bead (2743021446)	L2	5 Turns, 0.126″ ID, #20 AWG Enamel
Board — Glass Teflon®, 31 mils		Input/Output Connectors — Type N	

Note: Plated ceramic part locators (0.1″ x 0.15″) soldered onto Z6 and Z7.

Figure 1. 512 MHz Narrowband Test Circuit

TYPICAL CHARACTERISTICS

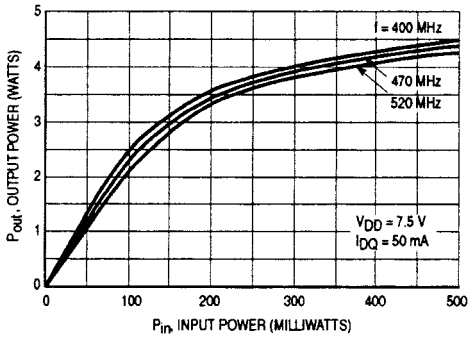


Figure 2. Output Power versus Input Power

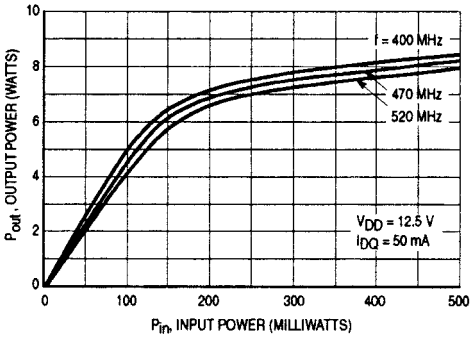


Figure 3. Output Power versus Input Power

TYPICAL CHARACTERISTICS

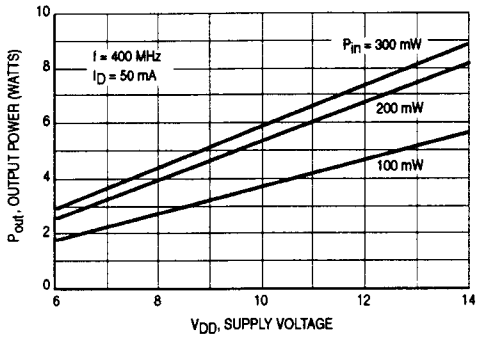


Figure 4. Output Power versus Supply Voltage

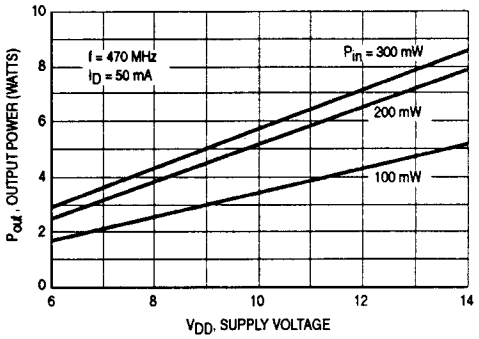


Figure 5. Output Power versus Supply Voltage

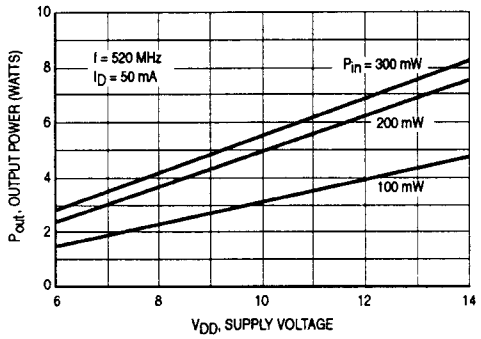


Figure 6. Output Power versus Supply Voltage

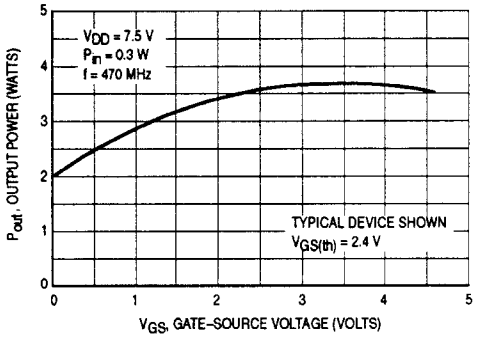


Figure 7. Output Power versus Gate Voltage

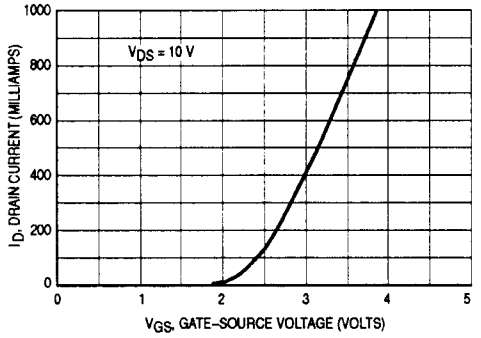


Figure 8. Drain Current versus Gate Voltage (Typical Device Shown)

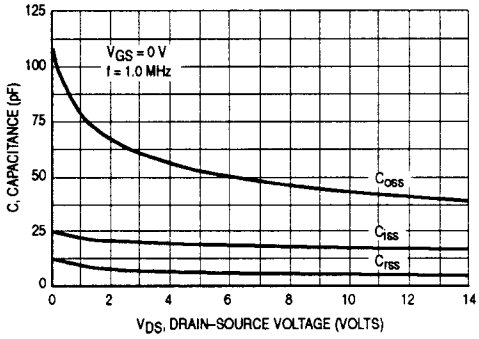


Figure 9. Capacitance versus Voltage

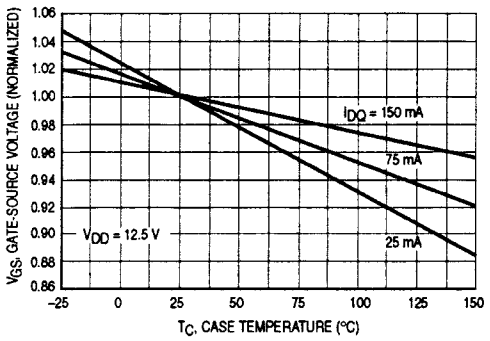


Figure 10. Gate–Source Voltage versus Case Temperature

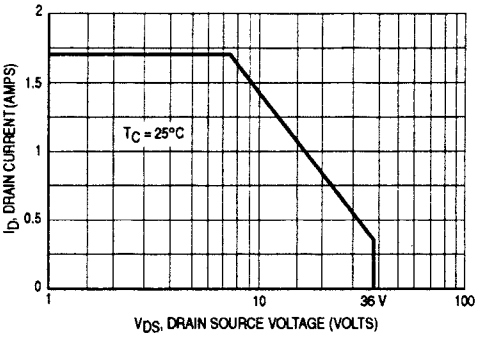
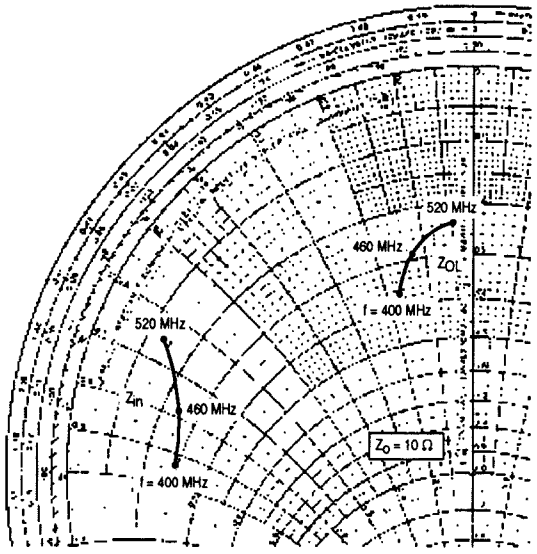


Figure 11. Maximum Rated Forward Biased Safe Operating Area



V_{DD} = 7.5 V, I_{DQ} = 50 mA, P_{out} = 3.0 W

f MHz	Z_{in} Ohms	Z_{OL}^* Ohms
400	2.8 – j9.2	3.6 – j1.7
430	2.7 – j8.5	3.3 – j1.5
460	2.5 – j7.8	2.7 – j1.1
490	2.0 – j7.2	2.5 – j0.8
520	1.3 – j6.5	2.4 – j0.5

Z_{in} = Conjugate of source impedance with parallel 35 Ω resistor and 47 pF capacitor in series with gate.

Z_{OL}^* = Conjugate of the load impedance at given output power, voltage, frequency, and η_D > 50%.

Note: Z_{ol}^* was chosen based on tradeoffs between gain, drain efficiency, and device stability.

Figure 12. Series Equivalent Input and Output Impedance

Table 1. Common Source Scattering Parameters (V_{DS} = 10 V)

I_D = 50 mA

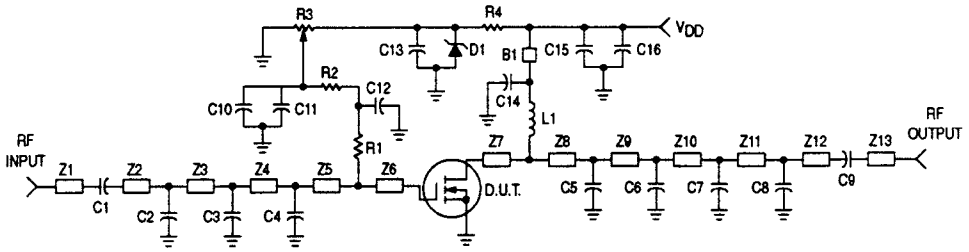
f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.69	−90	10.8	117	0.07	29	0.74	−119
100	0.58	−120	6.0	96	0.08	10	0.78	−146
200	0.58	−139	3.0	75	0.08	−7	0.81	−161
300	0.64	−147	1.9	61	0.07	−16	0.84	−166
400	0.70	−152	1.3	50	0.06	−21	0.86	−169
500	0.75	−157	0.99	41	0.05	−24	0.88	−172
700	0.82	−165	0.61	28	0.03	−15	0.92	−176
850	0.86	−171	0.45	21	0.02	13	0.94	−179
1000	0.89	−176	0.34	16	0.02	47	0.95	178

I_D = 500 mA

f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.76	−124	15.0	109	0.04	23	0.76	−151
100	0.72	−150	7.9	94	0.04	12	0.81	−165
200	0.72	−163	4.0	80	0.04	6	0.83	−172
300	0.73	−168	2.6	71	0.04	5	0.84	−175
400	0.75	−171	1.9	62	0.04	7	0.85	−176
500	0.77	−173	1.5	55	0.03	12	0.86	−178
700	0.81	−177	0.97	42	0.03	29	0.89	−180
850	0.84	−180	0.75	35	0.03	44	0.90	178
1000	0.86	177	0.60	29	0.04	55	0.92	176

I_D = 1.0 A

f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.80	−125	14.6	110	0.04	23	0.75	−155
100	0.76	−150	7.8	95	0.04	10	0.81	−167
200	0.76	−164	3.9	81	0.04	1	0.83	−173
300	0.77	−169	2.6	71	0.04	−3	0.84	−175
400	0.79	−172	1.9	63	0.03	−5	0.85	−176
500	0.80	−174	1.4	56	0.03	−5	0.86	−177
700	0.83	−178	0.95	43	0.03	−1	0.88	−179
850	0.85	179	0.73	35	0.02	9	0.90	179
1000	0.87	177	0.58	28	0.02	22	0.91	178



C1, C9	100 pF 100 mil Chip
C2	16 pF, 100 mil Chip
C3	24 pF, 100 mil Chip
C4	68 pF, 100 mil Chip
C5	51 pF, 100 mil Chip
C6	39 pF, 100 mil Chip
C7	6.2 pF, 100 mil Chip
C8	9.1 pF, 100 mil Chip
C10, C15	39000 pF, 100 mil Chip
C11, C16	10 µF, 50 V Electrolytic
C12	10000 pF, 100 mil Chip
B1	Fair Rite Products Short Ferrite Bead (2743021446)
C13	0.1 µF, 100 mil Chip
C14	160 pF, 100 mil Chip
R1	43 Ω, 0.1 W Chip Resistor
R2	1000 Ω, 0.1 W Chip Resistor
R3	10 kΩ Potentiometer
R4	3000 Ω, 0.1 W Chip Resistor
L1	5 Turns, 0.126″ ID, #20 AWG Enamel
Z1 to Z13	See Photomaster
D1	1N4734 Motorola Zener
Board — G10, 1/32″	
Input/Output Connectors — SMA	

Figure 13. Schematic of Broadband Demonstration Amplifier

PERFORMANCE CHARACTERISTICS OF BROADBAND DEMONSTRATION AMPLIFIER

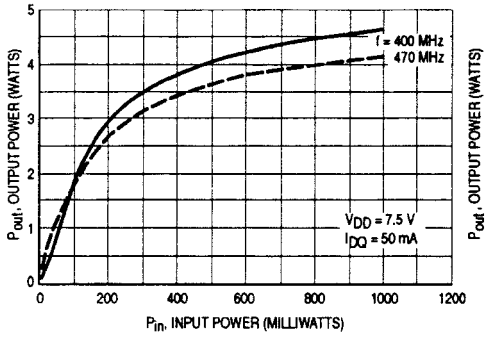


Figure 14. Output Power versus Input Power

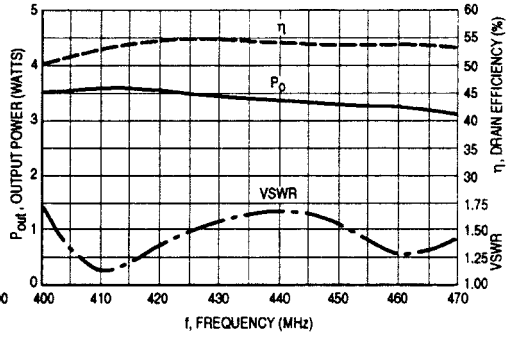


Figure 15. Output Power, Drain Efficiency and VSWR versus Frequency

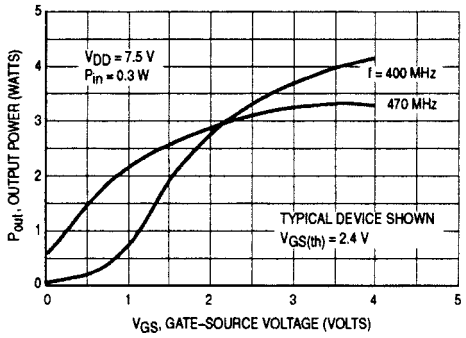


Figure 16. Output Power versus Gate Voltage

DESIGN CONSIDERATIONS

The MRF5003 is a common–source, RF power, N–Channel enhancement mode, Metal–Oxide Semiconductor Field–Effect Transistor (MOSFET). Motorola RF MOSFETs feature a vertical structure with a planar design. Motorola Application Note AN211A, "FETs in Theory and Practice", is suggested reading for those not familiar with the construction and characteristics of FETs.

This surface mount packaged device was designed primarily for VHF and UHF power amplifier applications. Manufacturability is improved by utilizing the tape and reel capability for fully automated pick and placement of parts.

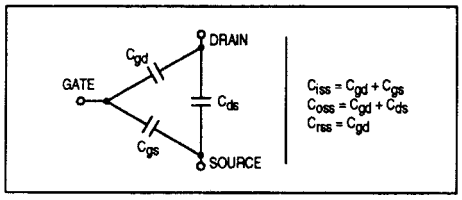
The major advantages of RF power MOSFETs include high gain, simple bias systems, relative immunity from thermal runaway, and the ability to withstand severely mismatched loads without suffering damage.

MOSFET CAPACITANCES

The physical structure of a MOSFET results in capacitors between all three terminals. The metal oxide gate structure determines the capacitors from gate–to–drain (C_{gd}), and gate–to–source (C_{gs}). The PN junction formed during fabrication of the RF MOSFET results in a junction capacitance from drain–to–source (C_{ds}). These capacitances are characterized as input (C_{iss}), output (C_{oss}) and reverse transfer (C_{rss}) capacitances on data sheets. The relationships between the inter–terminal capacitances and those given on data sheets are shown below. The C_{iss} can be specified in two ways:

1. Drain shorted to source and positive voltage at the gate.
2. Positive voltage of the drain in respect to source and zero volts at the gate.

In the latter case, the numbers are lower. However, neither method represents the actual operating conditions in RF applications.



DRAIN CHARACTERISTICS

One critical figure of merit for a FET is its static resistance in the full–on condition. This on–resistance, $R_{DS(on)}$, occurs in the linear region of the output characteristic and is specified at a specific gate–source voltage and drain current. The drain–source voltage under these conditions is termed $V_{DS(on)}$. For MOSFETs, $V_{DS(on)}$ has a positive temperature coefficient at high temperatures because it contributes to the power dissipation within the device.

GATE CHARACTERISTICS

The gate of the RF MOSFET is a polysilicon material, and is electrically isolated from the source by a layer of oxide. The input resistance is very high — on the order of $10^9\ \Omega$ — resulting in a leakage current of a few nanoamperes.

Gate control is achieved by applying a positive voltage to the gate greater than the gate–to–source threshold voltage, $V_{GS(th)}$.

Gate Voltage Rating — Never exceed the gate voltage rating. Exceeding the rated V_{GS} can result in permanent damage to the oxide layer in the gate region.

Gate Termination — The gates of these devices are essentially capacitors. Circuits that leave the gate open–circuited or floating should be avoided. These conditions can result in turn–on of the devices due to voltage build–up on the input capacitor due to leakage currents or pickup.

Gate Protection — These devices do not have an internal monolithic zener diode from gate–to–source. If gate protection is required, an external zener diode is recommended with appropriate RF decoupling.

Using a resistor to keep the gate–to–source impedance low also helps dampen transients and serves another important function. Voltage transients on the drain can be coupled to the gate through the parasitic gate–drain capacitance. If the gate–to–source impedance and the rate of voltage change on the drain are both high, then the signal coupled to the gate may be large enough to exceed the gate–threshold voltage and turn the device on.

DC BIAS

Since the MRF5003 is an enhancement mode FET, drain current flows only when the gate is at a higher potential than the source. See Figure 8 for a typical plot of drain current versus gate voltage. RF power FETs operate optimally with a quiescent drain current (I_{DQ}), whose value is application dependent. The MRF5003 was characterized at I_{DQ} = 50 mA, which is the suggested value of bias current for typical applications. For special applications such as linear amplification, I_{DQ} may have to be selected to optimize the critical parameters.

The gate is a dc open circuit and draws no current. Therefore, the gate bias circuit may generally be just a simple resistive divider network. Some special applications may require a more elaborate bias system.

GAIN CONTROL

Power output of the MRF5003 may be controlled from its rated value down to zero (negative gain) with a low power dc control signal, thus facilitating applications such as manual gain control, ALC/AGC and modulation systems. Figure 16 is an example of output power variation with gate–source bias voltage. This characteristic is very dependent on frequency and load line.

MOUNTING

The specified maximum thermal resistance of 14°C/W assumes a majority of the 0.100″ x 0.200″ source contact on the back side of the package is in good contact with an appropriate heat sink. In the test fixture shown in Figure 1, the device is clamped directly to a copper pedestal. In the demonstration amplifier, the device was mounted on top of the G10 circuit board and heat removal was accomplished through several solder filled plated through holes. As with all RF power devices, the goal of the thermal design should be to minimize the temperature at the back side of the package.

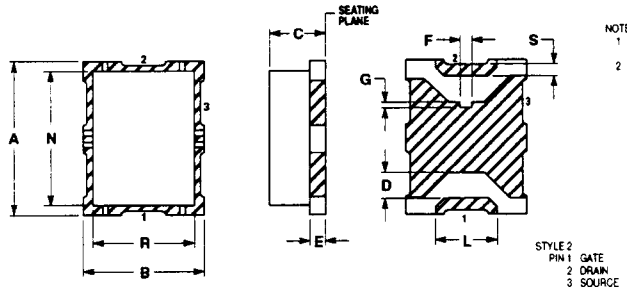
AMPLIFIER DESIGN

Impedance matching networks similar to those used with bipolar transistors are suitable for the MRF5003. For examples see Motorola Application Note AN721, "Impedance Matching Networks Applied to RF Power Transistors." Both small–signal S–parameters and large–signal impedances are provided. While the S–parameters will not produce an exact design solution for high power operation, they do yield a good first approximation. This is an additional advantage of RF power MOSFETs.

Since RF power MOSFETs are triode devices, they are not unilateral. This coupled with the very high gain of the MRF5003 yield a device capable of self oscillation. Stability may be achieved by techniques such as drain loading, input shunt resistive loading, or output to input feedback. Different stabilizing techniques were applied to the test fixture and demonstration amplifiers. The RF test fixture implements a parallel resistor and capacitor in series with the gate while the demonstration amplifier utilizes a 43 Ω shunt resistor from gate to ground. Both circuits have a load line selected for a higher efficiency, lower gain, and more stable operating region.

Two port stability analysis with the MRF5003 S–parameters provides a useful tool for selection of loading or feedback circuitry to assure stable operation. See Motorola Application Note AN215A, "RF Small–Signal Design Using Two–Port Parameters," for a discussion of two port network theory and stability.

PACKAGE DIMENSIONS



NOTES
1 DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2 CONTROLLING DIMENSION: INCH.

	INCHES		MILLIMETERS	
DIM	MIN	MAX	MIN	MAX
A	0.260	0.270	6.60	6.86
B	0.200	0.210	5.08	5.33
C	0.090	0.104	2.29	2.64
D	0.040	0.050	1.02	1.27
E	0.022	0.028	0.56	0.71
F	0.015	0.025	0.38	0.64
G	0.005	0.015	0.13	0.38
L	0.100	0.110	2.54	2.79
N	0.226	0.236	5.74	5.99
R	0.166	0.176	4.22	4.47
S	0.025	0.035	0.64	0.89

CASE 430–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters can and do vary in different applications. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and Ⓜ are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

2-3-7 GaAs MESFETs

Introduction. More than twenty years ago, for RF applications, the JFET was the dominant FET. As mentioned earlier, its major disadvantage is the fact that the f_T cutoff frequency, defined as

$$f_T = g_m / 2\pi C_{GS} \tag{2-262}$$

limits the operating range of this semiconductor device to 1 GHz at most. Also, the JFET's temperature sensitivity and loose tolerances in transconductance and other RF-related values made it difficult to use. Finally, no RFICs made use of this technology. While at the moment silicon-based MOS technology probably is the most cost-effective and is produced in the largest quantities, it does not rival the metal-semiconductor FET designed on GaAs technology. Figure 2-115 shows the structure of the silicon MOSFET, silicon JFET, and the GaAs MESFET.

Large-Signal Behavior of GaAs MESFETs. The mathematics for the large-signal behavior of the GaAsFET is basically quite similar to that for the JFET; however, the computations for JFET channel current, diode currents, and capacitance are much simpler than for the GaAsFET. Temperature effects are embedded in all the equations for all the transistors mentioned so far. In our opinion, the modified Materka model used by Ansoft and a large number of users is the most complete one; however, the following models are supported by most CAD tools:

- Chalmers (Angelov)
- Curtice–Ettenberg cubic
- Curtice quadratic
- IAF (Berroth)
- ITT PFET and TFET
- Modified Materka–Kacprzak
- Physics-based MESFET
- Raytheon (Statz)
- TriQuint (TOM-1, TOM-2, and TOM-3)
- With some restrictions, metal-insulator-semiconductor FETs (MISFETs), modulation-doped FETs (MODFETs), and high-electron-mobility transistors (HEMTs)

The main advantages of GaAsFET technology derive from the fact that it uses a metal-semiconductor junction with a barrier voltage of 0.8 V, its input capacitance is typically less than 0.2 pF, and the reverse feedback capacitance is less than 0.02 pF, or roughly 10% of the input capacitance. As a result of this, $f_{\max}$ is approximately five times higher than f_T. Table 2-16 shows a comparison of silicon BJT and GaAsFET technologies. While $f_{\max}$ for the bipolar transistor is

$$f_{\max} \simeq \sqrt{\frac{f_t}{8\pi r'_{bb} C_C}} \tag{2-263}$$

the $f_{\max}$ determination for the GaAsFET is given by

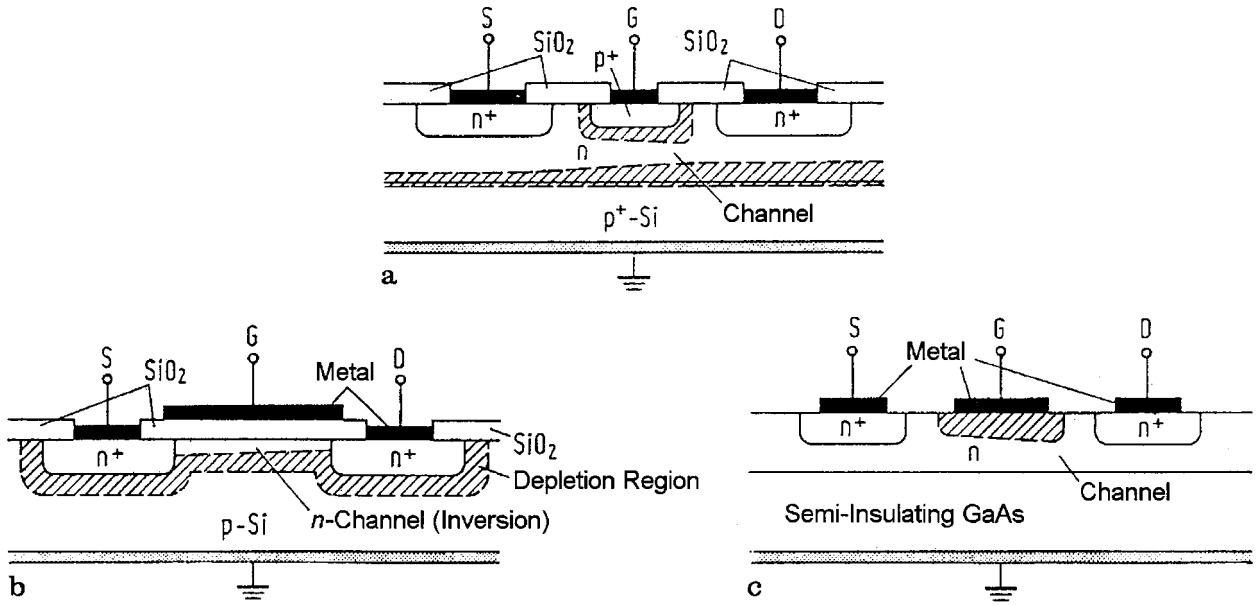


Figure 2-115 This picture is provided to differentiate between the structure of the three technologies. (a) The JFET and (b) the MOSFET are based on silicon; (c) the GaAsFET is based on semi-insulating GaAs. While the MOS transistor has silicon dioxide (SiO_2) as gate insulation, both the JFET and GaAs MESFET have a diode from gate to source, with a barrier voltage of 0.65 V for silicon and 0.8 V for GaAs.

Table 2-16 Comparison of Si BJT, SiGe HBT, and GaAsFET technologies

	Si Bipolar	SiGe HBT	GaAs FET
f_T	25 GHz	100 GHz	22 GHz
$f_{\max}$	40 GHz	200 GHz	110 GHz
Features	Low cost, low 1/*f* noise (5 kHz)	Low 1/*f* noise, very low distortion, high cost, IC only	Highest flexibility, lowest NF_0, well established

$$f_{\max} = \frac{f_T}{2}\sqrt{\frac{R_0}{R_i + R_s + R_g}} \tag{2-264}$$

As a sample calculation,

$$f_{\max}\ \frac{21.9\text{ GHz}}{2}\sqrt{\frac{450}{1 + 1.5 + 2}} = 110\text{ GHz} \tag{2-265}$$

The three major drawbacks of the GaAsFET are:

1. Much higher flicker corner frequency (somewhere between 10 and 100 MHz), probably due to a lack of surface passivation.
2. Much higher output conductance. This tends to load down any circuit connected to the drain. On the other hand, since the transconductance is quite high for even low currents, these devices have very high gains at low frequencies, which can make them quite unstable. In the saturated mode it is not uncommon to find a drain–source resistance of 100–500 Ω, while BJTs and JFETs offer values of several kilohms and higher.
3. Because of the very high flicker corner frequency (from 10 to 100 MHz), MESFETs are really not useful for low-noise mixers and oscillators, and unless there are no devices available in the frequency range above 30 GHz, they should be avoided for these applications.

As to the MESFET's construction and dc properties, Figure 2-116 shows a MESFET's cross section and dc *I*–*V* characteristics.

It was outlined in the beginning that while the GaAsFET is a close relative to the JFET and MOSFET, its actual behavior was found to be described best by a set of analytic equations. The first such model was the one by Curtice in the form of quadratic and cubic models, but it does not have enough derivatives to give enough insight into subtle things such as third- and higher-order intermodulation distortion and accurate harmonic generation. Other researchers have addressed various areas, but we still find that the Materka model has the best approximation, especially using the two-transistor approach outlined later [20, 21]. The large-signal topology for all FETs consists of an intrinsic model with some extrinsic parameters, further complicated by the package, as shown by Figures 2-117 and 2-118. Table 2-17 lists their keywords. The actual intrinsic model and its parameter definition depend on the particular model and since designs using GaAsFET will always be done using CAD tools, we will not go into any details of the equations but will list them. They are not dissimilar from the JFET and MOSFET equations.

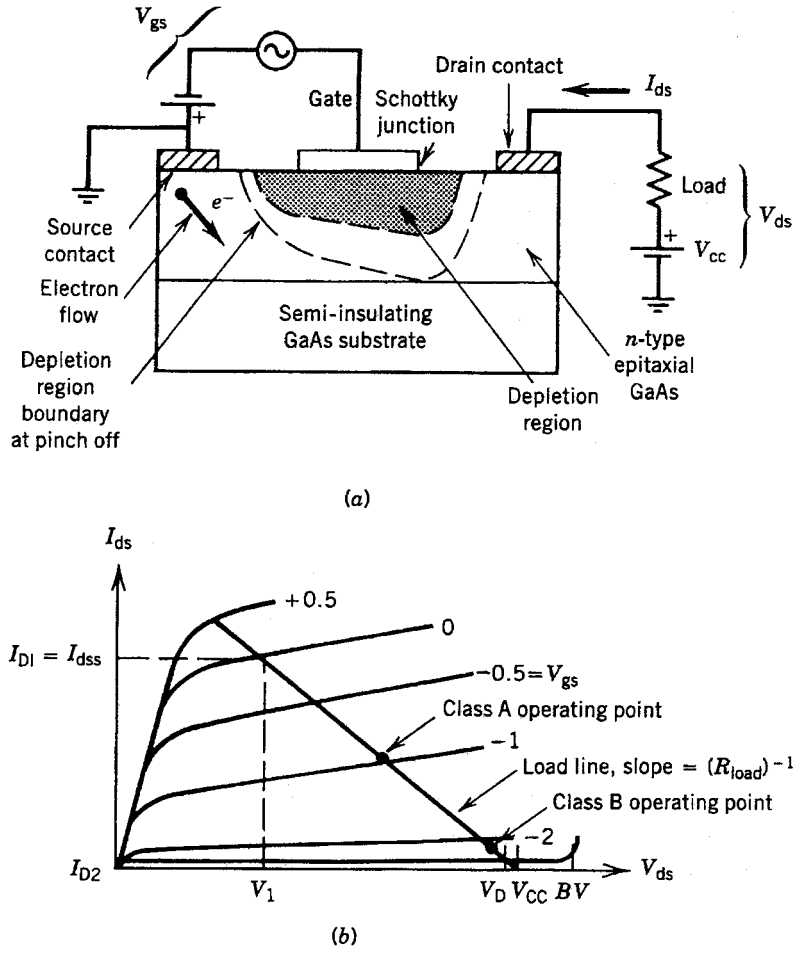


Figure 2-116 (a) Cross section and bias circuit and (b) dc *I–V* curve, including ac load line, for a MESFET.

The Modified Materka-Kacprzak Model Figure 2-119 shows the intrinsic model of the Materka FET. Table 2-18 lists its keywords.

Large-Signal Equations

Device Equations

V_{gsi} = Intrinsic gate–source voltage
V_{dsi} = Intrinsic drain–source voltage
V_1 = Voltage across C_{GS} and R_i
V_{gdi} = Intrinsic gate–drain voltage
$V_T = kT_J/q$ (thermal voltage)
k = Boltzmann's constant
q = Electron charge
T_J = Analysis temperature (kelvins)

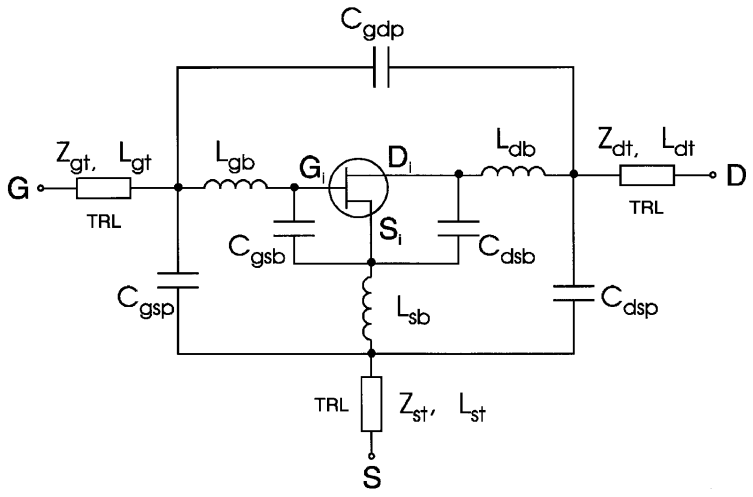


Figure 2-117 MESFET extrinsic model.

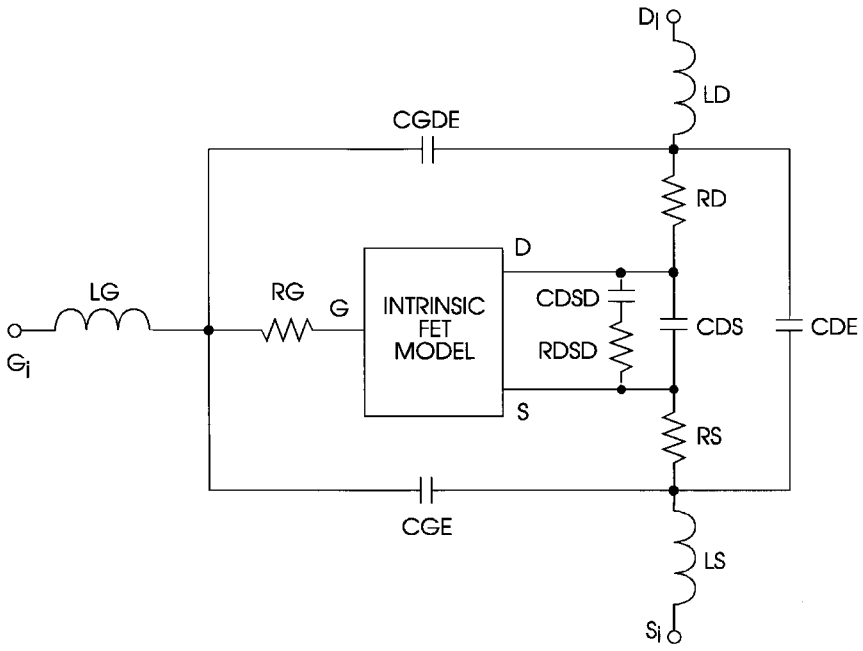


Figure 2-118 MESFET package model.

Table 2-17 Large-signal MESFET model keywords

Keyword	Description	Unit	Default
	Extrinsic Parameters		
RG	Gate bulk and ohmic resistance	ohm	0.0
RD	Drain bulk and ohmic resistance	ohm	0.0
RS	Source bulk and ohmic resistance	ohm	0.0
LG	Gate-lead inductance (metallization)	henry	0.0
LD	Drain-lead inductance (metallization)	henry	0.0
LS	Source-lead inductance (via)	henry	0.0
CDS	Drain–source capacitance	farad	0.0
CDSD	Low-frequency trapping capacitor	farad	0.0
RDSD	Channel trapping resistance	ohm	∞
CGE	Gate–source electrode capacitance	farad	0.0
CDE	Drain–source electrode capacitance	farad	0.0
CGDE	Gate–drain electrode capacitance	farad	0.0
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate-to-source package capacitance	farad	0.0
CDSP	Drain-to-source package capacitance	farad	0.0
CGDP	Gate-to-drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for $\varepsilon_r = 1$	meter	0.0

Channel Current

$$I_{ds} = \mathrm{IDSS}\left(1 + \mathrm{SS}\,\frac{V_{dsi}}{\mathrm{IDSS}}\right)\left(1 - \frac{V_{gsi}(t-T)}{\mathrm{VP0} + \mathrm{GAMA}V_{dsi}}\right)^{(\mathrm{E}+\mathrm{KE}V_{gsi}(t-T))}$$

$$\times \tanh\left(\frac{\mathrm{SL}V_{dsi}}{\mathrm{IDSS}(1 - \mathrm{KG}V_{gsi}(t-T))}\right)$$

Diode

$$I_{gd} = I_{gdc} - \begin{cases} \mathrm{IB0}\,\exp[-\mathrm{AFAB}(V_{gdi} + \mathrm{VBC})] \\ \dfrac{\mathrm{GMAX}}{4}\,\{\tanh[\mathrm{K1D}(V_{gsi} - \mathrm{K2D})] - 1\}\,[V_{gdi} + \mathrm{VBC} - \sqrt{(V_{gdi} + \mathrm{VBC})^2 + \mathrm{K3D}}] \end{cases}$$

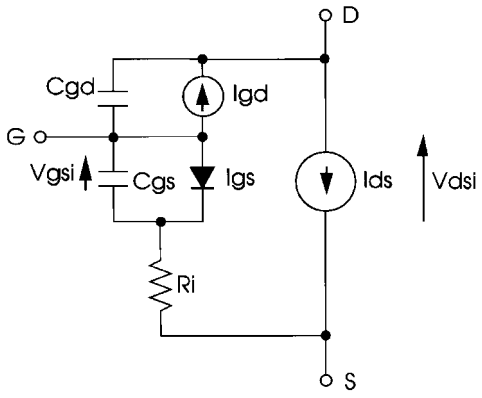


Figure 2-119 Intrinsic model of the modified Materka–Kacprzak MESFET.

where

$$I_{gdc} = \text{IG0}[\exp(\text{AFAB}V_{gdi}) - 1]$$

Channel Resistance

$$R_i = \begin{cases} \text{R10}(1 - \text{KR}V_{gsi}), & \text{KR}V_{gsi} < 1.0 \\ 0, & \text{KR}V_{gsi} \geq 1.0 \end{cases}$$

Capacitance Model

$$C_{gs} = \text{CGS0}\frac{F1F2}{\sqrt{1 - \dfrac{V_{\text{new}}}{\text{VBI}}}} + \text{CGD0}\ F3$$

$$C_{gd} = \text{CGS0}\frac{F1F3}{\sqrt{1 - \dfrac{V_{\text{new}}}{\text{VBI}}}} + \text{CGD0}\ F2$$

where

$$F1 = \frac{1}{2}\left(1 + \frac{V_{\text{eff}} - V_T}{\sqrt{(V_{\text{eff}} - V_T)^2 + \delta^2}}\right)$$

$$F2 = \frac{1}{2}\left(1 + \frac{V_{gsi} - V_{gdi}}{\sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}}\right)$$

Table 2-18 Materka–Kacprzak MESFET model keywords

Keyword	Description	Unit	Default
	Area, Noise, and Name		
AREA	Area multiplier		1.0
KFN	Flicker noise coefficient (Materka model only)[a]		0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
	Channel Current Model		
IDSS	Drain saturation current for $V_{GS} = 0$	ampere	0.1
VP0	Pinchoff voltage for $V_{DS} = 0$	volt	–2.0
GAMA	Voltage slope parameter of pinchoff voltage	/volt	0.0
E	Constant part of power-law parameter		2.0
KE	Dependence of power law on V_{GS}	/volt	0.0
SL	Slope of the $V_{GS} = 0$ drain characteristic in the linear region	ampere/volt	0.15
KG	Drain dependence on V_{GS} in the linear region	/volt	0.0
SS	Slope of the drain characteristic in the saturated region	ampere/volt	0.0
T	Channel transit-time delay	second	0.0
IG0	Diode saturation current	ampere	0
AFAG	Slope factor of forward diode current	/volt	38.696
IB0	Breakdown saturation current	ampere	0
AFAB	Slope factor of breakdown current	/volt	0
VBC	Breakdown voltage	volt	∞
GMAX	Breakdown conductance	ampere/volt	0
K1D	Fitting parameter	/volt	0
K2D	Fitting parameter	volt	0
K3D	Fitting parameter	$volt^2$	0
R10	Intrinsic channel resistance for $V_{GS} = 0$	ohm	0.0
KR	Slope factor of intrinsic channel resistance	/volt	0.0
	Materka Capacitance Model		
C10	Gate–source Schottky barrier capacitance for $V_{GS} = 0$	farad	0.0
K1	Slope parameter of gate–source capacitance	/volt	1.25
MGS	Gate–source grading coefficient		0.5
C1S	Constant parasitic component of gate–source capacitance	farad	0.0
CF0	Gate–drain feedback capacitance for $V_{GD} = 0$	farad	0.0
KF	Slope parameter of gate–drain feedback capacitance	/volt	1.25
MGD	Gate–drain grading coefficient		0.5
FCC	Forward-bias depletion capacitance coefficient		0.8

[a]The flicker noise parameter of the Materka model is KFN so as not to conflict with the KF parameter in the capacitance model.

$$F3 = \frac{1}{2}\left(1 - \frac{V_{gsi} - V_{gdi}}{\sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}}\right)$$

$$V_{\text{new}} = \begin{cases} A1, & A1 < \text{VMAX} \\ \text{VMAX}, & A1 \geq \text{VMAX} \end{cases}$$

$$A1 = \frac{1}{2}\left(V_{\text{eff}} + V_T + \sqrt{(V_{\text{eff}} - V_T)^2 + \delta^2}\right)$$

$$V_{\text{eff}} = \frac{1}{2}\left(V_{gsi} + V_{gdi} + \sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}\right)$$

$$V_T = \text{VP0} + \text{GAMA}\, V_{dsi}$$

$$\delta = 0.2$$

Some of the modifications to the Materka model have been done by Ansoft under various Department of Defense contracts, and by Raytheon and Texas Instruments under similar contracts, with Ansoft being a subcontractor.

The most relevant equation is really the channel current. Its derivatives are largely responsible for the accuracy of the intermodulation distortion, power-added efficiency, and, of course, its dc *I–V* curves.

Enhancement/Depletion FETs To make the designer's life more difficult, it turns out that there are two types of GaAsFETs:

1. *Depletion FETs (DFETs).* Most similar to the JFET; here V_G must be negative to control the device. They are the most commonly produced and are the FET type most referred to in this book.

On one hand:

- They require a negative gate voltage with respect to the source.
- Self-bias allows operation from a single supply voltage.

On the other hand:

- For low-voltage operation, a negative voltage generator may be required.
- Supply voltage must be doubled to accommodate full-swing operation.

2. *Enhancement FETs (EFETs).* Most similar to the MOSFET; here V_G must be positive to bring life to the device. Practically speaking, EFETs are used mostly in integrated circuits; they're typically not available in discrete, packaged form.

On one hand:

- They need only positive supply for biasing.
- They provide higher g_m/mA (for same device width)—5.1 mS versus 3.9 mS at 8 mA.

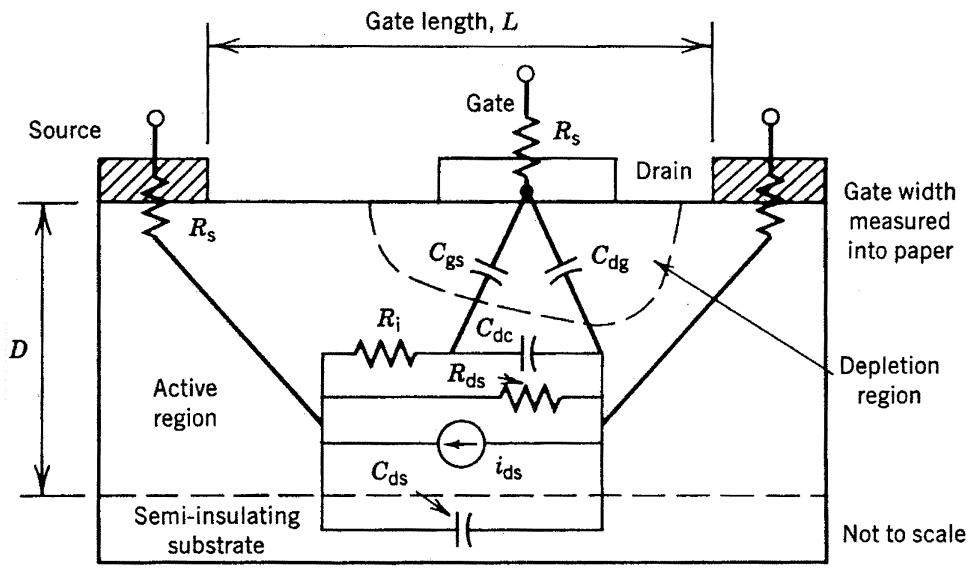


Figure 2-120 Location of lumped-element components for a MESFET.

- They are good for low-power LNAs, giving slightly better NF that DFETs, an NF of 1 dB at 1 GHz, and $I_{dd} < 10$ mA.

On the other hand:

- They have a very limited gate bias range (V_{GS} between 0.15 and 0.7 V).
- The gate conduction degrades NF and input impedance.
- The gate capacitance is higher than that of DFETs.
- The linearity is not as good as that of DFETs.

With today's technologies, all GaAs devices are *n*-channel; we have not seen any *p*-channels yet. More information about biasing will be given in the next chapter.

Figure 2-120 is a lumped-element, two-part equivalent circuit of a MESFET showing the location of lumped-element components.

2-3-8 Small-Signal GaAs MESFET Model

Figure 2-121 shows the applicable linear equivalent circuit for a MESFET and Table 2-19 lists its keywords. As with the MOS transistors, there is a gallium-arsenide dual-gate MOSFET available that is used mostly in special circuits, such as preamplifiers and mixers, whose IF has to be significantly higher than the flicker corner frequency, for example, higher than 200 MHz.

The following two datasheets* show typical GaAsFETs, both low-noise and medium-power. Such devices are, of course, available from a large variety of manufacturers—to name at least two, Siemens and the Nippon Electric Company, which is represented in the United States by California Eastern Laboratories.

*Reproduced with permission.

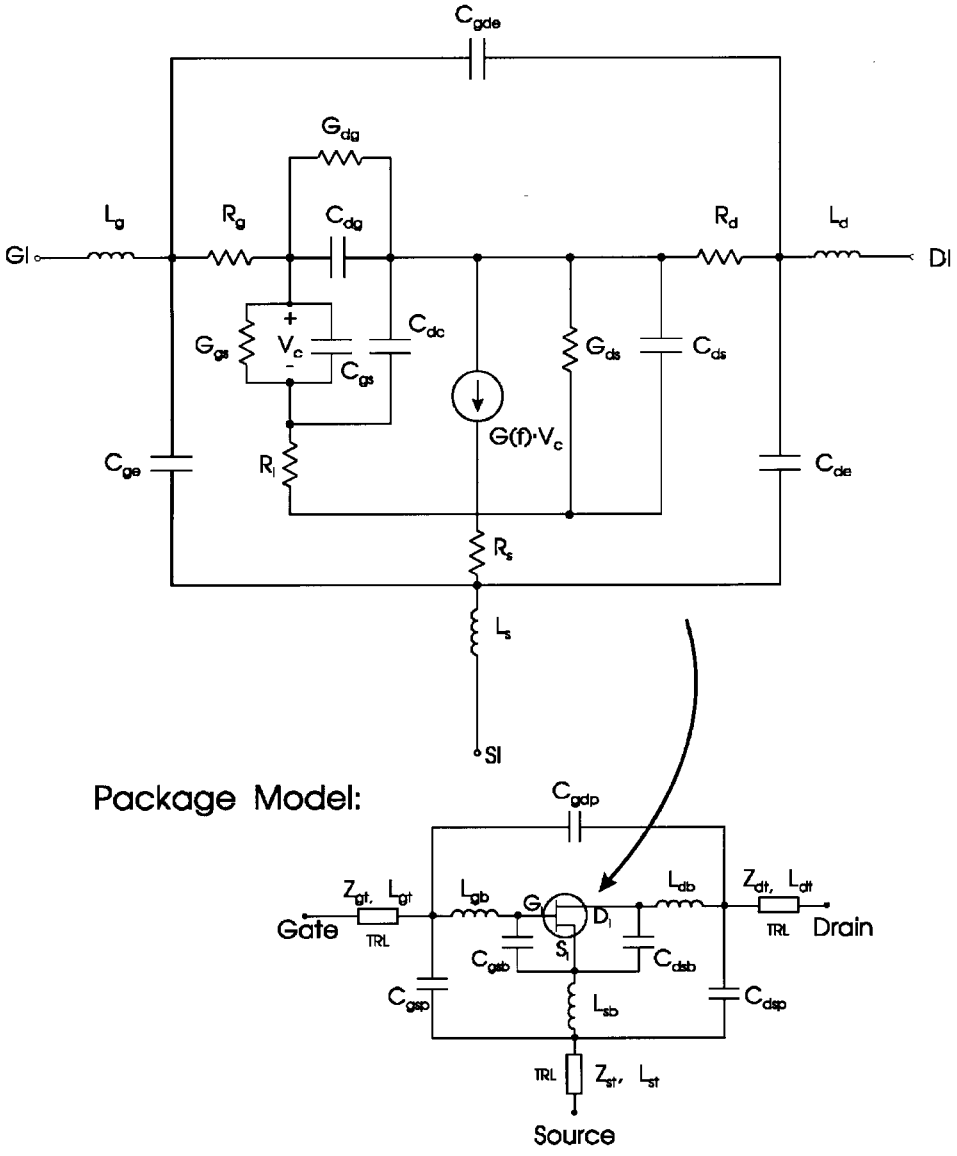


Figure 2-121 Small-signal model of a MESFET.

Table 2-19 Small-signal MESFET model keywords

Keyword	Description	Unit	Default
G	Transconductance at dc, G_0[a]	/ohm	
CGS	Gate–source capacitance	farad	
F	3-dB Roll-off frequency	hertz	∞
T	Time delay	second	0.0
TDS	Drain–source time delay	second	0.0
GGS	Gate–source conductance	/ohm	0.0
CDG	Drain–gate capacitance	farad	0.0
CDC	Dipole-layer capacitance	farad	0.0
CDS	Drain–source capacitance	farad	0.0
GDS	Drain–source conductance	/ohm	0.0
RI	Channel resistance	ohm	0.0
RG	Gate resistance	ohm	0.0
RD	Drain resistance	ohm	0.0
RS	Source resistance	ohm	0.0
CGE	External gate capacitance	farad	0.0
CDE	External drain capacitance	farad	0.0
LG	Gate-lead inductance	henry	0.0
LD	Drain-lead inductance	henry	0.0
LS	Source-lead inductance	henry	0.0
CGDE	External gate–drain capacitance	farad	0.0
GDG	Gate–drain conductance	/ohm	0.0
TJ	Chip temperature	kelvin	298
	Package Parasitics		
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate-to-source package capacitance	farad	0.0
CDSP	Drain-to-source package capacitance	farad	0.0
CGDP	Gate-to-drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for at $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for at $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for at $\varepsilon_r = 1$	meter	0.0
FC	Corner frequency of flicker (1/*f*) noise[b]	hertz	10 MHz
FCP	Shape factor of the 1/*f* noise response	1.0	
label	User-defined term that refers to temperature coefficient		

[a]The transconductance of this model may approximately be described by $g_m = G\dfrac{e^{-j\omega T}}{1 = j(f/F)}$ where $\omega = 2\pi f$ and f is frequency.

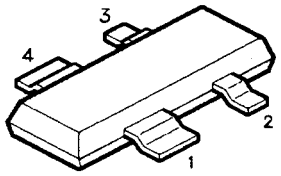
[b]The flicker noise frequency dependence is given by $1/(f/F_c)^{\mathrm{FCP}}$

SIEMENS AlGaAs / InGaAs HEMT CFY 77

Datasheet

Features

* Very low noise
* Very high gain
* For low noise front end amplifiers up to 20 GHz
* For DBS down converters



ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering code (taped)	Package 1)
CFY77-08	HG	Q62702-F1549	MW-4
CFY77-10	HH	Q62702-F1559	MW-4

Maximum ratings	Symbol		Unit
Drain-source voltage	V_{DS}	3.5	V
Drain-gate voltage	V_{DG}	4.5	V
Gate-source voltage	V_{GS}	-3.0	V
Drain current	I_D	60	mA
Channel temperature	T_{Ch}	150	°C
Storage temperature range	T_{stg}	-65...+150	°C
Total power dissipation ($T_s \leq 51$°C) 2)	P_{tot}	180	mW
Thermal resistance			
Channel-soldering point source	R_{thChS}	550	K/W

1) Dimensions see chapter Package Outlines
2) T_S: Temperature measured at soldering point

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Electrical characteristics at T_A = 25°C

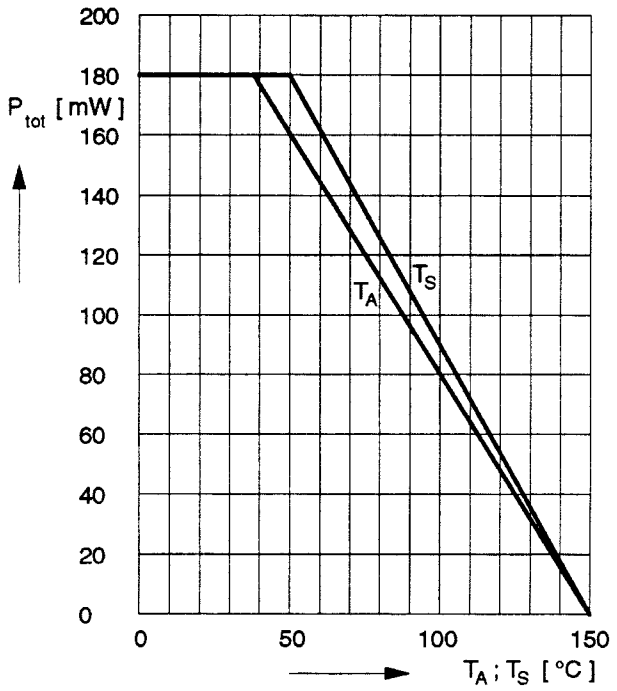
unless otherwise specified

Characteristics	Symbol	min	typ	max	Unit
Drain-source saturation current V_{DS} = 2 V V_{GS} = 0 V	I_{DSS}	15	30	60	mA
Pinch-off voltage V_{DS} = 2 V I_D = 1 mA	$V_{GS(P)}$	-2	-0.7	-0.2	V
Gate leakage current V_{DS} = 2 V I_D = 15 mA	I_G	-	0.05	2	µA
Transconductance V_{DS} = 2 V I_D = 15 mA	g_m	50	65	-	mS
Noise figure V_{DS} = 2 V I_D = 15 mA f = 12 GHz	F				dB
CFY77-08 CFY77-10		- -	0.7 0.9	0.8 1	
Associated gain V_{DS} = 2 V I_D = 15 mA f = 12 GHz	G_a				dB
CFY77-08 CFY77-10		10 9.5	10.5 10	- -	

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Total Power Dissipation $P_{tot} = f(T_s;T_A)$

Package mounted on alumina



Typical Common Source Noise Parameters

I_D = 15 mA U_{DS} = 2.0 V Z_0 = 50 Ω

f	F_{min}	G_a	Γ_{opt}		R_n	r_n	N	$F_{50\Omega}$
GHz	dB	dB	MAG	ANG	Ω			dB
2	0.36	19.4	0.79	27	13.7	0.274	0.03	1.2
4	0.44	15.9	0.72	60	10.1	0.202	0.04	1.1
6	0.51	13.9	0.63	92	5.85	0.117	0.05	1.05
8	0.58	12.4	0.56	134	2.35	0.047	0.06	1.0
10	0.65	11.2	0.52	180	1.1	0.022	0.07	1.0
12	0.72	10.4	0.54	-135	2.9	0.058	0.08	1.1
14	0.80	9.7	0.59	-108	7.15	0.143	0.10	1.5

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Typical Common Source S-Parameters

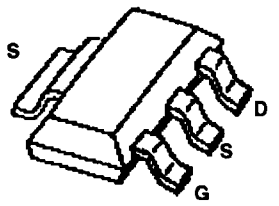
I_D = 15 mA U_D = 2.0 V Z_o = 50 Ω

	S11		S21		S12		S22	
GHz	Mag	Ang	Mag	Ang	Mag	Ang	Mag	Ang
1	0.98	-22.8	5.55	159.6	0.030	87.5	0.633	-16.3
2	0.94	-46.1	5.40	139.3	0.053	57.8	0.60	-32.5
3	0.88	-68.4	5.09	120.1	0.074	44.9	0.54	-48.0
4	0.82	-90.6	4.77	101.2	0.089	30.7	0.48	-63.3
5	0.77	-110.8	4.45	84.0	0.101	18.1	0.42	-77.5
6	0.72	-131.4	4.16	67.3	0.112	7.9	0.35	-92.6
7	0.66	-153.6	3.88	50.2	0.119	-3.3	0.28	-110.8
8	0.63	-175.2	3.58	34.5	0.122	-12.7	0.22	-132.0
9	0.62	164.4	3.29	18.9	0.120	-22.0	0.16	-157.3
10	0.62	145.0	3.01	4.0	0.119	-29.5	0.14	177.3
11	0.64	128.3	2.76	-10.3	0.119	-37.4	0.15	136.2
12	0.64	113.1	2.51	-23.5	0.114	-44.0	0.18	115.4
13	0.66	101.3	2.32	-35.7	0.114	-47.3	0.23	100.9
14	0.67	89.4	2.18	-48.2	0.116	-53.1	0.25	91.0
15	0.69	73.6	2.06	-62.4	0.116	-58.6	0.28	75.4
16	0.73	59.2	1.85	-75.9	0.115	-65.8	0.36	57.1
17	0.76	51.7	1.65	-86.5	0.112	-69.4	0.39	53.1
18	0.78	45.4	1.56	-96.7	0.115	-72.3	0.42	43.8
19	0.77	36.2	1.51	-108.6	0.121	-76.7	0.44	38.8

SIEMENS GaAs FET **CLY 15**

Datasheet

* Power amplifier for mobile phones
* For frequencies from 400 MHz to 2.5 GHz
* Operating voltage range: 2.7 to 6 V
* P_{out} at V_D=3V, f=1.8 GHz typ. 31.5 dBm
* Efficiency better 50%



ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering code (taped)	Package 1)
CLY 15	CLY 15	Q62702-L99	SOT 223

Maximum ratings	Symbol		Unit
Drain-source voltage	V_{DS}	9	V
Drain-gate voltage	V_{DG}	12	V
Gate-source voltage	V_{GS}	-6	V
Drain current	I_D	5	A
Channel temperature	T_{Ch}	150	°C
Storage temperature	T_{stg}	-55...+150	°C
Total power dissipation (Ts ≤ 80°C) Ts: Temperature at soldering point	P_{tot}	4.7	W
Thermal resistance Channel-soldering point (GND)	R_{thChS}	< 15	K/W

1) Dimensions see chapter Package Outlines

SIEMENS GaAs FET **CLY 15**

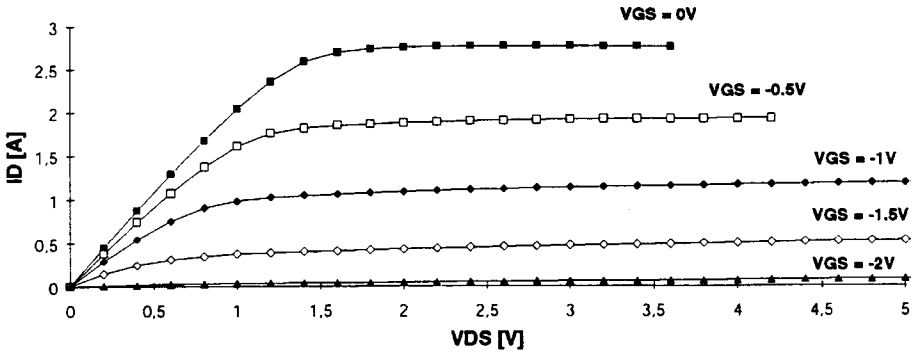
Electrical characteristics (T_A = 25°C , unless otherwise specified)

Characteristics	Symbol	min	typ	max	Unit
Drain-source saturation current *) V_{DS} = 3V V_{GS} = 0V	I_{DSS}	2.4	3.2	4.8	A
Cut-off current V_{DS} = 3V V_{GS} = -3.8V	I_D	-	-	400	µA
Gate cut-off current V_{DS} = 3V V_{GS} = -3.8V	I_G	-	20	70	µA
Pinch-off Voltage VDS=3V ID=400µA	$V_{GS(p)}$	-3.8	-2.8	-1.8	V
Small Signal Gain *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 5dBm	G	-	6	-	dB
Output Power *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 29dBm	P_o	32	32.5	-	dBm
Output Power *) V_{DS} = 5V I_D = 1.4A f = 1.8GHz P_{in} = 30 dBm	P_o	34.5	35	-	dBm
1dB-Compression Point *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz	P_{1dB}	-	31.5	-	dBm
1dB-Compression Point *) V_{DS} = 5V I_D = 1.4A f = 1.8GHz	P_{1dB}	-	34.5	-	dBm
Power Added Efficiency *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 29dBm	η_D	45	50	-	%

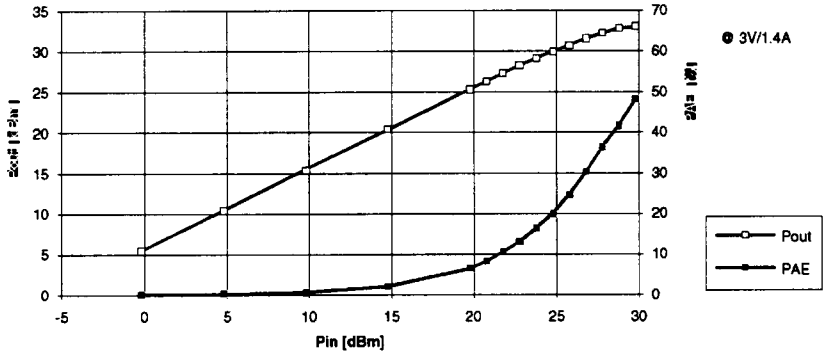
*) pulsed measurement; duty cycle 1:10; t_{on} = 1ms, power matching conditions.

SIEMENS GaAs FET **CLY 15**

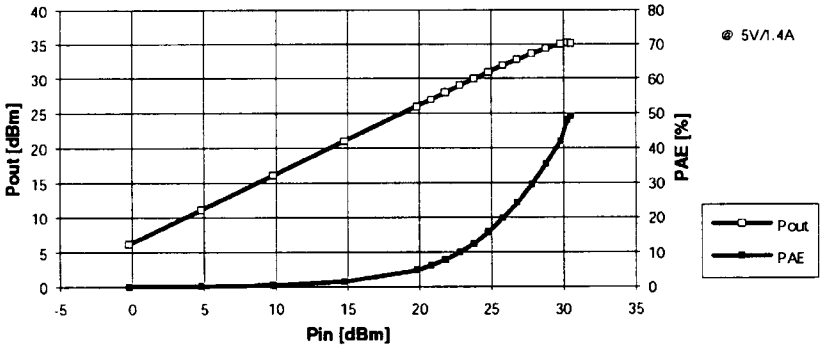
Output Characteristics



Power Characteristics



Power Characteristics



SIEMENS GaAs FET **CLY 15**

typ. Common Source S-Parameter

V_{DS} = 3V I_D=1.4A Z_O=50Ω

f	S11		S21		S12		S22	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
0.200	0.91	-150.9	5.69	99.7	0.01	48.5	0.90	176.0
0.250	0.91	-160.6	4.63	93.5	0.02	45.9	0.87	173.8
0.300	0.90	-167.9	3.89	88.7	0.02	46.8	0.88	171.8
0.350	0.90	-173.7	3.34	84.6	0.02	47.4	0.87	170.8
0.400	0.90	-178.7	2.92	80.9	0.02	47.5	0.87	168.8
0.450	0.90	176.9	2.60	77.5	0.02	47.6	0.87	167.3
0.500	0.90	173.0	2.34	74.6	0.02	48.1	0.87	165.8
0.550	0.90	169.5	2.12	71.4	0.02	47.7	0.87	164.2
0.600	0.90	166.1	1.95	68.7	0.03	47.0	0.87	162.8
0.650	0.90	163.1	1.79	66.1	0.03	47.1	0.87	161.2
0.700	0.90	160.0	1.66	63.5	0.03	46.6	0.87	159.7
0.750	0.90	157.2	1.54	60.9	0.03	45.6	0.87	158.3
0.800	0.90	154.6	1.45	58.6	0.03	45.0	0.87	156.9
0.850	0.90	152.0	1.36	56.1	0.03	43.9	0.87	155.6
0.900	0.90	149.3	1.28	53.8	0.04	43.0	0.87	154.0
0.950	0.90	146.9	1.21	51.5	0.04	41.9	0.87	152.6
1.000	0.90	144.5	1.15	49.0	0.04	41.0	0.87	151.3
1.200	0.91	135.2	0.95	40.3	0.05	36.1	0.87	145.8
1.400	0.91	126.7	0.81	31.8	0.05	31.9	0.88	140.1
1.600	0.92	118.5	0.70	23.8	0.06	26.1	0.88	134.7
1.800	0.92	110.6	0.61	16.3	0.06	20.8	0.88	129.7
2.000	0.93	103.2	0.55	8.7	0.06	15.6	0.89	124.3
2.200	0.93	96.3	0.49	2.1	0.07	10.4	0.88	119.1
2.400	0.93	89.3	0.44	-4.1	0.07	5.2	0.90	114.4
2.600	0.94	82.8	0.40	-10.0	0.07	0.2	0.90	109.3
2.800	0.94	77.0	0.37	-14.9	0.07	-4.2	0.90	104.5
3.000	0.94	71.3	0.34	-19.6	0.08	-9.7	0.91	99.8
3.200	0.93	66.0	0.32	-23.4	0.08	-15.0	0.92	95.1
3.400	0.92	61.6	0.31	-26.8	0.08	-19.4	0.93	90.8
3.600	0.91	57.3	0.30	-29.7	0.07	-23.7	0.92	87.0
3.800	0.90	53.1	0.31	-33.1	0.07	-28.1	0.93	83.1
4.000	0.89	49.2	0.32	-38.1	0.07	-31.9	0.93	79.8
4.200	0.86	46.4	0.34	-44.9	0.07	-35.4	0.92	76.4
4.400	0.83	44.7	0.36	-55.4	0.07	-37.5	0.92	73.4
4.600	0.89	44.2	0.07	-36.2	0.07	-38.1	0.92	71.0
4.800	0.83	43.7	0.34	-80.6	0.07	-39.4	0.92	68.2
5.000	0.85	42.2	0.30	-92.1	0.07	-40.3	0.92	65.2
5.200	0.88	39.4	0.27	-100.8	0.07	-42.5	0.92	62.2
5.400	0.89	36.5	0.24	-107.8	0.07	-45.0	0.92	58.7
5.600	0.90	33.1	0.22	-113.6	0.07	-48.6	0.92	55.7
5.800	0.91	29.6	0.19	-118.9	0.07	-51.4	0.92	52.1
6.000	0.92	26.4	0.18	-124.4	0.07	-54.4	0.92	48.0

SIEMENS GaAs FET **CLY 15**

typ. Common Source S-Parameter
V_{DS} = 5V I_D=1.4A Z_O=50Ω

f	S11		S21		S12		S22	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
0.200	0.90	-151.1	7.61	98.8	0.01	46.5	0.84	176.7
0.250	0.89	-160.6	6.18	92.4	0.02	43.4	0.82	174.7
0.300	0.89	-167.8	5.19	87.5	0.02	46.5	0.82	172.9
0.350	0.89	-173.7	4.45	83.3	0.02	46.0	0.82	171.8
0.400	0.88	-178.7	3.90	79.4	0.02	46.5	0.82	169.7
0.450	0.89	177.0	3.47	75.9	0.02	47.3	0.82	168.3
0.500	0.88	173.2	3.11	72.8	0.02	47.9	0.82	166.7
0.550	0.88	169.6	2.82	69.5	0.02	47.8	0.82	165.5
0.600	0.89	166.4	2.59	66.6	0.03	47.4	0.82	163.9
0.650	0.88	163.1	2.38	63.9	0.03	47.4	0.82	162.6
0.700	0.89	160.3	2.20	61.1	0.03	46.5	0.82	161.0
0.750	0.89	157.5	2.05	58.4	0.03	45.6	0.82	159.6
0.800	0.89	154.9	1.91	55.9	0.03	45.3	0.82	158.0
0.850	0.89	152.1	1.79	53.2	0.03	44.8	0.82	156.8
0.900	0.89	149.7	1.69	50.7	0.03	43.9	0.82	155.4
0.950	0.89	147.1	1.59	48.4	0.04	42.7	0.82	154.1
1.000	0.89	144.7	1.51	45.7	0.04	42.0	0.82	152.8
1.200	0.89	135.5	1.24	36.2	0.04	37.8	0.83	147.3
1.400	0.90	127.1	1.04	27.1	0.05	32.2	0.83	141.9
1.600	0.91	119.1	0.90	18.3	0.05	27.4	0.84	136.8
1.800	0.92	111.1	0.78	10.1	0.06	22.5	0.84	131.6
2.000	0.92	103.7	0.68	2.1	0.06	18.2	0.85	126.3
2.200	0.93	96.6	0.61	-5.1	0.06	12.5	0.86	121.3
2.400	0.93	89.8	0.54	-12.0	0.07	7.3	0.86	116.1
2.600	0.93	83.2	0.48	-18.6	0.07	2.6	0.87	111.4
2.800	0.93	77.3	0.43	-24.0	0.07	-2.6	0.88	106.3
3.000	0.93	71.8	0.39	-29.3	0.07	-7.2	0.89	101.8
3.200	0.92	66.6	0.37	-33.5	0.07	-12.1	0.90	97.2
3.400	0.92	61.8	0.34	-37.2	0.07	-16.8	0.91	92.5
3.600	0.91	57.9	0.32	-40.5	0.07	-21.1	0.91	88.8
3.800	0.90	54.1	0.32	-43.9	0.07	-24.9	0.92	85.1
4.000	0.88	50.5	0.31	-48.3	0.07	-27.6	0.92	81.4
4.200	0.87	47.8	0.32	-53.8	0.07	-31.5	0.92	78.1
4.400	0.86	45.7	0.32	-60.9	0.07	-33.4	0.92	74.9
4.600	0.85	43.4	0.32	-68.9	0.07	-35.4	0.92	72.2
4.800	0.85	42.3	0.31	-77.5	0.07	-37.2	0.92	69.3
5.000	0.86	40.3	0.30	-86.7	0.07	-39.3	0.92	66.2
5.200	0.87	37.7	0.28	-94.5	0.07	-41.6	0.92	63.1
5.400	0.88	35.2	0.26	-101.8	0.07	-44.0	0.92	59.6
5.600	0.89	32.1	0.24	-108.4	0.08	-48.5	0.92	56.6
5.800	0.90	29.0	0.22	-114.5	0.08	-50.3	0.92	53.0
6.000	0.90	26.0	0.21	-121.1	0.08	-54.0	0.93	49.0

2-4 PARAMETER EXTRACTION OF ACTIVE DEVICES

2-4-1 Introduction

We have already seen that all linear models that are used on the market are really derived from large-signal models, and it is our point that the quality of the model depends mostly on the quality of the parameters used to describe the model.

As far as modeling is concerned there are two options. The first option is a physics-based model, which can be applied to items such as diodes and bipolar transistors, including its most advanced versions, such as HBTs, SiGe versions, and other future derivatives. The bipolar transistor is essentially a combination of two diodes whereby the base–emitter junction generates the electrons or holes (depending on how the reader is accustomed to viewing the process) and then transfers those to the collector. The emitter, as its name states, emits the charged particles and the collector collects them, minus some current losses, which are expressed in the current gain of the device. A current gain (β) of 100 means that the collector receives 1% less than the emitter emits.

The second modeling option involves the construction of an analytical mathematic equivalent for the device modeled. Earlier in this chapter, we went through a detailed physics-based derivation about the inner workings of these devices, explaining how, for a given dc bias point, one can simplify them to a small-signal equivalent circuit. This is possible since we assume that the RF current and voltages are less than 1% of the dc equivalent currents and voltages. Reducing a device's complex nonlinear equivalent circuit to a quasistatic linear equivalent circuit allows its manufacturer to generate and publish sets of *S* parameters for it. These parameters are bias, voltage, and temperature dependent. At higher frequencies, parasitics play an enormous role in determining a device's RF performance.

A device operating at signal levels that are more than 1% of its dc equivalent currents and voltages must be evaluated in terms of large-signal performance. Most of the large-signal-equivalent circuits for nonbipolar devices, such as FETs (MOS, silicon JFETs, and members of the GaAsFET family), are analytic approximations that essentially bear no physical insight into the inner workings of the transistors. This fact is particularly painful because to look at high-order intermodulation distortion, one needs to have third and fourth derivatives of mathematically continuous equations. The SPICE approach cannot easily be translated into modern harmonic-balance simulators because such cop-outs as IF THEN ELSE statements in programming provide everything else but a continuous model. In the case of our own work, we used complicated curve-fitting equations to obtain analytic equations and their derivatives to accurately provide a large-signal model that is valid over a wide range of dc and RF.

2-4-2 Typical SPICE Parameters

Since we are about to evaluate bipolar microwave transistors, junction FETs, MOSFETs (model level 3), and GaAsFETs, Tables 2-20, 2-21, 2-22, and 2-23 list typical parameters for the devices we have used. These parameters can be obtained by the Scout program with the appropriate measurements. The BSIM model for MOSFETs, applicable for submicron technology transistors, requires an enormous level of parameter extraction and has not fully been validated for the LDMOS-type transistors currently favored for RF and microwave applications.

Table 2-20 BFR193W BJT

IS = 2.738E – 16	BF = 125	NF = 0.95341
VAF = 24	IKF = 0.26949	ISE = 1.0627E – 14
NE = 1.935	BR = 14.267	NR = 1.4289
VAR = 3.8742	IKR = 0.037925	ISC = 3.7409E – 17
NC = 0.94371	RB = 15	IRB = 0.00091763
RBM = 1.8368	RE1 = 0.76534	RC2 = 0.11938
CJE = 1.1824E – 15	VJE = 0.70276	MJE = 0.48654
TF = 1.8828E – 11	XTF = 0.69477	VTF = 0.8
ITF = 0.00096893	PTF = 0	CJC = 9.3503E – 13
VJC = 1.1828	MJC = 0.30002	XCJC = 0.053563
TR = 1.0037E – 09	VJS = 0.75	MJS = 0
XTB = 0	EG = 1.11	XTI = 3
FCC = 0.72063	LB = 0.57E – 9	LC = 0.00E – 9
LE = 0.43E – 9	CBCP = 0.101E – 12	CCEP = 0.175E – 12
CBEP = 0.061E – 12	VCMX = 10 V	

The meaning and significance of the various parameters is best explored in a book on SPICE or semiconductor physics [22–24].

2-4-3 Noise Modeling

Diode Noise Model. The noise model for the diodes (Figure 2-122 and Table 2-24) consists of two contributions: the shot noise and the flicker noise. The shot noise is computed automatically and does not require any parameters. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications).
2. Using bias-dependent flicker noise coefficients (specifying KF and AF at multiple bias points).

The noise generators in the diode noise model are the series parasitic resistance, R_S, and the intrinsic junction. Figure 2-122 illustrates the intrinsic junction noise generator. Let Δf be the bandwidth (usually normalized to 1 Hz). The intrinsic noise generator has a mean-square value of

$$\langle i_{Dn}^2 \rangle = 2qI_D\,\Delta f + \mathrm{KF}\,\frac{I_D^{\mathrm{AF}}}{f^{\mathrm{FCP}}}\,\Delta f \tag{2-266}$$

Notes on the Diode Noise Model:

1. Shot noise is always present unless the SN parameter is set to zero. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.

Table 2-21 2SK125 JFET[a]

IDSS	= 0.5250E – 01	VP0	= –0.3111E + 01	GAMA	= –0.1867E – 01	E	= 0.1520E + 01
KE	= –0.3856E – 03	SL	= 0.2818E – 01	KG	= –0.2398E + 00	T	= 0.0000E + 00
SS	= 0.7448E – 04	IG0	= 0.2000E – 14	AFAG	= 0.3846E + 02	IB0	= 0.1000E – 04
AFAB	= 0.3800E + 02	VBC	= 0.3000E + 02	R10	= 0.1711E + 02	KR	= 0.0000E + 00
C10	= 0.6609E – 11	K1	= 0.1675E + 01	C1S	= 0.6818E – 33	CF0	= 0.7261E – 11
KF	= 0.1156E + 01	RG	= 0.5000E + 00	RD	= 0.1542E + 01	RS	= 0.1333E + 01
LG	= 0.6098E – 09	LD	= 0.5159E – 08	LS	= 0.1482E – 08	CDS	= 0.4813E – 16
CGE	= 0.1590E – 11	CDE	= 0.3394E – 26	CGSP	= 0.8282E – 13	CDSP	= 0.4832E – 12
ZGT	= 0.5000E + 02	LGT	= 0.4712E – 01	ZDT	= 0.5000E + 02	LDT	= 0.3998E – 01
CGDP	= 0.3653E – 12	ZST	= 0.5000E + 02	LST	= 0.1495E – 01	CGDE	= 0.3831E – 12
CGSB	= 0.3120E – 13	CDSB	= 0.5896E – 12	VDMX	= 10		

[a]The currently implemented model for the junction FET is too primitive for serious RF applications. We have therefore taken the approach (liberty) of using the Materka parameter extraction approach for silicon junction FETs. This has resulted in unparalleled high-quality parameters; in particular, the knee voltage behavior has significantly improved, as well as the overall frequency response.

Table 2.22 GaAs MESFET

IDSS = 0.1077E+00	VP0 = –0.1800E+01	GAMA = –0.5741E-01	E = 0.1290E+01
KE = –0.1155E-01	SL = 0.1652E+00	KG = –0.1782E+00	T = 0.0000E+00
SS = –0.1208E+01	IG0 = 0.2130E-11	AFAG = 0.2740E+02	IB0 = 0.5680E-09
AFAB = 0.1826E+01	VBC = 0.9000E+01	R10 = 0.8382E+01	KR = 0.6359E+00
C10 = 0.5964E-12	K1 = 0.1296E+01	C1S = 0.0000E+00	CF0 = 0.6110E-13
KF = 0.9775E+00	RG = 0.1996E+01	RD = 0.1296E+01	RS = 0.1234E+01
CDS = 0.7852E-13	CDSD = 0.1000E-07	RDSD = 0.1581E+03	CGE = 0.1609E-12
CDE = 0.8674E-13	VDMX = 8		

Table 2.23 The 1 μm × 750 μm level 3 LDMOS FET

CBD = 0.863E-12	CGD0 = 166E-12	CGS0 = 246E-12	GAMA = 0.211
IS = 6.53E-16	KAPA = 0.809	MJ = 0.536	NSUB = 1E+15
PB = 0.71	PBSW = 0.71	PHI = 0.579	RD = 39
RS = 0.1	THET = 0.588	TOX = 4E-8	U0 = 835
VMAX = 3.38E5	VT0 = 2.78	XQC = 0.41	

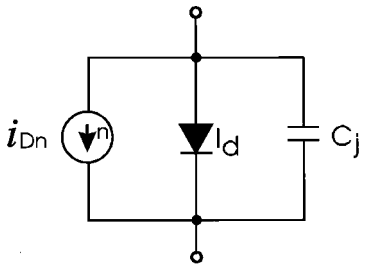


Figure 2-122 Equivalent noise circuit for a diode chip.

Table 2.24 Diode noise model keywords

Keyword	Description	Unit	Default
ID	Required bias current for the data point	Ampere	
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner fequency	hertz	

2. If the value of KF is specified as zero, then the flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.
3. The corner frequency noise model uses the system noise floor to internally compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the diode parameters and *kT*.
4. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

BJT Noise Model. The noise model for the Gummel–Poon BJT model consists of two contributions: shot noise and flicker noise. The shot noise is computed automatically and does not require any parameters. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications) (see Table 2-25).
2. Using bias-dependent flicker noise coefficients (specifying KF and AF at multiple bias points).

Option 1: *Specifying the Bias-Independent Flicker Noise Coefficient*. This option involves the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model.

1. Shot noise is always present unless it is turned off. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying The Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows a bias-dependent flicker noise coefficient (i.e., KF and AF vary with the bias point).

Notes on the BJT Noise Model:

Table 2.25 BJT noise model keywords

Keyword	Description	Unit	Default
IB	Required base bias current for the data point	ampere	
VCE	Required collector–emitter voltage for the data point	volt	
VBS	Base–substrate voltage required for LPNP type when four nodes are used	volt	
VCS	Collector–substrate voltage required for NPN or PNP type when four nodes are used	volt	
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

1. KF, AF, and FC can be specified as bias dependent. If only one set of noise data is specified, the corresponding bias point is not meaningful because all parameters are considered constant over all bias values. However, the bias point is needed for the program to identify the data as bipolar noise data.
2. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
3. This noise model of course considers the actual operating temperature, which must be supplied to the model.

Figure 2-123 shows the BJT noise model. Let Δf be the bandwidth (usually normalized to a 1-Hz bandwidth). The noise generators introduced in the intrinsic device are shown below and have mean-square values of

$$\langle i_{bn}^2 \rangle = 2qI_B \,\Delta f + \text{KF}\,\frac{I_B^{\text{AF}}}{f^{\text{FCP}}}\,\Delta f \tag{2-267}$$

$$\langle i_{cn}^2 \rangle = 2qI_C \Delta f \tag{2-268}$$

$$\langle i_{R_{bb}}^2 \rangle = \frac{4kT}{R_{bb}}\,\Delta f \tag{2-269}$$

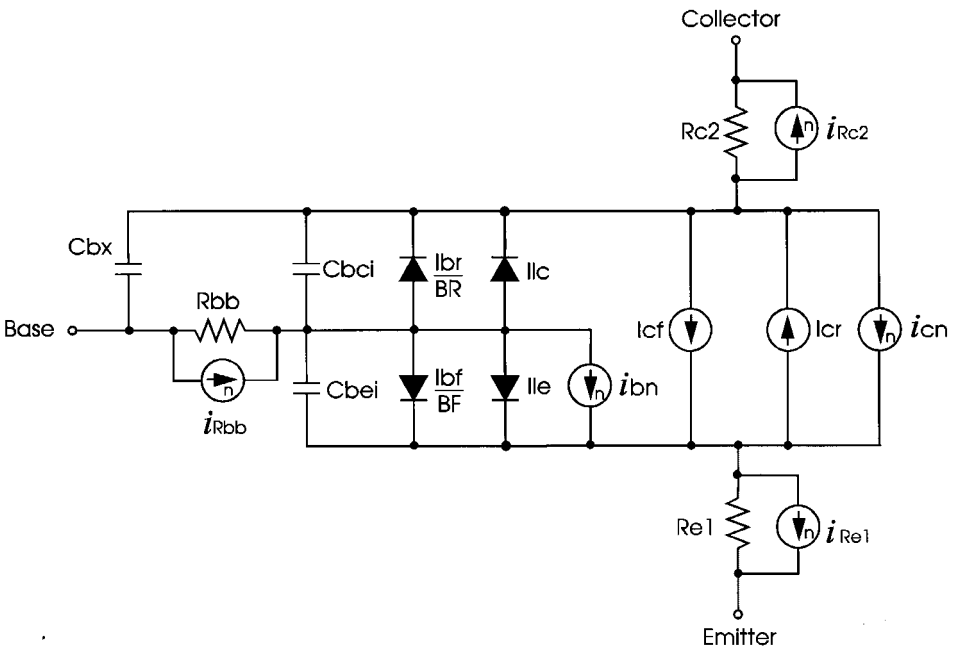


Figure 2-123 BJT noise model (not showing extrinsic parasitics). Current sources with *n* are noise sources.

$$\langle i_{R_{e1}}^2 \rangle = \frac{4kT}{R_{e1}} \Delta f \tag{2-270}$$

$$\langle i_{R_{c2}}^2 \rangle = \frac{4kT}{R_{c2}} \Delta f \tag{2-271}$$

$$I_B = \frac{I_{bf}}{BF} + I_{le} \tag{2-272}$$

$$I_C = I_{cf} - I_{cr} \tag{2-273}$$

JFET and MESFET Noise Models. The noise model for the FETs consists of two contributions: the shot noise and the flicker noise. There are two options to specify noise in the FET model:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (Table 2-26) to determine the flicker noise (this option is usually sufficient for most applications). The shot noise will be automatically computed using the SPICE equation.
2. Using bias-dependent flicker noise coefficients through a reference in the DATA block (specifying KF and AF at multiple bias points) and specifying the four noise parameters [$F_{\min}$, Γ_{opt} (magnitude and phase), and R_n] at multiple bias points.

Option 1: *Specifying the Enhanced SPICE Noise Model.* Option 1 is the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model. The drain noise model has the form

$$\langle |I_{dn}|^2 \rangle = 4K_B T \frac{2g_m}{3} \Delta f + \text{KF} \frac{|I_D|^{\text{AF}}}{f^{\text{FCP}}} \Delta f \tag{2-274}$$

Table 2.26 FET noise model keywords

Keyword	Description	Unit	Default
FN	Noise data measurement frequency	hertz	1.0 GHz
VGS	Required gate–source voltage for the data point	volt	
VDS	Required drain–source voltage for the data point	volt	
FMIN	Required minimum noise figure in dB at FN		
MGO	Required magnitude of optimum noise reflection coefficient at FN		
PGO	Required phase of optimum noise reflection coefficient at FN		
RN	Required normalized noise resistance at FN		
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

where the shot noise is derived from g_m and the flicker noise is proportional to KF and the drain channel current, I_D, and inversely proportional to frequency. The AF and FCP parameters tailor the flicker noise dependence on bias and frequency, respectively.

1. Shot noise is always present unless it is turned off. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, then flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying the Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows the specification of the complex bias-dependent nature of the shot noise and flicker noise. At high frequencies, the equivalent noise sources are correlated (the SPICE noise model does not account for this correlation). The complete evaluation of the shot noise sources can be determined from the four noise parameters. Since these are functions of bias, they can be specified over the (V_{GS}, V_{DS}) bias plane.

Additionally, a bias-dependent flicker noise coefficient (i.e., KF and AF vary with current) can be specified.

The MESFET noise model uses the four measured noise data [$F_{\min}$, Γ_{opt} (magnitude and phase), and R_n] at one frequency and multiple arbitrary bias points. The program uses these data and the FET model parameters to de-embed the noise data to an intrinsic noise model. The intrinsic model is accurate at all frequencies and, therefore, can predict the noise performance at all frequencies given data at just one frequency point. Built-in bias-dependent characteristics are used if multi-bias noise data are not provided.

Notes on the FET Noise Model

1. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
2. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

Noise in a MESFET is produced by sources intrinsic to the device. The same approach, but with different flicker corner frequencies, is highly applicable to JFETs and MOSFETs. For more detail as to simulation, see the Element library book for the active-device portion of Ansoft's Serenade Design Environment product. The equivalent noisy circuit of an intrinsic FET is represented in Figure 2-124.

The intrinsic FET is internally represented as a noiseless nonlinear two-port with one equivalent noise current connected across the gate–source terminal, and one across the drain–source terminal. The correlations of the gate and drain noise current sources are

$$\langle |I_{gn}|^2 \rangle = 4K_B T \, \Delta f \frac{\omega^2 C_{gs}^2}{g_m} R \tag{2-275}$$

$$\langle |I_{gn}|^2 \rangle = 4K_B T \, \Delta f \, g_m P \tag{2-276}$$

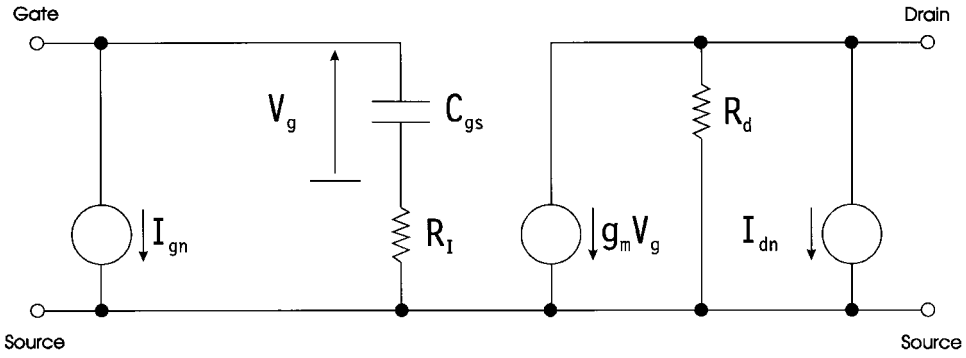


Figure 2-124 Equivalent noise circuit of an intrinsic FET device.

$$\langle I_{gn} I_{dn}^{*} \rangle = 4K_B T\, \Delta f\, j\omega C_{gs} \sqrt{PR}\, C \tag{2-277}$$

The correlation matrix of the noise current sources is

$$C_{dc}(\omega) = \frac{2}{\pi} K_B T\, d\omega \begin{bmatrix} \dfrac{\omega^2 C_{gs}^2}{g_m} R & -j\omega C_{gs}\sqrt{PR}\, C \\ j\omega C_{gs}\sqrt{PR}\, C & g_m P \end{bmatrix} \tag{2-278}$$

The gate and drain noise parameters R and P and the correlation coefficient C are related to the physical noise sources acting in the channel and are functions of the device structure and bias noise parameters. By defining measured noise parameters, $F_{\min}$, R_n, and Γ_{opt}, and using a noise de-embedding procedure, the parameters R, P, and C and the intrinsic noise correlation matrix of a FET device as functions of device bias are determined by the program.

In addition to the noise sources shown above, the flicker (1/f) noise can also be modeled by means of a noise current source connected in parallel with the intrinsic drain port. The flicker noise component in a narrow band, Δf, is expressed in the form

$$\langle |I_f|^2 \rangle = \mathrm{Q}\, \Delta f \frac{|I_D|^{\mathrm{AF}}}{f^{\mathrm{FCP}}} \tag{2-279}$$

where I_D is the instantaneous value of the channel current, and Q, AF, and FCP are empirical parameters. In most practical cases, AF and FCP are directly obtained from measurements (typically, AF = 2 and FCP = 1), while Q is not. In Ansoft's Serenade Design Environment, Q is either provided directly using KF or is computed by providing the flicker corner frequency (FC). FC is the frequency at which the flicker noise equals the shot/diffusion noise. The corner frequency is defined by the equation

$$\mathrm{Q} \frac{|I_D|^{\mathrm{AF}}}{f_{\mathrm{C}}^{\mathrm{FCP}}} = g_m P \tag{2-280}$$

Given the corner frequency FC and the measurement bias point V_{gs} and V_{ds}, the program automatically computes I_D, g_m, and P, and finally Q.

More information on FET noise modeling can be found in the literature [25–31].

MOSFET Noise Model. The MOSFET noise model (Figure 2-125 and Table 2-27) consists of two contributions: the shot noise and the flicker noise. The shot noise is computed automatically and does not require any parameters. It can be turned off by specifying SN = 0. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications).

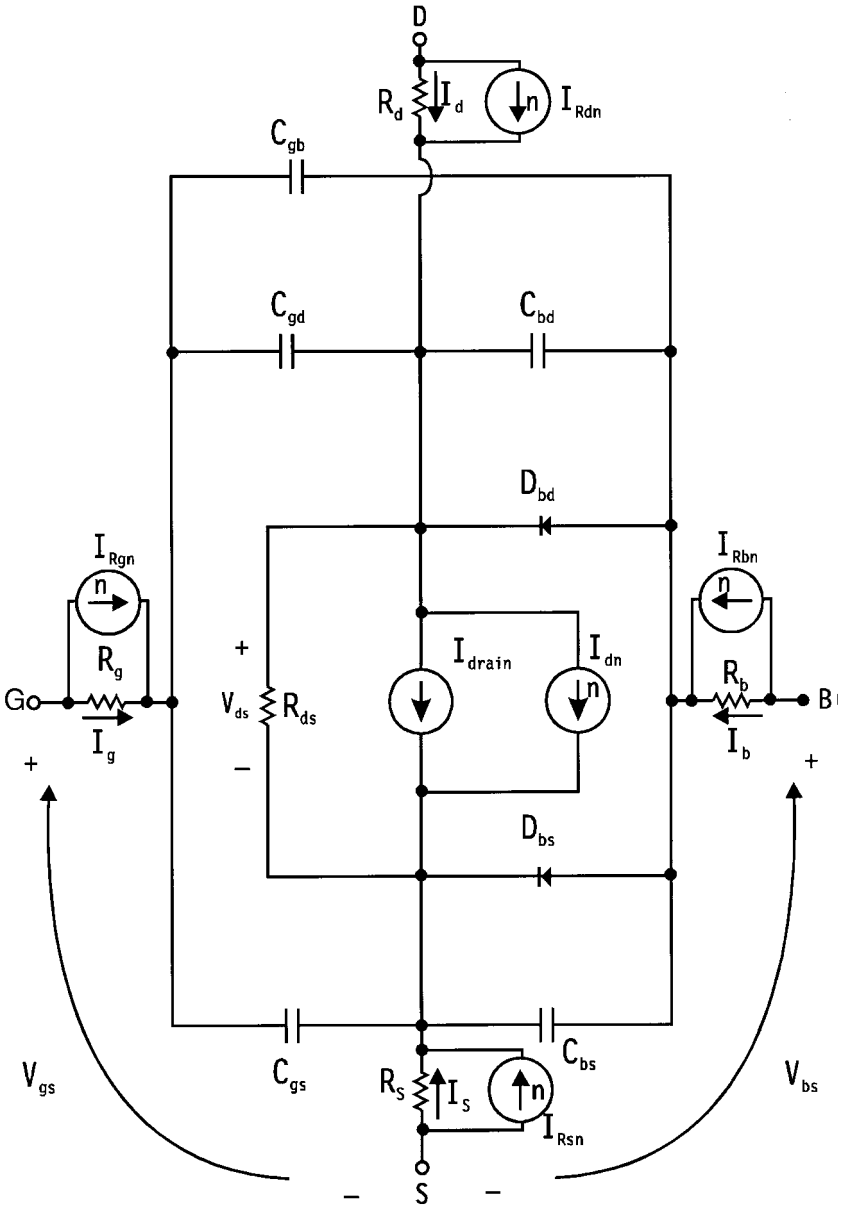


Figure 2-125 Equivalent noise circuit of an intrinsic MOSFET device.

Table 2.27 MOSFET noise model keywords

Keyword	Description	Unit	Default
VGS	Required gate–source bias for the data point	volt	
VDS	Required drain–source for the data point	volt	
VBS	Required drain–bulk for the data point	volt	
KF	Flicker noise coefficient		1.0E–13
AFT	Bias exponent of the flicker noise model		2.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

2. Using bias-dependent flicker noise coefficients through a reference (specifying KF and AF at multiple bias points).

Option 1: *Specifying the Enhanced SPICE Noise Model.* This option is the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model (the flicker noise is considered bias dependent).

1. Shot noise is always present unless the SN parameter is set to zero. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, then the flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying the Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows a bias-dependent flicker noise coefficient (i.e., KF and AF vary with drain current). The MOSFET noise model data are given and referenced by a model parameter.

Notes on the MOSFET Noise Model

1. KF, AF, and FC can be specified as bias dependent. If only one set of noise data is specified, the corresponding bias point is not meaningful because all parameters are considered constant over all bias values. However, the bias point is needed for the program to identify the data as MOSFET noise data.
2. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
3. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

Let Δf be the bandwidth (normalized to 1 Hz). The noise generators introduced in the intrinsic device are shown below and have mean-square values of

$$\langle i_{dn}^2 \rangle = \frac{8kTg_m}{3}\Delta f + \mathrm{KF}\frac{I_D^{\mathrm{AF}}}{f^{\mathrm{FCP}}}\Delta f$$

$$\langle i_{Rgn}^2 \rangle = 4\frac{kT}{R_g}\Delta f$$

$$\langle i_{Rdn}^2 \rangle = 4\frac{kT}{R_d}\Delta f$$

$$\langle i_{Rsn}^2 \rangle = 4\frac{kT}{R_s}\Delta f$$

$$\langle i_{Rbn}^2 \rangle = 4\frac{kT}{R_b}\Delta f$$

We include this MOSFET noise model (used for quite awhile) for completeness. At the moment, we do not know which MOSFET noise model the industry will settle on in the future.

Modern CAD tools, such as Ansoft's Serenade product, use these models to generate quite accurate noise data based on a good linear equivalent model. Internally Serenade uses the noise-correlation-matrix method (first introduced by Russer). The same noise-correlation techniques apply to BJTs, all FETs except MOS, and HBTs (such as SiGe devices).

2-4-4 Scalable Device Models

Since diodes and transistors are scalable, here are guidelines for how to scale them:

Microwave Diode

$$I_D = Area \times I_D$$
$$C_j = Area \times C_j$$
$$R_D = R_D/\text{Area}$$

PIN Diode

$$I_D = \text{Area} \times I_D$$
$$C_j = \text{Area} \times C_j$$
$$R_S = R_S/\text{Area}$$
$$R_{\max} = R_{\max}/\text{Area}$$

Bipolar

$$I_{bf} = \text{Area} \times I_{bf} \qquad I_{br} = \text{Area} \times I_{br}$$
$$I_{le} = \text{Area} \times I_{le} \qquad I_{lc} = \text{Area} \times I_{lc}$$
$$I_{cf} = \text{Area} \times I_{cf} \qquad I_{cr} = \text{Area} \times I_{cr}$$
$$C_{bc} = \text{Area} \times C_{bc} \qquad C_{be} = \text{Area} \times C_{be}$$
$$C_{bx} = \text{Area} \times C_{bx} \qquad R_{bb} = R_{bb}/\text{Area}$$
$$\text{RB2} = \text{RB2/Area} \qquad \text{RC2} = \text{RC2/Area}$$
$$\text{RE1} = \text{RE1/Area} \qquad I_{jss} = I_{jss} \times \text{Area}$$
$$C_{jss} = C_{jss} \times \text{Area}$$

Materka FET

$$\begin{aligned}
\text{IGSS} &= \text{IGSS} \times \text{Area} \\
\text{CGS0} &= \text{CGS0} \times \text{Area} \\
\text{CGS1} &= \text{CGS1} \times \text{Area} \\
\text{CDVC} &= \text{CDVC} \times \text{Area} \\
\text{CDVS} &= \text{CDVS} \times \text{Area} \\
R_G &= R_G \times \text{Area}/(\text{Number of fingers} \times 2) \\
R_D &= R_D/\text{Area} \\
R_S &= R_S/\text{Area}
\end{aligned}$$

Generating a Databank for Parameter Extraction. Given the fact that one can measure with an automated system (Rohde & Schwarz, Hewlett-Packard) the bias-dependent *S* parameters of any active device, bipolar and FET, and furthermore, that Hewlett-Packard makes a computer-controllable dc *I–V* curve tracer, we can generate a huge databank that allows us to have corresponding values for both RF and dc operating points. To validate this, one needs to run a large-signal simulator, such as Ansoft's Serenade Design Environment, or its equivalent from other manufacturers, enter the large-signal parameters, and generate a set of *S*-parameters that correspond to the dc values for which we have measurements. The agreement between measured and predicted points for both RF and dc values is a good measure to evaluate the accuracy of the parameter extraction and the simulator's performance. We cannot stress often enough, however, the fact that while most simulator models are quite good, the weakest point of the link is the parameter extraction program.

Currently, the best parameter extraction for bipolar transistors and FETs, excluding the BSIM model, is implemented in the Scout program by Ansoft, based on modified techniques found in the literature. These techniques are based on the fact that present-day nonlinear modeling of microwave devices, especially FETs, is not exactly adequate to describe all effects found within these devices. In the case of FETs, the most troublesome parameter is R_{DS}, which is really more bias dependent than practically all existing models take into consideration. Techniques used in published extraction methods extract nonlinear models independently for each section of a device model, and then assume that the proper response will be generated when all parts are put together. For instance, many extraction techniques independently fit the *I–V* data of the section describing the *I–V* curves and then fit the *C–V* data to the equivalent small-signal models. This results in an approximation to the bias-dependent response of the complete model.

In the technique used in the Scout program, model parameters are extracted simultaneously so the effects between model sections are accounted for and the device is treated as a whole. The most critical area happens to be in the linear region, typically around the pinchoff area, and for very small supply voltages. The Scout techniques use measured data taken at a wide variety of bias ranges and frequencies to define the model response. The possible solution region for the complete model becomes well identified and uniqueness of solution is vastly improved. The Scout program is interactive in nature and displays simultaneously measured versus predicted data. This allows for fine-tuning of the parameters.

Over the years, the industry has adapted a large number of models. Those supported by the Scout parameter extraction program include:

- *npn* Gummel–Poon BJT
- *pnp* Gummel–Poon BJT

- *npn* heterojunction BJT
- *pnp* heterojunction BJT
- General-purpose *n*-channel JFET
- General-purpose *p*-channel JFET
- Chalmers (Angelov) MESFET/HEMT
- Curtice–Ettenberg cubic MESFET/HEMT
- Curtice quadratic MESFET/HEMT
- IAF (Berroth) MESFET/HEMT
- Modified Materka–Kacprzak MESFET/HEMT
- Raytheon (Statz) MESFET/HEMT
- TriQuint (TOM-1 and TOM-2) MESFET/HEMT
- Ansoft physics-based MESFET

The Ansoft physics-based MESFET model is unique and politically sensitive because its users are essentially required to enter all device fabrication parameters to get the appropriate answer, and most, if not all, companies will stay away from handing out such company proprietary recipes. The validation of this model was done using manufacturing data for transistors that were developed under government contract and therefore accessible for the team members. Compact Software/Ansoft over the last ten years has been involved in a large number of government-funded research and development programs, and practically all modern modeling implementations and their validation were done in conjunction with the largest semiconductor manufacturers.

As far as further validation is concerned, besides our looking at the matching of dc *I–V* curves and *S* parameters, important parameters such as 1-dB compression point, power-added efficiency, and harmonic content are important criteria to evaluate the quality of such an undertaking. As far as hardware is concerned, one needs to have a network analyzer, such as the HP8510 or its equivalent. Its selection is determined by the highest cutoff frequency where measurements are necessary. A dc feed capability, up to several amperes, through the measuring ports, and appropriate power supplies are essential, as is a dc *I–V* curve tracer (HP4145). The built-in various optimizers of the program establish the best possible match between the selected parameters and the measured results.

The area of proper modeling has always been fascinating because the active-device model can make or break first-pass success. Figure 2-126 shows the doubler gain comparison of a model supplied by NEC (the device manufacturer, parameter extraction done with Scout, and parameter extraction supplied from the *Microwave Engineering Europe* (*MEE*) magazine in its CAD review of May 1994). It is somewhat unclear where the *MEE* model came from. The actual task given by *MEE* was to simulate a frequency doubler against measured data. This type of simulation required two areas of high precision, one being the model and the other being the electromagnetic modeling of the discontinuities of the circuit—in this case, a stub element. The circuit itself was very simple. The manufacturer-supplied data gave the poorest results (we believe the parameter extraction came from a third party not quite up to speed in this area). The results that came from the Scout parameter extraction and the third one that *MEE* supplied were fairly close, yet the Scout solution, including the electromagnetic simulation part, gave the best answer. More simulators have since appeared on the market, and it would be interesting to revisit this topic. Figure 2-127 shows a physical layout of the doubler circuit.

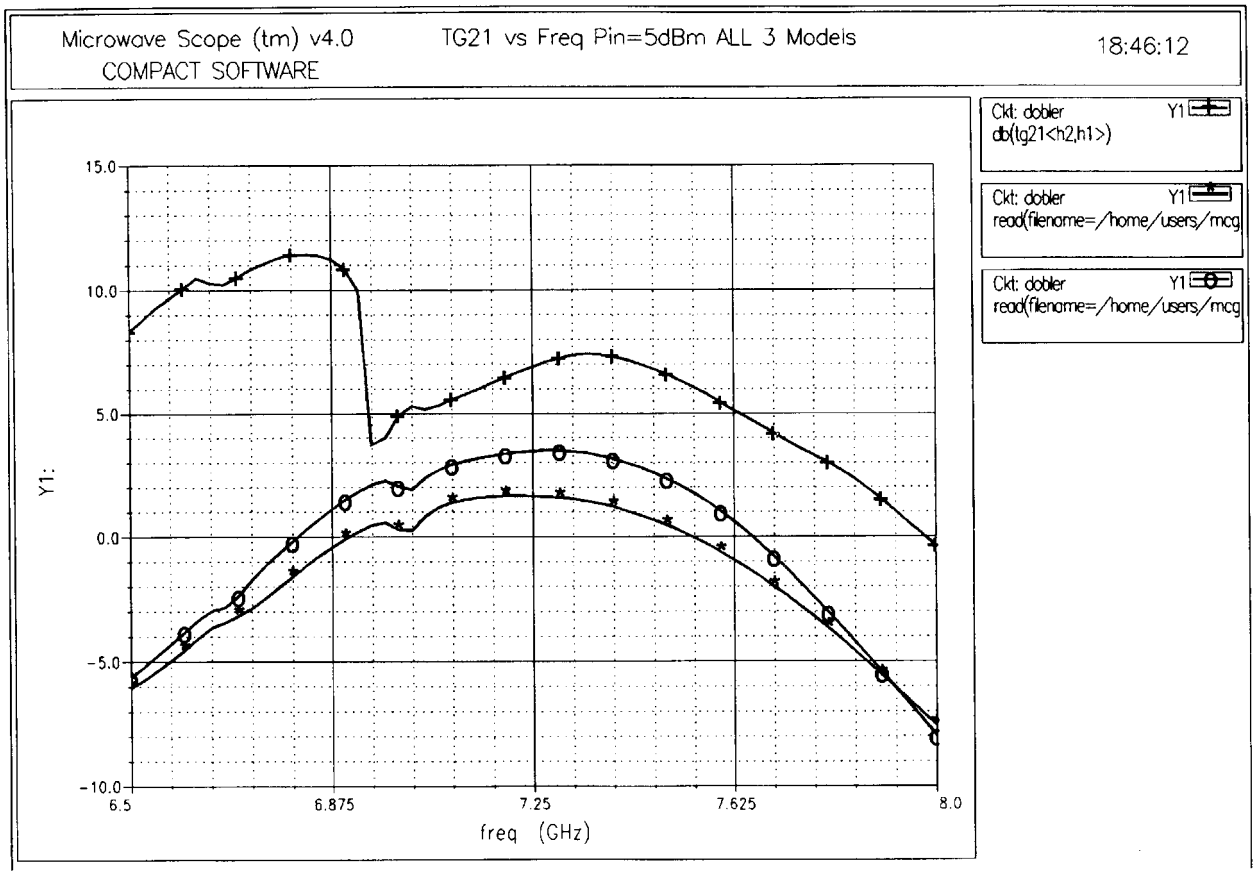


Figure 2-126 Gain comparison of a MESFET doubler based on the three device models used: NEC (+), Ansoft (O), and *Microwave Engineering Europe* (*).

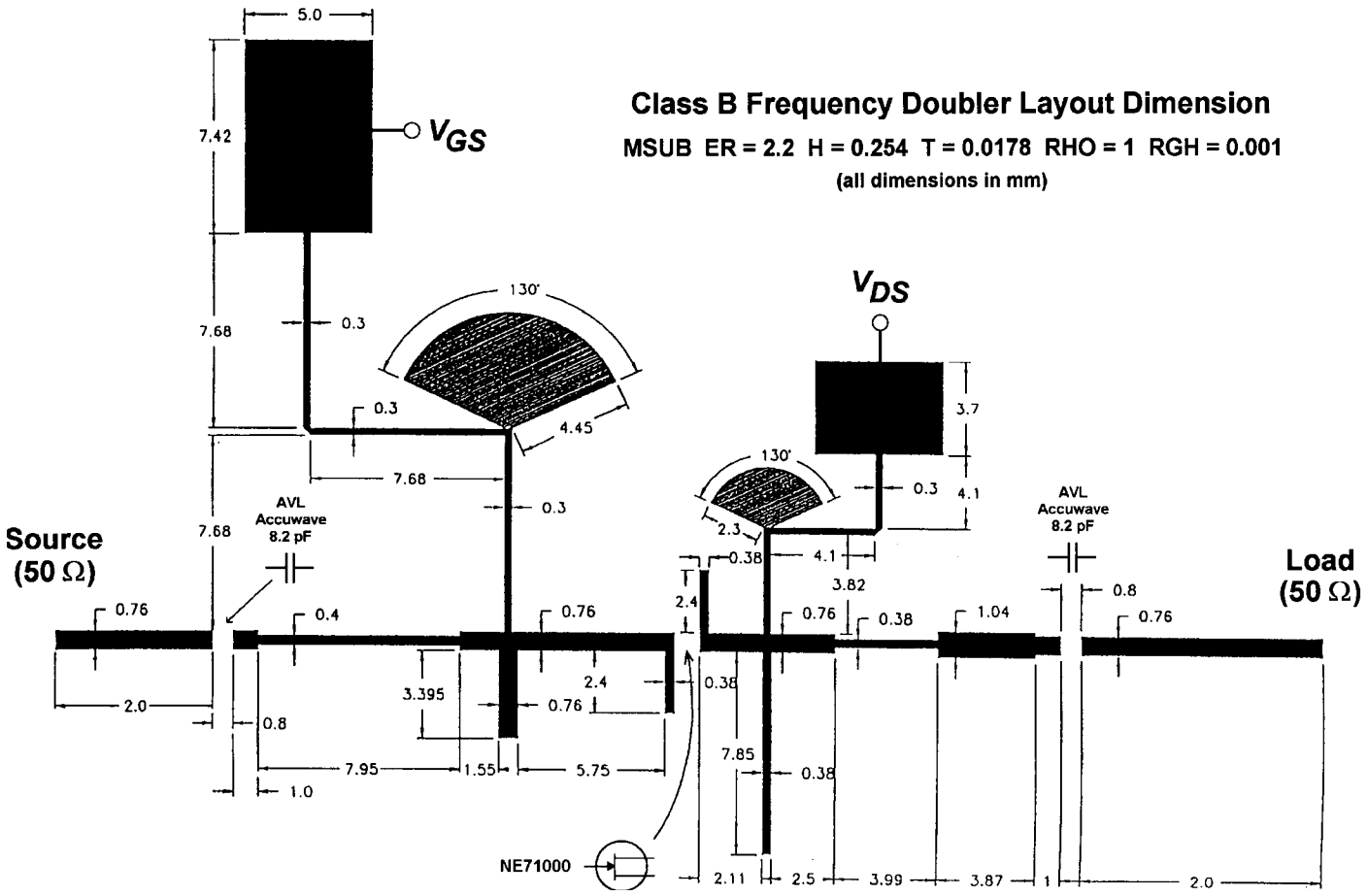


Figure 2-127 Layout of the MESFET doubler.

Test Setup. The measuring equipment required for the actual extraction includes:

- Semiconductor parameter analyzer (HP4145)
- Vector network analyzer (ZVR, HP8510, W360)
- Bias supplies and meters
- Power RF source, power meters, spectrum analyzer, and so on

The required software tools include:

- GPIB controller: Compact NETCOM (network analyzer control program)
- Parameter extraction performed by Scout
- Small-signal and harmonic-balance validation: Serenade nonlinear analysis (Microwave Harmonica) or its equivalent

After we have assembled all these tools, the device extraction process can begin. It will follow the flow as shown in Figure 2-128. Depending on the numbers for the dc/ac measurement, this process can take up to two hours. The Scout program, with its built-in optimizer, can take another two hours. It is therefore reasonable to assume a thorough parameter extraction of a device will take about one day.

Parameter Extraction. As we have outlined, the parameter extraction program Scout, which is optimized for high-frequency and millimeter-wave application, is a Windows-based parameter extraction and optimization program. It requires a network-communication program to acquire dc and *S* parameter measurement data. It supports interactive device fitting and optimization and supports the various models listed earlier. For ease of use, it has an attractive graphical user interface with multiple windows, as can be seen from Figure 2-129. Hewlett-Packard offers more general RF parameter extraction programs, but we did not have access to those. We also expect that some other companies will develop similar tools. An interesting program named HarPE, developed by Professor John Bandler in Canada, may have been withdrawn from the market after his company merger with HP. The program was acclaimed to provide quite good results using measured harmonics. One of the unfortunate deficiencies of most BJT parameter extraction programs known to us is the fact that (as one

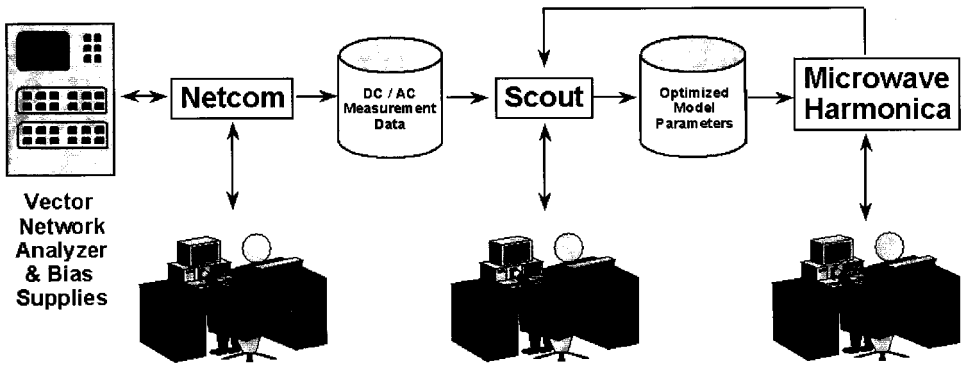


Figure 2-128 Flowchart of the parameter extraction process.

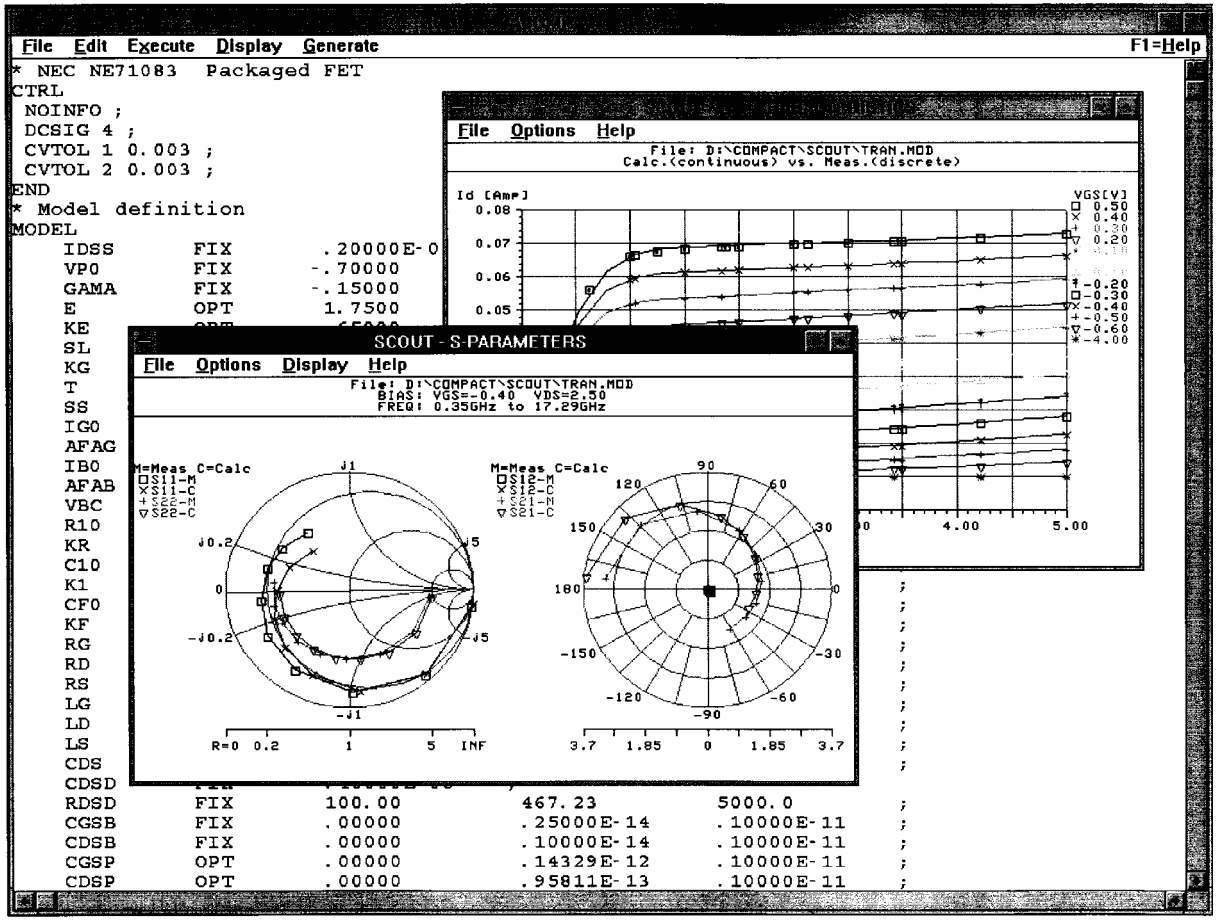


Figure 2-129 The Scout user interface.

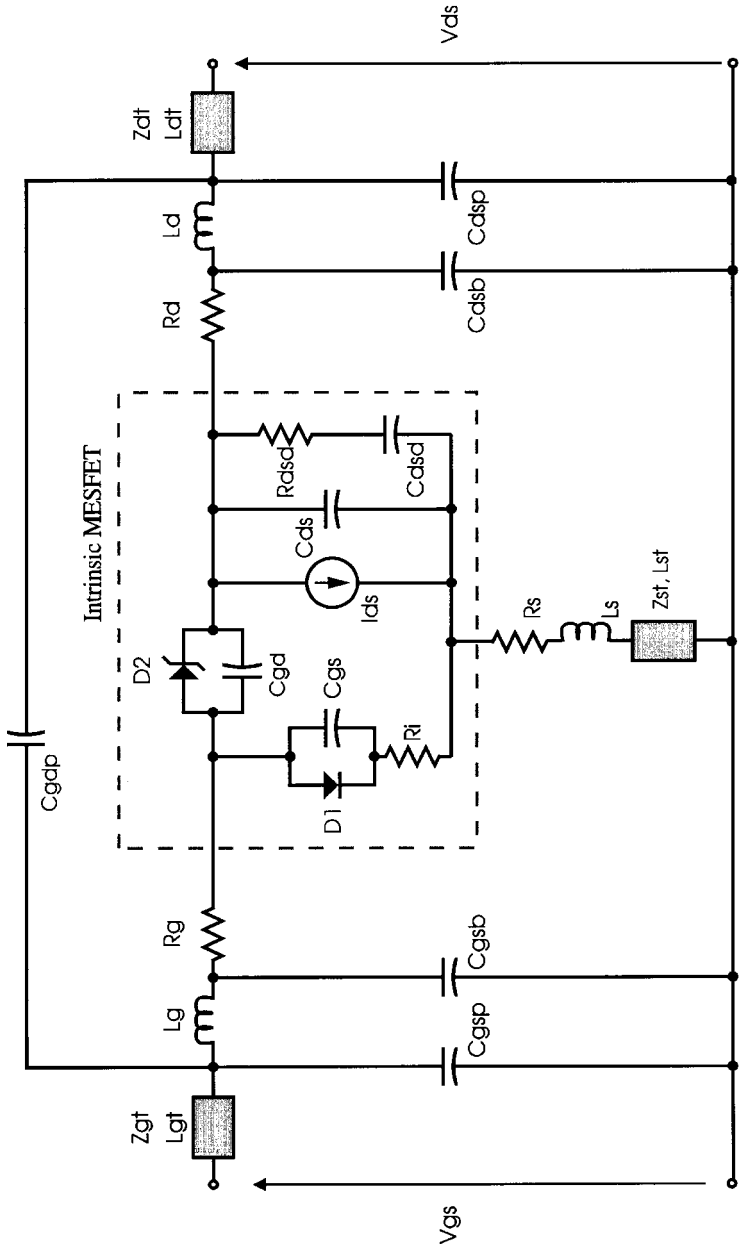


Figure 2-130 Intrinsic model and complete chip/package model.

example) the sensitivity to determine the actual base-spreading resistor for bipolar transistors is poor. The only useful workaround is considering the noise figure of the device, given the fact that the midrange noise factor of any bipolar transistor can be calculated from

$$F = 1 + \frac{R'_{bb}}{R_g} + \frac{0.5(26\ \text{mV}/I_C)}{R_g}$$

(approximate equation at medium frequencies) (2-281)

By solving the equation for R'_{bb}, since we know that R_g, the generator impedance, is 50 Ω, and the emitter diffusion resistance (26 mV/I_C) can be determined from the bias point, we can already eliminate the need to extract R'_{bb} from the HP test equipment. By using HP's program, errors of a factor up to 10 are not uncommon because of the low sensitivity for this parameter. Another headache in the case of the BJT is the determination of the excess phase (PTF) and forward and reverse transit time (TF and TR, respectively). The lack of quality modeling can best be found by examining the measured versus predicted values of S_{22}. Again, practically all the parameter extraction programs we have seen so far suffer from this deficiency.

MESFETs. In the case of a member of the MESFET family, a model as shown in Figure 2-130 is used. It consists of an intrinsic model and a complete chip/package parasitic model. In the case of the MESFET, to obtain first-class results, the following measurements are necessary:

- Dc *I*–*V* measurements
- Dc diode measurements
- Cold-FET and forward-biased-gate *S* parameters
- *S*-parameter measurements over bias
- Harmonic power measurements
- *Z*-Parameter Equations for Figure 2-131

In order to simplify the modeling, it is useful to do cold-FET [Figure 2-131 and Eqs. (2-282), (2-283), and (2-284)] and pinched-FET [Figure 2-132 and Eqs. (2-285), (2-286), and (2-287)] measurements to derive some of the intrinsic or extrinsic values.

$$Z_{11} = R_s + R_g + R_{gs} + \alpha_g R_{ch} + j\omega(L_g + L_s) \tag{2-282}$$

$$Z_{12} = Z_{21} = R_s + \alpha R_{ch} + j\omega L_s \tag{2-283}$$

$$Z_{22} = R_s + R_d + 2\alpha R_{ch} + j\omega(L_s + L_d) \tag{2-284}$$

Capacitance Equations for Figure 2-132

$$C_b = -\frac{\text{Im}(Y_{12})}{\omega} \tag{2-285}$$

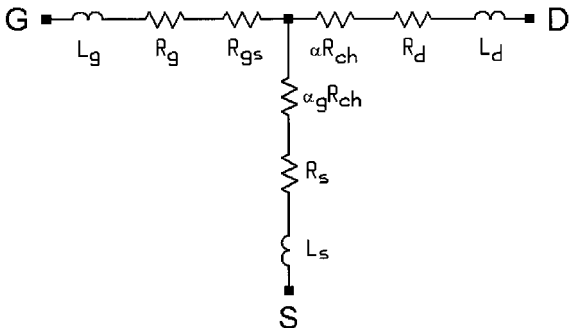


Figure 2-131 Forward-biased gate FET model for cold-FET (V_{DS} = 0, V_{GS} 0) measurements. This approach allows direct extraction of parasitic resistances and inductances.

$$C_{pg} = \frac{\mathrm{Im}(Y_{11})}{\omega} - 2C_b \tag{2-286}$$

$$C_{pd} + C_{ds} = \frac{\mathrm{Im}(Y_{22})}{\omega} - C_b \tag{2-287}$$

For the highest accuracy, the following extraction strategies are recommended:

- Pre-extraction using basic dc and *S* parameters—for example, C_{GS} from S_{11}
- Six extraction steps to determine parameters:
  Basic dc model
  Advanced dc model
  Basic ac model
  Advanced ac model
  Package determination
  Diode characterization
- Large-signal validation
  Power compression and bias shift
  Harmonic power comparison
  Intermodulation characterization

As an example, we are going to use the popular and well-characterized NE71000, which is not a particularly modern device, but has been used for various research projects. Figure

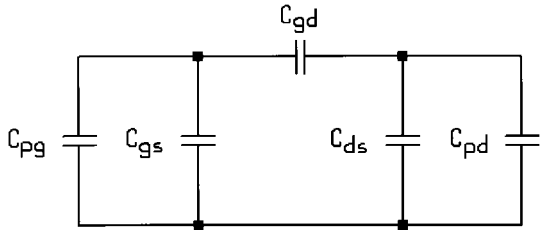


Figure 2-132 Pinched-FET (V_{DS} = 0, V_{GS} *VP*) model. This approach allows direct extraction of parasitic gate and drain capacitances.

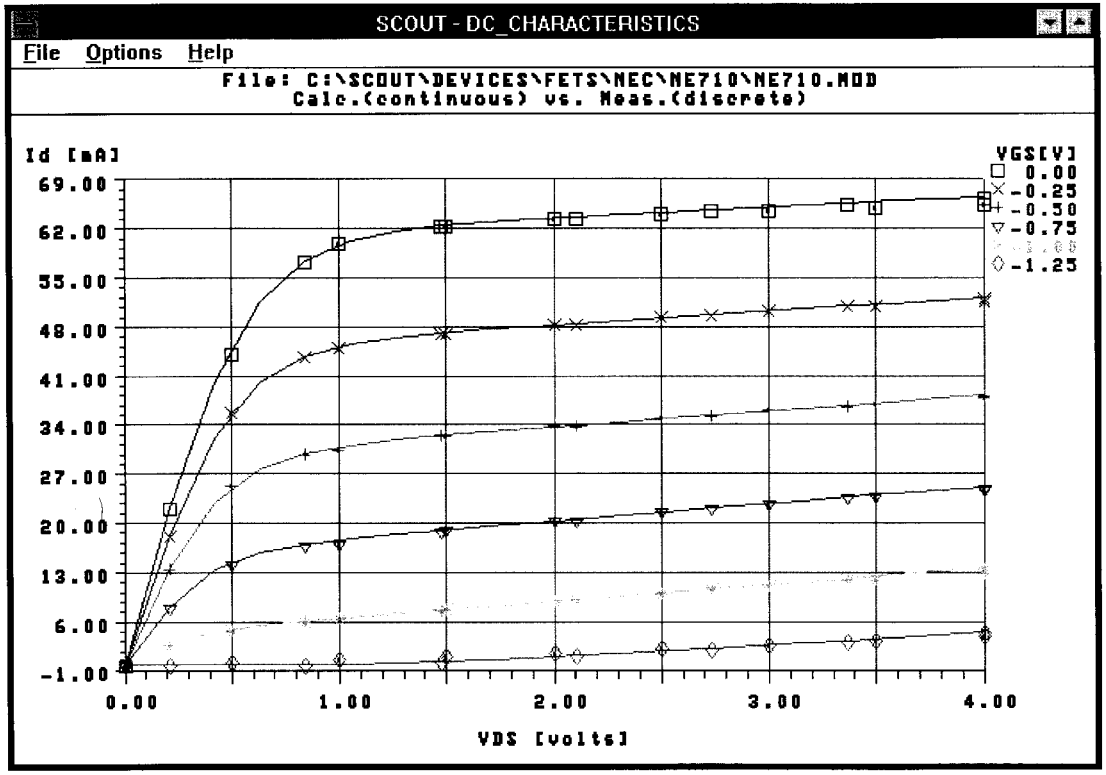


Figure 2-133 Measured NE71000 dc *I–V* curves versus those simulated using the Ansoft Materka model.

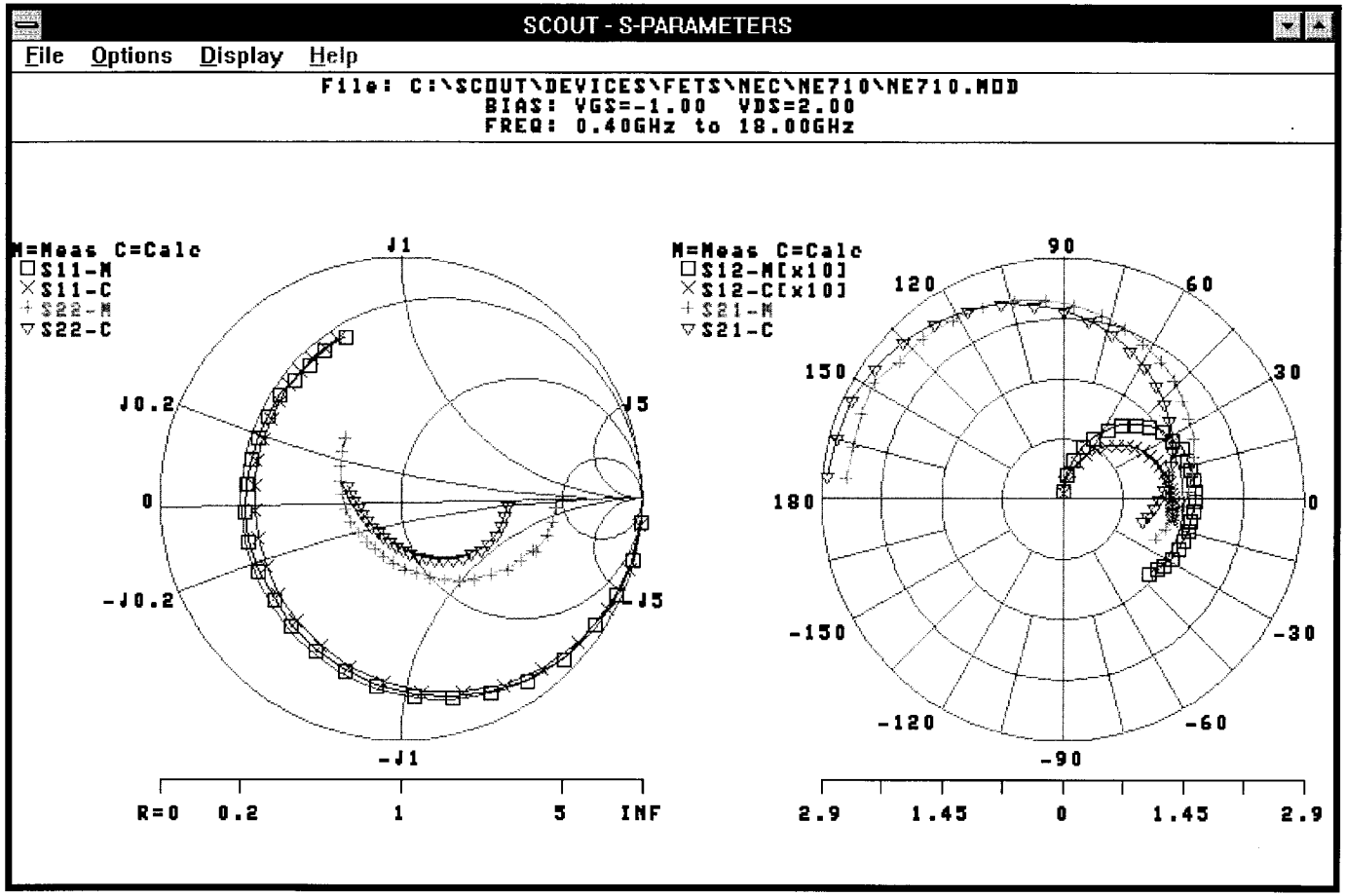


Figure 2-134 Measured versus calculated NE71000 *S* parameters for $V_{GS} = -1.00$ and $V_{DS} = 2.00$.

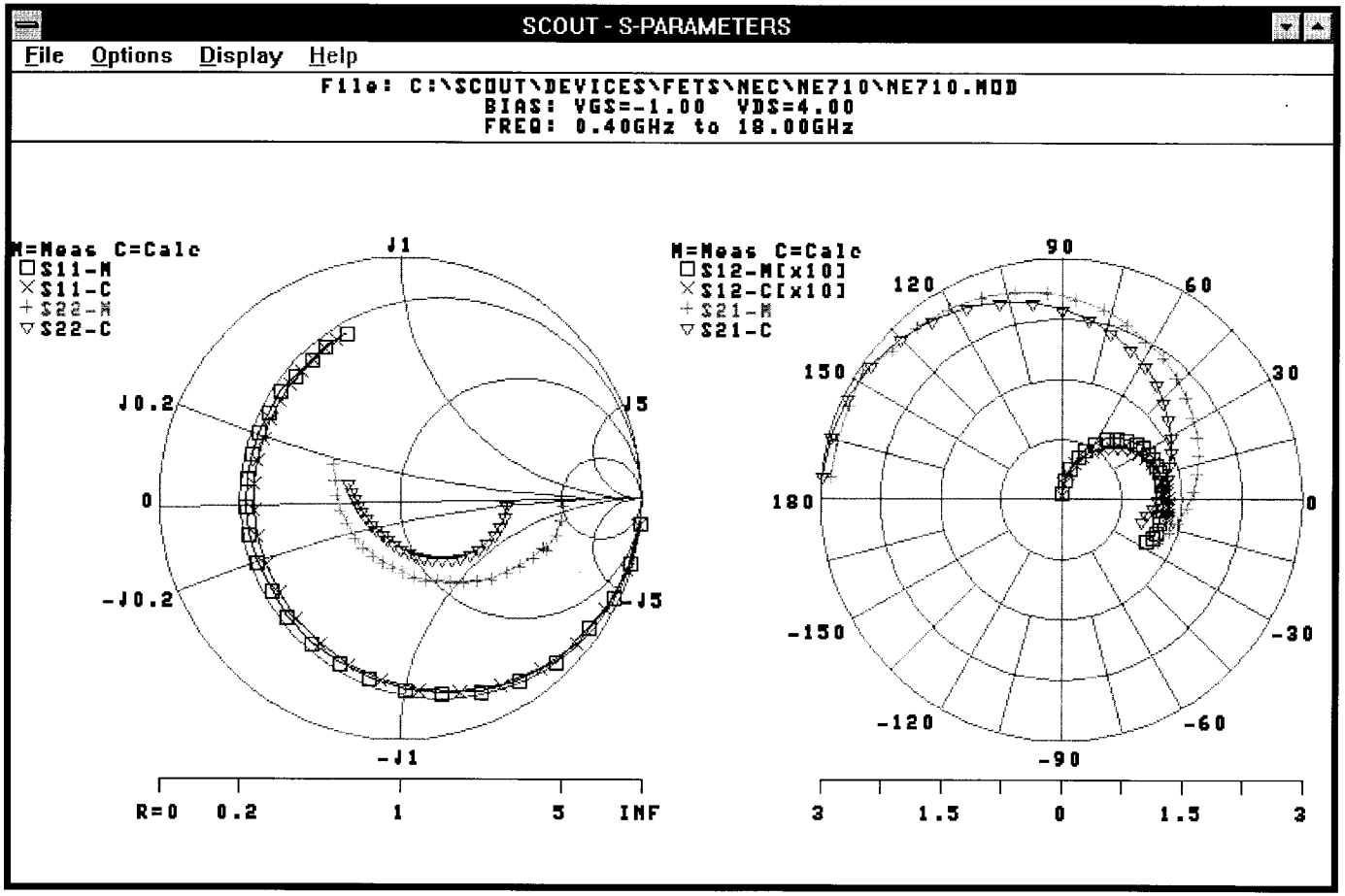


Figure 2-135 Measured versus calculated NE71000 *S* parameters for V_{GS} = –1.00 and V_{DS} = 4.00.

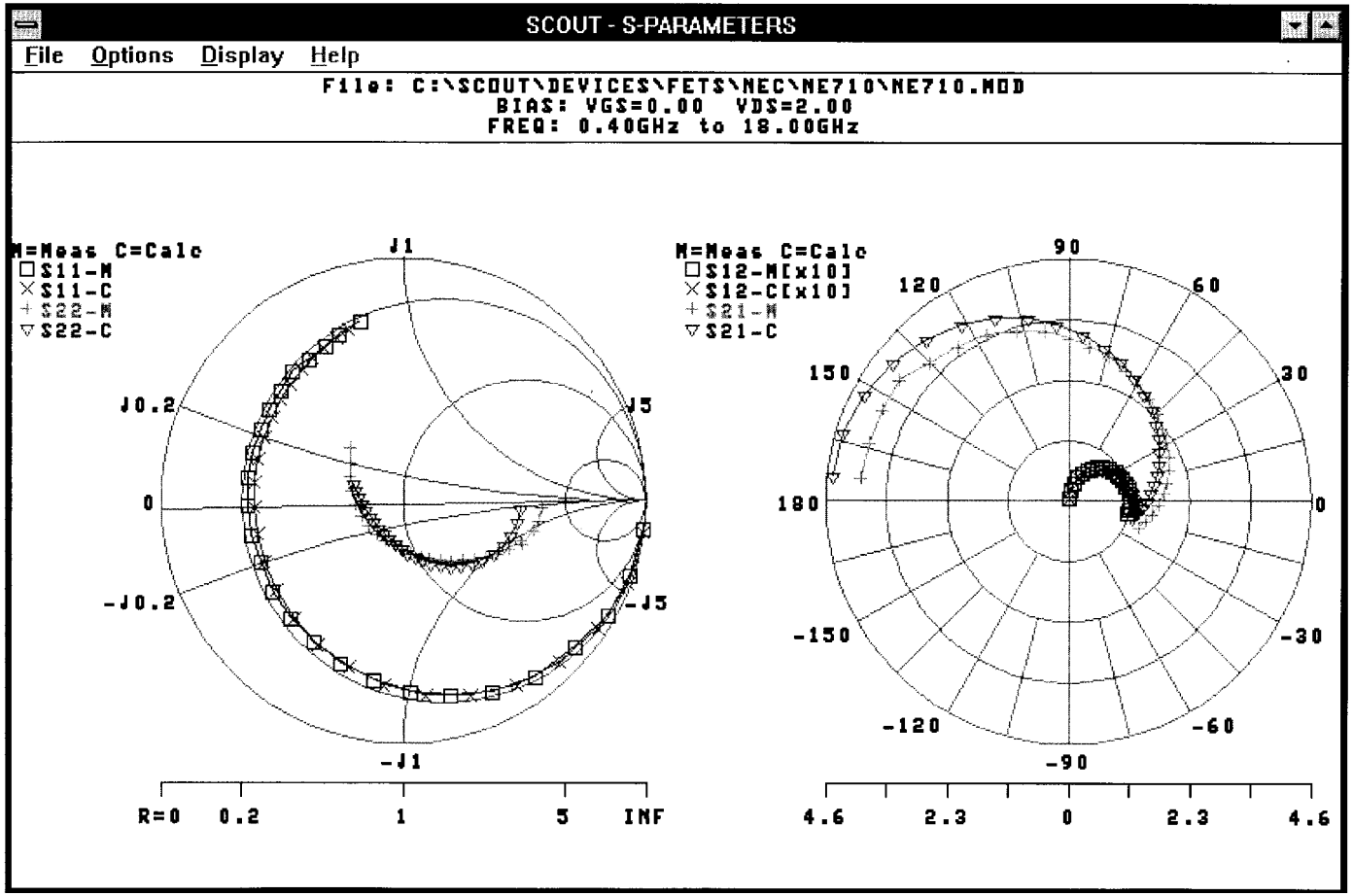


Figure 2-136 Measured versus calculated NE71000 *S* parameters for V_{GS} = 0.00 and V_{DS} = 2.00.

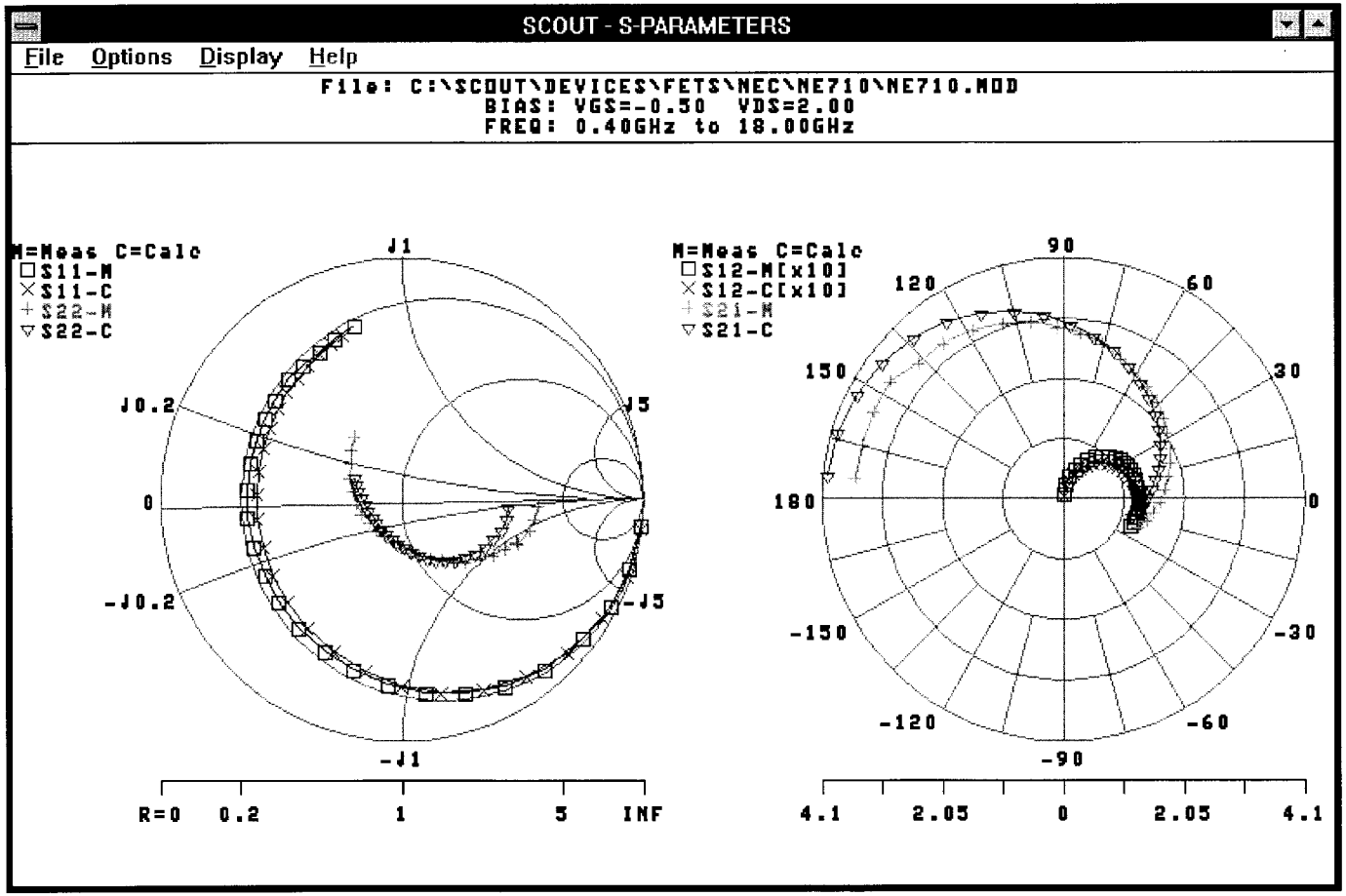


Figure 2-137 Measured versus calculated NE71000 *S* parameters for V_{GS} = –0.50 and V_{DS} = 2.00.

2-133 shows measured versus simulated dc I–V curves for the NE71000. The agreement between measured and predicted dc I–V curves is quite good. In a similar fashion, we obtain the S-parameter fits over bias. They are shown in Figures 2-134 through 2-137.

As mentioned above, 1-dB compression, power saturation, and drain current are functions of the drive. This particular model point is not particularly impressive. We believe this has to do with the original Materka model.

Having brought up the subject of models, Figure 2-138 compares measured data with that obtained with the Angelov, cubic, quadratic, TriQuint, Raytheon, and Materka models. The figure shows that there is no perfect model yet. To put this in perspective, in Figure 2-139 we show the dc I–V error bars for these six models. This graph shows that the least errors for the dc measurement versus modeling are obtained with the Curtice cubic model, followed by the modified Materka model (implemented in Serenade) and TriQuint models. The difference in S parameter matching for RF application is significantly more relevant for the designer.

Things become much more interesting if we compare the S-parameter errors between the Materka shown in Figure 2-140, the Curtice cubic model (Figure 2-141), and the TOM model (Figure 2-142). Improved modeling regarding the R_{DS} is being developed, and we will discuss this at the end of this chapter.

Since the trend is to go to low-voltage power RF operation, several issues must be battled:

- Models are most inaccurate at low V_{DS} and near pinchoff.
- Modeling of low-voltage operation is adversely impacted.
- Thermal effects must be considered, but older models do not account for this.

A Case Study. As a case study, we will consider the Oki KGF1608 power MESFET (0.5 μm × 28 mm). Figures 2-143 and 2-144 show the dc I–V curves and S parameters, respectively. The resolution of the S parameters, based on the scale used, is not particularly impressive.

Figure 2-145 shows the dc I–V curves, with load line, for the KGF1608; the self-heating effect of the transistor can be seen in the droop of the top two curves. Figures 2-146 and 2-147 show drain current and output power versus drive, respectively. The test conditions for these three figures were f = 850 MHz, V_{DS} = 3.4 V, and V_{DS} = 2.025 V.

2-4-5 Conclusions

From all plots of dc and ac error contours, dc I–V curves, and dc error bar charts we can draw the following conclusions.

Dc errors

- *In the Saturation Region*. All the models considered work very well, with a relative error of less than 15%. This means that all models are suitable for typical applications where the device operates in the saturation region.
- *In the Linear Region.* Though the errors in the saturation region are similar for different models, the errors in the linear region are quite different. The Materka, Curtice cubic, and TriQuint models behave better than other models in the linear region. The relative errors of these three models are less than 10%, while other models have more than 60% relative errors at some points.

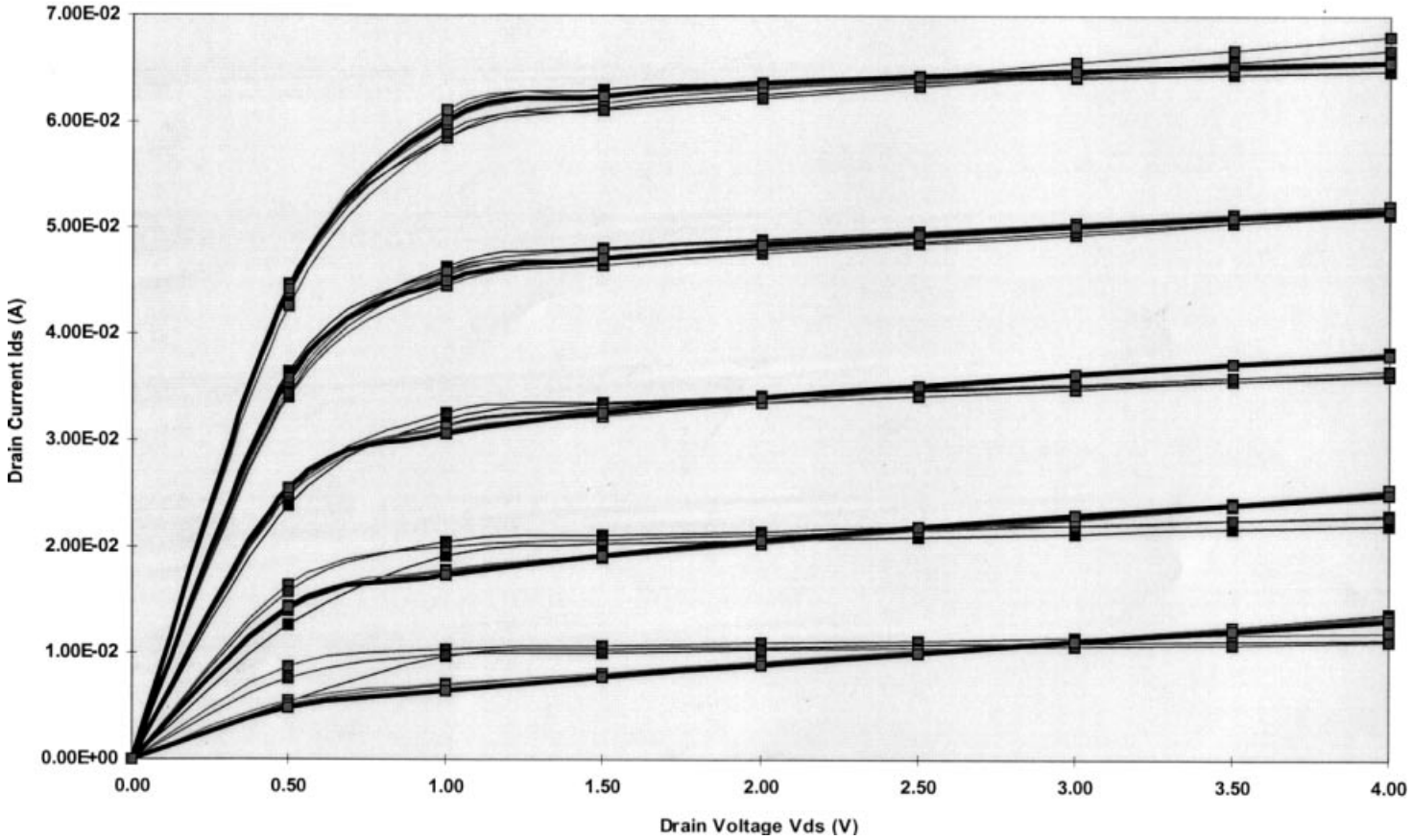


Figure 2-138 Comparison of dc models: Chalmers (Angelov), Curtice cubic, Curtice quadratic, Materka, Raytheon, and TOM (TriQuint).

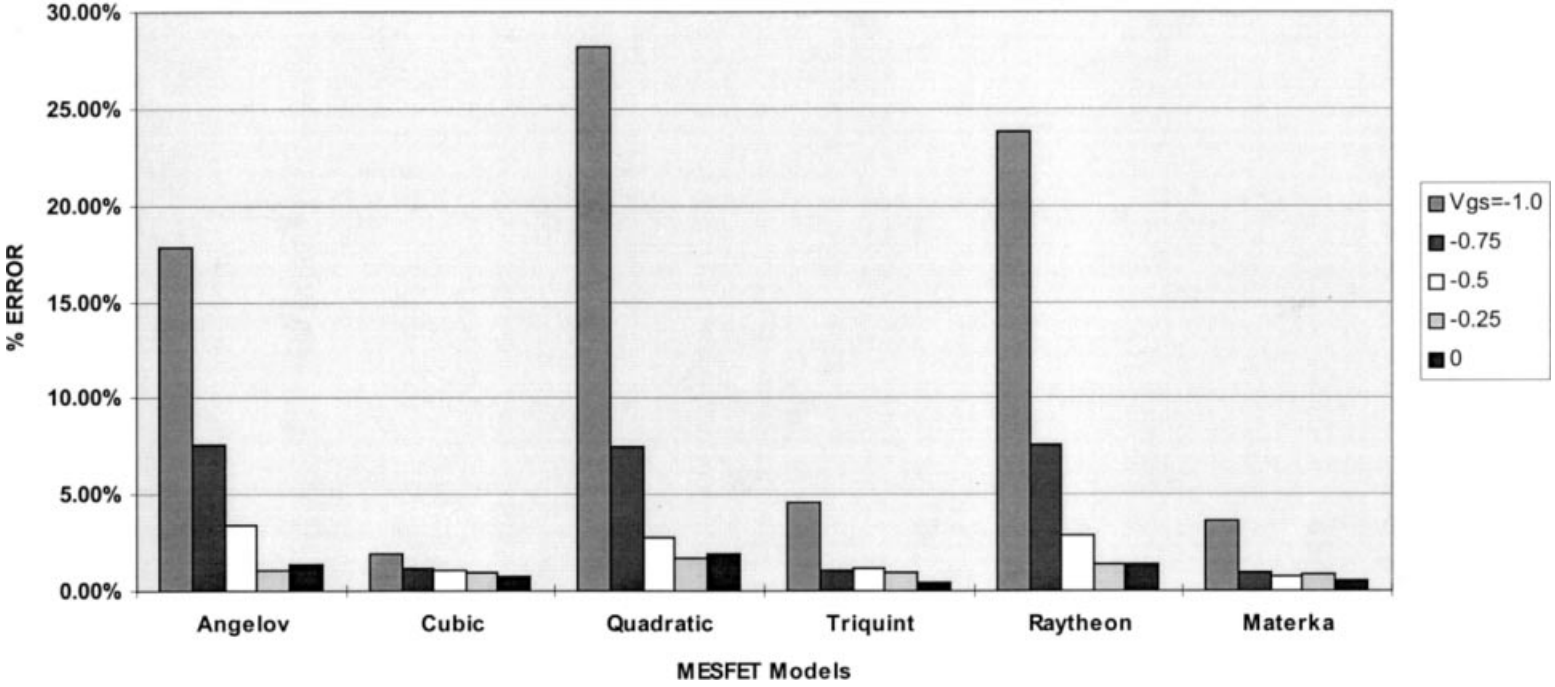


Figure 2-139 Dc *I–V* error bars: Chalmers (Angelov), Curtice cubic, Curtice quadratic, Materka, Raytheon, and TOM.

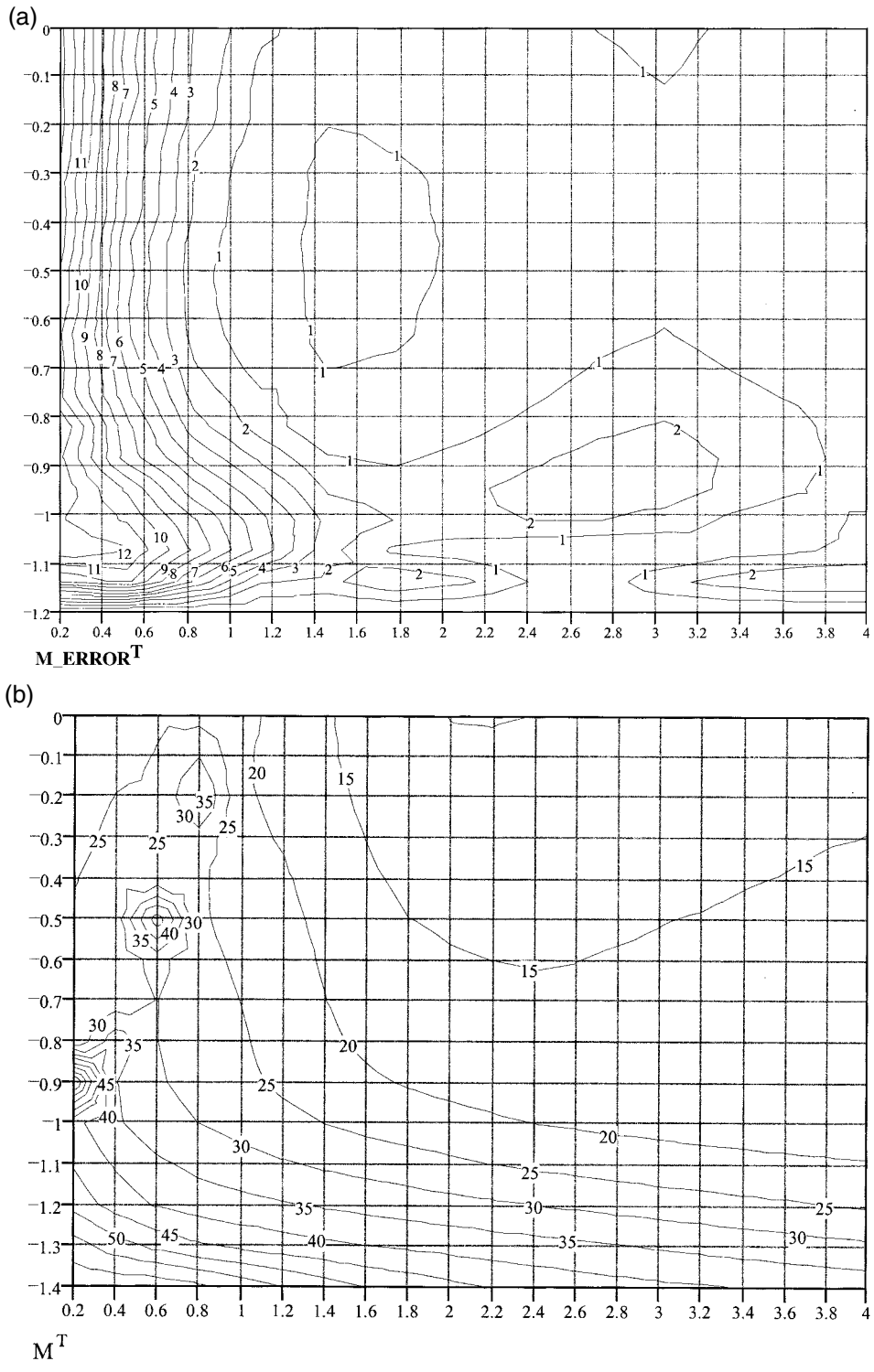


Figure 2-140 NE71000, Materka model, % error contours for (a) dc *I–V* and (b) *S* parameters.

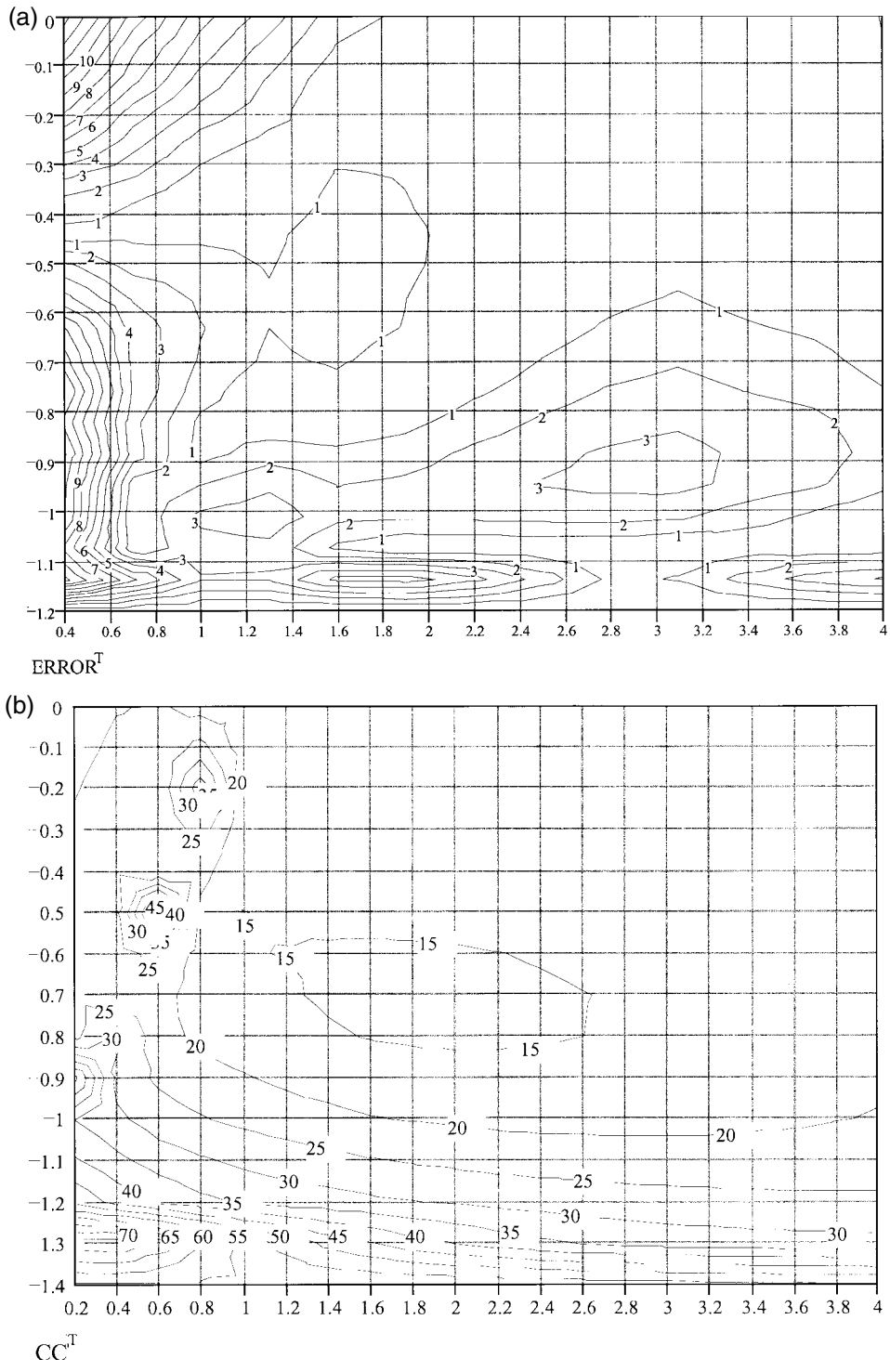


Figure 2-141 NE71000, Curtice cubic model, % error contours for (a) dc *I–V* and (b) *S* parameters.

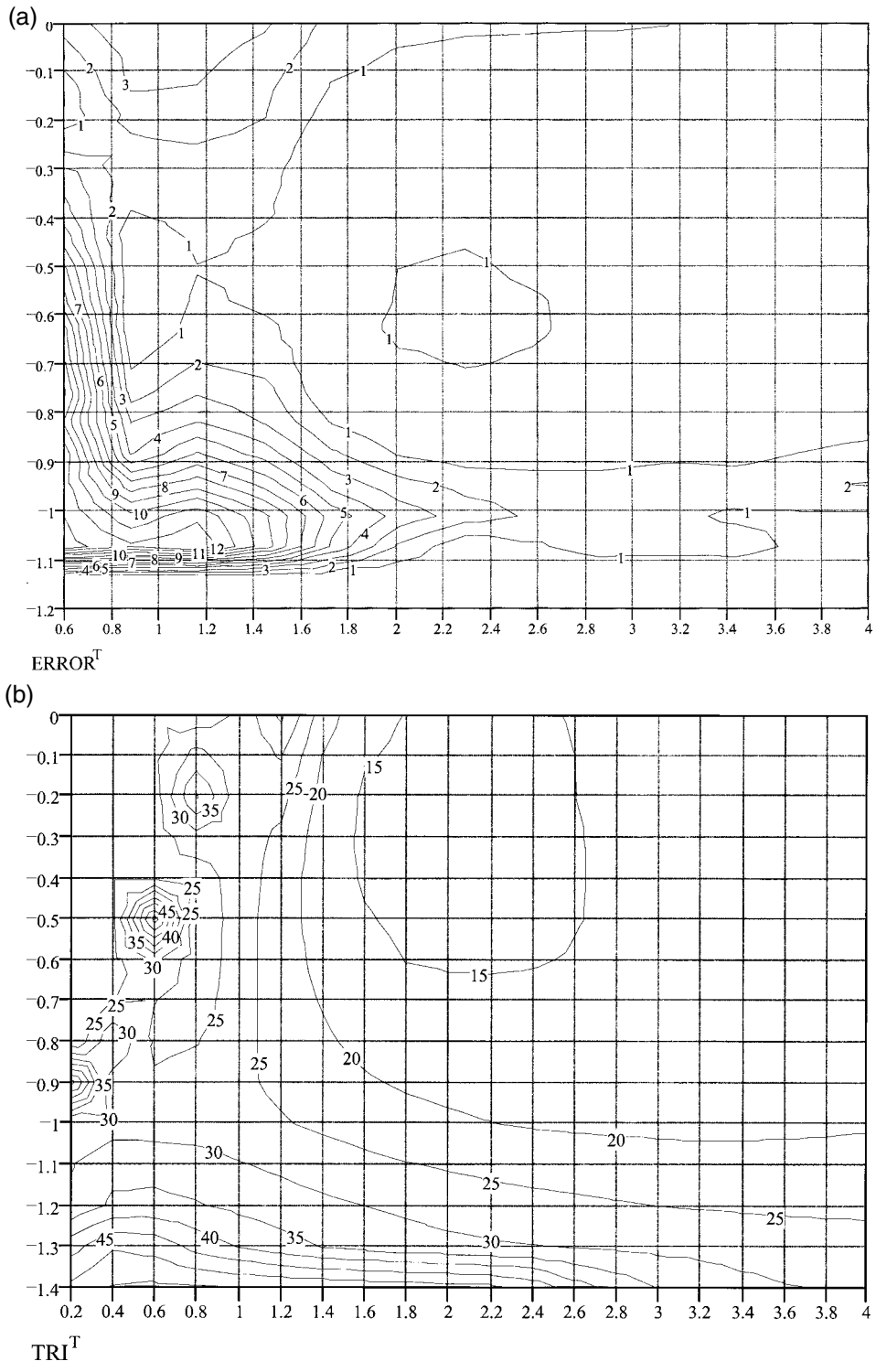


Figure 2-142 NE71000, TOM model, % error contours for (a) dc *I–V* and (b) *S* parameters.

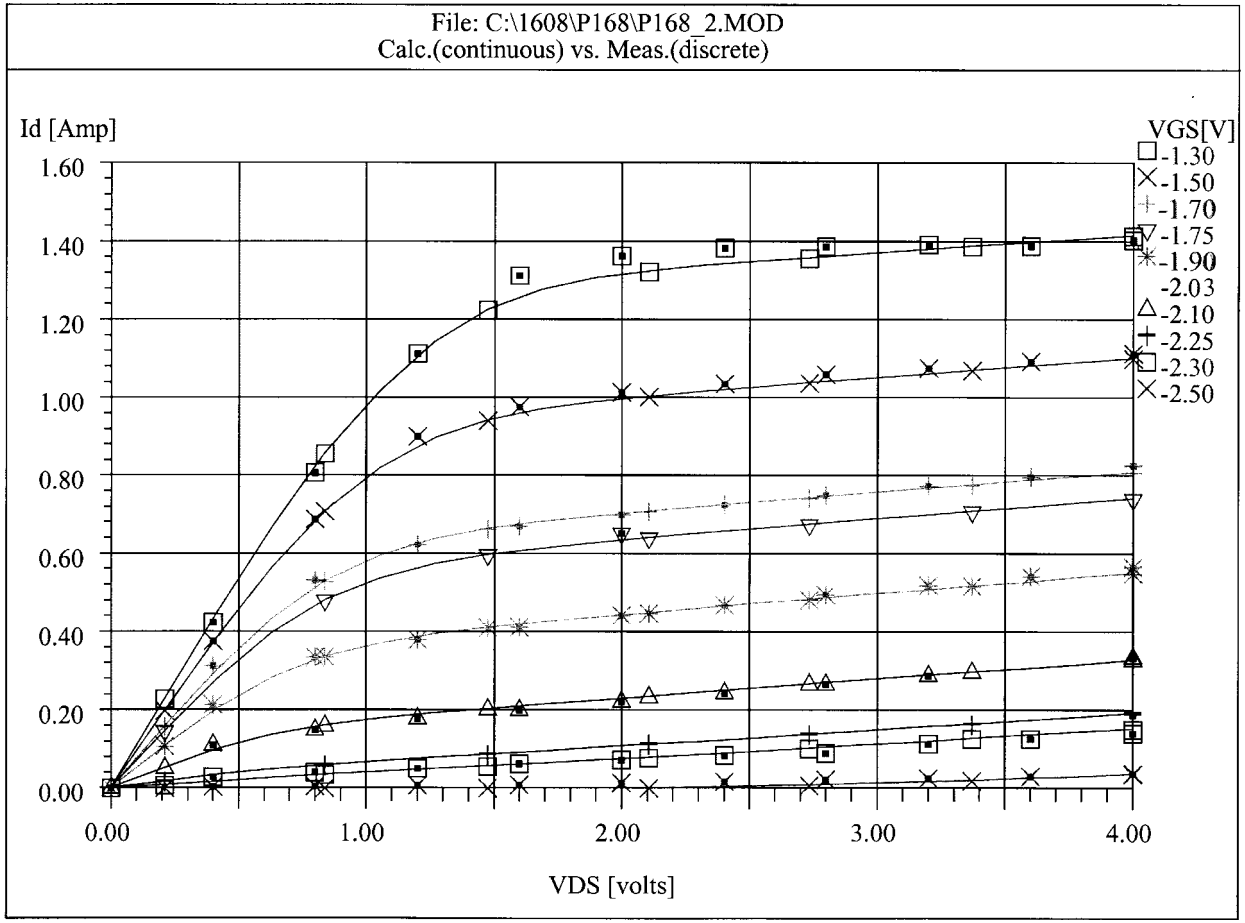


Figure 2-143 Calculated versus measured dc *I–V* curves for the Oki KGF1608 package MESFET.

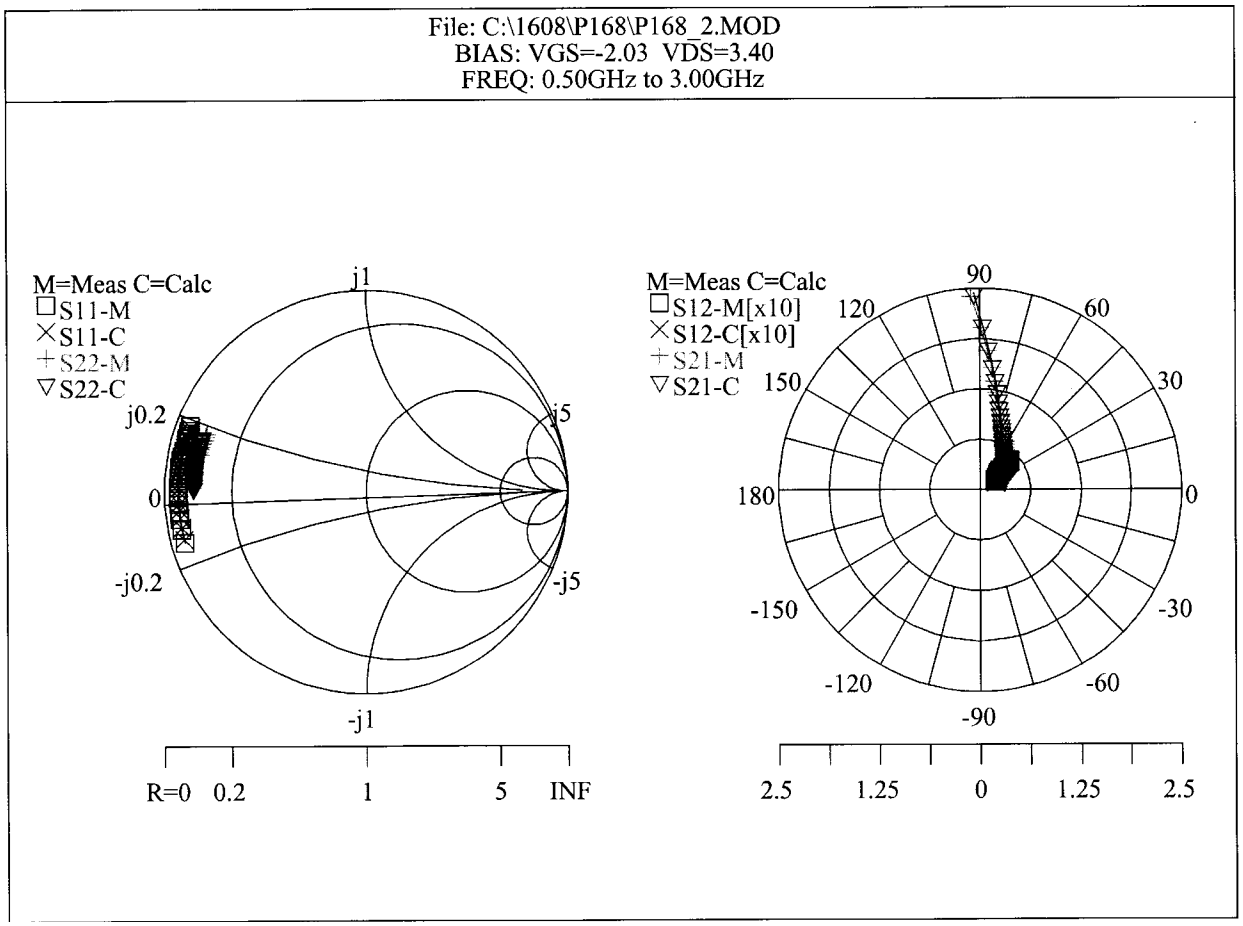


Figure 2-144 Calculated versus measured *S* parameters for the Oki KGF1608 package MESFET.

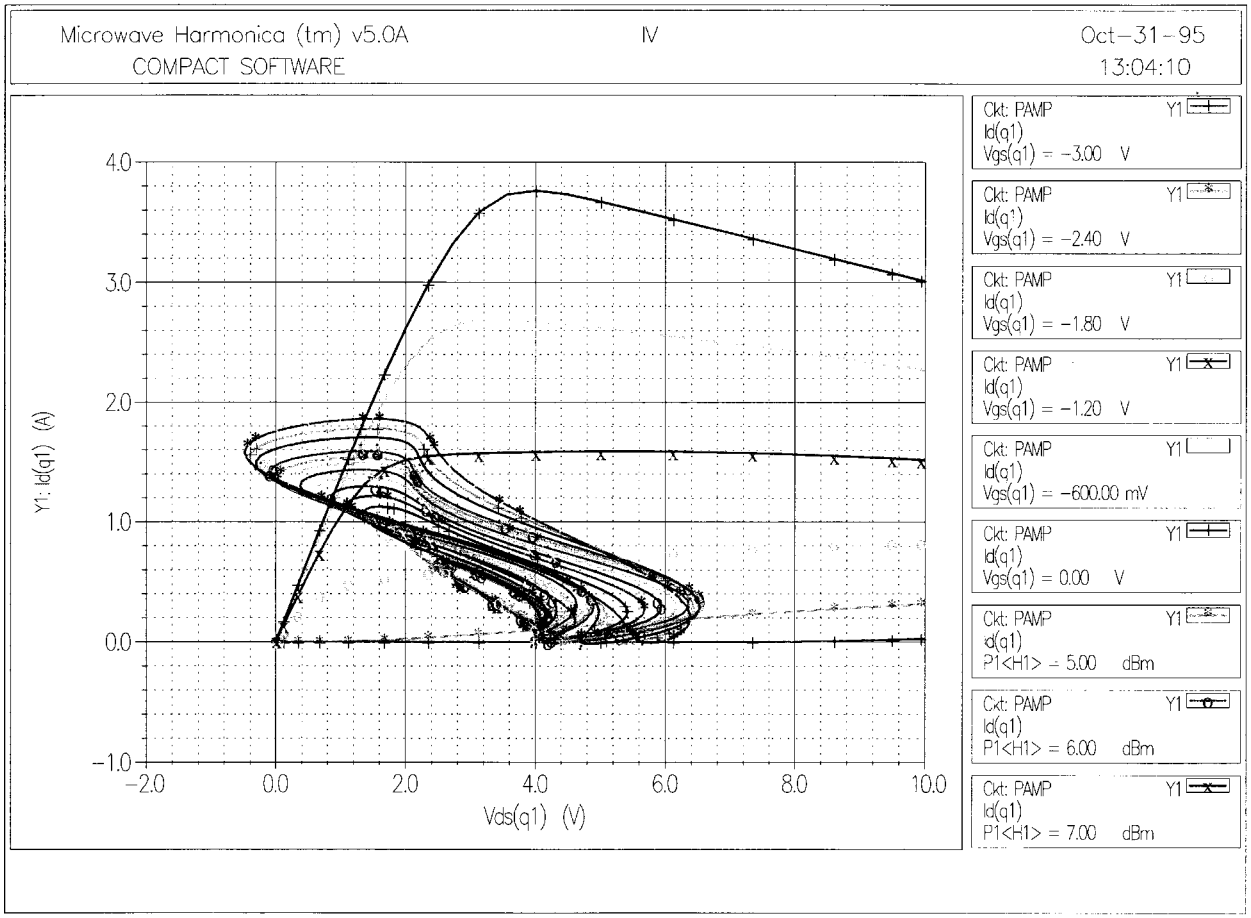


Figure 2-145 Load line and dc *I–V* curves for the KGF1608 MESFET.

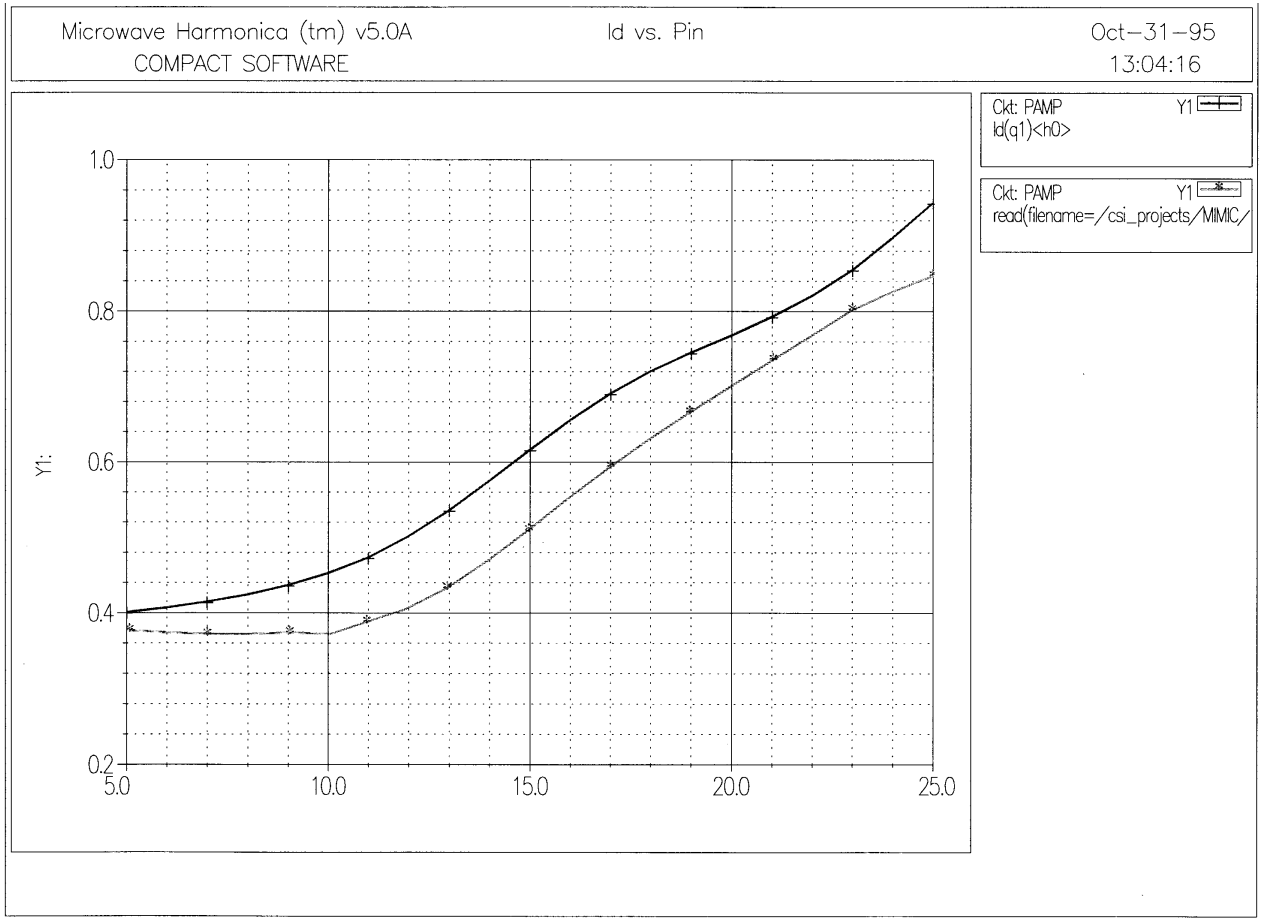


Figure 2-146 Measured versus calculated drain current for the KGF1608 MESFET.

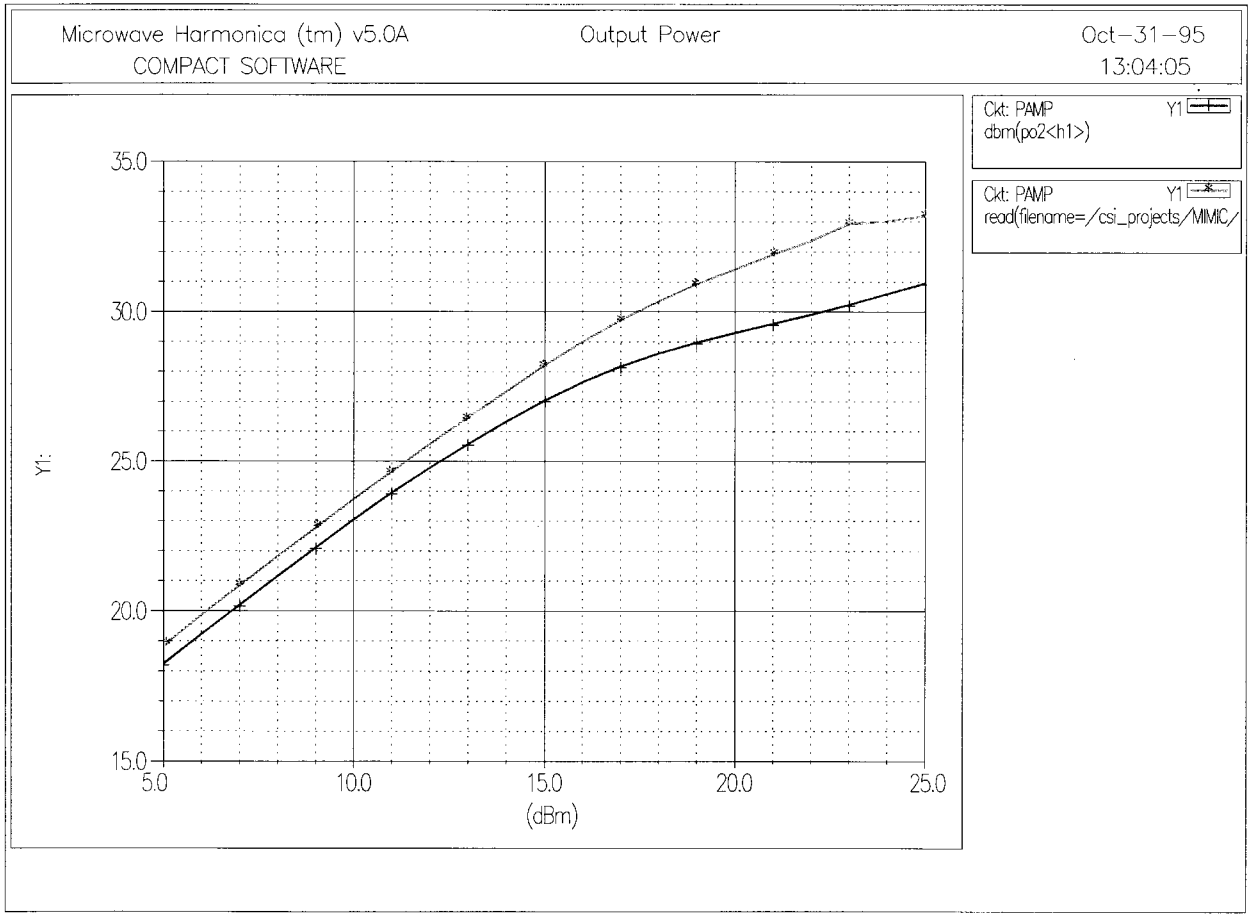


Figure 2-147 Measured versus calculated output power for the KGF1608 MESFET.

ac Errors. From the ac error contour plots we can see that the Chalmers (Angelov) model emerges as the most accurate for ac small-signal operation. The comparison also shows that all models considered are good for ac operations biased in the saturation region with a typical relative error of less than 15%. However, the ac relative errors are much higher in the linear region, particularly for the Curtice quadratic model, which exhibits more than 80% error at some points.

Notes. The above conclusions are based on the matching between the simulated responses of different models and the measured data from a typical NE71000 device; the models may behave differently for a different device. It is not possible to draw any broad-based conclusions as to what model is the best. In order to select the most suitable model for your application, you must consider all the factors, such as bias range, frequency range, and small-signal or large-signal operation, according to your experiences in device modeling.

2-4-6 Device Libraries

All these measurements lead to a nonlinear device library. Figures 2-148 and 2-149 shows a typical datasheet for a device in the library with measured and modeled data. The previous high-power example does not always lead to acceptable results as the models are far from being fully developed. We also show the simulation results using the physics-based model for FETs.

Physics-Based MESFET Modeling. We mentioned a physics-based model earlier. Figure 2-150 shows the measured versus predicted dc *I–V*, I/O power, and *S* parameters for an Alpha Industries MBE MESFET developed under a government contract. The simulation does not handle the trapped energy, which explains the negative slope on the third trace of the output dc *I–V* curve.

The process for bipolar transistors is quite similar and less tricky than the ones just described. Good examples can be found in the user manual for Compact Scout.

2-4-7 A Novel Approach for Simulation at Low Voltage and Near Pinchoff Voltage

As mentioned earlier, the most critical zone is the area of operating the FET both close to the pinchoff voltage and at low operating voltages. A crucial reason for the inadequate results is the modeling of the drain–source resistance as a function of bias point. This is particularly important in calculating the power-added efficiency. When doing your modeling, there is always a temptation to look at both high currents and operating voltages. Figure 2-151 shows a reasonable agreement between measured and predicted dc *I–V* performance. However, the very moment the transistor is modeled close to the linear region (Figure 2-152), things change drastically and the results from 0 to 6 mA have little resemblance to those from 0 to 50 mA. The modified Materka model (Figures 2-153 and 2-154) gives a significantly better match than the Angelov model.

In order to overcome the modeling inaccuracy for the FETs, Thomas Meier of Siemens AG in Munich has come up with the concept of combining two transistors in parallel, whereby the second one is not really connected for dc (Figure 2-155). Figure 2-156 compares measured and simulated R_{DS} using this new approach.

NE67300

Manufacturer : **NEC**
Model Keyword : **NE - 67300**
Device Type : **Mesfet**
Model Type : **Materka**
Package Type : **Chip**

Electrical Characteristics:
Idss : 40mA
Vp : 1.1V
Gain : 8.5dB @ 18GHz
P1db : N/A
Fmin : N/A

Model Data Range:
Vgs : -1.2V to 0.0V
Vds : 0.0V to 4.0V
Freq: 0.4GHz to 18GHz

Verification :
Type : Power sweep, 50Ω
Vgs : -0.6V Vds : 3.0V Ids : 30mA Freq: 6.0GHz

Comments : The S-parameters that appear in the figure were measured and simulated with wire bonds present. The length of each wire bond is approximately 30 mils. To obtain the S-parameters for the manufacture's data sheets, the wire bond inductances would be removed from the equivalent circuit.

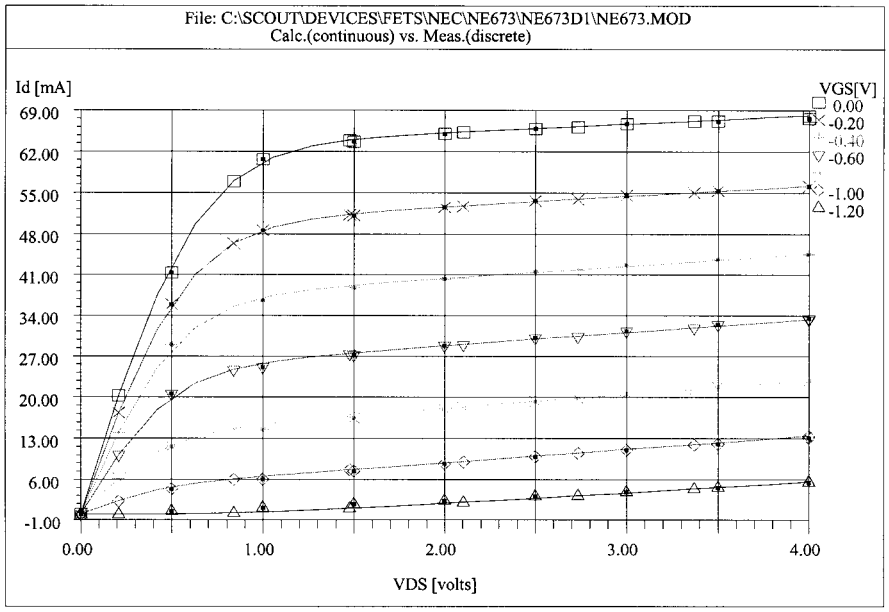
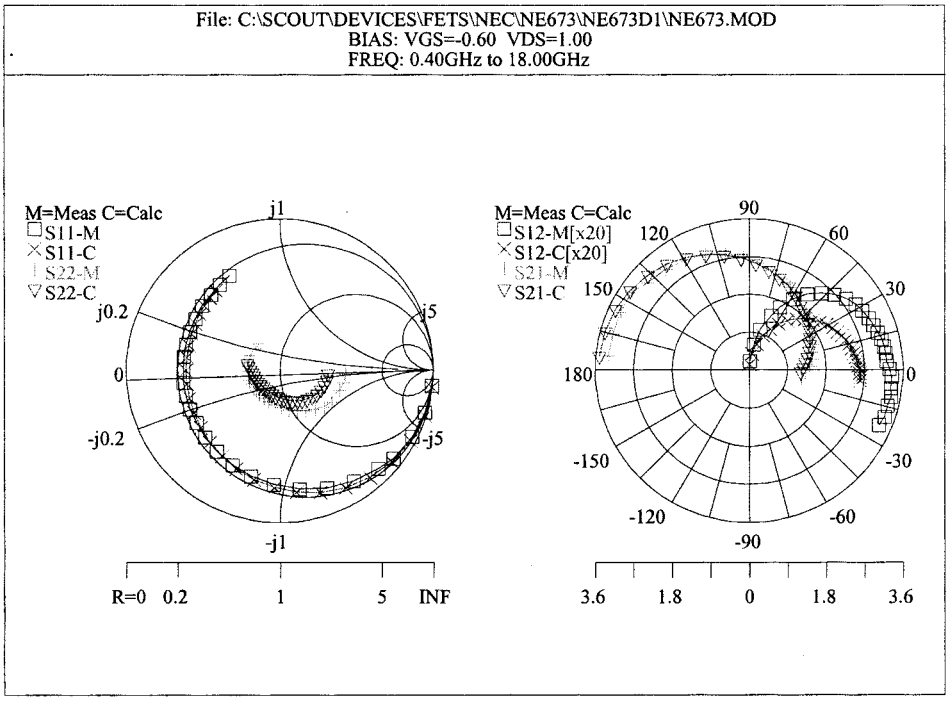


Figure 2-148 Page 1 of a nonlinear device library datasheet for the NE67300.



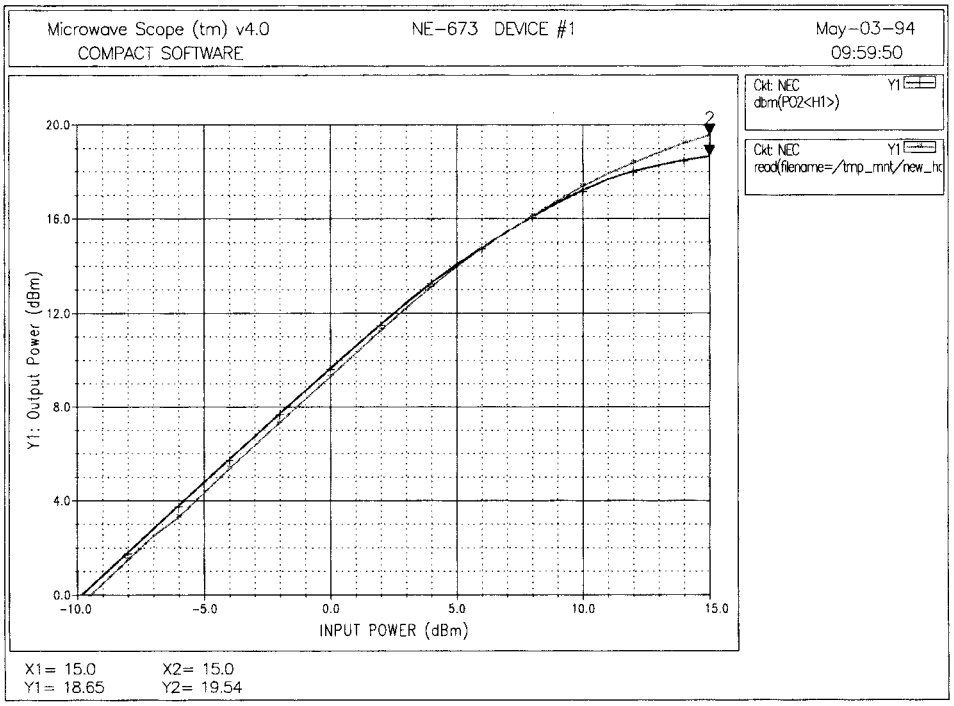


Figure 2-149 Page 2 of a nonlinear device library datasheet for the NE67300.

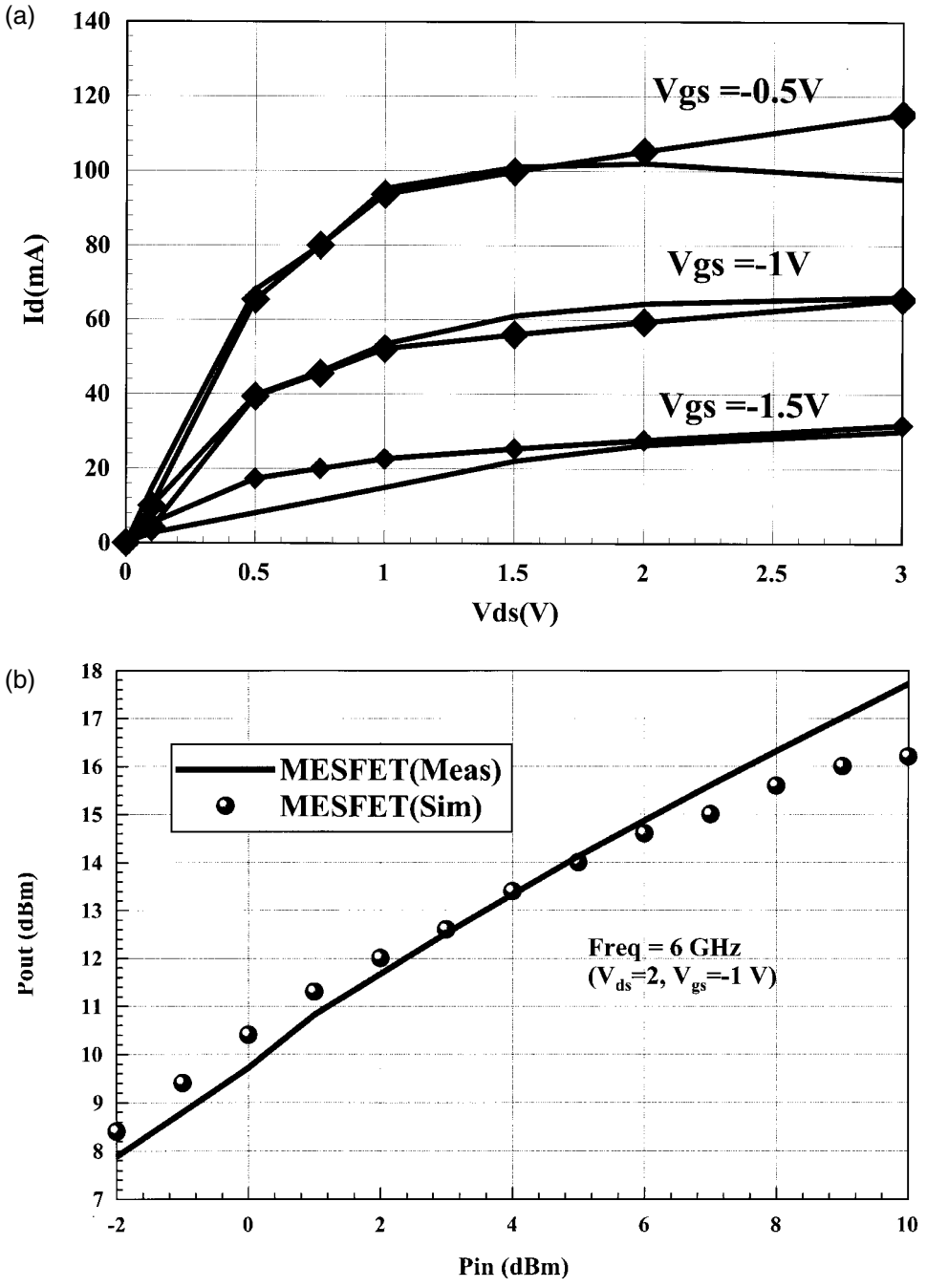


Figure 2-150 Comparison of measured data with PHYSFET and TOM models for an Alpha Industries MBE MESFET (0.25 μm × 400 μm) at $V_{GS} = -1$, $V_{DS} = 2$.

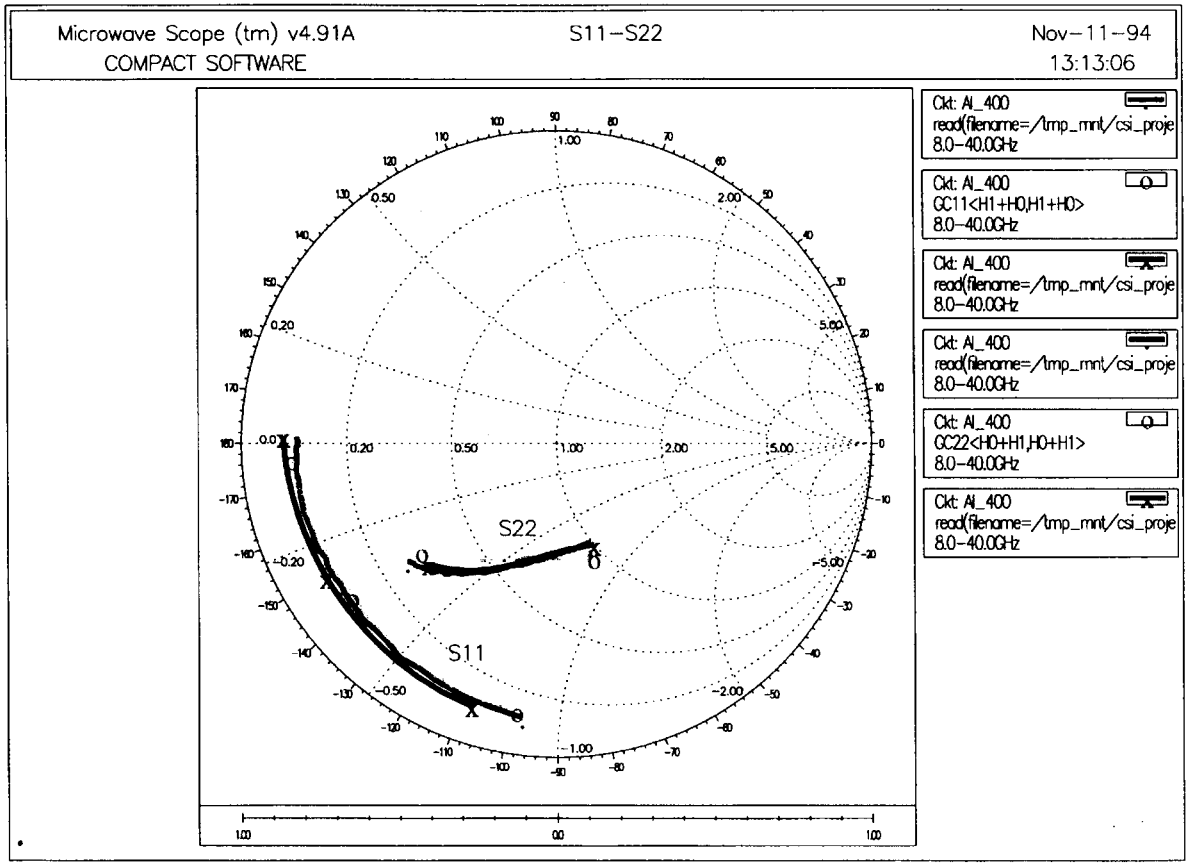


Figure 2-150 (*Continued*)

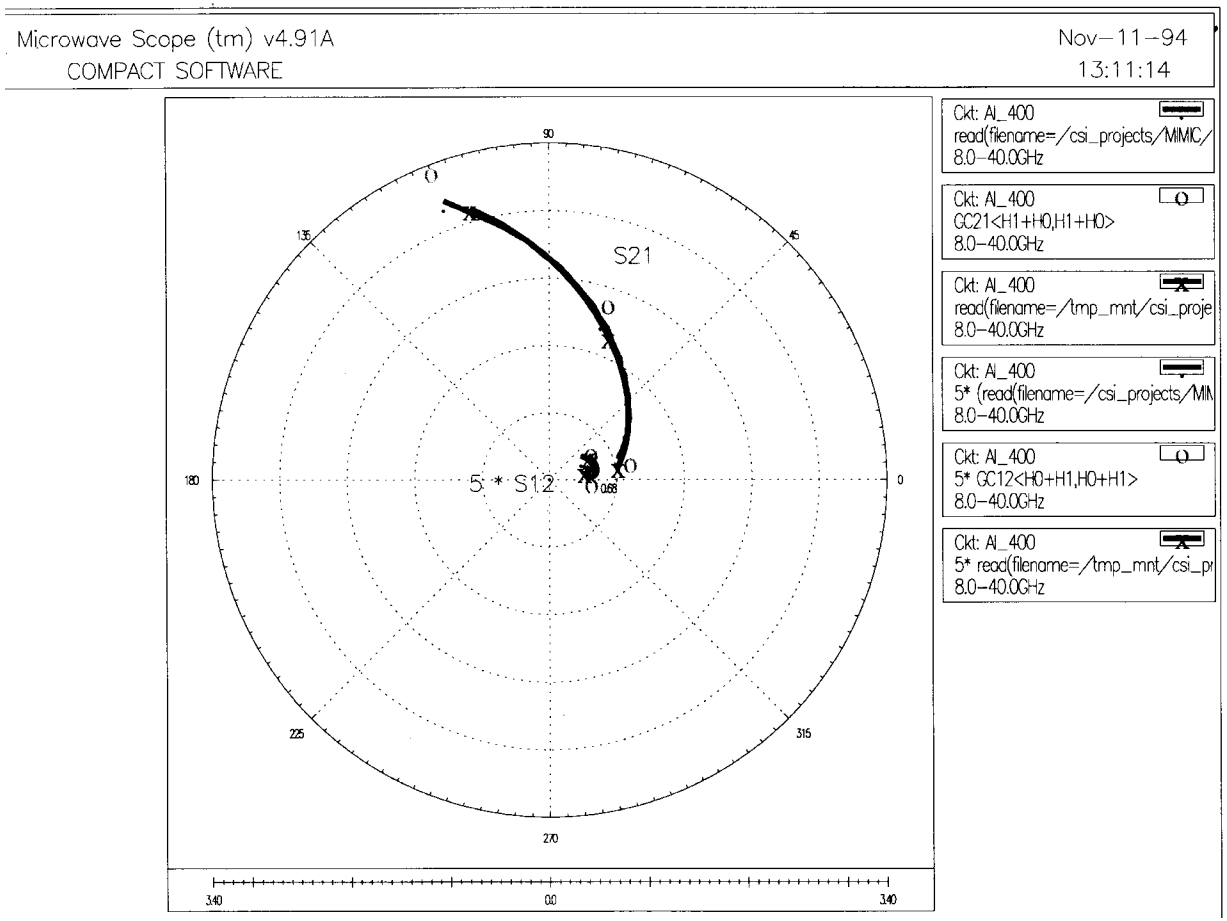


Figure 2-150 (*Continued*)

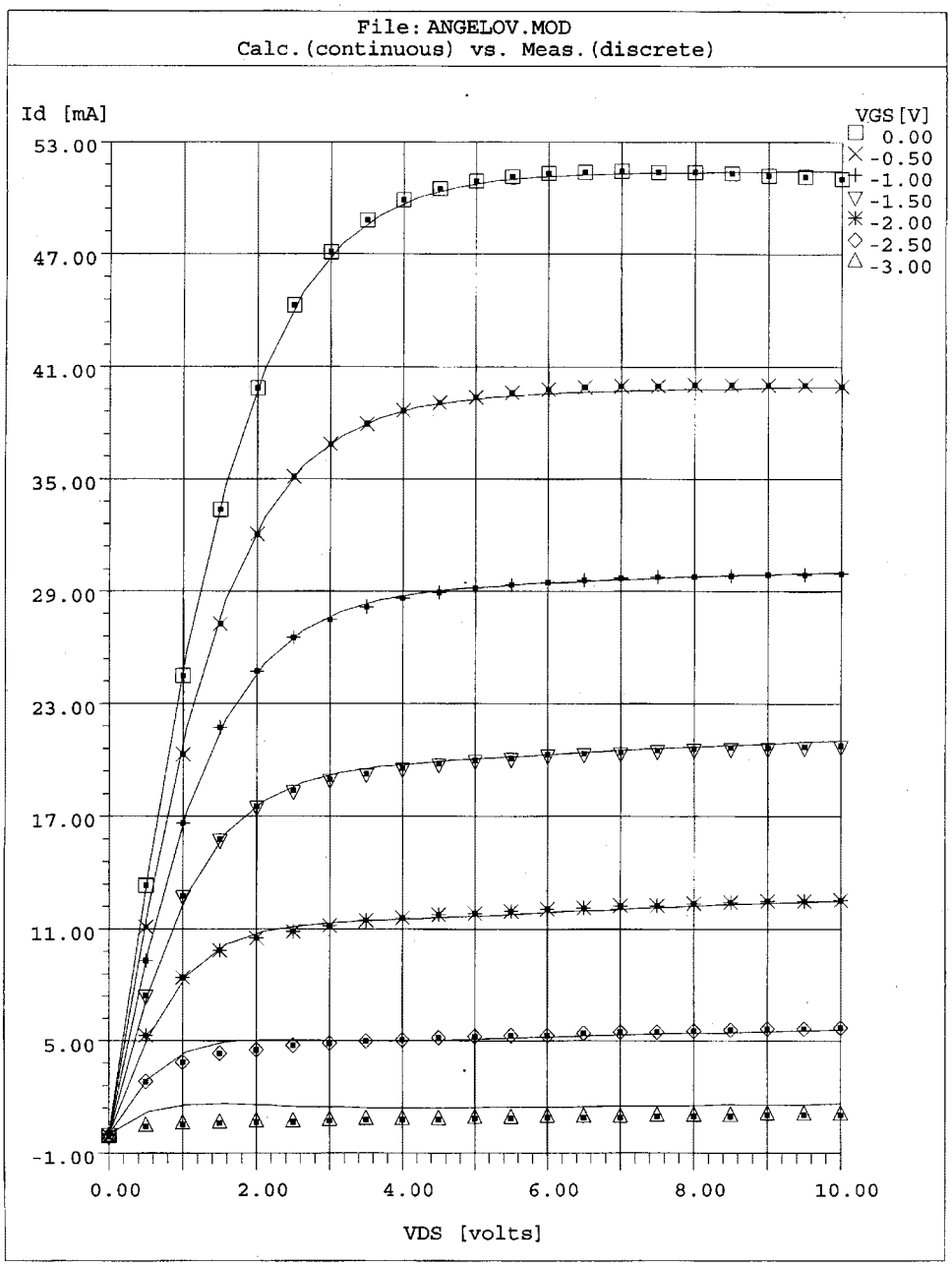


Figure 2-151 High-I_D dc I–V curves for the Angelov FET model.

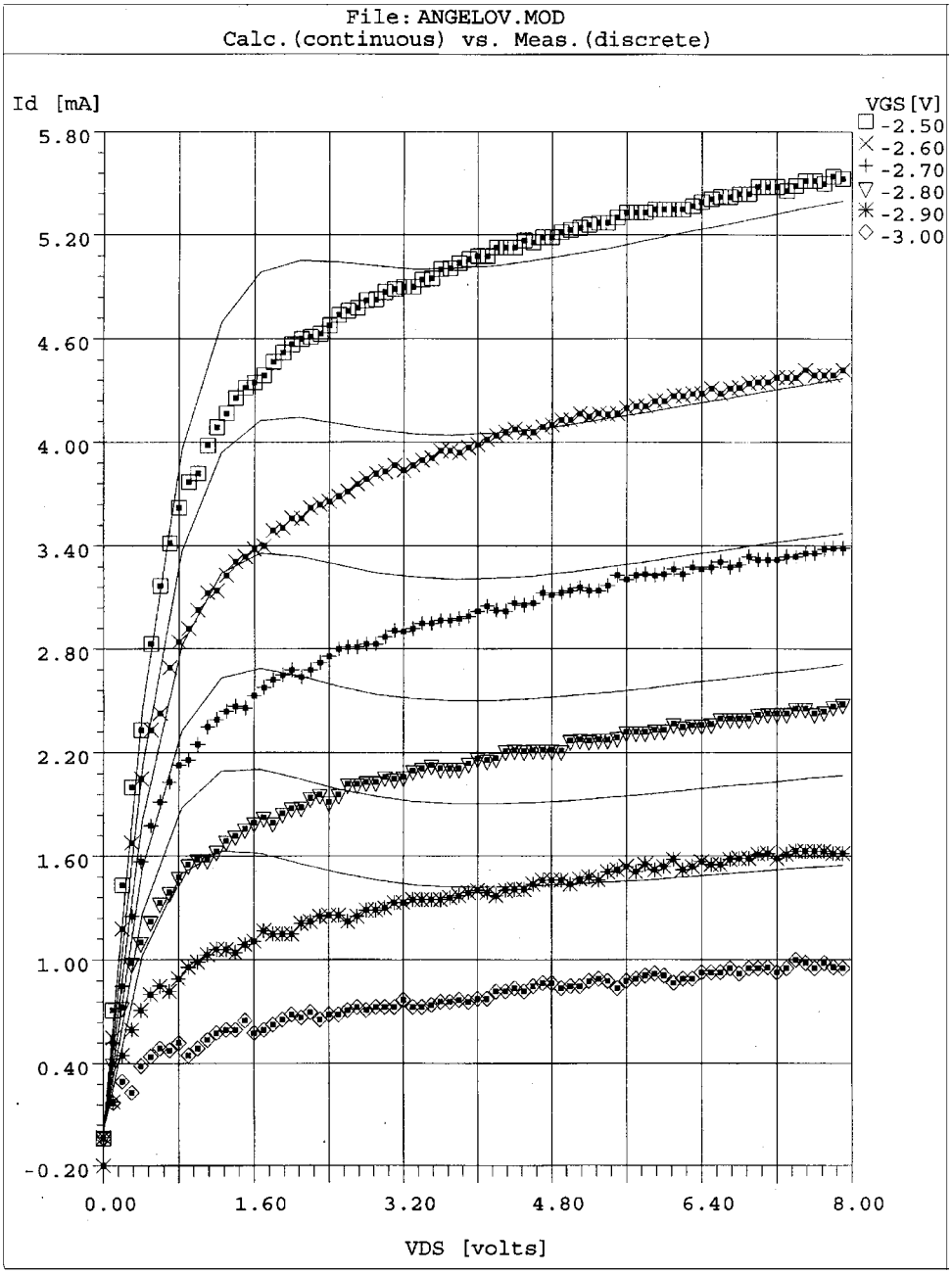


Figure 2-152 Low-I_D dc I–V curves for the Angelov FET model.

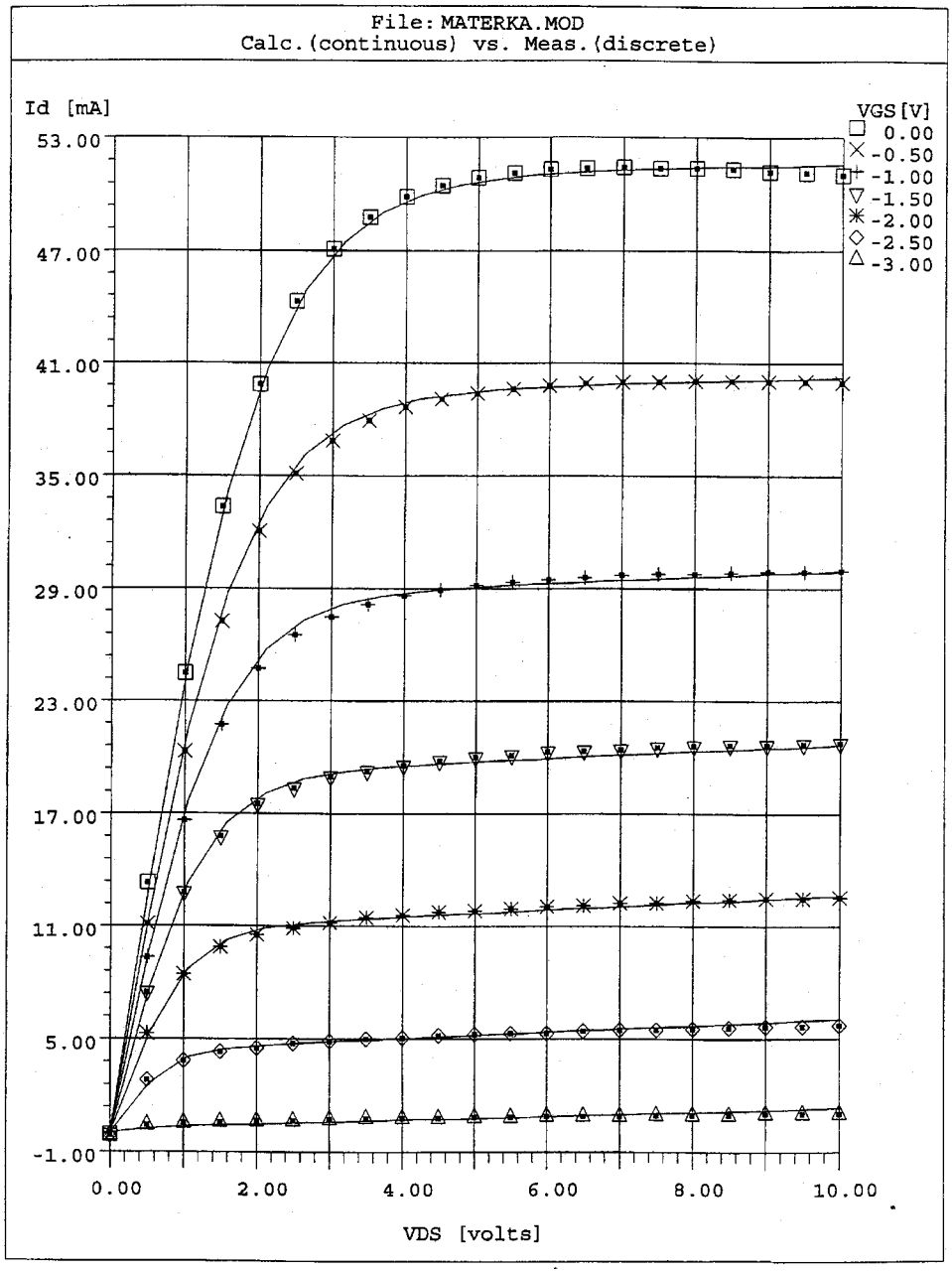


Figure 2-153 High-I_D dc I–V curves for the modified Materka FET model.

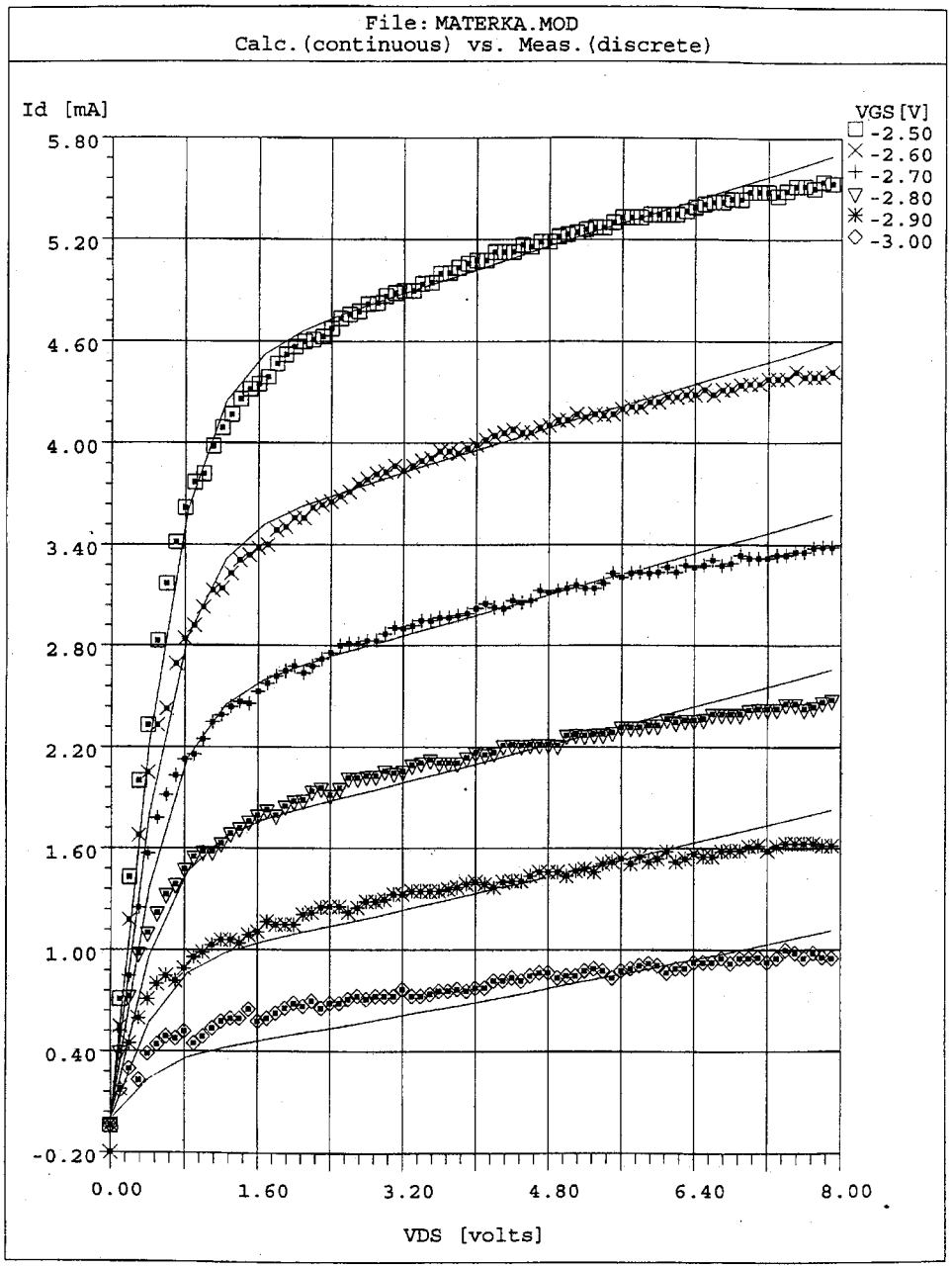


Figure 2-154 Low-I_D dc I–V curves for the modified Materka FET model.

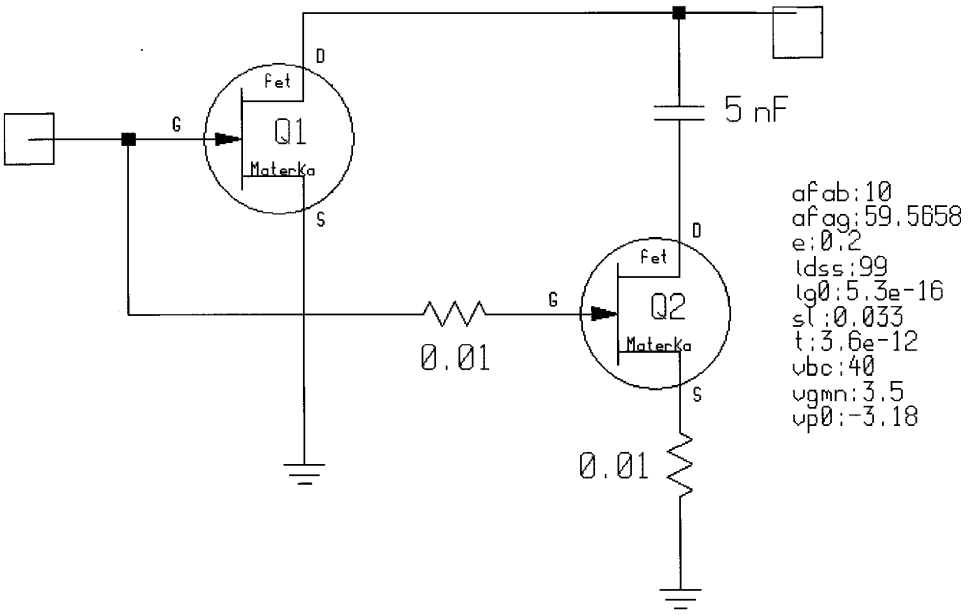


Figure 2-155 By putting two transistors in parallel, with the second one having no dc connection at the drain, one can simulate a dynamic R_{DS} quite well. Regardless of parameter values used for Q1, the parameters listed are necessary to get an accurate R_{DS} result.

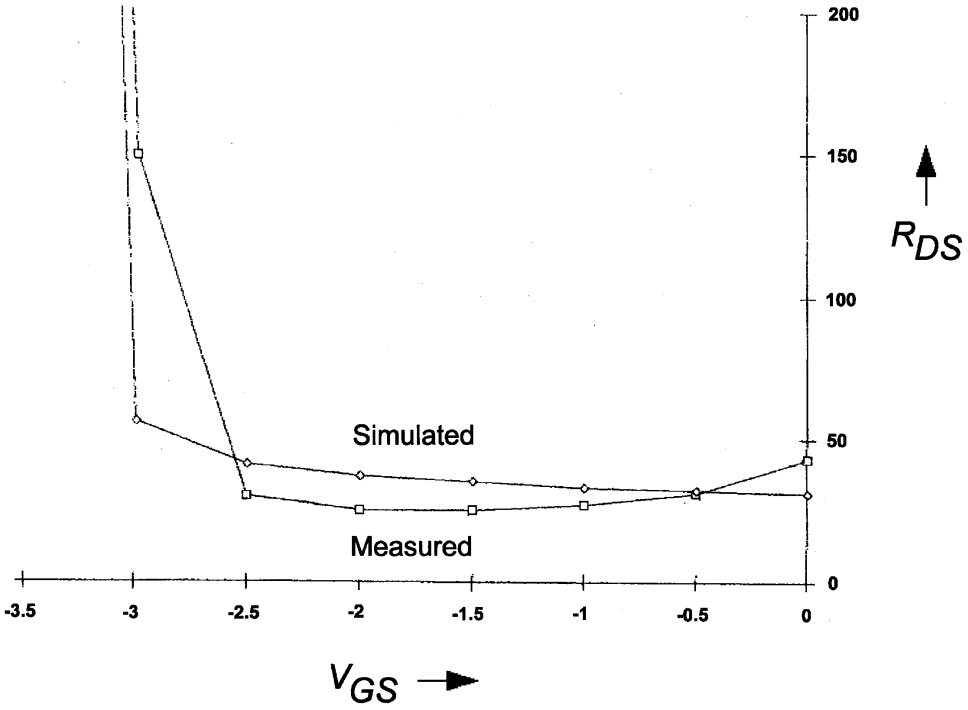


Figure 2-156 Measured versus simulated R_{DS} for the paralleled MESFET approach.

One might ask the question: "Why not fix the existing models?" It turns out that in order to get the best possible convergence and speed, one has to use analytic derivatives for voltage and current; at least the first and second derivatives, if not the third, are required. The programming of such a model becomes so difficult, however, and the calculation of its derivative is so highly likely to have subtle errors, that it turns out that the parallel configuration shown in Figure 2-155 is the best possible compromise for fast computation and accurate results.

A similar approach for bipolar transistors has been discussed by Philips with their MEXTRAM model. However, we have not seen any parameter extraction for this model and also we have not seen any industry acceptance for it. While, in principle, it will cover all derivatives of all members of the BJT family, its topology consists out of three discrete transistors connected and we doubt very much that a reliable parameter extraction program/mathematically continuous model for nonlinear simulation is supplied by the university that developed this jointly with Philips. At the time of writing this book, Philips has not yet provided us with the code for both, and until this happens, we are not in a position to qualify this any further.

2-4-8 Example: Improving the BFR193W Model

We have already mentioned the difficulty encountering the proper base-spreading resistor values and others by using standard parameter extraction programs. Table 2-28 is a copy from the Siemens semiconductor book showing Gummel–Poon Berkeley SPICE syntax parameters for the BFR193W BJT.

The first thing that will happen is that the observant reader will notice that the minimum possible base-spreading resistor, RBM (valid at large current densities), is larger than the actual base-spreading resistor, RB, "claimed to be measured" at 1 Ω (see Table 2-28).

Another headache is caused by the measured noise performance of the same device at an operating point of 10 mA. Using the above-mentioned simplified equation

Table 2.28. BFR193W transistor chip data: SPICE parameters (Gummel–Poon model, Berkeley–SPICE 2G.6 syntax)

IS	=	0.2738	fA	BF	=	125	—	NF	=	0.95341	—
VAF	=	24	V	IKF	=	0.26949	A	ISE	=	10.627	fA
NE	=	1.935	—	BR	=	14.267	—	NR	=	1.4289	—
VAR	=	3.8742	V	IKR	=	0.037925	A	ISC	=	0.037409	fA
NC	=	0.94371	—	RB	=	1	Ω	IRB	=	0.91763	mA
RBM	=	1.8368	Ω	RE	=	0.76534	Ω	RC	=	0.11938	Ω
CJE	=	1.1824	fF	VJE	=	0.70276	V	MJE	=	0.48654	—
TF	=	18.828	ps	XTF	=	0.69477	—	VTF	=	0.8	V
ITF	=	0.96893	mA	PTF	=	0	deg	CJC	=	935.03	fF
VJC	=	1.1828	V	MJC	=	0.30002	—	XCJC	=	0.053563	—
TR	=	1.0037	ns	CJS	=	1	fF	VJS	=	0.75	V
MJS	=	0	—	XTB	=	0	—	EG	=	1.1	eV
XTI	=	3	—	FC	=	0.72063	—	TNOM	=	300	K

All parameters are ready to use, no scaling is necessary. Extracted on behalf of SIEMENS Small Signal Semiconductors by Institut für Mobil-und Satellitenfunktechnik (IMST).

$$F = 1 + \frac{R'_{bb}}{R_g} + \frac{0.5(26\ \text{mV}/I_C)}{R_g}$$

(approximate equation at medium frequencies) (2-288)

we obtain the following results:

1. Inserting 1 Ω (value for RB provided by IMST) and R_B = 15 Ω (measured by the authors), here are the interesting results:

$$F = 1 + \frac{1}{50} + \frac{2.6}{100} = 1.046 \tag{2-289}$$

as noise factor or, as noise figure, F (dB) = 0.2 dB (IMST).

2. Or using the measured 15 Ω, one obtains

$$f = 1 + \frac{15}{50} + \frac{2.6}{100} = 1.326 \tag{2-290}$$

as noise factor or, as noise figure, F (dB) = 1.22.

The Siemens datasheet claims a noise figure of 1.3 dB, and therefore the measured versus calculated data differ by 0.08 dB. Since our noise figure was calculated well below f_T, which is seen from the fact that there are no frequency-dependent terms in the noise figure calculation, it is not unusual that the 900-MHz noise figure is 1.3 dB or slightly higher. This simple example shows how easy it is to provide measurement data that are inconsistent with the laws of physics. It is therefore highly recommended to add the noise figure as an additional parameter in the parameter extraction to reduce the number of variables—specifically, those for which the standard extraction has little sensitivity.

The collector–emitter junction capacitance is another important parameter that frequently gets measured incorrectly. The way around this is to remember the definition

$$f_{\max} = \sqrt{\frac{f_T}{8\pi R'_{bb} C_c}} \tag{2-291}$$

f_T can be measured from the 3-dB point of the emitter current gain (β) and R'_{bb} was just computed above, so the equation can be solved for C_c, as there is a relationship between measured $g_{\max}$ and $f_{\max}$. There are several correlations between the parameters of the equivalent circuit as published in different places, and therefore one can restrict the degree of freedom within the optimization process and obtain much better results. This approach is applicable to all types of transistors, bipolar and FET, and can be used to improve the accuracy of parameters that not only affect a transistor's RF performance but also its noise performance.

REFERENCES

1. H. A. Watson, *Microwave Semiconductor Devices and Their Circuit Applications* . New York: McGraw-Hill, 1969, Chap. 2.
2. A. S. Grove, *op. cit.,* Chap. 4.
3. A. S. Grove, *op. cit.,* Chap. 7.
4. P. E. Gray, D. DeWitt, A. R. Boothroyd, and J. G. Gibbons, *Physical Electronics and Circuit Models of Transistors* . New York: Wiley, 1964, p. 10.
5. P. E. Gray et al., *op. cit.,* p. 129.
6. P. E. Gray et al., *op. cit.,* p. 180.
7. Charles A. Liechti, "Microwave Field-Effect Transistors—1976," *IEEE Trans. Microwave Theory Tech.* , **MTT-24**, 279–300, June 1976.
8. Edward A. Wolff and Roger Kaul, *Microwave Engineering and Systems Applications* . New York: Wiley, 1988, p. 395.
9. Arthur D. Evans, ed., *Designing with Field-Effect Transistors* . New York: McGraw-Hill, 1981.
10. R. D. Middlebrook, "A Simple Derivation of Field-Effect Transistor Characteristics," *Proc. IEEE*, **50**, 1146–1147, 1963.
11. Tzu-Hwa and Craig P. Snapp, "Low Noise Microwave Bipolar Transistor with Sub-Half Micrometer Emitter Width," *IEEE Trans. Electron Devices* , **ED-25**, 723–730, June 1978.
12. Ulrich L. Rohde, "Designing a Match Low Noise Amplifier Using CAD Tools," *Microwave J.* , **20**(10) 154–160, October 1986.
13. R. A. Pucel, W. Strubel, R. Hallgren, and U. L. Rohde, "A General Noise De-embedding Procedure for Packaged Two-Port Linear Active Devices," *IEEE Trans. Microwave Theory Tech.* , **40**(11), 2013–2024, November 1992.
14. R. A. Pucel and U. L. Rohde, "An Exact Expression for Noise Resistance R_n of a Bipolar Transistor for the Hawkins Noise Model," *IEEE Microwave and Guided Wave Letters* , **3**(2), 35–37, February 1993.
15. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits.* New York: Wiley, 1986, p. 386.
16. D. Frohman-Bentchkowsky and A. S. Grove, "Conductance of MOS Transistors in Saturation," *IEEE Trans. Electron Devices* , **ED-16**, 108–113, January 1969.
17. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits* . New York: Wiley, 1986, p. 480.
18. "New LDMOST Technology Superb for RF Applications," *Philips Semiconductors World News,* Vol. 7, Issue 3, Article No. 29, http://www-us2.semiconductors.philips.com/publications/content/file_344.html.
19. Y. P. Tsividis, *Operation and Modeling of the MOS Transistor* . New York: McGraw-Hill, 1987, p. 136.
20. A. Materka and T. Kacprzak, "Computer Calculation of Large-Signal GaAs FET Amplifier Characteristics," *IEEE Trans. Microwave Theory Tech.* , **MTT-33**(2), 129–135, February 1985.
21. R. E. Anholt and S. E. Swirhun, "Experimental Investigation of the Temperature Dependence of GaAs FET Equivalent Circuits," *IEEE Trans. Electron Devices* , **ED-39**(9), 2029–2036, September 1992.
22. Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits* . New York: Wiley, 1986.
23. Paul W. Tuinenga, *SPICE: A Guide to Circuit Simulation and Analysis Using PSpice ®*. Englewood Cliffs, NJ: Prentice Hall, 1992.
24. Andrei Vladimirescu, *The SPICE Book* . New York: Wiley, 1994.

25. H. Statz, H. Haus, and R. A. Pucel, "Noise Characteristics of Gallium Arsenide Field-Effect Transistors," *IEEE Trans. Electron Devices,* **ED-21**, 549–562, September 1974.

26. V. Rizzoli, F. Mastri, and C. Cecchetti, "Computer-Aided Noise Analysis of MESFET and HEMT Mixers," *IEEE Trans. Microwave Theory Tech.* , **MTT-37**, 1401–1410, September 1989.

27. V. Rizzoli, F. Mastri, and D. Masotti, "General-Purpose Noise Analysis of Forced Nonlinear Microwave Circuits," *Military Microwave*, 1992.

28. Ulrich L. Rohde and Chao-Ren Chang, "Parameter Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits," *Microwave J.,* **36**(5), 222–239, May 1993.

29. Ulrich L. Rohde, "Parameters Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits Like Mixers and Oscillators," *23rd European Microwave Conference,* Madrid, Spain, September 6–9, 1993.

30. Ulrich L. Rohde, Chao-Ren Chang, and J. Gerber, "Design and Optimization of Low-Noise Oscillators Using Non-linear CAD Tools," presented at the *Frequency Control Symposium,* Boston, June 1–2, 1994.

31. Ulrich L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design* . New York: Wiley, 1997, Appendix B-3, "Bias Independent Noise Model."

FURTHER READING

George Anderson, "Moto Aims FETs at UHF," *Electron. Eng. Times*, 96, September 2, 1996.

Jean–Jacques Bouny, "Advantages of LDMOS in High Power Linear Amplification," *Microwave Eng. Eur.*, 37–40, April 1996.

Noël Boutin, "Develop SPICE High Frequency Small Signal Transistor Models from Data Sheets," *Appl. Microwave & Wireless*, 30–36, April 1999.

V. Danelon, F. Aniel, J. Benchimol, J. Mba, M. Riet, P. Crozat, G. Vernet, and R. Adde, "Parameters of InP-Based Double-Heterojunction Base Collector Self-Aligned Bipolar Transistors," *IEEE Microwave Guided Wave Lett.*

Suman Datta, Shen Shi, Kenneth P. Roenker, Marc M. Cahay, and William Stanchina, "Simulationand Design of InAlAs/InGaAs *pnp* Heterojunction Bipolar Transistors," *IEEE Trans. Electron Devices*, **ED-45**(8), 1634–1643, August 1998.

Russell J. Elias, Gordon C. Ma, and Laurence S. Golonka, "Statistical Optimization High Frequency LDMOS Devices Via Hyper-Fractionated Experimental Designs," *1996 IEEE/CPMT International Electronics Manufacturing Technology Symposium,* October 14–16, 1996, pp. 462–466.

Al Franceschino, "LDMOS Devices Enhance Cellular Amplifier Design," *Microwaves & RF*, **35**(10), 67–68, 70, 72, 74, October 1996.

GaAs IC Symposium Technical Digest 1995 , IEEE Gallium Arsenide Integrated Circuit Symposium, San Diego, CA, October 29 to November 1, 1995.

Hewlett-Packard, *Communications Components Designer's Catalog* , Hewlett-Packard, Jan Jose, CA, 1997.

ITT Capacitance Diodes, Tuner Diodes, Diode Switch, PIN Diodes, Basics and Applications , International Telephone and Telegraph Corporation, 1976.

Aron Kain, *Final Report for Bias Dependence Noise Modeling of Heterojunction Bipolar Transistors* , USAF SBIR Phase II (PIIN), F33615-95-C-1707, November 1997. Issued by USAF/AFMC/ASC, Wright Laboratory WL/AAKE BLD 7, 2530 C ST, Wright-Patterson AFB, OH 45433-7607.

X. McDade and X. Schiavone, "Switching Time Performance of Microwave PIN Diodes," *Microwave J.*, **17**, September 1974.

Danny Molezion, "A 1 watt Power Amplifier Design for 800–900 MHz Applications Using Discrete LDMOS FETs," *Wireless Technology '96 Conference & Exposition Conference Proceedings,* October 7–11, 1996, pp. 260–267.

Nadir Rahman, "A Survey of Existing Literature on Reverse Modeling of Field Effect Transistors," http://www-mtl.mit.edu/CAPAM/memos/95-3/95-3.html.

Jin-Ho Shin, Jiyoung Kim, Yuji Chugn, Joonwoo Lee, Youngsuk Suh, Kyu Hwan Ahn, and Bumman Kim, "Low-Frequency Noise Characterization of Self-Aligned AlGaAs/GaAs Heterojunction Bipolar Transistors with a Noise Corner Frequency Below 3 kHz," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, pp. 1604–1613, November 1998.

K. Y. Toh, P. K. Ko, and R. G. Meyer, "An Engineering Model for Short-Channel MOS Devices," *IEEE J. Solid-State Circuits*, **23**, 950–958, August 1988.

Robert J. Zavrel, Jr., "Silicon–Germanium and the World of RF Design," *RF Des.*, 46–50, June 1997.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

2

MODELS FOR ACTIVE DEVICES

Because we are dealing with amplifiers, both small-signal and power, oscillators, and active mixers, active devices will play a major role in circuit design. At our frequency range of interest (about 1000 MHz and above), distributed elements predominate, and even in device modeling we must consider some distributed effects such as packaging.

As we dive into providing functional explanations of the active devices used in practical circuits, there is some temptation not to provide enough information on semiconductor behavior. Because manufacturers of microwave products will use the services of a foundry or specify semiconductors from their suppliers, we will cover the device operation in depth. That said, for the beginner looking into designing circuits from high frequencies to low microwave frequencies, we will provide examples of a variety of devices unique to this frequency range. Some treatment of device physics will therefore be useful in providing a basic understanding of the nonlinear behavior of active devices and modeling in preparing for practical circuits using a computer-aided-design (CAD) tool. For CAD applications, we assume the user will have access to nonlinear CAD software, such as a high-performance SPICE program or a harmonic-balance program, such as Ansoft Serenade 8.0 or its derivatives. It is interesting that most semiconductor houses today still only offer very limited nonlinear data sets, while in the linear area a large body of useful data is made available, mostly on CDs. What CAD does for us—and this applies to SPICE and harmonic-balance programs alike—is predict the dc bias point of a device in the context of its surrounding circuitry. Configuring the simulation to drive the device into large-signal operation provides us with insight into the various forms of distortion and compression.

In evaluating active devices for RF circuit design, we have a few different technologies to consider. First, we will examine the semiconductor junction diode—a basic *pn* junction—and venture from there to the bipolar transistor. An investigation of metal-oxide semiconductor (MOS) devices, another important set of semiconductors for medium frequencies, will follow. We will also examine the gallium arsenide metal-gate Schottky field-effect transistor (GaAs MESFET) and members of its family. Since critical large-signal parameters for some devices may not be available from their manufacturers, we will examine

parameter-extraction techniques and problems. For various device types, we will provide insight into the noise properties of the device.

2-1 DIODES*

The diode model contains a nonlinear current source that follows the Shockley equation:

$$\text{Current} = \text{IS}\left(e^{V_j/\text{N}V_t} - 1\right) \tag{2-1}$$

where V_j = voltage across the junction
V_t = thermal voltage (= kT/q)

These values, with the model parameters IS and N, are used to model the current–voltage effects of the semiconductor junction. This does not include the nonideal operation of real diodes. For example, at low currents (less than 1 nA), other semiconductor processes that increase the flow of currents become noticeable.

By setting IS to different values, we can obtain the characteristics of other devices, such as a Schottky barrier diode, or a silicon diffused-junction diode. High-current effects are modeled, grossly, by including a series resistance that is intended to combine the effects of bulk resistance (the material on each side of the junction) and high-level injection. At high currents, the observed diode current stops following the Shockley form

$$I_{\text{forward}} = \text{IS} \times e^{V_j/\text{N}V_t} \tag{2-2}$$

and approaches a modified form

$$I_{\text{forward}} = \text{IS} \times e^{V_j/2\text{N}V_t} \tag{2-3}$$

2-1-1 Large-Signal Diode Model

Three diode models are used in the industry:

- Microwave diode model
- PIN diode model
- Enhanced SPICE diode model

Figure 2-1 shows the large-signal microwave diode model. Its keywords appear in Table 2-1. This model can also be used to simulate varactor diodes.

Table 2-2 lists SPICE parameters for a selection of Schottky mixer diodes by Alpha.

In most cases, the diode capacitance is modeled by a voltage-dependent capacitor, which is connected in parallel with the nonlinear current generator described previously, to represent the charge-storage effects of the junction. There are two components to this charge

- The reverse-voltage capacitive effect of the depletion region
- The forward-voltage charge represented by mobile carriers in the diode junction

*Portions of this chapter's diode coverage are based on material in Alpha Industries Semiconductor Application Reports 80800, 80200, and 80500. Used with permission.

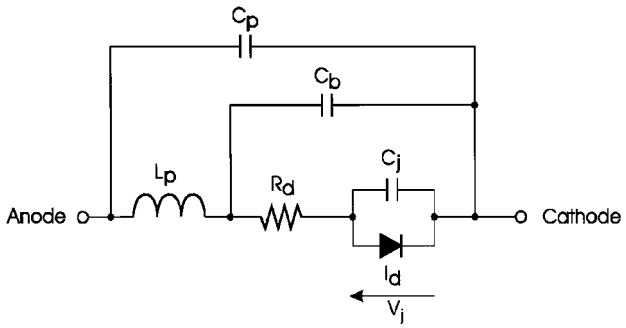


Figure 2-1 The large-signal microwave diode model. This model is temperature dependent.

Reverse-voltage capacitance follows the simple approximation that the depletion region (the area of the junction that is depleted of carriers) serves as the gap between the "plates" of a capacitor. This region varies in thickness, and therefore the capacitance varies with applied voltage. For a step (abrupt) junction, or linearly graded junction, the capacitance approximation is

Table 2-1 Large-signal microwave diode model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
JS	Saturation current	ampere	0
ALFA	Slope factor of conduction current	/volt	38.696
JB	Breakdown saturation current	ampere	10 mA
VB	Breakdown voltage	volt	$-\infty$
E	Power-law parameter of breakdown current	—	10.0
CT0	Zero-bias depletion capacitance	farad	0
FI	Built-in barrier potential	volt	0.8
GAMA	Capacitance power-law parameter	0.5	
GC1	Varactor capacitance polynomial coefficient 1	/volt	0.0
GC2	Varactor capacitance polynomial coefficient 2	/volt^2	0.0
GC3	Varactor capacitance polynomial coefficient 3	/volt^3	0.0
CD0	Zero-bias diffusion capacitance (*pn* diodes)	farad	0
AFAC	Slope factor of diffusion capacitance	/volt	38.696
R0	Bias-dependent part of series resistance in forward-bias condition	ohm	0
T	Intrinsic time constant of depletion layer for abrupt-junction diodes	second	0
KF	Flicker noise coefficient	—	0.0
AF	Flicker noise exponent	—	1.0
FCP	Flicker noise frequency shape factor	—	1.0
AREA	Area multiplier	—	1.0
	Extrinsic Model		
CP	Package parasitic capacitance	farad	0.0
CB	Beam-lead parasitic capacitance	farad	0.0
LP	Package parasitic inductance	henry	0.0

Table 2-2 Diode SPICE parameters

Parameter	Unit	SMS1546	SMS3922	SMS3923	SMS3924	SMS3926	SMS3927	SMS3928	SMS7621	SMS7630
I_S	A	3E–7	3E–8	5E–9	2E–11	2.5E–07	1.3E–09	9E–13	4E–8	5E–06
R_S	Ω	4	9	11	11	4	4	4	12	30
n	—	1.04	1.08	1.05	1.08	1.04	1.04	1.04	1.05	1.05
T_d	S	1E–11	8E–11	8E–11	8E–11	1E–11	1E–11	1E–11	1E–11	1E–11
C_{J0}	PF	0.38	0.9	0.93	1.6	0.42	0.39	0.39	0.1	0.14
m	—	0.36	0.26	0.24	0.4	0.32	0.37	0.42	0.35	0.4
E_G	EV	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
X_{ti}	—	2	2	2	2	2	2	2	2	2
F_C	—	0.5	0.5	0.5	0.5	0.5	0.5	0.5	5	0.5
B_J	V	3	20	46	100	2	3	4	3	2
I_{BV}	A	1E–5	1E–5	1E–5	1E–5	1.00E–05	1.00E–05	1.00E–05	1E–5	0.0001
V_J	—	0.51	0.65	0.15	0.84	0.495	0.595	0.800	0.51	0.34

$$\text{Capacitance} = \frac{C_{J0}}{\left(1 - V_j/\phi\right)^M} \tag{2-4}$$

where C_{J0} is the zero-bias value, ϕ (phi) is the junction barrier potential, and M is the grading coefficient that varies ($\frac{1}{2}$ is used for step junctions and $\frac{1}{3}$ is used for linearly graded junctions, and most junctions are somewhere in between).

There is often confusion about the barrier potential, which appears in the capacitance equation. From capacitance measurements, ϕ (model parameter VJ, and not to be confused with V_j in the equations) takes on a value of nearly 0.7 volt for regular (silicon) junction diodes and a range of 0.58–0.85 volt for various Schottky barrier diodes. This value is sometimes confused with the forward-current voltage drop of the diode or the energy gap of the material; it is *neither* of these.

Varying M generates a variety of reverse-bias capacitance characteristics. Inspection of the capacitance formula reveals that it predicts infinite capacitance for a forward bias, which is not the case for a real junction. Several depletion-capacitance formulas have been proposed that more correctly fit observed operation; however, SPICE uses a simple approach: For forward biases beyond some fraction (set by the parameter FC) of the value for ϕ, the capacitance is calculated as the linear extrapolation of the capacitance at the departure. This provides a continuous numerical result and does not affect circuit operation significantly because, for forward bias, the device capacitance is normally dominated by diffusion capacitance.

The diffusion charge (and therefore the capacitance) varies with forward current and is simply modeled as a transit time (model parameter TT) for the carriers to cross the diffusion region of the junction. The total charge is

$$\text{Diffusion charge} = \text{Device current} \times \text{Transit time} \tag{2-5}$$

and capacitance is the derivative, with respect to bias, of this:

$$\text{Diffusion capacitance} = \text{TT}\,\frac{\text{IS}}{\text{N}V_t}\,e^{V_j/\text{N}V_t} \tag{2-6}$$

Diffusion charge manifests itself as the *storage time* of a switching diode, which is the time required to discharge the diffusion charge in the junction, which must happen before the junction can be reverse-biased (switched off). Storage time is normally specified as the time to discharge the junction so that it is supporting only a fraction (typically 10%) of the initial reverse current. First, a forward current is supplied to the device to charge the junction. Then, as quickly as possible, a reverse current is supplied to the device. Internally, the junction is still forward-biased to a voltage nearly the same as before the switch in current; the junction is still conducting at the forward-current rate. This internal current adds to the external current as the total current discharging the junction. As the junction voltage decreases, the internal current falls off exponentially (according to the Shockley equation). This system is a relatively simple differential equation that can be solved to an explicit equation for the TT parameter (assuming complete discharge) as follows:

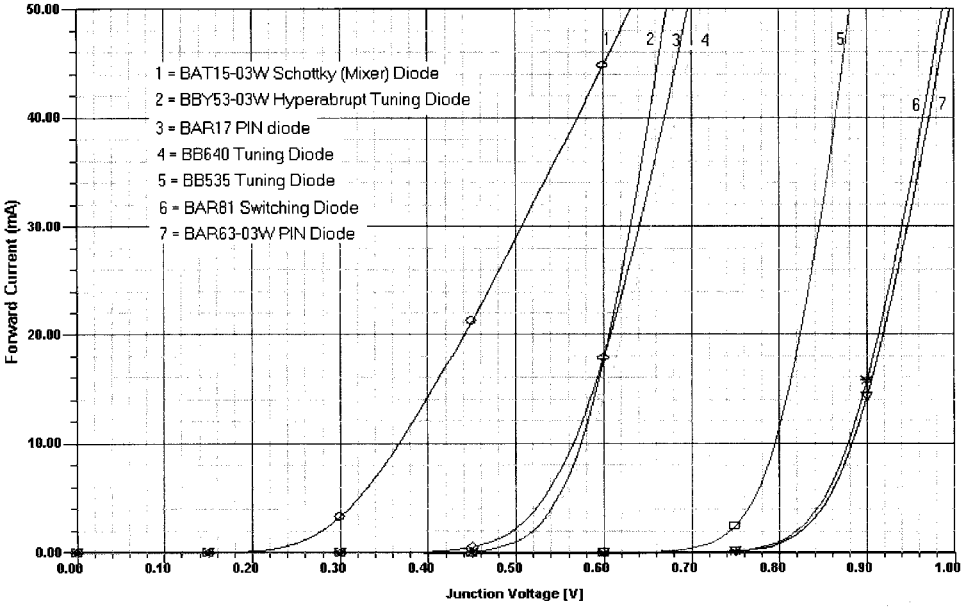


Figure 2-2 The dc *I–V* curves for seven diodes, showing the various barrier voltages that result from different doping profiles.

$$\text{Transit time} = \frac{\text{Storage time}}{\ln\left(\dfrac{I_F - I_R}{-I_R}\right)} \tag{2-7}$$

The diffusion charge dominates the reverse recovery characteristic of the diode. During the last part of the recovery, as the junction becomes reverse-biased, the depletion capacitance dominates. This causes the small tail at the end of the discharge cycle. Total capacitance is taken to be the sum of these capacitances: the depletion approximation dominates for reverse bias as the device current is small, and the diffusion proximity dominates for forward bias as the device current is large.

A special case for diode application is the switching diode, and its description and application will be part of a later chapter. Figure 2-2 shows the dc *I–V* curves, which indicate the different voltage potentials that are a result of the different doping profiles.

2-1-2 Mixer and Detector Diodes

Electrical Characteristics and Physics of Schottky Barriers. Schottky barrier diodes differ from junction diodes in that current flow involves only one type of carrier instead of both types. That is, in *n*-type Schottky barrier diodes the forward current consists of holes flowing from the *n*-type material into the metal.

Diode action results from a contact potential set up between the metal and the semiconductor, similar to the voltage between the two metals in a thermocouple. When metal is brought into contact with an *n*-type semiconductor (during fabrication of the chip), electronics diffuse out of the semiconductor into the metal, leaving a region under the contact that

has no free electrons ("depletion layer"). This region contains donor atoms that are positively charged (because each lost its excess electron), and this charge makes the semiconductor positive with respect to the metal. Diffusion continues until the semiconductor is so positive with respect to the metal that no more electrons can go into the metal. The internal voltage difference between the metal and the semiconductor is called the contact potential and is usually in the range of 0.3–0.8 volt for typical Schottky diodes. A cross section is shown in Figure 2-3.

When a positive voltage is applied to the metal, the internal voltage is reduced, and electrons can flow into the metal. The process is similar to thermionic emission of electrons from the hot cathode of a vacuum tube, except that the electrons are "escaping" into a metal instead of into a vacuum. Unlike the vacuum tube case, room temperature is "hot" enough for this to happen if enough voltage is applied. However, only those electronics whose thermal energy happens to be many times the average can escape, and these "hot electrons" account for all the forward current from the semiconductor into the metal.

One important thing to note is that there is no flow of minority carriers from the metal into the semiconductor and thus no neutral plasma of holes and electrons is formed. Therefore, if the forward voltage is removed, current stops "instantly," and reverse voltage can be established in a few picoseconds. There is no delay effect to charge storage as in junction diodes. This accounts for the exclusive use of barrier diodes in microwave mixers, where the diode must switch conductance states at microwave oscillator rates.

The current–voltage relationship for a barrier diode is described by the Richardson equation (which also applies to thermionic emission from a cathode). The derivation is given in many textbooks (e.g., Sze).

$$I = AA_{\mathrm{RC}}T^2 \exp\left(-\frac{q\phi_B}{kT}\right)\left[\exp\left(\frac{qV_J}{kT}\right) - M\right] \tag{2-8}$$

where A = area (cm^2)

A_{RC} = modified Richardson constant ($\mathrm{A}/K^2 \cdot \mathrm{cm}^2$)

k = Boltzmann's constant

T = absolute temperature (K)

ϕ_{B} = barrier height in volts

V_J = external voltage across the depletion layer (positive for external voltage) = $V - IR_S$

R_S = series resistance

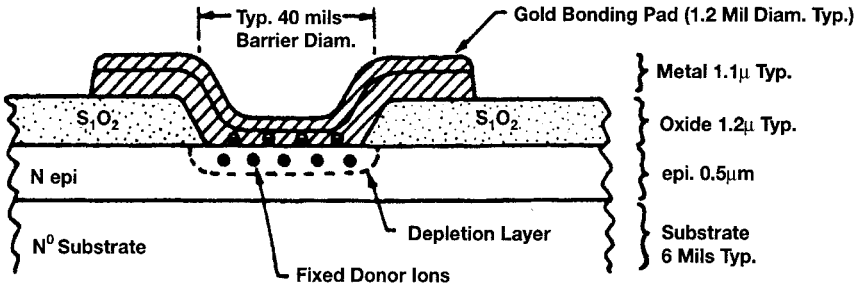


Figure 2-3 Schottky diode chip cross section.

M = avalanche multiplication factor
I = diode current in amperes (positive forward current)

The barrier height, ϕ_B, is typically a few tenths of a volt higher than the contact potential, ϕ_C (about 0.15 V higher than ϕ_C for silicon). This equation agrees well with experimental data for diodes without surface leakage, but it is difficult to use because A_{RC}, ϕ_B, and M are all dependent on applied voltage.

The major cause for variation in ϕ_B with voltage is the "image effect," in which the barrier height is lowered as the electric field near the metal is increased, especially at the edges. A better equation for circuit designers to use is one in which all parameters are independent of voltage and current. The simplest one that agrees reasonably well with Richardson's equation is

$$I = I_S\left[\exp\left(\frac{V_J}{0.028}\right) - 1 + \frac{K}{\left(1 - V_B/V\right)}\right] \tag{2-9}$$

where I_S = saturation current (a temperature-dependent quantity)
0.028 = nkT/q at room temperature (n = 1.08)
n = forward slope factor (derived from the variation of ϕ_B with forward voltage)
K = reverse slope factor (expressing the variation of ϕ_B with reverse voltage)
V_B = breakdown voltage (the voltage at which M = 1)

As before, V and I are considered positive for forward bias and negative for reverse bias.

Typical ranges for these parameters for microwave Schottky and point-contact mixer diodes are:

I_S = 10^{-12} to 10^{-5} A
n = 1.04–1.10
R_S = 2–20 Ω
K = 8–100
V_B = 2–20 V

The quantities I_S and 0.028 are strongly temperature dependent, while both R_S and V_B increase with temperature to a slight degree. R_S increases with current at high current levels (due to carrier velocity saturation) but is essentially independent of current at 10 mA and below for mixer diodes. Thus, for normal mixer and detector operation, R_S can be considered constant.

Agreement between Eqs. (2-8) and (2-9) is not perfect but Eq. (2-9) is much easier to use and is preferred by most circuit designers. A comparison of the two equations near zero bias gives the following relationship between zero-bias barrier height, ϕ_0, and saturation current:

$$I_S = AA_{RC}T^2\exp - \frac{q\phi_0}{kT}$$

$$\cong \left(\frac{10^7 A}{\text{cm}^2}\right)A\exp\left(-\frac{\phi_0}{0.026}\right) \quad \text{(for } n \text{ silicon at room temperature)} \tag{2-10}$$

Small-Signal Parameters. By combining Eqs. (2-9) and (2-10), the values of the parameters in Eq. (2-9) can be derived from a few simple measurements. Many specific equations can be derived, but the following are commonly used for production measurements:

$$R_S = \frac{V_{F10} - V_{F1} - 0.065}{0.009} \quad (\text{for } n = 1.08) \tag{2-11}$$

$$\phi_0 = \frac{V_{F1} - 0.001R_S + 0.28 + 0.12\log_{10} D}{1.08} \tag{2-12}$$

$$n = \frac{V_{F1} - V_{F0.1} - 0.0009R_S}{0.060} \tag{2-13}$$

$$K = \left(\frac{I_{R1}}{I_S} - 1\right)V_B \tag{2-14}$$

$$I_S = \exp\left(-\frac{V_{F1}}{0.028} + \frac{R_S}{28}\right) \quad (\text{in mA}) \tag{2-15}$$

where $V_{F0.1}$, V_{F1}, and V_{F10} are the forward voltages at 0.1, 1, and 10 mA, respectively, and I_{R1} is the reverse current at 1 volt. (The derivation of these equations requires that I_S be small compared to 0.1 mA.) The quantity D is the diameter of the metal–silicon contact in mils.

Measuring V_{F1} at 1 mA and 10 mA instead of some other current levels leads to the best accuracy for typical mixer diodes.

The total dynamic resistance for a forward-biased diode is given by

$$R_T = \frac{dV}{dI} = R_S + \frac{nkT}{q(I + I_S)} = R_S + R_B \tag{2-16}$$

and

$$R_B = \frac{28}{I + I_S} \quad \text{at room temperature (with } I \text{ and } I_S \text{ in mA, } n = 1.08) \tag{2-17}$$

This equation is also good at zero bias (unless K is very large or there is significant surface leakage). That is,

$$R_0 = R_S + \frac{28}{I_S} \tag{2-18}$$

For reverse voltages of a few volts, the dynamic resistance is dominated by the K term:

$$R_R = \text{Reverse resistance} = \frac{dV}{dI} \cong \frac{V_B}{KI_S} \tag{2-19}$$

For typical values of I_S, R_0 is larger than 5000 Ω and R_R is larger than 100 kΩ. For some zero-bias Schottky applications, it is desirable for R_0 to be made smaller than this.

The factors that determine R_S are (1) the thickness of the epitaxial layer, (2) the epitaxial layer doping level (N_D), (3) the barrier diameter, (4) the substrate resistivity ("spreading resistance"), (5) the contact resistances of the metals used for the barrier and the substrate contact, and (6) the resistance associated with the bonding wire or whisker. The barrier height is about 0.15 volt higher than the contact potential between the barrier metal and the semiconductor and is influenced by the method used to apply the metal, conditions at the edge of the junction, and the doping level. Saturation current depends on barrier height, junction area, and temperature; and the slope factors, n and K, depend on doping level, punchthrough voltage, and edge conditions.

Junction Capacitance. The capacitance of a Schottky barrier chip results mainly from two sources: the depletion layer under the metal–semiconductor contact and the capacitance of the oxide layer under the bonding pad (the so-called overlay capacitance). The bonding pad is required because the typical Schottky barrier diameter is so small that it is impractical to bond directly to the metal on the junction. If the semiconductor epitaxial layer is uniformly doped, the capacitance–voltage characteristic is similar to that of a textbook "abrupt-junction" diode.

$$C_J = \frac{\varepsilon_S \varepsilon_0 A'}{X_D} + C_0 \tag{2-20}$$

$$X_D = \sqrt{\frac{2\varepsilon_S \varepsilon_0 (\phi_C - V)}{qN}} \tag{2-21}$$

where ϕ = contact potential
C_0 = overlay (bonding pad) capacitance
ε_S = dielectric constant of the semiconductor ($\cong$ 12 for Si or GaAs)
N = doping level for the epitaxial layer
A' = effective contact area, including fringing corrections

In practical terms, the capacitance can be related to the 0-V barrier capacitance defined by

$$C_{B0} = \frac{\varepsilon_S \varepsilon_0 A'}{X_{D0}} \tag{2-22}$$

where

$$X_{D0} = \sqrt{\left(\frac{1.3 \times 10^{15}}{N_D}\right)\phi_C} \quad \text{(in } \mu\text{m)} \tag{2-23}$$

The resulting C–V relationship can be written

$$C_J = \frac{C_{B0}}{\sqrt{1 - V/\phi_C}} + C_0 \tag{2-24}$$

The contact potential, ϕ_C, is related to the barrier height as follows:

$$\phi_C = \phi_B - 0.026\left[1 + L_n\left(\frac{N_C}{N}\right)\right] \tag{2-25}$$

The theoretical meaning of these terms can be clarified by looking at Figure 2-4.

Parameter Trade-Offs

Barrier Height. The barrier height of a Schottky diode is important because it directly determines the forward voltage. In order to get a good noise figure, the LO drive voltage, V_L, must be large compared to V_T, which is essentially V_{F1}. Normally, it is best to have a low

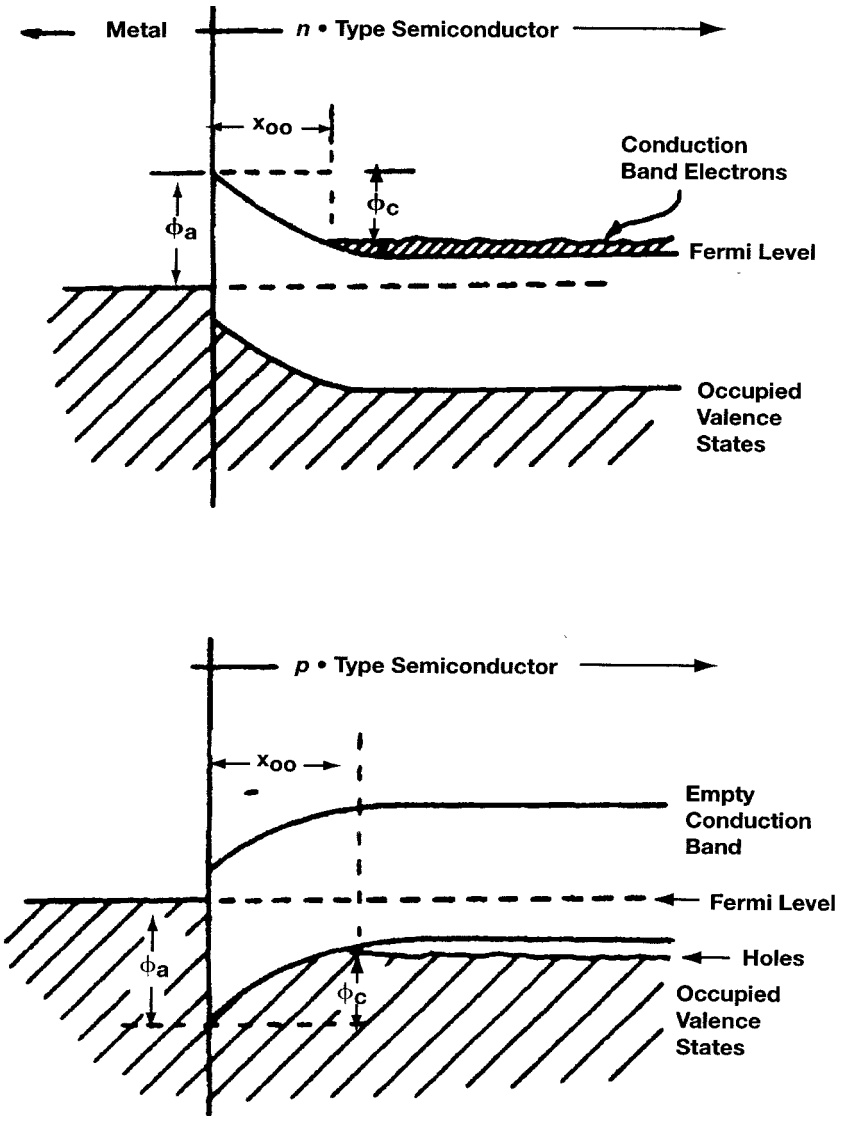


Figure 2-4 Schottky diode band diagrams.

Table 2-3 LO power and V_{F1} for various applications

Type	Typical V_{F1}	LO Power	Application
"Zero bias"	0.10–0.25	< 0.1 mW	Mainly for detectors
Low barrier	0.25–0.35	0.2–2 mW	Low-drive mixers
Medium barrier	0.35–0.50	0.5–10 mW	General purpose
High barrier	0.50–0.80	> 10 mW	High dynamic range

forward voltage (low V_{F1}), or low drive) diode, to reduce the amount of LO power needed. However, if high dynamic range is important, high LO power is needed, and the diode can have a higher V_F and should also have a high V_B (see Table 2-3).

Noise Figure Versus LO Power. At low LO drive levels, the noise figure is poor because of poor conversion loss, due to a too-low conduction angle. At very high LO drive levels, the noise figure again increases due to diode heating, excess noise, and reverse conduction.

If a high LO drive level is needed, for example, to get higher dynamic range, a high V_B (> 5 V) should be specified. However, Nature requires that you pay for this with higher R_S (lower f_C), so the noise figure will be degraded compared to what could be obtained with diodes designed for lower LO drive. Forward voltage and breakdown are basically independent parameters, but high breakdown is not needed or desirable unless high LO power is used.

Such a high-breakdown diode will have low reverse current (which is important only if the diode has to run hot).

Silicon Versus GaAs. Typical silicon Schottky diodes have cutoff frequencies in the 800–200-GHz range, which is good enough through the Ku band.

At the Ku band and above, or for image-enhanced mixers, higher f_C may be needed, which calls for the use of GaAs diodes. These have lower R_S due to higher mobility, which translates to cutoff frequencies in the 400–1000-GHz range.

However, if your IF is low, be careful; GaAs diodes have high 1/*f* noise. They also have high V_{F1}, so more LO power is required.

C_J Versus Frequency. There is quite a lot of latitude in choosing C_J. However, in general, the capacitive reactance should be a little lower than the transformed line impedance (Z_0). If Z_0 is not known, a good way to start is to use $X_C = 100\ \Omega$. Experience has shown that most practical mixers use an X_C near this value (a little higher in waveguide, and lower in 50-Ω systems). This translates to the following rule of thumb for choosing the junction capacitance of a diode for operation at frequency *f* (in GHz):

$$C_{J0} \approx \frac{100}{\omega}$$

$$\approx \frac{1.6}{f} \quad \text{(in pF)} \tag{2-26}$$

In order to evaluate possible tolerances, we show the range of forward current as a function of diode voltage (Figure 2-5), the junction capacitance as a function of the bias voltage

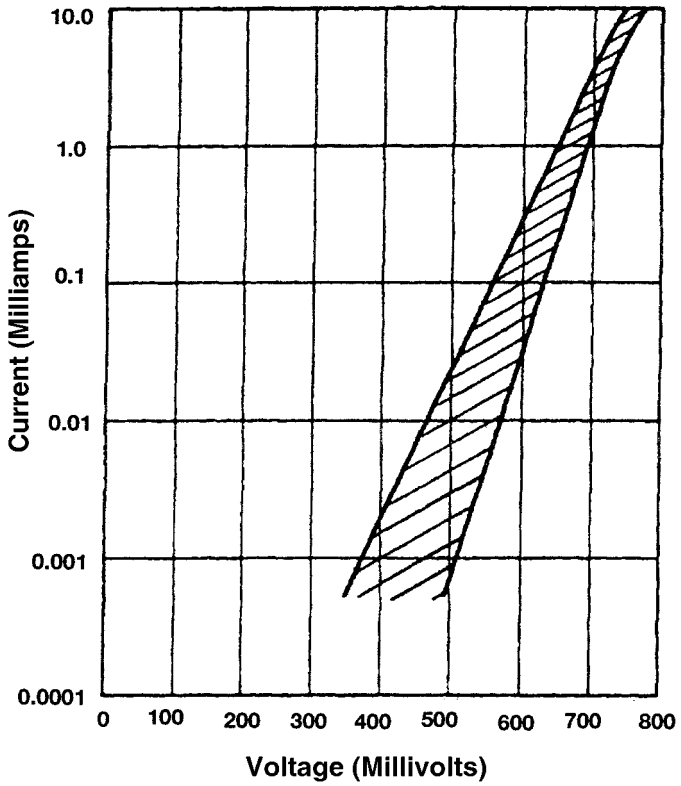


Figure 2-5 Forward dc characteristic curve range—current versus voltage.

(Figure 2-6), and finally some important RF parameters, such as noise figure and IF impedance, as a function of the local-oscillator drive (Figure 2-7).

Mixer Diodes. As an example of some of the parameters for mixer diodes, Table 2-4 gives data on some of the X-band mixer diodes. NF is measured at 9.375 GHz.

Linear Diode Model. Figure 2-8 shows the linear diode model. Its keywords appear in Table 2-5.

The circuit simulators, such as those supplied by Ansoft and Hewlett-Packard, provide a model library that has SPICE-type parameters for diodes (regular diodes, varactor diodes, and PIN diodes) as well as bipolar transistors and FETs, which will be discussed later.

2-1-3 PIN Diodes

Introduction. The PIN diode, in comparison with other microwave semiconductor devices, is fairly easy to understand. This makes possible the reduction of its complex behavior to relatively simple terms.

We do not attempt to describe the many possible microwave circuits in which PIN diodes are used. Rather, we attempt to explain the behavior of the diode in all aspects, giving the facts and some of the theory behind the facts. We offer the circuit designer the opportunity

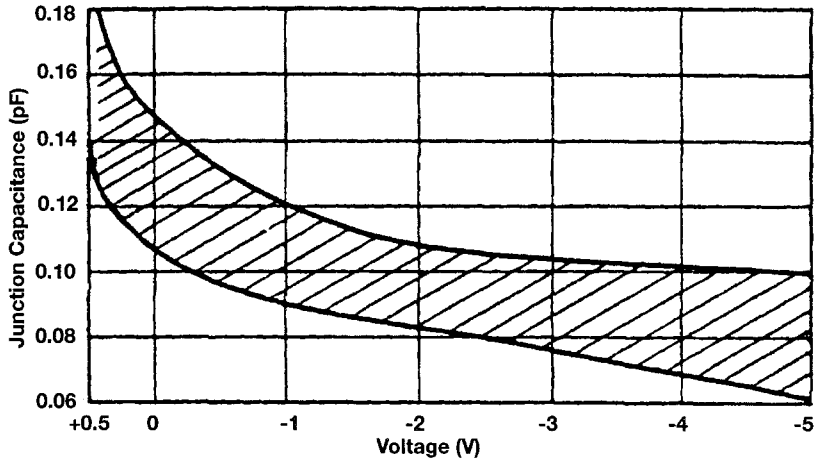


Figure 2-6 Junction capacitance range versus voltage.

to understand the PIN, so that he/she can understand its behavior in circuits. We assume the reader knows the circuit equations; to that knowledge we hope to add diode equations.

Most of the material presented consists of generalized data and explanations of the behavior of PIN diodes; we conclude with a brief description of circuit performance, test methods, and some hints on proper PIN specification writing.

The user can then evaluate the trade-offs involved in diode design and performance and be able to select the most nearly optimum diode from the wide range of diodes offered.

Large-Signal PIN Diode Model. Figure 2-9 shows the large-signal model for a PIN diode. Table 2-6 lists its keywords.

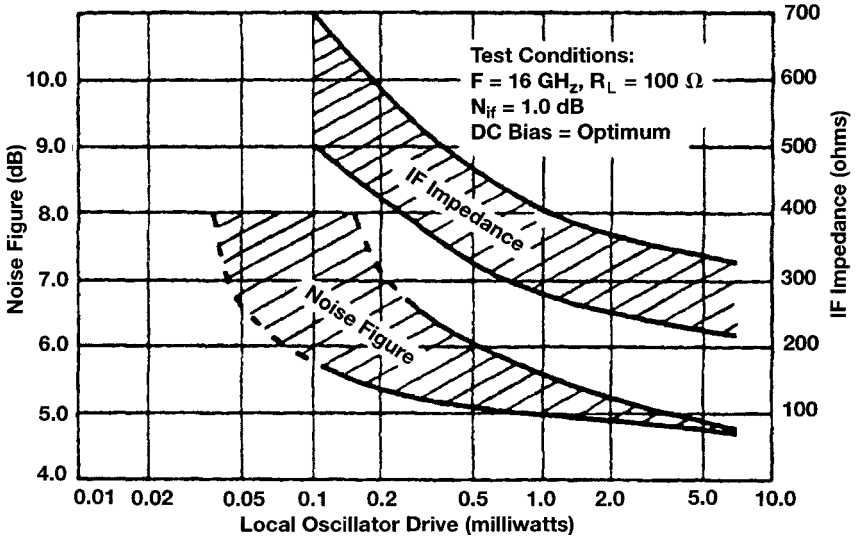


Figure 2-7 RF parameters versus local-oscillator drive level.

Table 2-4 X-band mixer diode data

Material	Barrier	Typical V_F (@1 mA)	Typical F_{C0} (GHz)	Typical R_S (Ω)	Typical C_{J0} (pF)	Maximum NF (dB)
n GaAs	High	0.70	1000	—	0.15	5.0[a]
n GaAs (BL)	High	0.70	500	—	0.15	6.0[a]
n GaAs (chip)	High	0.70	1000	—	0.15	5.3[a]
n Silicon (BL)	Low	0.28	150	6	0.20	6.5
n Silicon (quad)	Low	0.28	150	6	0.20	6.5
p Silicon (BL)	Low	0.28	150	12	0.20	6.5
n Silicon (BL)	High	0.60	100	8	0.20	6.5
n Silicon (quad)	High	0.60	100	8	0.20	6.5
n Silicon	Low	0.28	200	6	0.15	5.5
p Silicon	Low	0.28	200	18	0.14	6.0
p Silicon	Medium	0.40	150	150	0.12	6.5
n Silicon	Low	0.28	150	8	0.18	6.5
p Silicon	Low	0.28	150	12	0.18	6.5

[a]Specified for N_F = 1.0 dB.

Basic Theory: Variable Resistance. Intrinsic or "pure" silicon, as it can be grown in a laboratory, is an almost lossless dielectric. Some of is physical properties include:

Dielectric constant (relative)	12
Dielectric strength	400 V/mil (approximate)
Specific density	2.3
Specific heat	0.72 J/g · °C
Thermal conductivity	1.5 W/cm · °C
Resistivity	300,000 Ω/cm

Since a PIN diode is valuable essentially because it is a variable resistor, let us concentrate initially on the resistivity. Consider a volume comparable to a typical PIN diode chip, say, 20 mils in diameter and 2 mils thick. This chip has a dc resistance of about 0.75 MΩ. High

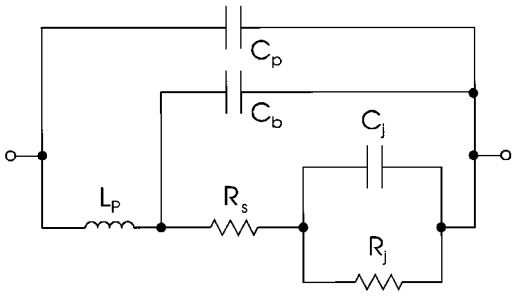


Figure 2-8 The linear diode model. This model is temperature dependent.

Table 2-5 Linear diode model keywords

Keyword	Description	Unit	Default
LP	Package inductance	henry	0.0
CB	Beam-lead capacitance	farad	0.0
CP	Package capacitance	farad	0.0
RS	Contact resistance	ohm	0.0
RJ	Junction resistance	ohm	
CJ	Junction capacitance	farad	

resistivity in any material indicates that most of the likely carriers of electric charge, electrons and holes, are tightly held in the crystal lattice and cannot "conduct."

In real life there are impurities, typically boron, that cannot be segregated out of the crystal. Such impurities contribute carriers, holes or electrons, that are not very tightly bound to the lattice and therefore lower the resistivity of the silicon.

Through various techniques we can adjust the level of impurities, called *dopants*, to produce resistivities ranging from 10 kΩ/cm (for good PIN diodes) to 0.001 Ω/cm (for substrates).

If the impurity adds "electrons" to the crystal, it is called a *donor*; if it adds holes it is called an *acceptor*. Boron adds holes, hence it is an acceptor, and the silicon plus boron combination is called *p*-type, or *positive*, because it has an excess of positive carriers. Phosphorus, on the other hand, is a *donor*, adding electrons, and the corresponding mix is *n*-type, or *negative*.

Notes on the PIN Diode Model

1. The PIN diode model is used to model a bias-dependent RF resistance for use in PIN diode circuits such as attenuators and switches. The resistance varies from RMAX to RS using the R function above. A typical R versus *I* characteristic is shown in Figure 2-10 with parameters IS = 5.96 nA, RS = 2.016, RMAX = 6500, K1 = 0.1272, K2 = 1.0, N = 2.077.
2. The transit-time parameter, TT, can also be used to approximately model a switching PIN diode's reverse-recovery time—a value often provided by diode manufacturers.
3. Diode breakdown can be modeled by specifying IBV and BV parameters.

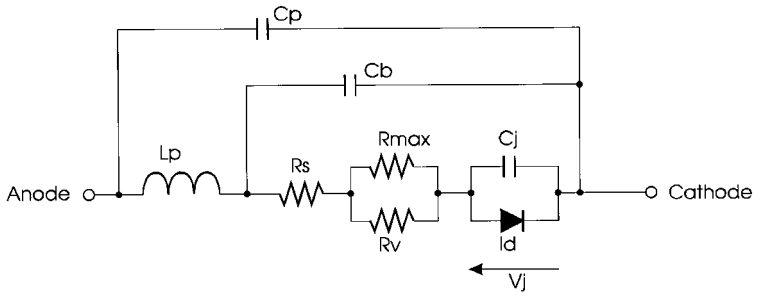


Figure 2-9 The large-signal PIN diode model. This model is temperature dependent.

Table 2-6 PIN diode model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
IS	Saturation current	ampere	1.0E–14
N	Emission coefficient		1.0
IBV	Magnitude of current at the reverse breakdown voltage	ampere	1.0E–10
BV	Magnitude of the reverse breakdown voltage	volt	∞
FC	Coefficient for forward-bias depletion capacitance		0.5
CJ0	Zero-bias *pn* junction capacitance	farad	0.0
VJ	Built-in junction potential	volt	1.0
M	*pn* Junction grading coefficient		0.5
GC1	Varactor capacitance polynomial coefficient 1	/volt	0.0
GC2	Varactor capacitance polynomial coefficient 2	/volt^2	0.0
GC3	Varactor capacitance polynomial coefficient 3	/volt^3	0.0
TT	Transit time	second	0.0
K1	Variable resistance coefficient	volt	0.0
K2	Variable resistance current exponent		1.0
RMAX	Maximum resistance of PIN intrinsic region	ohm	0.0
KF	Flicker noise coefficient		0.0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
AREA	Area multiplier		1.0
	Extrinsic Model		
RS	Series resistance (minimum resistance of PIN diode)	ohm	0.0
CP	Package parasitic capacitance	farad	0.0
CB	Beam-lead parasitic capacitance	farad	0.0
LP	Package parasitic inductance	henry	0.0

4. The reverse-bias capacitance characteristics can be more accurately modeled than the common expression derived from *pn* junction theory. The capacitance grading coefficient exponent can be expressed as a polynomial function of voltage by specifying values for GC1, GC2, and GC3.
5. The PIN diode model was derived from J. Walston, "SPICE Circuit Yields Recipe for PIN Diode," *Microwaves & RF*, November 1992.
6. Following Simon M. Sze, *Physics of Semiconductor Devices* (New York: John Wiley, 1982), the variable resistance may be modeled by setting

$$\text{K1} = \frac{3}{8} V_T \frac{W^2}{D_a \tau_a} \qquad \text{K2} = 1.0$$

where W is the width of the intrinsic region, D_a is the ambipolar diffusion coefficient, τ_a is the ambipolar lifetime, and V_T is the thermal voltage.

There are many concepts important to the physicist, but not to the diode user, that elaborate on the impact of impurities on the behavior of silicon. These can be studied in Watson [1]. The more carriers added, the lower the resistivity.

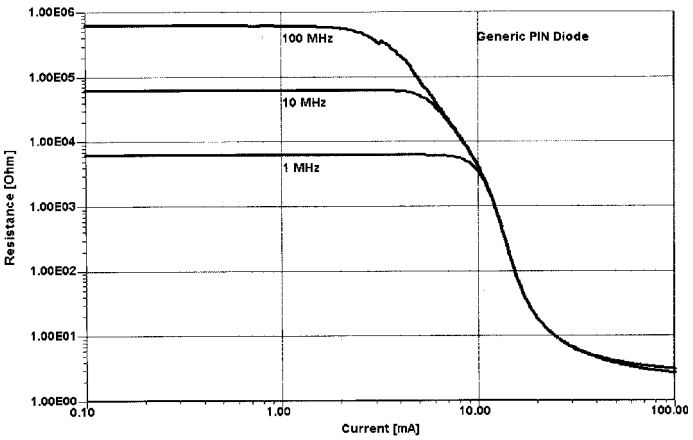


Figure 2-10 Simulated PIN diode resistance as a function of dc at 1, 10, and 100 MHz.

If one wished to vary the resistance of a given diode, in principle he/she could bring it into a semiconductor laboratory, add or subtract carriers as desired, and perhaps even make the process reversible. However, this is a slow, expensive, and impractical way to make a variable resistor; one would be better advised to take a wrench and a soldering iron and replace a component.

The PIN diode derives its value from the fact that the free-charge-carrier concentration in silicon, and hence its resistance, can be varied electronically by means of current from a simple bias supply. This can be done rapidly (in nanoseconds in some cases), reversibly, repeatably, and accurately. The thing that makes this possible is called a *junction*, the interface between the relatively pure silicon in the middle of the PIN diode (the I stands for *intrinsic*) and the heavily doped layers on either end, p^+ and n^+. The p^+ region is rich in holes; the n^+ region is rich in electrons. Both of these regions have low resistance. The I region is the variable resistive element in the diode (see Figure 2-11). In the absence of any external bias, internal effects within the crystal keep the charges fixed; the resistance of the I region is high.

When the p^+ region (anode) is biased positively with respect to the n^+ region (cathode), the interface potential "barrier" is overcome, and direct current flows in the form of holes

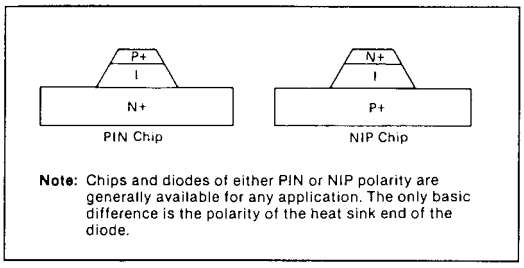


Figure 2-11 General outline of PIN diode construction.

streaming from p^+ toward n^+, with electrons moving in the opposite direction; we say that free carriers have been injected into the I region. The resistance of the I region becomes low.

The number of free carriers within the I region determines the resistivity of the region and thus the resistance of the diode.

Consider "one hole" and "one electron" drifting in opposite directions in the I region under the impetus of the applied field. Under certain conditions, imperfections in the silicon may cause these carriers to recombine. They are no longer available to constitute current or to lower the resistivity of the I region.

It can be shown that the amount of "recombination" between holes and electrons that continuously takes place in a semiconductor is governed by a property of the lattice called *lifetime*. In fact, lifetime is defined as the reciprocal of recombination rate.

Thus, $Q_S = Q_0 \exp(-t/T_L)$, where Q_S is the total amount of free charge "stored" in the I region, and T_L is the lifetime, or the mean time between recombination events. In the steady-state condition, the bias supply must deliver current to maintain a constant Q_S. The required current is

$$I_{\text{dc}} = \frac{Q_S d}{dt} = \frac{-Q_S}{T_L} \tag{2-27}$$

or $Q_S = I_{\text{dc}} T_L$, dropping the minus sign.

Ignoring some details that are not crucial to this section, we can now calculate the resistance of a given diode, of area A and thickness W. (W stands for *base width*, the width or thickness of the intrinsic layer.) The p^+ and n^+ regions have essentially zero resistance, as they are very heavily doped.

The resistivity of a given material is inversely proportional to the number of free carriers, N, and the mobility (not quite the same as velocity) of the carriers. Thus,

$$\rho = \frac{1}{q(\mu_n N + \mu_p P)} \tag{2-28}$$

where μ_n and μ_p are mobilities of electrons and holes, and N and P are numbers of electrons and holes. Simplifying, we find

$$\rho = \frac{C}{Q_S d} \tag{2-29}$$

where C is a collection of constants, and $Q_S d$ is the stored charge density (numbers per unit volume). For our piece of silicon, volume is WA, and

$$Q_S d = \frac{Q_S}{WA} \tag{2-30}$$

The resistivity is

$$\rho = \frac{CWA}{Q_S} \tag{2-31}$$

and the resistance is

$$R_S = \rho \frac{W}{A} = \frac{CW^2}{I_{\text{dc}} T_L} \tag{2-32}$$

This is a fundamental equation in PIN diode theory and design.

Rigorous analysis shows that

$$R = \frac{2kT/q}{I_f} \sinh\left(\frac{W}{2\sqrt{DT_L}} \tan^{-1}\left[\sinh\frac{W}{2\sqrt{DT_L}}\right]\right) \tag{2-33}$$

where k = Boltzmann's constant
T = temperature (in kelvins)
D = diffusion coefficient = $\mu kt/q$

For most PIN diodes, W/DT_L is less than unity, and the equation simplifies to the simple equation above.

Typical data on R_S as a function of bias current are shown in Figure 2-12. Many design choices are available, as the data indicate. Many combinations of W and T_L have been developed to satisfy the full range of applications.

Breakdown Voltage, Capacitance, Q Factor. The previous section on R_S explained how a PIN can become a low resistance, or a "short." This paragraph will describe the other state: a high impedance, or an "open." Clearly, the better PIN diode is the one that has the better on/off ratio at the frequency and power level of interest.

If we return to the undoped or intrinsic I region, we note that it is an almost lossless dielectric. As such, it has a dielectric strength of about 400 V/mil, and all PIN diodes have

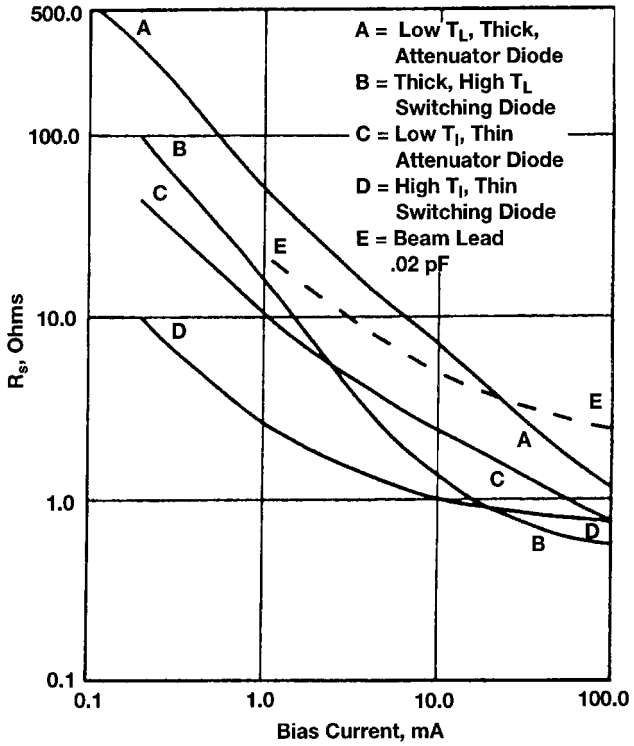


Figure 2-12 Typical series resistance as a function of bias (1 GHz).

a parameter called V_b, breakdown voltage, which is a direct measure of the width of the I region. Voltage in excess of this parameter results in a rapid increase in current flow (called avalanche current); see Figure 2-13. When the negative bias voltage is below the bulk breakdown of the I region, a few nanoamperes will be drawn. As V_b is approached, the leakage current increases often gradually, as is exaggerated in the curve. This current is primarily caused by less-than-perfect diode fabrication, although there is some contribution from temperature. Typically, the leakage current occurs at the periphery of the I region. For this reason, various *passivation* materials (silicon dioxide, silicon nitride, hard glass) are grown or deposited to protect and stabilize this surface and minimize leakage. These techniques have been well advanced over the years, and PIN diode reliability has improved as a result.

Most diodes are specified in terms of minimum V_b for a nominal leakage, usually 10 μA.

It will be noted later that RF voltage swings in excess of the rated V_b are permitted, for the mechanisms causing leakage current do not always respond at radiofrequencies. However, bulk breakdown is effectively instantaneous, and that voltage should never be exceeded.

The next characteristic of our "open" circuit is the capacitance. In simplest form, the capacitance of a PIN diode is determined by the area and width of the I region and the dielectric constant of silicon; however, we have discussed the fact that intrinsic material does contain some carriers and therefore has some conductivity. An *E* field could not exist unless all these carriers were swept out, or depleted.

Application of a reverse bias accomplishes this. At zero bias, the excess carriers on either side of the junction are separated, held apart, by "built-in" fields. This is the contact potential (about 0.5 V for silicon). If there are only a few excess carriers in the I region, this "potential" can separate the charges more easily. The junction "widens" in the sense that, starting at the p^+ and I interface, there is a region of no free carriers called the *depletion zone*. Beyond this depletion zone, the I region still contains the free charges with which it started. With the application of reverse bias, the depletion zone widens.

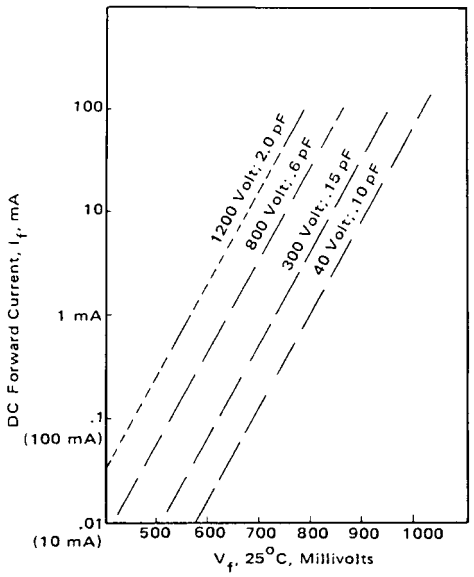


Figure 2-13 Current versus voltage for various PIN diodes.

Eventually, at a bias equal to a so-called *punchthrough* voltage (V_{PT}), the depletion zone fills the entire I region. At this voltage, the 1-MHz capacitance bottoms out and the diode Q reaches its maximum. Figure 2-14 illustrates the equivalent circuit of the I region before punchthrough.

Some very interesting facts can be derived from this model. Consider the undepleted region; this is a lossy dielectric consisting of a volume (area A, length ℓ) of silicon of permittivity 12 and resistivity ε. The capacitance is

$$12\frac{\varepsilon_0 A}{\ell} \tag{2-34}$$

and the admittance is

$$2\pi\frac{12\varepsilon_0 A}{\ell} \tag{2-35}$$

The resistance is $\varepsilon\ell/A$ and the conductance is $A/\varepsilon\ell$.

At very low frequencies, the undepleted zone looks like a pure resistor. At very high frequencies it looks like a lossy capacitor. The "crossover" frequency depends on the resistivity of the I-region material. For ε of 160 Ω/cm, the frequency is 1 GHz. Higher resistivity is generally used for PINs—say, 1000 Ω/cm, and the crossover frequency is 160 MHz.

Diode manufacturers measure junction capacitance at 1 MHz; clearly, what is measured is the depletion-zone capacitance.

If the I region thickness is W and the depletion is X_d, the undepleted region is $(W - X_d)$.

The capacitance of the depleted zone is, proportionally,

$$\frac{1}{X_d} \tag{2-36}$$

and of the undepleted zone,

$$\frac{1}{W - X_d} \tag{2-37}$$

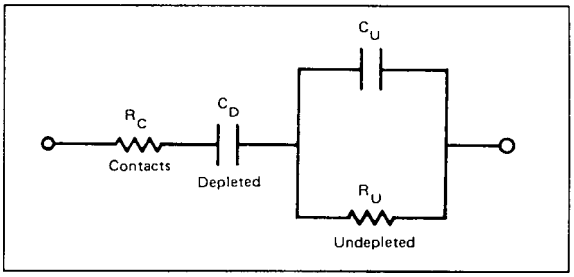


Figure 2-14 Equivalent circuit of I region before punchthrough.

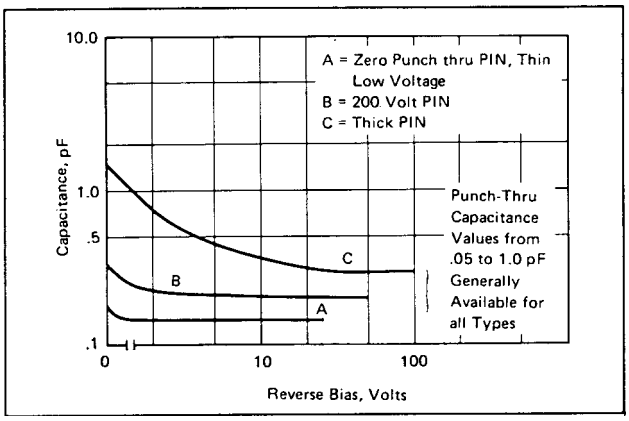


Figure 2-15 Typical 1-MHz capacitance.

The 1-MHz capacitance as a function of reverse bias is seen in Figure 2-15. The 1-MHz capacitance decreases with bias until punchthrough, where $X_d = W$. However, at microwave frequencies well above the crossover, the junction looks like two capacitors in series:

$$C_T = \frac{C_d C_u}{C_d + C_u} \propto \frac{1}{W} \tag{2-38}$$

That is, the microwave capacitance tends to be constant, independent of X_d and bias voltage.

However, since the undepleted zone is lossy, an increase in bias up to the punchthrough voltage reduces the loss.

At any given frequency, the equivalent network can now be drawn as Figure 2-16. R_v is now the equivalent series resistance of the undepleted region. Typical R_v data are shown in Figure 2-17. An alternate equivalent network is shown in Figure 2-18, and typical R shunt data are shown in Figure 2-19.

A good way to understand the effects of series resistance is to observe the insertion loss of a PIN chip shunt mounted in a 50-Ω line, as shown in Figure 2-20. An accepted way to include reverse loss in the figure of merit of a PIN diode is to write

$$Q = \frac{1}{2\pi + \sqrt{R_s R_v C}} \tag{2-39}$$

where R_s and R_v are measured under the expected forward- and reverse-bias conditions at the frequency of interest.

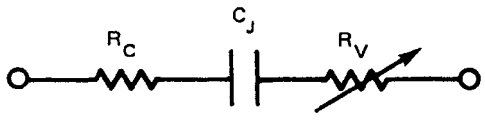


Figure 2-16 Simplified equivalent series circuit.

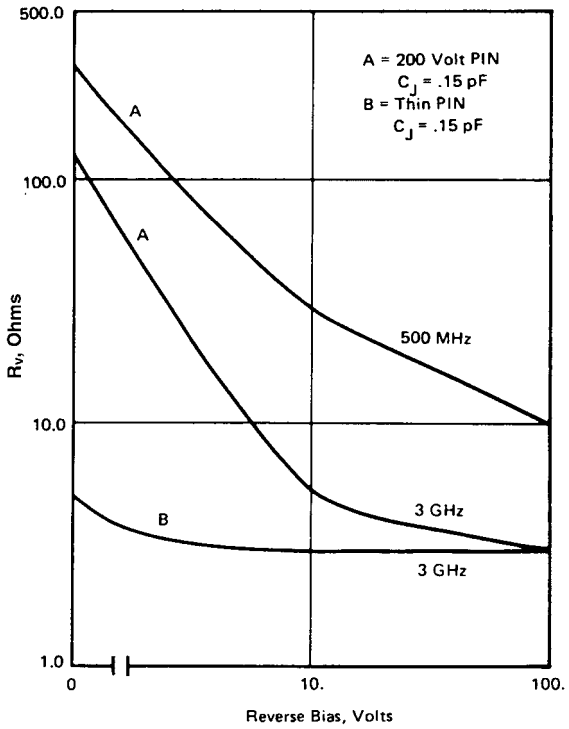


Figure 2-17 Reverse series resistance R_V.

The punchthrough voltage is a function of the resistivity and thickness of the I region. It is advisable to measure loss as a function of bias voltage and RF voltage to determine if the correct diode has been selected for your application. (*Note*: At frequencies below crossover, and diodes with thin I regions, the effective junction capacitance can increase substantially at low forward bias, on the order of 1–200 μA.)

Incidentally, if you are working with PIN or NIP chips that do not have an opaque covering, note that PIN diodes are photosensitive. Incident light causes photogeneration of carriers in the I region, increasing the chip's insertion loss.

PIN Diode Applications. If the intrinsic zone is thick (10–100 μm), we then have a high-reverse-voltage rectifier with a low forward-voltage drop at high current or, in other

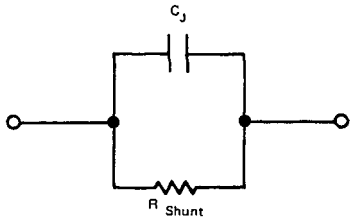


Figure 2-18 Simplified equivalent circuit, shunt.

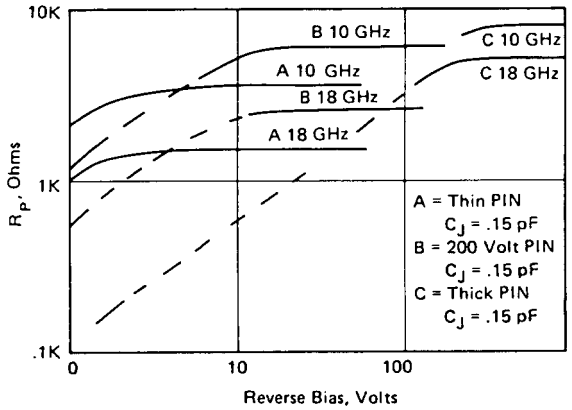


Figure 2-19 Reverse shunt resistance, R_p.

words, a highly efficient rectifier. The low forward voltage results from the fact that the conductivity of the I zone can be modulated by large amounts of charge carriers injected from the *p* and the *n* zones.

Another application of PIN diodes is the high-frequency (HF) field. Here, the fact is exploited that, due to the long carrier lifetime at frequencies beyond approximately 10 MHz, a rectifying effect will no longer occur and the PIN diode rather behaves like a real resistance, the magnitude of which depends on the forward direct current passed by the device and

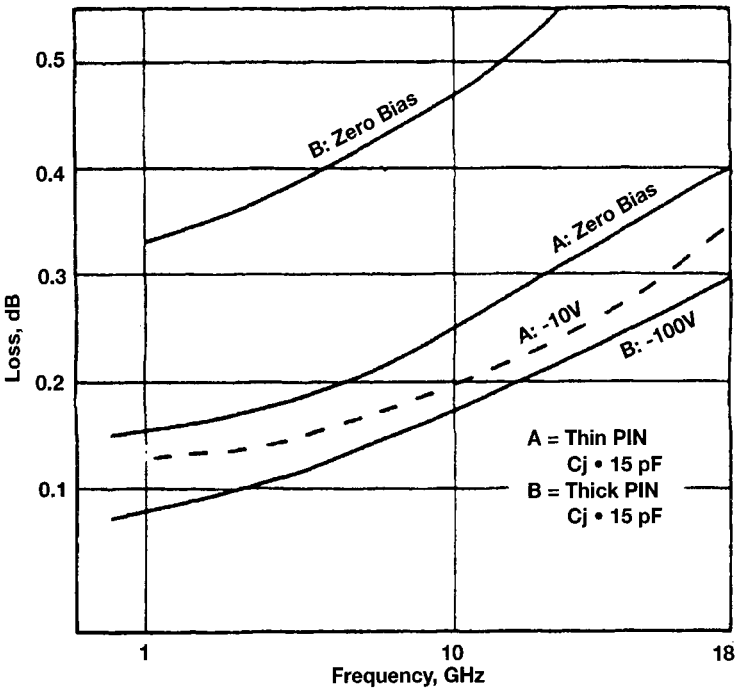


Figure 2-20 Insertion loss versus frequency.

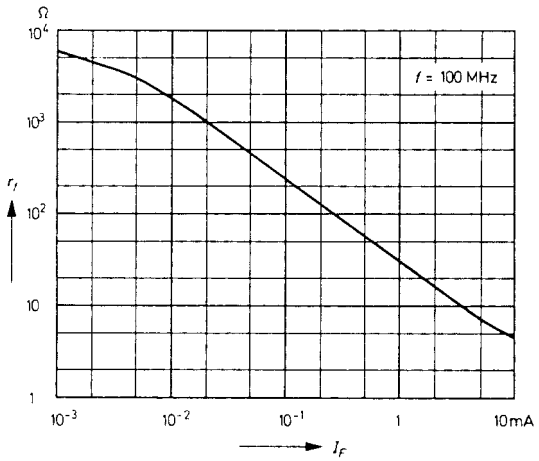


Figure 2-21 Forward resistance versus forward current.

produces an equal effect on both half waves of the HF signal. In view of this behavior, the PIN diode can be used as a switch or a variable resistor for HF signals. Thus it becomes possible, for example, to subject an HF signal to amplitude modulation by means of an AF-controlled PIN diode.

An important application of PIN diodes that has found favor in recent times is for dc-operated attenuators in TV tuners and antenna distribution amplifiers. Figure 2-21 shows the real HF forward resistance r_f as a function of the forward current I_f, measured at 100 MHz. Figures 2-22 and 2-23 show second-order IMD and cross-modulation for PIN diodes.

Applying the PIN in Amplitude Control of High-Frequency Signals. In conventional transistorized TV tuners, automatic gain control (AGC) is usually achieved by varying the emitter current of the input transistor. This method exploits the fact that the input transistor exhibits maximum gain at a certain level of emitter current and that this gain decreases when the emitter current is either increased or decreased relative to this point. Since modulation capacity and cross-modulation resistance grow with the emitter current, "upward" gain control has lately become the preferred solution, and according to it, gain reduction is achieved by an increase of the input transistor's emitter current. An input stage based on this design, therefore, is least resistant to cross-modulation when receiving a signal from a weak transmitter because it is operating at or below the point at which AGC action begins. This effect is particularly disadvantageous when the signal of a weak transmitter is to be received in the presence of a strong local transmitter. The unsatisfactory cross-modulation properties of such an input stage manifest themselves more and more frequently in the form of image perturbations as the TV reception band is crowded by an increasing number of transmitters. Another disadvantage is the variation in the transistor parameters that results from the control action. The variations affect antenna matching at the input and the response of the RF filter at the output.

To obviate these problems, methods being adopted today provide for the input stage to be equipped with cross-modulation-resistant transistors with fixed bias, gain control being provided by a variable attenuator preceding the transistor(s). The input-filter termination

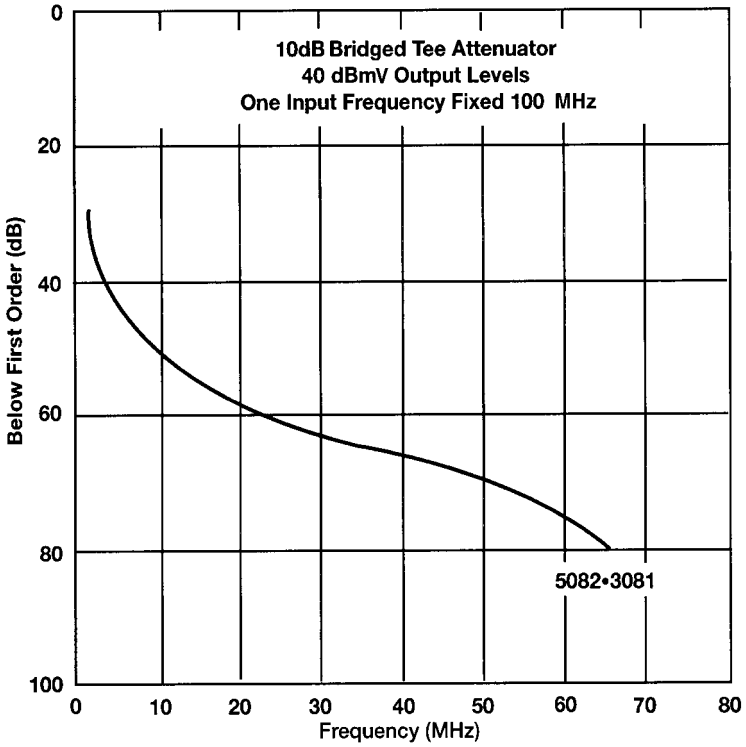


Figure 2-22 Second-order IMD in a PIN diode. Equation (1-48) can be used to determine the diode's IP_2 from the absolute values of its fundamental and IM_2 outputs.

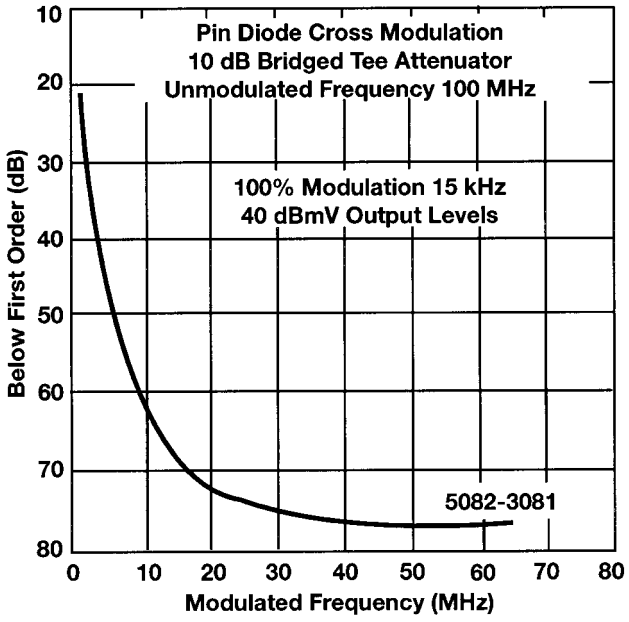


Figure 2-23 Cross-modulation in a PIN diode. Equation (2-61) relates cross-modulation level to IMD.

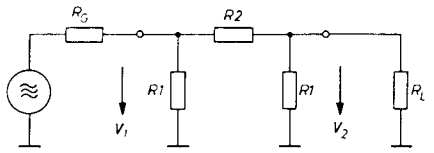


Figure 2-24 Basic circuit of a π-mode filter attenuator.

provided by this attenuator must be independent of the degree of attenuation to preserve the filter's response characteristics.

The first two requirements are met by the properties of the PIN diode and the latter requirement can be met by designing a PIN diode control in the form of a π network, a schematic of which is shown in Figure 2-24. The attenuation of a matched π network is expressed by the following equations:

$$\frac{a}{\text{dB}} = 20 \log \frac{Z + R_2}{R_2 - Z} \tag{2-40}$$

$$\frac{a}{\text{dB}} = 20 \log \left(\frac{R_1}{Z} + \sqrt{\frac{R_1^2}{Z^2} + 1} \right) \tag{2-41}$$

Figure 2-25 plots the relationship between the resistance levels of R_1 and R_2 and the input and output voltages of the π network. Figure 2-26 shows the circuit diagram of an implementation of the π network composed of PIN diodes. The resistance characteristic required according to Figure 2-25 is achieved approximately by varying the control current I_{co}.

If the control current I_{co} in Figure 2-26 equals zero, then—due to the fact that $R_1 = R_2$ and $R_3 = R_4$—the auxiliary voltage V_h causes forward currents of equal magnitude to flow through diodes D_1 and D_3. The voltage drops across R_1 and R_2 are of equal magnitude, so that no current passes through diode D_2. Therefore, the latter presents a high resistance, whereas D_1 and D_3 present a low resistance. Under these conditions, the network produces maximum attenuation. A control current I_{co} reduces the forward currents through D_1 and D_3 and allows a forward current to be passed by diode D_2. When it has attained its maximum value, diodes

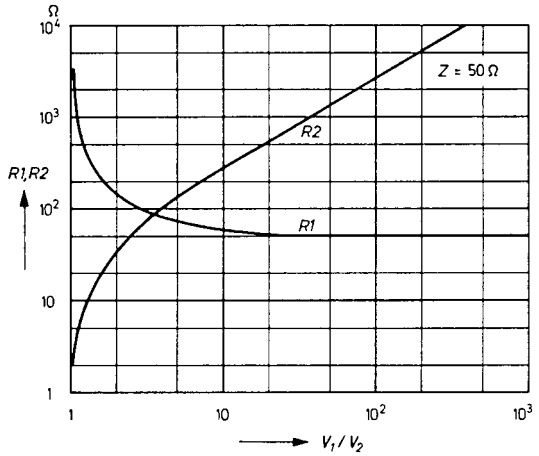


Figure 2-25 Values of R_1 and R_2 in Figure 2-24 versus relation of input and output voltage.

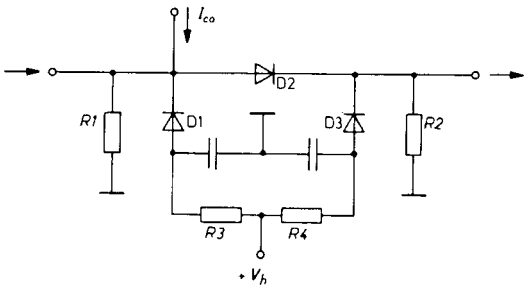


Figure 2-26 PIN diode π-mode attenuator.

D_1 and D_3 are blocked—that is, they are high-ohmic—and D_2 is low-ohmic. Under these conditions, the network produces its minimum attenuation.

If the maximum attenuation of the π network is to be increased, it is possible to make the two shunt diodes D_1 and D_3 more low-ohmic than the impedance Z, although this would reduce the reflection coefficient of attenuation. For example, if a reflection coefficient of about 7 dB is accepted over the entire control range, then a maximum attenuation of 25–30 dB could be achieved in the VHF range. The series inductances of the transverse diodes and the shunt capacitance of the series diode are responsible for this relatively unsatisfactory value. Improved attenuation is brought about with the integrated π network described in the next section.

Example: A PIN Diode π Network for TV Tuners. The since-discontinued TDA1053 is an example of an integrated PIN π network attenuator intended for TV-tuner use. It comprises three silicon planar PIN diodes connected to form a π network (Figure 2-27) and serves for the electronic amplitude control of the input signals of TV tuners and antenna distribution amplifiers in the 40–1000-MHz range. Its input and output impedances remain constant over the entire control range. This can also be achieved with discrete diodes.

The TDA1053 was normally supplied with vertical leads. The characteristics stated below apply to devices of this configuration.

MAXIMUM RATINGS OF INDIVIDUAL DIODES

Reverse voltage	V_R	30	V
Forward current at $T_{amb} = 25$ °C	I_F	50	mA
Junction temperature	T_j	125	°C
Storage temperature range	T_s	−55 to +125	°C

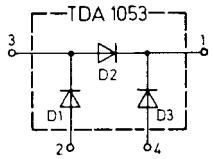


Figure 2-27 Internal circuitry of the TDA1053.

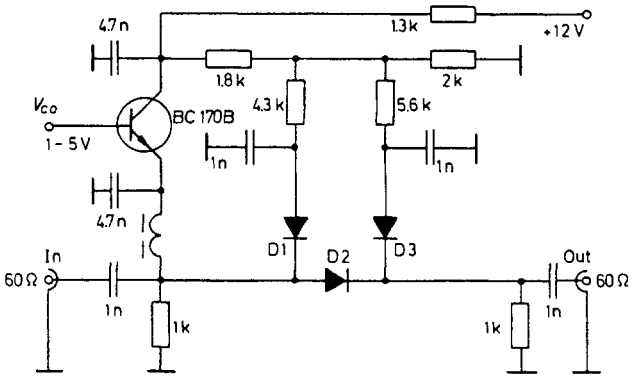


Figure 2-28 Test and application circuit.

MAXIMUM RATINGS OF THE π NETWORK

Ambient operating temperature range when operating according to Figure 2-28	T_{amb}	100	°C

CHARACTERISTICS OF INDIVIDUAL DIODES AT T_{amb} = 25 °C

Forward voltage at I_F = 50 mA	V_F	< 1.2	V
Forward current at V_R = 15 V	I_R	< 500	nA
Differential forward resistance:			
At I_F = 10 mA, f = 100 MHz	r_f	5	Ω
At I_F = 10 µA, f = 100 MHz	r_f	1.4	kΩ

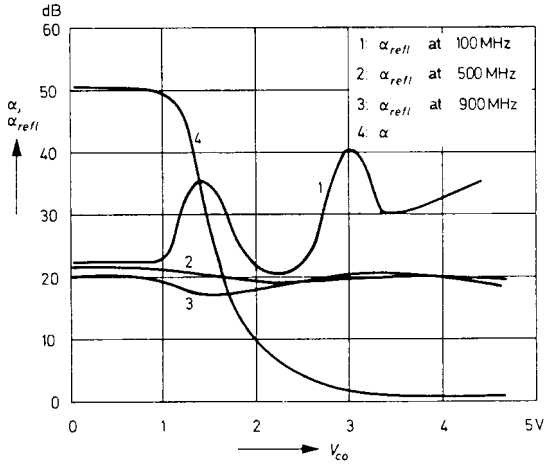


Figure 2-29 Attenuation and reflection attenuation versus control voltage.

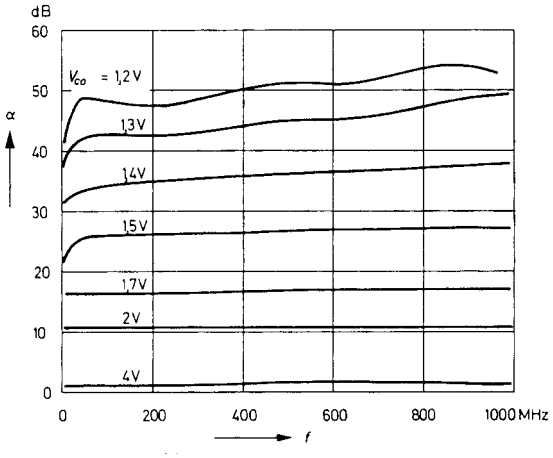


Figure 2-30 Attenuation versus frequency.

CHARACTERISTICS IN THE TEST CIRCUIT (FIGURE 2-28) AT T_{amb} = 25 °C

Voltage for 1% cross-modulation	V_{cr}	1	V
Attenuation in the 40–1000-MHz range:			
At V_{co} = 1.5 V (1–2 V)	a_{max}	45 (> 36)	dB
At V_{co} = 5 V (4–5 V)	a_{min}	1.5 (< 2)	dB
Reflection coefficient in the 40–1000-MHz range over the entire control range, depending on circuit design	a_{refl}	20 (> 16)	dB

These data reveal that the compact design of the three PIN diodes in a common 50B4 plastic package guarantees favorable values for minimum and maximum attenuation, as well as reflection attenuation. The test and application circuit shown in Figure 2-28 also comprises the transistor control-signal amplifier. The typical characteristic of the attenuation and the reflection attenuation for this circuit are shown in Figure 2-29, as a function of the control voltage V_{co}. Figure 2-30 shows the attenuation at different control voltages, as a function of frequency.

2-1-4 Tuning Diodes

Introduction. In recent years, continuous development of tuning diodes—also known as *varactors* or *varicaps*—together with increased commercial and military use has led to substantial improvement in Q, reproducibility, and reliability. Concurrently, new techniques for producing and controlling a hyperabrupt dopant profile in the semiconductor permit the capacitance–voltage law to be much faster than the classical square root or cube root behavior.

Current tuning diode materials include silicon and gallium arsenide; silicon is favored for low cost and lower Q applications from HF through microwave frequencies. Hyperabrupt varactors, also of silicon, are finding many applications in commercial television tuners, where their high tuning ratios, linear tuning, and low cost are needed. New developments include low-capacitance hyperabrupts for microwave and wireless applications.

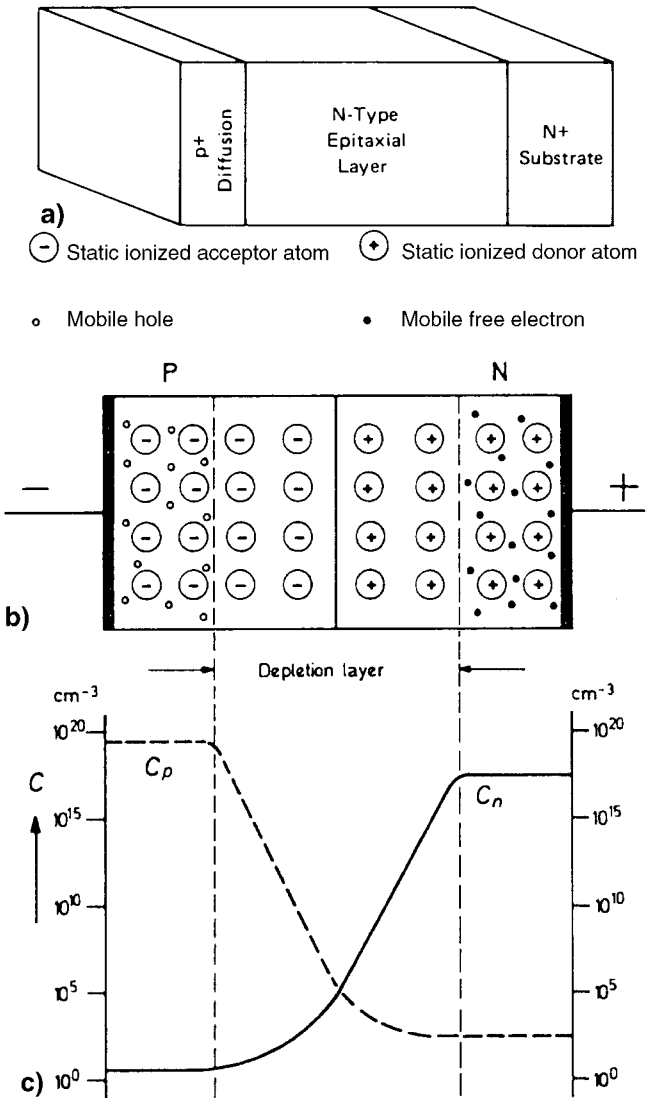


Figure 2-31 (a) Basic PIN structure; (b) cross section of a reverse-biased *pn* junction; and (c) density distribution of free charge carriers.

Gallium arsenide is used with high operating frequency and dictates the highest Q possible, as in parametric amplifiers and millimeter multipliers.

This section will acquaint the reader with tuning diodes: how they work, and what they can or cannot be expected to do in an electronic circuit. The basic properties of a tuning diode will be described in terms of the parameters that manufacturers use in characterizing them. The following topics will also be addressed:

- Capacitance ratio with respect to voltage and voltage breakdown
- Q as a function of design and operating conditions

- Stability—leakage current, temperature coefficient, and post-tuning drift
- Distortion products
- Packaging parasitics
- Applications—suggestions on how to specify a varactor

Tuning Diode Physics

Introduction. All junction diodes are made up of the same physical parts: a *pn* junction, a carefully controlled epitaxial layer, and a very-low-resistance substrate. These parts are shown in Figure 2-31.

No matter what type of junction device we are discussing—a tuning diode, a step-recovery diode, or a PIN diode—these parts are all present; the main difference between these devices is the resistivity and thickness of the epitaxial layer. Tuning diodes and multiplier diodes need epitaxial layers where both the resistivity and thickness are carefully controlled.

Abrupt Junction. An abrupt-junction diode is one in which the p^+ diffused region of the diode is much more highly doped than the epitaxial layer. Also, the high doping drops to the doping level of the epitaxial layer in a distance that is short compared to the epitaxial layer thickness, and the doping level of the epitaxial layer is constant over its thickness. This is shown in Figure 2-32, with the corresponding C–V curve shown in Figure 2-33. When these requirements are satisfied, the diode capacity, diode area, epitaxial layer doping level, and diode voltage are related by

$$\frac{C(V)}{A} = K\left(\frac{N}{V+\varphi}\right)^n \tag{2-42}$$

where $C(V)$ = capacitance of the diode at voltage V
A = area of the diode
N = doping level of the epitaxial layer

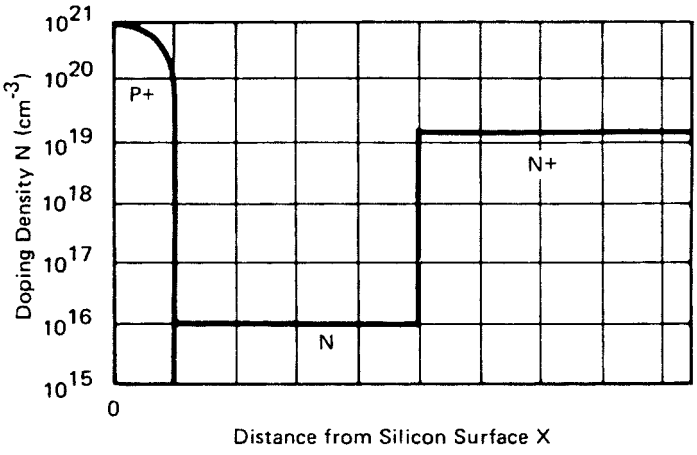


Figure 2-32 *N–X* abrupt junction.

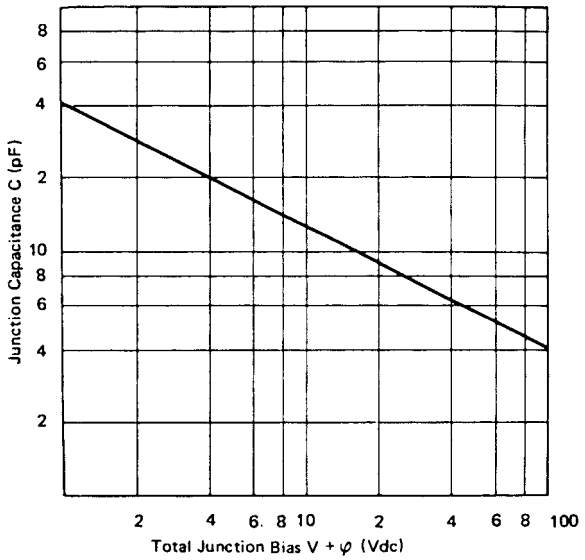


Figure 2-33 Capacitance versus total junction bias for abrupt-junction diode.

V = voltage applied to the diode
φ = built-in potential of the diode (0.6–0.8 V)
n = slope of diode C–V curve; $n \approx 0.5$ for an abrupt-junction diode
K = constant

As a consequence of the physical properties of a *pn* junction, a depletion layer is formed between the *p* and *n* regions whose width depends on the voltage applied to the diode. The capacitance of the diode is inversely proportional to the width of the depletion layer; that is, $C \sim 1/\ell_0$. In addition, the series resistance of the diode is proportional to the width of the undepleted epitaxial layer. Thus, as diode reverse bias is increased, the depletion layer increases, causing a decrease in capacitance and a decrease in series resistance. As the diode reverse bias is increased further, a point is reached where the electric field caused by the reverse bias reaches a critical level, and current through the diode increases rapidly; this is

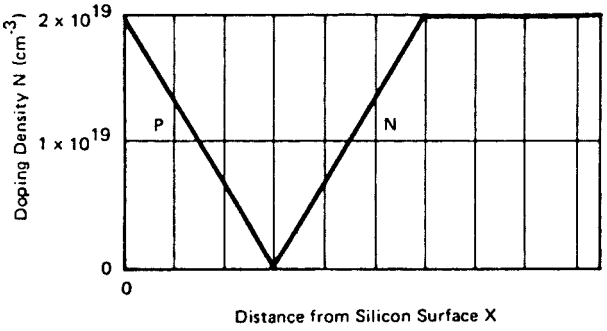


Figure 2-34 *N–X* linearly graded junction.

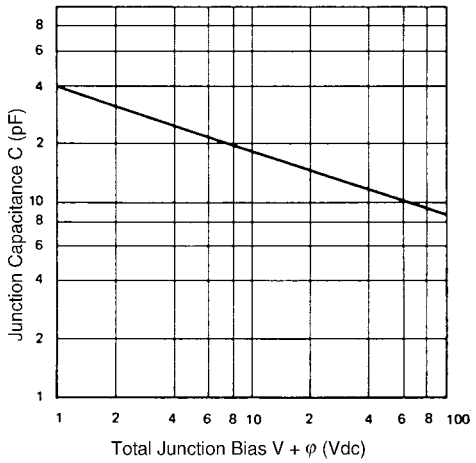


Figure 2-35 Capacitance versus total junction bias for linearly graded junction.

breakdown voltage of the diode. If, at the breakdown voltage, the epitaxial layer is not completely depleted, the diode will have excessive series resistance. Conversely, if the epitaxial layer is depleted before the breakdown voltage is reached, no further capacitance decrease occurs after the total depletion, and a condition called *punchthrough* occurs.

While, in the ideal case, voltage breakdown will occur just as the epitaxial layer is totally depleted, this seldom occurs in practice, and we generally have a condition of either punchthrough or excess series resistance.

Linearly Graded Junction. If, instead of the junction profile shown in Figure 2-32, we have a p^+ region and an n region whose doping levels increase linearly with distance from the pn junction as shown in Figure 2-34, with its corresponding C–V curve in Figure 2-35, then we have what is called a linearly graded junction diode. This diode follows Eq. (2-42) with the exception that the exponent n is equal to $\frac{1}{3}$. This means that, for a given voltage change, the linearly graded junction will have a smaller capacitance change than an abrupt-junction diode. Since, in most cases, the designer is looking for the maximum

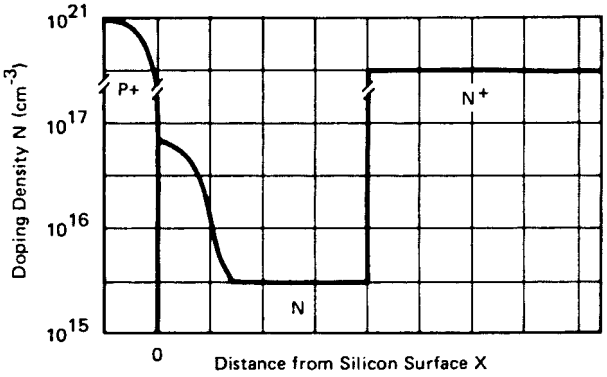


Figure 2-36 *N–X* hyperabrupt-junction diode.

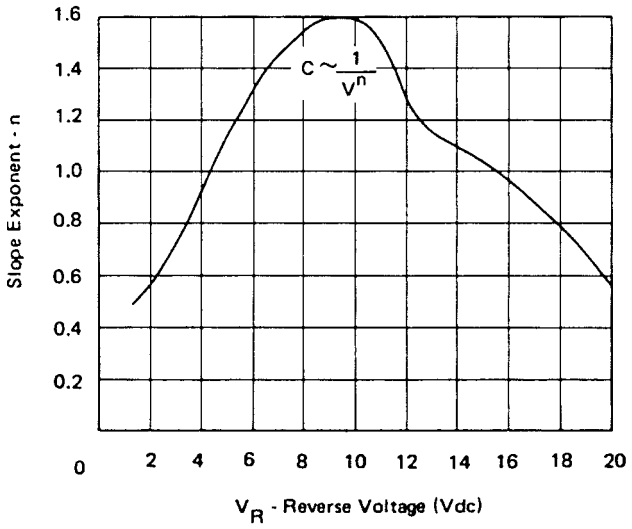


Figure 2-37 Typical *n* versus reverse voltage characteristic for hyperabrupt-junction diode.

capacitance change obtainable, the linearly graded junction is not used as a tuning diode. This structure found its greatest use several years ago as a "cube law" multiplier, but even this use has decreased as new structures have been developed.

Hyperabrupt Junction. The hyperabrupt-junction diode provides a greater capacitance change than the abrupt-junction diode for a given voltage change, as well as a linear frequency versus voltage characteristic over a limited voltage range. The structure of the hyperabrupt-junction diode is shown in Figure 2-36 and can be seen to be an abrupt-junction diode with an additional, increased doping level at the *pn* junction. This diode also follows Eq. (2-42) with the exception that n is not a function of voltage and is generally in the range of 0.5–2. A typical curve of n versus voltage is shown in Figure 2-37.

The C–V curve in a hyperabrupt-junction diode is shown in Figure 2-38 and is seen to start at a high value of capacitance per unit area at low bias (high epitaxial doping) and change to a lower value of capacitance per unit area (low epitaxial doping) at high bias. The details of the curve depend on details of the shape of the more highly doped region near the *pn* junction.

Unfortunately, with a hyperabrupt-junction diode, you must settle for a lower Q than an abrupt-junction diode with the same breakdown voltage and same capacitance at 4 V.

It should be noted that any diode that has an n value that exceeds 0.5 at any bias voltage is, by definition, a hyperabrupt-junction diode. Thus, the hyperabrupt-junction diode family can have an infinite number of different C–V curves. Since the abrupt-junction diode has a well-defined C–V curve, the capacitance value at one voltage is sufficient to define the C–V capacitance at any other voltage. This is not the case for the hyperabrupt-junction diode. In order to adequately define the C–V characteristics of a hyperabrupt-junction diode, two and sometimes three points on the curve must be specified.

Silicon Versus Gallium Arsenide. Everything mentioned so far applies to both silicon and gallium arsenide (GaAs) diodes. The main difference between silicon and GaAs from a

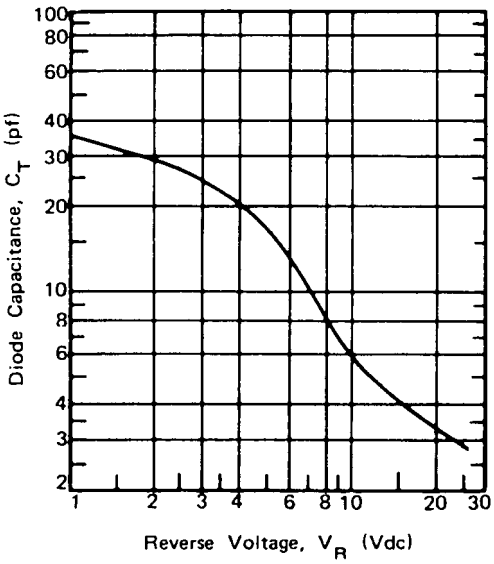


Figure 2-38 Capacitance versus junction bias for hyperabrupt-junction diode.

user's point of view is that higher Q can be obtained from GaAs devices. This is due to the lower resistivity of GaAs from a given doping level N. The resistivity of the epitaxial layer, or substrate, of a diode is given by

$$\rho = \frac{1}{Ne\mu} \tag{2-43}$$

where ρ = the resistivity
N = the doping level of the layer
e = the charge on an electron
μ = is the mobility of the charge carriers in the layer

Gallium arsenide has a mobility about four times that of silicon and, thus, a lower resistivity and higher Q for a given doping level N. Since diode capacitance is proportional to $\sqrt{N}$, independent of resistivity, a silicon diode and a GaAs diode of equal area and doping will have a capacitance difference proportional to the square root of the dielectric constant ratio. This gives the GaAs diode a 5% higher capacitance and is thus of little practical significance. The penalty paid for using GaAs is an unpassivated diode and a more expensive diode due to higher material and processing costs. If the higher Q of the GaAs device is not really needed, a substantial price saving will be obtained by using a silicon device.

Planar Versus Mesa Construction. The two basic construction techniques used to manufacture tuning diodes are planar and mesa; cross sections for both of these devices are shown in Figure 2-39. The planar process, which is the backbone of the integrated circuit industry, lends itself to large-volume production techniques and is the one used for the 1N series of tuning diodes. Mesa processing, on the other hand, requires more

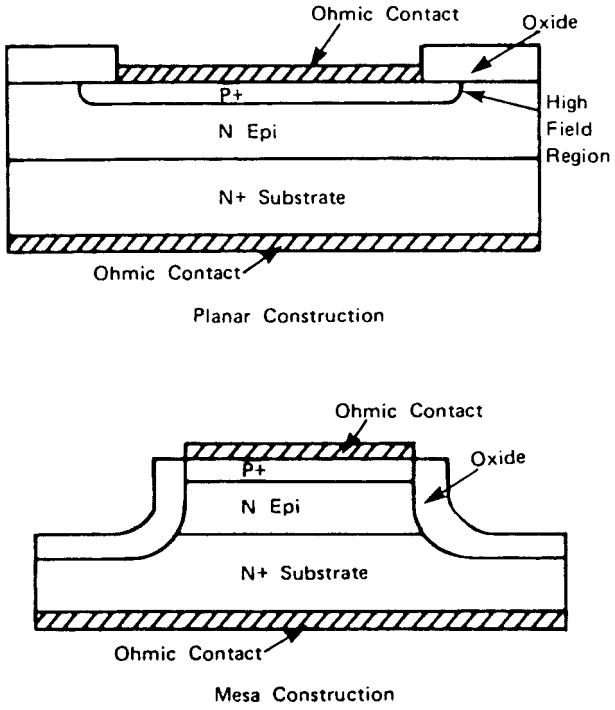


Figure 2-39 Cross sections of planar and mesa devices.

processing steps and is generally done on a wafer-by-wafer basis. This results in a more costly process and thus a more expensive diode. All microwave tuning diodes are of mesa design because of greatly higher Q. Due to the relatively small radius of curvature at the junction edge of a planar diode, the electric field in this area is greater than the electric field in the center (flat) portions of the junction. As a result, the breakdown voltage of the diode is determined by both the epitaxial resistivity and the radius of curvature of the junction edge. Thus, for a given breakdown voltage, a planar diode must use higher resistivity epitaxial material than a mesa diode, which has a completely flat junction. The end result is that the planar diode has a greater series resistance than a mesa diode for the same capacitance and breakdown voltage, and thus lower Q.

Capacitance

Capacitance Ratio. From the user's point of view, the capacitance ratio is simply the capacitance change available in the circuit. Thus, a user tuning from, say, –4 to –45 volts defines capacitance ratio as

$$R = \frac{C_T(-4)}{C_T(-45)} \tag{2-44}$$

where C_T includes C_J plus C_P plus C_F. The manufacturer, however, defines C_T as $C_J + C_P$.

To explore the significance of this difference, let's take two examples, one with a large C_J and another with a small C_J, in chip, package, and "typical" fringe situations. Both are 45-V tuning varactors.

Device	C_{J0}	C_{J45}	Ratio
A	0.6 pF	0.1 pF	6.0
B	15.0 pF	2.5 pF	6.0

Put both devices in a standard 023 package with C_P (strap and ceramic) of 0.18 pF:

Device	C_{T0}	C_{T45}	Ratio
C	0.78 pF	0.28 pF	2.75
D	15.18 pF	2.68 pF	5.67

Notice the drop in capacitance ratio, especially for the low-C_J diode. If we now add a typical 0.04 pF for external fringe capacitance, we get the following results:

Device	C_{T0}	C_{T45}	Ratio
C	0.82 pF	0.32 pF	2.56
D	15.22 pF	2.72 pF	5.6

The reduction in capacitance ratio, and thus circuit tuning capability, by the fringing fields is quite obvious and amounts to 7% in this example.

Because of the often stringent specifications on tuning ratio, it is mandatory that the manufacturer and customer clearly agree on the exact design of the holder used to measure the varactor in question.

Having described how to measure capacitance, it is relatively easy to describe the results. The section on diode physics described the various types of "laws," or C–V curves, and we will not repeat them here. Nonetheless, several important points must be covered.

The first is "available capacitance swing." The laws indicate a steadily decreasing capacitance with voltage, which indicates that the epitaxial region is widening and the electric field is increasing. (For an abrupt junction, since $C \propto 1/\sqrt{V}$, the depletion zone with W is increasing as $\sqrt{V}$, and the electric field V/W increases as $\sqrt{V}$.)

Two things can happen:

1. The junction width widens so that the entire intrinsic region is depleted. The capacitance bottoms out, resulting in voltage punchthrough.
2. The electric field exceeds the dielectric strength of silicon (or GaAs), and "solid-state discharge" or "avalanche" current is drawn.

The diode impedance drops, the varactor no longer "varacts," and circuit operation ceases. Moreover, if more than a few milliamperes of current is drawn, localized overheating may destroy the diode, resulting in breakdown voltage. All varactors are characterized for breakdown voltage—for example, 45 V minimum.

The theoretical tuning varactor is designed so that the punchthrough occurs at a voltage equal to the voltage breakdown of the diode. Logically, this means that in order to obtain

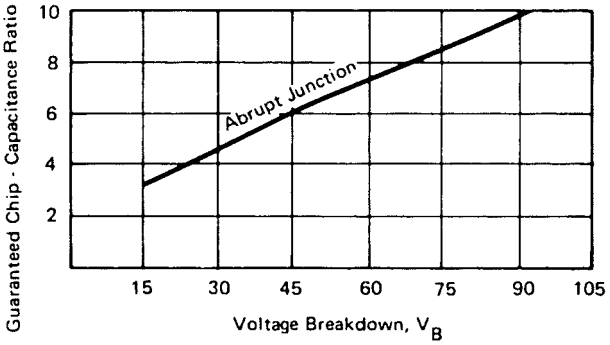


Figure 2-40 Capacitance ratio versus breakdown voltage.

greater tuning ratios, it is necessary to be able to increase the depletion-layer width without reaching punchthrough or breakdown. You must have a thicker epitaxial region to make this possible.

Figure 2-40 shows catalog ratio values, from zero bias to breakdown, as a function of breakdown voltage necessary. In the next section, on Q, we will discuss other elements in your choice of V_B.

To complete this section, we should mention that semiconductor processing control has been refined so well that capacitance tracking to within ±1% over the full range from zero to breakdown is now readily obtainable in production quantities.

Temperature Coefficient of Capacitance (T_{CC}). Unfortunately, since most datasheets give the value of T_{CC} at 4 V, it is sometimes assumed that this value applies at all bias voltages. This is not the case. Consider Eq. (2-45), a rewritten form of Eq. (2-42):

$$C(V) = \frac{C(O)}{(V + \varphi)^n} \tag{2-45}$$

Taking the derivative of this with respect to temperature T we have

$$\frac{dC(V)}{dT} = \frac{+nC(O)}{(v + \varphi)(v + \phi)^n} \frac{d\varphi}{dT} \tag{2-45a}$$

or, after substituting Eq. (2-45):

$$T_{CC} = \frac{1}{C(V)} \frac{dC(V)}{dT} = \frac{-n}{(V + \varphi)} \frac{d\varphi}{dT} \tag{2-46}$$

As a first approximation, we can say that $d\varphi/dT = -2.3$ mV/°C over the temperature range of interest.

From Eq. (2-46) we can draw the following conclusions:

1. The temperature coefficient is inversely proportional to the applied voltage.
2. The temperature coefficient is directly proportional to the diode slope, n.

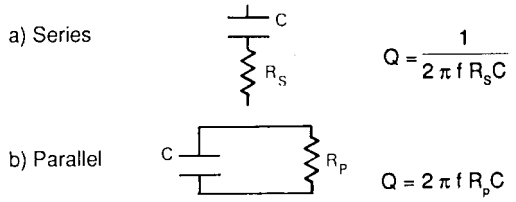
For an abrupt-junction diode that has a constant value of n (0.5), the temperature coefficient has a smooth curve in the form $K/(v + \varphi)$. However, in the case of hyperabrupt-junction diodes, n is a function of voltage, and the shape of the T_{CC} curve depends on the details of the $n(V)$ curve. A typical T_{CC} curve for an abrupt-junction diode is shown in Figure 2-41 and for an Alpha DKV6520 series hyperabrupt-junction diode in Figure 2-42. The inflection in the hyperabrupt T_{CC} is due to the fact that in this voltage range $n(V)$ is increasing faster than $1/V$, giving an increase in T_{CC}. It should also be noted, however, that over the range of the T_{CC} minimum the temperature coefficient is relatively constant, and operation in this area may be advantageous in some applications where a restricted tuning range can be used.

Q Factor or Diode Loss

Definitions. The classical definition of the Q of any device or circuit is

$$Q = \frac{2\pi \times \text{Energy stored}}{\text{Energy dissipated per cycle}} \tag{2-47}$$

For a capacitor, two formulations are possible:



Clearly, the two definitions must be equal at any frequency, which establishes

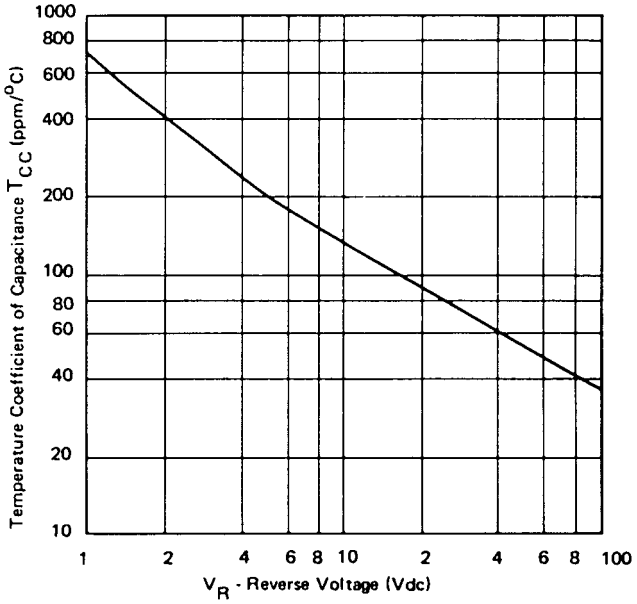


Figure 2-41 Temperature coefficient of capacitance versus tuning voltage for an abrupt-junction diode.

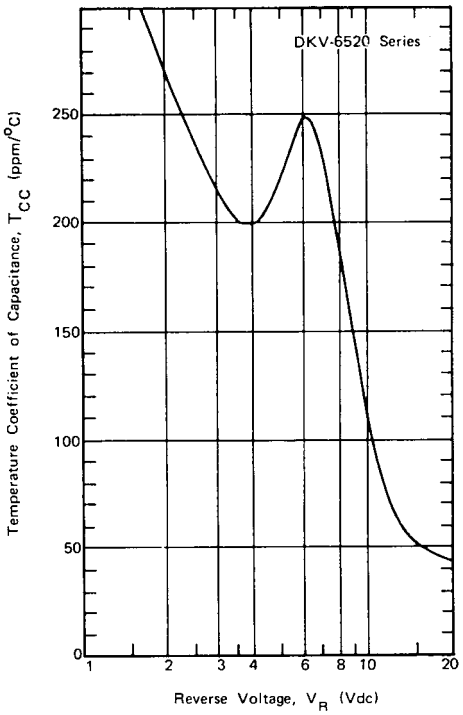


Figure 2-42 Temperature coefficient of capacitance versus tuning voltage ($T_A = 25$ °C) hyperabrupt-junction diode.

$$R_P = \frac{1}{(2\pi f)^2 C^2 R_S} \tag{2-48}$$

In the case of a high-Q tuning diode, the proper physical model is the series configuration, because the depleted region is almost perfectly pure capacitance, and the undepleted region, due to its relatively low resistivity, is almost a pure resistor in series with the capacitance. Furthermore, the contact resistances are also clearly in series. Q, then, for a tuning diode, is given by

$$Q_{(-v)} = \frac{1}{2\pi f_o R_{(-v)} C_{(-v)}} \tag{2-49}$$

where f_o is the operating frequency, $C_{(-v)}$ is the junction capacitance, and $R_{(-v)} = R(\text{epi}) + R_{C'}$, the sum of the resistance of the undepleted epitaxial layer and the fixed contact resistance.

Cutoff frequency, f_c, is defined as that frequency at which Q equals unity. Thus,

$$f_{c_{(-v)}} = \frac{1}{2\pi R_{(-v)} C_{(-v)}} \tag{2-50}$$

Historically, the tuning diode business has specified Q at 50 MHz, despite the fact that Q values of microwave diodes are so high that it is almost impossible to measure them at 50

Table 2-7 Parameters for 0.6-pF diode

Breakdown Voltage V_B	Resistivity ($\Omega \cdot$ cm)	Junction Diameter (mills)	Epitaxial Region (μm)	Depleted Epitaxial $V = -4$ V (μm)	Undepleted Epitaxial $V = -4$ V (μm)	R Undepleted (Ω)	R_{sp} (Ω)	Q_4
30	0.31	2.4	1.37	0.54	0.83	0.81	0.18	5200
45	0.52	2.8	2.25	0.73	1.52	1.86	0.15	2600
60	0.74	3.2	3.20	0.90	2.30	3.24	0.13	1500
90	1.25	3.7	5.27	1.21	4.06	7.17	0.10	700

MHz. Instead, as discussed below, Q is measured at microwave frequencies (e.g., 1–3 GHz) and related to 50 MHz by the relationship

$$Q_{(f1)} = Q_{(f2)} \frac{f_2}{f_1} \tag{2-51}$$

which derives quickly from the assumption that f_c is independent of the measuring frequency.

Since both junction capacitance and epitaxial layer resistance are functions of the applied bias, it is not possible to calculate Q as a function of bias from a measurement of capacitance alone. Catalog specifications typically show Q at –4 V, together with the capacitance at two or more voltages.

Relative to $Q_{(-4)}$, Q increases faster than the reduction in capacitance for bias greater than 4 V and, conversely, decreases faster for bias less than 4 V.

In the following section, we will discuss the diode design parameters that determine Q. Following this, we will describe some elementary Q measurement techniques.

Causes. In the discussion of device physics, the resistivity of the epitaxial region was discussed, together with its impact on punchthrough and breakdown. For example, Table 2-7 supplies typical resistivity and relative parameters of 0.6 pF ($C_{J(4)}$) diodes of different breakdowns. If we remember that if at any bias lower than breakdown the epitaxial region is not completely depleted, it follows that the undepleted portion presents a resistance in series with the pure capacitance of the depleted zone. The magnitude of this "undepleted" resistance is also shown in the table.

The entry R_{sp} ($R_{spreading}$) is the series resistance between the epitaxial region and the low-resistivity substrate. The calculations are for idealized cylindrical epitaxial regions of uniform resistivity, low-resistivity contact on the anode (top), and low-resistivity substrate on the cathode. This resistance is constant, independent of bias; also shown are epitaxial thickness region and width of the depletion zone at –4 V bias.

Note the substantial reduction in Q for high-voltage diodes caused by the increased epitaxial layer resistance; this is true for any value of capacitance or any type of junction. For greater voltage breakdown, the epitaxial region thickness must be increased, which requires an increase in epitaxial layer resistivity; the higher the resistivity of the undepleted zone, multiplied by the fact that it is much wider for high-voltage diodes, means the resistance increases substantially.

Consequently, a rule of thumb emerges: For maximum Q, never choose a tuning diode with a voltage breakdown in excess of what is needed for the necessary tuning range. If the

Table 2-8 ***Q*** **versus bias for 0.6-pF diode**

Breakdown Voltage V_B	C_{J0}	$Q(0)$	$C_J(-4)$	$Q(-4)$	$C_J(-10)$	Q (−10)	$C_J(-30)$	$Q(-30)$	$C_J(-45)$	$Q(-45)$	$C_J(-60)$	$Q(-60)$	$C_J(-90)$	Q(−90)
30	1.43	1700	0.6	5200	.4	10k	0.23	64k						
45	1.43	850	0.6	2600	.4	5k	0.23	20k	0.19	90k				
60	1.43	550	0.6	1500	.4	3k	0.23	9k	0.19	23k	0.17	170k		
90	1.43	270	0.6	700	.4	1.3k	0.23	3.5	0.19	6k	0.17	10k	0.14	220k

Table 2-9 Capacitance ratios (C_{J0}/C_JV_B)

Breakdown Voltage V_B	Optimum Ratio	Minimum Guaranteed Ratio	$Q_{(-4)}$ Typical
30	6.2	4.5	3000
45	7.5	6.0	2500
60	8.4	7.5	1400
90	10.2	8.7	650

required tuning range is an octave, requiring a 4:1 ratio, the selection of a 30-V diode will result in diode losses half those of a 60-V diode.

Table 2-8 lists capacitance and Q for each of these chips as a function of bias. Please remember that Q is calculated at 50 MHz.

Table 2-9 rewrites the data of Table 2-8 to show available capacitance ratios between zero bias and breakdown. The first column is the theoretical optimum, as tabulated. The second column is the typical catalog specification. The reduction in tuning ratio below theoretical optima is caused by nonideal junction fabrication. The junctions are never perfectly abrupt.

Although the tables and numbers given refer to abrupt-junction silicon devices, the principle applies without exception to all types of tuning diodes. For comparison, Table 2-10 lists available ratios and Q values for a number of different varactors. The high Q values for GaAs and the low values for hyperabrupts are apparent.

One last point: The Q values and series resistance refer to chips only. The effects of package parasitics will be discussed later, but it is important to consider circuit contact losses here. For low-capacitance diodes—for example, $C_{J4} = 0.6$ pF—the epitaxial region contributes a high value of resistance and dominates Q except at punchthrough. Diode contact losses are less significant.

Post-Tuning Drift. Post-tuning drift (PTD) is the change in oscillator frequency with time after the tuning voltage has stabilized. The minimization of PTD has assumed greater importance with the design of more sophisticated electronic countermeasure systems, where rapid, accurate frequency changes are required.

Table 2-10 Comparative Tuning Diodes

Type	C_{J0}	Breakdown Voltage V_B	$Q_{(-4)}$[a]	Ratio C_{J0}/C_JV_B[a]
Silicon abrupt	1.0	30	5,000	4.5
Silicon abrupt	2.5	30	4,600	4.5
Silicon abrupt	5.0	30	3,800	4.5
Gallium arsenide abrupt	1.0	25	10,000	3.6
Gallium arsenide abrupt	0.5	10	17,000	2.5
Silicon hyperabrupt	50.0	22	300	17.0
Silicon hyperabrupt	2.5	22	500	14

[a] Minimum guaranteed.

Post-tuning drift can be characterized as short-term and long-term. Short-term PTD occurs in the time range of tens of nanoseconds to a few seconds, while long-term PTD occurs in the time range of seconds to minutes, hours, or days.

Short-term PTD is mainly dependent on the thermal properties of the diode and is improved by high Q (low power loss) and flip chip construction. Long-term PTD depends on oxide stability and freedom of mobile charge in the oxide. It should be noted that actual oscillation frequency change may occur even with a perfect tuning diode because of variation with frequency in the power dissipated by the diode, changes in the diode heat-sink temperature, and frequency changes due to other circuit elements. Less than 0.01% short- and long-term PTD can be obtained.

Distortion Products. Inasmuch as nonlinear components generate harmonics and other distortion products, an understanding of this mechanism is of prime interest to the circuit designer. In some instances, the distortion products are the desired end result of the circuit design, as in frequency multipliers, where harmonics of the input signal frequency are the required output signal. For other applications, such as tuning-diode-tuned linear circuits, distortion products are extremely undesirable, and in some instances the end product specification may set a maximum limit on the distortion products allowed.

Cross-Modulation. Cross-modulation is the transfer of the modulation on one signal to another signal and is caused by third-order and higher odd-order nonlinearities in the behavior of the device. Rewriting Eq. (2-42) we have

$$C(V) = \frac{C_0}{(1 + V/\varphi)^n} \tag{2-52}$$

where C_0 = capacitance
V = applied voltage = $V_0 + v$
V_0 = dc applied voltage
v = ac applied voltage

Then, for a desired signal of

$$S_1 = v_1 \sin \omega_1 t \tag{2-53}$$

and a second, amplitude-modulated signal of

$$S_2 = v_2(1 + m \cos \omega_m t)\sin \omega_2 t \tag{2-54}$$

it can be shown that the cross-modulation, γ, defined by

$$\text{Output signal} \sim v_1 \sin \omega_1 t + \gamma \sin (\omega_1 \pm \omega_m)t \tag{2-55}$$

is found to be

$$\gamma = \frac{n(n+1)mv_2^2}{4(V_0+\varphi)^2} \tag{2-56}$$

From this equation it can be seen that cross-modulation is:

- Proportional to the square of the interfering signal
- Directly proportional to the interfering signal's modulation index, m
- Independent of the strength of the desired signal
- Independent of the frequencies of the desired and interfering signals (assuming that the nonlinearity to which both signals are subjected is sufficiently frequency-indiscriminate so this is the case)
- Present for all values of n; that is, no value of n gives zero cross-modulation

Solving Eq. (2-56) for the signal level v_2 required to produce cross-modulation of value γ, we have

$$v_2 = \frac{2(V_0+\phi)\gamma}{n(n+1)^m} \tag{2-57}$$

The interfering signal levels required to produce 1% cross-modulation from a 30%-modulated interfering signal applied to an abrupt-junction diode and a hyperabrupt-junction diode are shown in Figure 2-43.

From this figure, it can be seen that the hyperabrupt-junction diode is more susceptible to cross-modulation than the abrupt-junction diode in the region of maximum slope of the

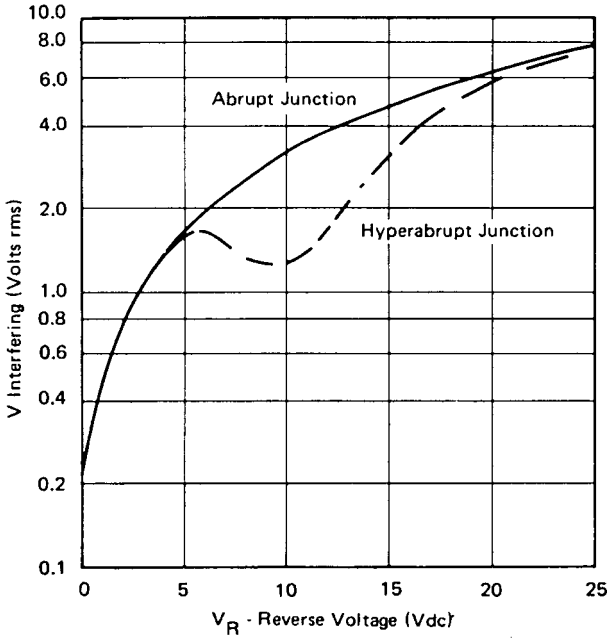


Figure 2-43 Interfering signal (30% amplitude modulated) level versus bias for 1% cross-modulation for abrupt-junction and hyperabrupt-junction diodes.

hyperabrupt-junction diode. For many applications, however, distortion products will be generated by other devices, such as transistors, at signal levels considerably below those given in Figure 2-43.

Intermodulation. Intermodulation is the production of undesired frequencies in the form

$$\sin(2\omega_1 t - \omega_2 t) \quad \text{and} \quad \sin(\omega_1 t - 2\omega_2 t) \tag{2-58}$$

from an input signal in the form of

$$v(\cos \omega_1 t + \cos \omega_2 t) \tag{2-59}$$

From an analysis similar to that done for cross-modulation, it can be shown that

$$\text{Intermodulation} = \frac{n(n+1)v^2}{8(V_0 + \varphi)^2} \tag{2-60}$$

or

$$\text{Cross-modulation} = (2m) \times \text{Intermodulation} \tag{2-61}$$

Harmonic Distortion. Harmonic distortion products are integral multiples of the signal frequencies and decrease in amplitude as the harmonic number decreases. Due to passband considerations and amplitude decrease with harmonic number, the second harmonic is the one of prime concern. Again, it can be shown that the second harmonic, v_2, of a signal of amplitude v_1 is

$$v_2 = \frac{n}{3(V_0 + \phi)} v_1^2 \tag{2-62}$$

Figure 2-44 shows the signal level required to produce 10% second-harmonic distortion in an abrupt-junction and a hyperabrupt-junction diode. Again, as in the case of cross-modulation, the hyperabrupt-junction diode is slightly worse than the abrupt-junction diode in the region of maximum slope of the hyperabrupt-junction diode.

Reduction of Distortion Products. In some cases, the signal levels applied to the diode generate distortion products larger than desirable for the circuit application. In this case, significant reduction in the distortion products can be achieved by using two diodes in a back-to-back configuration, as shown in Figure 2-45. Analysis shows that the fundamental signal components through the diodes are in phase and add, while some distortion products are out of phase and cancel, thus improving distortion performance.

Since the gradient of the electrical field produced in the depletion layer by a reverse bias applied to the device is proportional to the space charge density, the following equations can be written for the junction width W as a function of the reverse bias V_R:

For an Abrupt *pn* Junction (Alloyed Diodes)

$$W = {}_2\sqrt{2\frac{\varepsilon_r \varepsilon_0}{q}\left(\frac{1}{C_p} + \frac{1}{C_n}\right)(V_R + V_D)} \tag{2-63}$$

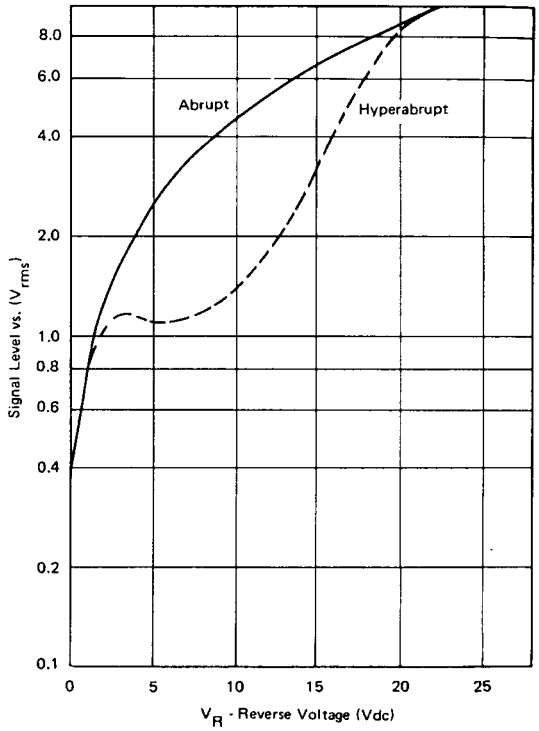


Figure 2-44 Signal level versus reverse voltage for 10% harmonic distortion.

For Linear *pn* Junctions (Single-Diffused Diodes, such as BA110–BA112)

$$W = \sqrt[3]{12\frac{\varepsilon_r \varepsilon_0}{aq}(V_R + V_D)} \tag{2-64}$$

where a is the impurity gradient within the depletion layer, ε_0 is the absolute dielectric constant ($8.85 \times 10^{-14}\ A_s/V_{cm}$, and $\varepsilon_r \approx 12$, the relative dielectric constant of silicon.

The junction capacitance, which is inversely proportional to the junction width, therefore varies in alloyed diodes with the square root, and in single-diffused diodes with the cube root of the externally applied reverse bias, and can be calculated from the general equation

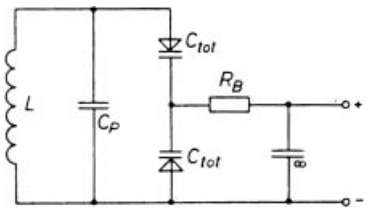


Figure 2-45 Back-to-back diodes. C_P is a fixed parallel capacitance, R_B is a bias decoupling resistor, and the capacitor marked ∞ is a low-impedance bypass.

$$C = \frac{\varepsilon_r \varepsilon_0 S}{W} \tag{2-65}$$

where S is the surface area of the *pn* junction. By way of approximation, we can also use the equation

$$C = \frac{K}{(V_R + V_D)^n} \tag{2-66}$$

where all constants and all parameters determined by the manufacturing process are contained in K. The exponent n is a measure of the slope of the capacitance/voltage characteristics and is 0.5 for alloyed diodes, 0.33 for single-diffused diodes, and (on average) 0.75 for tuner diodes with a hyperabrupt *pn* junction. Figure 2-46 shows the capacitance/voltage characteristics of an alloyed, a diffused, and a tuner diode.

Recently, an equation has been indicated which, although purely formal, describes the practical characteristics better than Eq. (2-66):

$$C = C_0 \left(\frac{A}{A + V_R} \right)^m \tag{2-67}$$

where C_0 is the capacitance at $V_R = 0$, and A is a constant whose dimension is a voltage. The exponent m is much less dependent on voltage than the exponent n in Eq. (2-66).

Equations (2-63) through (2-67) express the pure junction capacitance of the capacitance diode, but to this must still be added a constant capacitance, determined by structure parameters, in order to obtain the diode capacitance C_{tot}, which interests the user. With high inverse voltages—that is, low junction capacitance—a difference will therefore arise between the theoretical capacitance/voltage characteristic according to Eq. (2-66) and the practical characteristic, as shown in Figure 2-47.

The operating range of a capacitance diode or its useful capacitance ratio

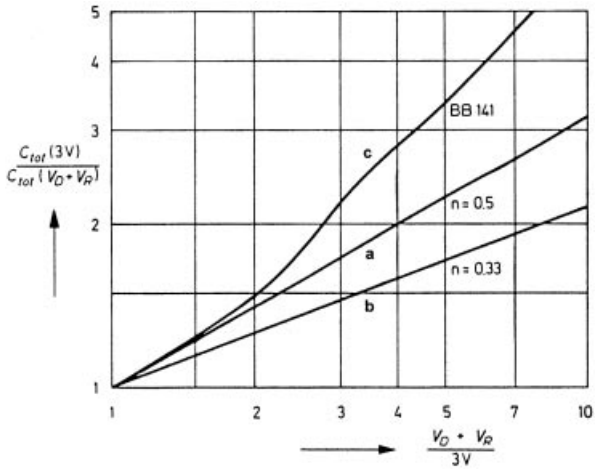


Figure 2-46 Capacitance/voltage characteristic for (a) an alloyed capacitance diode, n = 0.5; (b) a diffused capacitance diode, n = 0.33; and (c) a wide-range tuner diode (BB141).

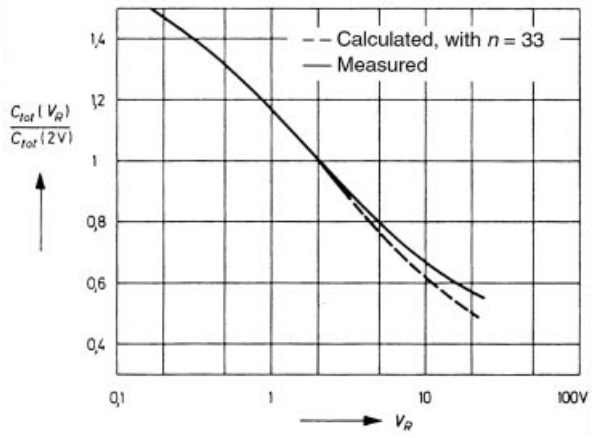


Figure 2-47 Capacitance/voltage characteristic of the BA110 diode.

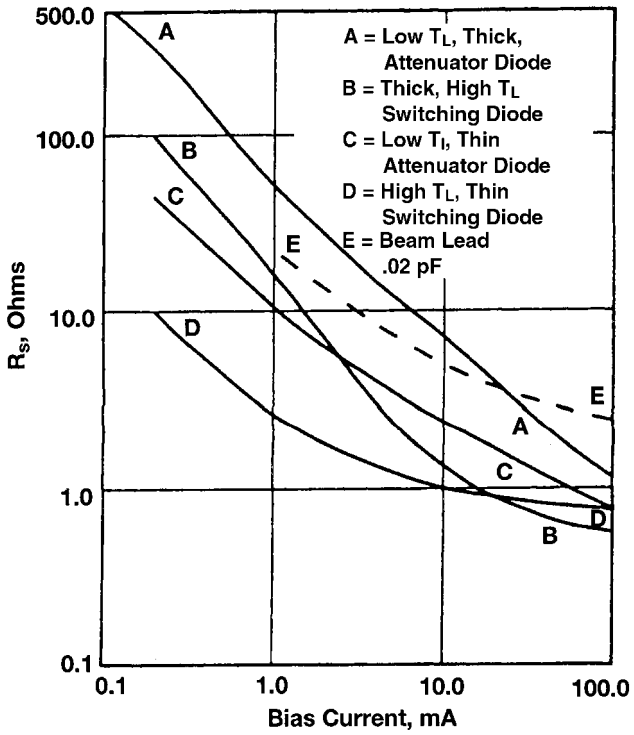


Figure 2-48 Basic current/voltage and capacitance/voltage characteristics.

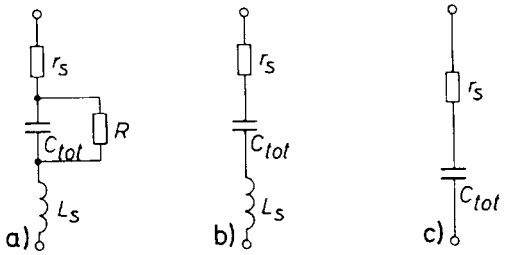


Figure 2-49 Equivalent circuits for capacitance diodes: (a) complete circuit, (b) simplified circuit, and (c) further simplification for low frequencies.

$$\frac{C_{\max}}{C_{\min}} = \frac{C_{\text{tot}}(V_{R\min})}{C_{\text{tot}}(V_{R\max})} \tag{2-68}$$

is limited by the fact that the diode must not be driven by the alternating voltage superimposed on the tuning voltage either into the forward mode or the breakdown mode. Otherwise, rectification, which would shift the diode's bias and considerably affect its figure of merit, would take place. Figure 2-48 plots the capacitance/voltage characteristics of a capacitance diode to clarify the relationship. The useful operating range lies between the voltages

$$V_{\min} > \hat{V} - V_F \quad \text{and} \quad V_{\max} < V_{(\text{BR})R} - \hat{V} \tag{2-69}$$

As has already been indicated, the exponent n for large-capacitance-ratio or tuning diodes used today for TV tuners is not constant but voltage-dependent and subject to manufacturing tolerances. This means that the capacitance/voltage characteristic of these diodes is likewise subject to manufacturing tolerances. Since, in a TV tuner, it is necessary for two or three circuits to be tuned uniformly, tuner diodes must be selected empirically for identical characteristics and supplied in equipment lots.

Electrical Properties of Tuning Diodes. In this section, the electrical properties of capacitance diodes are described, with reference to data published on the tuner diodes BB141 and BB142.

Equivalent Circuit. Since a capacitance diode is not an ideal capacitor, it is useful to introduce an equivalent circuit that can serve as a basis for discussing the diode's electrical properties. Figure 2-49 shows various types of equivalent circuits.

The complete equivalent circuit (Figure 2-49a), which conforms closely to physical conditions, comprises, in addition to the diode capacitance C_{tot}, the series resistance r_S and the series inductance L_S, and also the reverse resistance $R = dV_R/dI_R$. At higher frequencies, this resistance can usually be disregarded, so that the equivalent circuit is simplified in terms of Figure 2-49b—the configuration usually employed. In many cases, the series inductance can also be disregarded, in which case we obtain Figure 2-49c.

Capacitance. Figure 2-50 shows the capacitance/voltage characteristic borrowed from the datasheets of diodes BB141 and BB142. As this characteristic, like most characteristics published in datasheets, represents merely a typical curve, it would not convey sufficient

information for the user to dimension a tuner. Therefore, the datasheets contain additional data on capacitance and useful capacitance ratio:

Capacitance		BB141A (pF)	BB141B (pF)	BB142 (pF)
At $V_R = 1$ V	C_{tot}	16	19	17
At $V_R = 3$ V	C_{tot}	11	13	12
At $V_R = 25$ V	C_{tot}	2–2.35	2.25–2.65	2–3

$$\text{Useful capacitance ratio} = \frac{C_{tot}(3\text{ V})}{C_{tot}(25\text{ V})} = 4 \text{ to } 6$$

On the basis of the spread of the diode capacitance C_{tot} guaranteed for $V_R = 25$ V, it is possible to calculate the tuner with the aid of the spread guaranteed for the capacitance ratio.

Three further capacitance graphs are shown in Figures 2-51 through 2-53. Figure 2-51 shows a (typical) curve representing the normalized slope as a function of the reverse voltage, Figure 2-52 the normalized capacitance as a function of the junction temperature, and Figure 2-53 the temperature coefficient of the capacitance as a function of the reverse voltage.

The variation of diode capacitance with ambient temperature, as shown in Figures 2-52 and 2-53, is really only a function of the temperature dependence of the diffusion voltage [see Eq. (2-66)]. The diffusion voltage drops with rising temperatures by about 2 mV/°C , which means that the diode capacitance rises with temperature. The influence of the diffusion voltage, and thus the temperature coefficient of the capacitance, decreases as the reverse voltage rises. It is therefore advisable to run a capacitance diode with as high a reverse voltage as the required capacitance ratio permits. Compensation for the dependence of capacitance on temperature will be discussed later.

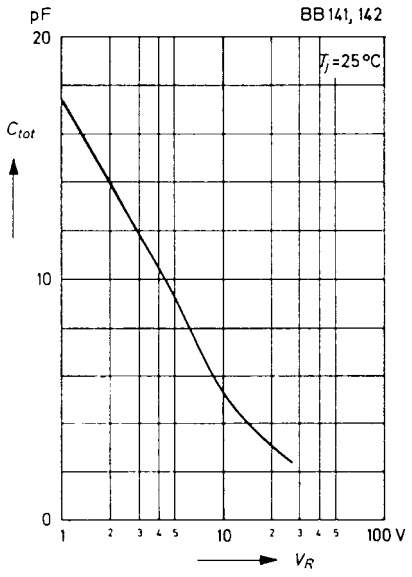


Figure 2-50 Capacitance/voltage characteristic of diodes BB141 and BB142.

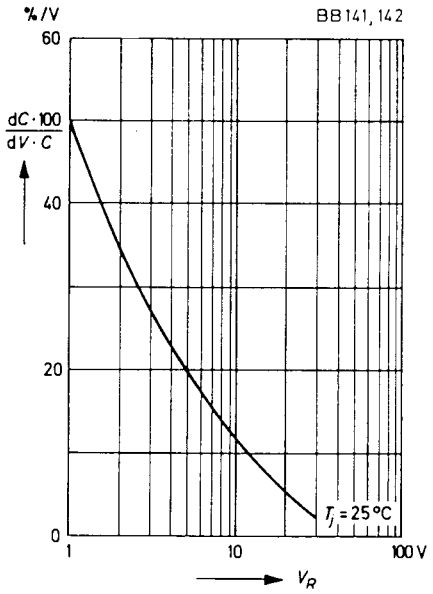


Figure 2-51 Slope (normalized) as a function of the reverse voltage.

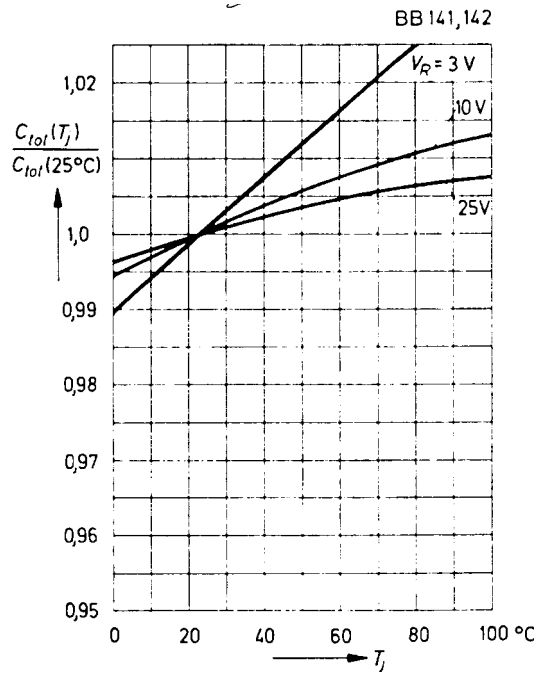


Figure 2-52 Capacitance (normalized) as a function of the junction temperature.

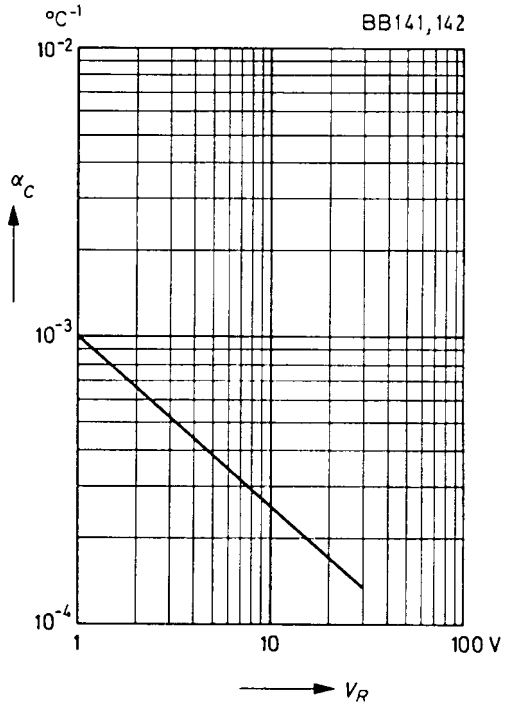


Figure 2-53 Temperature coefficient of capacitance as a function of the reverse voltage.

Series Resistance, Figure of Merit (Q). Since capacitance diodes are intended for employment in resonant circuits, indications are required as to the circuit attenuation that they produce. Usually, the series resistance or the figure of merit that can be calculated therefrom is indicated. With reference to the equivalent circuit shown in Figure 2-49a, the figure of merit (or Q factor) is calculated:

$$Q = \frac{1}{\omega C_{\text{tot}} r_S + 1/\omega C_{\text{tot}} R} \tag{2-70}$$

Figure 2-54 shows the theoretical (normalized) Q as a function of frequency. With the frequency range of 10–1000 MHz, which is essential for tuner diodes, the parallel resistance R caused by the reverse current (see Figure 2-49b) can be disregarded, and Eq. (2-70) can be simplified to read as follows:

$$Q = \frac{1}{\omega C_{\text{tot}} r_S} \tag{2-71}$$

In Figure 2-55, Q is plotted as a function of frequency for the tuner diodes BB141 and BB142. The datasheet indicates, in addition to this curve, the value of the series resistance at $f = 470$ MHz and $C_{\text{tot}} = 9$ pF:

	BB141A, BB141B	BB142
Series resistance (r_S)	0.6 (< 0.8)	0.9 (< 1.2)

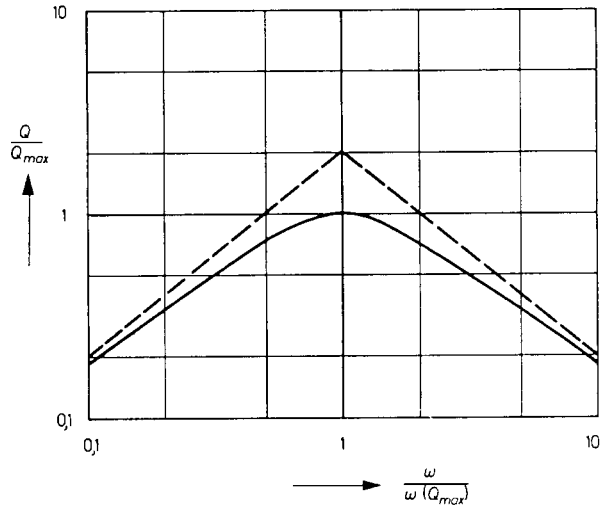


Figure 2-54 *Q* (normalized) as a function of frequency.

Since the diode capacitance and the series resistance decrease as the reverse voltage rises, Q is likewise dependent on the reverse voltage. It is smallest at a low reverse voltage, that is, at a low frequency. The series and parallel capacitances inevitable in a diode-tuned resonant circuit influence the effective Q of the tuning capacitance, as will be described later.

Series Inductance, Series Resonant Frequency, Cutoff Frequency for Q = 1. These three parameters define the behavior of the diode at high frequencies when the influence of

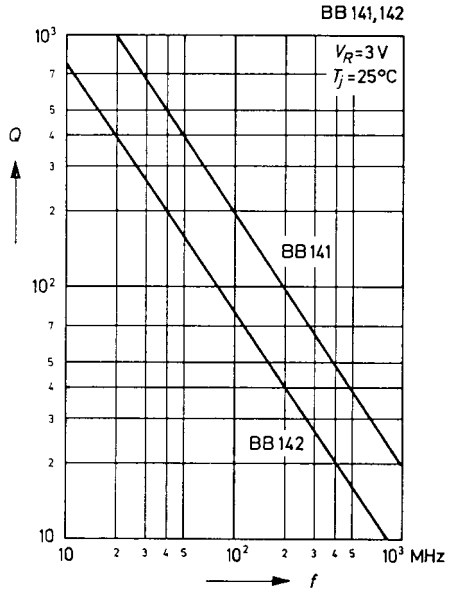


Figure 2-55 *Q* as a function of frequency.

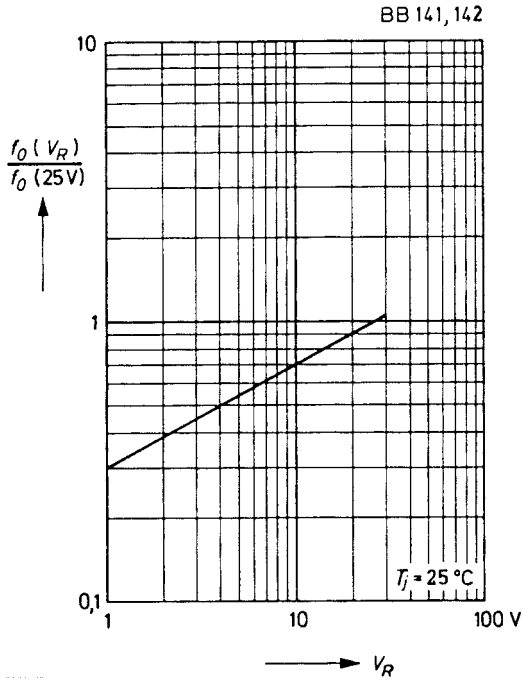


Figure 2-56 Series-resonant frequency (normalized) as a function of the reverse voltage.

the series inductance can no longer be neglected. The parameter on which the calculation is based is the series inductance L_S, which is determined by the connection leads. As indicated in Figure 2-49a, it is in series with the capacitance and, given a suitable frequency, forms a series-resonant circuit therewith. In that case

$$\omega L_S = \frac{1}{\omega C_{\text{tot}}} \tag{2-72}$$

and the series-resonant frequency is

$$f_0 = \frac{1}{2\pi\sqrt{C_{\text{tot}} L_S}} \tag{2-73}$$

Above this frequency, the impedance of the diode is inductive, but it depends on the reverse voltage. Figure 2-56 shows how the (normalized) series-resonant frequency depends on the reverse voltage for the tuner diodes BB141 and BB142.

The cutoff frequency for $Q = 1$ is defined as the frequency at which Q assumes the value of unity. In a rough calculation, the reactive impedance of the diode capacitance can be neglected and the frequency f_{Q1} obtained for the condition in which the inductive reactance equals the series resistance:

$$\omega L_S = r_S \tag{2-74}$$

$$f_{Q1} = \frac{r_S}{2\pi L_S} \tag{2-75}$$

For the tuner diodes BB141 and BB142, Figure 2-57 shows how the (normalized) cutoff frequency depends on the reverse voltage. The datasheets contain the following numerical information:

		BB141	BB142	
Series inductance measured 1.5 mm from the case	L_S	2.5	2.5	nH
Series-resonant frequency at $V_R = 25$ V	f_0	2	1.8	GHz
Cutoff frequency for $Q = 1$ at $V_R = 3$ V	f_{Q1}	24	16	GHz

Depending on the intended application, either the series-resonant frequency or the cutoff frequency for $Q = 1$ determines the maximum useful frequency range.

Leakage Current, Breakdown Voltage. These two parameters mark the dc performance of the diode when a reverse bias is applied to it. In the databook for the tuner diodes BB141 and BB142, the breakdown voltage is quoted as follows:

$$\text{Reverse breakdown voltage at } I_R = 100\ \mu\text{A}: \quad V_{(BR)R} > 30\ \text{V}$$

This means that the maximum reverse voltage that may be applied to the diode as a tuning voltage is 30 V.

The leakage current for the above-mentioned tuner diodes is guaranteed as follows:

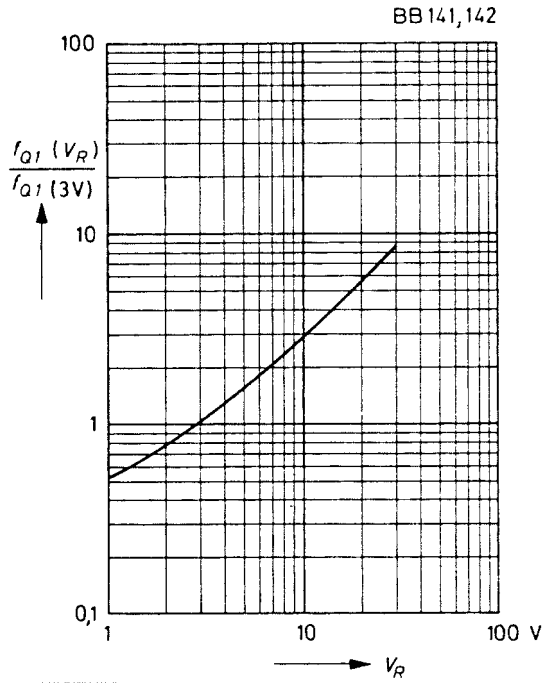


Figure 2-57 Cutoff frequency for $Q = 1$ (normalized) as a function of the reverse voltage.

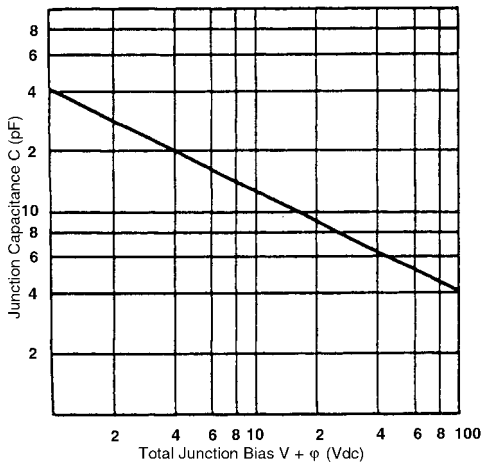


Figure 2-58 Leakage current as a function of reverse voltage.

$$\text{Leakage current at } V_R = 28 \text{ V:} \quad I_R < 50 \text{ nA}$$

Moreover, Figure 2-58 shows a graph that illustrates the leakage current as a function of the reverse voltage. Since this current is temperature dependent (as with every silicon diode)—it doubles with each temperature rise of about 10 °C—care should be taken when dimensioning the tuner circuit that the leakage current does not cause inadmissible voltage variations at increased ambient temperatures.

Matching of Tuner Diodes, Uniform Parameters. As has already been mentioned, the capacitance versus voltage characteristics of modern tuner diodes are subject to a certain scatter, so that it becomes necessary to test these diodes to obtain equipment lots empirically.

Diode-Tuned Resonant Circuits

Tuner Diode in the Parallel-Resonant Circuit. Figures 2-59, 2-60, and 2-61 illustrate three basic circuits for the tuning of parallel-resonant circuits by means of capacitance diodes.

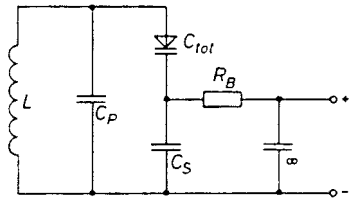


Figure 2-59 Parallel-resonant circuit with tuner diode, and bias resistor parallel to the series capacitor.

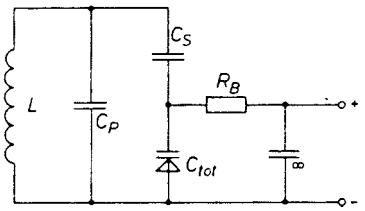


Figure 2-60 Parallel-resonant circuit with tuner diode, and bias resistor parallel to the diode.

In the circuit diagram of Figure 2-59, the tuning voltage is applied to the tuner diode via the input coil and the bias resistor R_B. Series-connected to the tuner diode is the series capacitor C_S, which completes the circuit for ac but isolates the cathode of the tuner diode from the coil and thus from the negative terminal of the tuning voltage. Moreover, a fixed parallel capacitance C_P is provided. The decoupling capacitor preceding the bias resistor is large enough for its value to be disregarded in the following discussion. Since for high-frequency purposes the biasing resistor is connected in parallel with the series capacitor, it is transformed into the circuit as an additional equivalent shunt resistance R_C. We have the equation

$$R_C = R_B\left(1 + \frac{C_S}{C_{\text{tot}}}\right)^2 \tag{2-76}$$

If in this equation the diode capacitance is substituted by the resonant circuit frequency ω, we obtain

$$R_C = R_B\left(\frac{\omega^2 LC_S}{1 - \omega^2 LC_P}\right)^2 \tag{2-77}$$

The resistive loss R_C caused by the bias resistor R_B is seen to be highly frequency dependent, and this may result in the bandwidth of the tuned circuit being independent of frequency if the capacitance of the series capacitor C_S is not chosen sufficiently high.

Figure 2-60 shows that the tuning voltage can also be applied directly and in parallel to the tuner diode. For the parallel loss resistance transformed into the circuit, we have the expression

$$R_C = R_B\left(\frac{1 + C_{\text{tot}}}{C_S}\right)^2 \tag{2-78}$$

and

$$R_C = R_B\left(\frac{\omega^2 LC_S}{\omega^2 L(C_S + C_P) - 1}\right)^2 \tag{2-79}$$

The influence of the bias resistor R_B in this case is larger than in the circuit of Figure 2-60, provided that

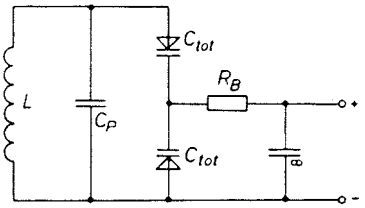


Figure 2-61 Parallel-resonant circuit with two tuner diodes.

$$C_S^2 > C_S(C_{\text{tot}} + C_P) + C_{\text{tot}}C_P \tag{2-80}$$

This is usually the case because the largest possible capacitance will be preferred for the series capacitor C_S, and the smallest for the shunt capacitance C_P. The circuit of Figure 2-59 is therefore normally preferred to that of Figure 2-60. An exception would be the case in which the resonant circuit is meant to be additionally damped by means of the bias resistor at higher frequencies.

In the circuit of Figure 2-61, the resonant circuit is tuned by two tuner diodes that are connected in parallel via the coil for tuning purpose, but series-connected in opposition for high-frequency signals. This arrangement has the advantage that the capacitance shift caused by the ac modulation (see section entitled "Modifying the Diode Capacitance by the Applied ac Voltage") takes effect in opposite directions in these diodes and therefore cancels itself. The bias resistor R_B, which applies the tuning voltage to the tuner diodes, is transformed into the circuit at a constant ratio throughout the whole tuning range. Given two identical, loss-free tuner diodes, we obtain the expression

$$R_C = 4R_B \tag{2-81}$$

Capacitances Connected in Parallel or Series with the Tuner Diode. Figures 2-59 and 2-60 show that a capacitor is usually in series with the tuner diode, in order to close the circuit for alternating current and, at the same time, to isolate one terminal of the tuner diode from the rest of the circuit with respect to direct current, so as to enable the tuning voltage to be applied to the diode. As far as possible, the value of the series capacitor C_S will be chosen such that the effective capacitance variation is not restricted. However, in some cases, as, for example, in the oscillator circuit of receivers whose intermediate frequency is of the order of magnitude of the reception frequency, this is not possible and the influence of the series capacitance will then have to be taken into account. By connecting the capacitor C_S, assumed to be lossless, in series with the diode capacitance C_{tot}, the tuning capacitance is reduced to the value

$$C^* = C_{\text{tot}} \frac{1}{1 + C_{\text{tot}}/C_S} \tag{2-82}$$

The Q of the effective tuning capacitance, taking into account the Q of the tuner diode, increases to

$$Q^* = Q\left(1 + \frac{C_{\text{tot}}}{C_S}\right) \tag{2-83}$$

The useful capacitance ratio is reduced to the value

$$\frac{C^*_{\max}}{C^*_{\min}} = \frac{C_{\max}}{C_{\min}} \frac{1 + C_{\min}/C_S}{1 + C_{\max}/C_S} \tag{2-84}$$

where $C_{\max}$ and $C_{\min}$ are the maximum and minimum capacitances of the tuner diode.

On the other hand, the advantage is gained that, due to capacitive potential division, the amplitude of the alternating voltage applied to the tuning diode is reduced to

$$\hat{v}^* = \hat{v} \frac{1}{1 + C_{\text{tot}}/C} \tag{2-85}$$

so that the lower value of the tuning voltage can be smaller, and this results in a higher maximum capacitance $C_{\max}$ of the tuner diode and a higher useful capacitance ratio. The influence exerted by the series capacitor, then, can actually be kept lower than Eq. (2-83) would suggest.

The parallel capacitance C_P that appears in Figures 2-59 to 2-61 is always present, since wiring capacitances are inevitable and every coil has its self-capacitance. By treating the capacitance C_P, assumed to be lossless, as a shunt capacitance, the total tuning capacitance rises value and, if C_S is assumed to be large enough to be disregarded, we obtain

$$C^* = C_{\text{tot}} \left(1 + \frac{C_P}{C_{\text{tot}}} \right) \tag{2-86}$$

The Q of the effective tuning capacitance, as derived from the Q of the tuner diode, is

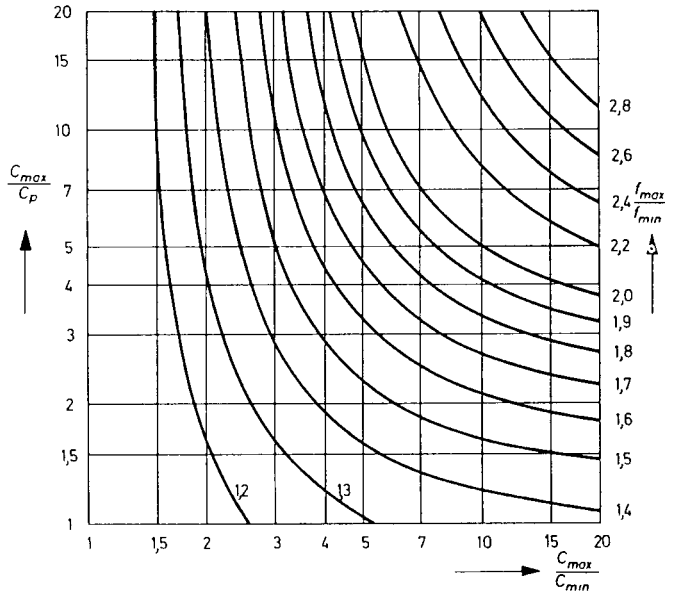


Figure 2-62 Diagram for determining the capacitance ratio and minimum capacitance.

$$Q^* = Q\left(1 + \frac{C_P}{C_{\text{tot}}}\right) \tag{2-87}$$

or, in other words, it rises with the magnitude of the parallel capacitance. The useful capacitance ratio is reduced:

$$\frac{C^*_{\text{max}}}{C^*_{\text{min}}} = \frac{C_{\text{max}}}{C_{\text{min}}} \frac{1 + C_P/C_{\text{max}}}{1 + C_P/C_{\text{max}}} \tag{2-88}$$

In view of the fact that even a comparatively small shunt capacitance reduces the capacitance ratio considerably, it is necessary to ensure low wiring and coil capacitances in the circuit-design stage.

TUNING RANGE. The frequency range over which a parallel-resonant circuit according to Figure 2-59 can be tuned by means of the tuner diode depends on the useful capacitance ratio of the diode and on the parallel and series capacitances present in the circuit.

The ratio can be found from

$$\frac{f_{\text{max}}}{f_{\text{min}}} = \sqrt{1 + \frac{\dfrac{C_{\text{max}}}{C_P\,(1 + C_{\text{max}}/C_S)}}{1 + \dfrac{C_{\text{max}}}{C_P\,(C_{\text{max}}/C_{\text{min}} + C_{\text{max}}/C_S)}}} \tag{2-89}$$

In many cases, the series capacitor can be chosen large enough for its effect to be negligible. In that case, Eq. (2-89) is simplified as follows:

$$\frac{f_{\text{max}}}{f_{\text{min}}} = \sqrt{\frac{1 + C_{\text{max}}/C_P}{1 + C_{\text{min}}/C_P}} \tag{2-89a}$$

From this equation, the diagram shown in Figure 2-62 is computed. With the aid of this diagram, the tuner diode parameters required for tuning a resonant circuit over a stipulated frequency range—that is, the maximum capacitance and the capacitance ratio—can be determined.

Whenever the series capacitance C_S cannot be disregarded, the effective capacitance ratio is reduced according to Eq. (2-84).

TRACKING. Some applications require the maintenance of a fixed frequency relationship between two or more tuned circuits as their tuning is simultaneously adjusted. Referred to as *tracking*, this technique requires narrow tolerances of capacitance versus tuning voltage. Minimizing tracking error requires special care if the tracking circuits must cover the same frequency span beginning at different start and end frequencies, as is necessary when simultaneously tuning oscillator and mixer/RF circuitry in a superheterodyne receiver. Then, tracking error must be minimized by means of series and shunt capacitances in accordance with methods known from variable capacitors. The frequency deviations that must be anticipated are summarized in the equation

$$\frac{df}{f} = -\frac{1}{2}\frac{dC_0}{C_0} - \frac{1}{2}\frac{d(L-L_0)}{L} - \frac{1}{2}\frac{dL_0}{L_0} + \frac{n}{2}\frac{dV_R}{V_R + V_D} \tag{2-90}$$

The spread of parameters dC_0/C_0 and dL_0/L_0 can only be compensated by varying the circuit inductances $d(L–L_0)/L$ or by varying the bias $dV_R/(V_R + V_D)$.

Modulating the Diode Capacitance by the Applied ac Voltage. In normal operation, the sum of the tuning voltage and the alternating signal voltage of the resonant circuits is applied to the tuner diode. The bias, and thus the capacitance, of the tuner diode therefore varies at the rhythm of the alternating voltage. Due the nonlinear character of the capacitance versus voltage curve, voltage distortions and capacitance shifts are inevitable, and these must be kept within adequate limits. This is done by maintaining the ac applied to the diode(s) at sufficiently low ac amplitude and by choosing an adequate minimum value for the tuning voltage. In the resonant circuit, a tuner diode is modulated predominantly by a current free from harmonics, according to the equation

$$i = \hat{i}\cos \omega t \tag{2-91}$$

The alternating voltage across the diode is

$$v = (V_R + V_D)\left[\left(1 + \frac{\hat{i}(1-n)}{\omega C_{\text{tot}} V_R}\sin \omega t\right)^{1/(1-n)} - 1\right] \tag{2-92}$$

An evaluation of this equation shows that especially the first harmonic makes its appearance. The capacitance shift caused by the alternating voltage superimposed on the tuning voltage is shown in Figure 2-63. However, the voltage distortion, and thus the capacitance shift, can be largely avoided if two tuner diodes are used, as in Figure 2-61.

COMPENSATING TEMPERATURE DEPENDENCE. As has already been mentioned, variations are caused in the capacitance of the diode mainly by the dependence of the diffusion voltage on

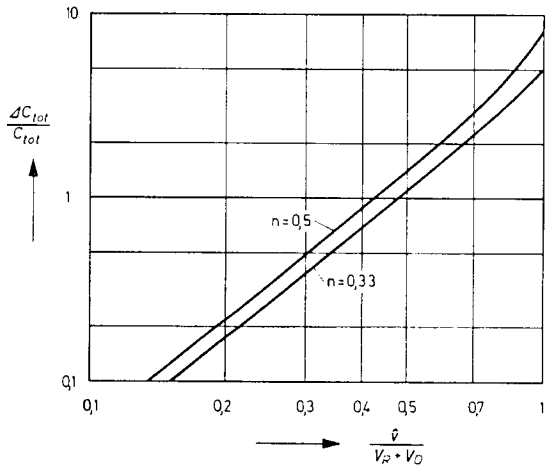


Figure 2-63 Capacitance increase as a function of the ac voltage drop across the tuner diode.

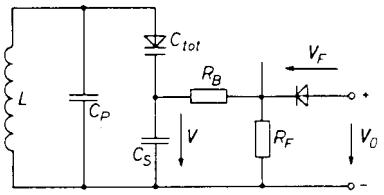


Figure 2-64 Temperature-compensation circuit with diode.

temperature. In a diode-tuned resonant circuit, the temperature coefficient of the resonant frequency, therefore, depends on the tuning voltage and thus on the resonant frequency. It is therefore impossible to compensate the temperature dependence of the resonant frequency by means of temperature-dependent capacitors—the method usually adopted for mechanically tuned circuits.

To achieve satisfactory compensation, the tuning voltage should be increased by an amount equal to that by which the diffusion voltage of the diode is reduced with rising temperature; that is, by approximately 2 mV/°C. This can be done easily by connecting a forward-biased silicon diode in series with the source of the tuning voltage, as shown in Figure 2-64. Since the forward voltage of the diode varies by –2 mV/°C with ambient temperature, the voltage

$$V = V_0 + V_D - V_F$$

which determines the bias of the tuner diode, and therefore also the diode capacitance, is almost temperature independent. In practice, it is advisable to feed this diode with additional current via the resistor R_F to keep its differential forward resistance sufficiently low and thus to prevent capacitive noise voltages from leaking into the tuning circuit.

To prevent the load placed by the resistor R_F on the source of the tuning voltage from fluctuating in a manner depending on the tuning voltage, a circuit such as shown in Figure 2-65 may be used in which the emitter diode of the transistor effects the compensation, and the only load placed on the source of the tuning voltage is the base current of this transistor.

Dynamic Stability. The resonant frequency of a diode-tuned RF circuit follows the equation

$$f = f_0 \sqrt{\left(\frac{V_R + V_D}{V_{R0} + V_D}\right)^n} \tag{2-93}$$

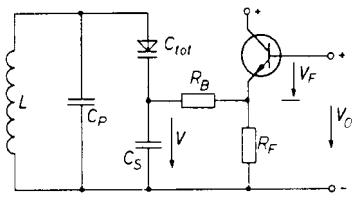


Figure 2-65 Temperature-compensation circuit with transistor.

The frequency f_0 is set with the aid of the circuit inductance when the tuning voltage equals V_{R0}. Since the signal amplitude in the circuit depends on the field strength of the received transmitter and since the resonant frequency of the circuit depends on the amplitude, undesirable feedback between the tuning voltage and the received alternating voltage may occur at the tuner diode. A disturbance that has a similar effect is the microphonic effect, which can sometimes be observed with mechanically variable capacitors. In an investigation of the feedback in diode-tuned resonant circuits, the following assumptions were made:

- The impedance of the antenna (or of the signal generator) is negligible.
- The full alternating voltage is applied to the diode (whereas, in practice, the series capacitor brings about a potential division).
- The capacitance versus voltage characteristic is an exponential function.
- The Q of this circuit is independent of the amplitude of the alternating voltage.

As a result of a high alternating voltage at the resonant circuit, three undesirable effects—harmonic generation, frequency shift, and cross-modulation—are encountered.

HARMONIC GENERATION. The nonlinearity of the capacitance versus voltage characteristic produces harmonics. However, the selectivity of the receiver reduces the effects of these harmonics to noise level.

FREQUENCY SHIFT. When a sinusoidal voltage v is applied to the tuner diode of the RF circuit, the variation of the diode capacitance does not follow the sine law, but the equation

$$\frac{C_\omega}{C} = \left(1 + \frac{v}{V_R + V_D} \sin \omega t\right)^{-n} \tag{2-94}$$

This results in a change of the resonant frequency of the circuit. The inductive slope of the resonant curve steepens and may even turn back. This could lead to bistable behavior of the resonant circuit. The change of frequency takes place as if the tuning voltage, and thus the resonant frequency, has decreased. The order of magnitude of this change of resonant frequency was determined with the aid of a computer from Fourier analysis of

$$\frac{\Delta f}{f} = 1 + \frac{1}{2\pi} \int_0^{2\pi} \left(1 + \frac{v}{V_R + V_D} \sin \omega t\right)^{n/2} d(\omega t) \tag{2-95}$$

Figure 2-66 shows the admissible alternating voltage at the diode, computed for a 2% frequency shift, as a function of the tuning voltage. The measured values differ by not more than 10% from the computed values. Since the Q of the circuits is about 50 for the VHF as well as the UHF range, a 2% frequency shift is permissible.

CROSS-MODULATION. Cross-modulation is a disturbance caused by nonlinearities of the characteristics of the components concerned. The modulation of an undesired amplitude-modulated signal is here transferred to the carrier of the amplitude-modulated intelligence signal. Cross-modulation is virtually independent of the amplitude of the intelligence signal.

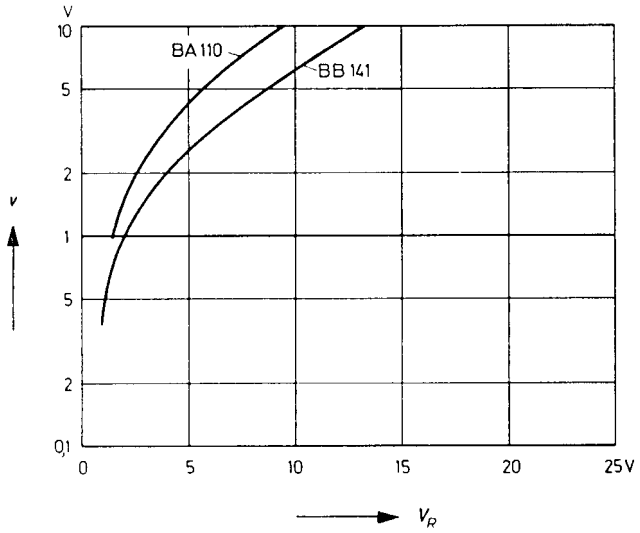


Figure 2-66 Admissible alternating voltage for a 2% frequency shift as a function of tuning voltage.

The cross-modulation factor for tuner diodes is

$$k = \frac{v^2}{(V_R + V_D)^2} m_v \frac{n(n+1)}{2} \tag{2-96}$$

Resolving this equation in terms of v, we obtain

$$v = \frac{V_R + V_D}{n} \sqrt{\frac{2kn}{m_v (n+1)}} \tag{2-97}$$

The usual definition of the cross-modulation percentage for measurement purposes is as follows: A 100% modulated noise signal ($m_v = 1$) of amplitude v is received simultaneously with the intelligence signal in the diode-tuned RF circuit. If a 1% modulation of the signal carrier by the noise signal takes place, this amounts to a cross-modulation factor $k = 0.01$. We thus obtain for tuner diodes

$$V = 0.14 \frac{V_R + V_D}{n} \sqrt{\frac{n}{n+1}} \tag{2-98}$$

Figure 2-67 gives the graphs for this equation. In practical operation, FM and VHF/UHF tuners equipped with the tuner diodes BB121 or BB141 exhibit merely the cross-modulation caused by the BJTs or FETs employed. For bipolar transistors, $v \approx 15$–100 mV; for FETs, 100–300 mV.

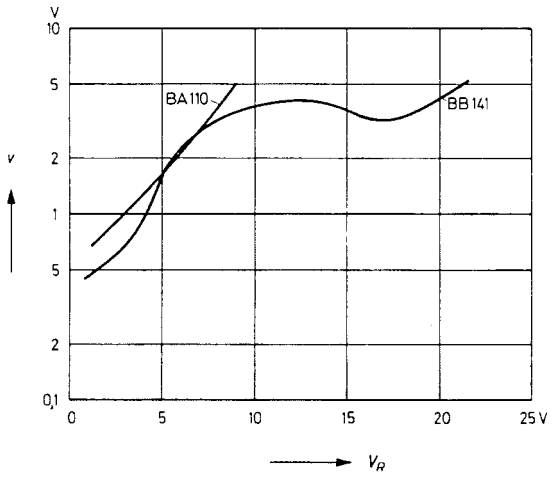


Figure 2-67 Graphical representation of Eq. (2-98).

Static Stability. To ensure trouble-free operation, the static parameters of the tuner diodes, especially the leakage current, must be taken into account. When differentiating the expression for the tuning voltage

$$V_R = V_B - I_R R_B \tag{2-99}$$

we obtain for a known leakage current I_R the tuning voltage variation

$$dV_R = dV_B - R_B I_R \left(\frac{dR_B}{R_B} + \frac{dI_R}{I_R} + dV_B \right) \tag{2-100}$$

This leads to a change in the resonant frequency

$$\frac{df}{f} = \frac{n}{2(V_R + V_D)} \left[dV_D - R_B I_R \left(\frac{dR_B}{R_B} + \frac{dI_R}{I_R} \right) + dV_B \right] \tag{2-101}$$

The resonant frequency is more stable the smaller the exponent n and the higher the tuning voltage V_R. The influence of the temperature dependence of the diffusion voltage V_D and its compensation have already been discussed. If the series bias resistance R_B is not too high, the effect of its variation ($dR_B/R_{B)}$ and the effect of the temperature dependence of the leakage current—doubling for every 10 °C temperature rise—may be neglected.

Generating the Tuning Voltage. The supply voltage V_B in Eq. (2-101) exerts an influence in many respects. The temperature coefficient of the voltage source—usually, for economic reasons, only a Zener diode is provided for stabilization—affects conditions not only with ambient temperature variations, but also when the Zener diode heats up after the set is switched on, or when its operating current varies. Because external temperature compensation of the operating voltage of the Zener diode cannot cover all detrimental effects and, moreover, increases the series bias resistance R_B, as well as any variations the latter may

undergo, the use of a temperature-compensated Zener diode, such as the 1N4065A (nominal Zener voltage, 33 V; temperature coefficient, 0.002 °C/W), is recommended.

Diode Switches. The diode switches described here differ somewhat from the switching diodes used in computer and pulse technology. Whereas in the case of switching diodes the signal itself triggers the switching operation—current does or does not pass through the switching diode depending on the signal level—diode switches allow an alternating current to be switched on or off by means of a direct voltage or a direct current. Figure 2-68 shows an example of the use of diode switches.

Diode Switch Technology. Special manufacturing techniques are needed to attain the low junction capacitance, low differential forward resistance r_F, and the low-inductance structure required for the special applications envisaged. Two conflicting requirements have to be met: On the one hand, a high-resistivity basic material should be used and the diode area should be kept as small as possible to minimize the junction capacitance for a given reverse voltage; on the other hand, the bulk resistance should be kept low. The size of the junction area, therefore, can represent only a compromise between these two requirements. To obtain a low bulk resistance R_B despite the use of high-resistivity silicon around the junction, the diodes are manufactured by the epitaxial process, characterized by a thin high-resistivity layer on a low-resistivity substrate. If the forward current I_F is sufficiently high, then the specific resistance of the epitaxial layer does not affect the bulk resistance, because under these conditions the resistivity of this layer is considerably reduced by being flooded with carriers.

The junction impedance Z_j is somewhat more critical (Figure 2-69). It is composed of the junction capacitance C_S, which does not depend on current, the current-dependent diffusion capacitance

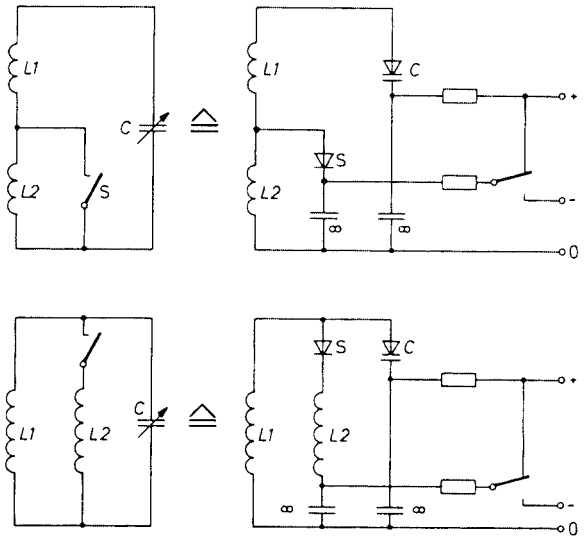


Figure 2-68 Comparison of mechanically and electronically tuned and switched resonant circuits.

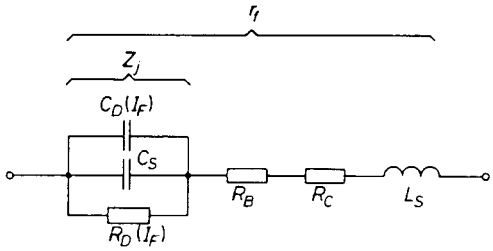


Figure 2-69 Equivalent circuit of diode switches in the frequency range from 1 MHz to 1 GHz.

$$C_D = \frac{\mathfrak{t}_p I_F}{V_T} \tag{2-102}$$

and the diffusion resistance

$$R_D = \frac{V_T}{I_F} \tag{2-103}$$

where $\mathfrak{t}_p$ is the lifetime of the minority charge carriers and

$$V_T = \frac{kT}{q} = 26 \text{ mV} \tag{2-104}$$

is the voltage equivalent of thermal energy. The resistive component of the impedance Z_j is calculated as follows:

$$\text{Re}(Z_j) = \frac{R_D}{1 + \omega^2 R_D^2 (C_S + C_D)^2} \tag{2-105}$$

As can be seen from the following approximations, this resistance becomes rather dependent on current at higher frequencies:

$$\text{For } I_F \to 0 \quad \text{Re}(Z_j) \approx \frac{I_F}{\omega^2 C_S^2 V_T} \tag{2-106}$$

$$\text{For } I_F \to \infty \quad \text{Re}(Z_j) \approx \frac{I_F}{\omega t_p I_F} \tag{2-107}$$

In Figure 2-70 the differential forward resistance is plotted as a function of the forward current. The only parameter by which the differential forward resistance can be reduced for a given current and a given frequency according to Eq. (2-107) is the carrier lifetime. Steps will have to be taken, therefore, to make the carrier lifetime as long as possible.

Because diode switches are designed for applications up to the gigahertz range, a geometry resulting in a low inherent capacitance and inductance had to be used. Furthermore, the device had to be assembled in such a way that virtually no additional contact resistance R_C was introduced. For this reason, a double-plug construction (hard-glass pressure contacts)

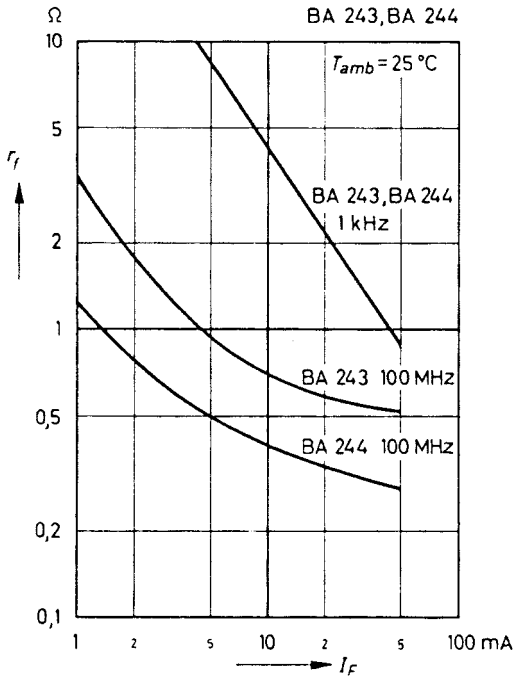


Figure 2-70 Differential forward resistance as a function of forward current.

was chosen to avoid the S- or U-bends customary with other encapsulations. The system is pressure-clamped between two stud-shaped enlargements of the connecting leads, the necessary high pressure being produced as the glass shrinks during cooling. This construction results in an encapsulation of short length (only 4 mm) and low series inductance, the advantages of which can be fully exploited because soldered connections to the leads can be made directly at the glass body.

Diode Switch Data. The silicon planar diode switches BA243 and BA244 are designed for electronic band selection in tuners in the frequency range from 10 to 1000 MHz. The differential forward resistance remains constant and is very low over a wide frequency and current range. The diode capacitance is likewise small and independent of voltage through a wide range. The devices are characterized by low-inductance geometry. The BA243 is intended for the VHF range, and the BA244 for the UHF range. See Table 2-11.

Resonant Circuits Incorporating Diode Switches. Figure 2-71 illustrates the function of the diode switch with reference to a parallel-resonant circuit whose resonant frequency can be switched by short-circuiting part of the circuit inductance. Such a switching procedure may be necessary for a VHF tuner, for example, to switch from band I to band III, because the frequencies of these bands are too far apart.

Capacitor C_K represents the capacitance of the resonant circuit. The symbol ∞ at capacitor C is meant to indicate that the value of this capacitor should be chosen to be much higher than the circuit capacitance. C completes the ac circuit when the inductor L_2 is shorted and

Table 2-11 BA243/244 specifications

Maximum Ratings			
Reverse voltage	V_R	20	V
Forward current at $T_A = 25$ °C	I_F	100	mA
Junction temperature	T_j	150	°C
Storage temperature range	T_s	–55 to +150	°C
Characteristics at $T_A = 25$ °C			
Forward voltage at $I_F = 100$ mA	V_F	<1	V
Leakage current at $V_R = 15$ V	I_R	< 100	nA
Dynamic forward resistance at $f = 50$–1000 MHz, $I_F = 10$ mA			
BA243	r_F	0.7 (< 1)	Ω
BA244	r_F	0.4 (< 0.5)	Ω
Relative variation of dynamic forward resistance with the variation of forward current in the range of $I_F + 2$–40 mA	$\frac{\Delta r_F 100}{r_F \Delta I_F}$	5	%/mA
Capacitance at $V_R = 15$ V, $f = 1$ MHz	C_{tot}	1.3 (< 2)	pF
Relative variation of capacitance with the variation of reverse voltage in the range of $V_R = 7$–20 V, $f = 100$ MHz	$\frac{\Delta C_{tot} 100}{C_{tot} \Delta V_R}$	1	%/V
Series inductance across case	L_S	2.5	nH

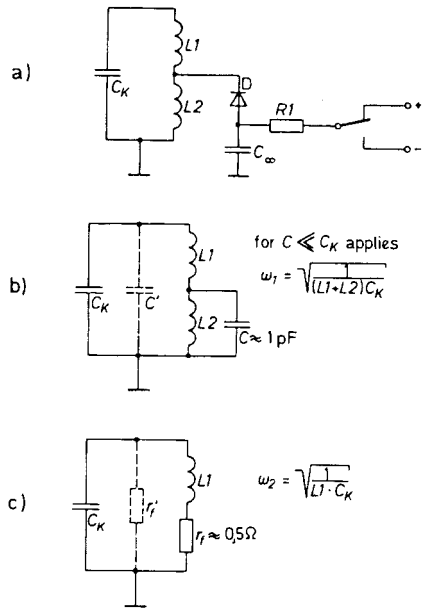


Figure 2-71 (a) Diode switch as bandswitch, shorting a partial inductance; (b) equivalent circuit with blocked diode switch; (c) equivalent circuit with conducting diode switch.

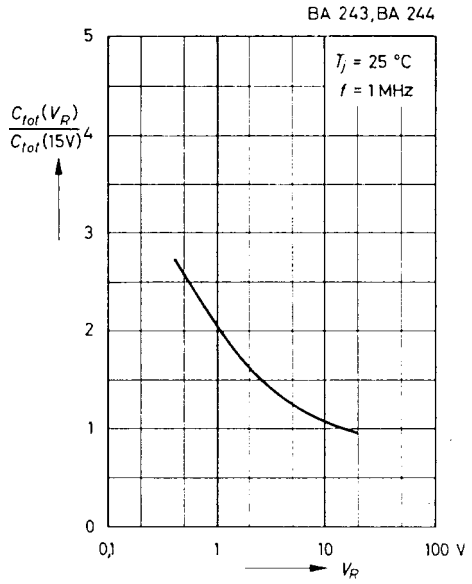


Figure 2-72 Capacitance as a function of the reverse voltage.

enables the switching or reverse voltage to be applied at a point of virtually zero RF potential. The resistor R_1 limits the diode current in the forward direction.

At a reverse voltage of $V_R = 15$ V, for example, the diode switch represents a 1-pF capacitor, as shown in Figure 2-71b. This represents parasitic capacitance C' in the circuit, which reduces the effective capacitance variation of the resonant circuit in the case of capacitive tuning. The capacitance C' is small enough to be negligible in most cases relative to the wiring and winding capacitances. Figure 2-72 shows the relative capacitance of the blocked-diode switches, as a function of the reverse voltage.

At a forward current of, for example, 10 mA, the diode has a differential resistance of approximately 0.5 Ω (see Figure 2-70). The effect of this resistance r_F on the resonant circuit is illustrated in Figure 2-71c. Conversion of a series loss resistance r_F into a parallel loss resistance r'_F is based on the following considerations: The equivalent parallel resistance r'_F affects the tuned circuit in the same way as the series resistance r_F; that is, the Q of the circuit should be the same in both cases. If one assumes that $\omega L >> r_F$, the equivalent of parallel resistance can be expressed as follows:

$$Q = \frac{\omega L_1}{r_F} = \frac{r'_F}{\omega L_1} \tag{2-108}$$

$$r'_f = \frac{(\omega L_1)^2}{r_F} \tag{2-109}$$

An evaluation of this equation shows that a differential resistance $r_F \approx 0.5$ Ω can be tolerated in most applications as regards amplification and selectivity.

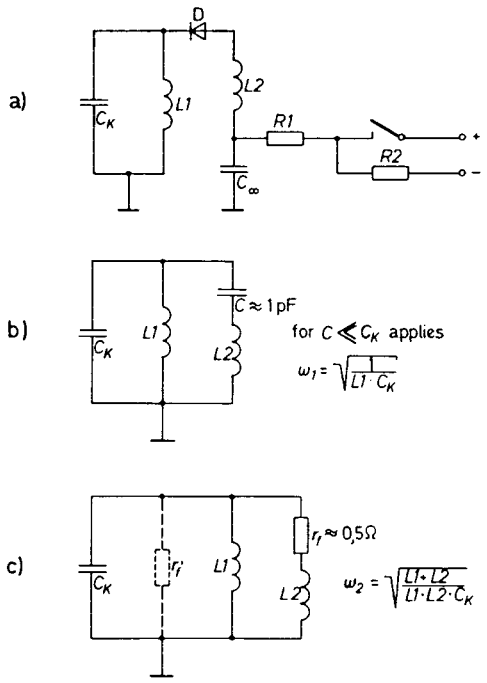


Figure 2-73 (a) Diode switch as a bandswitch, with an inductance in parallel; (b) equivalent circuit with blocked-diode switch; (c) equivalent circuit with conducting diode switch.

Instead of altering the resonant frequency of a tuned circuit by short-circuiting part of its inductance, the same result can be obtained by connecting a second inductance in parallel. The use of the diode switch for this purpose is illustrated in Figure 2-73. This differs from the one shown in Figure 2-71 in that the circuit inductance consists of either L_1 or of L_1 and L_2 in parallel. By applying a permanent negative bias to the diode via a high-value resistor R_2, it is possible to dispense with the use of one switch contact. Figure 2-73b shows the equivalent circuit for the diode under reverse-bias conditions (corresponding to a low resonant frequency), and Figure 2-73c shows the equivalent circuit under forward-bias conditions (corresponding to a higher resonant frequency).

In radio and television receivers, it usually is possible to derive the necessary reverse bias for the diode by rectification of the local oscillator signal, since the power required for blocking the diodes is extremely small. Note that under reverse-bias conditions the inductance L_2 and the capacitance C form a series-tuned acceptor circuit. However, this capacitance being extremely small, any undesirable effects due to its presence can normally be avoided. The same diode switching arrangements described for single-tuned circuits can also be employed in multisection bandpass and broadband filters.

In the following paragraph, a practical example for the use of diode switches will be given, and it will be seen that it is possible, in principle, to replace all of the mechanical RF switch contacts normally employed in receivers by diode switches. In cases where application of the diode dc potentials to a "cold" point in the circuit is not feasible, it may be necessary to apply them via *LC* or *RC* isolating networks.

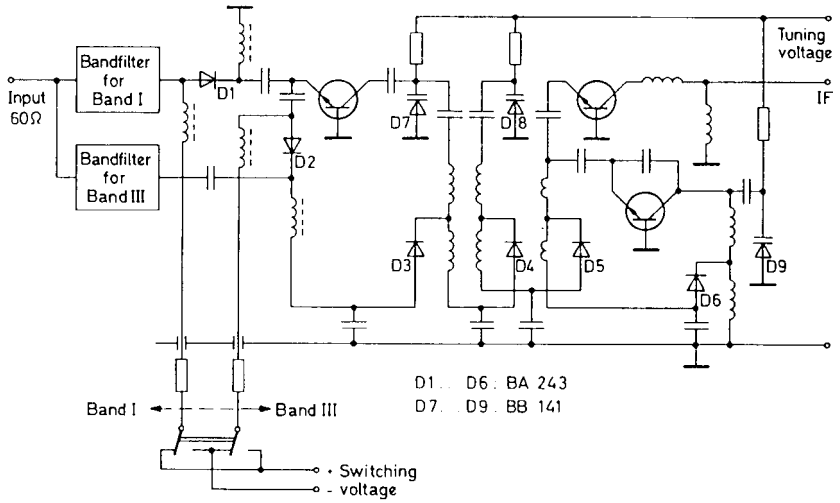


Figure 2-74 Schematic of a VHF television tuner with electronic band selection.

Use of the Diode Switch in a Television Receiver. Figure 2-74 shows the schematic of an electronically tuned and switched band I (48–62 MHz)/band III (175–224 MHz) TV tuner that requires switching of the input filter, bandpass filter, and oscillator circuit. These switching functions are performed by BA243 diode switches. As is customary in most European countries, in this example a broadband input filter between the antenna and the first RF transistor is used rather than a selective front-end circuit. In view of the need to correct noise and power matching over a relatively large frequency range, a switchable filter is employed in most tuners, switching being effected in this case by means of diodes D_1 and D_2.

Diodes D_3–D_6 are used for band-filter switching, for switching the coupling to the mixer transistor, and for switching the oscillator frequency. To minimize the bias-current requirements (each diode requires a bias current of approximately 10 mA), all diodes that are to be activated at the same time are connected in series. This arrangement requires only a few additional RF chokes, which are easily provided because of the high frequencies involved. The direct current required for the diode switches could be obtained by simple rectification, for example, from the heater circuit of the television receiver CRT, and the diode reverse bias could be obtained by rectification of the local oscillator signal, in which case the minute additional power requirement would have to be allowed for in the oscillator design.

A technically sophisticated design for a television tuner takes the form of a combined VHF/UHF unit incorporating electronically switched circuits in which several elements would be common. Such an all-band tuner can be assembled from only three transistors and a combined VHF/UHF band filter in $\lambda/2$ or $\lambda/4$ technology for UHF reception. The VHF circuits are short-circuited when the UHF position is selected.

Multistandard TV receivers require a large number of switch contacts in the tuner, the IF amplifier, the video amplifier, and the time base. The use of diode switches is particularly advantageous in this situation because it renders the physical layout of the equipment virtually independent of the location of the controls and obviates the need for long RF cable runs, which may be prone to interference.

2-2 BIPOLAR TRANSISTORS

2-2-1 Transistor Structure Types

All current RF silicon transistors are of the bipolar *npn* planar epitaxial type. Briefly, the significance of each of these terms is as follows.

Bipolar. In its broadest sense, it is the basic structure shown schematically in Figure 2-75, that is, the familiar three-semiconductor-region structure. *Bipolar* specifically means that the charge carriers of both negative (electron) and positive (hole) polarities are involved in the transistor action. In way of contrast, unipolar types include the junction-gate and insulated-gate field-effect transistors, which are basically one- or two-semiconductor-region structures in which carriers of a single polarity dominate.

npn. This abbreviation is for *negative–positive–negative* and identifies the regions of the structure as to polarity of the dominant or majority carrier in each region. The other polarity type is *pnp*. (See Figure 2-75.)

Silicon. Silicon (Si) is one of two elements from the fourth column of the periodic table that are in widespread use for transistor fabrication [the other is germanium (Ge)]. Other materials used include the compound gallium arsenide (GaAs). Although silicon is in predominant use because it results in the most favorable compromise among high-frequency, high-temperature, high-reliability, and ease-of-use attributes of the usable semiconductor material, silicon–germanium (SiGe) technology is increasingly important because it affords high-frequency performance unachievable with silicon alone.

Planar. The term planar denotes that both emitter–base and base–collector junctions of the transistor intersect the device surface in a common plane (hence, a better term might be *coplanar*). However, the real significance of the so-called planar structure is that the technique of diffusing dopants through an oxide mask, used in fabricating such a structure, results in junctions being formed beneath a protective oxide layer. These protected junctions are less prone to the surface problems sometimes associated with other types of structures, such as the mesa.

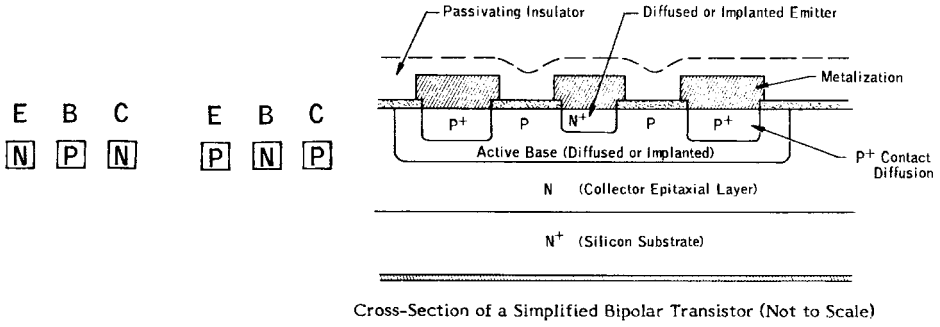


Figure 2-75 Transistor structure schematic.

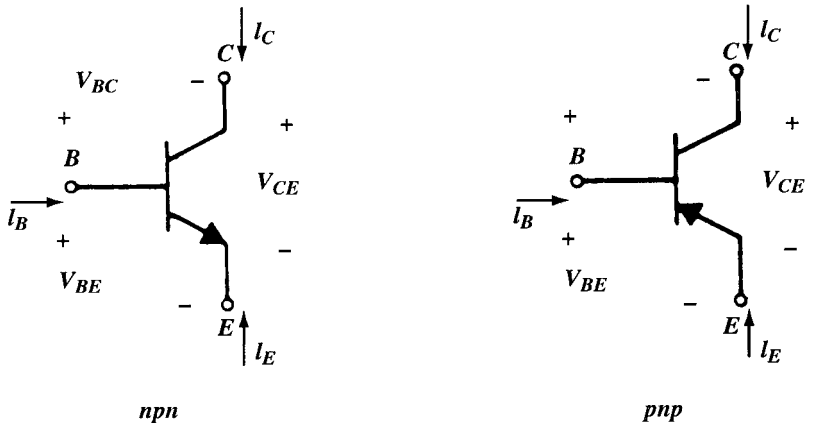


Figure 2-76 Bipolar transistor sign convention.

Epitaxial. This term, as it is commonly used, is actually a shortening of the term *epitaxial-collector*. That is, the collector region of the transistor is formed by the epitaxial technique, rather than by diffusion, which is commonly used to form the base and emitter regions. The epitaxial layer is formed by condensing a single-crystal film of semiconductor material on a wafer or substrate that is usually of the same material. Thus, an epitaxial (collector) transistor is one in which the collector region is formed on a low-resistivity substrate. Subsequently, the base and emitter regions are diffused into the "epi" layer. The epitaxial technique lends itself to precise tailoring of collector-region thickness and resistivity with consequent improved device performance and uniformity.

2-2-2 Large-Signal Behavior of Bipolar Transistors*

In this section, the large-signal or dc behavior of bipolar transistors is considered. Large-signal models are developed for the calculation of total currents and voltages in transistor circuits, and such effects as breakdown-voltage limitations, which are usually not included in models, are also considered. Second-order effects, such as current-gain variation with collector current and Early voltage, can be important in many circuits and are treated in detail.

The sign conventions used for bipolar transistor currents and voltages are shown in Figure 2-76. All bias currents for both *npn* and *pnp* transistors are assumed positive going into the device.

Electrical and Performance Characteristics and Specifications. Electrical characteristics may be described as uniquely defined, measurable electrical transistor properties that are not a function of the measuring circuit or apparatus (except insofar as standard terminations and measurement accuracy are concerned). Performance or operating characteristics are also electrical properties but they are, in general, not unique because their values depend on the measuring circuit (in particular, source and load impedance, which may be arbitrary). As might be expected, the terms are often used somewhat loosely (and sometimes

* Portions of this section, Sections 2-2-3 through 2-2-5, and Sections 2-3-3 through 2-3-6 are based on Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed., ©1993 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

interchangeably), especially in cases where only subtle differences are involved. The terms are generally used on transistor datasheets to segregate (for emphasis) under performance or operating characteristics those properties most directly applicable to the primary intended application.

Performance (Operating) Characteristics. Of the numerous performance characteristics that can be specified for high-frequency transistors, perhaps the most fundamental and pertinent characteristics are:

- Power gain and noise figure, for small-signal applications
- Power gain, power output, and efficiency, for large-signal applications

All of these characteristics are, of course, functions of frequency, bias temperature, and so on, and to completely characterize a transistor over its full frequency, bias, and temperature ranges would be prohibitively costly. Consequently, characterization data are given only for restricted ranges of these variables. These data should portray sufficiently the capabilities of a particular device for its intended applications. In the case of maximum ratings, some applications may require additional characterization by the user or through applications assistance from the manufacturer.

POWER GAIN

G_{max}. Of the various definitions for the measurement of power flow in an active two-port device, such as a transistor, two are unique enough to allow specification without recourse to specifying the complete measuring circuit in detail. One of these definitions is termed *maximum available gain* (G_{max}) and is the power gain obtained when the input and output ports are simultaneously conjugately matched to source and load impedances, respectively. Implicit in this definition is the assumption that the two-port is unconditionally stable; that is, that no combination of input/output tuning can result in increasing gain to the point of oscillation.

$|S_{21}|^2$. The other unique power gain is the gain realized when the transistor is inserted between a source and load with identical impedances (in practice, usually $50 + j0\ \Omega$). This particular insertion of transducer gain happens to coincide with the usual definition of the two-port forward scattering parameter, S_{21}. More precisely, it is equal to the magnitude squared of this parameter and is therefore often identified by the symbol $|S_{21}|^2$. For wideband applications, $|S_{21}|^2$ is important because wideband terminations "not too different from 50 Ω" are more easily realized than are wideband transforming networks that provide the matching required for G_{max}.

NOISE FIGURE. A common measure of the noise generated by an active two-port device—noise that sets a lower limit on amplifier sensitivity—is noise factor (F). This is defined as

$$F = \frac{\text{Input signal-to-noise ratio}}{\text{Output signal-to-noise ratio}} \tag{2-110}$$

or more generally as

$$F = \frac{\text{Total output noise power}}{\text{Output power due to source resistance}} \tag{2-111}$$

At high frequencies, spot noise factor or noise factor for a small fractional bandwidth (say, 1%) is used and is usually expressed as *noise figure* (*NF*), in decibels:

$$NF = 10 \log F \tag{2-112}$$

As already discussed, noise figure is a function of source impedance (as well as functions of frequency, bias, etc.) and hence, there is an infinity of noise figures associated with a given device corresponding to the infinity of possible impedances that may be presented to the device input. The only unique one, in the sense that it does not involve arbitrary source impedances, is $NF_{\min}$, the minimum noise figure obtained (at a given bias and frequency) when the input is tuned to optimize this parameter. This is the noise figure value commonly given on datasheets.

In practical amplifiers that involve more than one stage, *overall noise factor* (F_O) is given by

$$F_O = F_1 + \frac{F_2 - 1}{G_1} + \cdots + \frac{F_n - 1}{G_{(n-1)}} \tag{2-113}$$

where n = number of stages
G_n = gain of the nth stage
F_n = noise factor of the nth stage

This expression emphasizes the important fact that for low-noise amplifiers, the first stage must be designed for the lowest noise figure and highest gain possible. [Note that the noise contribution of the second stage is divided (reduced) by the gain of the first stage.] Since the optimum source impedance and bias currents for optimum gain and noise figure do not often coincide, very careful circuit design is required to minimize overall noise figure. For more accuracy, we recommend the noise correlation approach.

POWER OUTPUT. This characteristic is important for both amplifier and oscillator transistors. In both cases, it is extremely circuit sensitive. For amplifiers, maximum useful output is often limited to that power output level at which gain has compressed 1 dB, an indicator of the upper limit of the device's linearity range. For oscillators, it is merely a quantitative measure of RF power for given dc input power.

EFFICIENCY. In the most general sense, this characteristic expresses as a percentage the ratio of RF power output to the total circuit input power, both dc and RF. That expression is efficiency, η, defined as

$$\eta = \frac{P_O}{P_i + P_{\text{dc}}} \times 100 \tag{2-114}$$

where P_O = RF output power
P_i = RF input power
P_{dc} = total dc power input

Since for oscillator transistors there is no RF power input, and for amplifier transistors the maximum input RF power is calculable from the power gain and power output specifications,

the inclusion of P_i in efficiency is redundant. Moreover, since the major portion of the dc power is dissipated by the transistor collector, a more restricted definition of efficiency is pertinent. This parameter, termed *collector efficiency* (η_C), is given by

$$\eta_C = \frac{P_O}{P_{CC}} \times 100 \tag{2-115}$$

where $P_{CC} = V_{CC} \times I_{CC}$
V_{CC} = collector supply voltage
I_{CC} = collector supply current

Electrical Characteristics. Electrical characteristics may be conveniently classified into two main types, dc and ac.

dc CHARACTERISTICS. The importance of dc characteristics of high-frequency transistors lies primarily in biasing and reliability considerations. However, certain dc characteristics are also directly related to high-frequency performance. For example, high-frequency noise figure is affected by dc current gain. The dc characteristics discussed here are those usually found on high-frequency transistor datasheets.

$V_{(BR)CBO}$, I_{CBO}. These two parameters serve to characterize the reverse-biased collector–base *pn* junction and are defined as follows (with the aid of Figure 2-77a). The collector–base breakdown voltage, $V_{(BR)CBO}$, identifies the voltage at which collector current tends to increase without limit, usually due to the high electric field developed across the junction. This voltage sets a limit on the maximum transistor operating voltage and, as mentioned before under maximum ratings, usually is the basis for the collector–base maximum voltage rating. $V_{(BR)CBO}$ should be specified at a value of $I_C = I_{C1}$ in the figure, which is within the avalanche (or high slope) region of the device's reverse characteristic. Typical values of I_{C1} are in the 1–10-μA region for high-frequency transistors.

To further define the quality of the reverse *I–V* characteristic, a specification is usually placed on the collector cutoff current, I_{CBO}, which is measured at some value of collector–base voltage less than $V_{(BR)CBO}$. For a good-quality silicon junction ("sharp" instead of soft; see Figure 2-77a), I_{CBO} is in the nanoampere range.

$V_{(BR)EBO}$, I_{EBO}. These two parameters characterize the reverse-biased emitter–base *pn* junction in an analogous manner to the collector–base junction parameters $V_{(BR)CBO}$ and I_{CBO}, given above, and are shown in Figure 2-77b. No further discussion of these parameters will be given here.

$V_{(BR)CEO}$, I_{CEO}. The collector–emitter breakdown voltage and cutoff current are somewhat more complex in nature than the collector–base and emitter–base parameters. In the latter two, only a single *pn* diode is involved. In the collector–emitter case, two diodes are involved. Moreover, each is influenced by the other through transistor action, since the reverse current of the collector–base diode flows through the emitter–base junction as forward current. Thus, the collector–base reverse current is amplified by the dc current gain of the transistor, resulting in:

- I_{CEO} being greater than I_{CBO} (for a given voltage)
- Typically, the familiar negative-resistance region in the *I–V* characteristic as shown in Figure 2-77c.

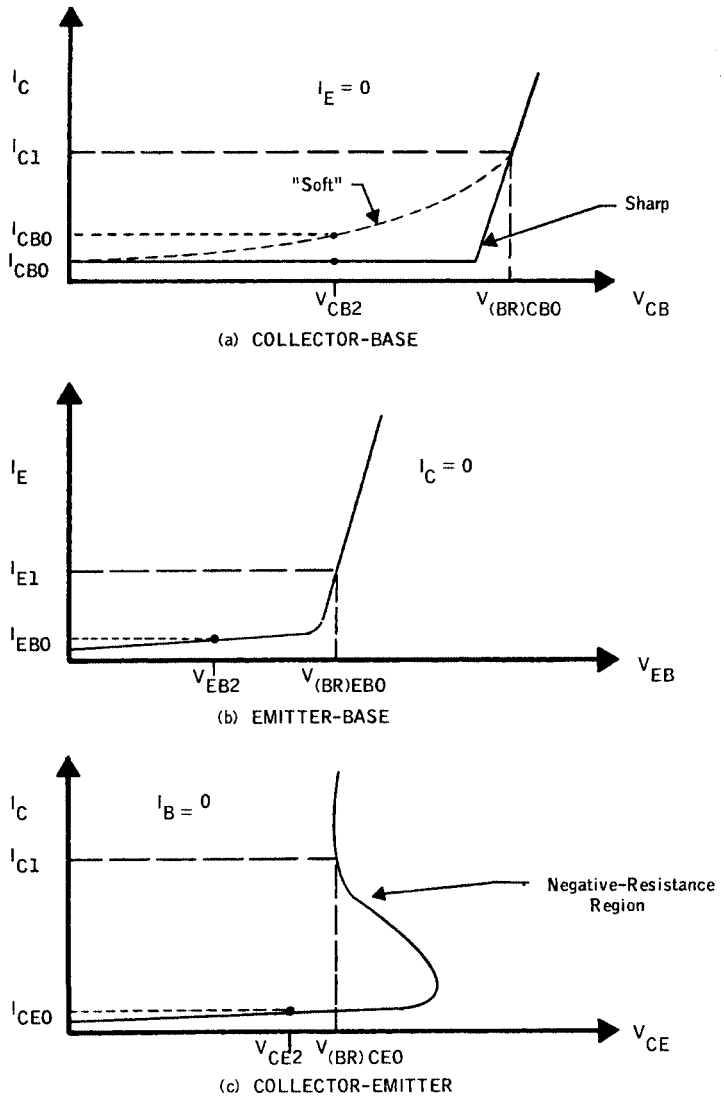


Figure 2-77 Transistor reverse *I–V* characteristics.

Consequently, $V_{(BR)CEO}$ is typically specified at collector currents one to three orders of magnitude higher than in the case of $V_{(BR)CBO}$ and $V_{(BR)EBO}$ to establish the minimum value of this characteristic.

h_{fe}. This parameter is simply the dc common-emitter current gain; that is, the ratio of collector current to base current at some specified collector voltage and current.

ac CHARACTERISTICS. Of the numerous ac characteristics defined for transistors, only relatively few are commonly used in characterizing high-frequency transistors. Some of the more pertinent parameters are briefly covered here.

S Parameters. By far the most useful and conveniently measured set of two-port parameters for transistor high-frequency (roughly 100 MHz and above) characterization is

the *S*-parameter or scattering-matrix set. These parameters completely and uniquely define the small-signal gain and input/output immittance properties of any linear "black box." (By definition, a transistor or any active device is linear under small-signal conditions.) However, these parameters reveal nothing (except possibly indirectly and approximately) about large-signal or noise behavior. Simply interpreted (more general definitions and other interpretations abound in the technical literature), the *S* parameters are merely insertion gains, forward and reverse, and reflection coefficients, input and output, with driven and nondriven ports both terminated in equal impedances, usually $50 + j0\ \Omega$. Such an interpretation tends to make *S* parameters very attractive, once some familiarity is gained, at high (especially microwave) frequencies, since the power flow or gain and reflection-coefficient concepts are more intuitively meaningful than voltage and current conceptual schemes. It should also be mentioned that *S*-parameters can be converted through straightforward matrix transformations to other two-port parameter sets, such as *h*-, *y*-, and *z*-parameters.

Proceeding with more specific definitions, the *S*-parameters are defined analytically by

$$b_1 = S_{11}a_1 + S_{12}a_2 \tag{2-116}$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \tag{2-117}$$

or, in matrix form,

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{2-118}$$

where, referring to Figure 2-78

a_1 = (incoming power at Port 1)$^{1/2}$
b_1 = (outgoing power at Port 1)$^{1/2}$
a_2 = (incoming power at Port 2)$^{1/2}$
b_2 = (outgoing power at Port 2)$^{1/2}$
E_1, E_2 = electrical stimuli at Port 1, Port 2

From the figure and defining linear equations, for $E_2 = 0$, then $a_2 = 0$, and (skipping through numerous rigorous steps)

$$S_{11} = \frac{b_1}{a_1}$$

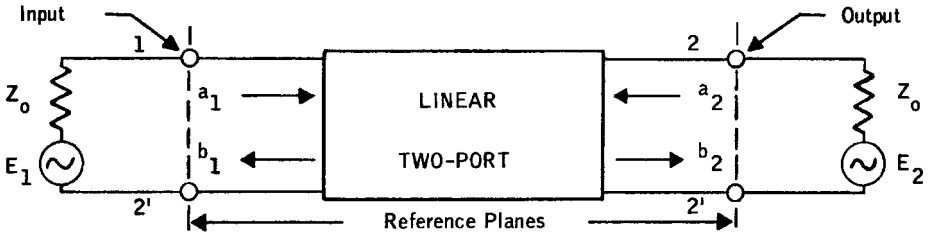


Figure 2-78 Two-port *S*-parameter definition.

$$= \left(\frac{\text{Outgoing input power}}{\text{incoming input power}}\right)^{1/2}$$

$$= \frac{\text{Reflected voltage}}{\text{Incident voltage}}$$

$$= \text{Input reflection coefficient} \tag{2-119}$$

$$S_{21} = \frac{b_2}{a_1}$$

$$= \left(\frac{\text{Outgoing output power}}{\text{incoming input power}}\right)^{1/2}$$

$$= \left(\frac{\text{Output power}}{\text{Available input power}}\right)^{1/2}$$

$$= (\text{Forward transducer gain})^{1/2} \tag{2-120}$$

or, more precisely in the case of S_{21}:

$$\text{Forward transducer gain} = G_{\text{TF}} = |S_{21}|^2 \tag{2-121}$$

$$Z_i = Z_o \tag{2-122}$$

Similarly, at Port 2 for $E_1 = 0$, then $a_1 = 0$, and

$$S_{22} = \frac{b_2}{a_2} = \text{Output reflection coefficient} \tag{2-123}$$

$$S_{12} = \frac{b_1}{a_2} = (\text{Reverse transducer gain})^{1/2} \tag{2-124}$$

$$G_{\text{TR}} = |S_{12}|^2 \tag{2-125}$$

Since many measurement systems display *S*-parameter magnitudes in decibels, the following relationships are particularly useful:

$$|S_{11}|_{\text{dB}} = 10\log|S_{11}|^2$$

$$= 20\log|S_{11}| \tag{2-126}$$

$$|S_{22}|_{\text{dB}} = 20\log|S_{22}| \tag{2-127}$$

$$|S_{21}|_{\text{dB}} = 10\log|S_{21}|^2$$

$$= 20 \log|S_{21}|$$

$$= 10 \log|G_{TF}| = |G_{TF}|_{dB} \tag{2-128}$$

$$|S_{12}|_{dB} = 10 \log|S_{12}|^2$$

$$= 20 \log|S_{12}|$$

$$= 10 \log|G_{TR}| = |G_{TR}|_{dB} \tag{2-129}$$

Figure 2-79 shows an example of the use of *S* parameters in transistor stage design.

Transition Frequency. One of the better known, but perhaps least understood, figures of merit for high-frequency transistors is the so-called transition frequency, f_T. Part of the misunderstanding that appears to exist is due to the use of the misleading (but common, for

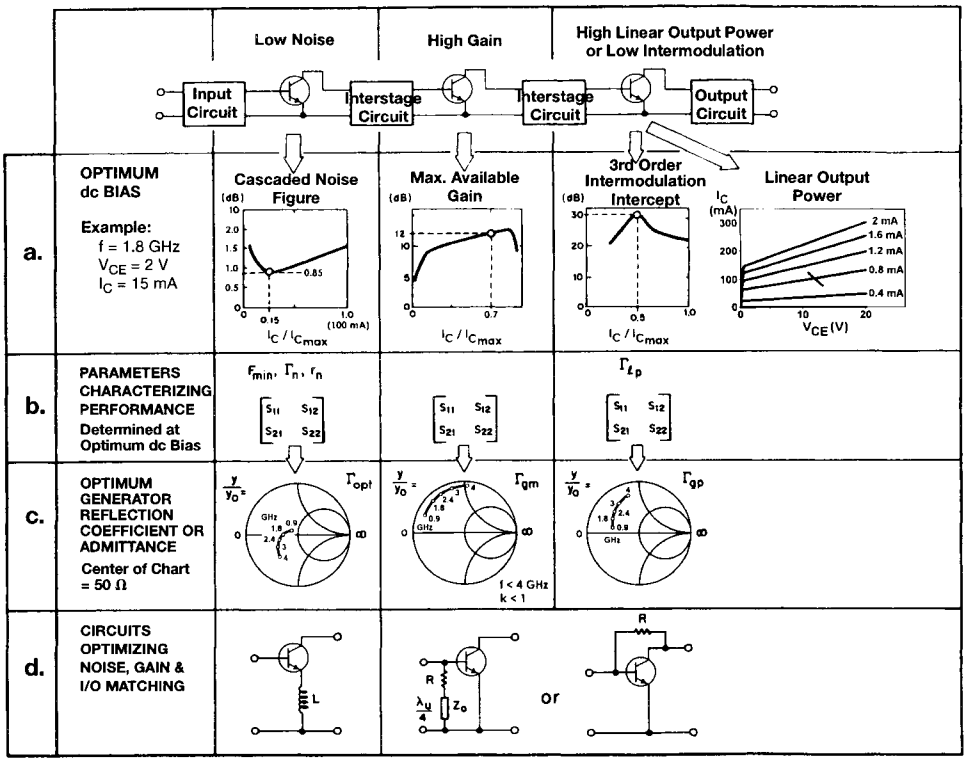


Figure 2-79 Key parameters in applying a BJT in low-noise front-end, high-gain, and linear-power stages. (a) The BJT characteristics that lead to optimum dc bias for a low-noise front end; (b) parameters characterizing device performance; (c) Γ_{opt} for the first stage, and S_{11} and S_{22} for the output stage as simulated by CAD software. The example is based on the Siemens BJTs BFP420 (low-noise stage), BFP450 (high-gain stage), and BFG235 (output stage). Circuits to optimize stage noise, gain, and I/O matching are shown in (d). (Presentation based on Figure 14 in Charles A. Liechti, "Microwave Field-Effect Transistors—1976," *IEEE Trans. Microwave Theory Tech.,* vol. MTT-24, pp. 279–300, June 1976.)

historical reasons) term *short-circuit gain–bandwidth product* for this parameter. By definition, f_T is that characteristic frequency described by the equation

$$f_T = h_{fe} \times f_{\text{meas}} \tag{2-130}$$

where h_{fe} is the magnitude of the small-signal common-emitter short-circuit current gain, and f_{meas} is the frequency of measurement, chosen such that

$$2 \le h_{fe} \le \frac{h_{fe0}}{2} \tag{2-131}$$

where h_{fe0} is the low-frequency value of h_{fe}.

To varying degrees of approximation, depending on the transistor type, f_T is the frequency at which h_{fe} approximates unity. It is not, in general, the frequency at which h_{fe} is precisely equal to unity. To clarify these points further, consider the plot of h_{fe} against frequency sketched in Figure 2-80.

At low frequencies, $f << f_B$, h_{fe} is constant and equal to h_{fe0}.

At f_B, h_{fe} has decreased to 0.707 h_{fe0}; that is, f_B is the 3-dB cutoff frequency for common-emitter short-circuit current gain, h_{fe}.

For frequencies such that

$$2f_B < f < f_T \tag{2-132}$$

h_{fe} varies inversely proportional to frequency. That is, the f_T defining relationship holds:

$$h_{fe} \times f_{\text{meas}} = \text{constant} = f_T \tag{2-133}$$

At frequencies approaching f_T, other parameters, especially package parasitics, can cause $|h_{fe}|$ to depart significantly from this $1/f$ variation. Therefore, the frequency, f_1, at which $|h_{fe}|$ actually equals unity can be somewhat different from f_T.

Applying these frequency-gain characteristics to the common-emitter wideband, low-pass amplifiers gives rise to the terminology of f_T being a "gain–bandwidth product." However,

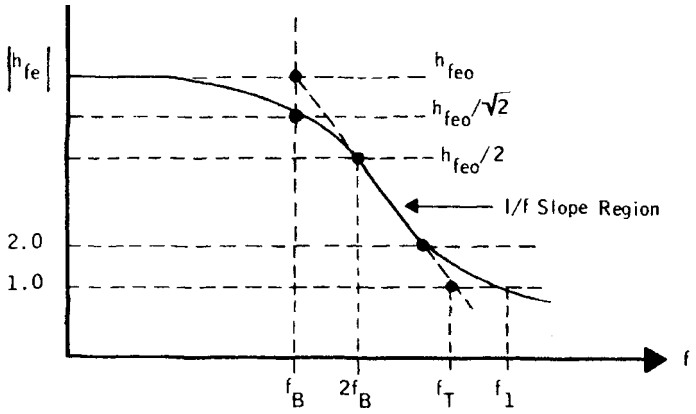


Figure 2-80 $|h_{fe}|$ frequency characteristics.

this is an optimistic approximation at best, since the product of low-frequency circuit gain and the 3-dB cutoff frequency is reduced from f_T by an amount that depends on circuit impedances.

The real significance of f_T lies in the fact that it is a measure of certain internal transistor parameters that do, in fact, affect high-frequency performance: for example, gain (though not in the convenient quantitative manner implied by the term "gain–bandwidth product"). In particular, good high-frequency noise performance requires that f_T be high. Thus, f_T is included on transistor datasheets primarily as a figure of merit, not as a parameter to be used directly in design.

Collector–Base Time Constant, $r_b'C_c$. This is an internal device parameter that relates only indirectly to high-frequency performance. It is primarily a measure of internal feedback within the transistor. It also relates to transistor high-frequency impedance. As the name says (in symbols), it is a measure of transistor base resistance and collector capacitance in combination; however, except for certain low-frequency transistors, it cannot be considered the simple two-element lumped *RC* time constant implied by the terminology. (In high-frequency transistors, base resistance and collector impedance must be considered to be distributed when considered in detail). As a figure of merit, it is included on transistor datasheets to indicate how well base resistance and collector impedance have been minimized. It also allows the estimation of certain gain properties of the transistor (see the $f_{\max}$ parameter, which we'll describe shortly).

Collector–Base Capacitance, C_{cb}. This parameter is simply the total collector–base *pn* junction capacitance measured at a low frequency (typically, 1 MHz), where it can be considered a single lumped element. For high-frequency transistors it is, of course, desirable that C_{cb} be small from bandwidth and stability considerations as well as from gain considerations alone.

Maximum Frequency of Oscillation, $f_{\max}$. This is another figure-of-merit parameter, as opposed to measurable parameters directly usable in the application of transistors. Its importance stems from the following approximate relationships (which will not be derived here):

$$f_{\max} \approx \left(\frac{f_T}{8\pi r_b' C_C} \right)^{1/2} \tag{2-134}$$

$$G_{\max} \approx \left(\frac{f_{\max}}{f_{\text{oper}}} \right)^2 \tag{2-135}$$

These expressions illustrate in a quantitative way the importance and the interrelationship between high f_T and low $r_b'C_C$ insofar as high-frequency gain is concerned. However, since they are approximations and since their derivation involves several assumptions that are not always valid, they must be interpreted with caution. For example, the expression for $G_{\max}$ is obviously not applicable at low frequencies, since as $f \to 0$, $G_{\max} \to \infty$, according to this expression. As a rule of thumb, the $G_{\max}$ expression is a reasonable approximation for frequencies such that

$$5 > \frac{f_{\max}}{f_{\text{oper}}} > 1 \tag{2-136}$$

For accurate analysis of transistor gain and stability, a complete set of two-port parameters must be employed in exact expressions, such as those from which the approximations shown above were derived.

Datasheet Transistor Specifications. The following is a reprint of a Siemens microwave transistor datasheet. This 25-GHz f_T transistor will soon be replaced by a 40-GHz family, based on a silicon–germanium design. Key parameters describing the transistor are:

- The dc parameters, such as maximum current and voltage and dissipation
- Noise figure as a function of generator impedance, bias point, and frequency
- Frequency-dependent behavior, typically expressed in *S* parameters, and also as a function of bias point and frequency
- Large-signal performance to be calculated based on a nonlinear model, such as the Gummel–Poon

The transistor description typically gives details about packaging, as the Gummel–Poon model is intrinsic (without parasitics). Not only does the package determine the parasitics, it also affects the device's ability to transfer heat to its heat sink. The following figures, reproduced with permission from the datasheet for the BFP420 transistor by Siemens (now Infineon Technologies), are a typical representation of the dc-I_C curves, *S* parameters, and noise parameters of a normal, packaged microwave transistor. Besides its dc parameters and capacitance, there are other critical parameters. A BJT's noise figure is a strong function of its dc current gain, and one can determine a 3-dB cutoff frequency referred to as f_T, which shows the gain roll-off as a function of collector current and frequency. Another important measure or figure of merit is $f_{\max}$, the maximum frequency of oscillation:

$$f_{\max} = \sqrt{\frac{|Y_{21}|}{16\pi R_{bb} C_{be} C_{bc}}} \tag{2-137}$$

Figure 2-81 shows a photograph of the die for the BFP420.

Since the intercept point is an important figure of merit, Figure 2-82 shows intercept point as a function of bias point for the BFR93AW BJT. We wish more manufacturers would supply this information.

2-2-3 Large-Signal Transistors in the Forward-Active Region

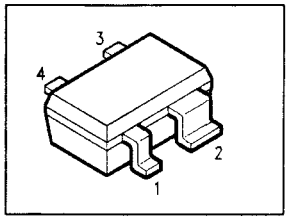
We saw in the diode case that its description is physics-based and it is, therefore, logical to apply the same modeling for the transistors. The first complete modeling for transistors was described in the Gummel–Poon model, but many engineers use the somewhat more simpler Ebers–Moll model. By combining two diodes, the parasitic elements, and the base separating resistor, we can show the "intrinsic" nonlinear model for the bipolar transistor.

Many equations that have been published contain some approximations, but for exact nonlinear analysis, we have to resort to the somewhat more complex or complete equations as shown in Appendix A for the bipolar model.

Recently, some other models have been proposed by researchers, such as the MEXTRAM models (Philips), but the trouble with these models is that while they may give quite good results for RF applications, it is virtually impossible to obtain parameters for them. So far the only industry-standard model is the Gummel–Poon, and to the authors' knowledge there

SIEMENS **SIEGET®25 BFP420**

NPN Silicon RF Transistor



- For High Gain Low Noise Amplifiers
- For Oscillators up to 10 GHz
- Noise Figure F = 1.05 dB at 1.8 GHz
  Outstanding G_{ms} = 20 dB at 1.8 GHz
- Transition Frequency f_T = 25 GHz
- Gold metallization for high reliability
- **SIEGET®25-Line**
  Siemens **G**rounded **E**mitter **T**ransistor-
  25 GHz f_T-**Line**

ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering Code (8-mm taped)	Pin Configuration 1	2	3	4	Package [1]
BFP420	AMs	Q62702-F1591	B	E	C	E	SOT343

Maximum Ratings

Parameter	Symbol		Unit
Collector-emitter voltage	V_{CEO}	4.5	V
Collector-base voltage	V_{CBO}	15	V
Emitter-base voltage	V_{EBO}	1.5	V
Collector current	I_C	35	mA
Base current	I_B	3	mA
Total power dissipation, Ts ≤ 107°C [2)3)]	P_{tot}	160	mW
Junction temperature	T_j	150	°C
Ambient temperature range	T_A	-65...+150°C	°C
Storage temperature range	T_{stg}	-65...+150°C	°C

Thermal Resistance

Junction-soldering point [2)]	$R_{th\ JS}$	270	K/W

1) For detailed information see chapter Package

2) T_S is measured on the emitter lead at the soldering point to the pcb.

3) P_{tot} due to Maximum Ratings.
At typical Ts ≤ 80°C: P_{tot} = 250 mW due to thermical characteristics.

SIEMENS **SIEGET®25 BFP420**

Electrical Characteristics

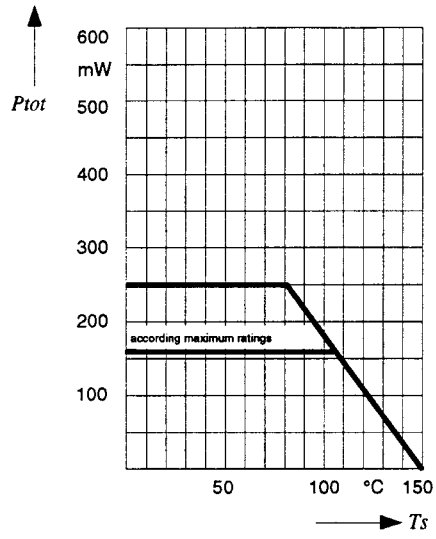
at T_A = 25 °C, unless otherwise specified.

Parameter	Symbol	Value			Unit
		min.	typ.	max.	
DC Characteristics					
Collector-emitter breakdown voltage I_C = 1 mA	$V_{(BR)CEO}$	4.5	5	6.5	V
Collector-cutoff current V_{CB} = 5 V, I_E = 0	I_{CBO}	-	-	200	nA
Emitter base cuttoff current V_{EB} = 1.5 V, I_C = 0	I_{EBO}	-	-	35	µA
DC current gain I_C = 20 mA, V_{CE} = 4 V	h_{FE}	50	80	150	
AC Characteristics					
Transition frequency I_C = 30 mA, V_{CE} = 3 V, f = 2 GHz	f_T	20	25	-	GHz
Collector-base capacitance V_{CB} = 2 V, V_{BE} = v_{be} = 0, f = 1 MHz	C_{cb}	-	0.15	0.24	pF
Collector-emitter capacitance V_{CE} = 2 V, V_{BE} = v_{be} = 0, f = 1 MHz	C_{ce}	-	0.41	-	pF
Emitter-base capacitance V_{EB} = 0.5 V, V_{CB} = v_{cb} = 0, f = 1 MHz	C_{eb}	-	0.55	-	pF
Noise figure I_C = 5 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}	F	-	1.05	1.4	dB
Power gain I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt} , Z_L = Z_{Lopt}	G_{ms} [1]	-	20	-	dB
Insertion power gain I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_L = 50Ω	$\lvert S_{21}\rvert^2$	14	17	-	dB
Third order intercept point at output I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}, Z_L = Z_{Lopt}	IP_3	-	22	-	dBm
1dB Compression point I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}, Z_L = Z_{Lopt}	P_{-1dB}	-	12	-	dBm

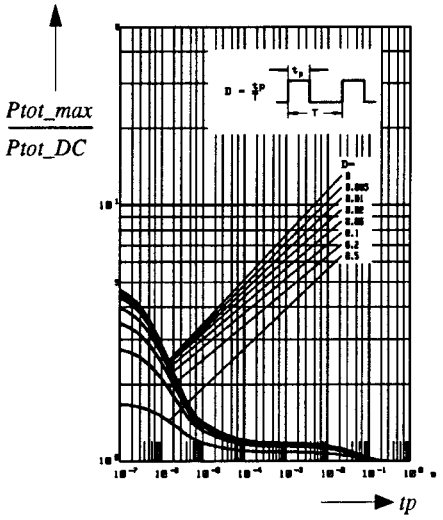
1) $G_{ms} = \left|\frac{S21}{S12}\right|$

SIEMENS **SIEGET®25 BFP420**

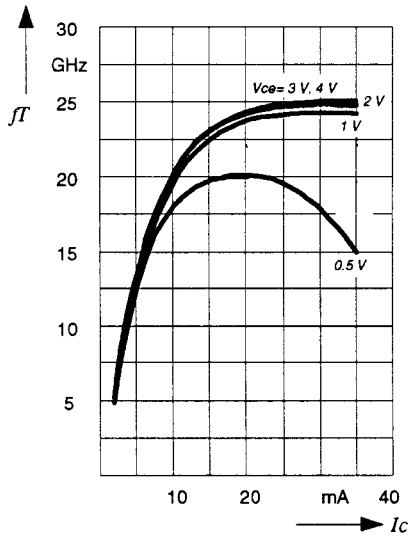
Total Power Dissipation
versus Soldering Point Temperature



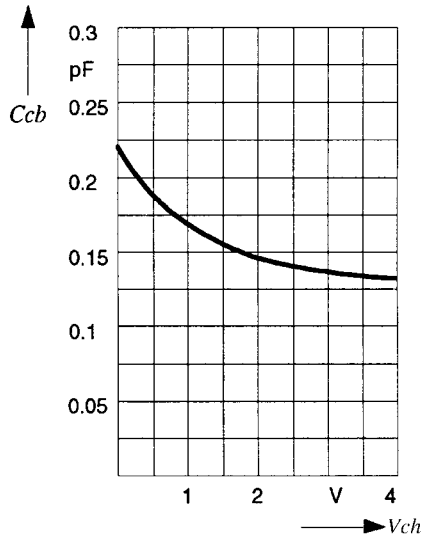
Permissible Pulse Power Dissipation
versus On-Time (V_{CE0max} = 4.5 V)



Transition Frequency
versus Collector Current
f = 2 GHz



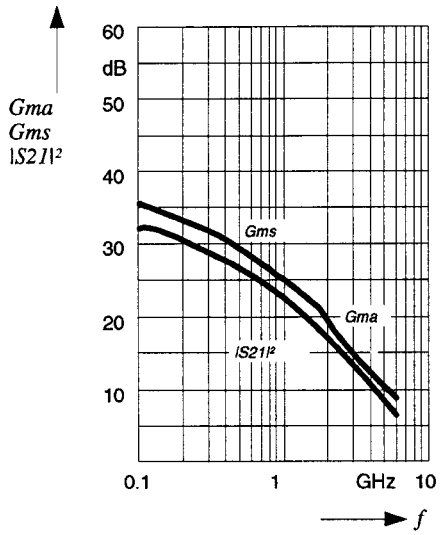
Collector-base Capacitance
versus Collector-base Voltage
V_{BE} = 0 V, f = 1MHz



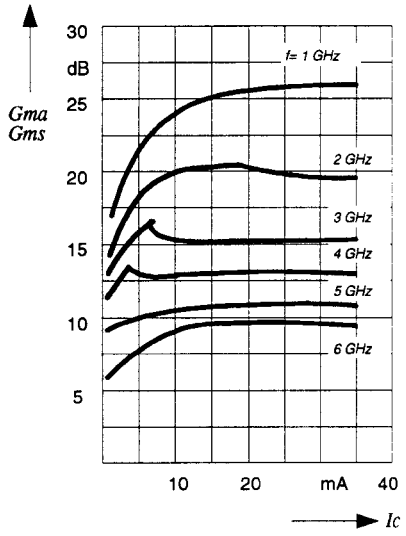
SIEMENS

SIEGET®25 BFP420

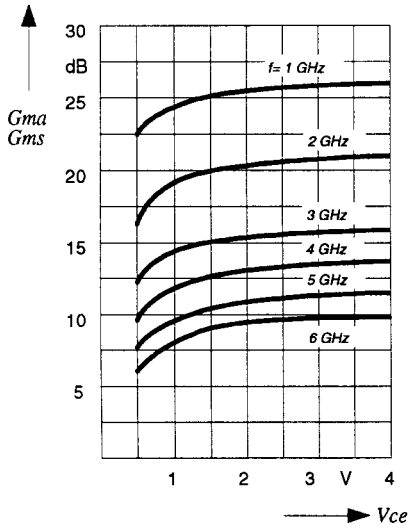
Power Gain
versus Frequency
V_{CE} = 2 V , I_C = 20 mA



Power Gain
versus Collector Current
V_{CE} = 2 V

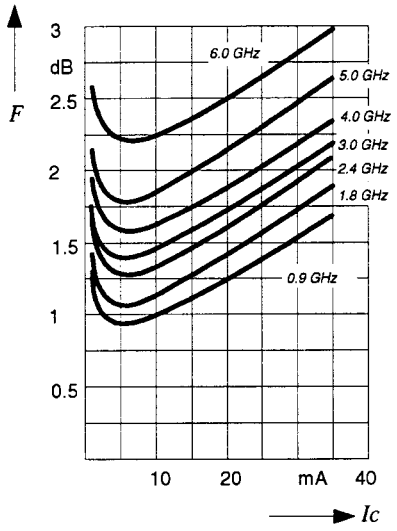


Power Gain
versus Collector Voltage
I_C = 20 mA

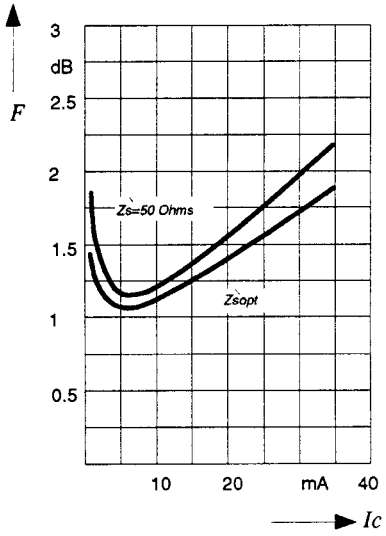


SIEMENS SIEGET®25 BFP420

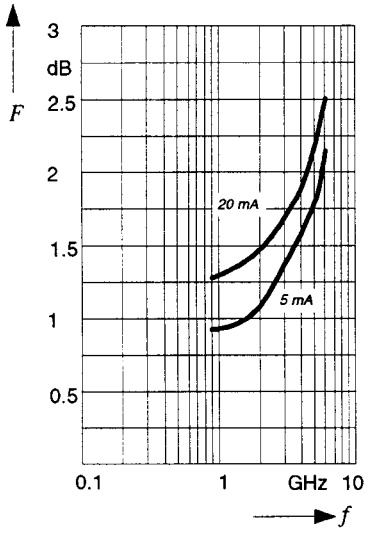
Noise Figure
versus Collector Current
V_{CE} = 2 V, Z_S = Z_{Sopt}



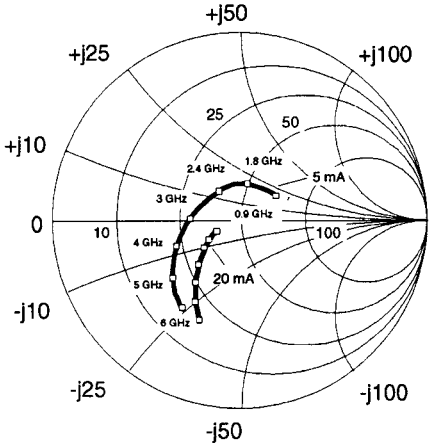
Noise Figure
versus Collector Current
VCE = 2 V, f = 1.8 GHz



Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA,
Z_S = Z_{Sopt}



Source Impedance for min.
Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA



SIEMENS **SIEGET®25 BFP420**

Common Emitter S-Parameters

f	S_{11}		S_{21}		S_{12}		S_{22}	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
V_{CE} = 2 V, I_C = 20 mA								
0.01	0.452	-2.3	37.62	178.3	0.0011	94.4	0.956	-0.6
0.1	0.447	-25.1	36.30	164.7	0.0068	82.5	0.941	-12.4
0.5	0.386	-101.1	23.41	121.0	0.0262	61.7	0.632	-47.2
1.0	0.378	-146.2	13.99	96.0	0.0395	57.8	0.395	-63.9
2.0	0.405	173.5	7.18	70.8	0.0664	54.0	0.222	-87.3
3.0	0.446	149.4	4.77	52.6	0.0949	47.1	0.133	-111.3
4.0	0.501	130.0	3.52	36.8	0.1206	38.5	0.133	-158.5
6.0	0.599	104.8	2.27	8.2	0.1646	18.9	0.196	142.0
8.0	0.700	78.5	1.51	-20.8	0.1800	-2.4	0.289	99.3
9.0	0.758	67.6	1.25	-34.4	0.1820	-13.0	0.379	84.1
10.0	0.800	62.0	1.04	-43.5	0.1800	-19.3	0.465	76.6
V_{CE} = 2 V, I_C = 5 mA								
0.01	0.790	-1.0	15.14	179.2	0.0012	83.4	0.988	-0.7
0.1	0.786	-11.6	14.98	171.8	0.0092	84.1	0.982	-6.5
0.5	0.702	-55.7	12.86	140.1	0.0398	62.8	0.857	-29.8
1.0	0.589	-99.1	9.63	112.6	0.0603	46.5	0.647	-48.6
2.0	0.507	-156.0	5.60	79.4	0.0798	34.6	0.401	-70.3
3.0	0.511	168.5	3.84	57.1	0.0957	29.8	0.267	-84.2
4.0	0.549	142.0	2.87	38.5	0.1121	25.1	0.207	-110.5
5.0	0.604	123.9	2.26	22.1	0.1285	19.4	0.150	-137.3
6.0	0.633	110.0	1.86	6.7	0.1442	13.1	0.173	-169.8

Common Emitter Noise Parameters

f	F_{min} 1)	G_a 1)	Γ_{opt}		R_N	r_n	$F_{50\Omega}$ 2)	$\lvert S_{21}\rvert^2$ 2)
GHz	dB	dB	MAG	ANG	Ω	-	dB	dB
V_{CE} = 2 V, I_C = 5 mA								
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	-175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	-157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	-142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	-123.0	15.0	0.30	3.30	5.3

1) Input matched for minimum noise figure, output for maximum gain 2) Z_S=Z_L=50 Ω

SIEMENS **SIEGET®25 BFP420**

SPICE Parameters:

Transistor Chip Data T502 (Berkeley-SPICE 2G.6 Syntax):

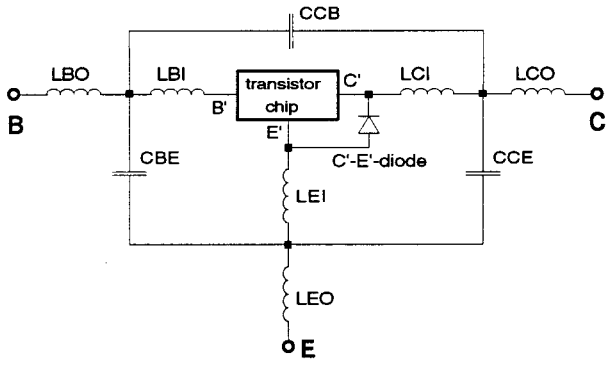
IS =	0.20045	fA	BF =	72.534	-	NF =	1.2432	-
VAF =	28.383	V	IKF =	0.48731	A	ISE =	19.049	fA
NE =	2.0518	-	BR =	7.8287	-	NR =	1.3325	-
VAR =	19.705	V	IKR =	0.69141	A	ISC =	0.019237	fA
NC =	1.1724	-	RB =	3.4849	OHM	IRB =	0.72983	mA
RBM =	8.5757	OHM	RE =	0.31111	OHM	RC =	0.10105	OHM
CJE =	1.8063	fF	VJE =	0.8051	V	MJE =	0.46576	-
TF =	6.7661	ps	XTF =	0.42199	-	VTF =	0.23794	V
ITF =	1.0	mA	PTF =	0	deg	CJC =	234.53	fF
VJC =	0.81969	V	MJC =	0.30232	-	XCJC =	0.3	-
TR =	2.3249	ns	CJS =	0	fF	VJS =	0.75	V
MJS =	0	-	XTB =	0	-	EG =	1.11	eV
XTI =	3.0	-	FC =	0.73234	-	Tnom =	300	K

C'-E'- Diode Data (Berkeley-SPICE 2G.6 Syntax):

IS =	3.5	fA	N =	1.02	-	RS =	10	Ω

All parameters are ready to use, no scaling is necessary. Caution: In these extracted parameters, RBM > the operating RB.

Package Equivalent Circuit:



SOT343-3:

LBI = 0.47 nH
LBO = 0.53 nH
LEI = 0.23 nH
LEO = 0.05 nH
LCI = 0.56 nH
LCO = 0.58 nH
CBE = 136 fF
CCE = 134 fF
CCB = 6.9 fF

valid up to 6 GHz

The SOT343 package has two emitter leads. To avoid high complexity of the package equivalent circuit, both leads are combined in one electrical connection.

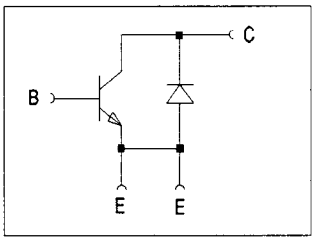
Extracted on behalf of SIEMENS Small Signal Semiconductors by:
Institut für Mobil- und Satellitenfunktechnik (IMST)
© 1996 SIEMENS AG

For non-linear simulation:

- Use transistor chip parameters in Berkeley SPICE 2G.6 syntax for all simulators.
- If you need simulation of the reverse characteristics, add the diode with the C'-E'-diode data between collector and emitter.
- Simulation of the package is not necessary for frequencies < 100 MHz.
  For higher frequencies add the wiring of the package equivalent circuit around the non-linear transistor and diode model.

Note:

- This transistor is constructed in a common emitter configuration. This feature causes an additional, reverse biased diode between emitter and collector, which does not effect normal operation.



Transistor Schematic Diagram

The common emitter configuration shows the following advantages:

- Higher gain because of lower emitter inductance.
- Power is dissipated via the grounded emitter leads, because the chip is mounted on the copper emitter leadframe.

Please note, that the broadest lead is the emitter lead.

- The AC-Characteristics are verified by random sampling.

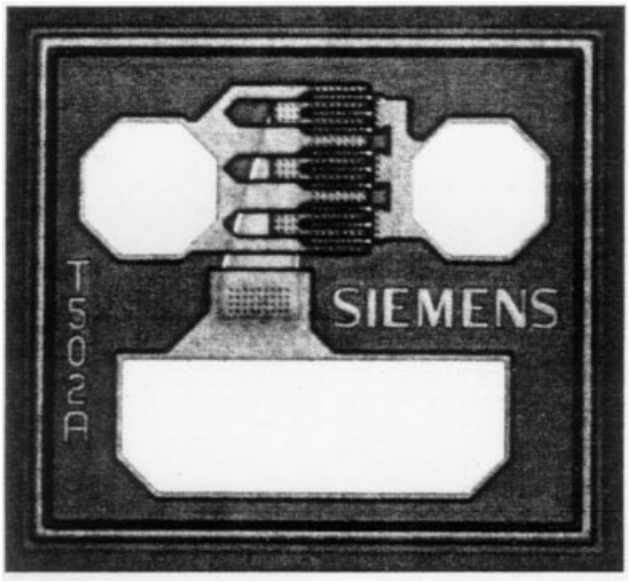


Figure 2-81 Photograph of the BFP420 die (310 × 290 μm). The bonding pad for the base is at the upper right; for the collector at the upper left. The large, lower pad connects to the emitter.

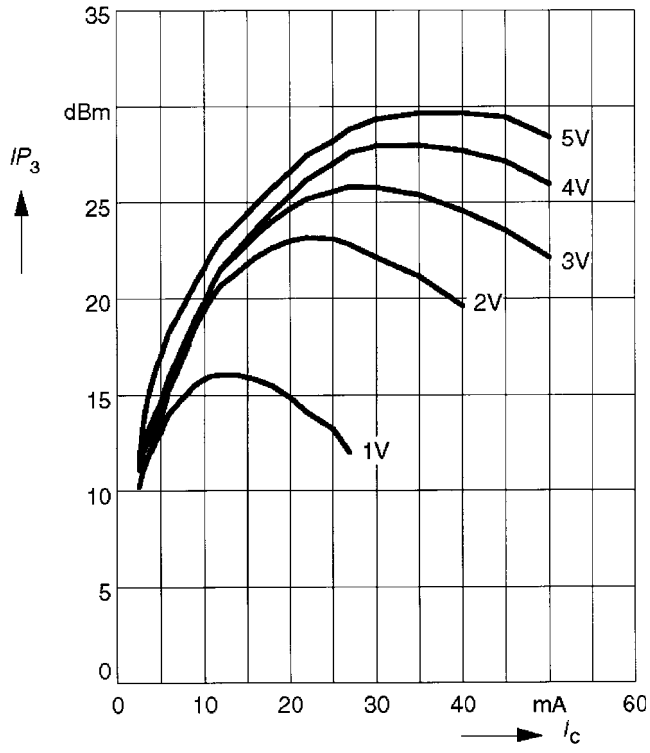


Figure 2-82 $IP_{3,out}$ versus I_C at five values of V_{CE} for the BRF93AW BJT.

are no accepted or affordable parameter extractor programs available, not even to consider the cost of hardware for such extractions. Hewlett-Packard recently published a system with a hardware cost in the vicinity of $1 million.

A typical *npn* planar bipolar transistor structure is shown in Figure 2-83a, where collector, base, and emitter are labeled *C*, *B*, and *E*, respectively. The impurity doping density in the base and the emitter of such a transistor is not constant but varies with distance from the top surface. However, many of the characteristics of such a device can be predicted by analyzing

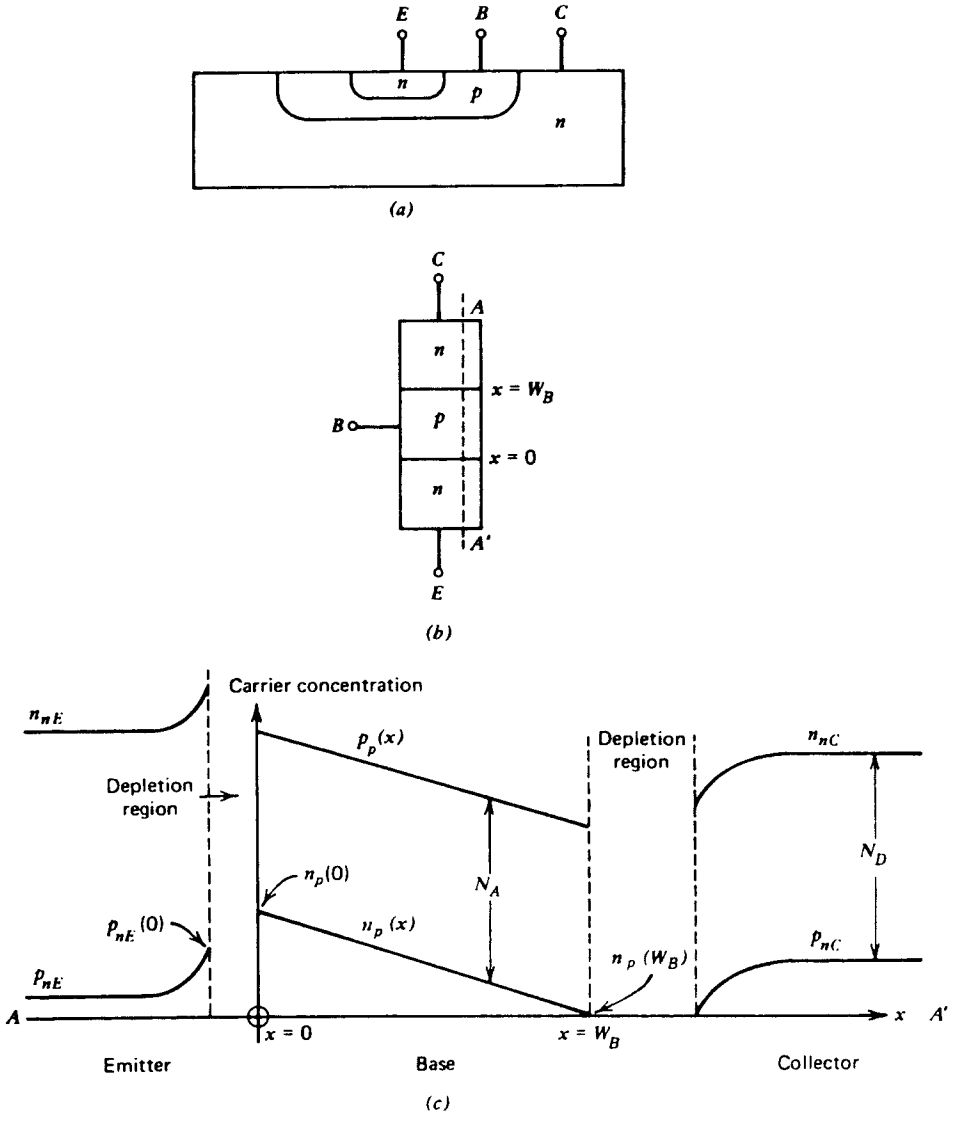


Figure 2-83 (a) Cross section of a typical *npn* planar bipolar transistor structure; (b) idealized transistor structure; (*c*) carrier concentrations along the cross section *AA*′ of the transistor in (b). Uniform doping densities are assumed. (Not to scale)

the idealized transistor structure shown in Figure 2-83b. In this structure, the base and emitter doping densities are assumed constant, and this is sometimes called a "uniform-base" transistor. Where possible in the following analyses, the equations for the uniform-base analysis are expressed in a form that applies also to nonuniform-base transistors.

A cross section AA' is taken through the device of Figure 2-83b and carrier concentrations along this section are plotted in Figure 2-83c. Hole concentrations are denoted by p and electron concentrations by n with subscripts p or n representing p-type or n-type regions. The n-type emitter and collector regions are distinguished by subscripts E and C, respectively. The carrier concentrations shown in Figure 2-83c apply to a device in the *forward-active region*. That is, the base–emitter junction is forward-biased and the base–collector junction is reverse-biased. The minority-carrier concentrations in the base at the edges of the depletion regions can be calculated from a Boltzmann approximation to the Fermi–Dirac distribution function to give [2]

$$n_p(0) = n_{po} \exp \frac{V_{BE}}{V_T} \tag{2-138}$$

$$n_p(W_B) = n_{po} \exp \frac{V_{BC}}{V_T} \simeq 0 \tag{2-139}$$

where W_B is the width of the base from the base–emitter junction depletion layer edge of the base–collector depletion layer edge and n_{po} is the equilibrium concentration of electrons in the base. Note that V_{BC} is negative for an *npn* transistor in the forward-active region and thus $n_p(W_B)$ is very small. Low-level injection conditions are assumed in the derivation of Eqs. (2-138) and (2-139). This means that the minority-carrier concentrations are always assumed much smaller than the majority-carrier concentrations.

If *recombination* of holes and electrons in the base is small, it can be shown that the minority-carrier concentration $n_p(x)$ in the base varies *linearly* with distance [3]. Thus, a straight line can be drawn joining the concentrations at $x = 0$ and $x = W_B$ in Figure 2-83c.

For charge neutrality in the base, it is necessary that

$$N_A + n_p(x) = p_p(x) \tag{2-140}$$

and thus

$$p_p(x) - n_p(x) = N_A \tag{2-141}$$

where $p_p(x)$ is the hole concentration in the base and N_A is the base doping density that is assumed constant. Equation (2-141) indicates that the hole and electron concentrations are separated by a constant amount and thus $p_p(x)$ also varies linearly with distance.

Collector current is produced by minority-carrier electrons in the base diffusing in the direction of the concentration gradient and being swept across the collector–base depletion region by the field existing there. The diffusion current density due to electrons in the base is

$$J_n = qD_n \frac{dn_p(x)}{dx} \tag{2-142}$$

where D_n is the diffusion constant for electrons. From Figure 2-83c,

$$J_n = -qD_n \frac{n_p(0)}{W_B} \tag{2-143}$$

If I_C is the collector current and is taken as positive flowing *into* the collector, it follows from Eq. (2-143) that

$$I_C = qA D_n \frac{n_p(0)}{W_B} \tag{2-144}$$

where A is the cross-sectional area of the emitter. Substitution of Eq. (2-138) into Eq. (2-144) gives

$$I_C = \frac{qAD_n n_{po}}{W_B} \exp \frac{V_{BE}}{V_T} \tag{2-145}$$

$$= I_S \exp \frac{V_{BE}}{V_T} \tag{2-146}$$

where

$$I_S = \frac{qA D_n n_{po}}{W_B} \tag{2-147}$$

and I_S is a constant used to describe the transfer characteristic of the transistor in the forward-active region. Equation (2-147) can be expressed in terms of the base doping density by noting that [4]

$$n_{po} = \frac{n_i^2}{N_A} \tag{2-148}$$

and substitution of Eq. (2-148) in Eq. (2-147) gives

$$I_S = \frac{qAD_n n_i^2}{W_B N_A} = \frac{qA\overline{D}_n n_i^2}{Q_B} \tag{2-149}$$

where $Q_B = W_B N_A$ is the number of doping atoms in the base per unit area of the emitter and n_i is the intrinsic carrier concentration in silicon. In this form Eq. (2-149) applies to both uniform- and nonuniform-base transistors and D_n has been replaced by $\overline{D}_n$, which is an average effective value of the electron diffusion constant in the base. This is necessary for nonuniform-base devices because the diffusion constant is a function of impurity concentration. Typical values of I_S as given by Eq. (2-149) are 10^{-14} to 10^{-16} A.

Equation (2-146) gives the collector current as a function of base–emitter voltage. The base current I_B is also an important parameter and, at moderate current levels, consists of two major components. One of these (I_{B1}) represents recombination of holes and electrons in the base and is proportional to the minority-carrier charge Q_e in the base. From Figure 2-83c, the minority-carrier charge in the base is

$$Q_e = \frac{1}{2} n_p(0) W_B q A \tag{2-150}$$

and we have

$$I_{B1} = \frac{Q_e}{\tau_b} = \frac{1}{2} \frac{n_p(0) W_B q A}{\tau_b} \tag{2-151}$$

where τ_b is the minority-carrier lifetime in the base. I_{B1} represents a flow of majority holes from the base lead into the base region. Substitution of Eq. (2-138) into Eq. (2-151) gives

$$I_{B1} = \frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} \exp \frac{V_{BE}}{V_T} \tag{2-152}$$

The second major component of base current (usually the dominant one in the integrated-circuit *npn* devices) is due to injection of holes from the base into the emitter. This current component depends on the gradient of minority-carrier holes into the emitter and is [5]

$$I_{B2} = \frac{q A D_p}{L_p} p_{nE}(0) \tag{2-153}$$

where D_p is the diffusion constant for holes and L_p is the diffusion length (assumed small) for holes in the emitter. $p_{nE}(0)$ is the concentration of holes in the emitter at the edge of the depletion region and is

$$P_{nE}(0) = p_{nEo} \exp \frac{V_{BE}}{V_T} \tag{2-154}$$

If N_D is the donor atom concentration in the emitter (assumed constant) then

$$p_{nEo} \simeq \frac{n_i^2}{N_D} \tag{2-155}$$

The emitter is deliberately doped much more heavily than the base, making N_D large and p_{nEo} small, so that the base-current component, I_{B2}, is minimized.

Substitution of Eqs. (2-154) and (2-155) into Eq. (2-153) gives

$$I_{B2} = \frac{q A D_p}{L_p} \frac{n_i^2}{N_D} \exp \frac{V_{BE}}{V_T} \tag{2-156}$$

The total base current, I_B, is the sum of I_{B1} and I_{B2}:

$$I_B = I_{B1} + I_{B2} = \left(\frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} + \frac{q A D_p}{L_p} \frac{n_i^2}{N_D} \right) \exp \frac{V_{BE}}{V_T} \tag{2-157}$$

Although this equation was derived assuming uniform base and emitter doping, it gives the correct functional dependence of I_B on device parameters for practical double-diffused

nonuniform-base devices. Second-order components of I_B, which are important at low current levels, are considered later.

Since I_C in Eq. (2-146) and I_B into Eq. (2-147) are both proportional to $\exp(V_{BE}/V_T)$ in this analysis, the base current can be expressed in terms of collector current as

$$I_B = \frac{I_C}{\beta_F} \tag{2-158}$$

where β_F is the forward current gain. An expression for β_F can be calculated by substituting Eqs. (2-145) and (2-157) into Eq. (2-158) to give

$$\beta_F = \frac{qAD_n n_{po}/W_B}{\frac{1}{2}\frac{n_{po}W_B qA}{\tau_b} + \frac{qAD_p n_i^2}{L_p N_D}} = \frac{1}{\frac{W_B^2}{2\tau_b D_n} + \frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \tag{2-159}$$

where Eq. (2-148) has been substituted for n_{po}. Equation (2-159) shows that β_F is maximized by minimizing the base width W_B and maximizing the ratio of emitter to base doping densities N_D/N_A. Typical values of β_F for *npn* transistors in integrated circuits are 50–500, whereas lateral *pnp* transistors have values of 10–100. Finally, the emitter current is

$$I_E = -(I_C + I_B) = -\left(I_C + \frac{I_C}{\beta_F}\right) = -\frac{I_C}{\alpha_F} \tag{2-160}$$

where

$$\alpha_F = \frac{\beta_F}{1+\beta_F} \tag{2-161}$$

The value of α_F can be expressed in terms of device parameters by substituting Eq. (2-159) into Eq. (2-161) to obtain

$$\alpha_F = \frac{1}{1+\frac{1}{\beta_F}} = \frac{1}{1+\frac{W_B^2}{2\tau_b D_n} + \frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \simeq \alpha_T \gamma \tag{2-162}$$

where

$$\alpha_T = \frac{1}{1+\frac{W_B^2}{2\tau_b D_n}} \tag{2-163}$$

$$\gamma = \frac{1}{1+\frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \tag{2-164}$$

The validity of Eq. (2-162) depends on $W_B^2/2\tau_b D_n << 1$ and $(D_p/D_n)(W_B/L_p)(N_A/N_D) << 1$, and this is always true if β_F is large [see Eq. (2-159)]. The term γ in Eq. (2-162) is

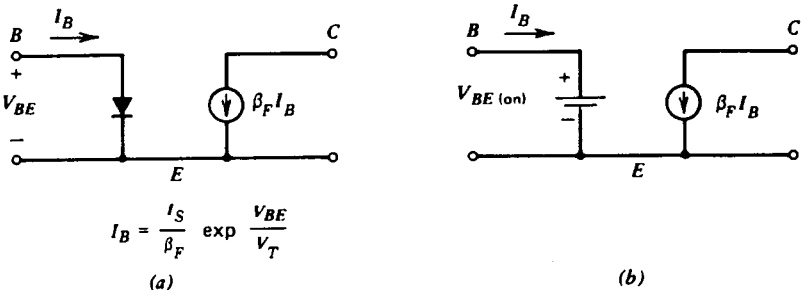


Figure 2-84 Large-signal models of *npn* transistors for use in bias calculations. (a) Circuit incorporating an input diode. (b) Simplified circuit with an input voltage source.

called the *emitter injection efficiency* and is equal to the ratio of the electron current (*npn* transistor) injected into the base from the emitter to the total hole and electron current crossing the base–emitter junction. Ideally $\gamma \to 1$, and this is achieved by making N_D/N_A large and W_B small. In that case very little reverse injection occurs from base to emitter.

The term α_T in Eq. (2-162) is called the *base transport factor* and represents the fraction of carriers injected into the base (from the emitter) that reach the collector. Ideally $\alpha_T \to 1$, and this is achieved by making W_B small. It is evident from the above development that fabrication changes that cause α_T and γ to approach unity also maximize the value of β_F of the transistor.

The results derived above allow formulation of a large-signal model of the transistor suitable for bias-circuit calculations with devices in the forward-active region. One such circuit is shown in Figure 2-84 and consists of a base–emitter diode to model Eq. (2-157) and a controlled collector-current generator to model Eq. (2-158). Note that the collector voltage ideally has no influence on the collector current and the collector node acts as a high-impedance current source. A simpler version of this equivalent circuit, which is often useful, is shown in Figure 2-84b, where the input diode has been replaced by a battery with a value $V_{BE(\text{on})}$, which is usually 0.6–0.7 V. This represents the fact that in the forward-active region the base–emitter voltage varies very little because of the steep slope of the exponential characteristic. In some circuits the temperature coefficient of $V_{BE(\text{on})}$ is important and a typical value for this is –2 mV/°C. The equivalent circuits of Figure 2-84 apply for *npn* transistors. For *pnp* devices the corresponding equivalent circuits are shown in Figure 2-85.

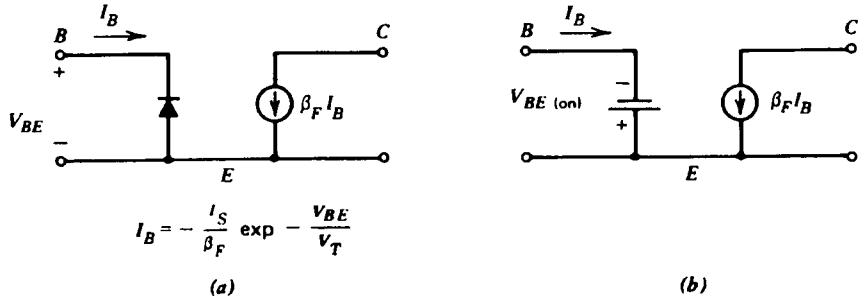


Figure 2-85 Large-signal models of *pnp* transistors corresponding to the circuits of Figure 2-84.

2-2-4 Effects of Collector Voltage on Large-Signal Characteristics in the Forward-Active Region

In the analysis of the previous section, the collector–base junction was assumed reverse-biased and ideally had no effect on the collector currents. This is a useful approximation for first-order calculations, but it is not strictly true in practice. There are occasions where the influence of collector voltage on collector current is important, and this will now be investigated.

The collector voltage has a dramatic effect on the collector current in two regions of device operation. These are the saturation (V_{CE} approaches zero) and breakdown (V_{CE} very large) regions that will be considered later. For values of collector–emitter voltage V_{CE} between these extremes, the collector current increases slowly as V_{CE} increases. The reason for this can be seen from Figure 2-86, which is a sketch of the minority-carrier concentration in the base of the transistor. Consider the effect of changes in V_{CE} on the carrier concentration for constant V_{BE}. Since V_{BE} is constant, the change in V_{CB} equals the change in V_{CE}, and this causes an increase in the collector–base depletion-layer width as shown. The change in the base width of the transistor, ΔW_B, equals the change in depletion-layer width and causes an increase ΔI_C in the collector current.

From Eqs. (2-146) and (2-149) we have

$$I_C = \frac{qA\bar{D}_n n_i^2}{Q_B} \exp \frac{V_{BE}}{V_T} \tag{2-165}$$

Differentiation of Eq. (2-165) yields

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{qA\bar{D}_n n_i^2}{Q_B^2} \left(\exp \frac{V_{BE}}{V_T} \right) \frac{dQ_B}{dV_{CE}} \tag{2-166}$$

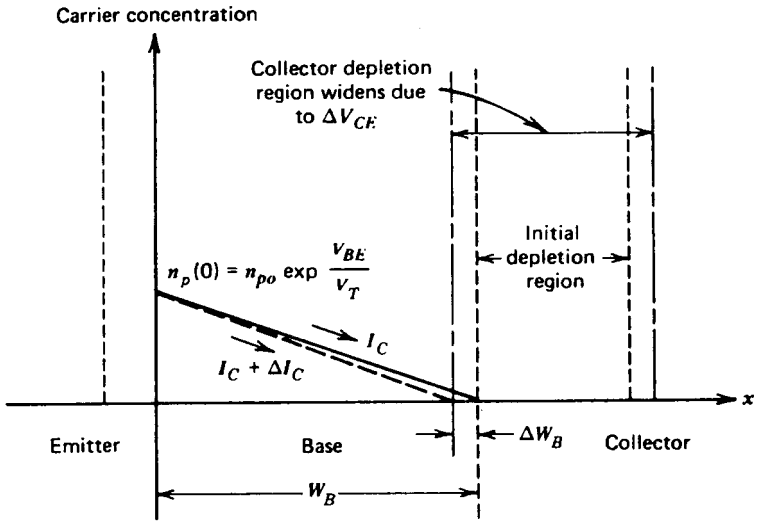


Figure 2-86 Effect of increases in V_{CE} on the collector depletion region and base width of a bipolar transistor.

and substitution of Eq. (2-165) in Eq. (2-166) gives

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{I_C}{Q_B}\frac{dQ_B}{dV_{CE}} \tag{2-167}$$

For a uniform-base transistor $Q_B = W_B N_A$ and Eq. (2-167) becomes

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{I_C}{W_B}\frac{dW_B}{dV_{CE}} \tag{2-168}$$

Note that since the base width *decreases* as V_{CE} increases, dW_B/dV_{CE} in Eq. (2-168) is negative and thus $\partial I_C/\partial V_{CE}$ is positive. dW_B/dV_{CE} is a function of the bias value of V_{CE}, but the variation is typically small for a reverse-biased junction and dW_B/dV_{CE} is often assumed constant. The resulting predictions agree adequately with experimental results.

Equation (2-168) shows that $\partial I_C/\partial V_{CE}$ is proportional to the collector-bias current and inversely proportional to the transistor base width. Thus narrow-base transistors show a greater dependence of I_C on V_{CE} in the forward-active region. The dependence of $\partial I_C/\partial V_{CE}$ on I_C results in typical transistor output characteristics as shown in Figure 2-87. In accordance with the assumptions made in the foregoing analysis, these characteristics are shown for constant values of V_{BE}. However, in most integrated-circuit transistors the base current is dependent only on V_{BE} and not on V_{CE}, and thus constant-base-current characteristics can often be used in the following calculation. The reason for this is that the base current is usually dominated by the I_{B2} component of Eq. (2-156), which has no dependence on V_{CE}. Extrapolation of the characteristics of Figure 2-87 back to the V_{CE} axis gives an intercept V_A called the Early voltage, where

$$V_A = \frac{I_C}{\dfrac{\partial I_C}{\partial V_{CE}}} \tag{2-169}$$

Substitution of Eq. (2-168) in Eq. (2-169) gives

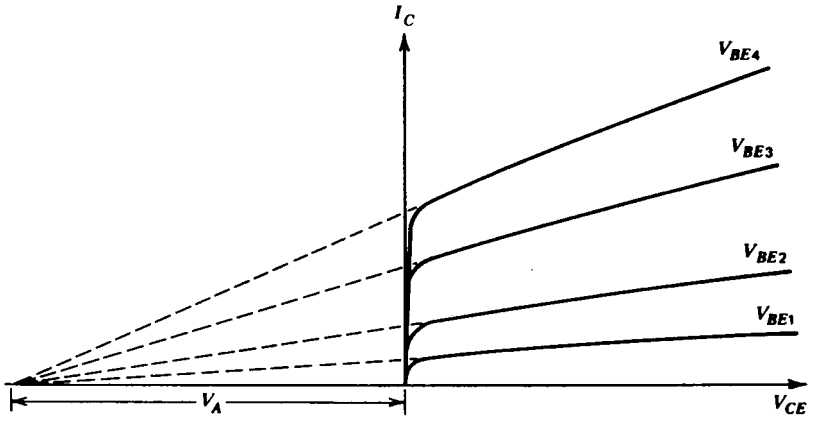


Figure 2-87 Bipolar transistor output characteristics showing the Early voltage, V_A.

$$V_A = -W_B \frac{dV_{CE}}{dW_B} \tag{2-170}$$

which is a constant, independent of I_C. Thus all the characteristics extrapolate to the same point on the V_{CE} axis. The variation of I_C with V_{CE} is called the Early effect and V_A is a common model parameter for circuit-analysis computer programs. Typical values of V_A for integrated-circuit transistors are 15–100 V. The inclusion of Early effect in dc bias calculations is usually limited to computer analysis because of the complexity introduced into the calculation. However, the influence of the Early effect is often dominant in small-signal calculations for high-gain circuits and this point will be considered later.

Finally, the influence of the Early effect on the transistor large-signal characteristics in the forward-active region can be represented approximately by modifying Eq. (2-146) to

$$I_C = I_S \left(1 + \frac{V_{BE}}{V_A}\right) \exp \frac{V_{BE}}{V_T} \tag{2-171}$$

This is a common means of representing the device output characteristics for computer simulation.

2-2-5 Saturation and Inverse Active Regions

Saturation is a region of device operation that is usually avoided in analog circuits because the transistor gain is very low in this region. Saturation is much more commonly encountered in digital circuits, where it provides a well-specified output voltage that represents a logic state. In saturation, both emitter–base and collector–base junctions are forward biased. Consequently, the collector–emitter voltage V_{CE} is quite small and usually in the range 0.05–0.3 V. The carrier concentrations in a saturated *npn* transistor with uniform base doping are shown in Figure 2-88. The minority-carrier concentration in the base of the edge of the depletion region is again given by Eq. (2-139) as

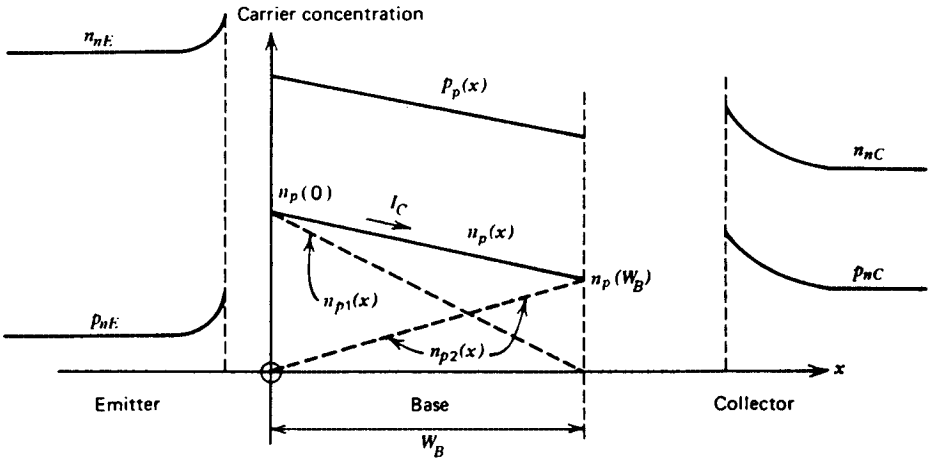


Figure 2-88 Carrier concentrations in a saturated *npn* transistor. (Not to scale)

$$n_p(W_B) = n_{po} \exp \frac{V_{BC}}{V_T} \tag{2-172}$$

but since V_{BC} is now positive, the value of $n_p(W_B)$ is no longer negligible. Consequently, changes in V_{CE} with V_{BE} held constant (which cause equal changes in V_{BC}) directly affect $n_p(W_B)$. Since the collector current is proportional to the slope of the minority-carrier concentration in the base [see Eq. (2-142)], it is also proportional to $[n_p(0) - (n_p(W_B)]$ from Figure 2-88. Thus, changes in $n_p(W_B)$ directly affect the collector current, and the collector node of the transistor appears to have a *low impedance*. As V_{CE} is decreased in saturation with V_{BE} held constant, V_{BC} increases as does $n_p(W_B)$ from Eq. (2-172). Thus, from Figure 2-88 the collector current decreases because the slope of the carrier concentration decreases. This gives rise to the saturation region of the $I_C - V_{CE}$ characteristic shown in Figure 2-89. The slope of the $I_C - V_{CE}$ characteristic in this region is largely determined by the resistance in series with the collector lead due to the finite resistivity of the *n*-type collector material. A useful model for the transistor in this region is shown in Figure 2-90

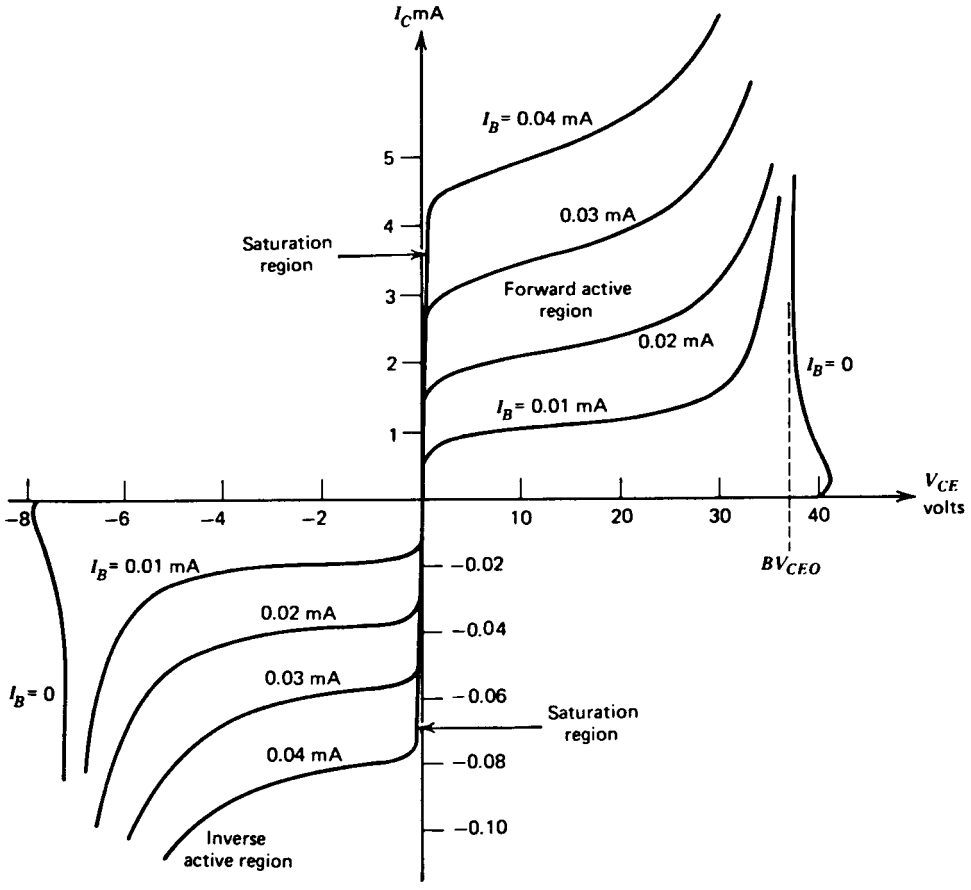


Figure 2-89 Typical $I_C - V_{CE}$ characteristics for an *npn* bipolar transistor. Note the different scales for positive and negative currents and voltages.

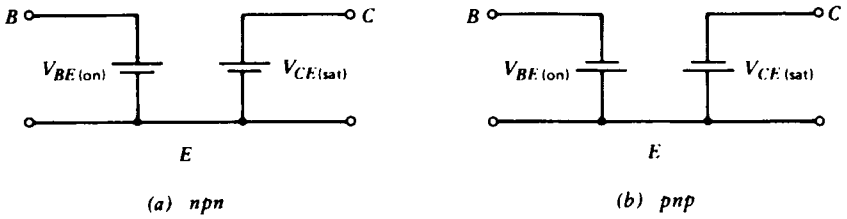


Figure 2-90 Large-signal models for bipolar transistors in the saturation region.

and consists of a fixed voltage source to represent $V_{BE(on)}$, and a fixed voltage source to represent the collector–emitter voltage $V_{CE(sat)}$. A more accurate but more complex model includes a resistor in series with the collector. This resistor can have a value ranging from 20 to 500 Ω, depending on the device structure.

An additional aspect of transistor behavior in the saturation region is apparent from Figure 2-88. For a given collector current, there is now a much larger amount of stored charge in the base than there is in the forward-active region. Thus the base-current contribution represented by Eq. (2-152) will be larger in saturation. In addition, since the collector–base junction is now forward biased, there is a new base-current component due to injection of carriers from the base to the collector. These two effects result in a base current I_B in saturation, which is larger than in the forward-active region for a given collector current I_C. Ratio I_C/I_B in saturation is often referred to as the *forced* β and is always less than β_F. As the forced β is made lower with respect to β_F, the device is said to be more *heavily saturated*.

The minority-carrier concentration in saturation, shown in Figure 2-88, is a straight line joining the two end points, assuming that recombination is small. This can be represented by a linear superposition of the two dotted distributions as shown. The justification for this is that the terminal currents depend *linearly* on the concentrations $n_p(0)$ and $n_p(W_B)$. This picture of device-carrier concentrations can be used to derive some general equations describing transistor behavior. Each of the distributions in Figure 2-88 is considered separately and the two contributions are combined. The *emitter* current that would result from $n_{p1}(x)$ above is given by the classical diode equation

$$I_{EF} = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) \tag{2-173}$$

where I_{ES} is a constant that is often referred to as the "saturation current" of the junction (no connection with the transistor saturation described above). Equation (2-173) predicts that the junction current is given by $I_{EF} \simeq I_{ES}$ with a reverse-bias voltage applied. However, in practice Eq. (2-173) is applicable only in the forward-bias region, since second-order effects dominate under reverse-bias conditions and typically result in a junction current several orders of magnitude larger than I_{ES}. The junction current that flows under reverse-bias conditions is often called the "leakage current" of the junction.

Returning to Figure 2-88, we can describe the *collector* current resulting from $n_{p2}(x)$ alone as

$$I_{CR} = -I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-174}$$

where I_{CS} is a constant. The total collector current I_C is given by I_{CR} plus the fraction of I_{EF} that reaches the collector (allowing for recombination and reverse emitter injection). Thus,

$$I_C = \alpha_F I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) - I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-175}$$

where α_F has been defined previously by Eq. (2-162). Similarly, the total emitter current is composed of I_{EF} plus the fraction of I_{CR} that reaches the emitter with the transistor acting in an inverted mode. Thus,

$$I_E = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + \alpha_R I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-176}$$

where α_R is the ratio of emitter to collector current with the transistor operating *inverted* (i.e., with the collector–base junction forward biased and emitting carriers into the base and the emitter–base junction reverse biased and collecting carriers). Typical values of α_R are 0.5–0.8. An inverse current gain β_R is also defined,

$$\beta_R = \frac{\alpha_R}{1 - \alpha_R} \tag{2-177}$$

and has typical values 1–5. This is the current gain of the transistor when operated inverted and is much lower than β_F because the device geometry and doping densities are designed to maximize β_F. The inverse-active region of device operation occurs for V_{CE} negative in an *npn* transistor and is shown in Figure 2-89. In order to display these characteristics adequately in the same figure as the forward-active region, the negative voltage and current scales have been expanded. The inverse-active mode of operation is rarely encountered in analog circuits.

Equations (2-175) and (2-176) describe *npn* transistor operation in the saturation region when V_{BE} and V_{BC} are both positive, and also in the forward-active and inverse-active regions. These equations are from the *Ebers–Moll* equations. In the forward-active region, they degenerate into a form similar to that of Eqs. (2-146), (2-158), and (2-160) derived earlier. This can be shown by putting V_{BE} positive and V_{BC} negative in Eqs. (2-175) and (2-176) to obtain

$$I_C = \alpha_F I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + I_{CS} \tag{2-178}$$

$$I_E = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) - \alpha_R I_{CS} \tag{2-179}$$

Equation (2-178) is similar in form to Eq. (2-146) except that leakage currents that were previously neglected have now been included. This minor difference is significant only at high temperatures or very low operating currents. Comparison of Eq. (2-178) with Eq. (2-146) allows us to identify $I_S = \alpha_F I_{ES}$ and it can be shown [6] in general that

$$\alpha_F I_{ES} = \alpha_R I_{CS} = I_S \tag{2-180}$$

where this expression represents a reciprocity condition. Using Eq. (2-180) in Eqs. (2-175) and (2-176) allows the Ebers–Moll equations to be expressed in the general form

$$I_C = I_S\left(\exp\frac{V_{BE}}{V_T} - 1\right) - \frac{I_S}{\alpha_R}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-181}$$

$$I_E = -\frac{I_S}{\alpha_F}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + I_S\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-182}$$

This form is often used for computer representation of transistor large-signal behavior.

The effect of leakage currents mentioned above can be further illustrated as follows. In the forward-active region, we have, from Eq. (2-179),

$$I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) = -I_E - \alpha_R I_{CS} \tag{2-183}$$

Substitution of Eq. (2-183) in Eq. (2-178) gives

$$I_C = -\alpha_F I_E + I_{CO} \tag{2-184}$$

where

$$I_{CO} = I_{CS}(1 - \alpha_R\alpha_F) \tag{2-185}$$

and I_{CO} is the collector–base leakage current with the emitter open. Although I_{CO} is given theoretically by Eq. (2-185), in practice, surface leakage effects dominate when the collector–base junction is reverse biased and I_{CO} is typically several orders of magnitude larger than the value given by Eq. (2-185). However, Eq. (2-184) is still valid if the appropriate measured value for I_{CO} is used. Typical values of I_{CO} are 10^{-10} to 10^{-12} A at 25 °C, and the magnitude doubles about every 8 °C. As a consequence, these leakage terms can become very significant at high temperatures. For example, consider the base current I_B. This is

$$I_B = -(I_C + I_E) \tag{2-186}$$

If I_E is calculated from Eq. (2-184) and substituted in Eq. (2-186), the result is

$$I_B = \frac{1 - \alpha_F}{\alpha_F} I_C - \frac{I_{CO}}{\alpha_F} \tag{2-187}$$

But from Eq. (2-161)

$$\beta_f = \frac{\alpha_F}{1 - \alpha_F} \tag{2-188}$$

and use of Eq. (2-188) in Eq. (2-187) gives

$$I_B = \frac{I_C}{\beta_F} - \frac{I_{CO}}{\alpha_F} \tag{2-189}$$

Since the two terms in Eq. (2-189) have opposite signs, the effect of I_{CO} is to *decrease* the magnitude of the external base current at a given value of collector current.

2-2-6 Small-Signal Models of Bipolar Transistors

Figure 2-91 shows the small-signal equivalent BJT model. It consists of an intrinsic model and a package model. Table 2-12 lists its keywords.

In order to better understand the difference between the various technologies, we show Table 2-13, which allows the reader to compare their performances.

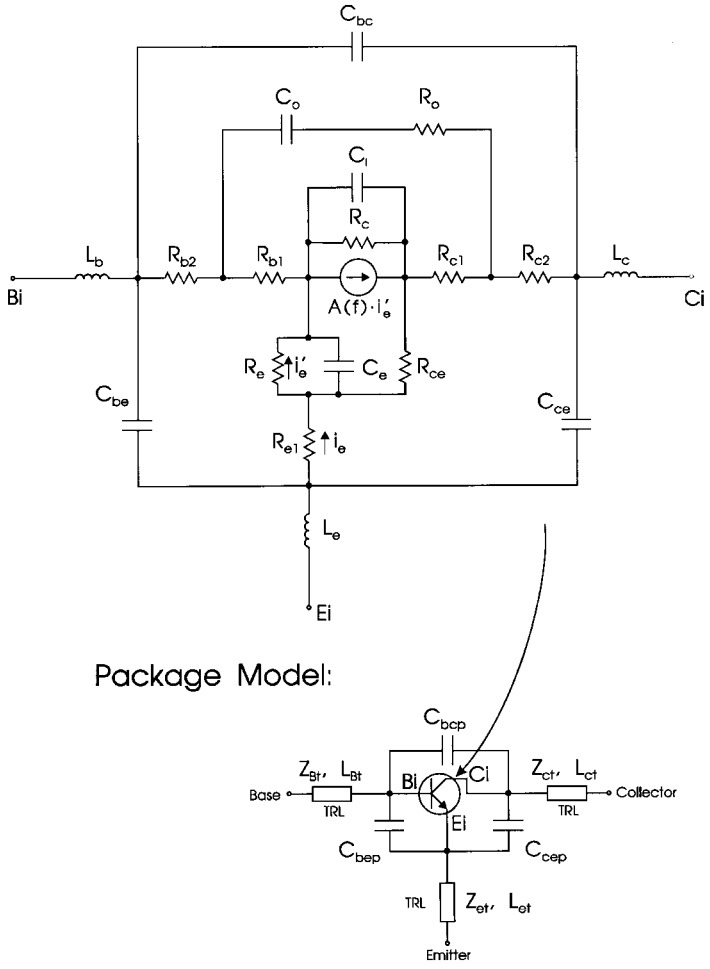


Figure 2-91 Linear BJT model.

Table 2-12 Small-signal equivalent BJT model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
A	Ratio of I_C to I_E at dc		0.0
RE	Emitter resistance	ohm	0.0
F	Current generator roll-off frequency	hertz	∞
T	Time delay	second	0.0
CE	Emitter capacitance	farad	0.0
CI	Collector capacitance	farad	0.0
RCE	Collector–emitter resistance	ohm	∞
RC	Collector resistance	ohm	∞
RO	Extrinsic base–collector resistance	ohm	0.0
CO	Extrinsic base–collector capacitance	farad	0.0
RB1	Intrinsic base resistance (R_{bb})	ohm	0.0
RC1	Parasitic collector resistance	ohm	0.0
RE1	Parasitic emitter resistance	ohm	0.0
RB2	Parasitic base resistance	ohm	0.0
RC2	Parasitic collector resistance	ohm	0.0
CBE	Base-to-emitter package capacitance	farad	0.0
CBC	Base-to-collector package capacitance	farad	0.0
CCE	Collector-to-emitter package capacitance	farad	0.0
LB	Base-lead inductance	henry	0.0
LC	Collector-lead inductance	henry	0.0
LE	Emitter-lead inductance	henry	0.0
TJ	Chip temperature	Kelvin	298
NFAC	Noise factor proportional to drive		1.0
FC	Flicker noise (1/*f* noise) corner frequency	hertz	50
	Package Model		
CBCP	Base-to-collector package capacitance	farad	0.0
CBEP	Base-to-emitter package capacitance	farad	0.0
CCEP	Collector-to-emitter package capacitance	farad	0.0
ZBT	Base transmission line impedance	ohm	50
ZCT	Collector transmission line impedance	ohm	50
ZET	Emitter transmission line impedance	ohm	50
LBT	Base transmission line length at $\varepsilon_r = 1$	meter	0.0
LCT	Collector transmission line length at $\varepsilon_r = 1$	meter	0.0
LET	Emitter transmission line length at $\varepsilon_r = 1$	meter	0.0

Notes:

1. $A \equiv \alpha = I_c/I_e$; β = dc current gain = $\alpha(1 - \alpha)$.
2. The bipolar current gain in this model is described by

$$\mathrm{A} = \mathrm{A}(0) = \frac{e^{-j\omega\mathrm{T}}}{1 + j\,(f/F)}$$

where $\omega = 2\,\pi f$ and f is frequency.

3. The current source is controlled by the current through R_e. The current generator has a cutoff frequency with respect to the total emitter current, I_E:

$$\mathrm{F} = g_m/2\pi C_e$$

where $g_m = 1/R_e$. This frequency becomes infinite for the default value for C_e (0.0). The parammeter F specifies the frequency roll-off for the current generator with respect to the current through R_e. Effectively, this frequency parameter may be used to model additional delays in the device.

Table 2-13 Overview of current BJT technologies including BiCMOS

Parameter	MOSAIC 3	MOSAIC 5	MOSAIC 5SE	RFBICMOS1	TFSOI
Emitter Size	$0.95 \times 3.2\ \mu m^2$	$0.3 \times 0.9\ \mu m^2$	$0.3 \times 0.9\ \mu m^2$	$0.4 \times 1.3\ \mu m^2$	$0.1 \times 1.2\ \mu m^2$
β	150	150	90	80	40–65
F_t (GHz)	14@ $V_{ce} = 1$ V	14@ $V_{ce} = 1$ V	27@ $V_{ce} = 1$ V	16@ $V_{ce} = 1$ V	8@ $V_{ce} = 2$ V
C_{je} (fF)	19.9	4.3	3.8	5.1	1.5
C_{jc} (fF)	25.1	2.8	2.8	3.7	0.34
C_{js} (fF)	49.6	4.3	3.6	7.4	0.95
R_e (Ω)	22	36	50	34	72
R_b (Ω)	462	578	583	462	n/a
BV_{ceo} (V)	6.5	7.0	4.8	6.0	>3.5
BV_{cbo} (V)	20	20	15	18	>10
BV_{ebo} (V)	5.2	4.2	3.8	6.0	4.3
V_A (V)	30	48	20	>25	25–40
Minimum T_{pd}	75 ps @ 1.5 mA	35 ps @ 800 μA	36 ps @ 800 μA	43 ps @ 300 μA	52 ps @ 25 μA

Parameter	TFSOI Ring Device	RFBICMOS3 Analog	RFBICMOS3 Digital	BiGCMOS3 Analog	BiGCMOS3 Digital
Emitter Size	$0.1 \times 7.2\ \mu m^2$	$0.25 \times 0.75\ \mu m^2$	$0.25 \times 0.75\ \mu m^2$	$0.2 \times 0.6\ \mu m^2$	$0.2 \times 0.6\ \mu m^2$
β	40–65	100	150	100	100
F_t (GHz)	8 @ $V_{ce} = 2$ V	15 @ $V_{ce} = 2$ V	25 @ $V_{ce} = 2$ V	15 @ $V_{ce} = 1$ V	35 @ $V_{ce} = 1$ V
C_{je} (fF)	3.4	1.8	1.8	1.7	1.7
C_{jc} (fF)	2.8	1.1	1.1	0.2	0.3
C_{js} (fF)	2.4	12.0	12.0	4	4
R_e (Ω)	26	140	140	113	113
R_b (Ω)	790	n/a	n/a	10 kΩ/sq.	15 kΩ/sq.
BV_{ceo} (V)	>3.5	6.5	4	6.5	4.8
BV_{cbo} (V)	>10	15	15	20	15
BV_{ebo} (V)	4.3	4	4	4	4
V_A (V)	30–40	40	25	40	20
Minimum T_{pd}	75 ps @ 100 μA	n/a	n/a	35 ps @ 50 μA	22 ps @ 200 μA

Parameter	AT&T BEST2[a]	Fujitsu POSET	HP HP-25	HP ISOSAT II[b]	Hitachi SMI
Emitter Size	$1.0 \times 2.0\ \mu m^2$	$0.1 \times 1.7\ \mu m^2$	$0.4 \times 20\ \mu m^2$	$0.6 \times ?\ \mu m^2$	$0.2 \times 2.0\ \mu m^2$
β	100	110	75	100	500
F_t (GHz)	23 @ $V_{ce} = 3$ V	38 @ $V_{ce} = 3$ V	28.5 @ $V_{ce} = 3$ V	14/F_{max} > 20 GHz	82@ $V_{ce} = 2$ V
C_{je} (fF)	5.0	3.5	42	n/a	2.9
C_{jc} (fF)	5.0	2.2	37	n/a	4.8
C_{js} (fF)	8.0	n/a	123	n/a	4.1
R_e (Ω)	25	25	n/a	n/a	30
R_b (Ω)	10 kΩ/sq.	960	100	n/a	240
BV_{ceo} (V)	5.5	5.4	5.2	15	2.5
BV_{cbo} (V)	11	18.2	15	25	6.4
BV_{ebo} (V)	8.0	4.5	>5	2.0	6.2
V_A (V)	24	n/a	n/a	25	8

(continued)

Table 2-13 (Continued)

Parameter	AT&T BEST2[a]	Fujitsu POSET	HP HP-25	HP ISOSAT II[b]	Hitachi SMI
Minimum T_{pd}	37 ps @ 2.1 mA 48 ps @ 600 µA	21.5 ps @ 320 µA	n/a	n/a	16 ps @ $I_g = 1.3$

Parameter	IBM SiGe[c]	Maxim GST-1[d]	Maxim GST-2[e]	Mitsubishi[f]	National ABIC V
Emitter Size	0.5×1.0 µm^2	1.2×2.6 µm^2	0.7×1.5 µm^2	0.5×5.7 µm^2	0.5×0.5 µm^2
β	100	75	190	90	85
F_t (GHz)	41 @ $V_{ce} = 1$ V	12.5	26 @ $V_{ce} = 4$ V	20	13 @ $V_{ce} = 2$ V
C_{je} (fF)	5.2	9	4	n/a	1.9
C_{jc} (fF)	3.4	15	7	9	3.7
C_{js} (fF)	3.6	11	9	n/a	9.6
R_e (Ω)	n/a	n/a	80	n/a	n/a
R_b (Ω)	300	600	710	155	383
BV_{ceo} (V)	3.3	>5	>4	6.2	8.5
BV_{cbo} (V)	9.5	>12	>14	n/a	20
BV_{ebo} (V)	n/a	>4	>4	n/a	7.5
V_A (V)	n/a	20	15	n/a	n/a
Minimum T_{pd}	n/a	n/a	n/a	n/a	38 ps @ 1.8 mA

Parameter	NEC BSA	NEC SiGe[g]	NEC NESAT III	Philips QBiC2	Philips PRET[g]
Emitter	0.4×7.6 µm^2	0.2×1.6 µm^2	0.6×40 µm^2	0.5×1.2 µm^2	0.2×0.9 µm^2
β	n/a	120	80	100	50
F_t (GHz)	43	51	20 @ $V_{ce} = 3$ V	17	14
C_{je} (fF)	20	n/a	120	2.4	1.5
C_{jc} (fF)	16	1.7	140	3.5	1.0
C_{js} (fF)	n/a	n/a	160	6.5	5.0
R_e (Ω)	n/a	80	n/a	120	120
R_b (Ω)	43	220	55	700 Ω/sq.	1 kΩ/sq.
BV_{ceo} (V)	3.3	2.7	6	4.5	7
BV_{cbo} (V)	n/a	n/a	17	13	12.5
BV_{ebo} (V)	2	2.2	4.5	5.5	5.5
V_A (V)	n/a	n/a	n/a	15	20
Minimum T_{pd}	29 ps @ 300 µA	19 ps @ 300 µA	n/a	n/a	n/a

Parameter	Philips SiGe[g]	Siemens B6HF	Siemens B6HF Plus	Siemens SiGe[g]	SONY[h]
Emitter Size	0.4×1.4 µm^2	0.8×3.3 (drawn) µm^2	0.25×2.7 µm^2	0.27×2.5 µm^2	$0.2 \times ?$ µm^2
β	100	>60	115	130	100
F_t (GHz)	45 @ $V_{ce} = 1.5$ V	25 @ $V_{cb} = 0$ V	50 @ $V_{cb} = 2$ V	61 @ $V_{cb} = 1$ V	30
C_{je} (fF)	9.1	11	8.5	7.5	n/a
C_{jc} (fF)	4.6	6.5	8.0	5.5	n/a
C_{js} (fF)	10.8	31	26.5	13	n/a

(continued)

Table 2-13 Overview of current BJT technologies including BiCMOS (Continued)

Parameter	Phillips SiGe[g]	Siemens B6HF	Siemens B6HF Plus	Siemens SiGe[g]	SONY[h]
R_e (Ω)	n/a	n/a	n/a	n/a	n/a
R_b (Ω)	n/a	165	14 kΩ/sq.	8 kΩ/sq.	n/a
BV_{ceo} (V)	3.2	>3.5	7	2.5	4.0
BV_{cbo} (V)	11	n/a	10.8	9.9	18
BV_{ebo} (V)	3.5	n/a	2.0	3.0	3.5
V_A (V)	70–100	n/a	20	130	10
Minimum T_{pd}	14 ps @ 6 mA	25 ps @ 1.2 mA (CML)	16 ps @ 2.0 mA (CML)	11 ps @ 1.9 mA (CML)	21 ps @ I_g = 3.5 mA (CML)

Parameter	STM HSB[i]	STM HSB2A	iSTM HSB3[g,j]	Toshiba	Toshiba Epi Base[g]
Emitter Size	$0.35 \times ? \ \mu m^2$	$0.35 \times ? \ \mu m^2$	$0.2 \times ? \ \mu m^2$	$0.4 \times 5.0 \ \mu m^2$	$0.6 \times 5.0 \ \mu m^2$
β	>50	>50	>50	60	105
F_t (GHz)	n/a	20	~50	22 @ V_{ce} = 2 V	52 @ V_{ce} = 3 V
C_{je} (fF)	n/a	n/a	n/a	9.0	16.1
C_{jc} (fF)	n/a	n/a	n/a	3.6	14.1
C_{js} (fF)	n/a	n/a	n/a	11	n/a
R_e (Ω)	n/a	n/a	n/a	14	22.7
R_b (Ω)	10 kΩ/sq.	n/a	n/a	90	205
BV_{ceo} (V)	4.5	n/a	n/a	6.2	4.1
BV_{cbo} (V)	n/a	n/a	n/a	14.8	9.5
BV_{ebo} (V)	n/a	n/a	n/a	4.2	n/a
V_A (V)	n/a	>30	n/a	n/a	26.5
Minimum T_{pd}	n/a	30 ps	n/a	19 ps @ 1.2 mA	n/a

[a] 1.0 µm Design Rule; NF = 3.25 dB, G_A = 12 dB at I_C = 2 mA, freq. = 2 GHz for $16 \times 1 \times 2 \ \mu m^2$; Resistors: 150 Ω/sq. (mono), 300 Ω/sq. (mono), 500 Ω/sq. (poly); Layers of metal, Schottky diodes, linear capacitor (silicided n^+ poly/n^+Si); Gate length = 0.68 µm (*n*-ch)/0.8 µm (*p*-ch), Gate Oxide = 12.5 nm (3 V)/15 nm (5 V), V_{th} = 0.65 V (*n*-ch)/–0.9 V (*p*-ch), Gm (µS/µm) = 204 (*n*-ch)/65 (*p*-ch).

[b] 0.6 µm Design Rule; Gold metallization, ILD: polyimide; Inductors: $Q > 10$ for $L \leq 10$ nH; Capacitors: $Q > 25$ for $C \leq 20$ pF; Schottky diodes (f_c = 100 GHz), Zener diodes, MIS capacitor, thin film resistors, and LPNP, 10 mask process.

[c] Analog devices using this process.

[d] LPNP F_t = 80 MHz; three-layer metallization; Schottky diodes (R_sC_j ~ 4 ps); Implanted resistors: 200 Ω/sq.; Poly resistors: 400/2000 Ω/sq.; NiCr thin film resistors (50 Ω/sq.) with laser trim option; MOS capacitor.

[e] LPNP F_t = 85 MHz; two-layer metallization; Schottky diodes (R_sC_j ~ 0.8 ps); Implanted resistors: 170/600/1200 Ω/sq.; NiCr thin film resistors (50 Ω/sq.) with laser trim option; MOS capacitor.

[f] L_n = 0.8 µm, T_{ox} = 18 nm, V_{th} = 0.76 V (nmos), V_{th} = –0.78 V (pmos), 2poly-MIM capacitor: 2.7 fF/µm^2, Schottkys, inductors.

[g] Not known if this process is in production.

[h] 0.8 µm Design Rules; 2.5 fF/µm^2 MIS capacitor; 600 MHz PJFET; 12L; F_t = 4 GHz VPNP; poly resistors 80/400/2000 Ω/sq.

[i] Targeted for 900-MHz GSM radio front-end blocks, a CT2 cordless single-chip radio, and DECT (digital enhanced cordless telephones).

[j] Targeted for multimode 900-MHz GSM, 1.8-GHz PCN and 1.9-GHz PCS based systems. Company is also evaluating integrating SiGe in HSB3.

2-3 FIELD-EFFECT TRANSISTORS

Both diodes and *pnp/npn* transistors work on the injection principle, which means the base–emitter junction generates either free electrons or other carriers, which are being "collected" by the collector connection and the base connection controls it. While this is a fairly complex process, its physics is well understood, and its latest spin-offs, heterojunction BJTs and silicon–germanium transistors, seem to offer a bright future.

In opposition to this, the field-effect transistor (FET) controls the electronic conduction in a solid by an electric field, and this concept actually predates the invention of the bipolar transistor. J. E. Lilienfeld filed for a patent on such a device in 1925 (U.S. Patent No. 1,745,175), and W. Shockley presented a comprehensive theory of the FET in 1952. It took until 1960 before the first commercially available FETs came on the market.

Several types of semiconductor materials have been used for making FETs: silicon, germanium, gallium arsenide, and others have been used. By far the most widely used is silicon, followed now by gallium arsenide.

The FET is a class of electronic semiconductor device in which the conduction of a "channel" between source (*S*) and drain (*D*) terminals is controlled by an electric field impressed on the channel via a gate (*G*) terminal. The conduction channel may utilize *n*-type carriers (electrons) or *p*-type carriers (holes). The electric field that controls the channel conduction may be introduced via a *pn* junction [for a "junction" FET (JFET)], a metal plate separated from the semiconductor channel by an oxide dielectric [for a metal-oxide semiconductor (MOS) FET], or a combination of the two methods. The polarity of the controlling electrical field is a function of the type of carriers in the channel. A FET "family tree" is shown in Figure 2-92.

For today's microwave and RF integrated circuits, the JFET with its high-frequency problems (low cutoff frequency) and large variation in parameters, such as transconductance and pinchoff voltage, has fallen out of favor and is mainly used in discrete circuits such as two-way radios or other RF applications that tolerate these wide variations and temperature dependence. The most-used FET types are MOS for small-signal applications, LDMOS (lateral diffused MOS) for power applications, and GaAs MESFET for lower-power applications (low-voltage applications) as well as high-power applications (recent advertisements by NEC have shown output power up to 100 W). Additional information on the FET family is shown in Figure 2-93.

All members of the FET family are high-impedance-input devices. Figure 2-94 shows the key operating parameters for MOSFETs and MESFETs in low-noise stages, high-gain stages, and linearized power stages. The linearization also improves the matching. This figure is similar to its counterpart in Section 2-2-2.

The following several pages present a typical example of a datasheet* for a JFET (U310) and the chip (NZA) on which it is based, followed by an example of a device with a much higher dynamic range (the CP640 family). The CP640 family and devices similar to them are being replaced by low-power LDMOS, which are not yet well-documented. At least one company, InterFET, is devoted entirely to producing discrete JFETs and JFET-related hybrid and small-scale-integration IC products.

* Reproduced with permission of Vishay-Siliconix.

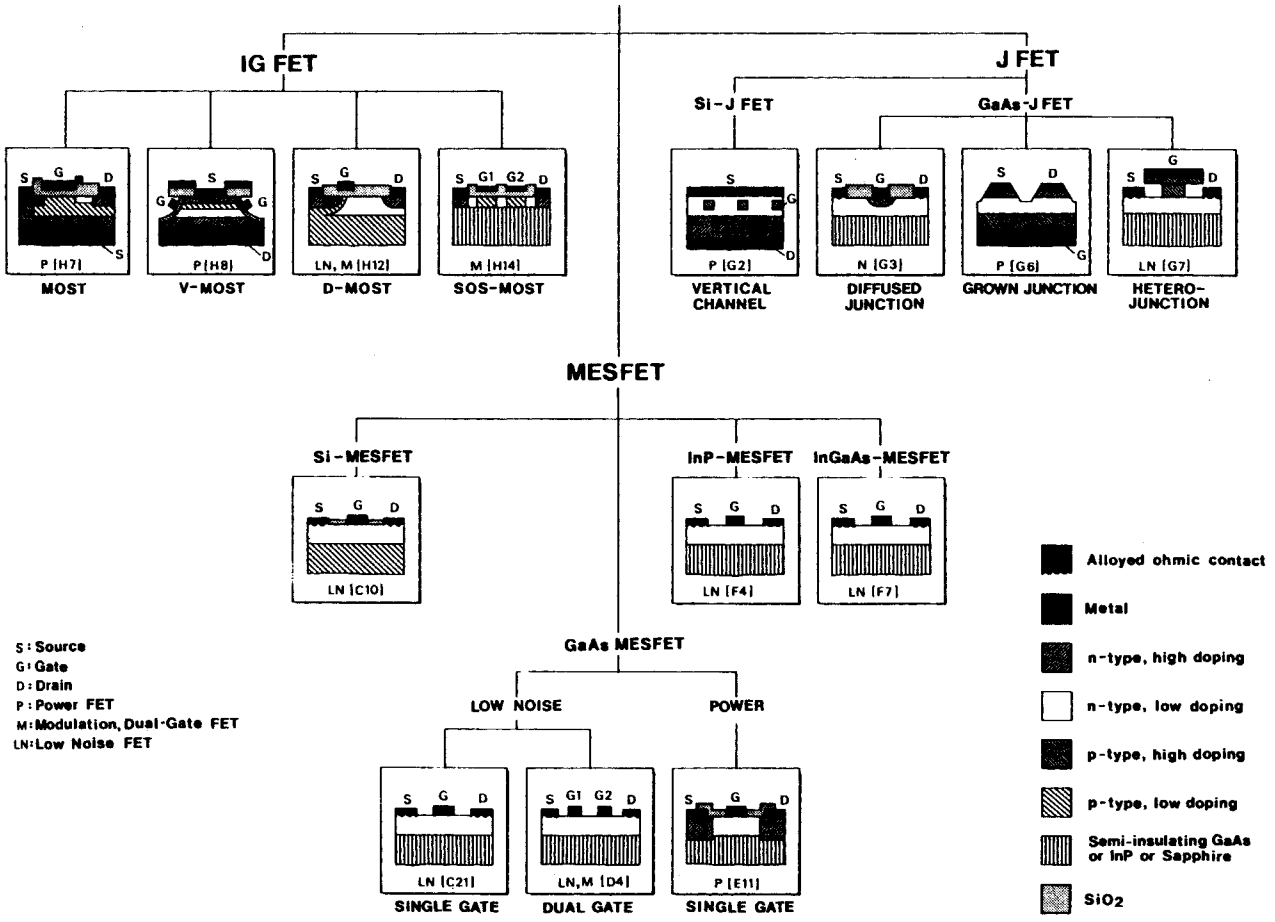


Figure 2-92 "Family tree" of microwave FET types showing cross sections of their structures [7]. Copyright ©1976 IEEE.

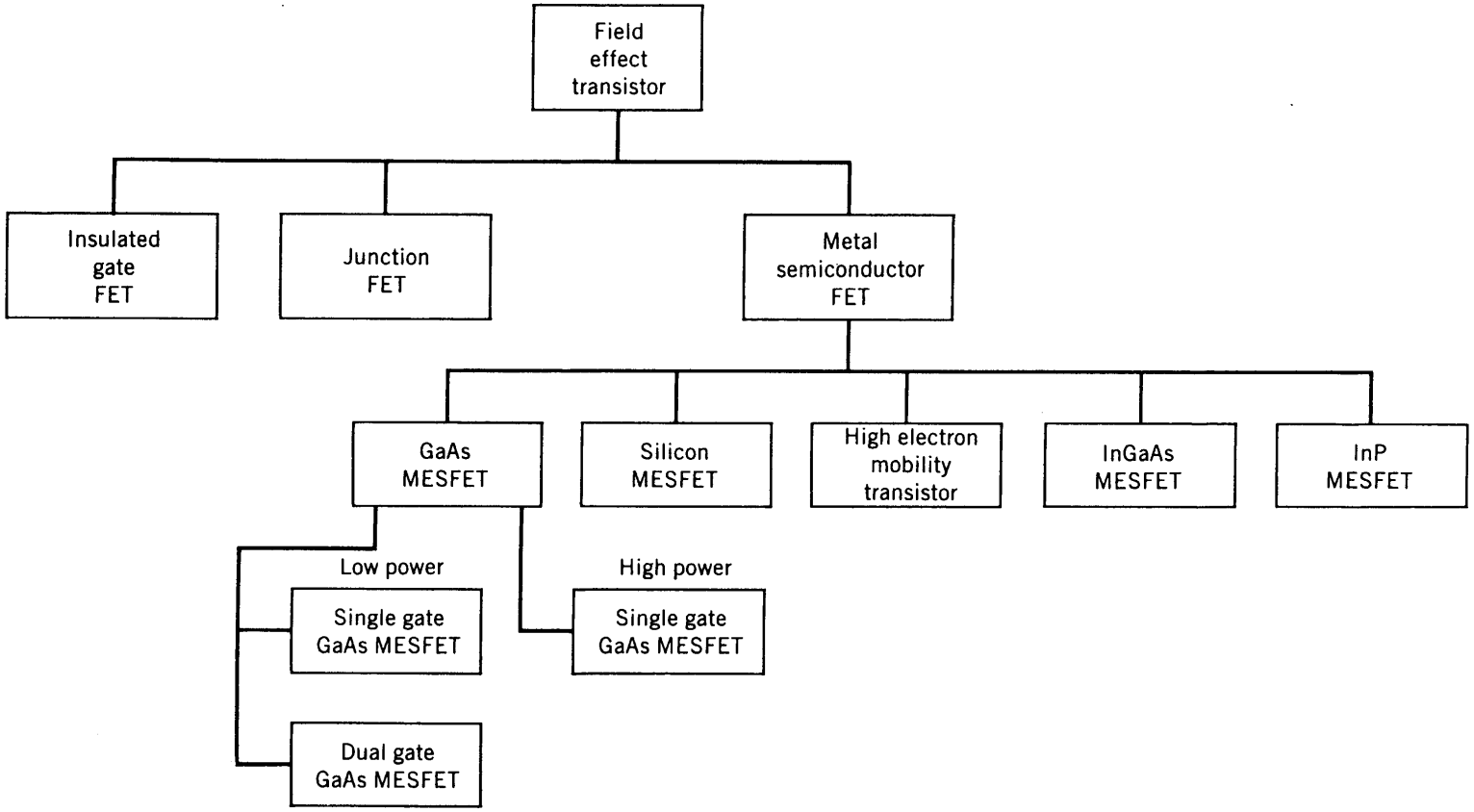


Figure 2-93 Another FET family tree [8].

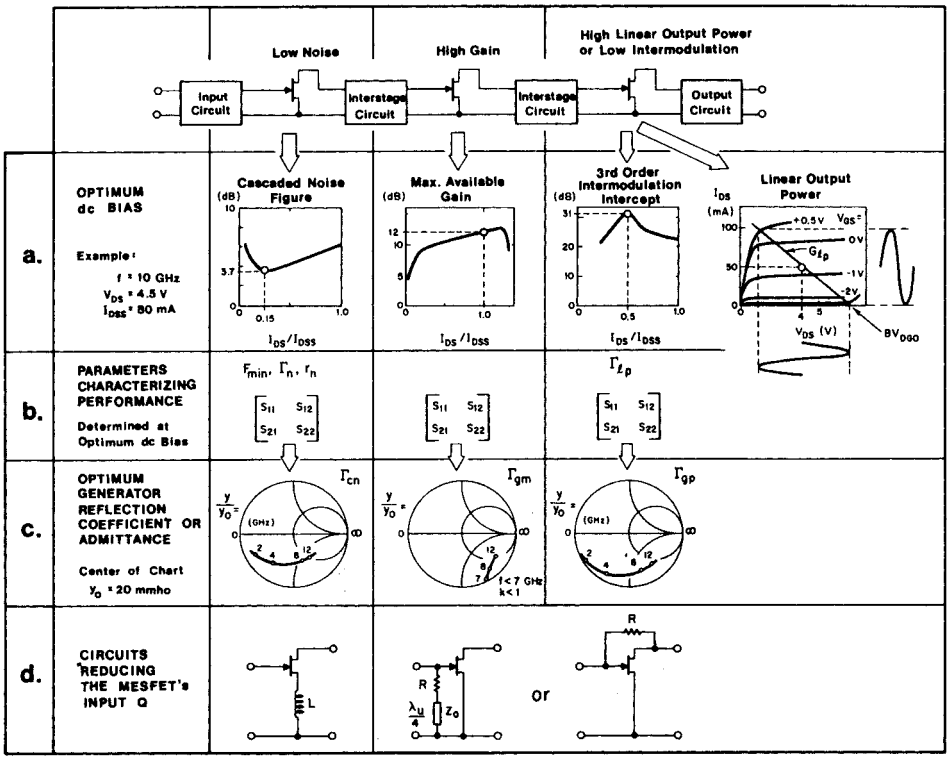


Figure 2-94 Key parameters in operating a MESFET in a low-noise front-end stage, a high-gain intermediate stage, and a linear power stage, showing (a) the device characteristics that lead to optimum dc bias for each service, (b) the parameters that characterize the device performance, (c) the optimum generator reflection coefficient that must be synthesized by each circuit, and (d) circuit arrangements that reduce the transistor's input *Q* or stabilize it at low frequencies. Although this example is based on an HP MESFET with a gate length of 1 μm and a gate width of 500 μm, this principle applies to all members of the FET family [7]. Copyright ©1976 IEEE.

n-channel JFETs *designed for . . .*

Siliconix

Performance Curves NZA See Section 4

- **VHF Amplifiers**
- **Front End High Sensitivity Amplifiers**
- **Oscillators**
- **Mixers**

BENEFITS

- Industry Standard
- High Power Gain
  16 dB at 105 MHz, Common-Gate
  11 dB at 450 MHz, Common-Gate
- Low Noise
  2.7 dB Noise Figure at 450 MHz
- Wide Dynamic Range
  Greater than 100 dB
- 75 Ω Input Match Common Gate

ABSOLUTE MAXIMUM RATINGS (25°C)

Gate-Drain or Gate-Source Voltage . . . . . -25 V
Gate Current . . . . . 20 mA
Total Power Dissipation at T_A = 25°C . . . . . 500 mW
Power Derating to 150°C . . . . . 4.0 mW/°C
Storage Temperature Range . . . . . -65 to +200°C
Lead Temperature
(1/16" from case for 10 seconds) . . . . . 300°C

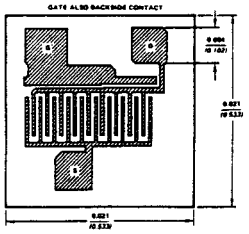
TO-52
See Section 6

ELECTRICAL CHARACTERISTICS (25°C unless otherwise noted)

		Characteristic		U308 Min	U308 Typ	U308 Max	U309 Min	U309 Typ	U309 Max	U310 Min	U310 Typ	U310 Max	Unit	Test Conditions	
1	STATIC	I_{GSS}	Gate Reverse Current			-150			-150			-150	pA	V_{GS} = -15 V,	
2						-150			-150			-150	nA	V_{GS} = 0	T_A = 125°C
3		BV_{GSS}	Gate-Source Breakdown Voltage	-25			-25			-25			V	I_G = -1 μA, V_{DS} = 0	
4		$V_{GS(off)}$	Gate-Source Cutoff Voltage	-1.0		-6.0	-1.0		-4.0	-2.5		-6.0		V_{DS} = 10 V, I_D = 1 nA	
5		I_{DSS}	Saturation Drain Current (Note 1)	12		60	12		30	24		60	mA	V_{DS} = 10 V, V_{GS} = 0	
6		$V_{GS(f)}$	Gate-Source Forward Voltage			1.0			1.0			1.0	V	I_G = 10 mA, V_{DS} = 0	
7	DYNAMIC	g_{fg}	Common-Gate Forward Transconductance (Note 1)	10	17		10	17		10	17		mmho	V_{DS} = 10 V, I_D = 10 mA	f = 1 kHz
8		g_{og}	Common-Gate Output Conductance			250			250			250	μmho		
9		C_{gd}	Drain-Gate Capacitance			2.5			2.5			2.5	pF	V_{GS} = -10 V, V_{DS} = 10 V	f = 1 MHz
10		C_{gs}	Gate-Source Capacitance			5.0			5.0			5.0			
11		$\overline{e_n}$	Equivalent Short Circuit Input Noise Voltage		10			10			10		$\frac{nV}{\sqrt{Hz}}$	V_{DS} = 10 V, I_D = 10 mA	f = 100 Hz
12	HI FREQ	g_{fg}	Common-Gate Forward Transconductance		15			15			15		mmho	V_{DS} = 10 V, I_D = 10 mA	f = 105 MHz
13					14			14			14				f = 450 MHz
14		g_{og}	Common-Gate Output Conductance		0.18			0.18			0.18				f = 105 MHz
15					0.32			0.32			0.32				f = 450 MHz
16		G_{pg}	Common-Gate Power Gain (Note 2)	14	16		14	16		14	16		dB		f = 105 MHz
17				10	11		10	11		10	11				f = 450 MHz
18		NF	Noise Figure		1.5	2.0		1.5	2.0		1.5	2.0			f = 105 MHz
19					2.7	3.5		2.7	3.5		2.7	3.5			f = 450 MHz

NOTES:
1. Pulse test duration = 2 ms.
2. Gain (G_{pg}) measured at optimum input noise match.

NZA



n-channel JFET
designed for . . .

- VHF/UHF Amplifiers
- Front End High Sensitivity Amplifiers
- Oscillators
- Mixers

BENEFITS

- Industry Standard
- High Power Gain
  16 dB at 100 MHz, Common Gate
  11 dB at 450 MHz, Common Gate

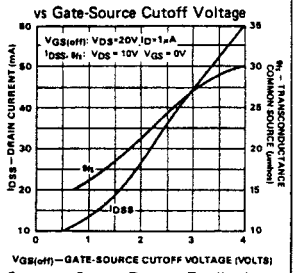
Siliconix

BENEFITS

- Low Noise
  3 dB Noise Figure at 450 MHz
- Wide Dynamic Range
  Greater Than 100 dB
- 75 Ohm Input Match Common Gate

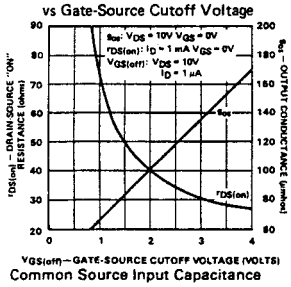
TYPE	PACKAGE	PRINCIPAL DEVICES
Single	TO-52	U308-10
Single	TO-72	U311
Single	TO-92	J308-10
Dual	TO-99	U430-1
Single	Chip	J308CHP-10CHP, U308CHP-10CHP, U311CHP
Dual	Chip	U430CHP-1CHP

Drain Current & Transconductance vs Gate-Source Cutoff Voltage

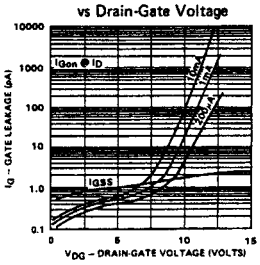


PERFORMANCE CURVES (25°C unless otherwise noted)

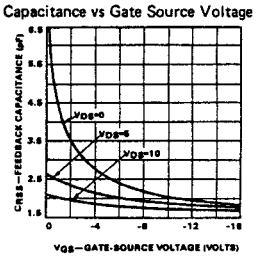
On Resistance & Output Conductance vs Gate-Source Cutoff Voltage



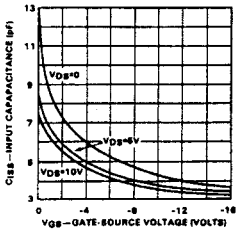
Gate Operating Current vs Drain-Gate Voltage



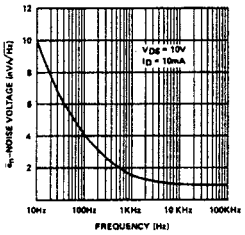
Common Source Reverse Feedback Capacitance vs Gate Source Voltage



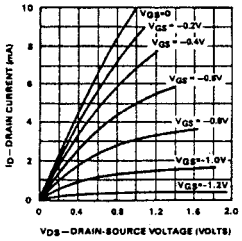
Common Source Input Capacitance vs Gate-Source Voltage



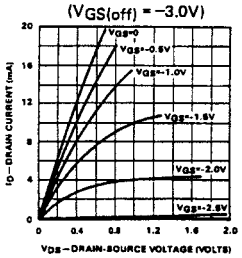
Noise Voltage vs Frequency



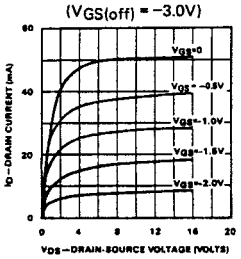
Output Characteristic ($V_{GS(off)} = -1.7V$)



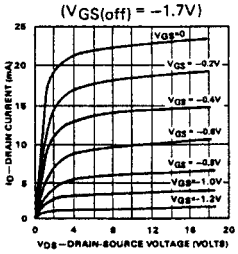
Output Characteristic ($V_{GS(off)} = -3.0V$)



Output Characteristic ($V_{GS(off)} = -3.0V$)



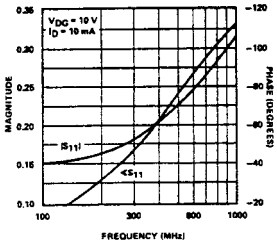
Output Characteristic ($V_{GS(off)} = -1.7V$)



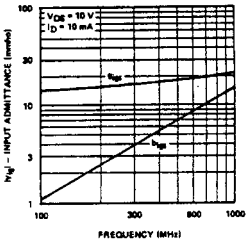
NZA

PERFORMANCE CURVES (Con't) (25°C unless otherwise noted)

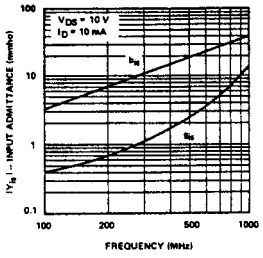
Forward Reflection Coefficient Common Gate



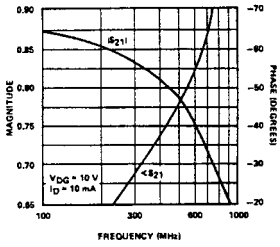
Input Admittance Common Gate



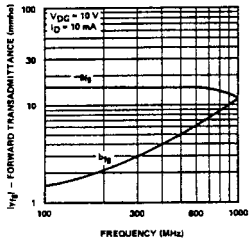
Input Admittance Common Source



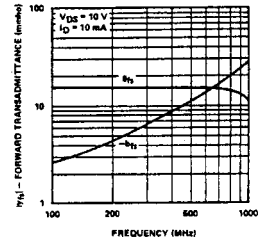
Forward Transmission Coefficient Common Gate



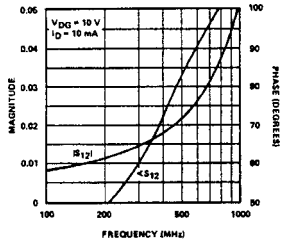
Forward Transfer Admittance Common Gate



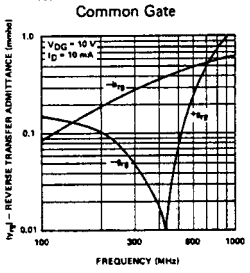
Forward Transfer Admittance Common Source



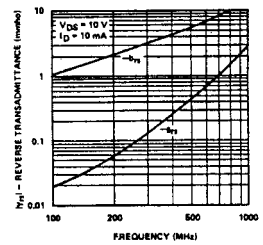
Reverse Transmission Coefficient Common Gate



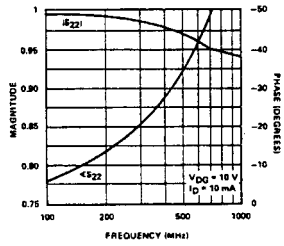
Reverse Transfer Admittance Common Gate



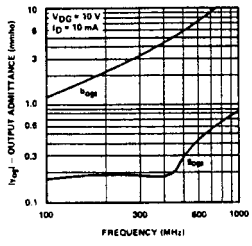
Reverse Transfer Admittance Common Source



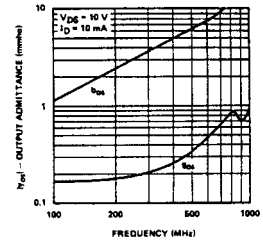
Reverse Reflection Coefficient Common Gate



Output Admittance Common Gate



Output Admittance Common Source



BROADBAND RF FET
SILICON EPITAXIAL JUNCTION N-CHANNEL FIELD EFFECT TRANSISTOR

HIGH DYNAMIC RANGE HF AND VHF AMPLIFIER
FOR USE IN COMMON GATE CONFIGURATION

- USABLE TO OVER 300 MHz
- 50 Ohm VSWR < 1.5:1 0.5-50 MHz (FIG. 1)
- LOW NOISE FIGURE — 2.2 dB TYPICAL @ 50 MHz
- INPUT Z CONSTANT 0.5-50 MHz
- HIGH IM INTERCEPT POINT — > + 40 dBm
- HIGH TRANSCONDUCTANCE — 100,000 μmhos (TYP.)
- 1 dB COMPRESSION POINT > + 20 dBm
- DYNAMIC RANGE > 140 dB (TO 1 dB COMPRESSION)
- HIGH VOLTAGE—TO 50 V.

CP640
CP664
CP665
CP666

ELECTRICAL DATA ABSOLUTE MAXIMUM RATINGS

PARAMETER	SYMBOL	CP 640	CP 664	CP 665	CP 666	UNITS
Drain to Source Voltage	BVDSO	20	30	40	50	Volts
Drain to Gate Voltage	BVDGO	20	30	40	50	Volts
Gate to Source Voltage	BVGSO	−15	−20	−20	−20	Volts
Peak Drain Current	ID	1.2	1.2	1.2	1.2	Amps
Power Dissipation 25 °C CASE	PD	8.0	8.0	8.0	8.0	Watts
Derating Factor (slope)	DF	22	22	22	22	°C/W
Junction Temp.(Oper. & Store)	TJ	−55 °C to +200 °C				

TYPICAL TWO TONE 3rd ORDER IM PRODUCTS — CIRCUIT FIGURE 1

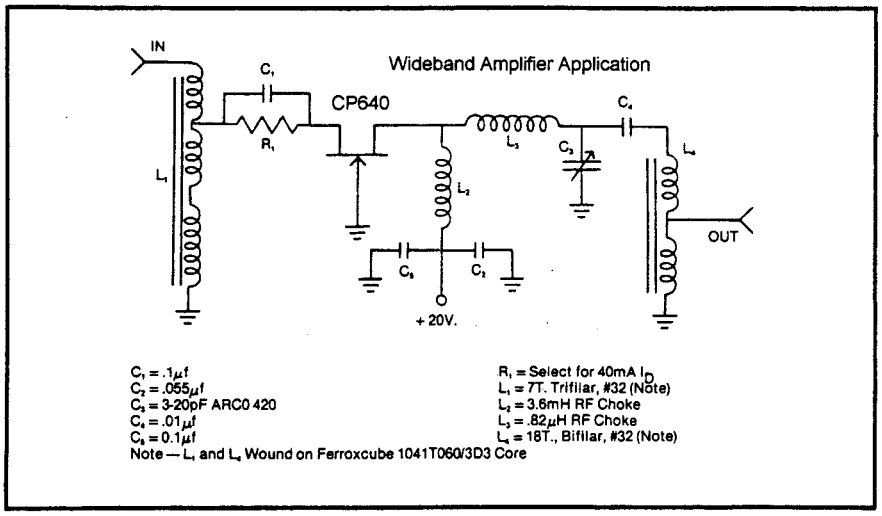
Tones at 3MHz/5MHz

Signal Level EMF	3rd Order Product
1 Volt	-44 dB
0.3 Volt	-75 dB
0.25 Volt (0dBM)	-80 dB

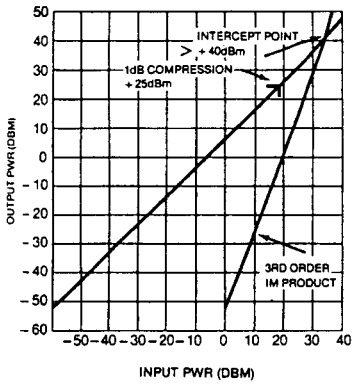
ELECTRICAL CHARACTERISTICS: T_{CASE} = 25 °C (UNLESS OTHERWISE STATED)

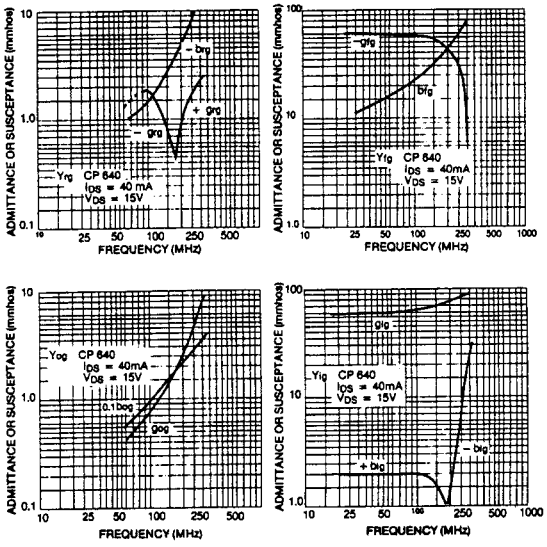
PARAMETERS	CONDITIONS		SYMBOL	Min.	Typ.	Max.	UNITS
Gate Leakage Current	V_{GS} = 15V, V_{DS} = 0	25 °C	I_{GSS}		5	100	nA
		150 °C	I_{GSS}			10	μA
Operating Transconductance	V_{DS} = 15V, I_{DS} = 40 mA		g_{fg}	40	60	80	mmho
Zero Bias Transconductance	V_{DS} = 15V, V_{GS} = 0(1)		g_{fg}	75	100	200	mmho
Gate-Source Cut-Off Voltage	V_{DS} = 5V, I_{DS} = 1.0 mA		$V_{GS(off)}$	2	5	10	Volts
Zero Bias Drain Current	V_{DS} = 15V, V_{GS} = 0 (1)		I_{DSS}	100	400	800	mA
Gate to Source Cap.	V_{GS} = −20V		C_{GS}		15	20	pf
Gate to Drain Cap.	V_{GD} = −20V		C_{GD}		15	20	pf
Power Gain	I_{DS} = 40mA, f = 50MHz, Fig. 1		Gpg	8	8.5	9.5	dB
Noise Figure	I_{DS} = 40mA, f = 30MHz, Fig. 1		N.F.		2.2	3.0	dB
Voltage Standing Wave Ratio	f = 0.5-50MHz, 50 Ω Source, Fig. 1		VSWR			1.5:1	
Common Gate Input Conductance	f = 0.5-50MHz, V_{DS} = 15, I_D = 40mA		g_{igs}		60		mmho
Common Gate Output Conductance	f = −50MHz, V_{DS} = 15, I_D = 40mA		g_{ogs}		0.4		mmho

(1)Pulse Measurement 1% Duty Cycle 10 mS Max.



TYPICAL INTERCEPT AND COMPRESSION POINT





2-3-1 Large-Signal Behavior of JFETs

A section view of a junction field-effect transistor (JFET) is shown in Figure 2-95. This structure contains, between the source and the drain contacts, an *n*-type "channel" embedded in a *p*-type silicon substrate. If it is assumed that the *pn* junction forms a barrier to current flow, then it can be seen that channel conduction is a function of the channel width, length, and thickness and of the density and mobility of the carriers. In this structure, current can flow equally well in either direction through the channel; that is, the drain can be positive or negative with respect to the source.

Since we have found that the SPICE models for JFETs are really incomplete and give poor answers, we will only briefly touch here on the JFET's performance. The modified Materka model, one of the best GaAsFET models, turns out to be best suited for JFETs because both transistors act similarly in the dc area. Both have a gate-to-source diode that becomes conductive above 0.7 V at the input and they are also quite similar in other respects. Their equivalent circuits are also quite similar, as parameter extractions for both approaches have shown.

In general, in a three-terminal device, the drain current I_D is a function of two variables: V_{DS} and V_{GS}. This function is best represented by families of characteristic curves, as shown in Figure 2-96. These curves are for the "common source" configuration, with the drain as the output and the gate as the input. They reveal that for this device, if V_{DS} is greater than about 2 V (but less than the drain breakdown), I_D is primarily determined by the gate voltage V_{GS}. Under these conditions it is valid for small signals to characterize the FET by a single transfer characteristic curve, commonly called the "forward transconductance curve," such as the upper curve shown in Figure 2-96b. Some of the relationships between the forward transfer curve and the output characteristic curves of Figure 2-96a will be examined. The value of V_{GS} that reduces I_D to approximately zero is the gate–source cutoff voltage, $V_{GS(\text{off})}$. With reference to the output curve, Figure 2-96a, note that the drain current at $V_{GS} = 0$ tends to become saturated at a drain voltage approximately equal in magnitude to $-V_{GS(\text{off})}$. This drain voltage is often referred to as the pinchoff voltage V_P; however, in this section pinchoff voltage is used interchangeably with gate–source cutoff voltage. V_P will have the same meaning as $V_{GS(\text{off})}$. The symbol I_{DSS} is commonly used to indicate the value of saturated drain current at $V_{GS} = 0$.

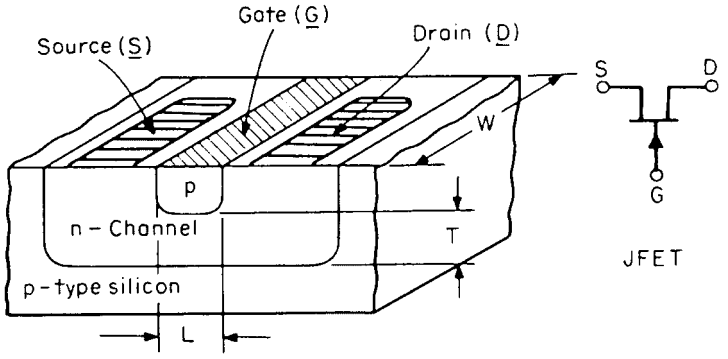


Figure 2-95 Junction field-effect transistor [9].

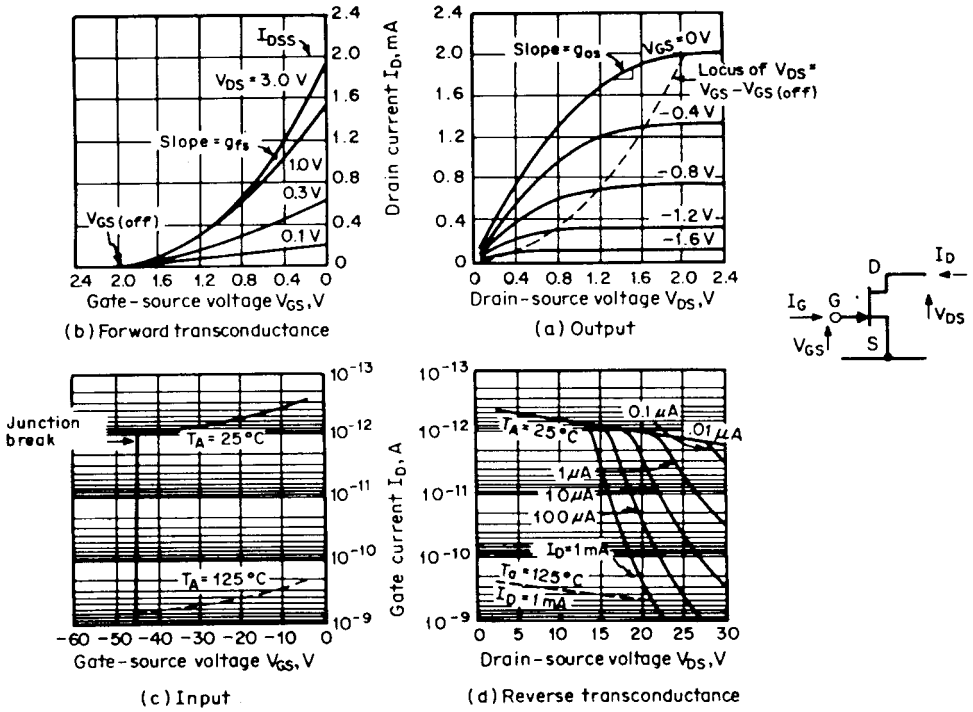


Figure 2-96 Static characteristics of an *n*-channel JFET.

The forward transconductance characteristic of Figure 2-96b can be approximated by a power law relation expressed as

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right)^n \tag{2-190}$$

if $V_{DS} \geq -V_{GS(\text{off})}$. By differentiation the small-signal transconductance g_{fs} is given by

$$g_{fs} = \frac{dI_D}{dV_{GS}} = -n\frac{I_{DSS}}{V_{GS(\text{off})}}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right)^{n-1} \tag{2-191}$$

Some texts indicate a value of $\frac{3}{2}$ for n; however, experimental measurements on a number of n-channel JFET geometries indicate that the exponent n is close to 2, which is the value derived in an approximate treatment by R. D. Middlebrook [10].

A useful relationship between g_{fso}, I_{DSS}, and $V_{GS(\text{off})}$ is derived from the ratio of Eqs. (2-190) and (2-191):

$$\frac{g_{gs}}{I_D} = n(V_{GS} - V_{GS(\text{off})})^{-1} \tag{2-192}$$

At $V_{GS} = 0$, $I_D = I_{DSS}$ and $g_{fs} = g_{fso}$. Using 2 as the value of the constant n leads to

$$g_{fso} = -2\frac{I_{DSS}}{V_{GS(\text{off})}} \tag{2-193}$$

For n-channel FETs I_{DSS} is positive and $V_{GS(\text{off})}$ is negative; for p-channel FETs I_{DSS} is negative and $V_{GS(\text{off})}$ is positive; thus g_{fs} is a positive quantity for both p- and n-channel FETs.

Equation (2-190) indicates that for $V_{GS} = V_{GS(\text{off})}$, $I_D = 0$. In a real device this does not happen. Starting from zero, as the gate voltage is made more negative, the drain current decreases until it reaches a very low value equal to the drain-leakage current. At this value, the source current will consist of source–gate leakage. Any further increase in the magnitude of the negative gate voltage will result in an increase in I_D leakage. For the small-signal device illustrated, this minimum I_D is on the order of 2×10^{-13} A.

For some types of applications of the FET, it is helpful to understand the characteristics at very low values of V_{DS}, such as those shown in Figure 2-97. A "very low value" is one that is small compared to the magnitude of $V_{GS} - V_{GS(\text{off})}$. In this region, V_{DS} is small enough to have little effect on channel thickness, so that the I_D/V_{DS} slope is nearly linear. Since the slope is a function of V_{GS}, the FET can be utilized as a voltage-controlled resistor. The conductance slope ($\Delta I_D/V_{DS}$) at $V_{DS} = 0$ is approximately a linear function of $V_{GS} - V_{GS(\text{off})}$.

If g_{ds} at $V_{GS} = 0$ is given the term g_{dso}, then

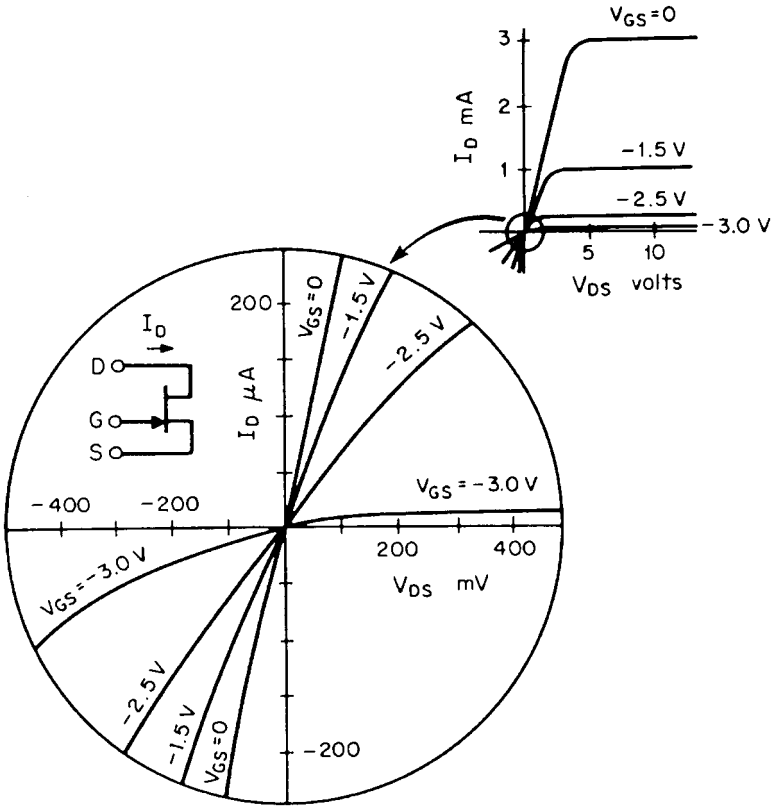


Figure 2-97 Enlargement of n-channel output characteristic around $V_{DS} = 0$.

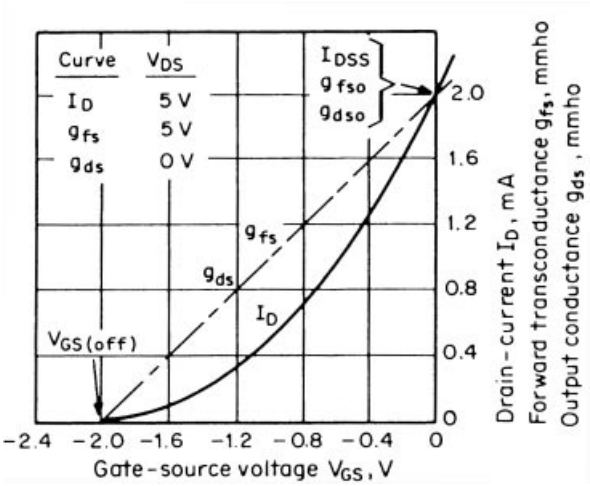


Figure 2-98 JFET I_D, g_{fs}, and g_{ds}.

$$g_{ds} = g_{dso}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right) \tag{2-194}$$

with $V_{DS} = 0$. A plot of this characteristic is shown in Figure 2-98, along with g_{fs} and I_D characteristics.

The relationship between g_{dso}, $V_{GS(\text{off})}$, and I_{DSS} is given by the equation

$$g_{dso} = -2\frac{I_{DSS}}{V_{GS(\text{off})}} \tag{2-195}$$

where I_{DSS} and $V_{GS(\text{off})}$ are as indicated in Figure 2-96. It is important to note that Eqs. (2-194) and (2-195) and the g_{ds} curve of Figure 2-98 are valid only for the case where V_{DS} is very small compared to V_P. Drain–source conductance at higher values of V_{DS} will be discussed later. FETs designed to be used as voltage-controlled resistors typically have a high $V_{GS(\text{off})}$ because the $V_{DS}/(V_{GS} - V_{GS(\text{off})})$ ratio should be low to keep distortion low. The actual large-signal JFET model as used in most simulators is shown in Figure 2-99, including a list of its intrinsic keywords (Table 2-14). Again, we had very little luck using this model as published by several SPICE CAD software manufacturers as offered. While we cannot judge on the quality of the parameters that came with the software, it was not possible to closely match even optimized sets of parameters to this large-signal model and obtain acceptable results. However, since the modified Materka model worked very well for this, we recommend that this popular JFET model not be used at frequencies above 10 MHz.

2-3-2 Small-Signal Behavior of JFETs

Figure 2-100 shows the small-signal equivalent transistor model we would recommend. It consists of an empirical FET model and a package model. Table 2-15 lists its keywords. Its advantage is that its built-in noise model is very accurate, even for JFETs and metal semiconductor FETs.

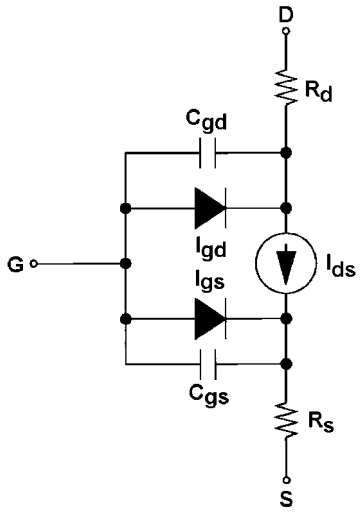


Figure 2-99 The nonlinear JFET model. Table 2-14 lists its intrinsic and extrinsic parameters. In our opinion, this model is limited to frequencies below 10 MHz; above this frequency, the modified Materka model should replace it.

Table 2-14 Large-signal JFET model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
NJF/PJF	Channel-type selection (NJF = *n*-channel, PJF = *p*-channel)		NJF
VT0	Threshold voltage	volt	–2.0
BETA	Transconductance coefficient	ampere/volt^2	1.0E–4
LAMB	Channel-length modulation	/volt	0.0
IS	Gate-junction saturation current	ampere	1.0E–14
PB	Gate-junction potential	volt	1.0
FC	Forward-bias depletion capacitance coefficient		0.5
CGS	Zero-bias gate–source junction capacitance	farad	0.0
CGD	Zero-bias gate–drain junction capacitance	farad	0.0
T	Channel transit-time delay	second	0.0
KF	Flicker noise coefficient		0.0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
SN	Switch to turn device shot noise on (1) or off (0)		1
AREA	Area multiplier		1.0
NOIS	Reference label to a set of noise data		
NAME	Required user-specified name up to 8 characters		
	Extrinsic Model		
RD	Drain ohmic resistance	ohm	0.0
RS	Source ohmic resistance	ohm	0.0

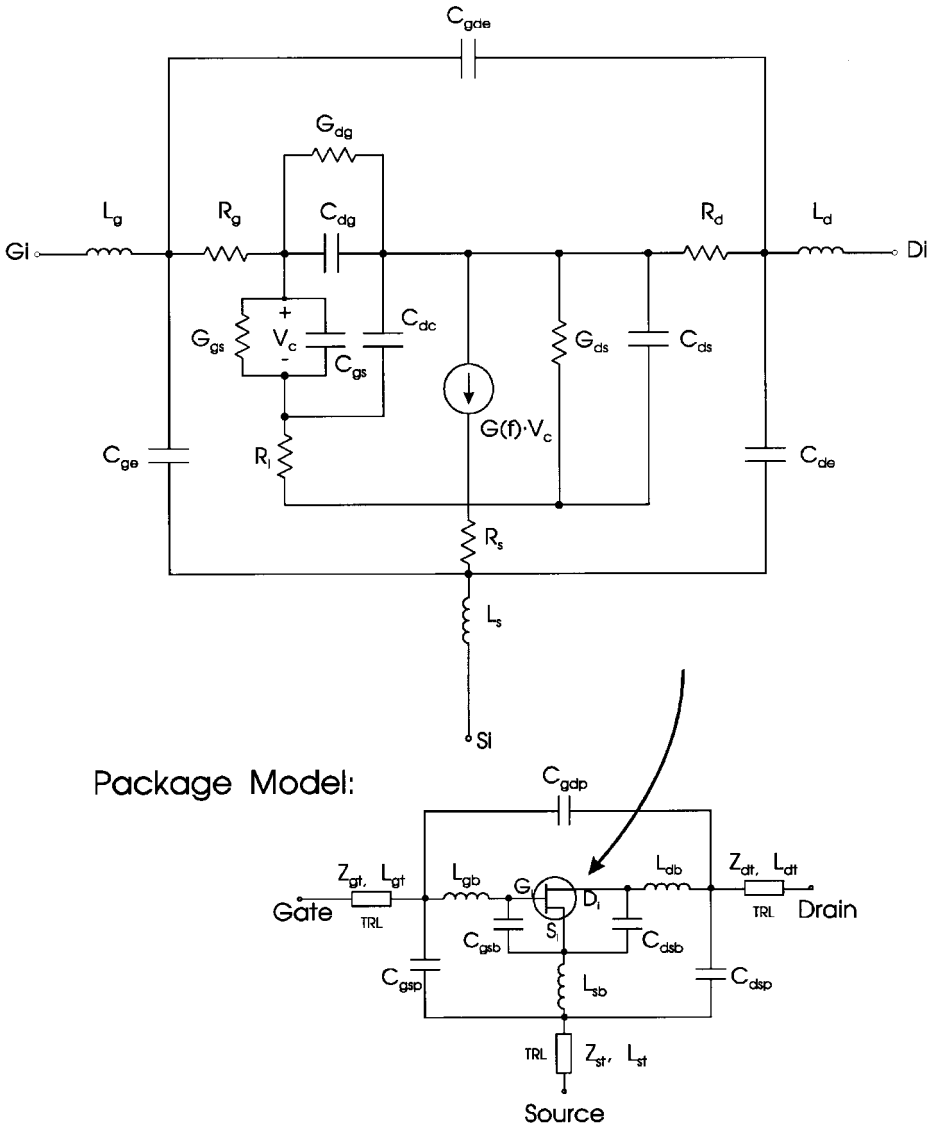


Figure 2-100 Linear FET model.

Electrical noise generated within a JFET is usually represented by equivalent noise sources, $\overline{e}_n$ and $\overline{i}_n$. Both noise voltage $\overline{e}_n$ and noise current $\overline{i}_n$ are frequency dependent and have the characteristics shown in Figure 2-101.

An equivalent noise circuit is shown in Figure 2-102. Above the frequency f_1, $\overline{e}_n$ is approximately given by

$$\overline{e}_n \cong \left(4KTB\,\frac{0.67}{g_{fs}}\right)^{1/2} \tag{2-196}$$

Table 2-15 Small-signal FET model keywords

Keyword	Description	Unit	Default
	Intrinsic model		
G	Transconductance at dc, G_0[a]	/ohm	
CGS	Gate–source capacitance	farad	
F	3-dB roll-off frequency	hertz	∞
T	Time delay	second	0.0
TDS	Drain–source time delay	second	0.0
GGS	Gate–source conductance	/ohm	0.0
CDG	Drain–gate capacitance	farad	0.0
CDC	Dipole layer capacitance	farad	0.0
CDS	Drain–source capacitance	farad	0.0
GDS	Drain–source conductance	/ohm	0.0
RI	Channel resistance	ohm	0.0
RG	Gate resistance	ohm	0.0
RD	Drain resistance	ohm	0.0
RS	Source resistance	ohm	0.0
CGE	External gate capacitance	farad	0.0
CDE	External drain capacitance	farad	0.0
LG	Gate-lead inductance	henry	0.0
LD	Drain-lead inductance	henry	0.0
LS	Source-lead inductance	henry	0.0
CGDE	External gate–drain capacitance	farad	0.0
GDG	Gate–drain conductance	/ohm	0.0
TJ	Chip temperature	kelvin	298
	Package Parasitics		
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate to source package capacitance	farad	0.0
CDSP	Drain to source package capacitance	farad	0.0
CGDP	Gate to drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for at $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for $\varepsilon_r = 1$	meter	0.0
FC	Corner frequency of flicker (1/*f*) noise[b]	hertz	10 MHz
FCP	Shape factor of the 1/*f* noise response		1.0
label	User-defined term that refers to temperature coefficient		

[a]The transconductance of this model may be approximately described by $g_m = G\,\frac{e^{-j\omega T}}{1 + j(f/F)}$ where $\omega = 2\pi f$ and f is frequency.
[b]The flicker noise frequency dependence is given by $1/(f/F_c)^{\mathrm{FCP}}$.

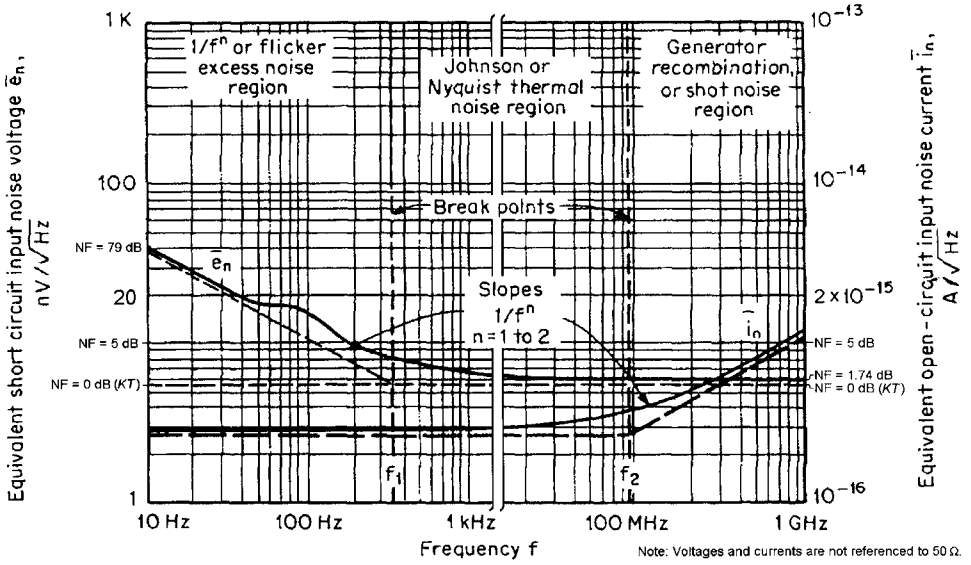


Figure 2-101 JFET noise characteristics.

where $K = 1.374 \times 10^{-23}$ J/K

T = absolute temperature in kelvins (273 K = 0 °C)

B = frequency range in hertz

g_{fs} = transconductance of FET

With the input short-circuited, the noise voltage across the load R_L resulting from the FET is

$$\text{Output noise voltage} = \bar{e}_n A_V \tag{2-197}$$

Below the frequency f_1, $\bar{e}_n$ increases proportional to $1/f^n$ and is expressed as

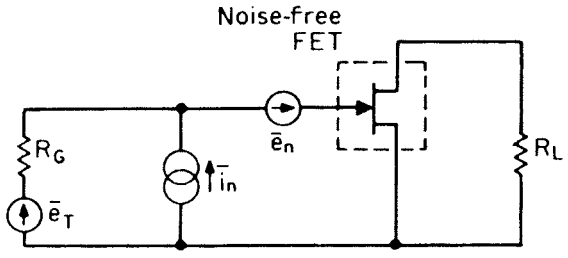


Figure 2-102 Equivalent noise circuit of the FET.

$$\bar{e}_n = \left[4KTB \left(\frac{0.67}{g_{fs}} \right) \left(1 + \frac{f_1}{f^n} \right)^{1/2} \right] \tag{2-198}$$

The low-frequency corner frequency f_1 for JFETs is typically in the 100-Hz to 1-kHz range, and the exponent n is usually between 1 and 2. As indicated by the equations, $\bar{e}_n$ is inversely proportional to the square root of g_{fs}.

The equivalent input noise current $\bar{i}_n$ is caused by the current in the gate-to-channel junction. Its approximate value below f_2 is

$$\bar{i}_n = (2qI_G B)^{1/2} \tag{2-199}$$

where $q = 1.602 \times 10^{-19}$
B = frequency range in hertz
I_G = dc gate current

This expression is fairly accurate when I_G is the result of the active device conductance. Typically, $\bar{i}_n$ will be lower than the calculated value because part of I_G is due to conductance across the device package.

At higher frequencies (above f_2),

$$\bar{i}_n = \left(\frac{4KTB}{R_P} \right)^{1/2} \tag{2-200}$$

where R_P is the real part of gate input impedance. R_P, in terms of Y_{11}, can range from several tens of megohms (at audiofrequencies) to 1 kΩ or less (at VHF/UHF frequencies). The high-frequency corner f_2 is typically in the range of 5–50 kHz and more than 100 MHz for high-frequency devices.

Another form of noise is known as "popcorn" or burst noise, the causes of which have not been completely identified. It shows up as a random short-duration step-function change in drain current, equivalent to an input gate–source voltage change of a few tenths of a microvolt. It is not unlike big bubbles on the surface of boiling water.

As far as the noise is concerned, we can also state that the Materka built-in proprietary noise model is significantly more accurate than the one mentioned above [11–14]. This model is implemented, with minimal documentation of its inner workings because of its proprietary nature, in the linear portion of the circuit simulator in Ansoft's Serenade Design Environment. In many cases, the manufacturer supplies only limited data, so it is most convenient that after SPICE-type parameter extraction is done by using the Scout program, one even gets good noise prediction for JFETs using the modified Materka model.

2-3-3 Large-Signal Behavior of MOSFETs

Metal-oxide-semiconductor field-effect transistors (MOSFETs) are important components in contemporary analog integrated circuits. Whereas initial applications were centered on all-MOS processes, combined bipolar and MOS processes give the designer the best of both worlds. A major advantage of MOS processes for realizing analog functions is that complex, dense digital functions can be realized on the same chip.

Transfer Characteristics of MOS Devices. A cross section of a typical enhancement-mode *n*-channel MOS transistor (NMOS) is shown in Figure 2-103a. Heavily doped *n*-type source and drain regions are fabricated in a *p*-type substrate (often called the body). A thick layer of silicon dioxide is grown over the substrate material and a conductive gate material (metal or polycrystalline silicon) covers this oxide between source and drain. The operation of the device is very similar to that of a JFET in that the gate–source voltage is used to modify the conductance of the region under the gate. This allows the gate voltage to control the current flowing between source and drain, giving gain in analog circuits and switching characteristics in digital circuits.

The enhancement-mode NMOS device of Figure 2-103a shows significant conduction between source and drain only when an *n*-type channel exists under the gate, and this is the origin of the "*n*-channel" designation. The term "enhancement mode" refers to the fact that no conduction occurs for $V_{GS} = 0$, and thus the channel must be "enhanced" to cause conduction. MOS devices can be made equally well by using an *n*-type substrate with a *p*-type conducting channel. Such devices are called enhancement-mode *p*-channel MOS transistors (PMOS). In technologies employing one or the other of these device types, the circuit symbol of Figure 2-103b is commonly used for either one. In complementary MOS technology (CMOS) both device types are present and the circuit symbols of Figure 2-103c will be used to distinguish them. In PMOS and NMOS technologies, the substrate is common to all devices, invariably connected to a dc power supply voltage, and usually not shown on the circuit diagram. In CMOS technology, however, devices of one type or another are fabricated in individual, separate isolation regions, which may or may not be connected to a power supply voltage. If these isolation regions *are* connected to the appropriate power supply, the symbols of Figure 2-103c will be used and the substrate connection will not be shown. If the individual isolation regions are connected elsewhere, however, the devices will be represented by the symbols of Figure 2-103d, where the substrate is labeled *B*. Finally, in NMOS technology an additional device type called a "depletion-mode" device is usually available. This is a conducting channel implanted between the source and drain so that conduction occurs for $V_{GS} = 0$. This device has characteristics that are almost identical to that of a JFET and will be represented by the symbol of Figure 2-103e.

The derivation of the transfer characteristics of the enhancement-mode NMOS device of Figure 2-103a begins by noting that with $V_{GS} = 0$, the source and drain regions are separated by back-to-back *pn* junctions. These junctions are formed between the *n*-type source and drain regions and the *p*-type substrate, resulting in an extremely high resistance (about 10^{12} Ω) between drain and source when the device is off.

Now consider substrate, source, and drain grounded and a positive voltage V_{GS} applied to the gate as shown in Figure 2-104. The gate and substrate then form the plates of a capacitor with the SiO_2 as a dielectric. Positive charge accumulates on the gate and the negative charge in the substrate. Initially, the negative charge in the *p*-type substrate is manifested by creation of a *depletion region* and resulting exclusion of holes under the gate. This is shown in Figure 2-104. The depletion-layer width *X* under the oxide is

$$X = \left(\frac{2\varepsilon\phi}{qN_A} \right)^{1/2} \tag{2-201}$$

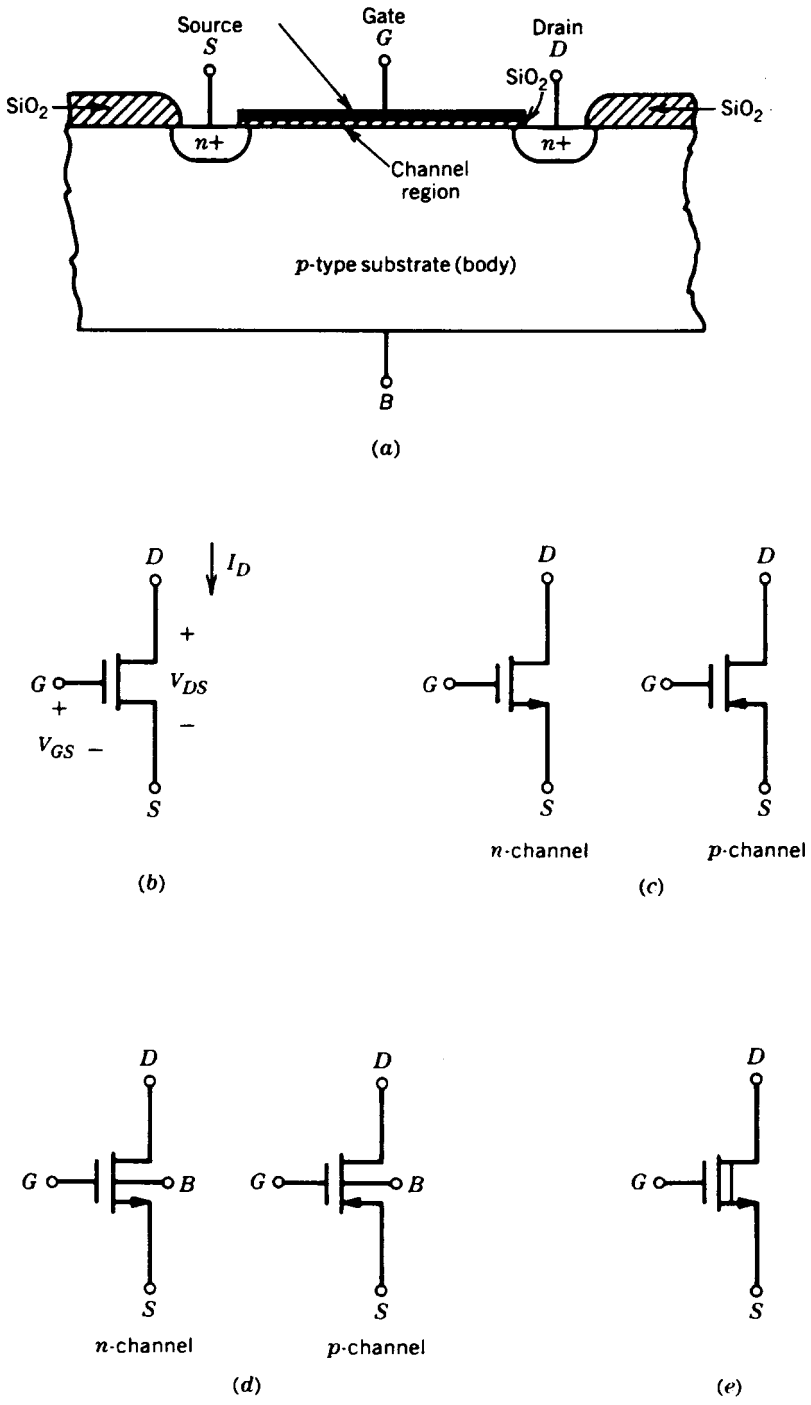


Figure 2-103 (a) Typical enhancement-mode NMOS structure. (b) Enhancement-mode NMOS or PMOS circuit symbol when one device type only is present. (c) NMOS and PMOS symbols used in CMOS circuits. (d) NMOS and PMOS symbols used when the substrate connection is nonstandard. (e) Depletion MOS device symbol.

where ϕ is the potential in the depletion layer at the oxide–silicon interface, N_A atoms/cm^3 is the doping density (assumed constant) of the p-type substrate, and ε is the permittivity of the silicon. The charge per unit area in this depletion region is

$$Q = qN_A X = \sqrt{2qN_A \varepsilon \phi} \tag{2-202}$$

When the potential in the silicon reaches a critical value equal to twice the Fermi level, $\phi_f \simeq 0.3$ V, a phenomenon known as "inversion" occurs [15]. Further increases in gate voltage produce no further changes in the depletion-layer width but instead a thin layer of electronics is induced in the depletion layer directly under the oxide. This produces a continuous n-type region with the source and drain regions and is the conducting channel between source and drain. This channel can then be modulated by increases or decreases in the gate voltage. In the presence of an inversion layer, and with no substrate bias, the depletion region contains a fixed charge

$$Q_{b0} = \sqrt{2qN_A \varepsilon 2\phi_f} \tag{2-203}$$

If a substrate bias voltage V_{SB} (source positive for n-channel devices) is applied between source and substrate, the potential required to produce inversion becomes $(2\phi_f + V_{SB})$ and the charge stored in the depletion region in general is

$$Q_b = \sqrt{2qN_A \varepsilon (2\phi_f + V_{SB})} \tag{2-204}$$

The gate voltage V_{GS}, required to produce an inversion layer, is called the threshold voltage V_t and can now be calculated. This voltage consists of several components. First, a voltage $[2\phi_f + (Q_b/C_{ox})]$ is required to sustain the depletion-layer charge Q_b, where C_{ox} is the gate oxide capacitance per unit area. Second, a work-function difference ϕ_{ms} exists between the

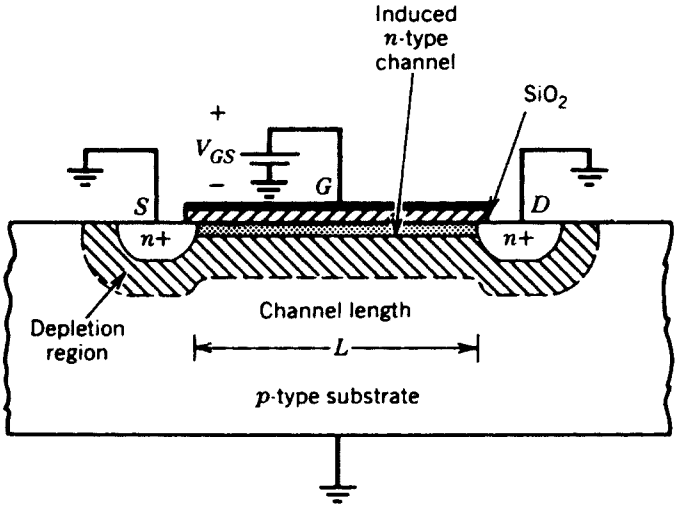


Figure 2-104 Idealized NMOS device cross section with positive V_{GS} applied, showing depletion regions and the induced channel.

gate metal and the silicon. Third, there is always charge density Q_{ss} (positive) in the oxide at the silicon interface. This is caused by crystal discontinuities at the Si–SiO_2 interface and must be compensated by a gate voltage contribution of $-Q_{ss}/C_{ox}$. Thus we have a threshold voltage

$$V_t = \phi_{ms} + 2\phi_f + \frac{Q_b}{C_{ox}} - \frac{Q_{ss}}{C_{ox}}$$

$$= \phi_{ms} + 2\phi_f + \frac{Q_{b0}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} + \frac{Q_b - Q_{b0}}{C_{ox}} \tag{2-205}$$

$$= V_{t0} + \gamma\left(\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f}\right) \tag{2-206}$$

where Eqs. (2-203) and (2-204) have been used and V_{t0} is the threshold voltage with $V_{SB} = 0$. The parameter γ is defined as

$$\gamma = \frac{1}{C_{ox}}\sqrt{2q\varepsilon N_A} \tag{2-207}$$

and

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \tag{2-208}$$

where ε_{ox} and t_{ox} are the permittivity and thickness of the oxide, respectively. A typical value of γ is 0.5 $V^{1/2}$ and $C_{ox} = 3.5 \times 10^{-4}$ pF/μ^2 for $t_{ox} = 0.1$ μm.

In practice, the value of V_{t0} is usually adjusted in processing by implanting additional impurities into the channel region. Extra *p*-type impurities are implanted in the channel to make $V_{t0} \simeq 0.5$–1.5 V for *n*-channel enhancement devices. By implanting *n*-type impurities in the channel region, a conducting channel can be formed, even for $V_{GS} = 0$. This MOS transistor is called a depletion device, with typical values of V_{t0} in the range –1 to –4 V. If Q_i is the charge density per unit area due to the implant, then the threshold voltage given by Eq. (2-205) is shifted by approximately Q_i/C_{ox}.

The preceding equations can now be used to calculate the large-signal characteristics of an NMOS transistor. For purposes of analysis the source is assumed grounded and bias voltages V_{GS}, V_{DS}, and V_{SB} are applied as shown in Figure 2-105. If V_{GS} is greater than V_t, a conducting channel exists and V_{DS} causes a larger reverse bias from drain to substrate than exists from source to substrate, and thus a wider depletion region exists at the drain. However, for simplicity, we assume that the voltage drop along the channel itself is small so that the depletion layer is constant along the channel.

At a distance *y* along the channel the voltage with respect to the source is $V(y)$ and the gate-to-channel voltage at that point is $V_{GS} - V(y)$. We assume this voltage exceeds the threshold voltage V_t, and thus the induced electron charge per unit area in the channel is

$$Q_I(y) = C_{ox}[V_{GS} - V(y) - V_t] \tag{2-209}$$

The resistance *dR* of a length *dy* of the channel is

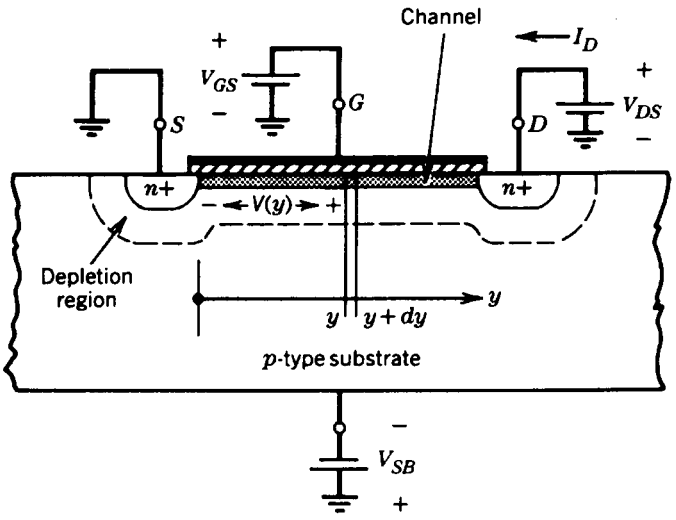


Figure 2-105 NMOS device with bias voltages applied.

$$dR = \frac{dy}{W\mu_n Q_I(y)} \tag{2-210}$$

where W is the width of the device, perpendicular to the plane of Figure 2-105, and μ_n is the average electron mobility of the channel.

The voltage drop dV along the length of the channel dy is

$$dV = I_D\, dR = \frac{I_D}{W\mu_n Q_I(y)}\, dy \tag{2-211}$$

If L is the total channel length, then substitution of Eq. (2-209) into Eq. (2-211) and integration gives

$$\int_0^L I_D\, dy = \int_0^{V_{DS}} W\mu_n C_{\text{ox}}(V_{GS} - V - V_t)\, dV \tag{2-212}$$

This results in

$$I_D = \frac{k'}{2}\frac{W}{L}[2(V_{GS} - V_t)V_{DS} - V_{DS}^2] \tag{2-213}$$

where

$$k' = \mu_n C_{\text{ox}} = \frac{\mu_n \varepsilon_{\text{ox}}}{t_{\text{ox}}} \tag{2-214}$$

Equation (2-213) is important and describes the I–V characteristics of an MOS transistor, assuming a continuous induced channel. A typical value of k' for $t_{\text{ox}} = 0.1$ μm is about 20 μA/V^2 for an n-channel device.

As the value of V_{DS} is increased, the induced conducting channel narrows at the drain end and Eq. (2-209) indicates that Q_I at the drain end approaches zero as V_{DS} approaches $(V_{GS} - V_t)$. This results in the same pinchoff phenomenon as occurs in a JFET, and further increases in V_{DS} produce little change in I_D. Equation (2-213) is thus no longer valid if V_{DS} is greater than $(V_{GS} - V_t)$. The value of I_D in this region is obtained by substituting $V_{DS} = (V_{GS} - V_t)$ in Eq. (2-213), giving

$$I_D = \frac{k'}{2}\frac{W}{L}(V_{GS} - V_t)^2 \tag{2-215}$$

for an MOS transistor in the pinchoff region. As in the case of a JFET, the drain current in the pinchoff region varies slightly as the drain voltage is varied. This is due to the presence of a depletion region between the physical pinchoff point in the channel at the drain end and the drain region itself. If this depletion-layer width is X_d, then the *effective* channel length is given by

$$L_{\text{eff}} = L - X_d \tag{2-216}$$

If L_{eff} is used in place of L in Eq. (2-215), we obtain a more accurate formula for the pinchoff region:

$$I_D = \frac{k'}{2}\frac{W}{L_{\text{eff}}}(V_{GS} - V_t)^2 \tag{2-217}$$

The fact that X_d (and thus L_{eff}) is a function of the drain–source voltage results in a variation of I_D and V_{DS} in the pinchoff region. Using Eqs. (2-216) and (2-217), we obtain

$$\frac{\partial I_D}{\partial V_{DS}} = -\frac{k'}{2}\frac{W}{L_{\text{eff}}^2}(V_{GS} - V_t)^2\frac{dL_{\text{eff}}}{dV_{DS}} \tag{2-218}$$

and thus

$$\frac{\partial I_D}{\partial V_{DS}} = \frac{I_D}{L_{\text{eff}}}\frac{dX_d}{dV_{DS}} \tag{2-219}$$

This equation is analogous to Eq. (2-168) for bipolar transistors. Following a similar procedure, a voltage analogous to the Early voltage can be defined as

$$V_A = \frac{I_D}{\partial I_D/\partial V_{DS}} \tag{2-220}$$

and thus

$$V_A = L_{\text{eff}}\left(\frac{dX_d}{dV_{DS}}\right)^{-1} \tag{2-221}$$

For MOS transistors, the most widely used parameter for the characterization of output resistance is

$$\lambda = \frac{1}{V_A} \tag{2-222}$$

As in the bipolar case, the large-signal properties of the transistor can be approximated by assuming that λ and V_A are constants, independent of bias conditions. Thus we can formulate a better approximation to the I–V characteristics as

$$I_D = \frac{k'}{2}\frac{W}{L}(V_{GS} - V_t)^2 (1 + \lambda V_{DS}) \tag{2-223}$$

In practical MOS transistors variation of X_d with voltage is complicated by the fact that the field distribution in the drain depletion region is not one dimensional. A vertical component in the field distribution is introduced by the potential difference between the gate and channel and the gate and drain. As a result, the calculation of λ from the device structure is quite different [16], and it is usually necessary to develop effective values of λ from experimental data. The parameter λ is a linear function of effective channel length and is an increasing function of the doping level in the channel. Typical values of λ are in the range 0.05–0.005 V^{-1}. Note that, in the case of a JFET, the pinchoff region for MOS devices is often called the *saturation* region.

A plot of I_D versus V_{DS} for an NMOS transistor is shown in Figure 2-106. Below pinchoff, the device behaves as a nonlinear voltage-controlled resistor, which is often called the *ohmic* or *triode* region. Above pinchoff, the device approximates a voltage-controlled current source. Note that for depletion MOS devices, V_t is negative and I_D is finite, even for $V_{GS} = 0$. For PMOS devices, all polarities of voltage and current are reversed.

The results as previously derived can be used to form a large-signal model of an MOS transistor. The model topology in the pinchoff region is the same as that for the JFET, but using Eq. (2-215) for the controlled-current generator.

MOS Device Voltage Limitations. The voltage limitations of MOS transistors depend on the gate length L. For small values of L (less than about 10 μm), the drain-depletion region

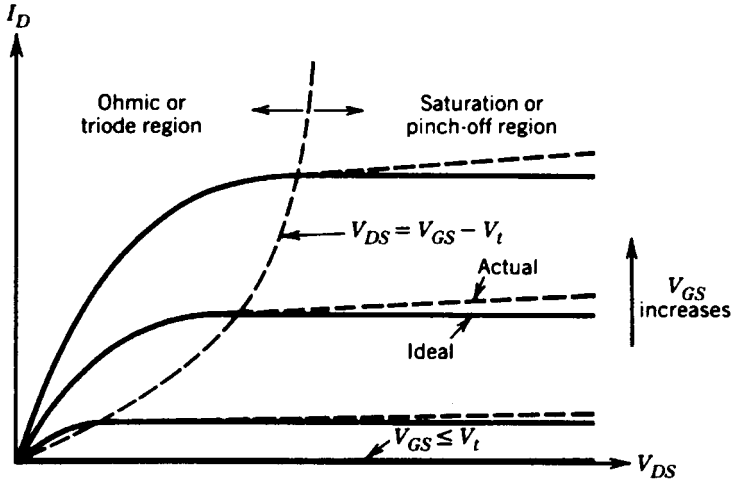


Figure 2-106 NMOS device characteristics.

exerts an appreciable influence on the channel. This causes I_D to rise with increasing V_{DS} in a similar fashion to the bipolar transistor curves of Figure 2-89.

For long channel lengths, the drain-depletion region has little effect on the channel and the I_D-versus-V_{DS} curves follow closely the ideal curves of Figure 2-106. Eventually the drain–substrate *pn*-junction breakdown voltage is exceeded and a sharp breakdown characteristic is obtained, which is similar to the JFET characteristic.

In addition to V_{DS} limitations, MOS devices must also be protected against excessive gate voltages. Typical gate oxides break down with about 25–50 V applied from gate to channel, and this process is destructive to the transistor.

2-3-4 Small-Signal Model of the MOS Transistor in Saturation

The preceding large-signal equations can now be used to derive the small-signal model of the MOS transistor in the saturation or pinchoff region. The important equations are (2-220) and (2-206). Note from Eq. (2-206) that the source–substrate voltage V_{BS} affects V_t, and thus I_D. This is due to the influence of the substrate acting as a second gate and is called *body effect*. As a consequence, I_D is a function of both V_{GS} and V_{BS}, and we require *two* transconductance generators in the small-signal model as shown in Figure 2-107. Variations in the voltage v_{bs} from source to body cause current $g_{mb}v_{bs}$ to flow from drain to source. Note that the body (or substrate) of an NMOS integrated circuit is usually connected to the most negative supply voltage and is thus an ac ground. However, the source connection can have a significant ac voltage impressed on it. Parasitic resistances due to the channel contact regions should be included in series with the source and drain of the model but are usually neglected in hand calculations. These resistances have an inverse dependence on channel width W and have typical values of 50–100 Ω for devices with W of about 1 μm.

The parameters of Figure 2-107 can be determined from Eq. (2-223) by differentiating.

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = k' \frac{W}{L} (V_{GS} - V_t)(1 + \lambda V_{DS}) \tag{2-224}$$

If $\lambda V_{DS} << 1$, this is often approximated as

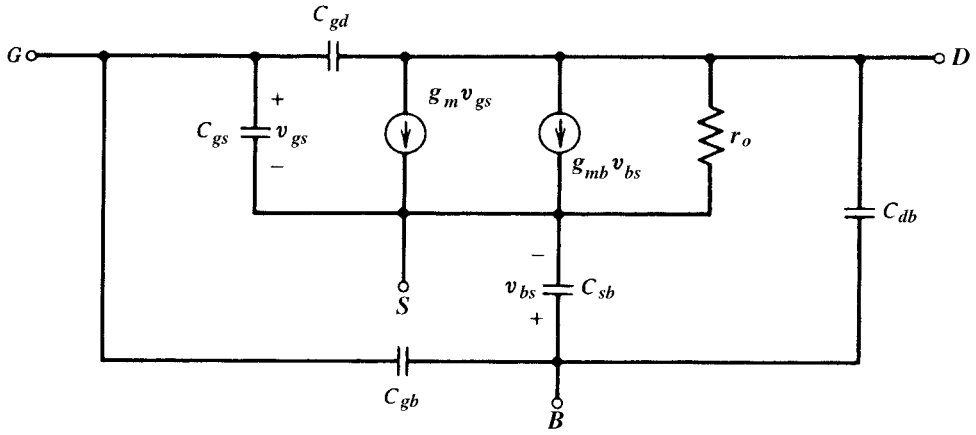


Figure 2-107 Small-signal MOS transistor equivalent circuit.

$$g_m = k' \frac{W}{L} (V_{GS} - V_t) = \sqrt{2k' \frac{W}{L} I_D} \tag{2-225}$$

Like the JFET and unlike the bipolar transistor, the transconductance of the MOSFET depends on both bias current and the *W*/*L* ratio (also on the oxide thickness via *k*′). Similarly,

$$g_{mb} = \frac{\partial I_D}{\partial V_{BS}} = -k' \frac{W}{L} (V_{GS} - V_t) (1 + \lambda V_{DS}) \frac{\partial V_t}{\partial V_{BS}} \tag{2-226}$$

From Eq. (2-206)

$$\frac{\partial V_t}{\partial V_{BS}} = -\frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} = -\chi \tag{2-227}$$

where this equation defines a factor χ. This parameter is equal to the rate of change of threshold voltage with body bias voltage. One obtains

$$\chi = \frac{C_{js}}{C_{\text{ox}}} \tag{2-228}$$

where C_{js} is the capacitance per unit area of the depletion region under the channel, assuming a one-sided step junction with a built-in potential $\psi_0 = 2\phi_f$.

Substitution of Eq. (2-227) into Eq. (2-226) gives

$$g_{mb} = \frac{\gamma k' (W/L) (V_{GS} - V_t) (1 + \lambda V_{DS})}{2\sqrt{2\phi_f + V_{SB}}} \tag{2-229}$$

Again, if $\lambda V_{DS} << 1$, we have

$$g_{mb} = \frac{\gamma \sqrt{k' (W/L) I_D}}{\sqrt{2(2\phi_f + V_{SB})}} \tag{2-230}$$

An important quantity is the ratio g_{mb}/g_m, and from Eqs. (2-224) and (2-229) we find

$$\frac{g_{mb}}{g_m} = \frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} = \chi \tag{2-231}$$

The factor χ is typically in the range 0.1–0.3.

Finally, the small-signal output resistance can be obtained directly from Eq. (2-218):

$$r_o = \left(\frac{\partial I_D}{\partial V_{DS}} \right)^{-1} = \frac{L_{\text{eff}}}{I_D} \left(\frac{dX_d}{dV_{DS}} \right)^{-1} \tag{2-232}$$

and using Eqs. (2-221) and (2-222), we find

$$r_o = \frac{1}{\lambda I_D} = \frac{V_A}{I_D} \tag{2-233}$$

Small-signal model capacitances are also shown in Figure 2-107. Of these, only the gate–source capacitance C_{gs} is intrinsic to the device operation in the saturation region. Capacitances C_{sb} and C_{db} are parasitic depletion-region capacitances between the substrate and the source and drain regions, respectively. These can be expressed as follows:

$$C_{sb} = \frac{C_{sb0}}{(1 + V_{SB}/\psi_0)^{1/2}} \tag{2-234}$$

$$C_{db} = \frac{C_{db0}}{(1 + V_{DB}/\psi_0)^{1/2}} \tag{2-235}$$

These capacitances are proportional to the gate and source region areas (including sidewalls), and C_{sb} also includes depletion-region capacitance from the induced channel in the body.

Capacitance C_{gb} between gate and substrate models parasitic oxide capacitance between the gate contact material and the substrate outside the active-device area. This is a constant capacitance and models coupling between polysilicon and metal interconnects and the underlying substrate. In fact, parasitic capacitance of this type underlies all polysilicon and metal traces on the chip and should be taken into account when simulating and calculating high-frequency circuit and device performance. Typical values depend on oxide thicknesses and range from about 0.04 to 0.15 fF per square micron of interconnect, with fringing effects becoming important for narrow lines (several microns or less in width).

Capacitances C_{gs} and C_{gd} exist from gate to source and drain, respectively. If C_{ox} is the oxide capacitance per unit area from gate to channel then the total capacitance under the gate is $C_{\text{ox}}WL$. This capacitance is intrinsic to the device operation and models the gate control of the channel conductance. In the ohmic region of device operation, this capacitance is split equally between source and drain so that $C_{gs} = C_{gd} = 1/2C_{\text{ox}}WL$. However, in the saturation region, the channel is very narrow at the drain end and the drain voltage exerts little influence on either the channel or the gate charge. As a consequence, the intrinsic portion of C_{gd} is essentially zero in the saturation region and C_{gd} then consists of a constant parasitic oxide-capacitance contribution due to gate overlap of the drain region. This is on the order of 1–10 fF for small devices.

In order to calculate the corresponding value of C_{gs} in the saturation region, we must calculate the total charge Q_T stored in the channel. This can be obtained by integrating Eq. (2-209) to obtain

$$Q_T = WC_{\text{ox}} \int_0^L [V_{GS} - V(y) - V_t]dy \tag{2-236}$$

Substituting for dy/dV from Eq. (2-211) into Eq. (2-236), we find

$$Q_T = \frac{W^2 C_{\text{ox}}^2 \mu_n}{I_D} \int_0^{V_{GS}-V_t} (V_{GS} - V - V_t)^2 dV \tag{2-237}$$

where the limit $y = L$ corresponds to $V = (V_{GS} - V_t)$ in saturation. Solution of Eq. (2-237) and use of Eqs. (2-214) and (2-215) gives

$$Q_T = \frac{2}{3} WLC_{\text{ox}}(V_{GS} - V_t) \tag{2-238}$$

and thus

$$C_{gs} = \frac{\partial Q_T}{\partial V_{GS}} = \frac{2}{3} WLC_{\text{ox}} \tag{2-239}$$

In addition, there is a contribution to C_{gs} from the constant parasitic oxide capacitance due to gate overlap of the source region.

The f_T of the MOSFET is given by

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{gs} + C_{gd} + C_{gb}} \tag{2-240}$$

The dependence of MOSFET f_T on device and process parameters can be seen by assuming that the intrinsic device capacitance G_{gs} dominates. Thus from Eq. (2-240) we have

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{gs}} \tag{2-241}$$

Substituting in Eq. (2-241) for g_m from Eq. (2-225) and C_{gs} from Eq. (2-239) we find for a MOSFET

$$f_T = 1.5 \frac{\mu_n}{2\pi L^2} (V_{GS} - V_t) \tag{2-242}$$

It is interesting to compare this with the intrinsic f_T of a bipolar transistor when parasitic depletion-layer capacitance is neglected. From

$$f_T = \frac{1}{2\pi\tau_F} \tag{2-243}$$

and substituting for τ_f and using the Einstein relationship $D_n/\mu_n = kT/q = V_T$, we find for a bipolar transistor

$$f_T = 2 \frac{\mu_n}{2\pi W_B^2} V_T \tag{2-244}$$

The similarity in form between Eqs. (2-242) and (2-244) is striking. In both cases the intrinsic device f_T increases as the inverse square of the critical device dimension across which carriers are in transit. The voltage V_t = 26 mV is fixed for a bipolar transistor, but MOSFET f_T can be increased by operating at high values of $(V_{GS} - V_t)$. Note that the base width W_B in a bipolar transistor is a vertical dimension determined by diffusions or implants and can typically be made much smaller than the channel length L of a MOSFET, which depends on surface geometry and photolithographic processes. Thus bipolar transistors generally have higher f_T than MOSFETs made with comparable processing. Finally, Eq. (2-242) was derived assuming that the MOSFET square law is valid. However, as discussed in Section 2-3-5, submicron MOSFETs depart significantly from square-law characteristics, and we find that for such devices f_T increases as L^{-1} rather than L^{-2}.

2-3-5 Short-Channel Effects in FETs

The evolution of integrated-circuit processing techniques has led to continuing reductions in both the horizontal and vertical dimensions of the active devices (the minimum allowed dimension of passive devices has also increased). This trend is driven primarily by economics in that more devices and circuits can be processed at one time on a given wafer. A second benefit has been that the frequency capability of the active devices continues to increase, as intrinsic f_T values increase with smaller dimensions while parasitic capacitances decrease.

Vertical dimensions such as the base width of a bipolar transistor in production processes may now be on the order of 0.05 μm or less, whereas horizontal dimensions such as bipolar emitter width or FET gate length may be significantly less than 1 μm. Even at these very small dimensions, the large-signal and small-signal models of bipolar transistors given in previous sections remain valid. However, significant short-channel effects become important in FETs of all types at channel lengths on the order of 1 μm or less and require modifications to the FET models given previously. The primary effect is to modify the classical FET square-law transfer characteristic in the saturation region to make the device more closely approach an ideal linear transfer function. Note, however, that even in processes with submicron capability, many of the FETs in a given analog circuit may be deliberately chosen to be larger than the minimum size and may be well approximated by the square-law model.

The most important short-channel effect in FETs is due to velocity saturation of carriers in the channel [17]. At low electric field values, the linear relation between carrier velocity and field implied by Eqs. (2-210) and (2-211) is valid. At high fields, however, the carrier velocities approach the thermal velocities and subsequently the carrier velocities increase more slowly with increasing field. This is illustrated in Figure 2-108, which shows typical measured electron drift velocity v_d versus tangential electric field strength magnitude $\mathscr{E}$ in an NMOS surface channel. Note that at low field values the velocity is proportional to the field, while at high fields the velocity approaches a constant value called the scattering-limited velocity v_{scl}. A first-order analytical approximation to this curve is

$$v_d = \frac{\mu_n \mathscr{E}}{1 + \mathscr{E}/\mathscr{E}_c} \tag{2-245}$$

where $\mathscr{E} \simeq 1.5 \times 10^6$ V/m and $\mu_n \simeq 0.07$ m^2/V · s is the low-field mobility. Equation (2-245) is also plotted in Figure 2-108. From Eq. (2-245), as $\mathscr{E} \to \infty$ we have $v_{\text{scl}} = \mu_n \mathscr{E}_c$. At the critical field value $\mathscr{E}_c$, the carrier velocity is a factor of 2 less than the low-field formula would predict. In a device with a channel length $L = 1$ μm, we need a voltage drop of only 1.5 V along the channel to have an average field equal to $\mathscr{E}_c$, and this condition is readily achieved in small MOSFETs. Similar results are found for PMOS devices.

As an example of the effects of velocity saturation on FET characteristics, we consider the example of the MOSFET. In the analysis of "Transfer Characteristics of MOS Devices" in Section 2-3-3, we now use the more general expression

$$I_D = WQ_I(y)v_d(y) \tag{2-246}$$

Substituting Eq. (2-245) into Eq. (2-246) and using

$$\mathscr{E} = \frac{dV}{dy} \tag{2-247}$$

for the magnitude of the field, we find that

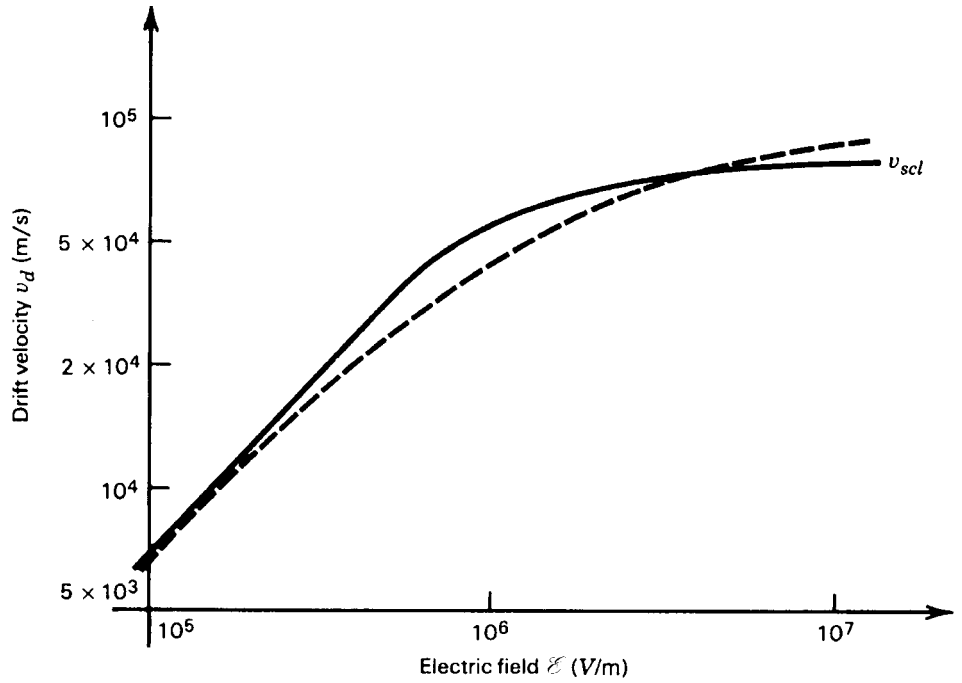


Figure 2-108 Typical measured electron drift velocity v_d versus tangent electric field $\mathscr{E}$ in an MOS surface channel (solid line). Also shown (dotted line) is the analytical approximation of Eq. (2.245) with $\mathscr{E}_c = 1.5 \times 10^6$ V/m and $\mu_n = 0.07$ m^2/V · s.

$$I_D\left(1 + \frac{1}{\mathscr{E}_c}\frac{dV}{dy}\right) = WQ_I(y)\mu_n\frac{dV}{dy} \tag{2-248}$$

Note that as $\mathscr{E}_c \to \infty$ and velocity saturation becomes negligible, Eq. (2-248) approaches the original equation (2-211).

Integrating Eq. (2-248) along the channel we obtain

$$\int_0^L I_D\left(1 + \frac{1}{\mathscr{E}_c}\frac{dV}{dy}\right)dy = \int_0^{V_{DS}} WQ_I(y)\mu_n dV \tag{2-249}$$

and thus

$$I_D = \frac{\mu_n C_{ox}}{2\left(1 + \dfrac{1}{\mathscr{E}_c}\dfrac{V_{DS}}{L}\right)}\frac{W}{L}[2(V_{GS} - V_t)V_{DS} - V_{DS}^2] \tag{2-250}$$

The quantity V_{DS}/L in Eq. (2-250) can be interpreted as the average field in the channel. If this is comparable to $\mathscr{E}_c$, the drain current for a given V_{DS} is less than the simple expression (2-213) would predict.

Equation (2-250) is valid in the triode region. The MOSFET transfer function in saturation can be obtained by using $V_{DS} = (V_{GS} - V_t)$ in Eq. (2-250) to obtain

$$I_D = \frac{k'}{2[1 + \theta(V_{GS} - V_t)]} \frac{W}{L} (V_{GS} - V_t)^2 \tag{2-251}$$

where $\theta = 1/L\mathscr{E}_c$ and has the dimension volts^{-1}. For $L = 1$ μm, a typical value is $\theta \approx 0.7\ \text{V}^{-1}$. Note that in the presence of velocity saturation effects, the device enters the saturation region for $V_{DS} < (V_{GS} - V_t)$. However, Eq. (2-251) still gives a good estimate of the saturation current.

Thus far, we have only considered the effects of the tangential field due to the V_{DS} along the channel when examining velocity saturation effects. However, there exists a normal field originating from the gate voltage that also inhibits channel carrier mobility. Since the normal field depends on the value of V_{GS}, we find that an empirical modification to θ in Eq. (2-251) can adequately model this effect. In practice, θ is determined by a best fit to measured device characteristics.

Returning to Eq. (2-251) we note that for very short channel lengths, θ becomes large and Eq. (2-251) reduces to

$$I_D \propto (V_{GS} - V_t) \tag{2-252}$$

Thus, the FET characteristics tend toward a *linear* transfer function as the channel length becomes very small (less than 1 μm).

Equation (2-251) has a simple circuit representation. Consider the circuit of Figure 2-109, where an ideal square-law MOSFET has a resistance R_{SX} in series with the source of the FET. Assume

$$I_D = \frac{\mu C_{\text{ox}}}{2} \frac{W}{L} (V'_{GS} - V_t)^2 \tag{2-253}$$

Now

$$V_{GS} = V'_{GS} + I_D R_{SX} \tag{2-254}$$

and substituting Eq. (2-254) into Eq. (2-253) we find that

$$I_D = \frac{\mu C_{\text{ox}}}{2} \frac{W}{L} (V_{GS} - I_D R_{SX} - V_t)^2 \tag{2-255}$$

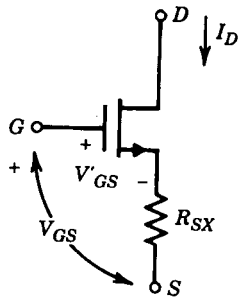


Figure 2-109 Model of velocity saturation of a MOSFET by addition of a series source resistance to an ideal square-law device.

Rearranging Eq. (2-255) we find

$$I_D = \frac{\mu C_{\text{ox}}}{2\left[1 + \mu C_{\text{ox}} \frac{W}{L} R_{SX}(V_{GS} - V_t)\right]} \frac{W}{L} (V_{GS} - V_t)^2 \tag{2-256}$$

This has exactly the same form as Eq. (2-251) if we identify

$$\theta = \mu C_{\text{ox}} \frac{W}{L} R_{SX} \tag{2-257}$$

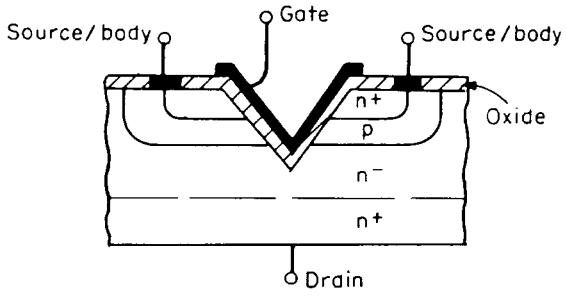
Substituting $\theta = 1/L\mathscr{E}_c$ into Eq. (2-257) we have

$$R_{SX} = \frac{1}{\mathscr{E}_c} \frac{1}{\mu C_{\text{ox}}} \frac{1}{W} \tag{2-258}$$

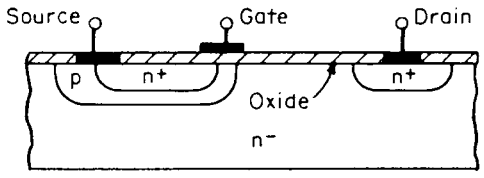
Thus, the influence of velocity saturation on the large-signal characteristics of a FET can be modeled to first order by a resistor R_{SX} in series with the source of an ideal square-law device. Note that R_{SX} varies inversely with W, as does the intrinsic physical series resistance due to source and drain contact regions. Typically, R_{SX} is larger than the physical series resistance. For $W = 2\ \mu\text{m}$, $\mu C_{\text{ox}} = 40\ \mu\text{A}/V^2$, and $\mathscr{E}_c = 1.5 \times 10^6$ V/m, we find $R_{SX} = 8\ \text{k}\Omega$.

The foregoing analysis has developed a modified large-signal mode for the MOSFET including velocity saturation effects. Small-signal MOSFET modeling for small devices can still be done using the equivalent circuit of Figure 2-107 if the values of g_m and g_{mb} are modified to account for the effects of velocity saturation using Eq. (2-251).

Figure 2-110 shows a cross section of VMOS (vertical MOS) and DMOS (double-diffused MOS) transistors. There is also a TMOS device—a Motorola variant of VMOS—and the



(a) VMOS (vertical channel MOS)



(b) DMOS (double-diffused MOS)

Figure 2-110 Cross section of (a) VMOS and (b) DMOS FETs.

now very popular LDMOS (lateral diffused MOS) power transistor (Figures 2-111 and 2-112). LDMOS FETs operate from a single supply (approximately 8 V and up) and are available for output powers surpassing 100 W at 2 GHz. The feedback capacitance is much lower than VMOS, TMOS, and earlier models, significantly improving circuit stability.

Following is a listing of Motorola high-power LDMOS transistors capable of operation up to 2 GHz:

RF High-Power LDMOS Transistors by Motorola
2.0 GHz, V_{DD} = 26 V, Class AB

Device	P_{out} Output Power (watts)	P_{in} Input Power Typical (watts)	G_{ps} (Typ.)/Freq. (dB/MHz)	IMDs (–dBc)	η Efficiency, Typ. %	θ_{JC} (°C/W)	Package/Style
MRF281S	4	0.2	13.6/2000	–30	40	8.75	458/1
MRF281Z	4	0.2	13.6/2000	–30	40	8.75	458A/1
MRF6525-5	5	0.4	11.5/2000	—	40	15	458A/1
MRF6525-10	10	1.0	10/2000	—	40	6.0	458A/1
MRF282S	10	0.5	13/2000	–30	35	2.9	458/1
MRF282Z	10	0.5	13/2000	–30	35	2.9	458A/1
MRF284	30	2.4	11.5/2000	–30	33	2.0	360B/1
MRF284S	30	2.4	11.5/2000	–30	33	2.0	360C/1
MRF286	60	4.75	11.4/2000	–30	33	.73	465/1
MRF286S	60	4.75	11.4/2000	–30	33	.73	465A/1

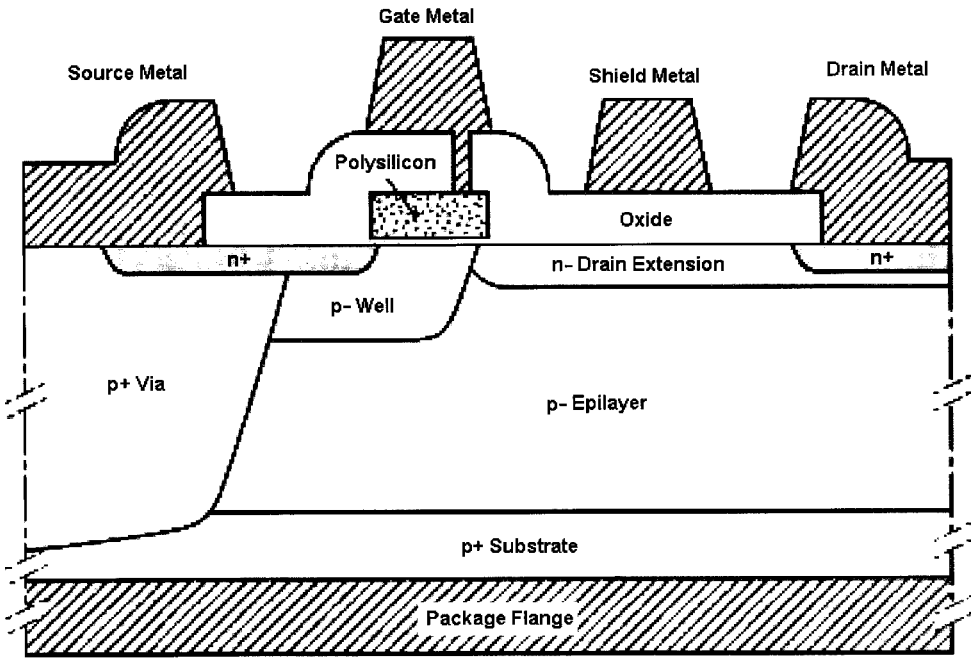


Figure 2-111 Cross section of a Philips LDMOS FET [18].

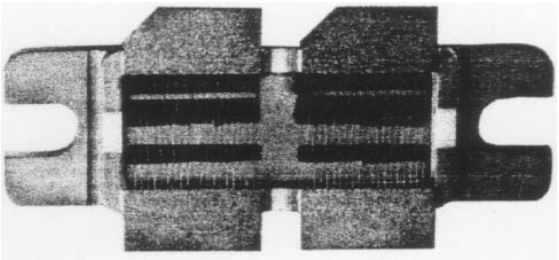


Figure 2-112 Photo of the underside of an LDMOS FET.

2-3-6 Small-Signal Models of MOSFETs

Subthreshold Conduction in MOSFETs. The MOSFET analysis of Section 2-3-3 focused on the normal region of operation where this is a well-defined conducting channel under the gate. Changes in the gate voltage are assumed to cause changes in the channel charge only, and not in the depletion region below. However, for gate voltages less than the extrapolated threshold voltage V_t, the applied gate potential affects both the depletion-region charge and the channel charge, which is then very small but not zero. The device can thus conduct finite (but small) current for $V_{GS} < V_t$, so that Eq. (2-215) is not valid in this region. The electrons in the n^+ source region of an NMOS transistor can overcome the potential barrier to the p-type substrate and enter the channel region. This process is very similar to the turn-on of a bipolar transistor, and in fact the MOSFET characteristics in this subthreshold region (also called weak inversion) are very similar to those of a bipolar transistor. Analysis [19] shows that in the subthreshold region of the MOSFET characteristics can be defined by the equation

$$I_D = k_x \frac{W}{L} e^{V_{GS}/nV_T}(1 - e^{-V_{DS}/V_T}) \tag{2-259}$$

where k_x depends on process parameters and $n \simeq 1.5$. The value of n is not equal to unity in this case (as it is for the bipolar transistor) because the applied voltage V_{GS} appears partially at the silicon surface and partially across the depletion layer.

To illustrate this effect, we show measured NMOS characteristics plotted on three different scales in Figure 2-113. In Figure 2-113a we show the transfer characteristic in the forward-active region plotted on linear scales. For this device W = 20 μm and L = 20 μm and short-channel effects are negligible. The same data are plotted in Figure 2-113b as $\sqrt{I_D}$ versus V_{GS}. The resulting straight line shows that the device characteristic is close to an ideal square law. Plots like this are commonly used to obtain V_t by extrapolation (0.7 V in this case) and also k' from the slope of the curve (54 μA/V^2 in this case). Note that near the threshold voltage, the curve deviates from the straight line representing the square law. This is the subthreshold region. The data are plotted a third time in Figure 2-113c on log–linear scales. The straight line obtained for $V_{GS} < V_t$ fits Eq. (2-259) with n = 1.5. At currents below 10^{-12} A, the effect of leakage currents becomes evident.

The major application of subthreshold operation is in very low power applications at relatively low signal frequencies. The limitation to low signal frequencies occurs because the MOSFET f_T becomes very small. Since the device capacitances at very low bias currents

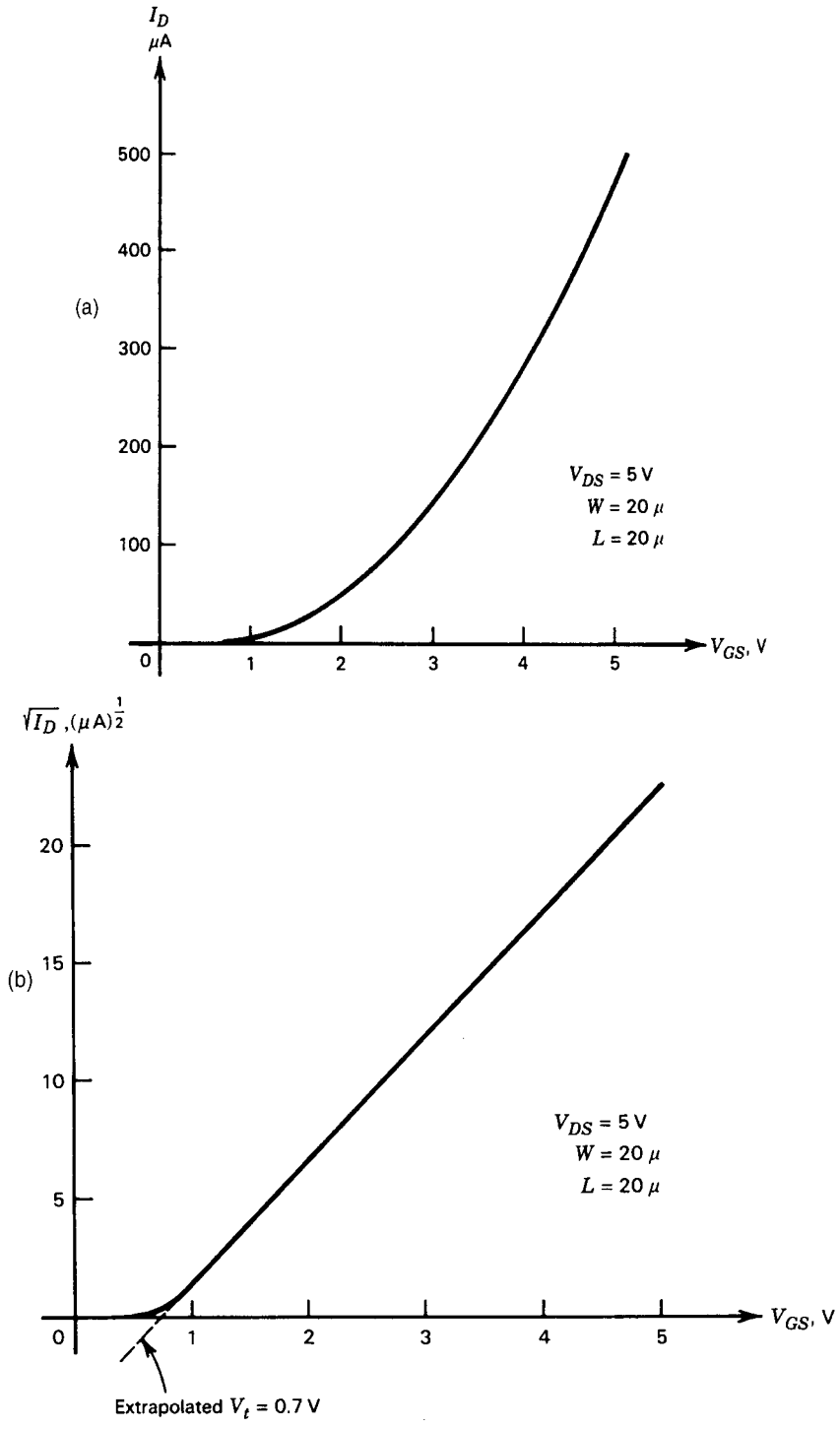


Figure 2-113 (a) Measured NMOS transfer characteristic in the forward-active region. (b) Data from part (a) plotted as $\sqrt{I_D}$ versus V_{GS} showing the square-law characteristic. (c) Data from part (a) plotted on log–linear scales showing the exponential characteristic of the subthreshold region.

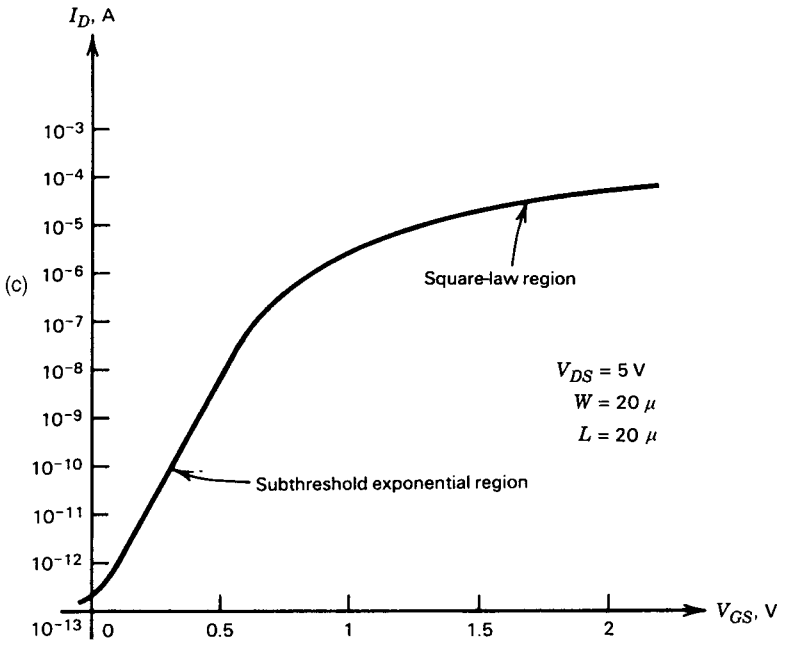


Figure 2-113. (*Continued*)

are essentially fixed, and the small-signal g_m calculated from Eq. (2-259) becomes proportional to I_D, we see that the value of f_T becomes very small at very low values of I_D.

Substrate Flow in MOSFETs. As the reverse-bias voltages on the device are increased, carriers traversing the depletion regions gain sufficient energy to create new hole–electron pairs in lattice collisions by a process known as *impact ionization.* Eventually, at high enough bias voltages, the process results in large avalanche currents. For collector–base bias voltages well below the breakdown value, a small enhanced current flow may occur across the collector–base junction due to this process with little apparent effect on the device characteristics.

Impact ionization also occurs in MOSFETs but has a significantly different effect on the device characteristics. This is because the channel electrons (for the case of NMOS) create hole–electron pairs in lattice collisions in the drain depletion region, and some of the resulting holes then flow to the substrate, creating a substrate current. (The electrons created

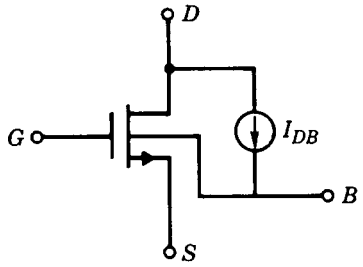


Figure 2-114 Representation of impact ionization in a MOSFET by a drain–substrate current generator.

in the process flow out the drain terminal.) The carriers created by impact ionization are thus not confined within the device as they are in a bipolar transistor. The effect of this phenomenon can be modeled by inclusion of a controlled current generator I_{DB} from drain to substrate, as shown in Figure 2-114 for an NMOS device. The magnitude of this substrate current depends on the voltage across the drain depletion region (which determines the energy of the ionizing channel electrons), and also on the drain current (which is the rate at which channel electrons enter the depletion region). It has been found empirically [20] that the current I_{DB} can be expressed as

$$I_{DB} = K_1(V_{DS} - V_{DS,\text{sat}})I_D e^{-[K_2/(V_{DS}-V_{DS,\text{sat}})]} \tag{2-260}$$

where K_1 and K_2 are process-dependent parameters and $V_{DS,\text{sat}}$ is the value of V_{DS} where the drain characteristics enter the saturation region. Typical values for NMOS devices are $K_1 = 5\ \text{V}^{-1}$ and $K_2 = 30$ V. The effect is generally much less significant in PMOS devices because the holes carrying the charge in the channel are much less efficient in creating hole–electron pairs than are energetic electrons.

The major impact of this phenomenon on circuit performance is that a parasitic resistance from drain to substrate now exists. Since the substrate of an NMOS device in a *p*-substrate process is an ac ground (the common substrate terminal must always be connected to the most negative supply voltage in the circuit), the parasitic resistance shunts the drain to ground and can be a limiting factor in many circuit designs. Differentiating Eq. (2-260) we find for the drain–substrate small-signal conductance

$$g_{db} = \frac{\partial I_{DB}}{\partial V_D}$$

$$= K_2 \frac{I_{DB}}{(V_{DS} - V_{DS,\text{sat}})^2} \tag{2-261}$$

where the gate and source are assumed held at fixed potentials.

In IC applications for RF/UHF circuits, many not-yet-solved problems remain:

1. The flicker corner frequency for MOS devices is somewhere between 100 kHz and 1 MHz.
2. Because of backplane problems in ICs, such as the substrate capacitance, also relative to the source, the performance of an IC-MOS/CMOS is not as good as expected if these effects are not considered.
3. Discrete devices, such as *n*-channel MOS devices, which are also available as dual-gate MOSFETs, can work well up to 1200 MHz. Such performance is not yet shown in silicon-based RFICs.
4. For RF power devices, such as LDMOS and others mentioned, promising results are only shown for discrete devices. There are no equivalent ICs on the market yet.

It will be interesting to see how this technology will work itself upward in frequency. Its main advantage is the BiCMOS process, which allows the mixing of analog and digital circuits nicely. Heterojunction transistors, such as SiGe, are inherently much more linear

and, for the same performance, require less current than MOS technology. On the other hand, we have not seen large-scale production SiGe-based ICs.

In the following five pages, we reproduce the datasheet for the Siemens (now Infineon Technologies) BF999 single-gate MOSFET.* Above 500 MHz, a dual-gate (also referred to as tetrode) MOSFET, such as the BF998, is preferred over a single-gate device. The main reasons for this are:

1. Internally, it is equivalent to two single-gate devices connected in a cascade arrangement, with the output device operating in the grounded-gate configuration, which entirely avoids the Miller effect and therefore remains stable at higher frequencies, operating up to 1.2 GHz.
2. The second gate can be used for AGC. As the Gate 2 voltage changes, the gain varies heavily.
3. By applying a local oscillator signal to Gate 2, this configuration can be used as a mixer.

The major drawbacks of these devices are their wide tolerances and temperature sensitivity. They follow to some degree the CMOS design but seem to have better performance.

The BF995, also called a tetrode, is in reality a dual-gate MOSFET specified to only 300 MHz; however, we are reproducing its datasheet* because the one more applicable for wireless circuits, the BF1005, does not have such a complete set of data.

To demonstrate the capabilities of low-voltage LDMOS transistors, we show the MRF5003 datasheet.† This device is intended for low-voltage applications down to 7.5 V. To the best of our knowledge, this is the lowest-voltage LDMOS transistor on the market. Other members of this family, which produce 5 W (MRF5005) and 15 W (MRF5015), begin to require higher supply voltages. Just to make sure it is not forgotten, the advantage of LDMOS is, besides lower-voltage operation, much-reduced feedback capacitance. LDMOS devices are now available that operate above 1 GHz and produce 60 W. There is now some competition with the recently introduced NEC GaAsFET, which goes up to 100 W in a similar frequency range.

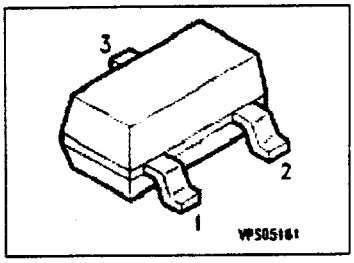
*Reproduced with permission.
†Copyright of Motorola, used by permission.

SIEMENS

Silicon N Channel MOSFET Triode

BF 999

- For high-frequency stages up to 300 MHz, preferably in FM applications



Type	Marking	Ordering Code (tape and reel)	Pin Configuration 1	2	3	Package
BF 999	LB	Q62702-F1132	G	D	S	SOT-23

Maximum Ratings

Parameter	Symbol	Values	Unit
Drain-source voltage	V_{DS}	20	V
Drain current	I_D	30	mA
Gate-source peak current	$\pm I_{GSM}$	10	
Total power dissipation, $T_A \leq 60$ °C	P_{tot}	200	mW
Storage temperature range	T_{stg}	– 55 ... + 150	°C
Channel temperature	T_{ch}	150	

Thermal Resistance

Junction - ambient [1]	$R_{th\,JA}$	≤ 450	K/W

[1] Package mounted on alumina 15 mm × 16.7 mm × 0.7 mm.

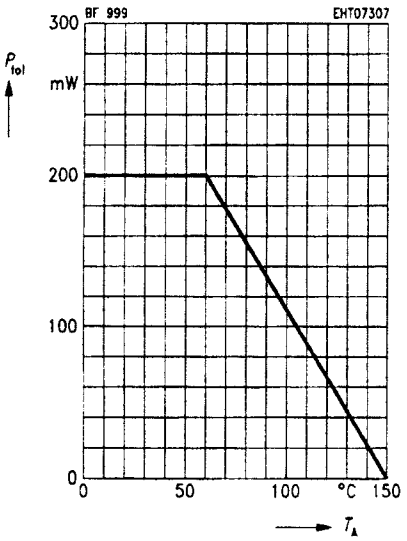
SIEMENS **BF 999**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

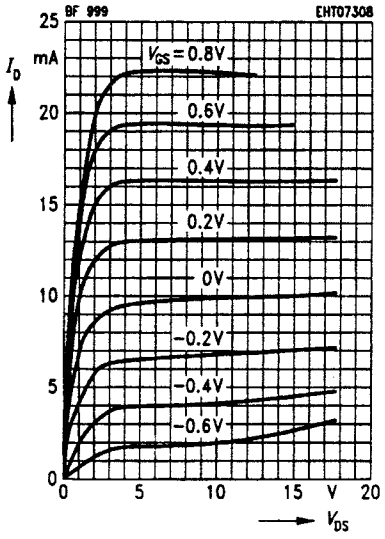
Parameter	Symbol	Values			Unit
		min.	typ.	max.	
DC Characteristics					
Drain-source breakdown voltage I_D = 10 µA, – V_{GS} = 4 V	$V_{(BR)\,DS}$	20	–	–	V
Gate-source breakdown voltage ± I_{GS} = 10 mA, V_{DS} = 0	± $V_{(BR)\,GSS}$	6.5	–	12	
Gate-source leakage current ± V_{GS} = 5 V, V_{DS} = 0	± I_{GSS}	–	–	50	nA
Drain current V_{DS} = 10 V, V_{GS} = 0	I_{DSS}	5	–	18	mA
Gate-source pinch-off voltage V_{DS} = 10 V, I_D = 20 µA	– $V_{GS\,(p)}$	–	–	2.5	V
AC Characteristics					
Forward transconductance V_{DS} = 10 V, I_D = 10 mA, f = 1 kHz	g_{fs}	14	16	–	mS
Gate input capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{gss}	–	2.5	–	pF
Reverse transfer capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{dg}	–	25	–	fF
Output capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{dss}	–	1	–	pF
Power gain (test circuit) V_{DS} = 10 V, I_D = 10 mA, f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS	G_p	–	25	–	dB
Noise figure (test circuit) V_{DS} = 10 V, I_D = 10 mA, f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS	F	–	1	–	

SIEMENS **BF 999**

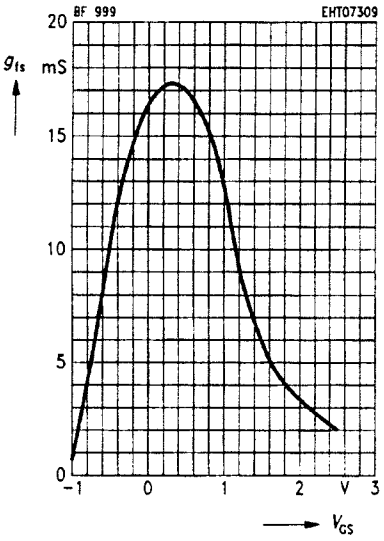
Total power dissipation $P_{tot} = f(T_A)$



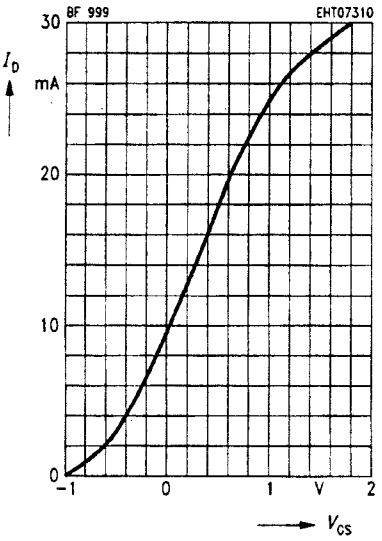
Output characteristics $I_D = f(V_{DS})$



Gate transconductance $g_{fs} = f(V_{GS})$
V_{DS} = 10 V, I_{DSS} = 10 mA, f = 1 kHz

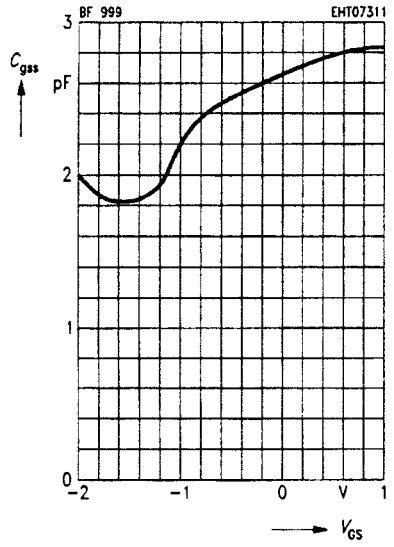


Drain current $I_D = f(V_{GS})$
V_{DS} = 10 V

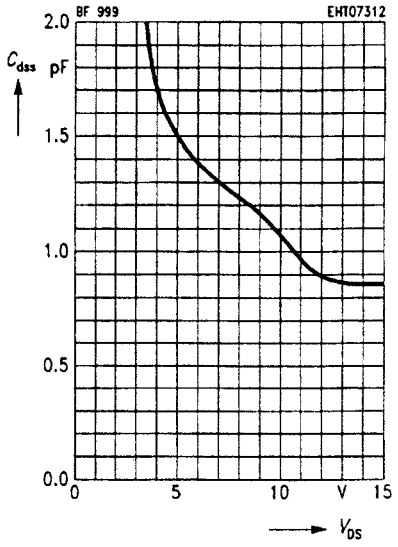


SIEMENS **BF 999**

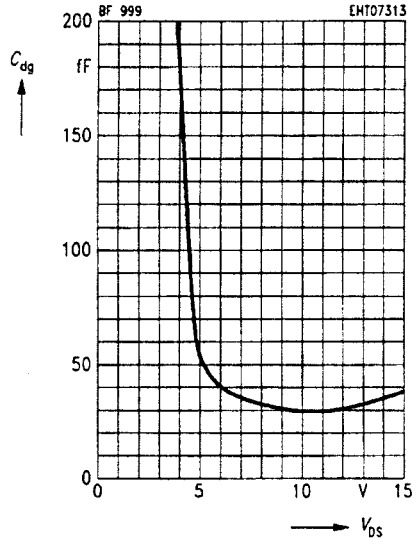
Gate input capacitance $C_{gss} = f(V_{GS})$
V_{DS} = 10 V, I_{DSS} = 10 mA, f = 1 MHz



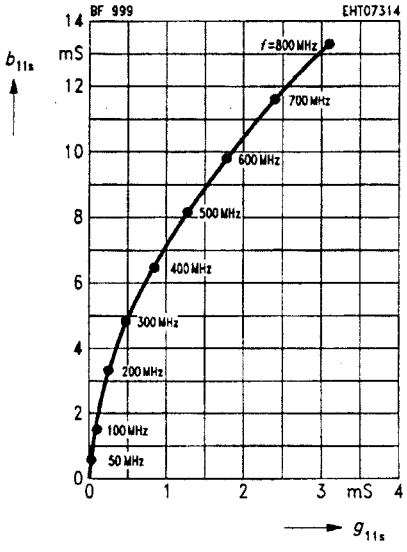
Output capacitance $C_{dss} = f(V_{DS})$
V_{GS} = 0, I_{DSS} = 10 mA, f = 1 MHz



Reverse transfer capacitance
$C_{dg} = f(V_{DS})$
I_{DSS} = 10 mA, f = 1 MHz, V_{GS} = 0

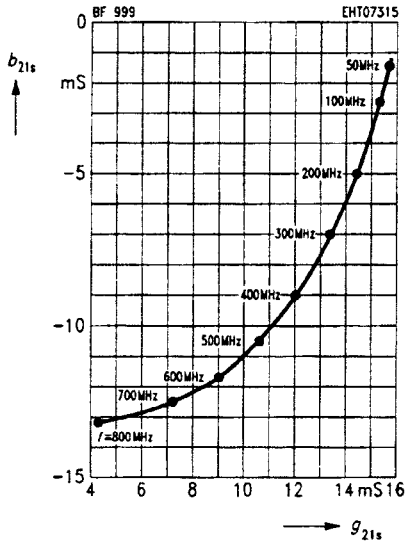


Gate input admittance y_{11s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)

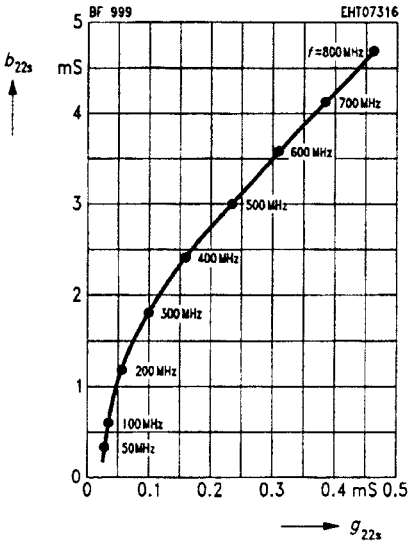


SIEMENS **BF 999**

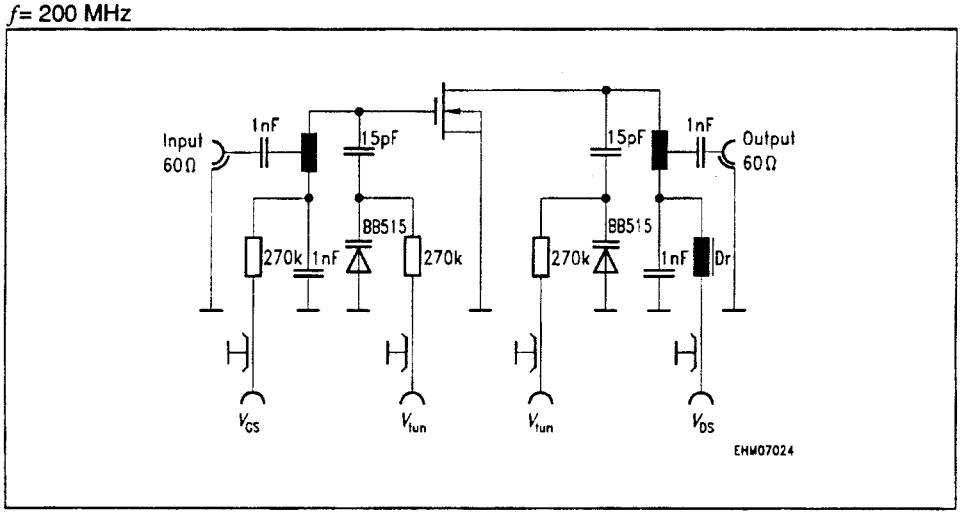
Gate forward transfer admittance y_{21s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)



Output admittance y_{22s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)



Test circuit for power gain and noise figure
f= 200 MHz

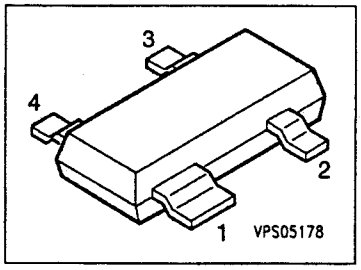


SIEMENS

Silicon N Channel MOSFET Tetrode

BF 995

- For input and mixer stages in FM and VHF TV tuners



Type	Marking	Ordering Code (tape and reel)	Pin Configuration 1	2	3	4	Package
BF 995	MB	Q62702-F936	S	D	G_2	G_1	SOT-143

Maximum Ratings

Parameter	Symbol	Values	Unit
Drain-source voltage	V_{DS}	20	V
Drain current	I_D	30	mA
Gate 1/gate 2 peak source current	$\pm I_{G1/2SM}$	10	
Total power dissipation, T_S < 76 °C	P_{tot}	200	mW
Storage temperature range	T_{stg}	– 55 ... + 150	°C
Channel temperature	T_{ch}	150	

Thermal Resistance

Junction - soldering point	$R_{th\,JS}$	< 370	K/W

SIEMENS **BF 995**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

Parameter	Symbol	Values			Unit
		min.	typ.	max.	
DC Characteristics					
Drain-source breakdown voltage I_D = 10 µA, – V_{G1S} = – V_{G2S} = 4 V	$V_{(BR)\,DS}$	20	–	–	V
Gate 1 source breakdown voltage ± I_{G1S} = 10 mA, V_{G2S} = V_{DS} = 0	± $V_{(BR)\,G1SS}$	8.5	–	14	
Gate 2 source breakdown voltage ± I_{G2S} = 10 mA, V_{G1S} = V_{DS} = 0	± $V_{(BR)\,G2SS}$	8.5	–	14	
Gate 1 source leakage current ± V_{G1S} = 5 V, V_{G2S} = V_{DS} = 0	± I_{G1SS}	–	–	50	nA
Gate 2 source leakage current ± V_{G2S} = 5 V, V_{G1S} = V_{DS} = 0	± I_{G2SS}	–	–	50	
Drain current V_{DS} = 15 V, V_{G1S} = 0, V_{G2S} = 4 V	I_{DSS}	4	–	20	mA
Gate 1 source pinch-off voltage V_{DS} = 15 V, V_{G2S} = 4 V, I_D = 20 µA	– $V_{G1S\,(p)}$	–	–	2.5	V
Gate 2 source pinch-off voltage V_{DS} = 15 V, V_{G1S} = 0, I_D = 20 µA	– $V_{G2S\,(p)}$	–	–	2.0	

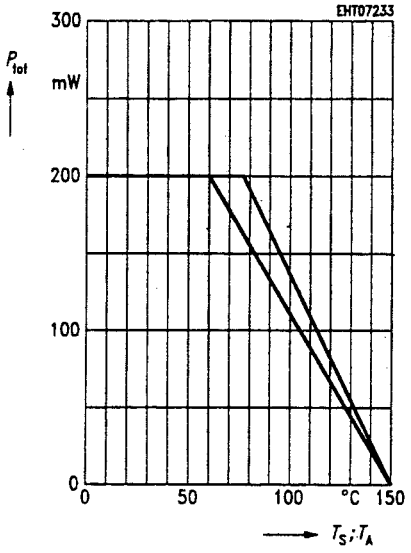
SIEMENS **BF 995**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

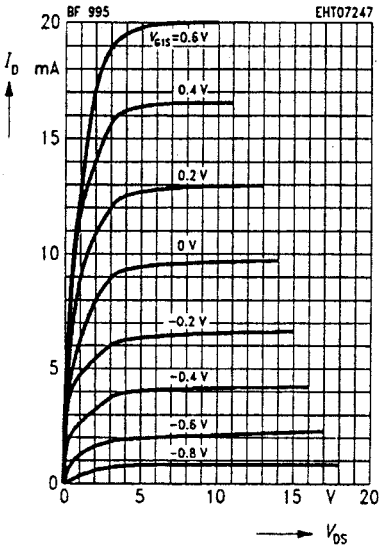
Parameter	Symbol	Values			Unit
		min.	typ.	max.	
AC Characteristics					
Forward transconductance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 kHz	g_{fs}	12	17	–	mS
Gate 1 input capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{g1ss}	–	3.6	–	pF
Gate 2 input capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{g2ss}	–	1.6	–	
Feedback capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{dg1}	–	25	–	fF
Output capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{dss}	–	1.6	–	pF
Power gain V_{DS} = 15 V, I_D = 10 mA f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS 2 Δf = 12 MHz (see test circuit 1)	G_{ps}	–	23	–	dB
Noise figure V_{DS} = 15 V, I_D = 10 mA f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS (see test circuit 1)	F	–	1.1	–	
Gain control range V_{DS} = 15 V, V_{G2S} = 4 ... – 2 V, f = 200 MHz (see test circuit 1)	ΔG_{ps}	–	50	–	
Mixer gain (additive) V_{DS} = 15 V, V_{G2S} = 6 V, R_S = 220 Ω f = 200 MHz, f_F = 36 MHz 2 Δf_F = 5 MHz, V_{osc} = 0.5 V (see test circuit 2)	G_{psc}	–	16	–	
Mixer gain (multiplicative) V_{DS} = 15 V, V_{G1S} = 1.7 V, V_{G2S} = 2.5 V R_S = 220 Ω, f = 200 MHz, f_F = 36 MHz 2 Δf_F = 5 MHz, V_{osc} = 2 V (see test circuit 3)	G_{psc}	–	18	–	

SIEMENS **BF 995**

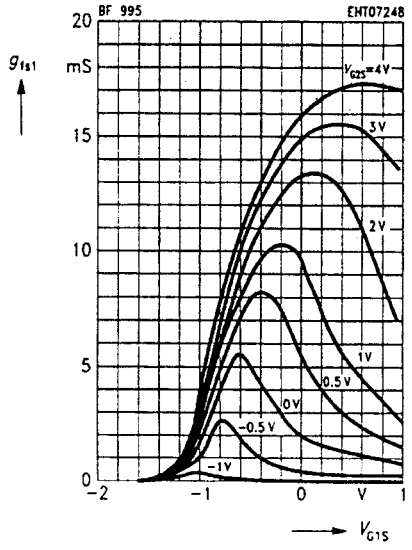
Total power dissipation $P_{tot} = f(T_A)$



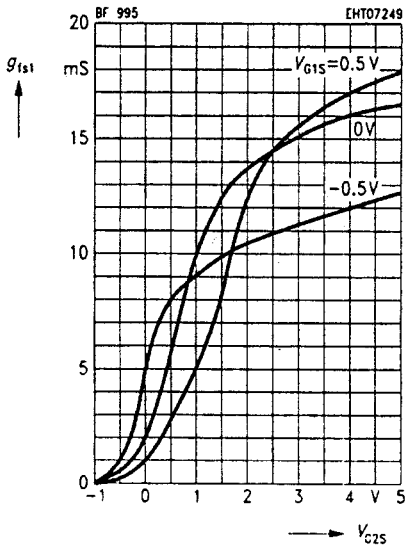
Output characteristics $I_D = f(V_{DS})$
V_{G2S} = 4 V



Gate 1 forward transconductance $g_{fs1} = f(V_{G1S})$
V_{DS} = 15 V, I_{DSS} = 10 mA, f = 1 kHz

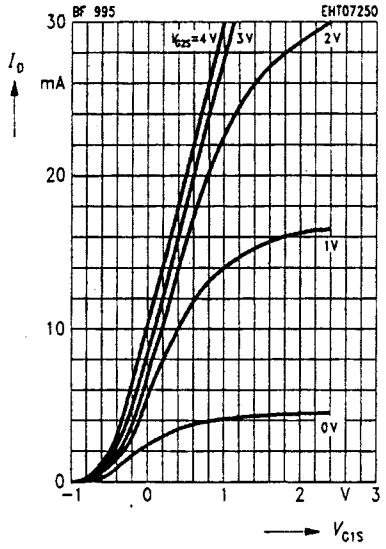


Gate 1 forward transconductance $g_{fs1} = f(V_{G2S})$
V_{DS} = 15 V, I_{DSS} = 10 mA, f = 1 kHz

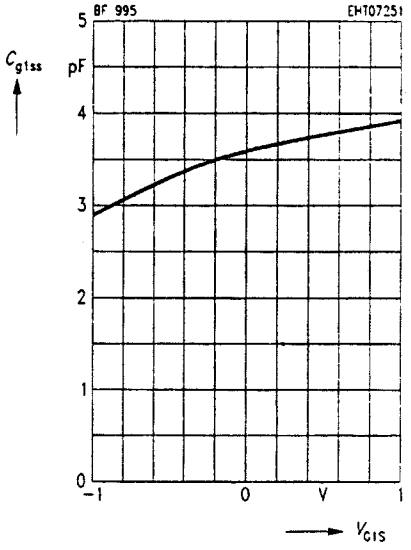


SIEMENS **BF 995**

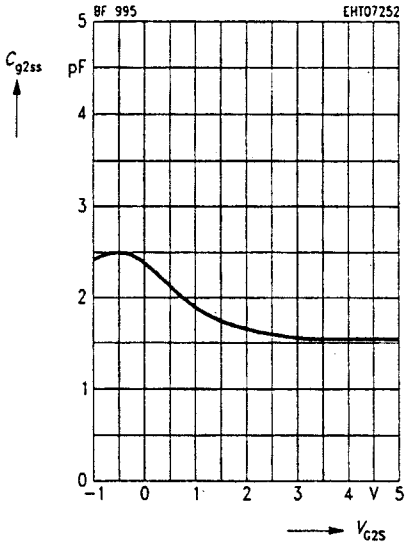
Drain current $I_D = f(V_{G1S})$
V_{DS} = 15 V



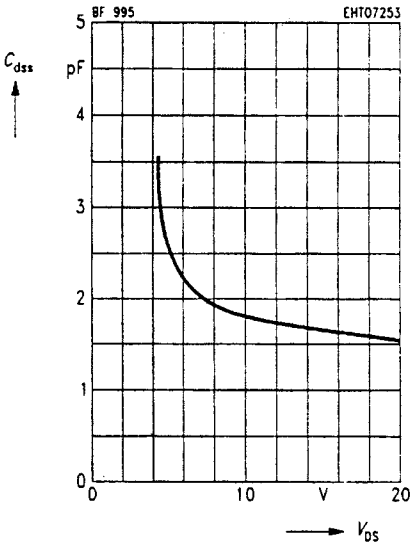
Gate 1 input capacitance $C_{g1ss} = f(V_{G1S})$
V_{G2S} = 4 V, V_{DS} = 15 V
I_{DSS} = 10 mA, f = 1 MHz



Gate 2 input capacitance $C_{g2ss} = f(V_{G2S})$
V_{G1S} = 0 V, V_{DS} = 15 V
I_{DSS} = 10 mA, f = 1 MHz

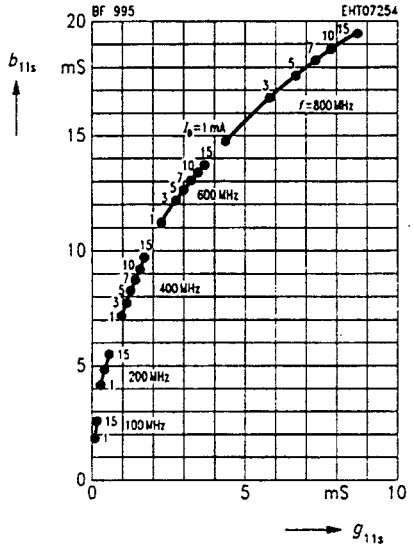


Output capacitance $C_{dss} = f(V_{DS})$
V_{G1S} = 0 V, V_{G2S} = 4 V
I_{DSS} = 10 mA, f = 1 MHz

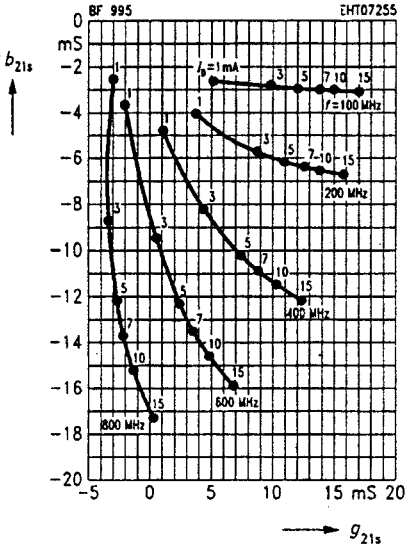


SIEMENS **BF 995**

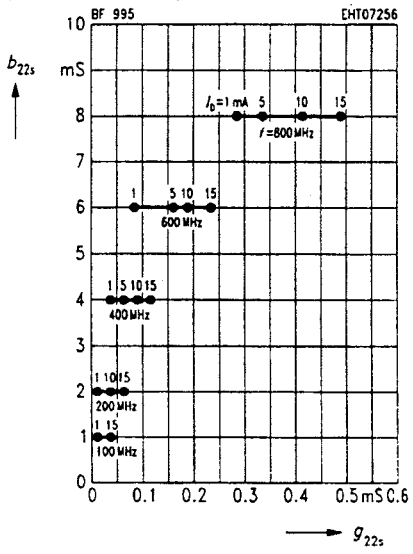
Gate 1 input admittance y_{11s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)



Gate 1 forward transfer admittance y_{21s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)

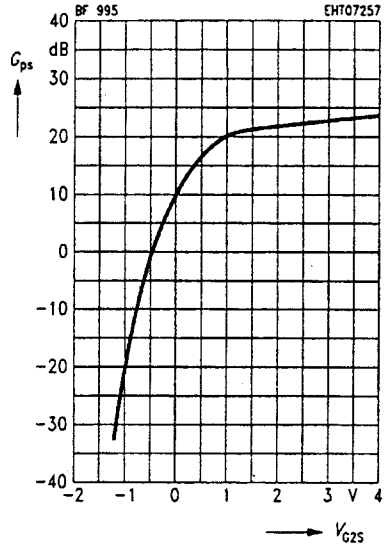


Output admittance y_{22s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)

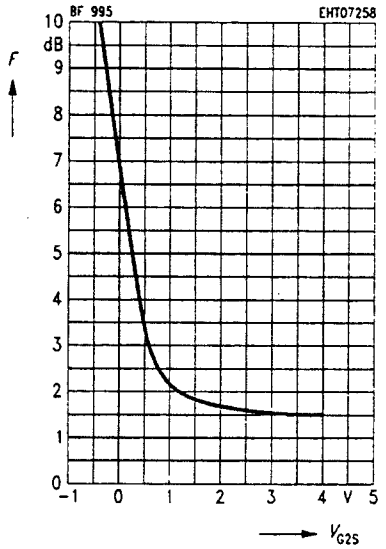


SIEMENS **BF 995**

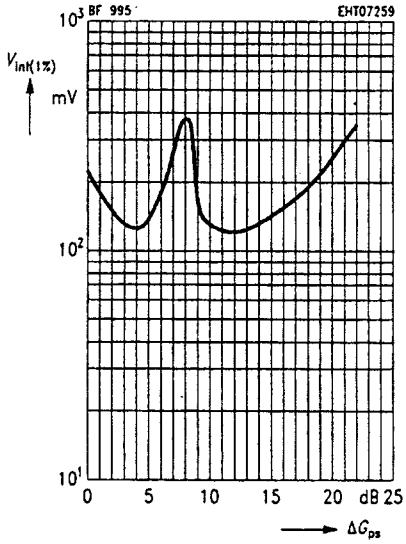
Power gain $G_{ps} = f(V_{G2S})$
V_{DS} = 15 V, V_{G1S} = 0 V, I_{DSS} = 10 mA
f = 200 MHz (see test circuit 1)



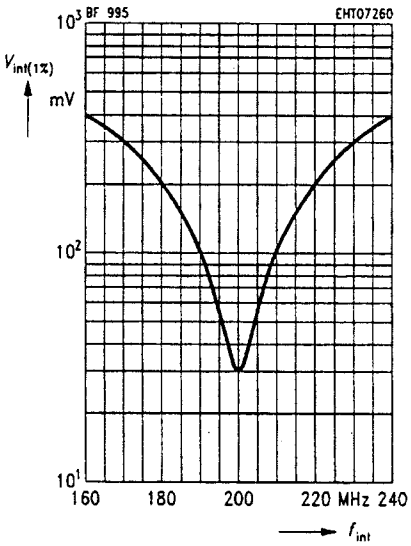
Noise figure $F = f(V_{G2S})$
V_{DS} = 15 V, V_{G1S} = 0 V, I_{DSS} = 10 mA
f = 200 MHz (see test circuit 1)



Interference voltage for 1% cross modulation $V_{int\,(1\%)} = f(\Delta G_{ps})$[1)]
V_{DS} = 15 V, V_{G1S} = 0, f = 200 MHz
f_{int} = 221 MHz (see test circuit 1)



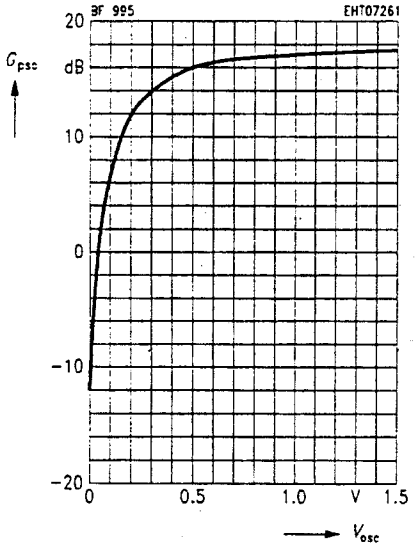
Interference voltage for 1% cross modulation $V_{int\,(1\%)} = f(f_{int})$[1)]
V_{DS} = 15 V, V_{G2S} = 4 V, V_{G1S} = 0
f = 200 MHz (see test circuit 1)



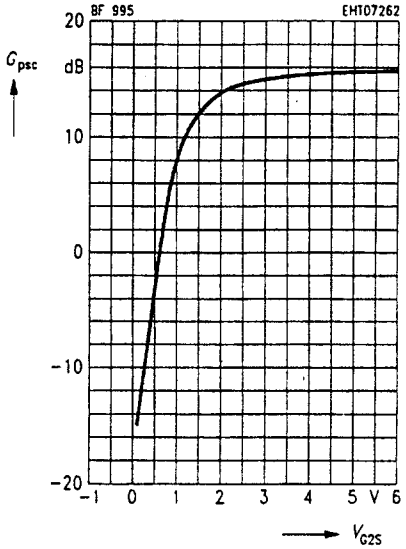
1) For footnote refer to the last page of this data sheet.

SIEMENS **BF 995**

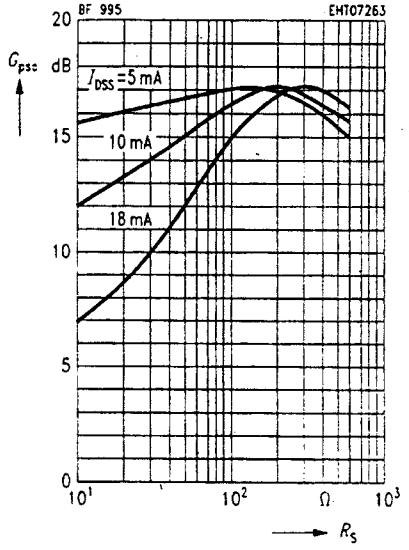
Mixer gain (additive) $G_{psc} = f(V_{osc})$
V_D = 15 V, V_{G1S} = 0, V_{G2S} = 6 V
R_S = 220 Ω, I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)



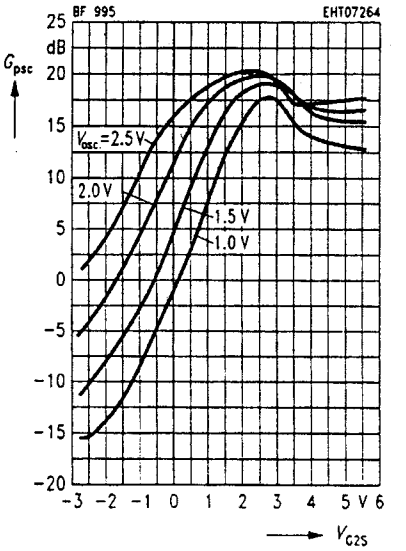
Mixer gain (additive) $G_{psc} = f(V_{G2S})$
V_D = 15 V, V_{G1S} = 0, R_S = 220 Ω
V_{osc} = 0.5 V, I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)



Mixer gain (additive) $G_{psc} = f(R_S)$
V_D = 15 V, V_{G1S} = 0, V_{G2S} = 6 V
V_{osc} = 0.5 V, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)

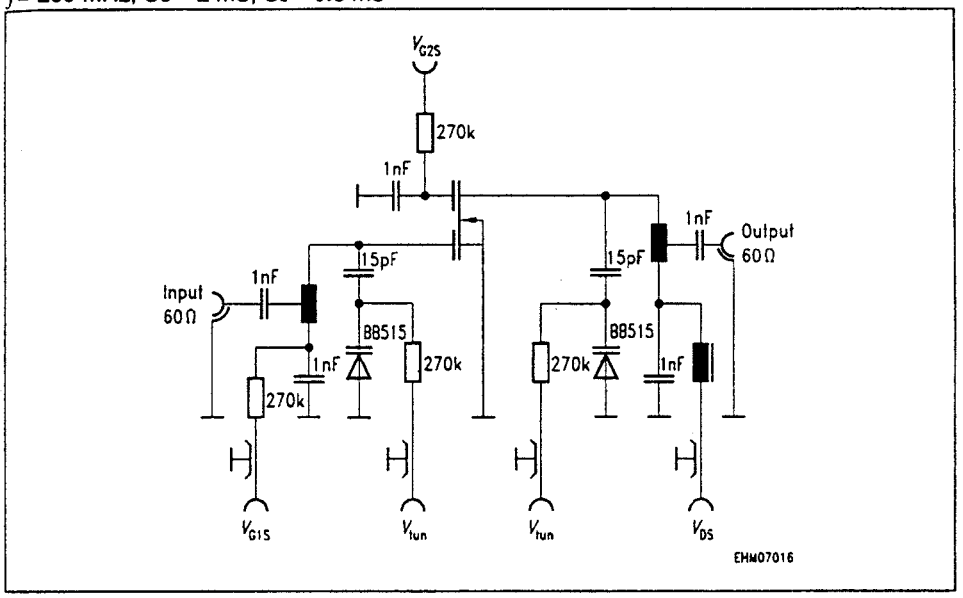


Mixer gain (multiplicative) $G_{psc} = f(V_{G2S})$
V_D = 15 V, V_{G1S} = 1.7 V, R_S = 200 Ω
I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 3)

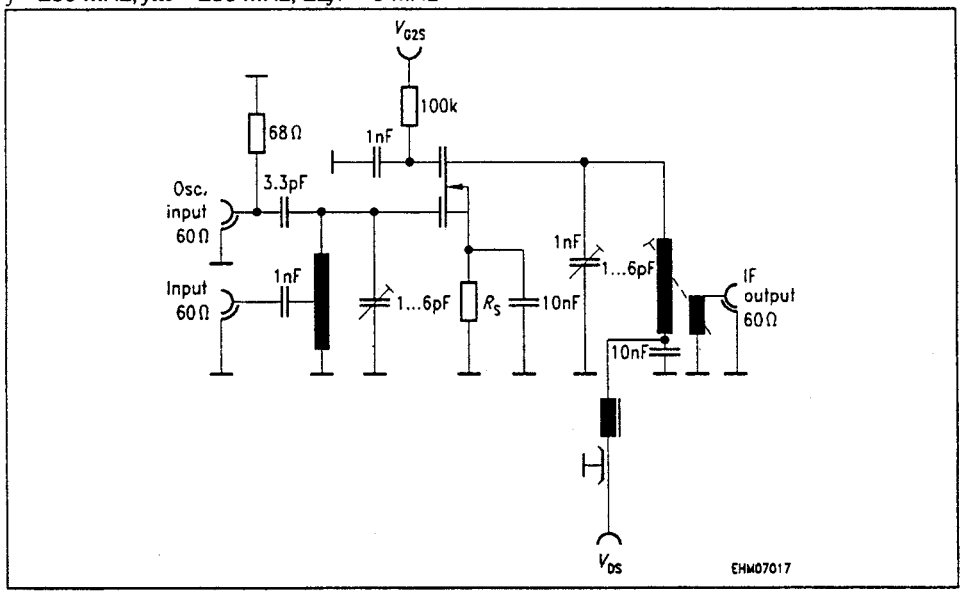


SIEMENS **BF 995**

Test circuit 1 for power gain, noise figure and cross modulation
f= 200 MHz, G_S = 2 mS, G_L = 0.5 mS

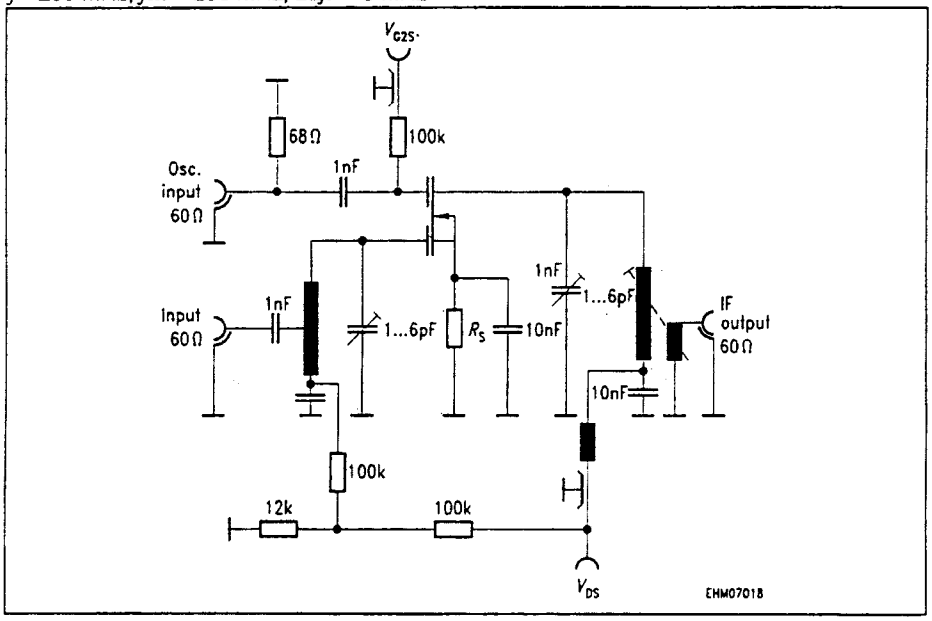


Test circuit 2 for mixer gain (additive)
f= 200 MHz, f_{osc} = 236 MHz, $2\Delta f_F$ = 5 MHz



SIEMENS **BF 995**

Test circuit 3 for mixer gain (multiplicative)
f = 200 MHz, f_{osc} = 236 MHz, $2\Delta f_F$ = 5 MHz



1) $V_{int\,(1\%)}$ is the rms value of half the emf (terminal voltage at matching) of a 100 % sine modulated TV carrier at an internal generator resistance of 60 Ω, causing 1 % amplitude modulation on the active carrier.

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF5003/D

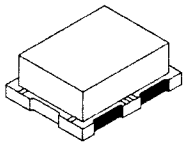
The RF MOSFET Line
RF Power Field Effect Transistor
N–Channel Enhancement–Mode

MRF5003

3.0 W, 7.5 V, 512 MHz
N–CHANNEL
BROADBAND
RF POWER FET

The MRF5003 is designed for broadband commercial and industrial applications at frequencies to 520 MHz. The high gain and broadband performance of this device makes it ideal for large–signal, common source amplifier applications in 7.5 Volt and 12.5 Volt mobile, portable, and base station FM equipment.

- Guaranteed Performance at 512 MHz, 7.5 Volts
  Output Power = 3.0 Watts
  Power Gain = 9.5 dB
  Efficiency = 45%
- Characterized with Series Equivalent Large–Signal Impedance Parameters
- S–Parameter Characterization at High Bias Levels
- Excellent Thermal Stability
- All Gold Metal for Ultra Reliability
- Capable of Handling 20:1 VSWR, @ 15.5 Vdc, 512 MHz, 2.0 dB Overdrive
- Suitable for 12.5 Volt Applications
- True Surface Mount Package
- Available in Tape and Reel by Adding R1 Suffix to Part Number.
  R1 Suffix = 500 Units per 16 mm, 7 inch Reel.
- Circuit board photomaster available upon request by contacting
  RF Tactical Marketing in Phoenix, AZ.



CASE 430–01, STYLE 2

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Drain–Source Voltage	V_{DSS}	36	Vdc
Drain–Gate Voltage (R_{GS} = 1.0 Meg Ohm)	V_{DGR}	36	Vdc
Gate–Source Voltage	V_{GS}	±20	Vdc
Drain Current — Continuous	I_D	1.7	Adc
Total Device Dissipation @ T_C = 25°C Derate above 25°C	P_D	12.5 0.07	Watts W/°C
Storage Temperature Range	T_{stg}	–65 to +150	°C
Operating Junction Temperature	T_J	200	°C

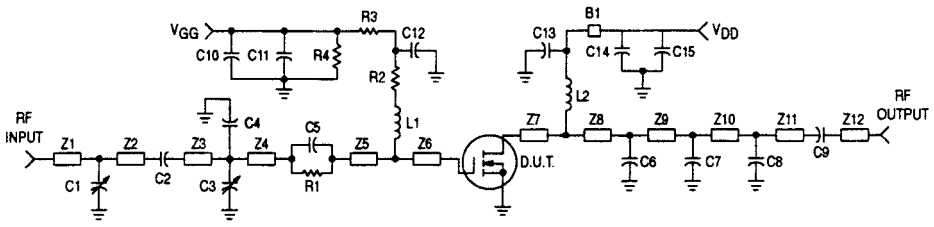
THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	14	°C/W

NOTE – **CAUTION** – MOS devices are susceptible to damage from electrostatic charge. Reasonable precautions in handling and packaging MOS devices should be observed.

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted.)

Characteristic		Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS						
Drain–Source Breakdown Voltage (V_{GS} = 0, I_D = 2.5 mAdc)		$V_{(BR)DSS}$	36	—	—	Vdc
Zero Gate Voltage Drain Current (V_{DS} = 15 Vdc, V_{GS} = 0)		I_{DSS}	—	—	1.0	mAdc
Gate–Source Leakage Current (V_{GS} = 20 Vdc, V_{DS} = 0)		I_{GSS}	—	—	1.0	μAdc
ON CHARACTERISTICS						
Gate Threshold Voltage (V_{DS} = 10 Vdc, I_D = 5.0 mAdc)		$V_{GS(th)}$	1.25	2.25	3.5	Vdc
Drain–Source On–Voltage (V_{GS} = 10 Vdc, I_D = 0.5 Adc)		$V_{DS(on)}$	—	—	0.375	Vdc
Forward Transconductance (V_{DS} = 10 Vdc, I_D = 0.5 Adc)		g_{fs}	0.6	—	—	mho
DYNAMIC CHARACTERISTICS						
Input Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{iss}	—	16.5	—	pF
Output Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{oss}	—	37	—	pF
Reverse Transfer Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{rss}	3.5	4.4	5.4	pF
FUNCTIONAL TESTS (In Motorola Test Fixture)						
Common–Source Amplifier Power Gain (V_{DD} = 7.5 Vdc, P_{out} = 3.0 W, I_{DQ} = 50 mA)	f = 512 MHz f = 175 MHz	G_{ps}	 9.5 —	 10.5 15	 — —	dB
Drain Efficiency (V_{DD} = 7.5 Vdc, P_{out} = 3.0 W, I_{DQ} = 50 mA)	f = 512 MHz f = 175 MHz	h	 45 —	 50 55	 — —	%



C1, C3, C7, C8	0 to 20 pF Johanson	Z1	0.350″ x 0.08″ Microstrip
C2, C9	56 pF, 100 mil Chip	Z2	0.190″ x 0.08″ Microstrip
C4	10 pF, 100 mil Chip	Z3	0.800″ x 0.08″ Microstrip
C5	47 pF, Miniature Clamped Mica Capacitor	Z4	0.380″ x 0.08″ Microstrip
C6	22 pF, 100 mil Chip	Z5	0.150″ x 0.08″ Microstrip
C10, C15	10 µF, 50 V, Electrolytic	Z6	0.285″ x 0.08″ Microstrip
C11, C14	0.1 µF, Capacitor	Z7	0.340″ x 0.08″ Microstrip
C12	1000 pF, 100 mil Chip	Z8	0.070″ x 0.08″ Microstrip
C13	160 pF, 100 mil Chip	Z9	0.280″ x 0.08″ Microstrip
R1	35 Ω, 1/4 W Carbon	Z10	0.840″ x 0.08″ Microstrip
R2	30 Ω, 0.1 W Chip	Z11	0.180″ x 0.08″ Microstrip
R3	1.0 kΩ, 0.1 W Chip	Z12	0.600″ x 0.08″ Microstrip
R4	1.0 MΩ, 1/4 W Carbon	L1	7 Turns, 0.076″ ID, #24 AWG Enamel
B1	Fair Rite Products Short Ferrite Bead (2743021446)	L2	5 Turns, 0.126″ ID, #20 AWG Enamel
Board — Glass Teflon®, 31 mils		Input/Output Connectors — Type N	

Note: Plated ceramic part locators (0.1″ x 0.15″) soldered onto Z6 and Z7.

Figure 1. 512 MHz Narrowband Test Circuit

TYPICAL CHARACTERISTICS

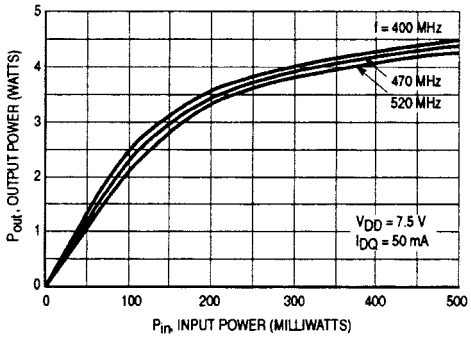


Figure 2. Output Power versus Input Power

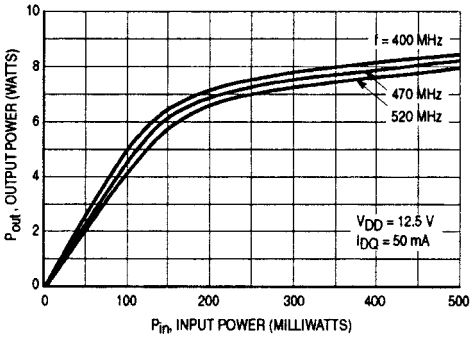


Figure 3. Output Power versus Input Power

TYPICAL CHARACTERISTICS

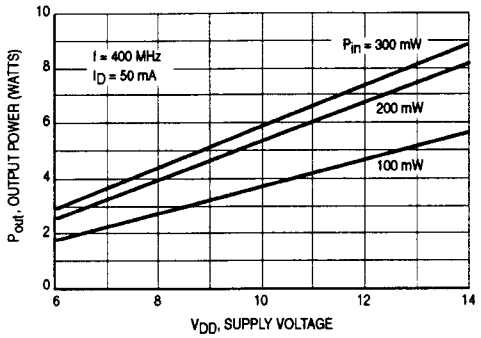


Figure 4. Output Power versus Supply Voltage

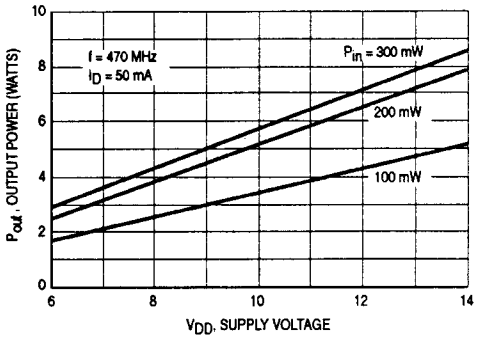


Figure 5. Output Power versus Supply Voltage

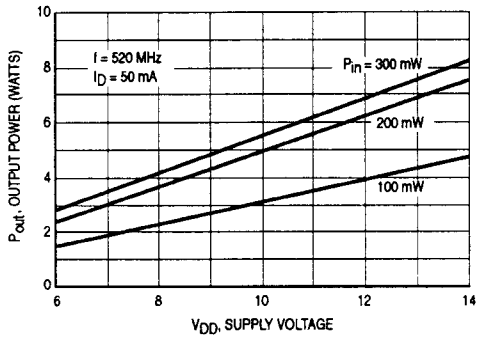


Figure 6. Output Power versus Supply Voltage

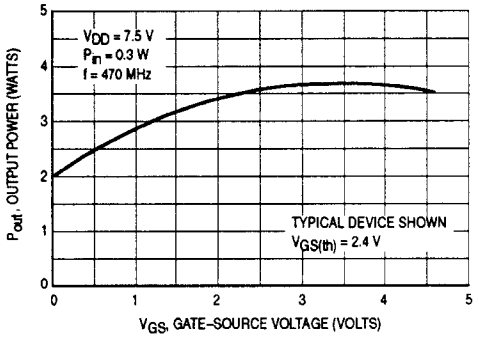


Figure 7. Output Power versus Gate Voltage

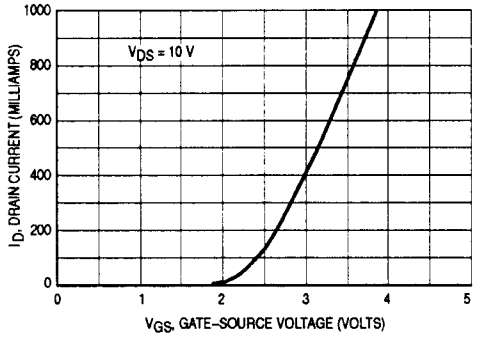


Figure 8. Drain Current versus Gate Voltage (Typical Device Shown)

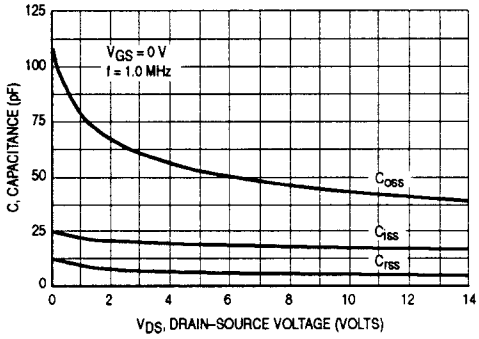


Figure 9. Capacitance versus Voltage

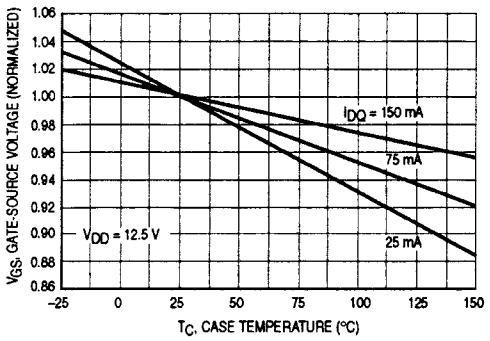


Figure 10. Gate–Source Voltage versus Case Temperature

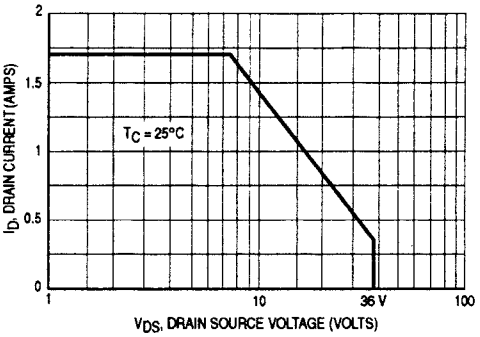
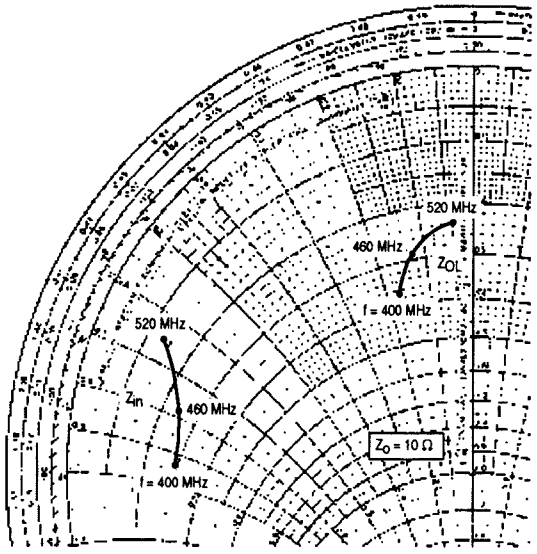


Figure 11. Maximum Rated Forward Biased Safe Operating Area



V_{DD} = 7.5 V, I_{DQ} = 50 mA, P_{out} = 3.0 W

f MHz	Z_{in} Ohms	Z_{OL}^* Ohms
400	2.8 – j9.2	3.6 – j1.7
430	2.7 – j8.5	3.3 – j1.5
460	2.5 – j7.8	2.7 – j1.1
490	2.0 – j7.2	2.5 – j0.8
520	1.3 – j6.5	2.4 – j0.5

Z_{in} = Conjugate of source impedance with parallel 35 Ω resistor and 47 pF capacitor in series with gate.

Z_{OL}^* = Conjugate of the load impedance at given output power, voltage, frequency, and η_D > 50%.

Note: Z_{ol}^* was chosen based on tradeoffs between gain, drain efficiency, and device stability.

Figure 12. Series Equivalent Input and Output Impedance

Table 1. Common Source Scattering Parameters (V_{DS} = 10 V)

I_D = 50 mA

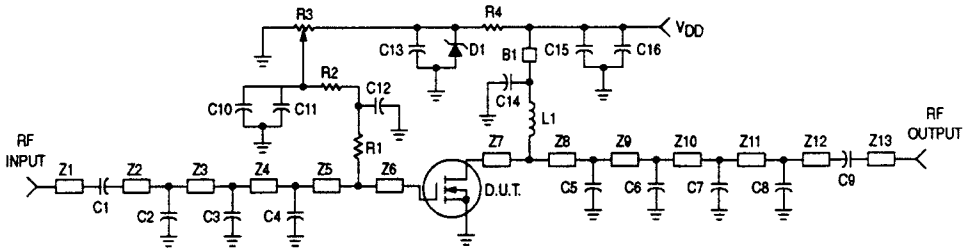
f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.69	−90	10.8	117	0.07	29	0.74	−119
100	0.58	−120	6.0	96	0.08	10	0.78	−146
200	0.58	−139	3.0	75	0.08	−7	0.81	−161
300	0.64	−147	1.9	61	0.07	−16	0.84	−166
400	0.70	−152	1.3	50	0.06	−21	0.86	−169
500	0.75	−157	0.99	41	0.05	−24	0.88	−172
700	0.82	−165	0.61	28	0.03	−15	0.92	−176
850	0.86	−171	0.45	21	0.02	13	0.94	−179
1000	0.89	−176	0.34	16	0.02	47	0.95	178

I_D = 500 mA

f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.76	−124	15.0	109	0.04	23	0.76	−151
100	0.72	−150	7.9	94	0.04	12	0.81	−165
200	0.72	−163	4.0	80	0.04	6	0.83	−172
300	0.73	−168	2.6	71	0.04	5	0.84	−175
400	0.75	−171	1.9	62	0.04	7	0.85	−176
500	0.77	−173	1.5	55	0.03	12	0.86	−178
700	0.81	−177	0.97	42	0.03	29	0.89	−180
850	0.84	−180	0.75	35	0.03	44	0.90	178
1000	0.86	177	0.60	29	0.04	55	0.92	176

I_D = 1.0 A

f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.80	−125	14.6	110	0.04	23	0.75	−155
100	0.76	−150	7.8	95	0.04	10	0.81	−167
200	0.76	−164	3.9	81	0.04	1	0.83	−173
300	0.77	−169	2.6	71	0.04	−3	0.84	−175
400	0.79	−172	1.9	63	0.03	−5	0.85	−176
500	0.80	−174	1.4	56	0.03	−5	0.86	−177
700	0.83	−178	0.95	43	0.03	−1	0.88	−179
850	0.85	179	0.73	35	0.02	9	0.90	179
1000	0.87	177	0.58	28	0.02	22	0.91	178



C1, C9	100 pF 100 mil Chip
C2	16 pF, 100 mil Chip
C3	24 pF, 100 mil Chip
C4	68 pF, 100 mil Chip
C5	51 pF, 100 mil Chip
C6	39 pF, 100 mil Chip
C7	6.2 pF, 100 mil Chip
C8	9.1 pF, 100 mil Chip
C10, C15	39000 pF, 100 mil Chip
C11, C16	10 µF, 50 V Electrolytic
C12	10000 pF, 100 mil Chip
B1	Fair Rite Products Short Ferrite Bead (2743021446)
C13	0.1 µF, 100 mil Chip
C14	160 pF, 100 mil Chip
R1	43 Ω, 0.1 W Chip Resistor
R2	1000 Ω, 0.1 W Chip Resistor
R3	10 kΩ Potentiometer
R4	3000 Ω, 0.1 W Chip Resistor
L1	5 Turns, 0.126″ ID, #20 AWG Enamel
Z1 to Z13	See Photomaster
D1	1N4734 Motorola Zener
Board — G10, 1/32″	
Input/Output Connectors — SMA	

Figure 13. Schematic of Broadband Demonstration Amplifier

PERFORMANCE CHARACTERISTICS OF BROADBAND DEMONSTRATION AMPLIFIER

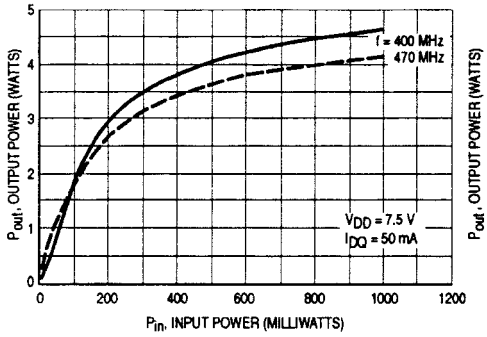


Figure 14. Output Power versus Input Power

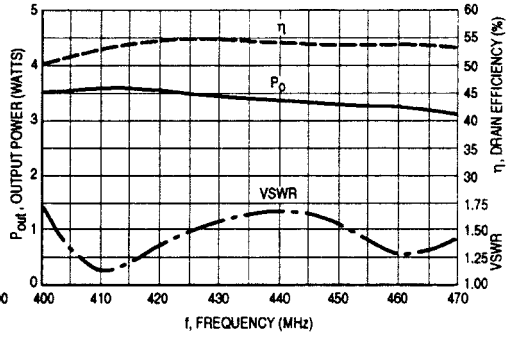


Figure 15. Output Power, Drain Efficiency and VSWR versus Frequency

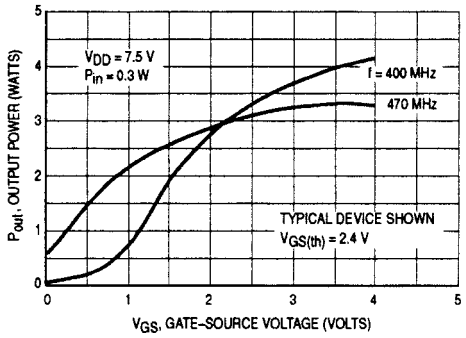


Figure 16. Output Power versus Gate Voltage

DESIGN CONSIDERATIONS

The MRF5003 is a common–source, RF power, N–Channel enhancement mode, Metal–Oxide Semiconductor Field–Effect Transistor (MOSFET). Motorola RF MOSFETs feature a vertical structure with a planar design. Motorola Application Note AN211A, "FETs in Theory and Practice", is suggested reading for those not familiar with the construction and characteristics of FETs.

This surface mount packaged device was designed primarily for VHF and UHF power amplifier applications. Manufacturability is improved by utilizing the tape and reel capability for fully automated pick and placement of parts.

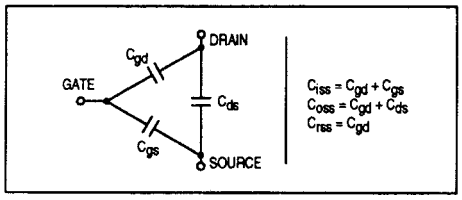
The major advantages of RF power MOSFETs include high gain, simple bias systems, relative immunity from thermal runaway, and the ability to withstand severely mismatched loads without suffering damage.

MOSFET CAPACITANCES

The physical structure of a MOSFET results in capacitors between all three terminals. The metal oxide gate structure determines the capacitors from gate–to–drain (C_{gd}), and gate–to–source (C_{gs}). The PN junction formed during fabrication of the RF MOSFET results in a junction capacitance from drain–to–source (C_{ds}). These capacitances are characterized as input (C_{iss}), output (C_{oss}) and reverse transfer (C_{rss}) capacitances on data sheets. The relationships between the inter–terminal capacitances and those given on data sheets are shown below. The C_{iss} can be specified in two ways:

1. Drain shorted to source and positive voltage at the gate.
2. Positive voltage of the drain in respect to source and zero volts at the gate.

In the latter case, the numbers are lower. However, neither method represents the actual operating conditions in RF applications.



DRAIN CHARACTERISTICS

One critical figure of merit for a FET is its static resistance in the full–on condition. This on–resistance, $R_{DS(on)}$, occurs in the linear region of the output characteristic and is specified at a specific gate–source voltage and drain current. The drain–source voltage under these conditions is termed $V_{DS(on)}$. For MOSFETs, $V_{DS(on)}$ has a positive temperature coefficient at high temperatures because it contributes to the power dissipation within the device.

GATE CHARACTERISTICS

The gate of the RF MOSFET is a polysilicon material, and is electrically isolated from the source by a layer of oxide. The input resistance is very high — on the order of $10^9\ \Omega$ — resulting in a leakage current of a few nanoamperes.

Gate control is achieved by applying a positive voltage to the gate greater than the gate–to–source threshold voltage, $V_{GS(th)}$.

Gate Voltage Rating — Never exceed the gate voltage rating. Exceeding the rated V_{GS} can result in permanent damage to the oxide layer in the gate region.

Gate Termination — The gates of these devices are essentially capacitors. Circuits that leave the gate open–circuited or floating should be avoided. These conditions can result in turn–on of the devices due to voltage build–up on the input capacitor due to leakage currents or pickup.

Gate Protection — These devices do not have an internal monolithic zener diode from gate–to–source. If gate protection is required, an external zener diode is recommended with appropriate RF decoupling.

Using a resistor to keep the gate–to–source impedance low also helps dampen transients and serves another important function. Voltage transients on the drain can be coupled to the gate through the parasitic gate–drain capacitance. If the gate–to–source impedance and the rate of voltage change on the drain are both high, then the signal coupled to the gate may be large enough to exceed the gate–threshold voltage and turn the device on.

DC BIAS

Since the MRF5003 is an enhancement mode FET, drain current flows only when the gate is at a higher potential than the source. See Figure 8 for a typical plot of drain current versus gate voltage. RF power FETs operate optimally with a quiescent drain current (I_{DQ}), whose value is application dependent. The MRF5003 was characterized at I_{DQ} = 50 mA, which is the suggested value of bias current for typical applications. For special applications such as linear amplification, I_{DQ} may have to be selected to optimize the critical parameters.

The gate is a dc open circuit and draws no current. Therefore, the gate bias circuit may generally be just a simple resistive divider network. Some special applications may require a more elaborate bias system.

GAIN CONTROL

Power output of the MRF5003 may be controlled from its rated value down to zero (negative gain) with a low power dc control signal, thus facilitating applications such as manual gain control, ALC/AGC and modulation systems. Figure 16 is an example of output power variation with gate–source bias voltage. This characteristic is very dependent on frequency and load line.

MOUNTING

The specified maximum thermal resistance of 14°C/W assumes a majority of the 0.100″ x 0.200″ source contact on the back side of the package is in good contact with an appropriate heat sink. In the test fixture shown in Figure 1, the device is clamped directly to a copper pedestal. In the demonstration amplifier, the device was mounted on top of the G10 circuit board and heat removal was accomplished through several solder filled plated through holes. As with all RF power devices, the goal of the thermal design should be to minimize the temperature at the back side of the package.

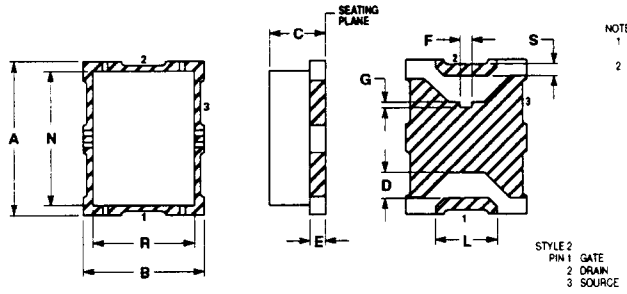
AMPLIFIER DESIGN

Impedance matching networks similar to those used with bipolar transistors are suitable for the MRF5003. For examples see Motorola Application Note AN721, "Impedance Matching Networks Applied to RF Power Transistors." Both small–signal S–parameters and large–signal impedances are provided. While the S–parameters will not produce an exact design solution for high power operation, they do yield a good first approximation. This is an additional advantage of RF power MOSFETs.

Since RF power MOSFETs are triode devices, they are not unilateral. This coupled with the very high gain of the MRF5003 yield a device capable of self oscillation. Stability may be achieved by techniques such as drain loading, input shunt resistive loading, or output to input feedback. Different stabilizing techniques were applied to the test fixture and demonstration amplifiers. The RF test fixture implements a parallel resistor and capacitor in series with the gate while the demonstration amplifier utilizes a 43 Ω shunt resistor from gate to ground. Both circuits have a load line selected for a higher efficiency, lower gain, and more stable operating region.

Two port stability analysis with the MRF5003 S–parameters provides a useful tool for selection of loading or feedback circuitry to assure stable operation. See Motorola Application Note AN215A, "RF Small–Signal Design Using Two–Port Parameters," for a discussion of two port network theory and stability.

PACKAGE DIMENSIONS



NOTES
1 DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2 CONTROLLING DIMENSION: INCH.

	INCHES		MILLIMETERS	
DIM	MIN	MAX	MIN	MAX
A	0.260	0.270	6.60	6.86
B	0.200	0.210	5.08	5.33
C	0.090	0.104	2.29	2.64
D	0.040	0.050	1.02	1.27
E	0.022	0.028	0.56	0.71
F	0.015	0.025	0.38	0.64
G	0.005	0.015	0.13	0.38
L	0.100	0.110	2.54	2.79
N	0.226	0.236	5.74	5.99
R	0.166	0.176	4.22	4.47
S	0.025	0.035	0.64	0.89

CASE 430–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters can and do vary in different applications. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and Ⓜ are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

2-3-7 GaAs MESFETs

Introduction. More than twenty years ago, for RF applications, the JFET was the dominant FET. As mentioned earlier, its major disadvantage is the fact that the f_T cutoff frequency, defined as

$$f_T = g_m / 2\pi C_{GS} \tag{2-262}$$

limits the operating range of this semiconductor device to 1 GHz at most. Also, the JFET's temperature sensitivity and loose tolerances in transconductance and other RF-related values made it difficult to use. Finally, no RFICs made use of this technology. While at the moment silicon-based MOS technology probably is the most cost-effective and is produced in the largest quantities, it does not rival the metal-semiconductor FET designed on GaAs technology. Figure 2-115 shows the structure of the silicon MOSFET, silicon JFET, and the GaAs MESFET.

Large-Signal Behavior of GaAs MESFETs. The mathematics for the large-signal behavior of the GaAsFET is basically quite similar to that for the JFET; however, the computations for JFET channel current, diode currents, and capacitance are much simpler than for the GaAsFET. Temperature effects are embedded in all the equations for all the transistors mentioned so far. In our opinion, the modified Materka model used by Ansoft and a large number of users is the most complete one; however, the following models are supported by most CAD tools:

- Chalmers (Angelov)
- Curtice–Ettenberg cubic
- Curtice quadratic
- IAF (Berroth)
- ITT PFET and TFET
- Modified Materka–Kacprzak
- Physics-based MESFET
- Raytheon (Statz)
- TriQuint (TOM-1, TOM-2, and TOM-3)
- With some restrictions, metal-insulator-semiconductor FETs (MISFETs), modulation-doped FETs (MODFETs), and high-electron-mobility transistors (HEMTs)

The main advantages of GaAsFET technology derive from the fact that it uses a metal-semiconductor junction with a barrier voltage of 0.8 V, its input capacitance is typically less than 0.2 pF, and the reverse feedback capacitance is less than 0.02 pF, or roughly 10% of the input capacitance. As a result of this, $f_{\max}$ is approximately five times higher than f_T. Table 2-16 shows a comparison of silicon BJT and GaAsFET technologies. While $f_{\max}$ for the bipolar transistor is

$$f_{\max} \simeq \sqrt{\frac{f_t}{8\pi r'_{bb} C_C}} \tag{2-263}$$

the $f_{\max}$ determination for the GaAsFET is given by

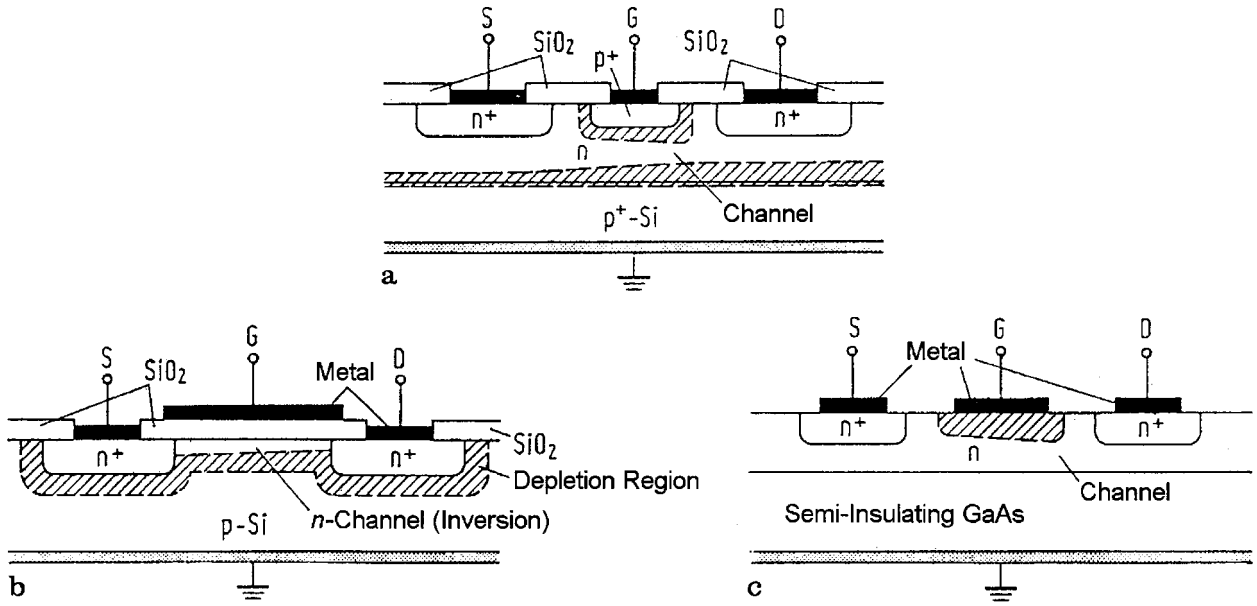


Figure 2-115 This picture is provided to differentiate between the structure of the three technologies. (a) The JFET and (b) the MOSFET are based on silicon; (c) the GaAsFET is based on semi-insulating GaAs. While the MOS transistor has silicon dioxide (SiO_2) as gate insulation, both the JFET and GaAs MESFET have a diode from gate to source, with a barrier voltage of 0.65 V for silicon and 0.8 V for GaAs.

Table 2-16 Comparison of Si BJT, SiGe HBT, and GaAsFET technologies

	Si Bipolar	SiGe HBT	GaAs FET
f_T	25 GHz	100 GHz	22 GHz
$f_{\max}$	40 GHz	200 GHz	110 GHz
Features	Low cost, low 1/*f* noise (5 kHz)	Low 1/*f* noise, very low distortion, high cost, IC only	Highest flexibility, lowest NF_0, well established

$$f_{\max} = \frac{f_T}{2}\sqrt{\frac{R_0}{R_i + R_s + R_g}} \tag{2-264}$$

As a sample calculation,

$$f_{\max}\ \frac{21.9\text{ GHz}}{2}\sqrt{\frac{450}{1 + 1.5 + 2}} = 110\text{ GHz} \tag{2-265}$$

The three major drawbacks of the GaAsFET are:

1. Much higher flicker corner frequency (somewhere between 10 and 100 MHz), probably due to a lack of surface passivation.
2. Much higher output conductance. This tends to load down any circuit connected to the drain. On the other hand, since the transconductance is quite high for even low currents, these devices have very high gains at low frequencies, which can make them quite unstable. In the saturated mode it is not uncommon to find a drain–source resistance of 100–500 Ω, while BJTs and JFETs offer values of several kilohms and higher.
3. Because of the very high flicker corner frequency (from 10 to 100 MHz), MESFETs are really not useful for low-noise mixers and oscillators, and unless there are no devices available in the frequency range above 30 GHz, they should be avoided for these applications.

As to the MESFET's construction and dc properties, Figure 2-116 shows a MESFET's cross section and dc *I*–*V* characteristics.

It was outlined in the beginning that while the GaAsFET is a close relative to the JFET and MOSFET, its actual behavior was found to be described best by a set of analytic equations. The first such model was the one by Curtice in the form of quadratic and cubic models, but it does not have enough derivatives to give enough insight into subtle things such as third- and higher-order intermodulation distortion and accurate harmonic generation. Other researchers have addressed various areas, but we still find that the Materka model has the best approximation, especially using the two-transistor approach outlined later [20, 21]. The large-signal topology for all FETs consists of an intrinsic model with some extrinsic parameters, further complicated by the package, as shown by Figures 2-117 and 2-118. Table 2-17 lists their keywords. The actual intrinsic model and its parameter definition depend on the particular model and since designs using GaAsFET will always be done using CAD tools, we will not go into any details of the equations but will list them. They are not dissimilar from the JFET and MOSFET equations.

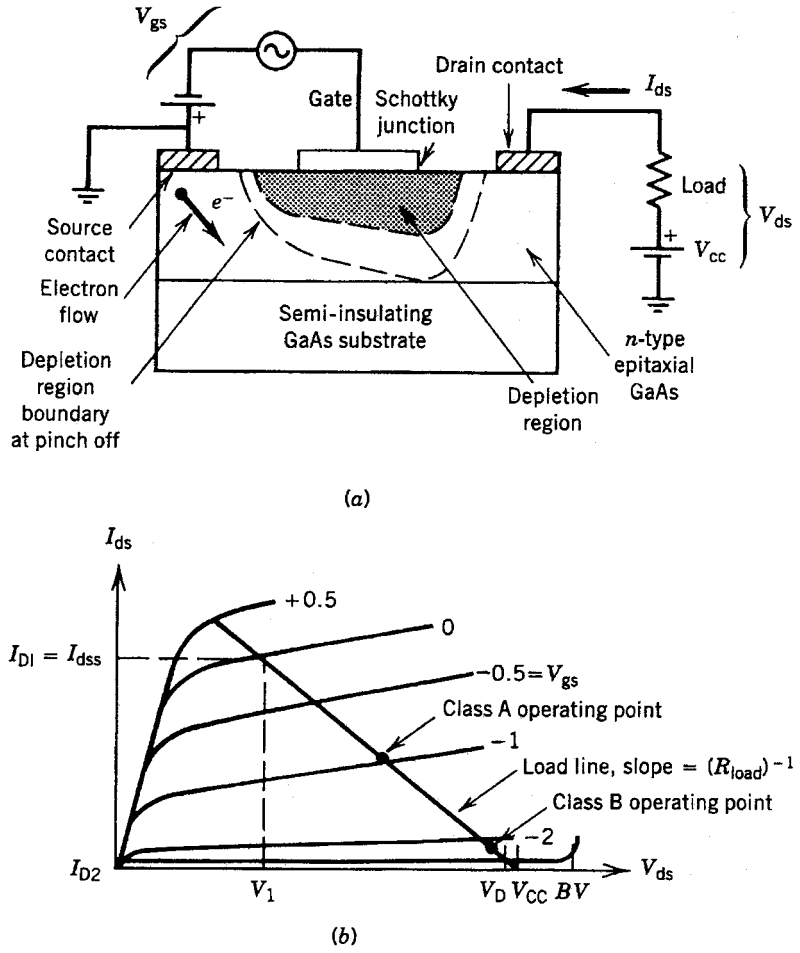


Figure 2-116 (a) Cross section and bias circuit and (b) dc *I–V* curve, including ac load line, for a MESFET.

The Modified Materka-Kacprzak Model Figure 2-119 shows the intrinsic model of the Materka FET. Table 2-18 lists its keywords.

Large-Signal Equations

Device Equations

V_{gsi} = Intrinsic gate–source voltage
V_{dsi} = Intrinsic drain–source voltage
V_1 = Voltage across C_{GS} and R_i
V_{gdi} = Intrinsic gate–drain voltage
$V_T = kT_J/q$ (thermal voltage)
k = Boltzmann's constant
q = Electron charge
T_J = Analysis temperature (kelvins)

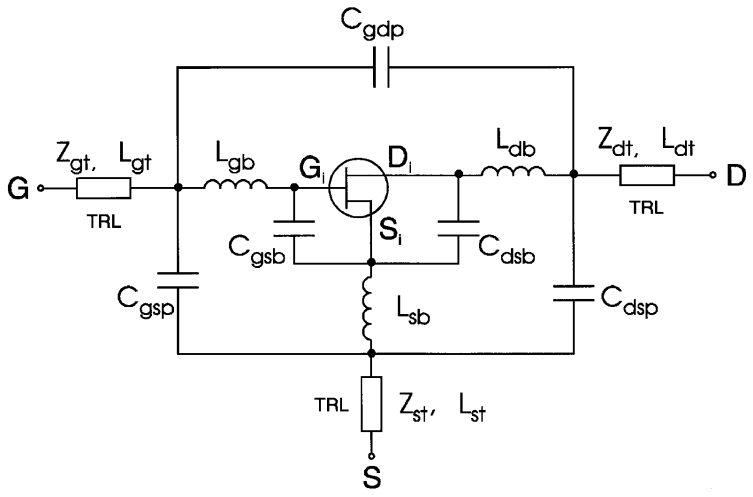


Figure 2-117 MESFET extrinsic model.

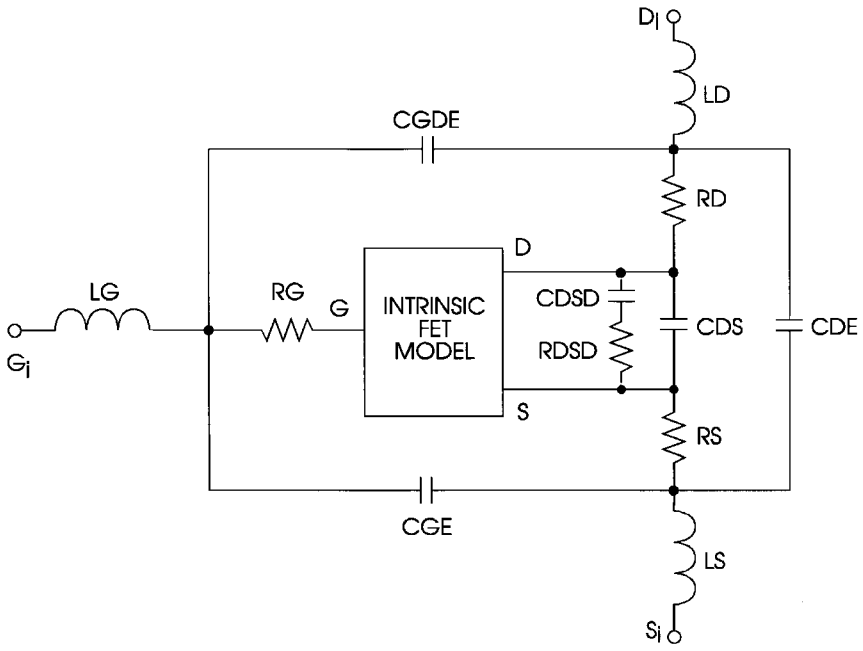


Figure 2-118 MESFET package model.

Table 2-17 Large-signal MESFET model keywords

Keyword	Description	Unit	Default
	Extrinsic Parameters		
RG	Gate bulk and ohmic resistance	ohm	0.0
RD	Drain bulk and ohmic resistance	ohm	0.0
RS	Source bulk and ohmic resistance	ohm	0.0
LG	Gate-lead inductance (metallization)	henry	0.0
LD	Drain-lead inductance (metallization)	henry	0.0
LS	Source-lead inductance (via)	henry	0.0
CDS	Drain–source capacitance	farad	0.0
CDSD	Low-frequency trapping capacitor	farad	0.0
RDSD	Channel trapping resistance	ohm	∞
CGE	Gate–source electrode capacitance	farad	0.0
CDE	Drain–source electrode capacitance	farad	0.0
CGDE	Gate–drain electrode capacitance	farad	0.0
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate-to-source package capacitance	farad	0.0
CDSP	Drain-to-source package capacitance	farad	0.0
CGDP	Gate-to-drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for $\varepsilon_r = 1$	meter	0.0

Channel Current

$$I_{ds} = \mathrm{IDSS}\left(1 + \mathrm{SS}\,\frac{V_{dsi}}{\mathrm{IDSS}}\right)\left(1 - \frac{V_{gsi}(t-T)}{\mathrm{VP0} + \mathrm{GAMA}V_{dsi}}\right)^{(\mathrm{E}+\mathrm{KE}V_{gsi}(t-T))}$$

$$\times \tanh\left(\frac{\mathrm{SL}V_{dsi}}{\mathrm{IDSS}(1 - \mathrm{KG}V_{gsi}(t-T))}\right)$$

Diode

$$I_{gd} = I_{gdc} - \begin{cases} \mathrm{IB0}\,\exp[-\mathrm{AFAB}(V_{gdi} + \mathrm{VBC})] \\ \dfrac{\mathrm{GMAX}}{4}\,\{\tanh[\mathrm{K1D}(V_{gsi} - \mathrm{K2D})] - 1\}\,[V_{gdi} + \mathrm{VBC} - \sqrt{(V_{gdi} + \mathrm{VBC})^2 + \mathrm{K3D}}] \end{cases}$$

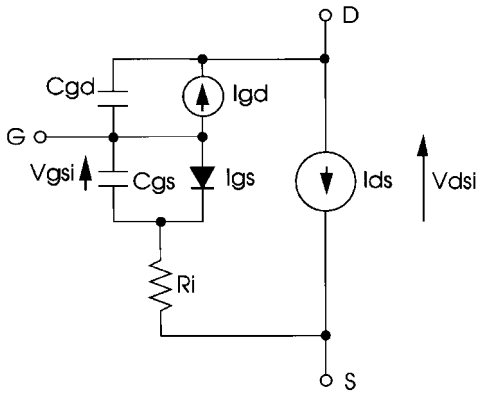


Figure 2-119 Intrinsic model of the modified Materka–Kacprzak MESFET.

where

$$I_{gdc} = \text{IG0}[\exp(\text{AFAB}V_{gdi}) - 1]$$

Channel Resistance

$$R_i = \begin{cases} \text{R10}(1 - \text{KR}V_{gsi}), & \text{KR}V_{gsi} < 1.0 \\ 0, & \text{KR}V_{gsi} \geq 1.0 \end{cases}$$

Capacitance Model

$$C_{gs} = \text{CGS0}\frac{F1F2}{\sqrt{1 - \dfrac{V_{\text{new}}}{\text{VBI}}}} + \text{CGD0}\ F3$$

$$C_{gd} = \text{CGS0}\frac{F1F3}{\sqrt{1 - \dfrac{V_{\text{new}}}{\text{VBI}}}} + \text{CGD0}\ F2$$

where

$$F1 = \frac{1}{2}\left(1 + \frac{V_{\text{eff}} - V_T}{\sqrt{(V_{\text{eff}} - V_T)^2 + \delta^2}}\right)$$

$$F2 = \frac{1}{2}\left(1 + \frac{V_{gsi} - V_{gdi}}{\sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}}\right)$$

Table 2-18 Materka–Kacprzak MESFET model keywords

Keyword	Description	Unit	Default
	Area, Noise, and Name		
AREA	Area multiplier		1.0
KFN	Flicker noise coefficient (Materka model only)[a]		0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
	Channel Current Model		
IDSS	Drain saturation current for $V_{GS} = 0$	ampere	0.1
VP0	Pinchoff voltage for $V_{DS} = 0$	volt	–2.0
GAMA	Voltage slope parameter of pinchoff voltage	/volt	0.0
E	Constant part of power-law parameter		2.0
KE	Dependence of power law on V_{GS}	/volt	0.0
SL	Slope of the $V_{GS} = 0$ drain characteristic in the linear region	ampere/volt	0.15
KG	Drain dependence on V_{GS} in the linear region	/volt	0.0
SS	Slope of the drain characteristic in the saturated region	ampere/volt	0.0
T	Channel transit-time delay	second	0.0
IG0	Diode saturation current	ampere	0
AFAG	Slope factor of forward diode current	/volt	38.696
IB0	Breakdown saturation current	ampere	0
AFAB	Slope factor of breakdown current	/volt	0
VBC	Breakdown voltage	volt	∞
GMAX	Breakdown conductance	ampere/volt	0
K1D	Fitting parameter	/volt	0
K2D	Fitting parameter	volt	0
K3D	Fitting parameter	$volt^2$	0
R10	Intrinsic channel resistance for $V_{GS} = 0$	ohm	0.0
KR	Slope factor of intrinsic channel resistance	/volt	0.0
	Materka Capacitance Model		
C10	Gate–source Schottky barrier capacitance for $V_{GS} = 0$	farad	0.0
K1	Slope parameter of gate–source capacitance	/volt	1.25
MGS	Gate–source grading coefficient		0.5
C1S	Constant parasitic component of gate–source capacitance	farad	0.0
CF0	Gate–drain feedback capacitance for $V_{GD} = 0$	farad	0.0
KF	Slope parameter of gate–drain feedback capacitance	/volt	1.25
MGD	Gate–drain grading coefficient		0.5
FCC	Forward-bias depletion capacitance coefficient		0.8

[a]The flicker noise parameter of the Materka model is KFN so as not to conflict with the KF parameter in the capacitance model.

$$F3 = \frac{1}{2}\left(1 - \frac{V_{gsi} - V_{gdi}}{\sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}}\right)$$

$$V_{\text{new}} = \begin{cases} A1, & A1 < \text{VMAX} \\ \text{VMAX}, & A1 \geq \text{VMAX} \end{cases}$$

$$A1 = \frac{1}{2}\left(V_{\text{eff}} + V_T + \sqrt{(V_{\text{eff}} - V_T)^2 + \delta^2}\right)$$

$$V_{\text{eff}} = \frac{1}{2}\left(V_{gsi} + V_{gdi} + \sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}\right)$$

$$V_T = \text{VP0} + \text{GAMA}\, V_{dsi}$$

$$\delta = 0.2$$

Some of the modifications to the Materka model have been done by Ansoft under various Department of Defense contracts, and by Raytheon and Texas Instruments under similar contracts, with Ansoft being a subcontractor.

The most relevant equation is really the channel current. Its derivatives are largely responsible for the accuracy of the intermodulation distortion, power-added efficiency, and, of course, its dc *I–V* curves.

Enhancement/Depletion FETs To make the designer's life more difficult, it turns out that there are two types of GaAsFETs:

1. *Depletion FETs (DFETs).* Most similar to the JFET; here V_G must be negative to control the device. They are the most commonly produced and are the FET type most referred to in this book.

On one hand:

- They require a negative gate voltage with respect to the source.
- Self-bias allows operation from a single supply voltage.

On the other hand:

- For low-voltage operation, a negative voltage generator may be required.
- Supply voltage must be doubled to accommodate full-swing operation.

2. *Enhancement FETs (EFETs).* Most similar to the MOSFET; here V_G must be positive to bring life to the device. Practically speaking, EFETs are used mostly in integrated circuits; they're typically not available in discrete, packaged form.

On one hand:

- They need only positive supply for biasing.
- They provide higher g_m/mA (for same device width)—5.1 mS versus 3.9 mS at 8 mA.

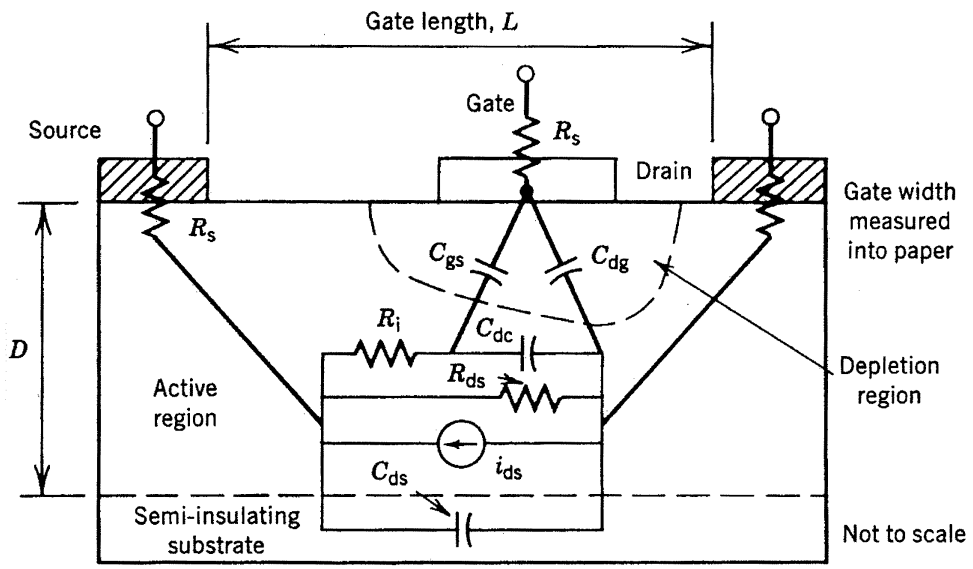


Figure 2-120 Location of lumped-element components for a MESFET.

- They are good for low-power LNAs, giving slightly better NF that DFETs, an NF of 1 dB at 1 GHz, and $I_{dd} < 10$ mA.

On the other hand:

- They have a very limited gate bias range (V_{GS} between 0.15 and 0.7 V).
- The gate conduction degrades NF and input impedance.
- The gate capacitance is higher than that of DFETs.
- The linearity is not as good as that of DFETs.

With today's technologies, all GaAs devices are *n*-channel; we have not seen any *p*-channels yet. More information about biasing will be given in the next chapter.

Figure 2-120 is a lumped-element, two-part equivalent circuit of a MESFET showing the location of lumped-element components.

2-3-8 Small-Signal GaAs MESFET Model

Figure 2-121 shows the applicable linear equivalent circuit for a MESFET and Table 2-19 lists its keywords. As with the MOS transistors, there is a gallium-arsenide dual-gate MOSFET available that is used mostly in special circuits, such as preamplifiers and mixers, whose IF has to be significantly higher than the flicker corner frequency, for example, higher than 200 MHz.

The following two datasheets* show typical GaAsFETs, both low-noise and medium-power. Such devices are, of course, available from a large variety of manufacturers—to name at least two, Siemens and the Nippon Electric Company, which is represented in the United States by California Eastern Laboratories.

*Reproduced with permission.

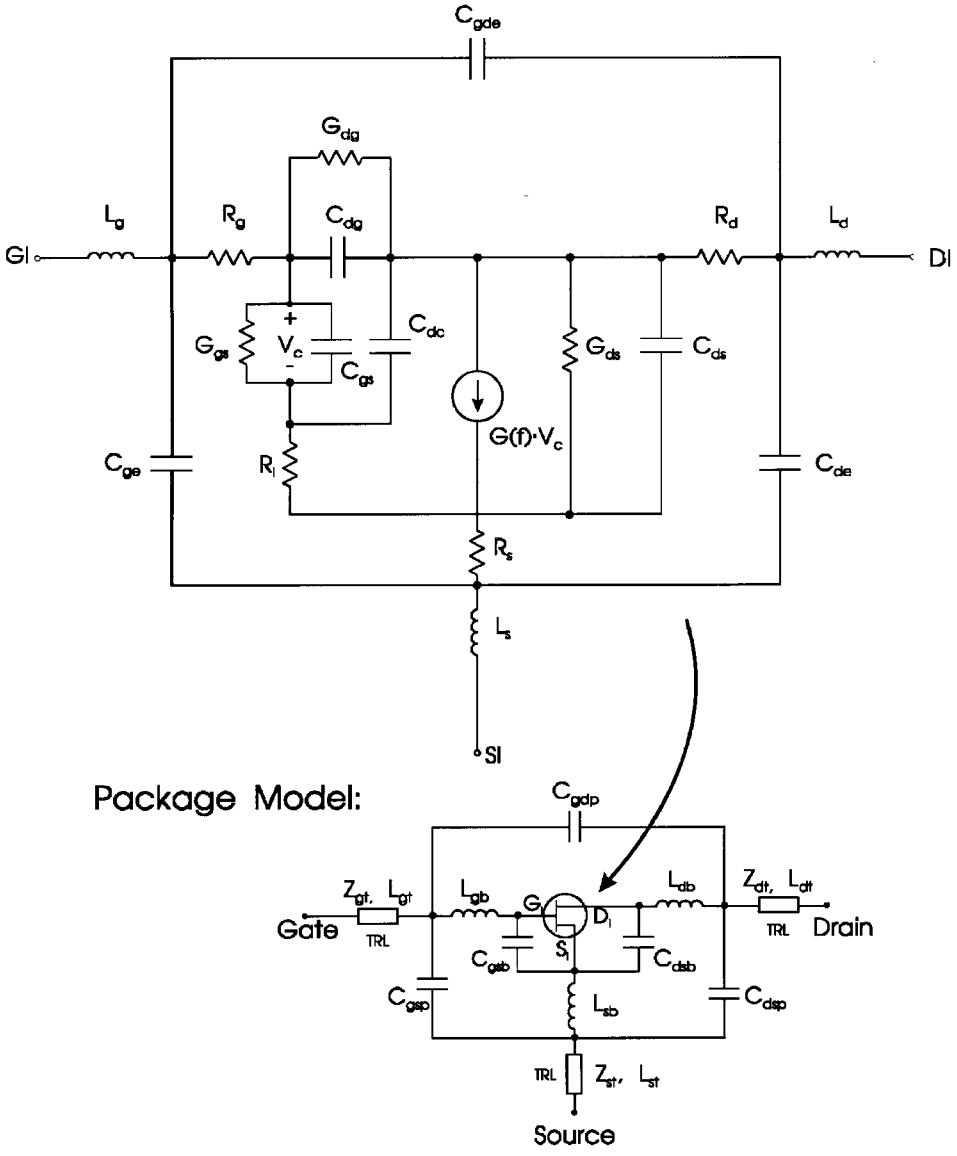


Figure 2-121 Small-signal model of a MESFET.

Table 2-19 Small-signal MESFET model keywords

Keyword	Description	Unit	Default
G	Transconductance at dc, G_0[a]	/ohm	
CGS	Gate–source capacitance	farad	
F	3-dB Roll-off frequency	hertz	∞
T	Time delay	second	0.0
TDS	Drain–source time delay	second	0.0
GGS	Gate–source conductance	/ohm	0.0
CDG	Drain–gate capacitance	farad	0.0
CDC	Dipole-layer capacitance	farad	0.0
CDS	Drain–source capacitance	farad	0.0
GDS	Drain–source conductance	/ohm	0.0
RI	Channel resistance	ohm	0.0
RG	Gate resistance	ohm	0.0
RD	Drain resistance	ohm	0.0
RS	Source resistance	ohm	0.0
CGE	External gate capacitance	farad	0.0
CDE	External drain capacitance	farad	0.0
LG	Gate-lead inductance	henry	0.0
LD	Drain-lead inductance	henry	0.0
LS	Source-lead inductance	henry	0.0
CGDE	External gate–drain capacitance	farad	0.0
GDG	Gate–drain conductance	/ohm	0.0
TJ	Chip temperature	kelvin	298
	Package Parasitics		
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate-to-source package capacitance	farad	0.0
CDSP	Drain-to-source package capacitance	farad	0.0
CGDP	Gate-to-drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for at $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for at $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for at $\varepsilon_r = 1$	meter	0.0
FC	Corner frequency of flicker (1/*f*) noise[b]	hertz	10 MHz
FCP	Shape factor of the 1/*f* noise response	1.0	
label	User-defined term that refers to temperature coefficient		

[a]The transconductance of this model may approximately be described by $g_m = G\dfrac{e^{-j\omega T}}{1 = j(f/F)}$ where $\omega = 2\pi f$ and f is frequency.

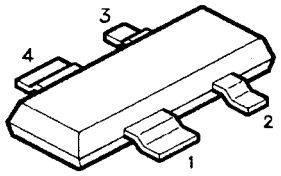
[b]The flicker noise frequency dependence is given by $1/(f/F_c)^{\mathrm{FCP}}$

SIEMENS AlGaAs / InGaAs HEMT CFY 77

Datasheet

Features

* Very low noise
* Very high gain
* For low noise front end amplifiers up to 20 GHz
* For DBS down converters



ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering code (taped)	Package 1)
CFY77-08	HG	Q62702-F1549	MW-4
CFY77-10	HH	Q62702-F1559	MW-4

Maximum ratings	Symbol		Unit
Drain-source voltage	V_{DS}	3.5	V
Drain-gate voltage	V_{DG}	4.5	V
Gate-source voltage	V_{GS}	-3.0	V
Drain current	I_D	60	mA
Channel temperature	T_{Ch}	150	°C
Storage temperature range	T_{stg}	-65...+150	°C
Total power dissipation ($T_s \leq 51$°C) 2)	P_{tot}	180	mW
Thermal resistance			
Channel-soldering point source	R_{thChS}	550	K/W

1) Dimensions see chapter Package Outlines
2) T_S: Temperature measured at soldering point

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Electrical characteristics at T_A = 25°C

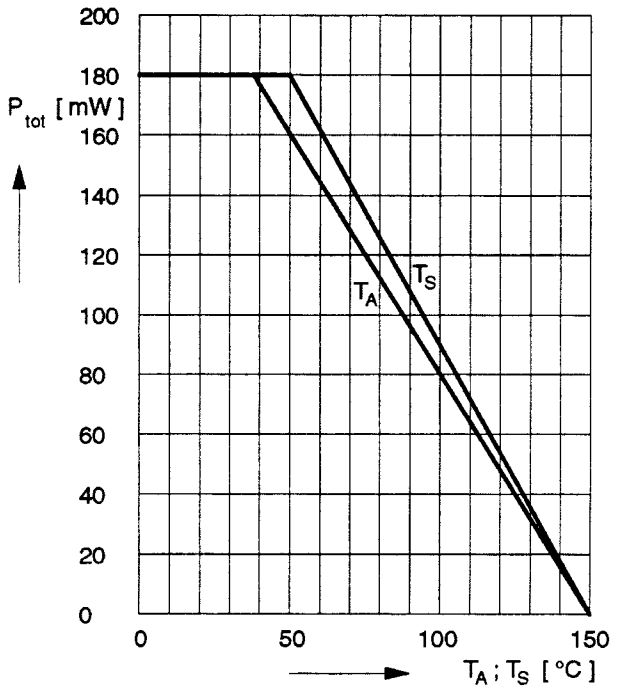
unless otherwise specified

Characteristics	Symbol	min	typ	max	Unit
Drain-source saturation current V_{DS} = 2 V V_{GS} = 0 V	I_{DSS}	15	30	60	mA
Pinch-off voltage V_{DS} = 2 V I_D = 1 mA	$V_{GS(P)}$	-2	-0.7	-0.2	V
Gate leakage current V_{DS} = 2 V I_D = 15 mA	I_G	-	0.05	2	µA
Transconductance V_{DS} = 2 V I_D = 15 mA	g_m	50	65	-	mS
Noise figure V_{DS} = 2 V I_D = 15 mA f = 12 GHz	F				dB
CFY77-08 CFY77-10		- -	0.7 0.9	0.8 1	
Associated gain V_{DS} = 2 V I_D = 15 mA f = 12 GHz	G_a				dB
CFY77-08 CFY77-10		10 9.5	10.5 10	- -	

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Total Power Dissipation $P_{tot} = f(T_s;T_A)$

Package mounted on alumina



Typical Common Source Noise Parameters

I_D = 15 mA U_{DS} = 2.0 V Z_0 = 50 Ω

f	F_{min}	G_a	Γ_{opt}		R_n	r_n	N	$F_{50\Omega}$
GHz	dB	dB	MAG	ANG	Ω			dB
2	0.36	19.4	0.79	27	13.7	0.274	0.03	1.2
4	0.44	15.9	0.72	60	10.1	0.202	0.04	1.1
6	0.51	13.9	0.63	92	5.85	0.117	0.05	1.05
8	0.58	12.4	0.56	134	2.35	0.047	0.06	1.0
10	0.65	11.2	0.52	180	1.1	0.022	0.07	1.0
12	0.72	10.4	0.54	-135	2.9	0.058	0.08	1.1
14	0.80	9.7	0.59	-108	7.15	0.143	0.10	1.5

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Typical Common Source S-Parameters

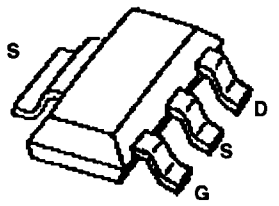
I_D = 15 mA U_D = 2.0 V Z_o = 50 Ω

	S11		S21		S12		S22	
GHz	Mag	Ang	Mag	Ang	Mag	Ang	Mag	Ang
1	0.98	-22.8	5.55	159.6	0.030	87.5	0.633	-16.3
2	0.94	-46.1	5.40	139.3	0.053	57.8	0.60	-32.5
3	0.88	-68.4	5.09	120.1	0.074	44.9	0.54	-48.0
4	0.82	-90.6	4.77	101.2	0.089	30.7	0.48	-63.3
5	0.77	-110.8	4.45	84.0	0.101	18.1	0.42	-77.5
6	0.72	-131.4	4.16	67.3	0.112	7.9	0.35	-92.6
7	0.66	-153.6	3.88	50.2	0.119	-3.3	0.28	-110.8
8	0.63	-175.2	3.58	34.5	0.122	-12.7	0.22	-132.0
9	0.62	164.4	3.29	18.9	0.120	-22.0	0.16	-157.3
10	0.62	145.0	3.01	4.0	0.119	-29.5	0.14	177.3
11	0.64	128.3	2.76	-10.3	0.119	-37.4	0.15	136.2
12	0.64	113.1	2.51	-23.5	0.114	-44.0	0.18	115.4
13	0.66	101.3	2.32	-35.7	0.114	-47.3	0.23	100.9
14	0.67	89.4	2.18	-48.2	0.116	-53.1	0.25	91.0
15	0.69	73.6	2.06	-62.4	0.116	-58.6	0.28	75.4
16	0.73	59.2	1.85	-75.9	0.115	-65.8	0.36	57.1
17	0.76	51.7	1.65	-86.5	0.112	-69.4	0.39	53.1
18	0.78	45.4	1.56	-96.7	0.115	-72.3	0.42	43.8
19	0.77	36.2	1.51	-108.6	0.121	-76.7	0.44	38.8

SIEMENS GaAs FET **CLY 15**

Datasheet

* Power amplifier for mobile phones
* For frequencies from 400 MHz to 2.5 GHz
* Operating voltage range: 2.7 to 6 V
* P_{out} at V_D=3V, f=1.8 GHz typ. 31.5 dBm
* Efficiency better 50%



ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering code (taped)	Package 1)
CLY 15	CLY 15	Q62702-L99	SOT 223

Maximum ratings	Symbol		Unit
Drain-source voltage	V_{DS}	9	V
Drain-gate voltage	V_{DG}	12	V
Gate-source voltage	V_{GS}	-6	V
Drain current	I_D	5	A
Channel temperature	T_{Ch}	150	°C
Storage temperature	T_{stg}	-55...+150	°C
Total power dissipation (Ts ≤ 80°C) Ts: Temperature at soldering point	P_{tot}	4.7	W
Thermal resistance Channel-soldering point (GND)	R_{thChS}	< 15	K/W

1) Dimensions see chapter Package Outlines

SIEMENS GaAs FET **CLY 15**

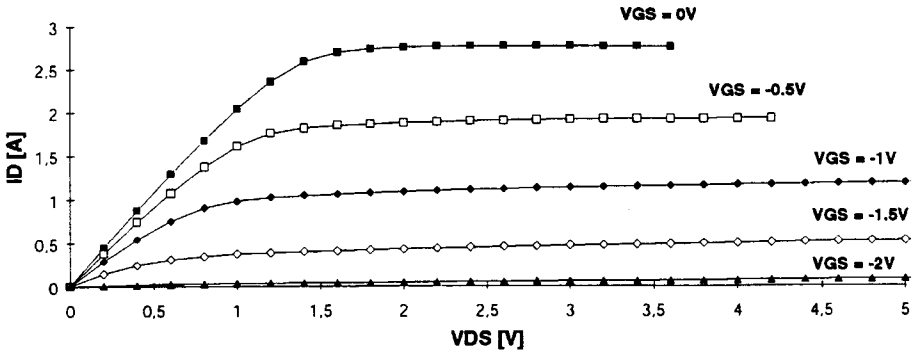
Electrical characteristics (T_A = 25°C , unless otherwise specified)

Characteristics	Symbol	min	typ	max	Unit
Drain-source saturation current *) V_{DS} = 3V V_{GS} = 0V	I_{DSS}	2.4	3.2	4.8	A
Cut-off current V_{DS} = 3V V_{GS} = -3.8V	I_D	-	-	400	µA
Gate cut-off current V_{DS} = 3V V_{GS} = -3.8V	I_G	-	20	70	µA
Pinch-off Voltage VDS=3V ID=400µA	$V_{GS(p)}$	-3.8	-2.8	-1.8	V
Small Signal Gain *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 5dBm	G	-	6	-	dB
Output Power *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 29dBm	P_o	32	32.5	-	dBm
Output Power *) V_{DS} = 5V I_D = 1.4A f = 1.8GHz P_{in} = 30 dBm	P_o	34.5	35	-	dBm
1dB-Compression Point *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz	P_{1dB}	-	31.5	-	dBm
1dB-Compression Point *) V_{DS} = 5V I_D = 1.4A f = 1.8GHz	P_{1dB}	-	34.5	-	dBm
Power Added Efficiency *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 29dBm	η_D	45	50	-	%

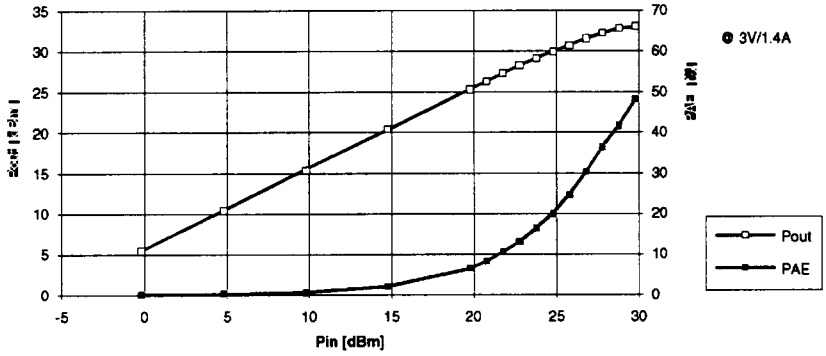
*) pulsed measurement; duty cycle 1:10; t_{on} = 1ms, power matching conditions.

SIEMENS GaAs FET **CLY 15**

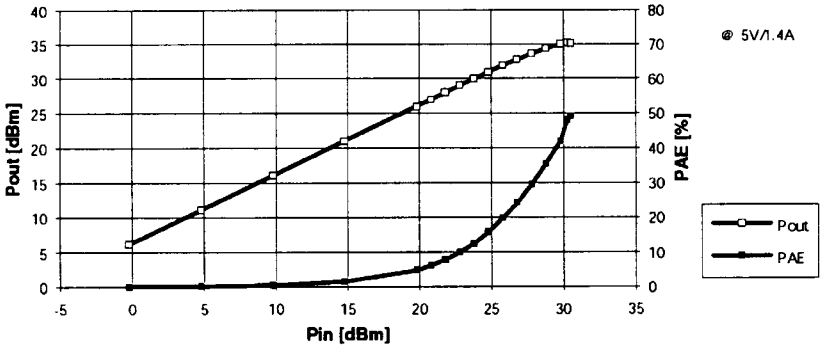
Output Characteristics



Power Characteristics



Power Characteristics



SIEMENS GaAs FET **CLY 15**

typ. Common Source S-Parameter

V_{DS} = 3V I_D=1.4A Z_O=50Ω

f	S11		S21		S12		S22	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
0.200	0.91	-150.9	5.69	99.7	0.01	48.5	0.90	176.0
0.250	0.91	-160.6	4.63	93.5	0.02	45.9	0.87	173.8
0.300	0.90	-167.9	3.89	88.7	0.02	46.8	0.88	171.8
0.350	0.90	-173.7	3.34	84.6	0.02	47.4	0.87	170.8
0.400	0.90	-178.7	2.92	80.9	0.02	47.5	0.87	168.8
0.450	0.90	176.9	2.60	77.5	0.02	47.6	0.87	167.3
0.500	0.90	173.0	2.34	74.6	0.02	48.1	0.87	165.8
0.550	0.90	169.5	2.12	71.4	0.02	47.7	0.87	164.2
0.600	0.90	166.1	1.95	68.7	0.03	47.0	0.87	162.8
0.650	0.90	163.1	1.79	66.1	0.03	47.1	0.87	161.2
0.700	0.90	160.0	1.66	63.5	0.03	46.6	0.87	159.7
0.750	0.90	157.2	1.54	60.9	0.03	45.6	0.87	158.3
0.800	0.90	154.6	1.45	58.6	0.03	45.0	0.87	156.9
0.850	0.90	152.0	1.36	56.1	0.03	43.9	0.87	155.6
0.900	0.90	149.3	1.28	53.8	0.04	43.0	0.87	154.0
0.950	0.90	146.9	1.21	51.5	0.04	41.9	0.87	152.6
1.000	0.90	144.5	1.15	49.0	0.04	41.0	0.87	151.3
1.200	0.91	135.2	0.95	40.3	0.05	36.1	0.87	145.8
1.400	0.91	126.7	0.81	31.8	0.05	31.9	0.88	140.1
1.600	0.92	118.5	0.70	23.8	0.06	26.1	0.88	134.7
1.800	0.92	110.6	0.61	16.3	0.06	20.8	0.88	129.7
2.000	0.93	103.2	0.55	8.7	0.06	15.6	0.89	124.3
2.200	0.93	96.3	0.49	2.1	0.07	10.4	0.88	119.1
2.400	0.93	89.3	0.44	-4.1	0.07	5.2	0.90	114.4
2.600	0.94	82.8	0.40	-10.0	0.07	0.2	0.90	109.3
2.800	0.94	77.0	0.37	-14.9	0.07	-4.2	0.90	104.5
3.000	0.94	71.3	0.34	-19.6	0.08	-9.7	0.91	99.8
3.200	0.93	66.0	0.32	-23.4	0.08	-15.0	0.92	95.1
3.400	0.92	61.6	0.31	-26.8	0.08	-19.4	0.93	90.8
3.600	0.91	57.3	0.30	-29.7	0.07	-23.7	0.92	87.0
3.800	0.90	53.1	0.31	-33.1	0.07	-28.1	0.93	83.1
4.000	0.89	49.2	0.32	-38.1	0.07	-31.9	0.93	79.8
4.200	0.86	46.4	0.34	-44.9	0.07	-35.4	0.92	76.4
4.400	0.83	44.7	0.36	-55.4	0.07	-37.5	0.92	73.4
4.600	0.89	44.2	0.07	-36.2	0.07	-38.1	0.92	71.0
4.800	0.83	43.7	0.34	-80.6	0.07	-39.4	0.92	68.2
5.000	0.85	42.2	0.30	-92.1	0.07	-40.3	0.92	65.2
5.200	0.88	39.4	0.27	-100.8	0.07	-42.5	0.92	62.2
5.400	0.89	36.5	0.24	-107.8	0.07	-45.0	0.92	58.7
5.600	0.90	33.1	0.22	-113.6	0.07	-48.6	0.92	55.7
5.800	0.91	29.6	0.19	-118.9	0.07	-51.4	0.92	52.1
6.000	0.92	26.4	0.18	-124.4	0.07	-54.4	0.92	48.0

SIEMENS GaAs FET **CLY 15**

typ. Common Source S-Parameter
V_{DS} = 5V I_D=1.4A Z_O=50Ω

f	S11		S21		S12		S22	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
0.200	0.90	-151.1	7.61	98.8	0.01	46.5	0.84	176.7
0.250	0.89	-160.6	6.18	92.4	0.02	43.4	0.82	174.7
0.300	0.89	-167.8	5.19	87.5	0.02	46.5	0.82	172.9
0.350	0.89	-173.7	4.45	83.3	0.02	46.0	0.82	171.8
0.400	0.88	-178.7	3.90	79.4	0.02	46.5	0.82	169.7
0.450	0.89	177.0	3.47	75.9	0.02	47.3	0.82	168.3
0.500	0.88	173.2	3.11	72.8	0.02	47.9	0.82	166.7
0.550	0.88	169.6	2.82	69.5	0.02	47.8	0.82	165.5
0.600	0.89	166.4	2.59	66.6	0.03	47.4	0.82	163.9
0.650	0.88	163.1	2.38	63.9	0.03	47.4	0.82	162.6
0.700	0.89	160.3	2.20	61.1	0.03	46.5	0.82	161.0
0.750	0.89	157.5	2.05	58.4	0.03	45.6	0.82	159.6
0.800	0.89	154.9	1.91	55.9	0.03	45.3	0.82	158.0
0.850	0.89	152.1	1.79	53.2	0.03	44.8	0.82	156.8
0.900	0.89	149.7	1.69	50.7	0.03	43.9	0.82	155.4
0.950	0.89	147.1	1.59	48.4	0.04	42.7	0.82	154.1
1.000	0.89	144.7	1.51	45.7	0.04	42.0	0.82	152.8
1.200	0.89	135.5	1.24	36.2	0.04	37.8	0.83	147.3
1.400	0.90	127.1	1.04	27.1	0.05	32.2	0.83	141.9
1.600	0.91	119.1	0.90	18.3	0.05	27.4	0.84	136.8
1.800	0.92	111.1	0.78	10.1	0.06	22.5	0.84	131.6
2.000	0.92	103.7	0.68	2.1	0.06	18.2	0.85	126.3
2.200	0.93	96.6	0.61	-5.1	0.06	12.5	0.86	121.3
2.400	0.93	89.8	0.54	-12.0	0.07	7.3	0.86	116.1
2.600	0.93	83.2	0.48	-18.6	0.07	2.6	0.87	111.4
2.800	0.93	77.3	0.43	-24.0	0.07	-2.6	0.88	106.3
3.000	0.93	71.8	0.39	-29.3	0.07	-7.2	0.89	101.8
3.200	0.92	66.6	0.37	-33.5	0.07	-12.1	0.90	97.2
3.400	0.92	61.8	0.34	-37.2	0.07	-16.8	0.91	92.5
3.600	0.91	57.9	0.32	-40.5	0.07	-21.1	0.91	88.8
3.800	0.90	54.1	0.32	-43.9	0.07	-24.9	0.92	85.1
4.000	0.88	50.5	0.31	-48.3	0.07	-27.6	0.92	81.4
4.200	0.87	47.8	0.32	-53.8	0.07	-31.5	0.92	78.1
4.400	0.86	45.7	0.32	-60.9	0.07	-33.4	0.92	74.9
4.600	0.85	43.4	0.32	-68.9	0.07	-35.4	0.92	72.2
4.800	0.85	42.3	0.31	-77.5	0.07	-37.2	0.92	69.3
5.000	0.86	40.3	0.30	-86.7	0.07	-39.3	0.92	66.2
5.200	0.87	37.7	0.28	-94.5	0.07	-41.6	0.92	63.1
5.400	0.88	35.2	0.26	-101.8	0.07	-44.0	0.92	59.6
5.600	0.89	32.1	0.24	-108.4	0.08	-48.5	0.92	56.6
5.800	0.90	29.0	0.22	-114.5	0.08	-50.3	0.92	53.0
6.000	0.90	26.0	0.21	-121.1	0.08	-54.0	0.93	49.0

2-4 PARAMETER EXTRACTION OF ACTIVE DEVICES

2-4-1 Introduction

We have already seen that all linear models that are used on the market are really derived from large-signal models, and it is our point that the quality of the model depends mostly on the quality of the parameters used to describe the model.

As far as modeling is concerned there are two options. The first option is a physics-based model, which can be applied to items such as diodes and bipolar transistors, including its most advanced versions, such as HBTs, SiGe versions, and other future derivatives. The bipolar transistor is essentially a combination of two diodes whereby the base–emitter junction generates the electrons or holes (depending on how the reader is accustomed to viewing the process) and then transfers those to the collector. The emitter, as its name states, emits the charged particles and the collector collects them, minus some current losses, which are expressed in the current gain of the device. A current gain (β) of 100 means that the collector receives 1% less than the emitter emits.

The second modeling option involves the construction of an analytical mathematic equivalent for the device modeled. Earlier in this chapter, we went through a detailed physics-based derivation about the inner workings of these devices, explaining how, for a given dc bias point, one can simplify them to a small-signal equivalent circuit. This is possible since we assume that the RF current and voltages are less than 1% of the dc equivalent currents and voltages. Reducing a device's complex nonlinear equivalent circuit to a quasistatic linear equivalent circuit allows its manufacturer to generate and publish sets of *S* parameters for it. These parameters are bias, voltage, and temperature dependent. At higher frequencies, parasitics play an enormous role in determining a device's RF performance.

A device operating at signal levels that are more than 1% of its dc equivalent currents and voltages must be evaluated in terms of large-signal performance. Most of the large-signal-equivalent circuits for nonbipolar devices, such as FETs (MOS, silicon JFETs, and members of the GaAsFET family), are analytic approximations that essentially bear no physical insight into the inner workings of the transistors. This fact is particularly painful because to look at high-order intermodulation distortion, one needs to have third and fourth derivatives of mathematically continuous equations. The SPICE approach cannot easily be translated into modern harmonic-balance simulators because such cop-outs as IF THEN ELSE statements in programming provide everything else but a continuous model. In the case of our own work, we used complicated curve-fitting equations to obtain analytic equations and their derivatives to accurately provide a large-signal model that is valid over a wide range of dc and RF.

2-4-2 Typical SPICE Parameters

Since we are about to evaluate bipolar microwave transistors, junction FETs, MOSFETs (model level 3), and GaAsFETs, Tables 2-20, 2-21, 2-22, and 2-23 list typical parameters for the devices we have used. These parameters can be obtained by the Scout program with the appropriate measurements. The BSIM model for MOSFETs, applicable for submicron technology transistors, requires an enormous level of parameter extraction and has not fully been validated for the LDMOS-type transistors currently favored for RF and microwave applications.

Table 2-20 BFR193W BJT

IS = 2.738E – 16	BF = 125	NF = 0.95341
VAF = 24	IKF = 0.26949	ISE = 1.0627E – 14
NE = 1.935	BR = 14.267	NR = 1.4289
VAR = 3.8742	IKR = 0.037925	ISC = 3.7409E – 17
NC = 0.94371	RB = 15	IRB = 0.00091763
RBM = 1.8368	RE1 = 0.76534	RC2 = 0.11938
CJE = 1.1824E – 15	VJE = 0.70276	MJE = 0.48654
TF = 1.8828E – 11	XTF = 0.69477	VTF = 0.8
ITF = 0.00096893	PTF = 0	CJC = 9.3503E – 13
VJC = 1.1828	MJC = 0.30002	XCJC = 0.053563
TR = 1.0037E – 09	VJS = 0.75	MJS = 0
XTB = 0	EG = 1.11	XTI = 3
FCC = 0.72063	LB = 0.57E – 9	LC = 0.00E – 9
LE = 0.43E – 9	CBCP = 0.101E – 12	CCEP = 0.175E – 12
CBEP = 0.061E – 12	VCMX = 10 V	

The meaning and significance of the various parameters is best explored in a book on SPICE or semiconductor physics [22–24].

2-4-3 Noise Modeling

Diode Noise Model. The noise model for the diodes (Figure 2-122 and Table 2-24) consists of two contributions: the shot noise and the flicker noise. The shot noise is computed automatically and does not require any parameters. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications).
2. Using bias-dependent flicker noise coefficients (specifying KF and AF at multiple bias points).

The noise generators in the diode noise model are the series parasitic resistance, R_S, and the intrinsic junction. Figure 2-122 illustrates the intrinsic junction noise generator. Let Δf be the bandwidth (usually normalized to 1 Hz). The intrinsic noise generator has a mean-square value of

$$\langle i_{Dn}^2 \rangle = 2qI_D\,\Delta f + \mathrm{KF}\,\frac{I_D^{\mathrm{AF}}}{f^{\mathrm{FCP}}}\,\Delta f \tag{2-266}$$

Notes on the Diode Noise Model:

1. Shot noise is always present unless the SN parameter is set to zero. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.

Table 2-21 2SK125 JFET[a]

IDSS	= 0.5250E – 01	VP0	= –0.3111E + 01	GAMA	= –0.1867E – 01	E	= 0.1520E + 01
KE	= –0.3856E – 03	SL	= 0.2818E – 01	KG	= –0.2398E + 00	T	= 0.0000E + 00
SS	= 0.7448E – 04	IG0	= 0.2000E – 14	AFAG	= 0.3846E + 02	IB0	= 0.1000E – 04
AFAB	= 0.3800E + 02	VBC	= 0.3000E + 02	R10	= 0.1711E + 02	KR	= 0.0000E + 00
C10	= 0.6609E – 11	K1	= 0.1675E + 01	C1S	= 0.6818E – 33	CF0	= 0.7261E – 11
KF	= 0.1156E + 01	RG	= 0.5000E + 00	RD	= 0.1542E + 01	RS	= 0.1333E + 01
LG	= 0.6098E – 09	LD	= 0.5159E – 08	LS	= 0.1482E – 08	CDS	= 0.4813E – 16
CGE	= 0.1590E – 11	CDE	= 0.3394E – 26	CGSP	= 0.8282E – 13	CDSP	= 0.4832E – 12
ZGT	= 0.5000E + 02	LGT	= 0.4712E – 01	ZDT	= 0.5000E + 02	LDT	= 0.3998E – 01
CGDP	= 0.3653E – 12	ZST	= 0.5000E + 02	LST	= 0.1495E – 01	CGDE	= 0.3831E – 12
CGSB	= 0.3120E – 13	CDSB	= 0.5896E – 12	VDMX	= 10		

[a]The currently implemented model for the junction FET is too primitive for serious RF applications. We have therefore taken the approach (liberty) of using the Materka parameter extraction approach for silicon junction FETs. This has resulted in unparalleled high-quality parameters; in particular, the knee voltage behavior has significantly improved, as well as the overall frequency response.

Table 2.22 GaAs MESFET

IDSS = 0.1077E+00	VP0 = –0.1800E+01	GAMA = –0.5741E-01	E = 0.1290E+01
KE = –0.1155E-01	SL = 0.1652E+00	KG = –0.1782E+00	T = 0.0000E+00
SS = –0.1208E+01	IG0 = 0.2130E-11	AFAG = 0.2740E+02	IB0 = 0.5680E-09
AFAB = 0.1826E+01	VBC = 0.9000E+01	R10 = 0.8382E+01	KR = 0.6359E+00
C10 = 0.5964E-12	K1 = 0.1296E+01	C1S = 0.0000E+00	CF0 = 0.6110E-13
KF = 0.9775E+00	RG = 0.1996E+01	RD = 0.1296E+01	RS = 0.1234E+01
CDS = 0.7852E-13	CDSD = 0.1000E-07	RDSD = 0.1581E+03	CGE = 0.1609E-12
CDE = 0.8674E-13	VDMX = 8		

Table 2.23 The 1 μm × 750 μm level 3 LDMOS FET

CBD = 0.863E-12	CGD0 = 166E-12	CGS0 = 246E-12	GAMA = 0.211
IS = 6.53E-16	KAPA = 0.809	MJ = 0.536	NSUB = 1E+15
PB = 0.71	PBSW = 0.71	PHI = 0.579	RD = 39
RS = 0.1	THET = 0.588	TOX = 4E-8	U0 = 835
VMAX = 3.38E5	VT0 = 2.78	XQC = 0.41	

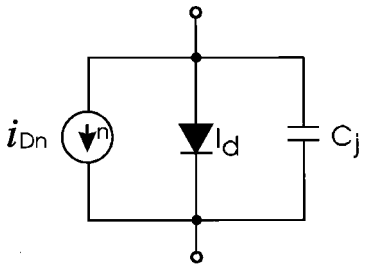


Figure 2-122 Equivalent noise circuit for a diode chip.

Table 2.24 Diode noise model keywords

Keyword	Description	Unit	Default
ID	Required bias current for the data point	Ampere	
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner fequency	hertz	

2. If the value of KF is specified as zero, then the flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.
3. The corner frequency noise model uses the system noise floor to internally compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the diode parameters and *kT*.
4. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

BJT Noise Model. The noise model for the Gummel–Poon BJT model consists of two contributions: shot noise and flicker noise. The shot noise is computed automatically and does not require any parameters. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications) (see Table 2-25).
2. Using bias-dependent flicker noise coefficients (specifying KF and AF at multiple bias points).

Option 1: *Specifying the Bias-Independent Flicker Noise Coefficient*. This option involves the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model.

1. Shot noise is always present unless it is turned off. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying The Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows a bias-dependent flicker noise coefficient (i.e., KF and AF vary with the bias point).

Notes on the BJT Noise Model:

Table 2.25 BJT noise model keywords

Keyword	Description	Unit	Default
IB	Required base bias current for the data point	ampere	
VCE	Required collector–emitter voltage for the data point	volt	
VBS	Base–substrate voltage required for LPNP type when four nodes are used	volt	
VCS	Collector–substrate voltage required for NPN or PNP type when four nodes are used	volt	
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

1. KF, AF, and FC can be specified as bias dependent. If only one set of noise data is specified, the corresponding bias point is not meaningful because all parameters are considered constant over all bias values. However, the bias point is needed for the program to identify the data as bipolar noise data.
2. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
3. This noise model of course considers the actual operating temperature, which must be supplied to the model.

Figure 2-123 shows the BJT noise model. Let Δf be the bandwidth (usually normalized to a 1-Hz bandwidth). The noise generators introduced in the intrinsic device are shown below and have mean-square values of

$$\langle i_{bn}^2 \rangle = 2qI_B \,\Delta f + \text{KF}\,\frac{I_B^{\text{AF}}}{f^{\text{FCP}}}\,\Delta f \tag{2-267}$$

$$\langle i_{cn}^2 \rangle = 2qI_C \Delta f \tag{2-268}$$

$$\langle i_{R_{bb}}^2 \rangle = \frac{4kT}{R_{bb}}\,\Delta f \tag{2-269}$$

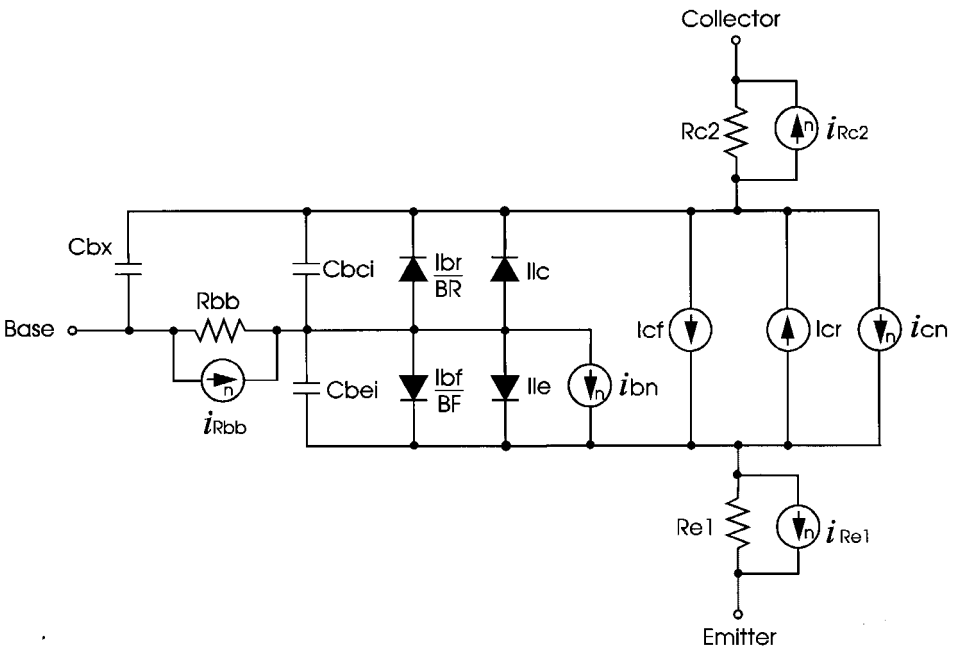


Figure 2-123 BJT noise model (not showing extrinsic parasitics). Current sources with *n* are noise sources.

$$\langle i_{R_{e1}}^2 \rangle = \frac{4kT}{R_{e1}} \Delta f \tag{2-270}$$

$$\langle i_{R_{c2}}^2 \rangle = \frac{4kT}{R_{c2}} \Delta f \tag{2-271}$$

$$I_B = \frac{I_{bf}}{BF} + I_{le} \tag{2-272}$$

$$I_C = I_{cf} - I_{cr} \tag{2-273}$$

JFET and MESFET Noise Models. The noise model for the FETs consists of two contributions: the shot noise and the flicker noise. There are two options to specify noise in the FET model:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (Table 2-26) to determine the flicker noise (this option is usually sufficient for most applications). The shot noise will be automatically computed using the SPICE equation.
2. Using bias-dependent flicker noise coefficients through a reference in the DATA block (specifying KF and AF at multiple bias points) and specifying the four noise parameters [$F_{\min}$, Γ_{opt} (magnitude and phase), and R_n] at multiple bias points.

Option 1: *Specifying the Enhanced SPICE Noise Model.* Option 1 is the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model. The drain noise model has the form

$$\langle |I_{dn}|^2 \rangle = 4K_B T \frac{2g_m}{3} \Delta f + \text{KF} \frac{|I_D|^{\text{AF}}}{f^{\text{FCP}}} \Delta f \tag{2-274}$$

Table 2.26 FET noise model keywords

Keyword	Description	Unit	Default
FN	Noise data measurement frequency	hertz	1.0 GHz
VGS	Required gate–source voltage for the data point	volt	
VDS	Required drain–source voltage for the data point	volt	
FMIN	Required minimum noise figure in dB at FN		
MGO	Required magnitude of optimum noise reflection coefficient at FN		
PGO	Required phase of optimum noise reflection coefficient at FN		
RN	Required normalized noise resistance at FN		
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

where the shot noise is derived from g_m and the flicker noise is proportional to KF and the drain channel current, I_D, and inversely proportional to frequency. The AF and FCP parameters tailor the flicker noise dependence on bias and frequency, respectively.

1. Shot noise is always present unless it is turned off. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, then flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying the Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows the specification of the complex bias-dependent nature of the shot noise and flicker noise. At high frequencies, the equivalent noise sources are correlated (the SPICE noise model does not account for this correlation). The complete evaluation of the shot noise sources can be determined from the four noise parameters. Since these are functions of bias, they can be specified over the (V_{GS}, V_{DS}) bias plane.

Additionally, a bias-dependent flicker noise coefficient (i.e., KF and AF vary with current) can be specified.

The MESFET noise model uses the four measured noise data [$F_{\min}$, Γ_{opt} (magnitude and phase), and R_n] at one frequency and multiple arbitrary bias points. The program uses these data and the FET model parameters to de-embed the noise data to an intrinsic noise model. The intrinsic model is accurate at all frequencies and, therefore, can predict the noise performance at all frequencies given data at just one frequency point. Built-in bias-dependent characteristics are used if multi-bias noise data are not provided.

Notes on the FET Noise Model

1. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
2. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

Noise in a MESFET is produced by sources intrinsic to the device. The same approach, but with different flicker corner frequencies, is highly applicable to JFETs and MOSFETs. For more detail as to simulation, see the Element library book for the active-device portion of Ansoft's Serenade Design Environment product. The equivalent noisy circuit of an intrinsic FET is represented in Figure 2-124.

The intrinsic FET is internally represented as a noiseless nonlinear two-port with one equivalent noise current connected across the gate–source terminal, and one across the drain–source terminal. The correlations of the gate and drain noise current sources are

$$\langle |I_{gn}|^2 \rangle = 4K_B T \, \Delta f \frac{\omega^2 C_{gs}^2}{g_m} R \tag{2-275}$$

$$\langle |I_{gn}|^2 \rangle = 4K_B T \, \Delta f \, g_m P \tag{2-276}$$

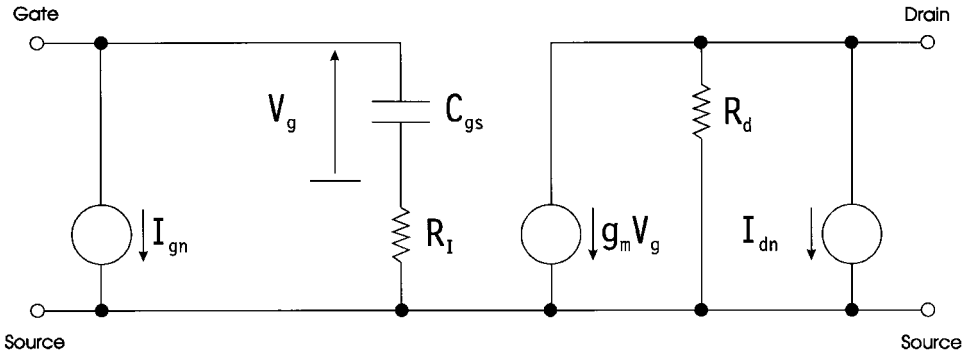


Figure 2-124 Equivalent noise circuit of an intrinsic FET device.

$$\langle I_{gn} I_{dn}^{*} \rangle = 4K_B T\, \Delta f\, j\omega C_{gs} \sqrt{PR}\, C \tag{2-277}$$

The correlation matrix of the noise current sources is

$$C_{dc}(\omega) = \frac{2}{\pi} K_B T\, d\omega \begin{bmatrix} \dfrac{\omega^2 C_{gs}^2}{g_m} R & -j\omega C_{gs}\sqrt{PR}\, C \\ j\omega C_{gs}\sqrt{PR}\, C & g_m P \end{bmatrix} \tag{2-278}$$

The gate and drain noise parameters R and P and the correlation coefficient C are related to the physical noise sources acting in the channel and are functions of the device structure and bias noise parameters. By defining measured noise parameters, $F_{\min}$, R_n, and Γ_{opt}, and using a noise de-embedding procedure, the parameters R, P, and C and the intrinsic noise correlation matrix of a FET device as functions of device bias are determined by the program.

In addition to the noise sources shown above, the flicker (1/f) noise can also be modeled by means of a noise current source connected in parallel with the intrinsic drain port. The flicker noise component in a narrow band, Δf, is expressed in the form

$$\langle |I_f|^2 \rangle = \mathrm{Q}\, \Delta f \frac{|I_D|^{\mathrm{AF}}}{f^{\mathrm{FCP}}} \tag{2-279}$$

where I_D is the instantaneous value of the channel current, and Q, AF, and FCP are empirical parameters. In most practical cases, AF and FCP are directly obtained from measurements (typically, AF = 2 and FCP = 1), while Q is not. In Ansoft's Serenade Design Environment, Q is either provided directly using KF or is computed by providing the flicker corner frequency (FC). FC is the frequency at which the flicker noise equals the shot/diffusion noise. The corner frequency is defined by the equation

$$\mathrm{Q} \frac{|I_D|^{\mathrm{AF}}}{f_{\mathrm{C}}^{\mathrm{FCP}}} = g_m P \tag{2-280}$$

Given the corner frequency FC and the measurement bias point V_{gs} and V_{ds}, the program automatically computes I_D, g_m, and P, and finally Q.

More information on FET noise modeling can be found in the literature [25–31].

MOSFET Noise Model. The MOSFET noise model (Figure 2-125 and Table 2-27) consists of two contributions: the shot noise and the flicker noise. The shot noise is computed automatically and does not require any parameters. It can be turned off by specifying SN = 0. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications).

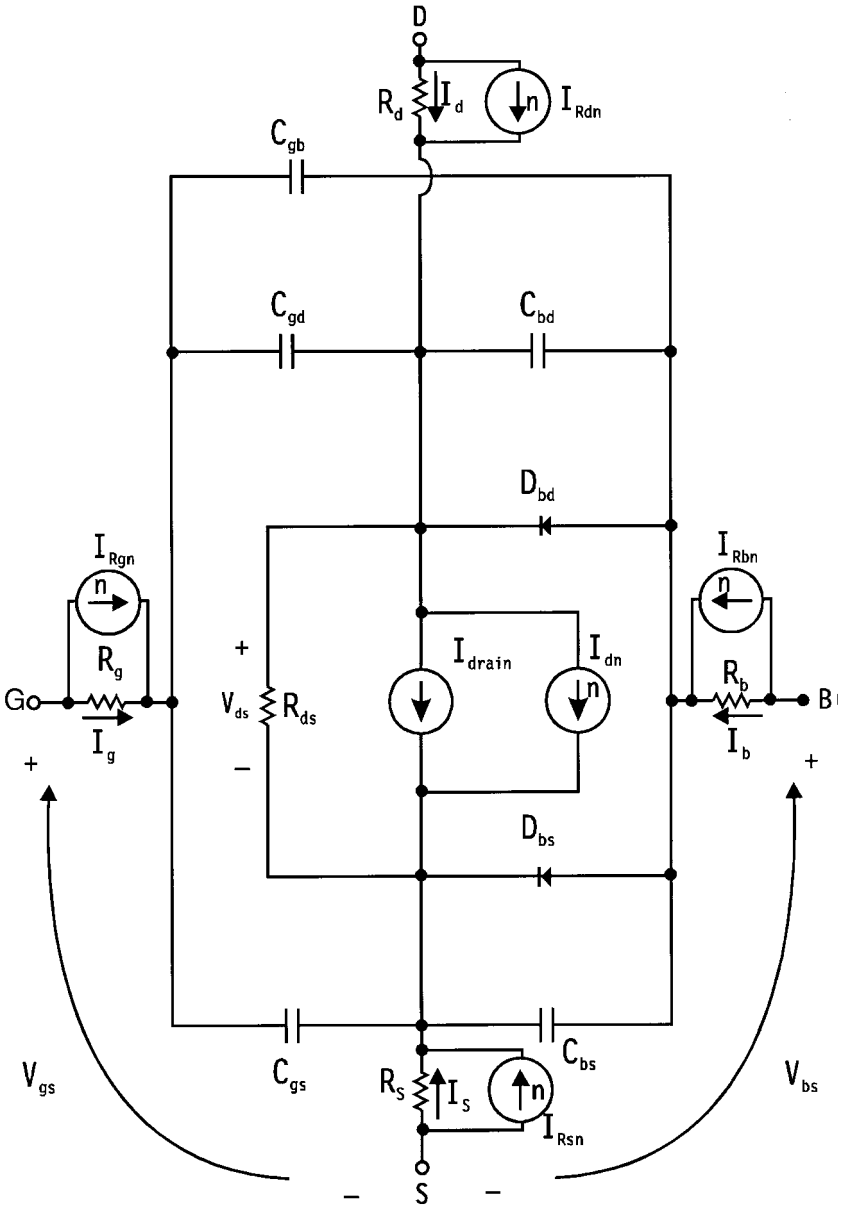


Figure 2-125 Equivalent noise circuit of an intrinsic MOSFET device.

Table 2.27 MOSFET noise model keywords

Keyword	Description	Unit	Default
VGS	Required gate–source bias for the data point	volt	
VDS	Required drain–source for the data point	volt	
VBS	Required drain–bulk for the data point	volt	
KF	Flicker noise coefficient		1.0E–13
AFT	Bias exponent of the flicker noise model		2.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

2. Using bias-dependent flicker noise coefficients through a reference (specifying KF and AF at multiple bias points).

Option 1: *Specifying the Enhanced SPICE Noise Model.* This option is the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model (the flicker noise is considered bias dependent).

1. Shot noise is always present unless the SN parameter is set to zero. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, then the flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying the Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows a bias-dependent flicker noise coefficient (i.e., KF and AF vary with drain current). The MOSFET noise model data are given and referenced by a model parameter.

Notes on the MOSFET Noise Model

1. KF, AF, and FC can be specified as bias dependent. If only one set of noise data is specified, the corresponding bias point is not meaningful because all parameters are considered constant over all bias values. However, the bias point is needed for the program to identify the data as MOSFET noise data.
2. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
3. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

Let Δf be the bandwidth (normalized to 1 Hz). The noise generators introduced in the intrinsic device are shown below and have mean-square values of

$$\langle i_{dn}^2 \rangle = \frac{8kTg_m}{3}\Delta f + \mathrm{KF}\frac{I_D^{\mathrm{AF}}}{f^{\mathrm{FCP}}}\Delta f$$

$$\langle i_{Rgn}^2 \rangle = 4\frac{kT}{R_g}\Delta f$$

$$\langle i_{Rdn}^2 \rangle = 4\frac{kT}{R_d}\Delta f$$

$$\langle i_{Rsn}^2 \rangle = 4\frac{kT}{R_s}\Delta f$$

$$\langle i_{Rbn}^2 \rangle = 4\frac{kT}{R_b}\Delta f$$

We include this MOSFET noise model (used for quite awhile) for completeness. At the moment, we do not know which MOSFET noise model the industry will settle on in the future.

Modern CAD tools, such as Ansoft's Serenade product, use these models to generate quite accurate noise data based on a good linear equivalent model. Internally Serenade uses the noise-correlation-matrix method (first introduced by Russer). The same noise-correlation techniques apply to BJTs, all FETs except MOS, and HBTs (such as SiGe devices).

2-4-4 Scalable Device Models

Since diodes and transistors are scalable, here are guidelines for how to scale them:

Microwave Diode

$$I_D = Area \times I_D$$
$$C_j = Area \times C_j$$
$$R_D = R_D/\text{Area}$$

PIN Diode

$$I_D = \text{Area} \times I_D$$
$$C_j = \text{Area} \times C_j$$
$$R_S = R_S/\text{Area}$$
$$R_{\max} = R_{\max}/\text{Area}$$

Bipolar

$$I_{bf} = \text{Area} \times I_{bf} \qquad I_{br} = \text{Area} \times I_{br}$$
$$I_{le} = \text{Area} \times I_{le} \qquad I_{lc} = \text{Area} \times I_{lc}$$
$$I_{cf} = \text{Area} \times I_{cf} \qquad I_{cr} = \text{Area} \times I_{cr}$$
$$C_{bc} = \text{Area} \times C_{bc} \qquad C_{be} = \text{Area} \times C_{be}$$
$$C_{bx} = \text{Area} \times C_{bx} \qquad R_{bb} = R_{bb}/\text{Area}$$
$$\text{RB2} = \text{RB2/Area} \qquad \text{RC2} = \text{RC2/Area}$$
$$\text{RE1} = \text{RE1/Area} \qquad I_{jss} = I_{jss} \times \text{Area}$$
$$C_{jss} = C_{jss} \times \text{Area}$$

Materka FET

$$\begin{aligned}
\text{IGSS} &= \text{IGSS} \times \text{Area} \\
\text{CGS0} &= \text{CGS0} \times \text{Area} \\
\text{CGS1} &= \text{CGS1} \times \text{Area} \\
\text{CDVC} &= \text{CDVC} \times \text{Area} \\
\text{CDVS} &= \text{CDVS} \times \text{Area} \\
R_G &= R_G \times \text{Area}/(\text{Number of fingers} \times 2) \\
R_D &= R_D/\text{Area} \\
R_S &= R_S/\text{Area}
\end{aligned}$$

Generating a Databank for Parameter Extraction. Given the fact that one can measure with an automated system (Rohde & Schwarz, Hewlett-Packard) the bias-dependent *S* parameters of any active device, bipolar and FET, and furthermore, that Hewlett-Packard makes a computer-controllable dc *I–V* curve tracer, we can generate a huge databank that allows us to have corresponding values for both RF and dc operating points. To validate this, one needs to run a large-signal simulator, such as Ansoft's Serenade Design Environment, or its equivalent from other manufacturers, enter the large-signal parameters, and generate a set of *S*-parameters that correspond to the dc values for which we have measurements. The agreement between measured and predicted points for both RF and dc values is a good measure to evaluate the accuracy of the parameter extraction and the simulator's performance. We cannot stress often enough, however, the fact that while most simulator models are quite good, the weakest point of the link is the parameter extraction program.

Currently, the best parameter extraction for bipolar transistors and FETs, excluding the BSIM model, is implemented in the Scout program by Ansoft, based on modified techniques found in the literature. These techniques are based on the fact that present-day nonlinear modeling of microwave devices, especially FETs, is not exactly adequate to describe all effects found within these devices. In the case of FETs, the most troublesome parameter is R_{DS}, which is really more bias dependent than practically all existing models take into consideration. Techniques used in published extraction methods extract nonlinear models independently for each section of a device model, and then assume that the proper response will be generated when all parts are put together. For instance, many extraction techniques independently fit the *I–V* data of the section describing the *I–V* curves and then fit the *C–V* data to the equivalent small-signal models. This results in an approximation to the bias-dependent response of the complete model.

In the technique used in the Scout program, model parameters are extracted simultaneously so the effects between model sections are accounted for and the device is treated as a whole. The most critical area happens to be in the linear region, typically around the pinchoff area, and for very small supply voltages. The Scout techniques use measured data taken at a wide variety of bias ranges and frequencies to define the model response. The possible solution region for the complete model becomes well identified and uniqueness of solution is vastly improved. The Scout program is interactive in nature and displays simultaneously measured versus predicted data. This allows for fine-tuning of the parameters.

Over the years, the industry has adapted a large number of models. Those supported by the Scout parameter extraction program include:

- *npn* Gummel–Poon BJT
- *pnp* Gummel–Poon BJT

- *npn* heterojunction BJT
- *pnp* heterojunction BJT
- General-purpose *n*-channel JFET
- General-purpose *p*-channel JFET
- Chalmers (Angelov) MESFET/HEMT
- Curtice–Ettenberg cubic MESFET/HEMT
- Curtice quadratic MESFET/HEMT
- IAF (Berroth) MESFET/HEMT
- Modified Materka–Kacprzak MESFET/HEMT
- Raytheon (Statz) MESFET/HEMT
- TriQuint (TOM-1 and TOM-2) MESFET/HEMT
- Ansoft physics-based MESFET

The Ansoft physics-based MESFET model is unique and politically sensitive because its users are essentially required to enter all device fabrication parameters to get the appropriate answer, and most, if not all, companies will stay away from handing out such company proprietary recipes. The validation of this model was done using manufacturing data for transistors that were developed under government contract and therefore accessible for the team members. Compact Software/Ansoft over the last ten years has been involved in a large number of government-funded research and development programs, and practically all modern modeling implementations and their validation were done in conjunction with the largest semiconductor manufacturers.

As far as further validation is concerned, besides our looking at the matching of dc *I–V* curves and *S* parameters, important parameters such as 1-dB compression point, power-added efficiency, and harmonic content are important criteria to evaluate the quality of such an undertaking. As far as hardware is concerned, one needs to have a network analyzer, such as the HP8510 or its equivalent. Its selection is determined by the highest cutoff frequency where measurements are necessary. A dc feed capability, up to several amperes, through the measuring ports, and appropriate power supplies are essential, as is a dc *I–V* curve tracer (HP4145). The built-in various optimizers of the program establish the best possible match between the selected parameters and the measured results.

The area of proper modeling has always been fascinating because the active-device model can make or break first-pass success. Figure 2-126 shows the doubler gain comparison of a model supplied by NEC (the device manufacturer, parameter extraction done with Scout, and parameter extraction supplied from the *Microwave Engineering Europe* (*MEE*) magazine in its CAD review of May 1994). It is somewhat unclear where the *MEE* model came from. The actual task given by *MEE* was to simulate a frequency doubler against measured data. This type of simulation required two areas of high precision, one being the model and the other being the electromagnetic modeling of the discontinuities of the circuit—in this case, a stub element. The circuit itself was very simple. The manufacturer-supplied data gave the poorest results (we believe the parameter extraction came from a third party not quite up to speed in this area). The results that came from the Scout parameter extraction and the third one that *MEE* supplied were fairly close, yet the Scout solution, including the electromagnetic simulation part, gave the best answer. More simulators have since appeared on the market, and it would be interesting to revisit this topic. Figure 2-127 shows a physical layout of the doubler circuit.

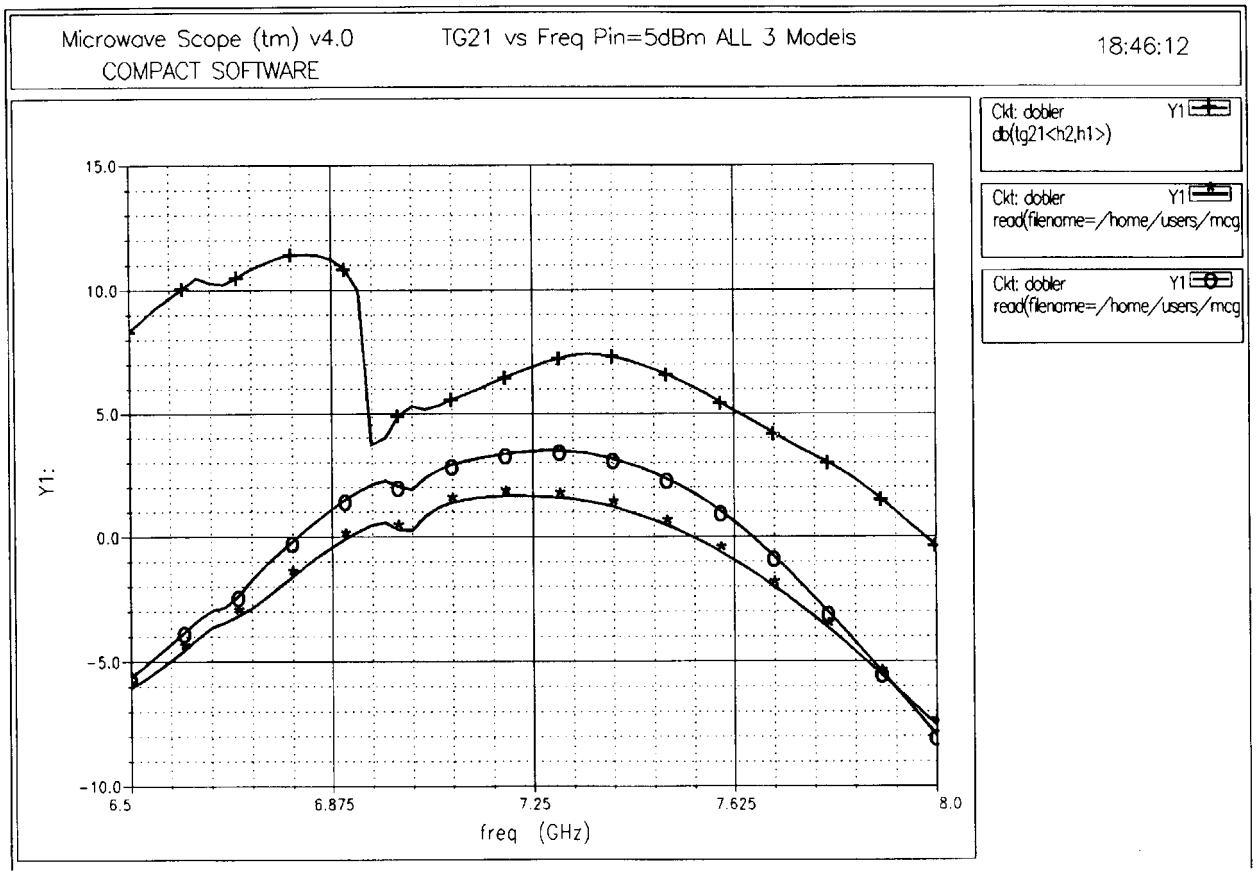


Figure 2-126 Gain comparison of a MESFET doubler based on the three device models used: NEC (+), Ansoft (O), and *Microwave Engineering Europe* (*).

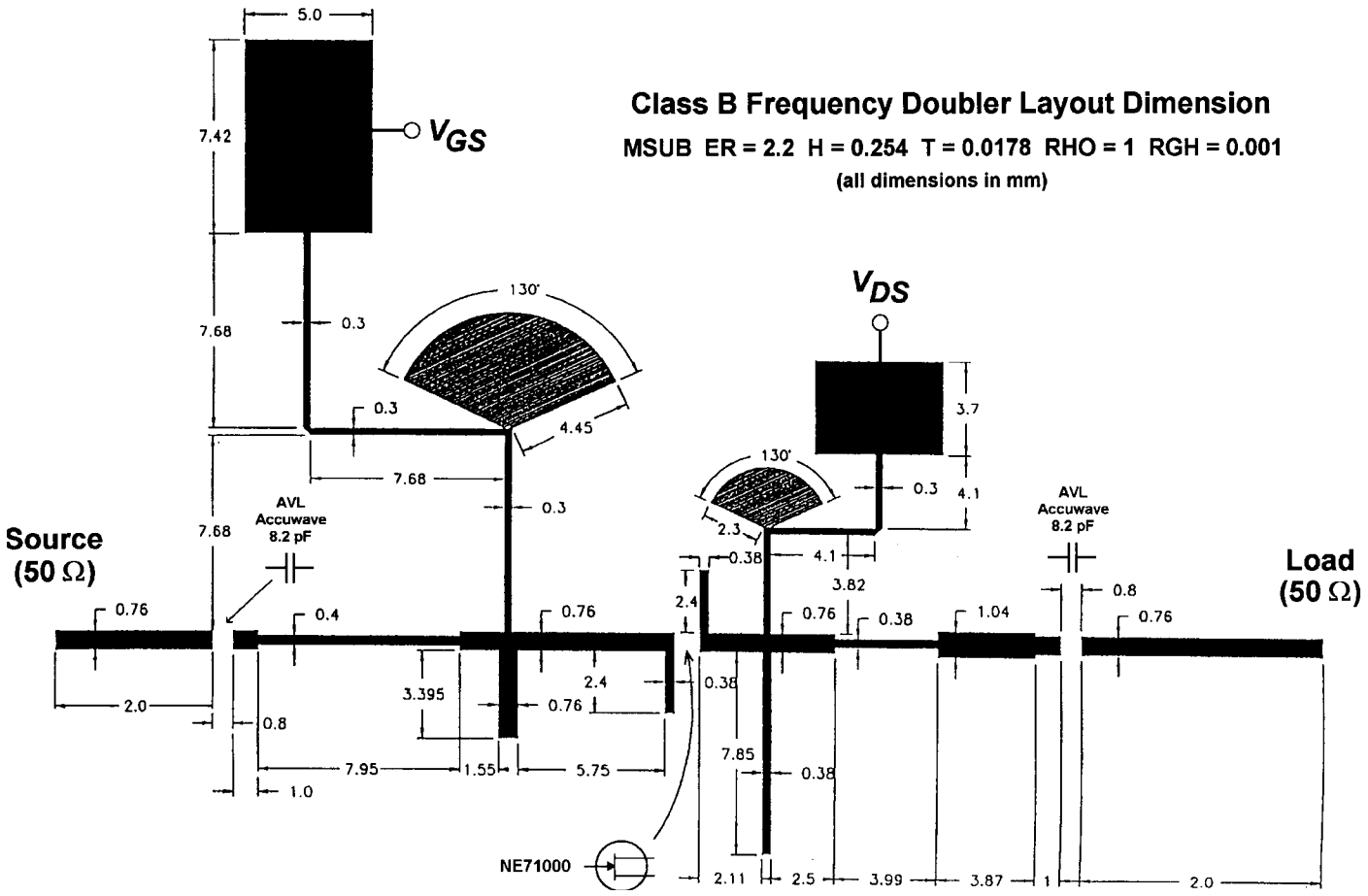


Figure 2-127 Layout of the MESFET doubler.

Test Setup. The measuring equipment required for the actual extraction includes:

- Semiconductor parameter analyzer (HP4145)
- Vector network analyzer (ZVR, HP8510, W360)
- Bias supplies and meters
- Power RF source, power meters, spectrum analyzer, and so on

The required software tools include:

- GPIB controller: Compact NETCOM (network analyzer control program)
- Parameter extraction performed by Scout
- Small-signal and harmonic-balance validation: Serenade nonlinear analysis (Microwave Harmonica) or its equivalent

After we have assembled all these tools, the device extraction process can begin. It will follow the flow as shown in Figure 2-128. Depending on the numbers for the dc/ac measurement, this process can take up to two hours. The Scout program, with its built-in optimizer, can take another two hours. It is therefore reasonable to assume a thorough parameter extraction of a device will take about one day.

Parameter Extraction. As we have outlined, the parameter extraction program Scout, which is optimized for high-frequency and millimeter-wave application, is a Windows-based parameter extraction and optimization program. It requires a network-communication program to acquire dc and *S* parameter measurement data. It supports interactive device fitting and optimization and supports the various models listed earlier. For ease of use, it has an attractive graphical user interface with multiple windows, as can be seen from Figure 2-129. Hewlett-Packard offers more general RF parameter extraction programs, but we did not have access to those. We also expect that some other companies will develop similar tools. An interesting program named HarPE, developed by Professor John Bandler in Canada, may have been withdrawn from the market after his company merger with HP. The program was acclaimed to provide quite good results using measured harmonics. One of the unfortunate deficiencies of most BJT parameter extraction programs known to us is the fact that (as one

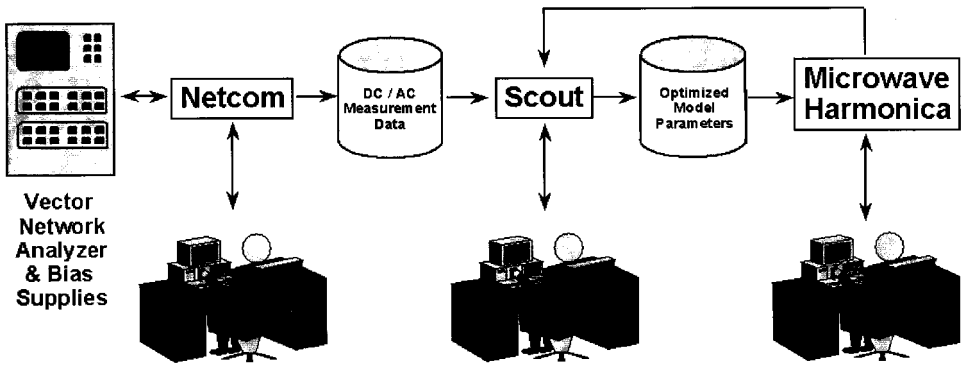


Figure 2-128 Flowchart of the parameter extraction process.

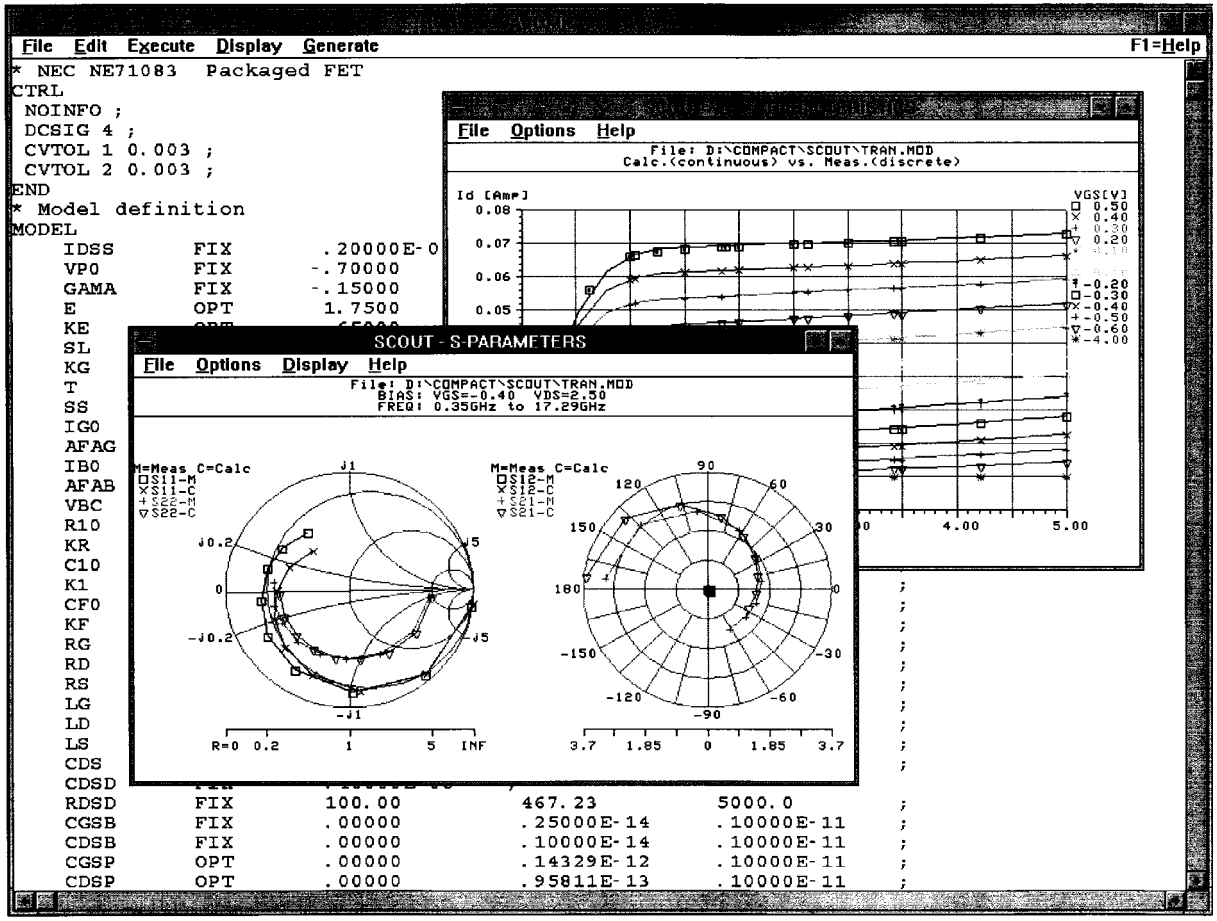


Figure 2-129 The Scout user interface.

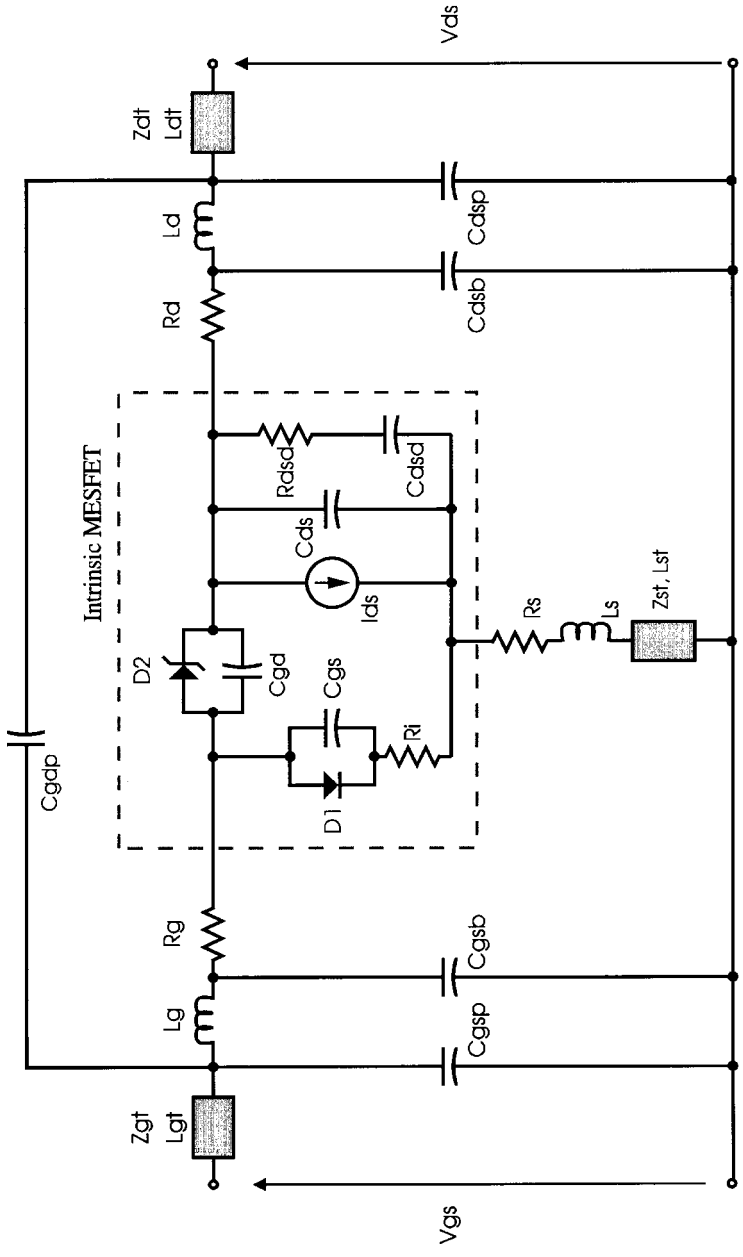


Figure 2-130 Intrinsic model and complete chip/package model.

example) the sensitivity to determine the actual base-spreading resistor for bipolar transistors is poor. The only useful workaround is considering the noise figure of the device, given the fact that the midrange noise factor of any bipolar transistor can be calculated from

$$F = 1 + \frac{R'_{bb}}{R_g} + \frac{0.5(26\ \text{mV}/I_C)}{R_g}$$

(approximate equation at medium frequencies) (2-281)

By solving the equation for R'_{bb}, since we know that R_g, the generator impedance, is 50 Ω, and the emitter diffusion resistance (26 mV/I_C) can be determined from the bias point, we can already eliminate the need to extract R'_{bb} from the HP test equipment. By using HP's program, errors of a factor up to 10 are not uncommon because of the low sensitivity for this parameter. Another headache in the case of the BJT is the determination of the excess phase (PTF) and forward and reverse transit time (TF and TR, respectively). The lack of quality modeling can best be found by examining the measured versus predicted values of S_{22}. Again, practically all the parameter extraction programs we have seen so far suffer from this deficiency.

MESFETs. In the case of a member of the MESFET family, a model as shown in Figure 2-130 is used. It consists of an intrinsic model and a complete chip/package parasitic model. In the case of the MESFET, to obtain first-class results, the following measurements are necessary:

- Dc *I*–*V* measurements
- Dc diode measurements
- Cold-FET and forward-biased-gate *S* parameters
- *S*-parameter measurements over bias
- Harmonic power measurements
- *Z*-Parameter Equations for Figure 2-131

In order to simplify the modeling, it is useful to do cold-FET [Figure 2-131 and Eqs. (2-282), (2-283), and (2-284)] and pinched-FET [Figure 2-132 and Eqs. (2-285), (2-286), and (2-287)] measurements to derive some of the intrinsic or extrinsic values.

$$Z_{11} = R_s + R_g + R_{gs} + \alpha_g R_{ch} + j\omega(L_g + L_s) \tag{2-282}$$

$$Z_{12} = Z_{21} = R_s + \alpha R_{ch} + j\omega L_s \tag{2-283}$$

$$Z_{22} = R_s + R_d + 2\alpha R_{ch} + j\omega(L_s + L_d) \tag{2-284}$$

Capacitance Equations for Figure 2-132

$$C_b = -\frac{\text{Im}(Y_{12})}{\omega} \tag{2-285}$$

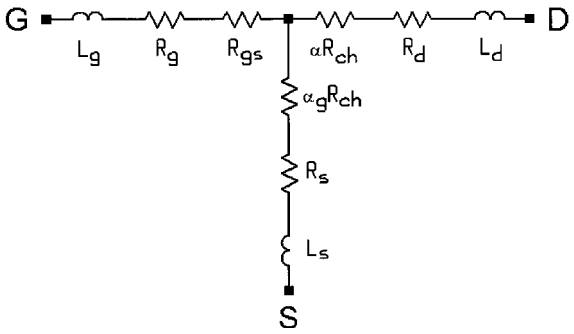


Figure 2-131 Forward-biased gate FET model for cold-FET (V_{DS} = 0, V_{GS} 0) measurements. This approach allows direct extraction of parasitic resistances and inductances.

$$C_{pg} = \frac{\mathrm{Im}(Y_{11})}{\omega} - 2C_b \tag{2-286}$$

$$C_{pd} + C_{ds} = \frac{\mathrm{Im}(Y_{22})}{\omega} - C_b \tag{2-287}$$

For the highest accuracy, the following extraction strategies are recommended:

- Pre-extraction using basic dc and *S* parameters—for example, C_{GS} from S_{11}
- Six extraction steps to determine parameters:
  Basic dc model
  Advanced dc model
  Basic ac model
  Advanced ac model
  Package determination
  Diode characterization
- Large-signal validation
  Power compression and bias shift
  Harmonic power comparison
  Intermodulation characterization

As an example, we are going to use the popular and well-characterized NE71000, which is not a particularly modern device, but has been used for various research projects. Figure

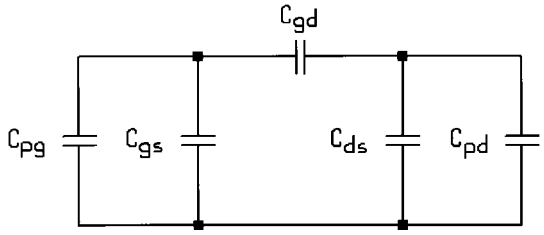


Figure 2-132 Pinched-FET (V_{DS} = 0, V_{GS} *VP*) model. This approach allows direct extraction of parasitic gate and drain capacitances.

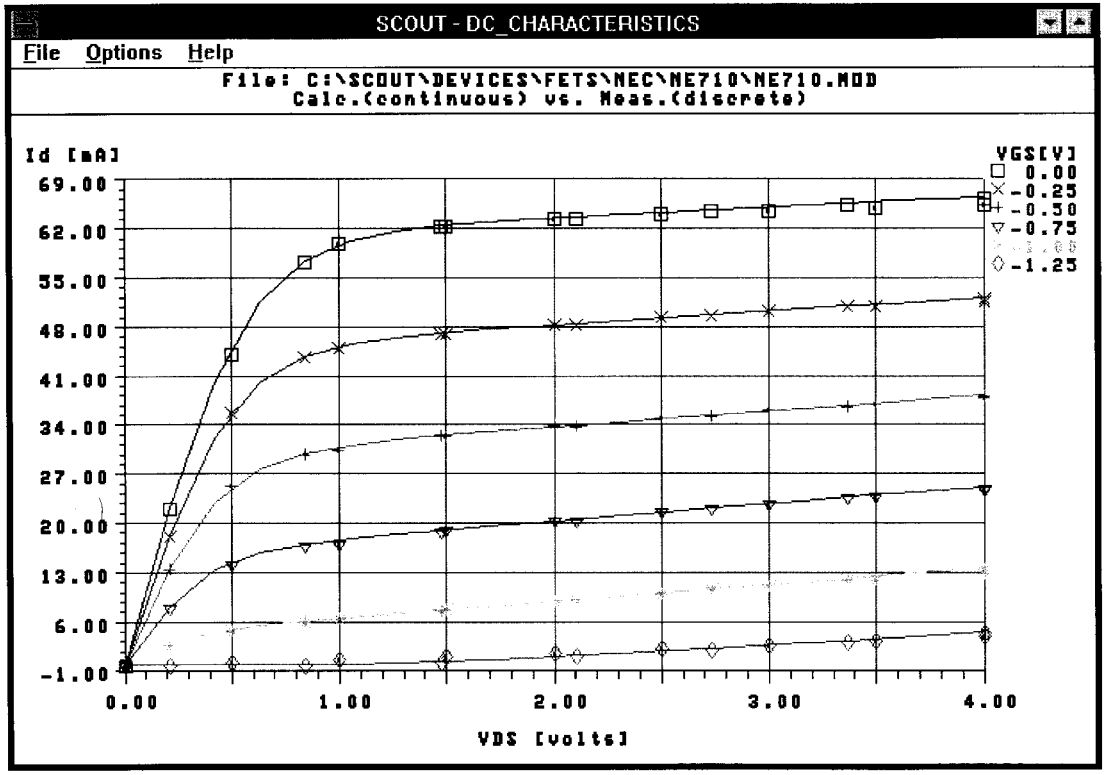


Figure 2-133 Measured NE71000 dc *I–V* curves versus those simulated using the Ansoft Materka model.

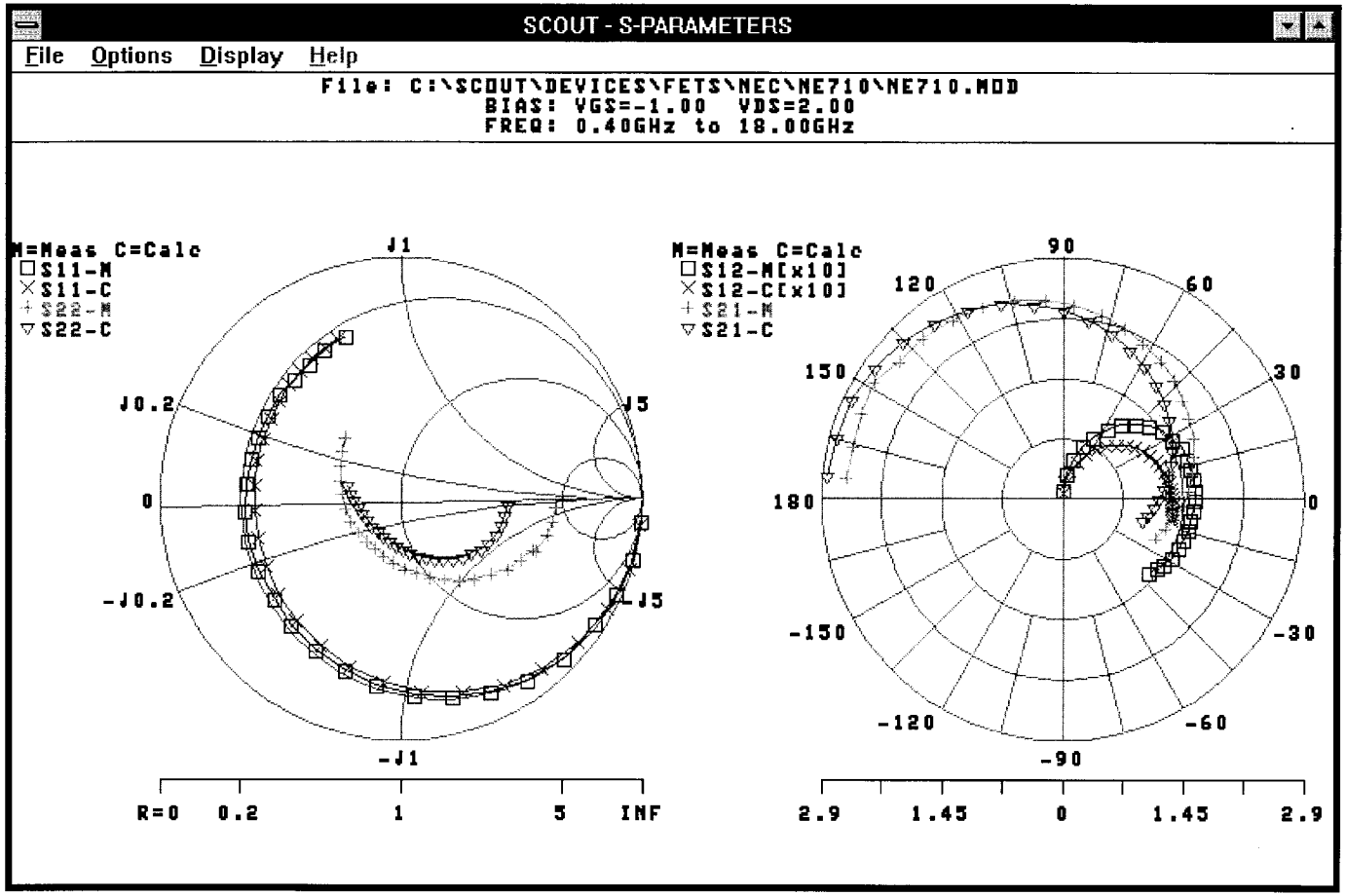


Figure 2-134 Measured versus calculated NE71000 *S* parameters for $V_{GS} = -1.00$ and $V_{DS} = 2.00$.

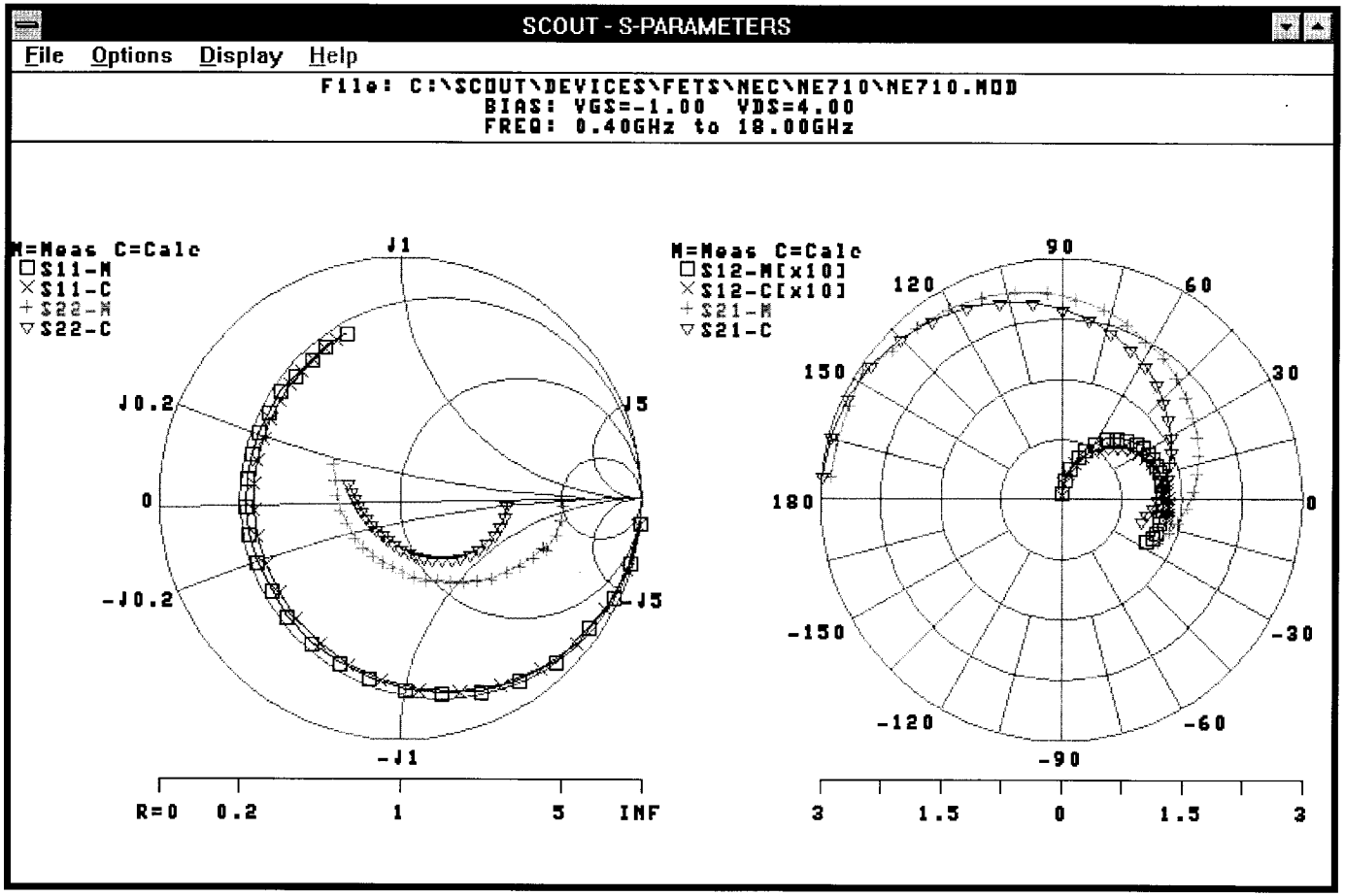


Figure 2-135 Measured versus calculated NE71000 *S* parameters for V_{GS} = –1.00 and V_{DS} = 4.00.

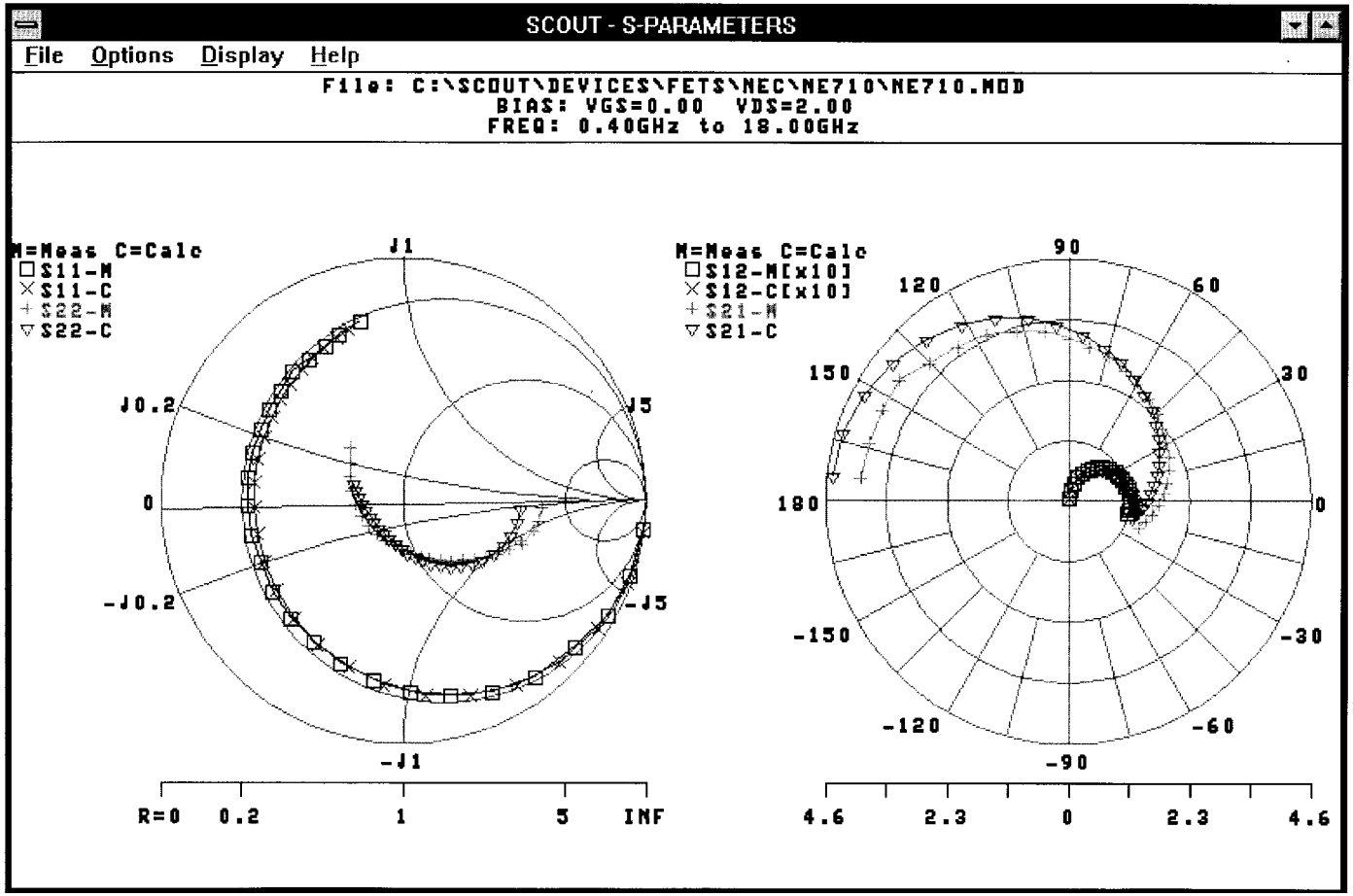


Figure 2-136 Measured versus calculated NE71000 *S* parameters for V_{GS} = 0.00 and V_{DS} = 2.00.

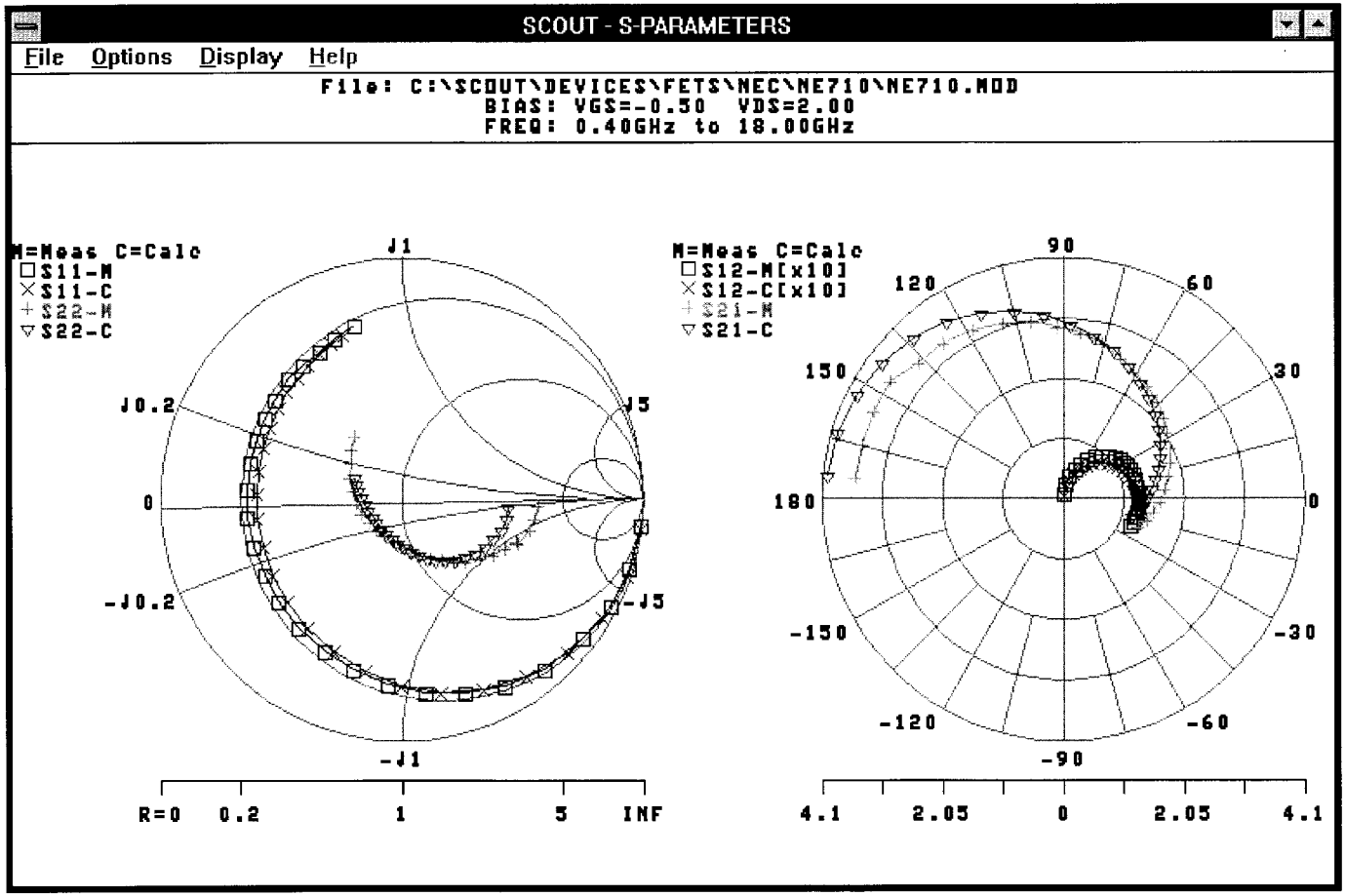


Figure 2-137 Measured versus calculated NE71000 *S* parameters for V_{GS} = –0.50 and V_{DS} = 2.00.

2-133 shows measured versus simulated dc I–V curves for the NE71000. The agreement between measured and predicted dc I–V curves is quite good. In a similar fashion, we obtain the S-parameter fits over bias. They are shown in Figures 2-134 through 2-137.

As mentioned above, 1-dB compression, power saturation, and drain current are functions of the drive. This particular model point is not particularly impressive. We believe this has to do with the original Materka model.

Having brought up the subject of models, Figure 2-138 compares measured data with that obtained with the Angelov, cubic, quadratic, TriQuint, Raytheon, and Materka models. The figure shows that there is no perfect model yet. To put this in perspective, in Figure 2-139 we show the dc I–V error bars for these six models. This graph shows that the least errors for the dc measurement versus modeling are obtained with the Curtice cubic model, followed by the modified Materka model (implemented in Serenade) and TriQuint models. The difference in S parameter matching for RF application is significantly more relevant for the designer.

Things become much more interesting if we compare the S-parameter errors between the Materka shown in Figure 2-140, the Curtice cubic model (Figure 2-141), and the TOM model (Figure 2-142). Improved modeling regarding the R_{DS} is being developed, and we will discuss this at the end of this chapter.

Since the trend is to go to low-voltage power RF operation, several issues must be battled:

- Models are most inaccurate at low V_{DS} and near pinchoff.
- Modeling of low-voltage operation is adversely impacted.
- Thermal effects must be considered, but older models do not account for this.

A Case Study. As a case study, we will consider the Oki KGF1608 power MESFET (0.5 μm × 28 mm). Figures 2-143 and 2-144 show the dc I–V curves and S parameters, respectively. The resolution of the S parameters, based on the scale used, is not particularly impressive.

Figure 2-145 shows the dc I–V curves, with load line, for the KGF1608; the self-heating effect of the transistor can be seen in the droop of the top two curves. Figures 2-146 and 2-147 show drain current and output power versus drive, respectively. The test conditions for these three figures were f = 850 MHz, V_{DS} = 3.4 V, and V_{DS} = 2.025 V.

2-4-5 Conclusions

From all plots of dc and ac error contours, dc I–V curves, and dc error bar charts we can draw the following conclusions.

Dc errors

- *In the Saturation Region*. All the models considered work very well, with a relative error of less than 15%. This means that all models are suitable for typical applications where the device operates in the saturation region.
- *In the Linear Region.* Though the errors in the saturation region are similar for different models, the errors in the linear region are quite different. The Materka, Curtice cubic, and TriQuint models behave better than other models in the linear region. The relative errors of these three models are less than 10%, while other models have more than 60% relative errors at some points.

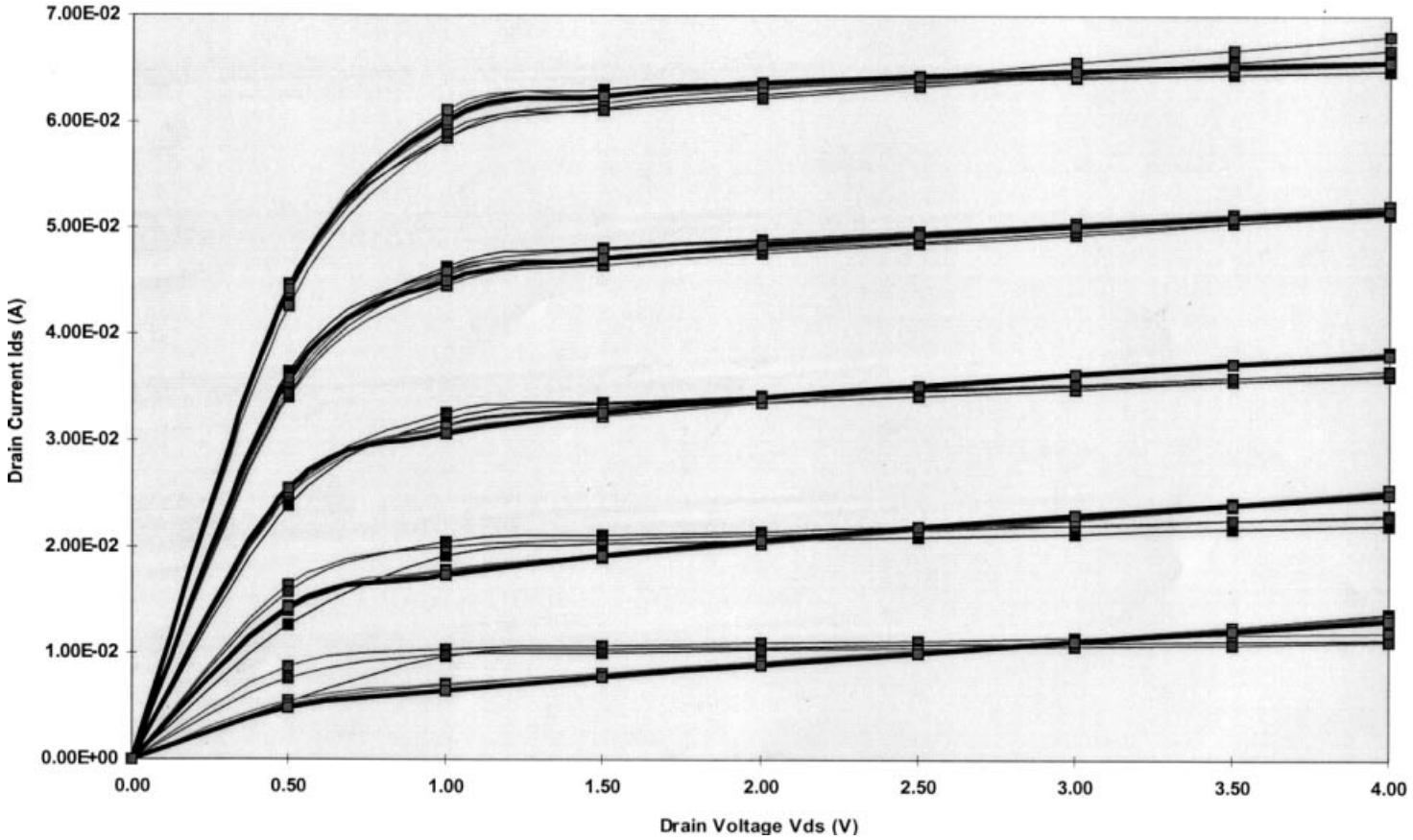


Figure 2-138 Comparison of dc models: Chalmers (Angelov), Curtice cubic, Curtice quadratic, Materka, Raytheon, and TOM (TriQuint).

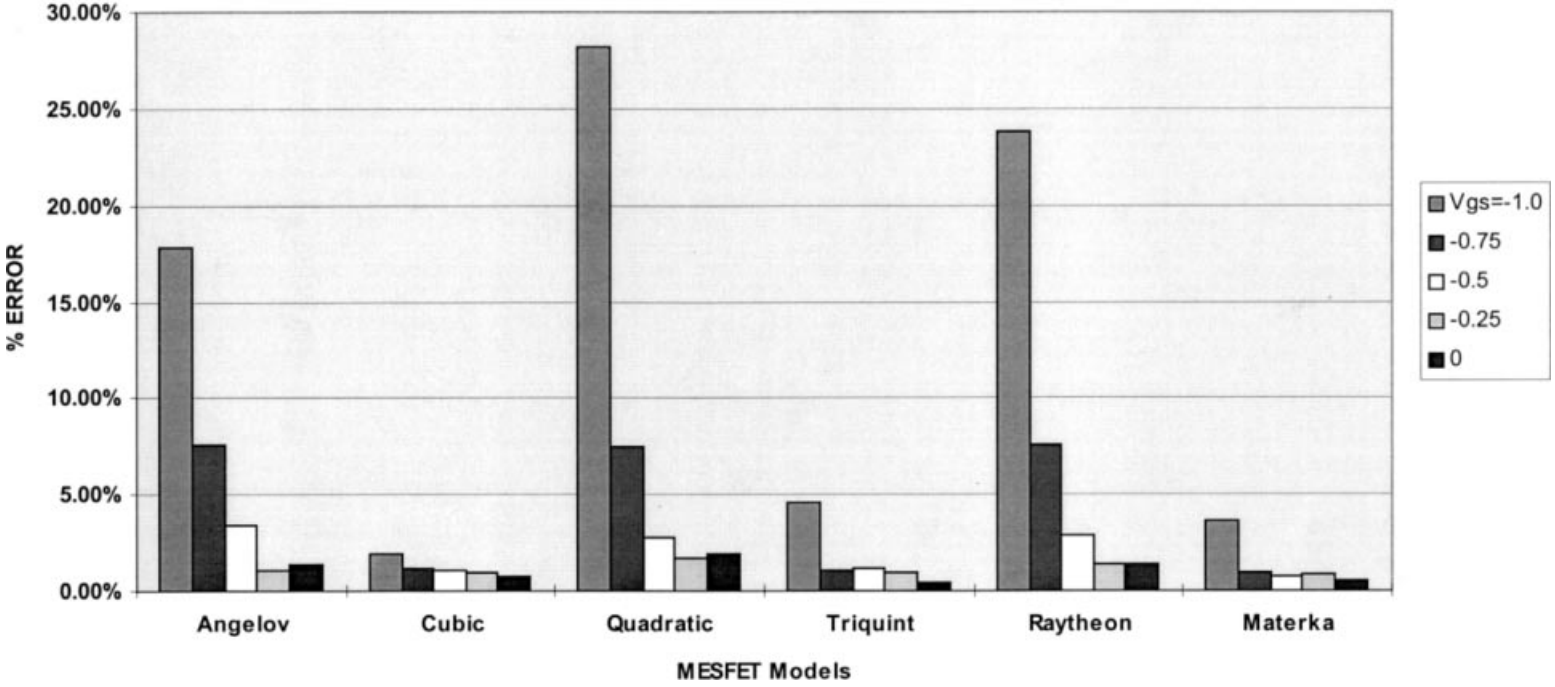


Figure 2-139 Dc *I–V* error bars: Chalmers (Angelov), Curtice cubic, Curtice quadratic, Materka, Raytheon, and TOM.

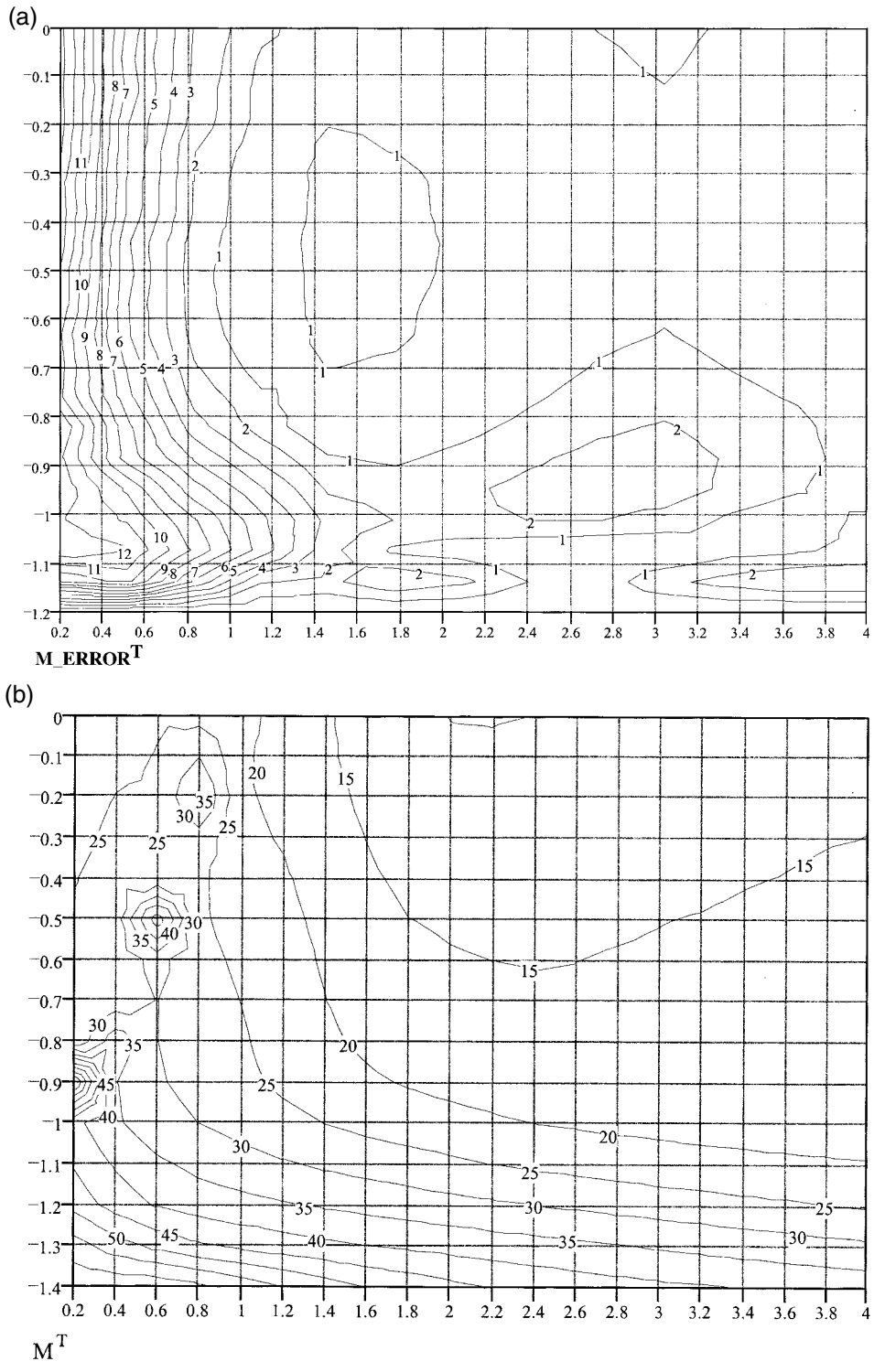


Figure 2-140 NE71000, Materka model, % error contours for (a) dc *I–V* and (b) *S* parameters.

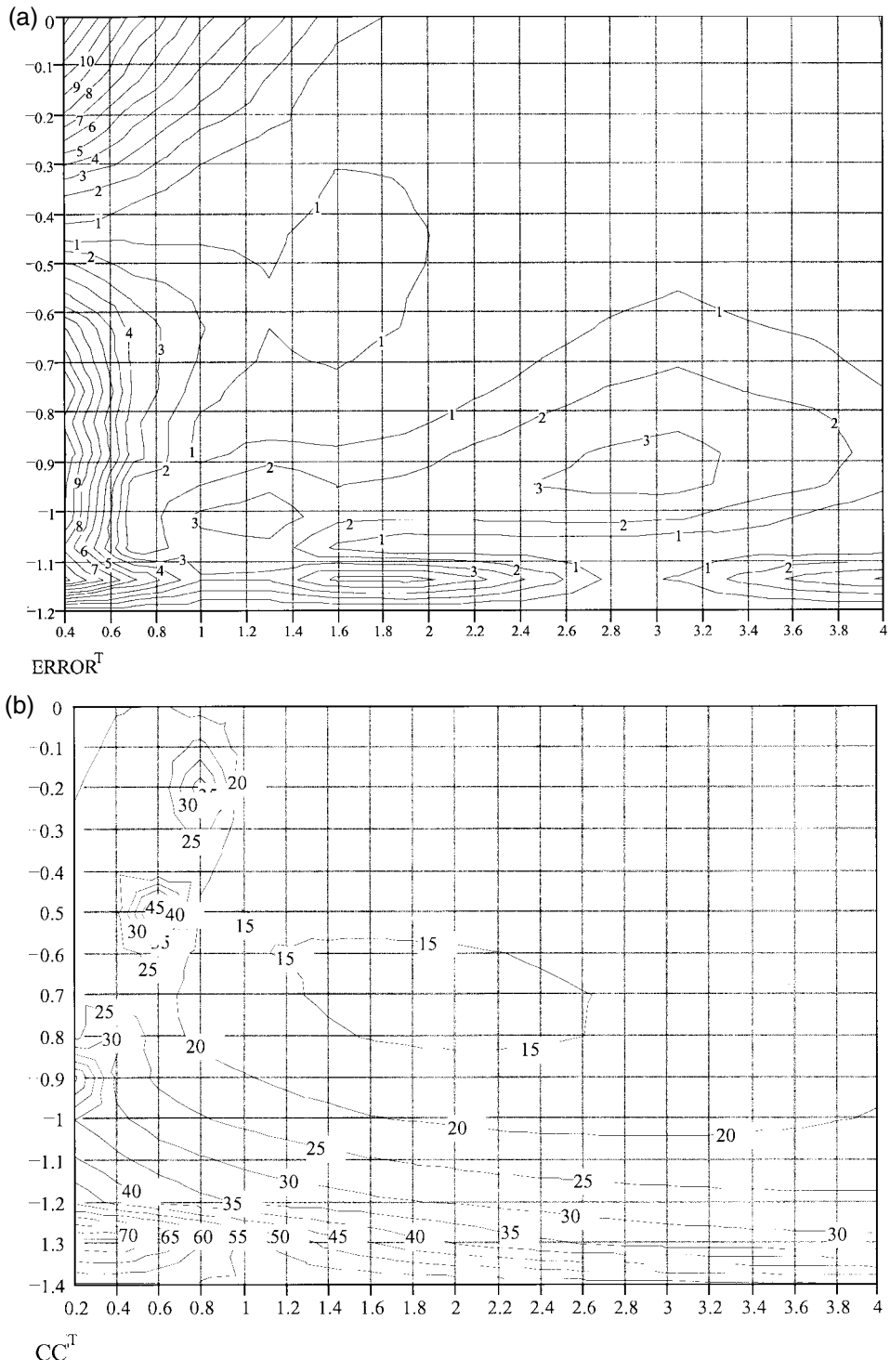


Figure 2-141 NE71000, Curtice cubic model, % error contours for (a) dc *I–V* and (b) *S* parameters.

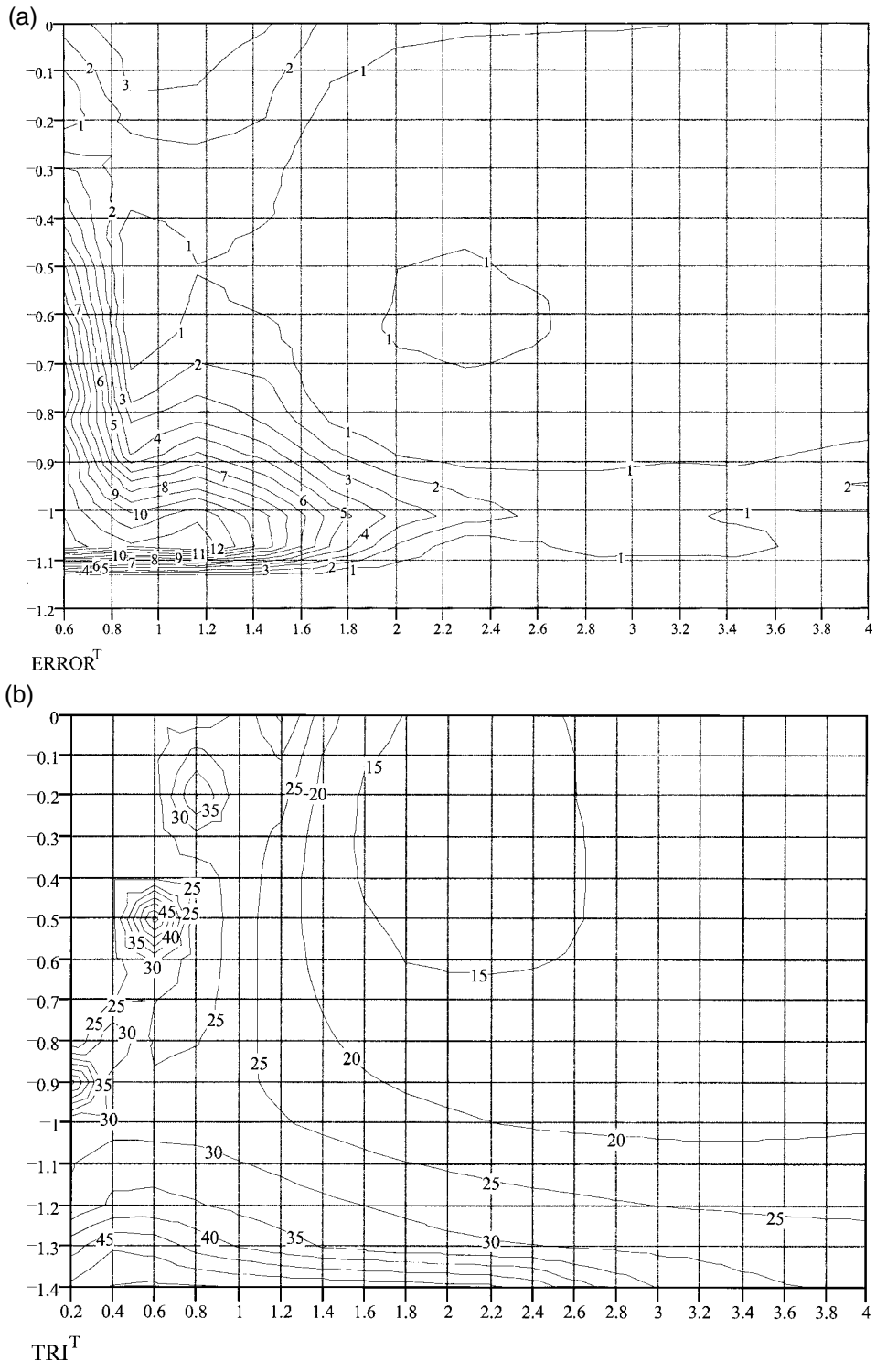


Figure 2-142 NE71000, TOM model, % error contours for (a) dc *I–V* and (b) *S* parameters.

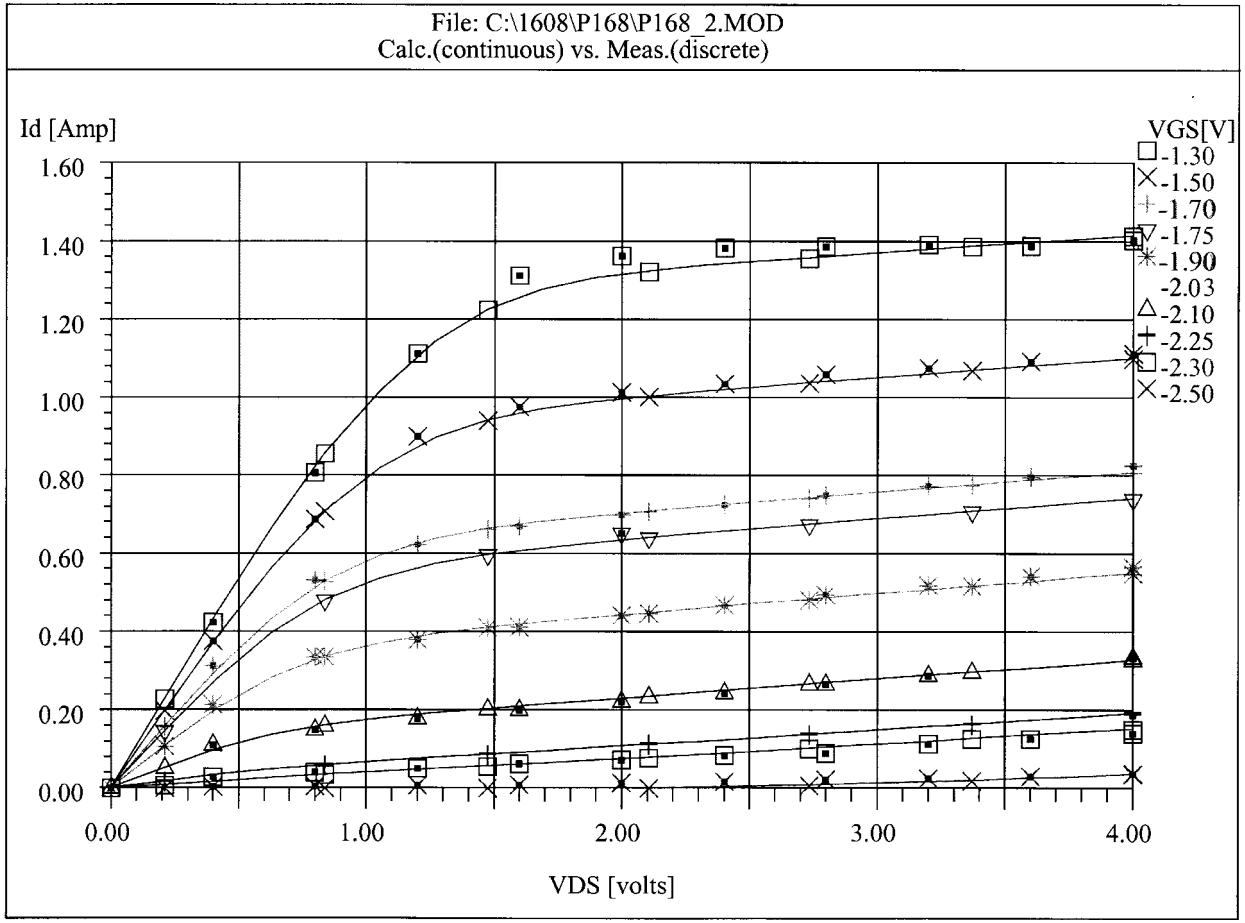


Figure 2-143 Calculated versus measured dc *I–V* curves for the Oki KGF1608 package MESFET.

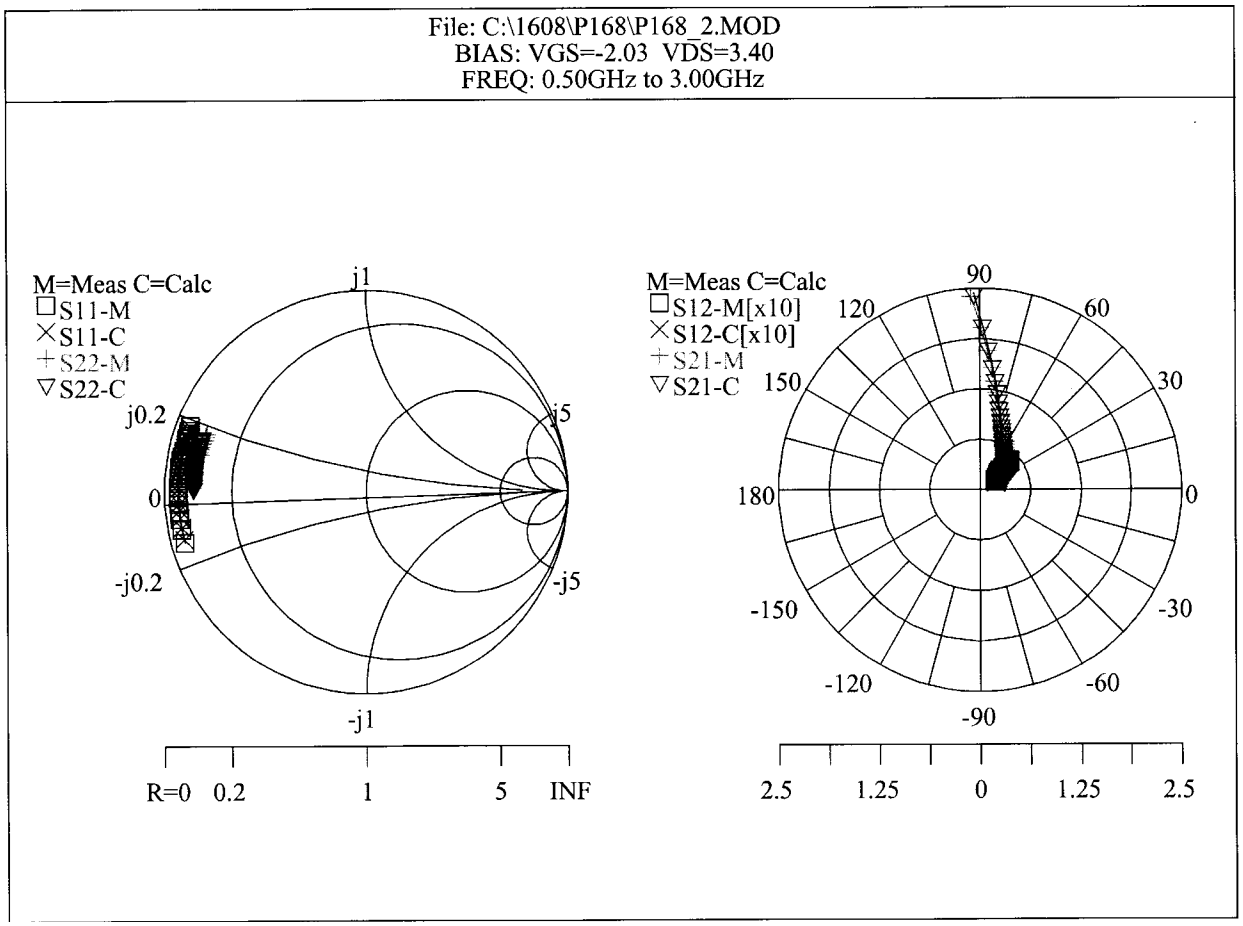


Figure 2-144 Calculated versus measured *S* parameters for the Oki KGF1608 package MESFET.

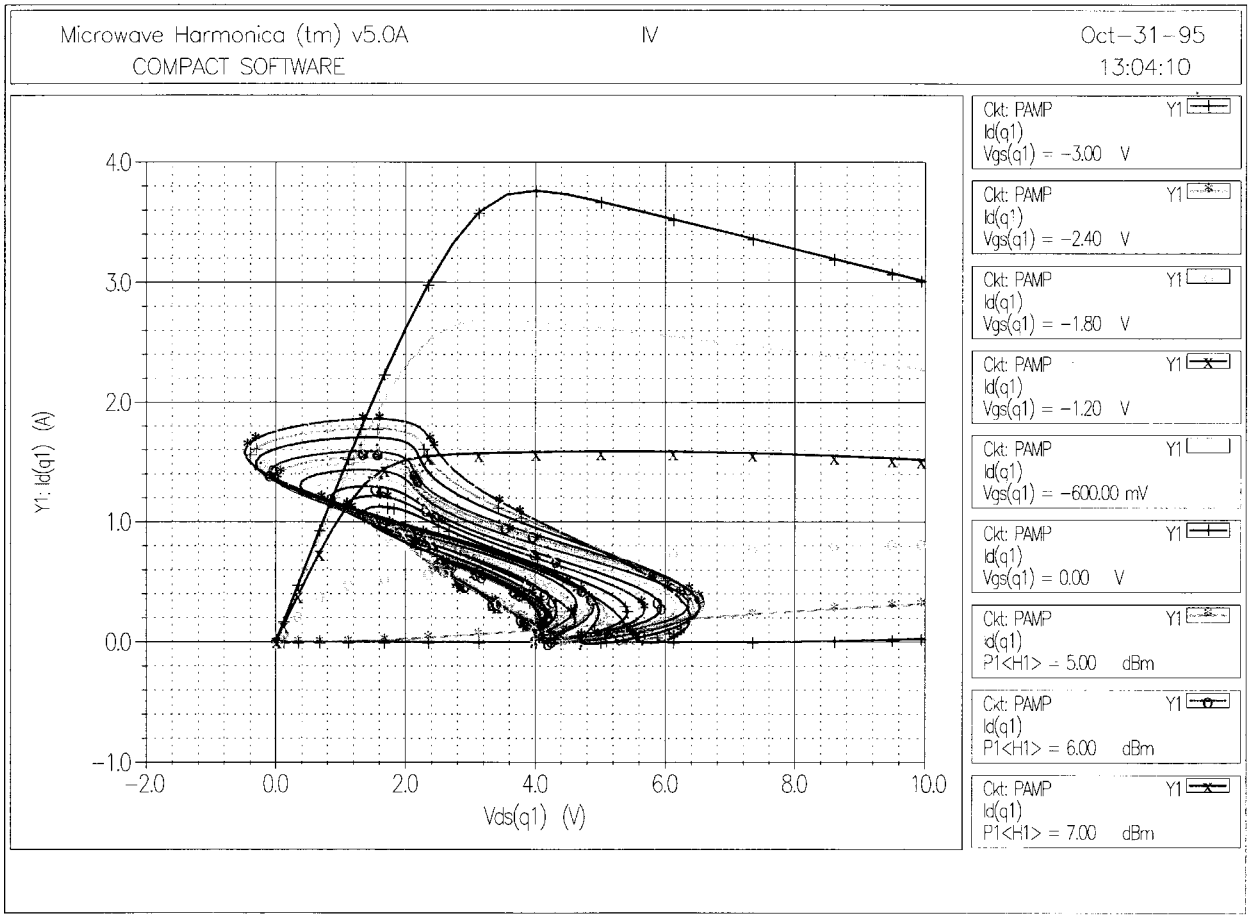


Figure 2-145 Load line and dc *I–V* curves for the KGF1608 MESFET.

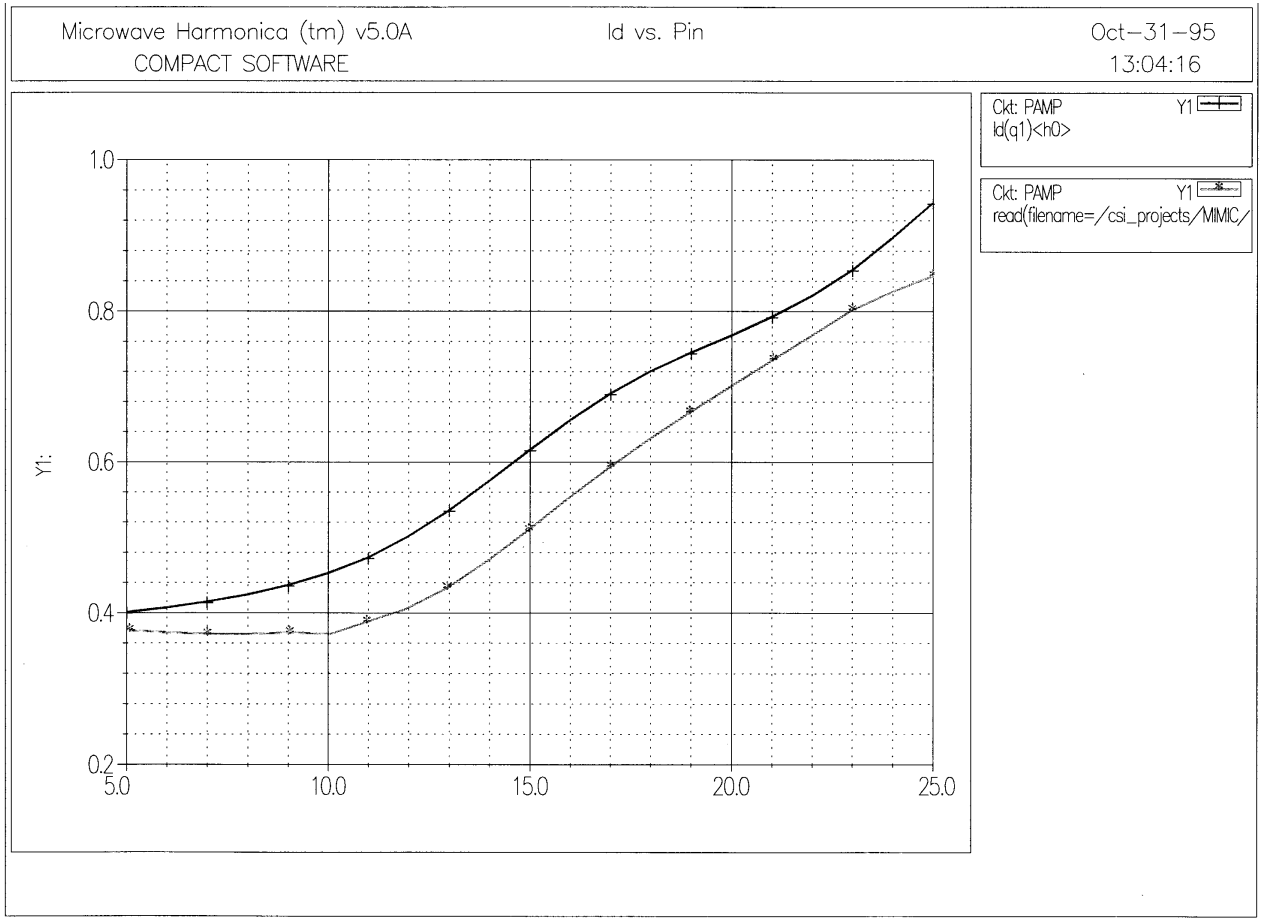


Figure 2-146 Measured versus calculated drain current for the KGF1608 MESFET.

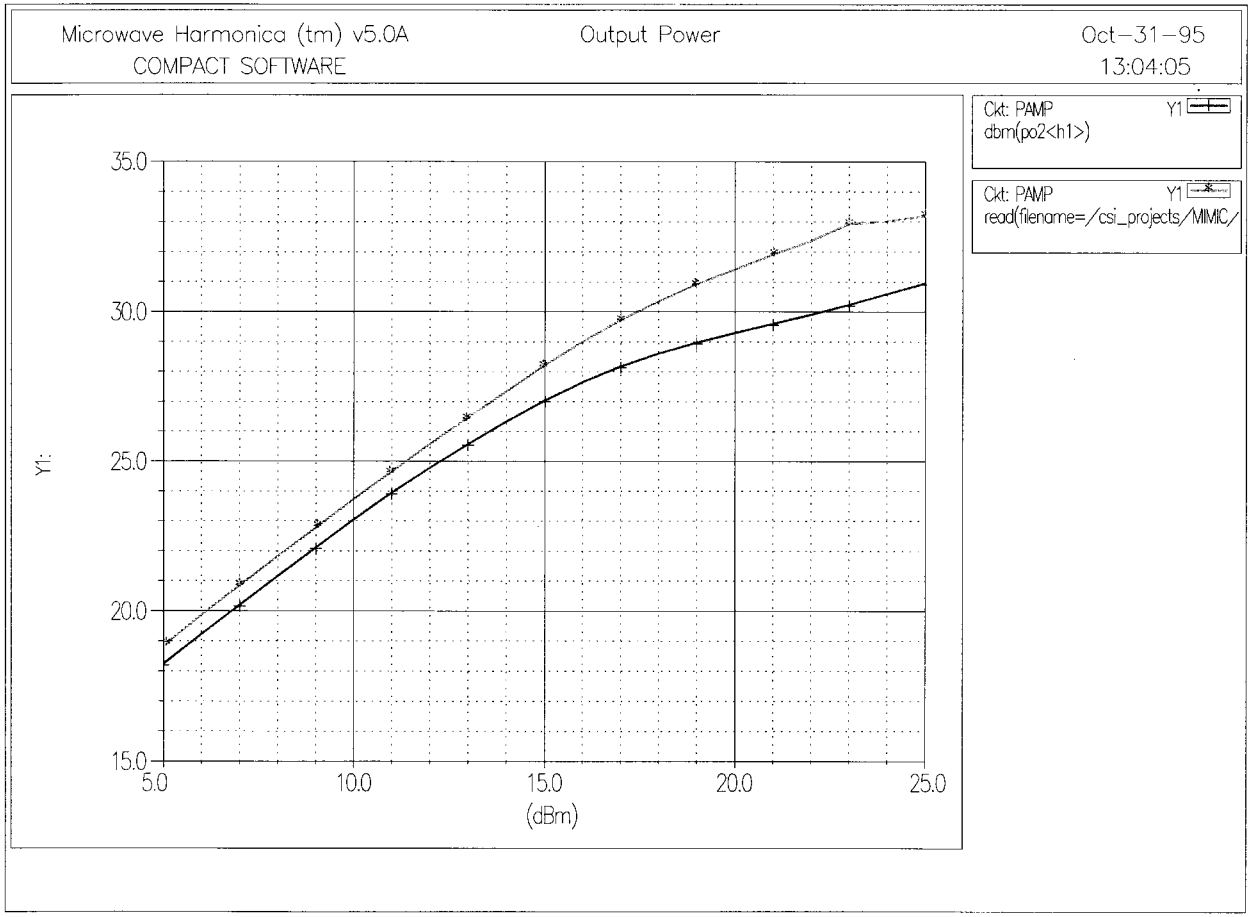


Figure 2-147 Measured versus calculated output power for the KGF1608 MESFET.

ac Errors. From the ac error contour plots we can see that the Chalmers (Angelov) model emerges as the most accurate for ac small-signal operation. The comparison also shows that all models considered are good for ac operations biased in the saturation region with a typical relative error of less than 15%. However, the ac relative errors are much higher in the linear region, particularly for the Curtice quadratic model, which exhibits more than 80% error at some points.

Notes. The above conclusions are based on the matching between the simulated responses of different models and the measured data from a typical NE71000 device; the models may behave differently for a different device. It is not possible to draw any broad-based conclusions as to what model is the best. In order to select the most suitable model for your application, you must consider all the factors, such as bias range, frequency range, and small-signal or large-signal operation, according to your experiences in device modeling.

2-4-6 Device Libraries

All these measurements lead to a nonlinear device library. Figures 2-148 and 2-149 shows a typical datasheet for a device in the library with measured and modeled data. The previous high-power example does not always lead to acceptable results as the models are far from being fully developed. We also show the simulation results using the physics-based model for FETs.

Physics-Based MESFET Modeling. We mentioned a physics-based model earlier. Figure 2-150 shows the measured versus predicted dc *I–V*, I/O power, and *S* parameters for an Alpha Industries MBE MESFET developed under a government contract. The simulation does not handle the trapped energy, which explains the negative slope on the third trace of the output dc *I–V* curve.

The process for bipolar transistors is quite similar and less tricky than the ones just described. Good examples can be found in the user manual for Compact Scout.

2-4-7 A Novel Approach for Simulation at Low Voltage and Near Pinchoff Voltage

As mentioned earlier, the most critical zone is the area of operating the FET both close to the pinchoff voltage and at low operating voltages. A crucial reason for the inadequate results is the modeling of the drain–source resistance as a function of bias point. This is particularly important in calculating the power-added efficiency. When doing your modeling, there is always a temptation to look at both high currents and operating voltages. Figure 2-151 shows a reasonable agreement between measured and predicted dc *I–V* performance. However, the very moment the transistor is modeled close to the linear region (Figure 2-152), things change drastically and the results from 0 to 6 mA have little resemblance to those from 0 to 50 mA. The modified Materka model (Figures 2-153 and 2-154) gives a significantly better match than the Angelov model.

In order to overcome the modeling inaccuracy for the FETs, Thomas Meier of Siemens AG in Munich has come up with the concept of combining two transistors in parallel, whereby the second one is not really connected for dc (Figure 2-155). Figure 2-156 compares measured and simulated R_{DS} using this new approach.

NE67300

Manufacturer : **NEC**
Model Keyword : **NE - 67300**
Device Type : **Mesfet**
Model Type : **Materka**
Package Type : **Chip**

Electrical Characteristics:
Idss : 40mA
Vp : 1.1V
Gain : 8.5dB @ 18GHz
P1db : N/A
Fmin : N/A

Model Data Range:
Vgs : -1.2V to 0.0V
Vds : 0.0V to 4.0V
Freq: 0.4GHz to 18GHz

Verification :
Type : Power sweep, 50Ω
Vgs : -0.6V Vds : 3.0V Ids : 30mA Freq: 6.0GHz

Comments : The S-parameters that appear in the figure were measured and simulated with wire bonds present. The length of each wire bond is approximately 30 mils. To obtain the S-parameters for the manufacture's data sheets, the wire bond inductances would be removed from the equivalent circuit.

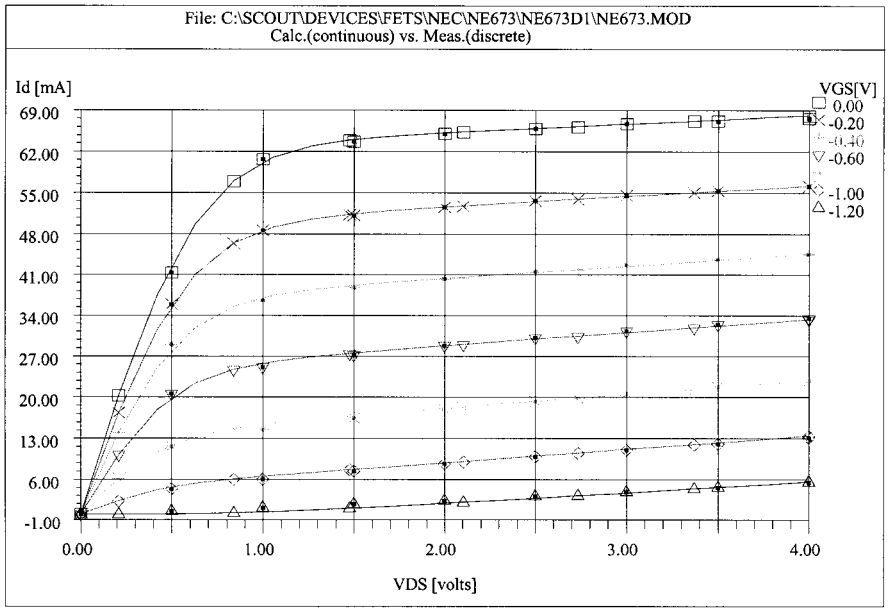
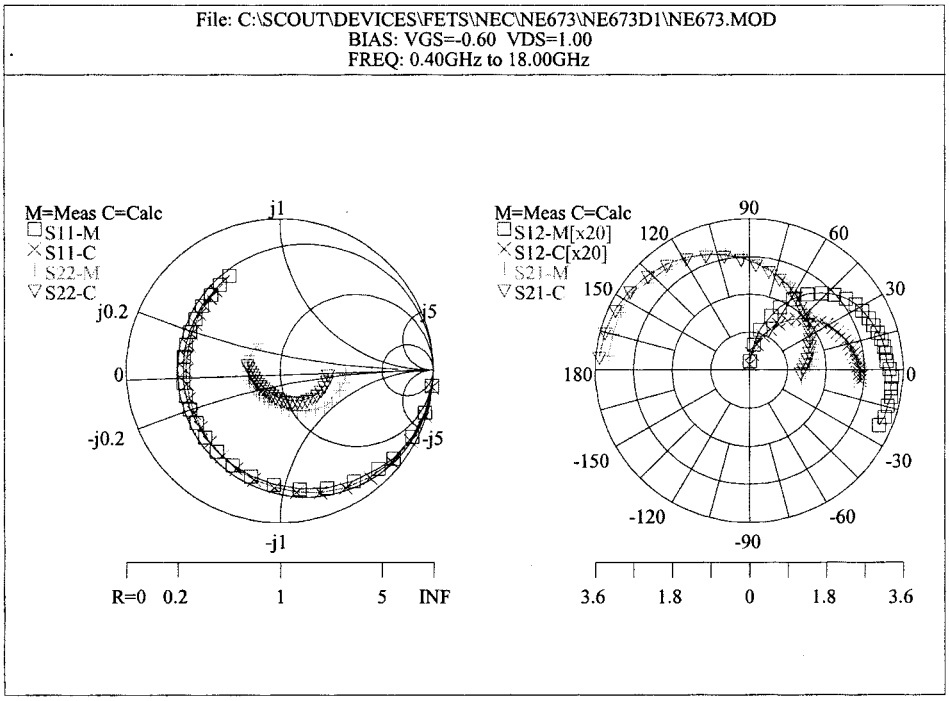


Figure 2-148 Page 1 of a nonlinear device library datasheet for the NE67300.



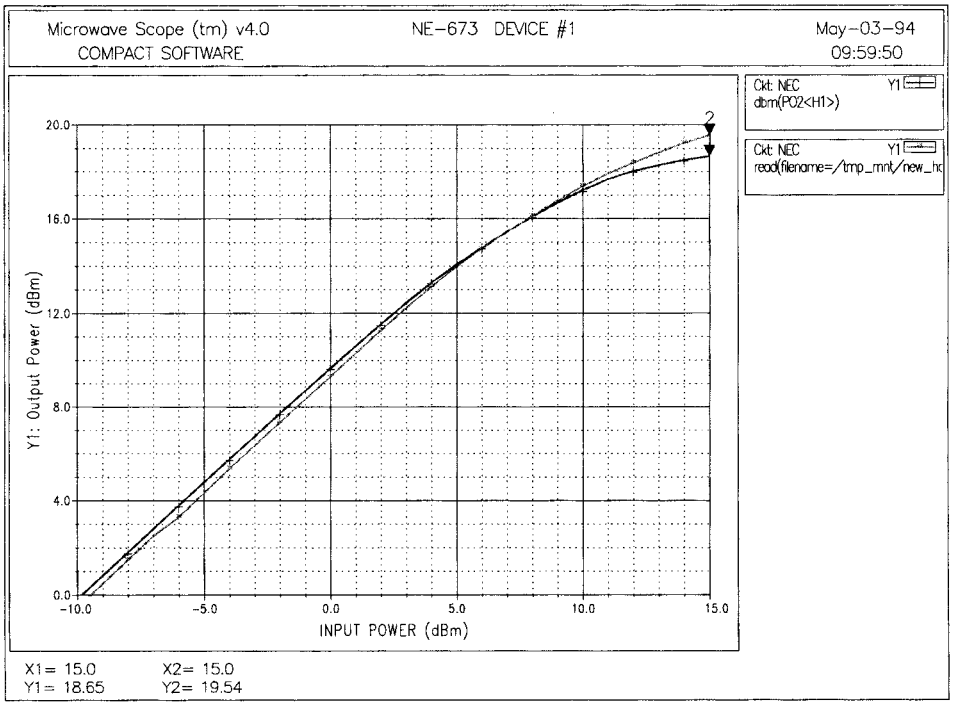


Figure 2-149 Page 2 of a nonlinear device library datasheet for the NE67300.

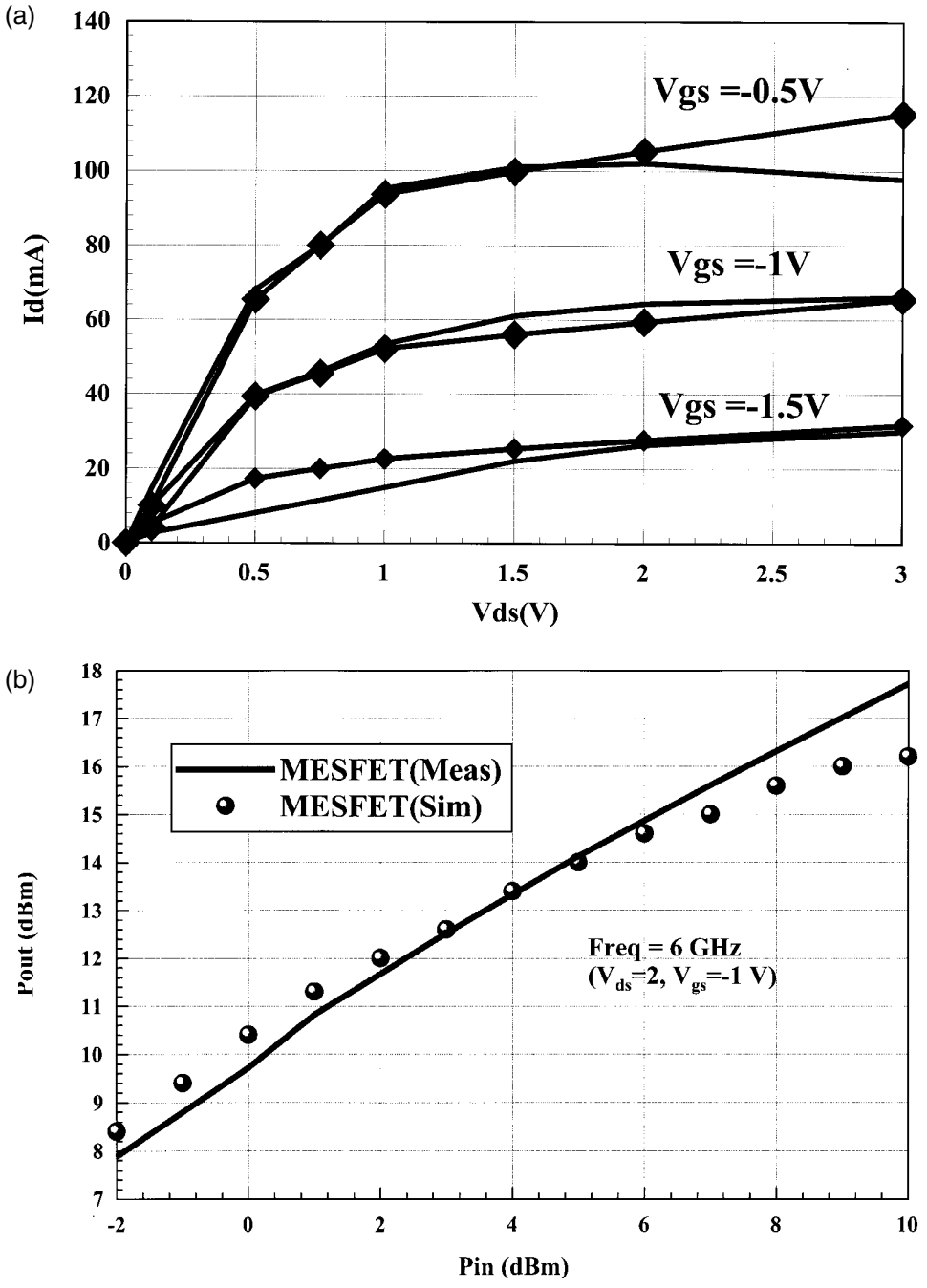


Figure 2-150 Comparison of measured data with PHYSFET and TOM models for an Alpha Industries MBE MESFET (0.25 μm × 400 μm) at $V_{GS} = -1$, $V_{DS} = 2$.

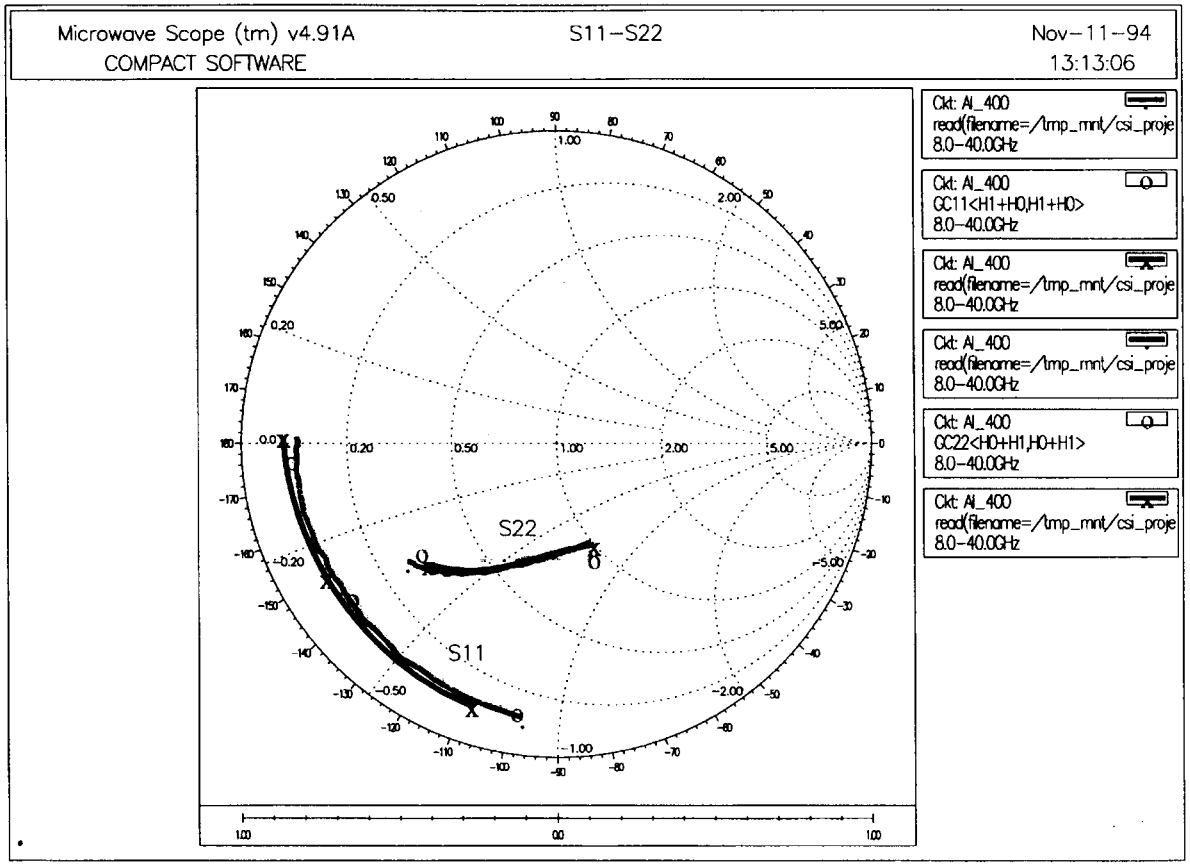


Figure 2-150 (*Continued*)

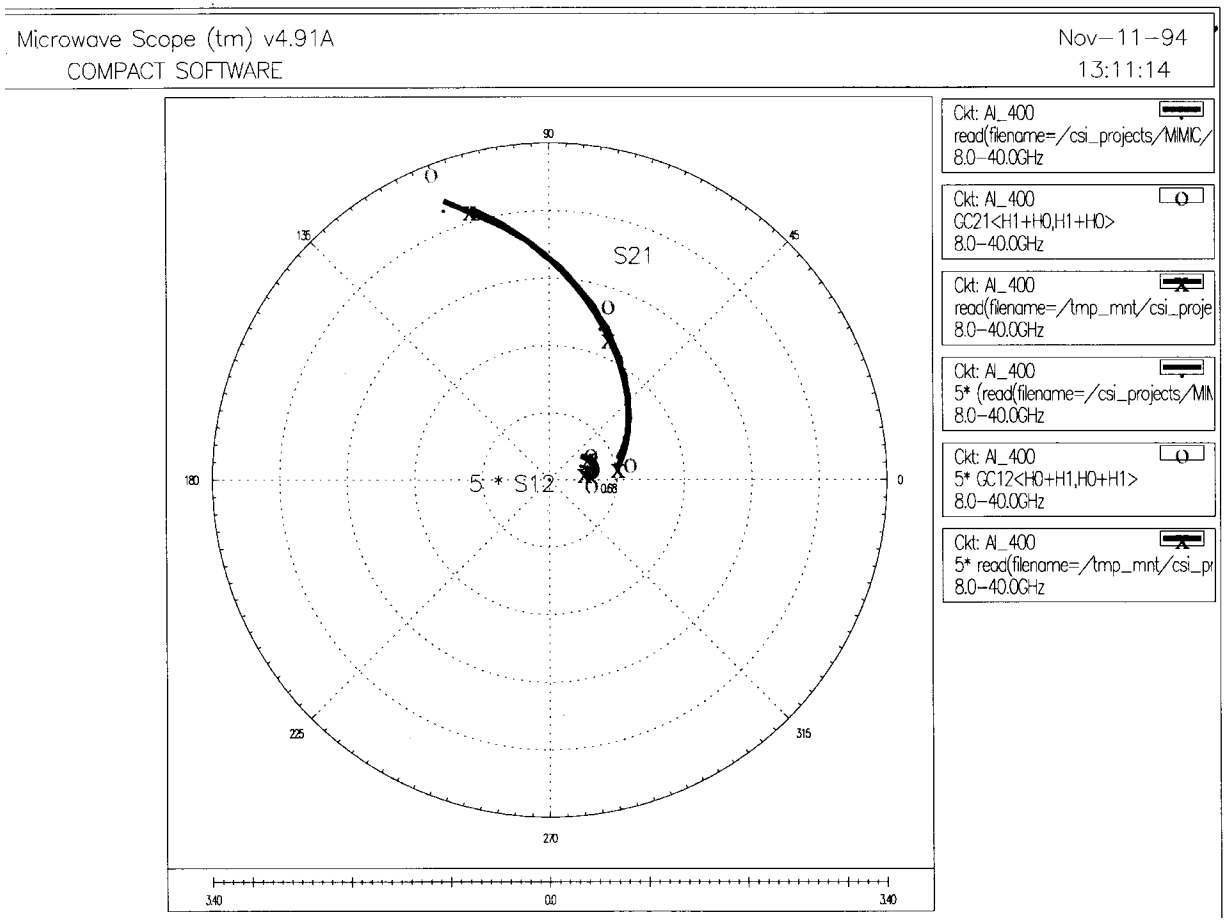


Figure 2-150 (*Continued*)

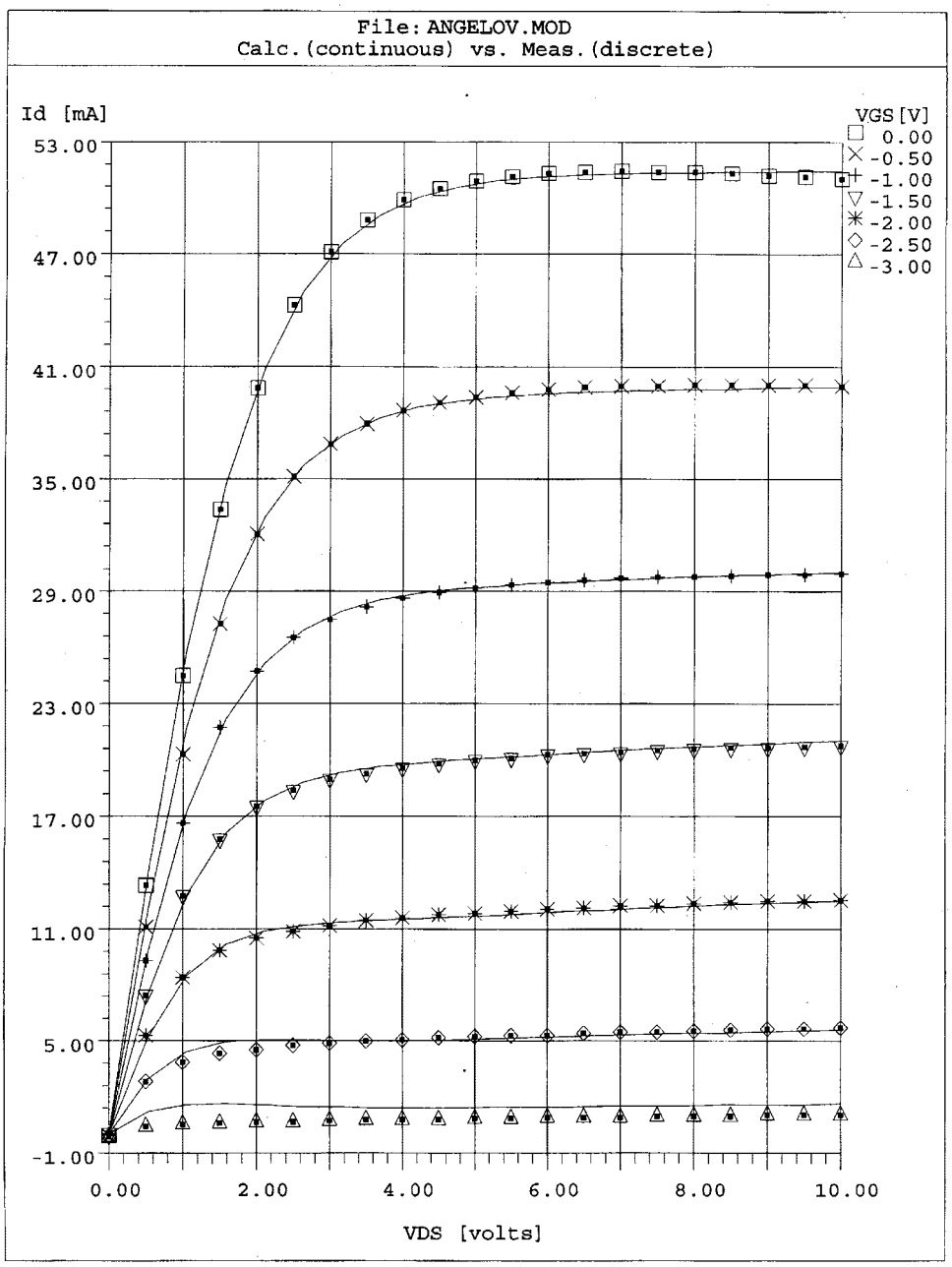


Figure 2-151 High-I_D dc I–V curves for the Angelov FET model.

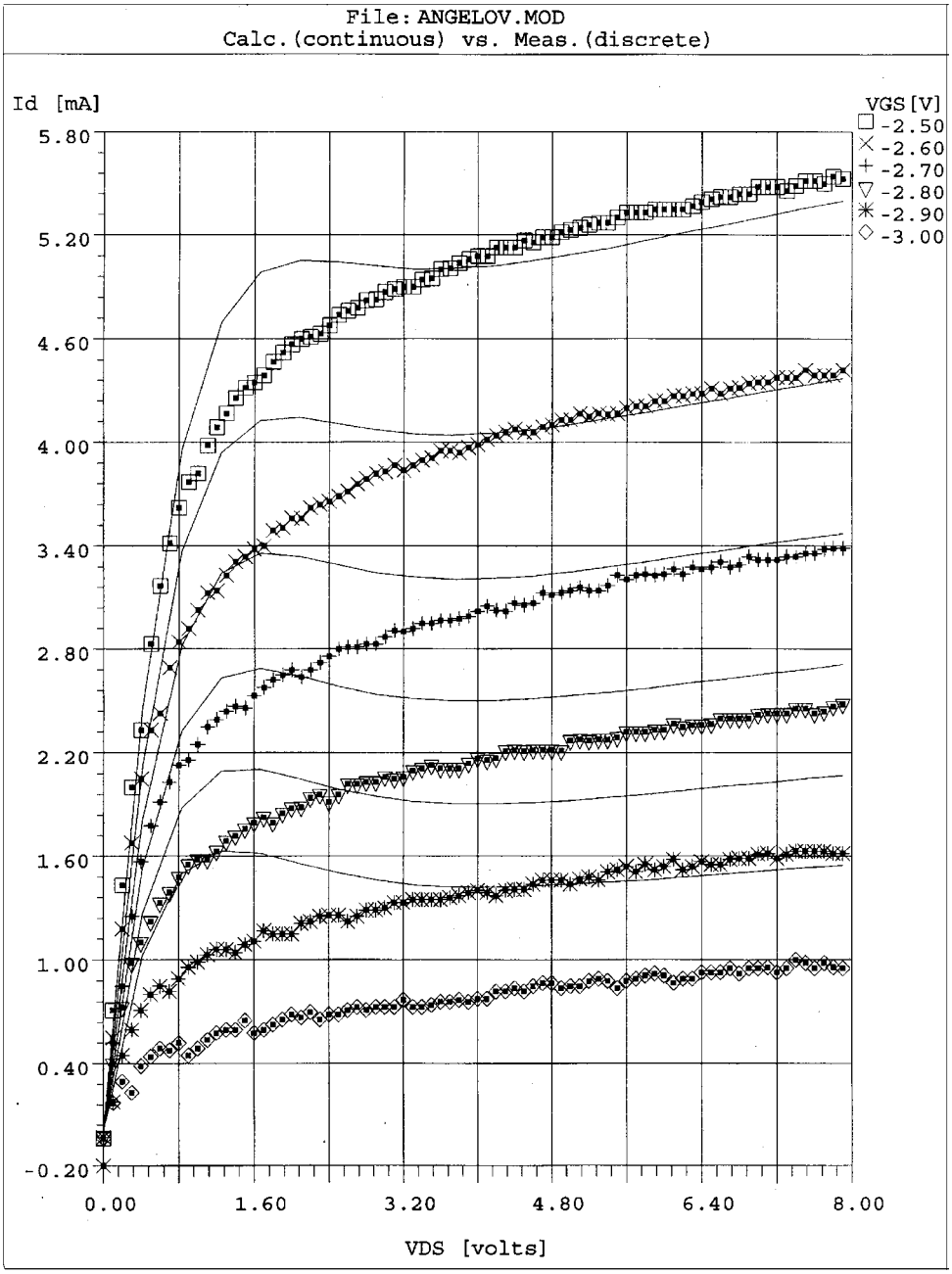


Figure 2-152 Low-I_D dc I–V curves for the Angelov FET model.

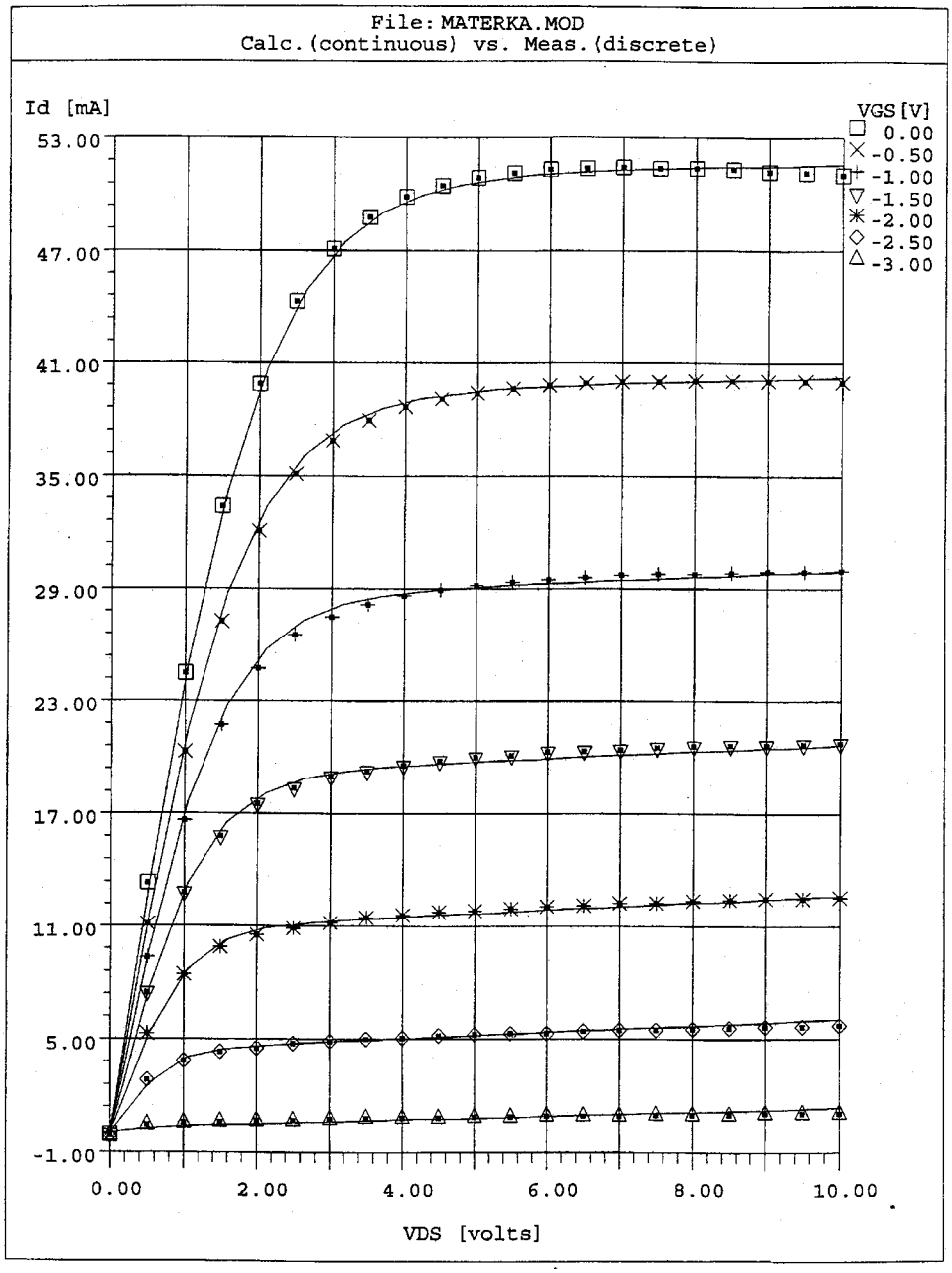


Figure 2-153 High-I_D dc I–V curves for the modified Materka FET model.

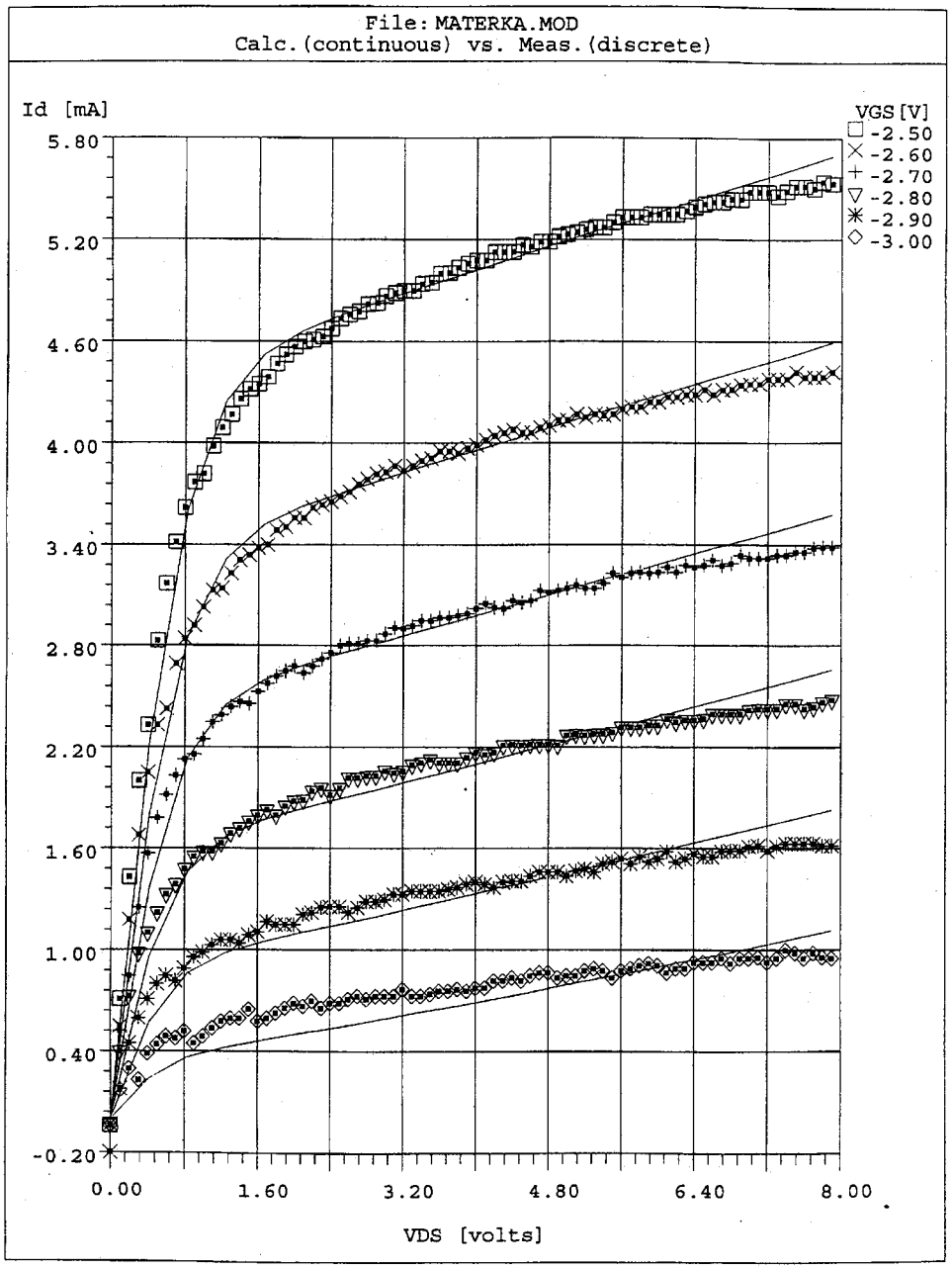


Figure 2-154 Low-I_D dc I–V curves for the modified Materka FET model.

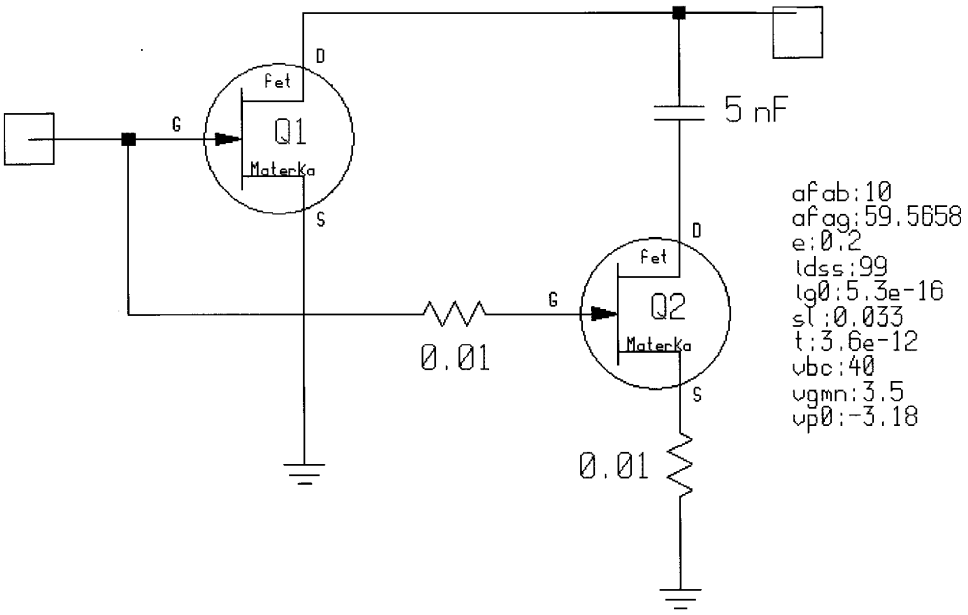


Figure 2-155 By putting two transistors in parallel, with the second one having no dc connection at the drain, one can simulate a dynamic R_{DS} quite well. Regardless of parameter values used for Q1, the parameters listed are necessary to get an accurate R_{DS} result.

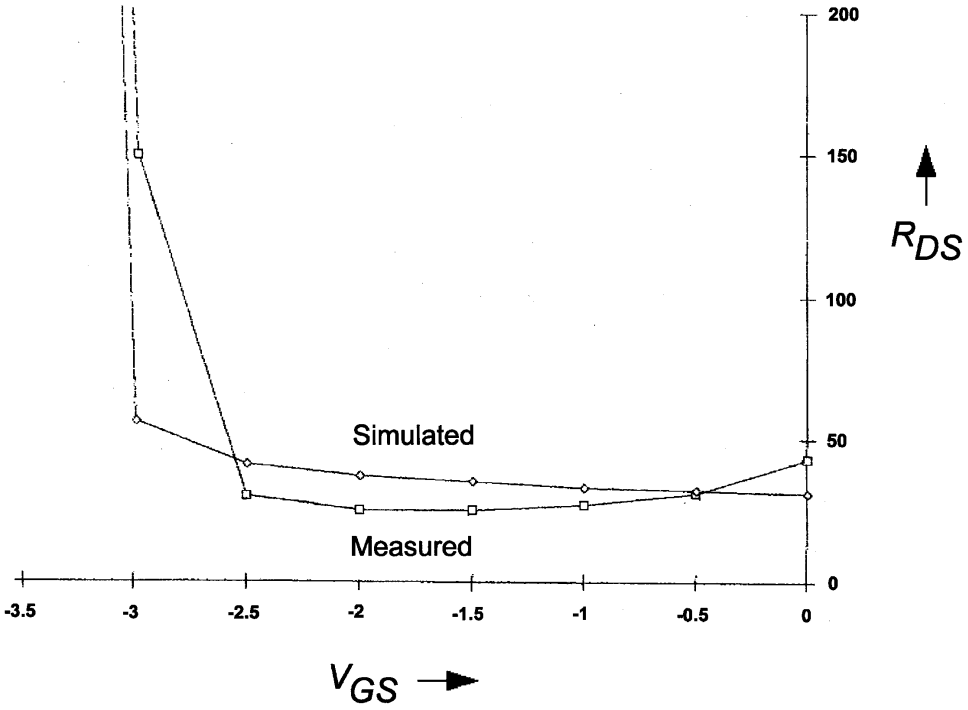


Figure 2-156 Measured versus simulated R_{DS} for the paralleled MESFET approach.

One might ask the question: "Why not fix the existing models?" It turns out that in order to get the best possible convergence and speed, one has to use analytic derivatives for voltage and current; at least the first and second derivatives, if not the third, are required. The programming of such a model becomes so difficult, however, and the calculation of its derivative is so highly likely to have subtle errors, that it turns out that the parallel configuration shown in Figure 2-155 is the best possible compromise for fast computation and accurate results.

A similar approach for bipolar transistors has been discussed by Philips with their MEXTRAM model. However, we have not seen any parameter extraction for this model and also we have not seen any industry acceptance for it. While, in principle, it will cover all derivatives of all members of the BJT family, its topology consists out of three discrete transistors connected and we doubt very much that a reliable parameter extraction program/mathematically continuous model for nonlinear simulation is supplied by the university that developed this jointly with Philips. At the time of writing this book, Philips has not yet provided us with the code for both, and until this happens, we are not in a position to qualify this any further.

2-4-8 Example: Improving the BFR193W Model

We have already mentioned the difficulty encountering the proper base-spreading resistor values and others by using standard parameter extraction programs. Table 2-28 is a copy from the Siemens semiconductor book showing Gummel–Poon Berkeley SPICE syntax parameters for the BFR193W BJT.

The first thing that will happen is that the observant reader will notice that the minimum possible base-spreading resistor, RBM (valid at large current densities), is larger than the actual base-spreading resistor, RB, "claimed to be measured" at 1 Ω (see Table 2-28).

Another headache is caused by the measured noise performance of the same device at an operating point of 10 mA. Using the above-mentioned simplified equation

Table 2.28. BFR193W transistor chip data: SPICE parameters (Gummel–Poon model, Berkeley–SPICE 2G.6 syntax)

IS	=	0.2738	fA	BF	=	125	—	NF	=	0.95341	—
VAF	=	24	V	IKF	=	0.26949	A	ISE	=	10.627	fA
NE	=	1.935	—	BR	=	14.267	—	NR	=	1.4289	—
VAR	=	3.8742	V	IKR	=	0.037925	A	ISC	=	0.037409	fA
NC	=	0.94371	—	RB	=	1	Ω	IRB	=	0.91763	mA
RBM	=	1.8368	Ω	RE	=	0.76534	Ω	RC	=	0.11938	Ω
CJE	=	1.1824	fF	VJE	=	0.70276	V	MJE	=	0.48654	—
TF	=	18.828	ps	XTF	=	0.69477	—	VTF	=	0.8	V
ITF	=	0.96893	mA	PTF	=	0	deg	CJC	=	935.03	fF
VJC	=	1.1828	V	MJC	=	0.30002	—	XCJC	=	0.053563	—
TR	=	1.0037	ns	CJS	=	1	fF	VJS	=	0.75	V
MJS	=	0	—	XTB	=	0	—	EG	=	1.1	eV
XTI	=	3	—	FC	=	0.72063	—	TNOM	=	300	K

All parameters are ready to use, no scaling is necessary. Extracted on behalf of SIEMENS Small Signal Semiconductors by Institut für Mobil-und Satellitenfunktechnik (IMST).

$$F = 1 + \frac{R'_{bb}}{R_g} + \frac{0.5(26\ \text{mV}/I_C)}{R_g}$$

(approximate equation at medium frequencies) (2-288)

we obtain the following results:

1. Inserting 1 Ω (value for RB provided by IMST) and R_B = 15 Ω (measured by the authors), here are the interesting results:

$$F = 1 + \frac{1}{50} + \frac{2.6}{100} = 1.046 \tag{2-289}$$

as noise factor or, as noise figure, F (dB) = 0.2 dB (IMST).

2. Or using the measured 15 Ω, one obtains

$$f = 1 + \frac{15}{50} + \frac{2.6}{100} = 1.326 \tag{2-290}$$

as noise factor or, as noise figure, F (dB) = 1.22.

The Siemens datasheet claims a noise figure of 1.3 dB, and therefore the measured versus calculated data differ by 0.08 dB. Since our noise figure was calculated well below f_T, which is seen from the fact that there are no frequency-dependent terms in the noise figure calculation, it is not unusual that the 900-MHz noise figure is 1.3 dB or slightly higher. This simple example shows how easy it is to provide measurement data that are inconsistent with the laws of physics. It is therefore highly recommended to add the noise figure as an additional parameter in the parameter extraction to reduce the number of variables—specifically, those for which the standard extraction has little sensitivity.

The collector–emitter junction capacitance is another important parameter that frequently gets measured incorrectly. The way around this is to remember the definition

$$f_{\max} = \sqrt{\frac{f_T}{8\pi R'_{bb} C_c}} \tag{2-291}$$

f_T can be measured from the 3-dB point of the emitter current gain (β) and R'_{bb} was just computed above, so the equation can be solved for C_c, as there is a relationship between measured $g_{\max}$ and $f_{\max}$. There are several correlations between the parameters of the equivalent circuit as published in different places, and therefore one can restrict the degree of freedom within the optimization process and obtain much better results. This approach is applicable to all types of transistors, bipolar and FET, and can be used to improve the accuracy of parameters that not only affect a transistor's RF performance but also its noise performance.

REFERENCES

1. H. A. Watson, *Microwave Semiconductor Devices and Their Circuit Applications* . New York: McGraw-Hill, 1969, Chap. 2.
2. A. S. Grove, *op. cit.,* Chap. 4.
3. A. S. Grove, *op. cit.,* Chap. 7.
4. P. E. Gray, D. DeWitt, A. R. Boothroyd, and J. G. Gibbons, *Physical Electronics and Circuit Models of Transistors* . New York: Wiley, 1964, p. 10.
5. P. E. Gray et al., *op. cit.,* p. 129.
6. P. E. Gray et al., *op. cit.,* p. 180.
7. Charles A. Liechti, "Microwave Field-Effect Transistors—1976," *IEEE Trans. Microwave Theory Tech.* , **MTT-24**, 279–300, June 1976.
8. Edward A. Wolff and Roger Kaul, *Microwave Engineering and Systems Applications* . New York: Wiley, 1988, p. 395.
9. Arthur D. Evans, ed., *Designing with Field-Effect Transistors* . New York: McGraw-Hill, 1981.
10. R. D. Middlebrook, "A Simple Derivation of Field-Effect Transistor Characteristics," *Proc. IEEE*, **50**, 1146–1147, 1963.
11. Tzu-Hwa and Craig P. Snapp, "Low Noise Microwave Bipolar Transistor with Sub-Half Micrometer Emitter Width," *IEEE Trans. Electron Devices* , **ED-25**, 723–730, June 1978.
12. Ulrich L. Rohde, "Designing a Match Low Noise Amplifier Using CAD Tools," *Microwave J.* , **20**(10) 154–160, October 1986.
13. R. A. Pucel, W. Strubel, R. Hallgren, and U. L. Rohde, "A General Noise De-embedding Procedure for Packaged Two-Port Linear Active Devices," *IEEE Trans. Microwave Theory Tech.* , **40**(11), 2013–2024, November 1992.
14. R. A. Pucel and U. L. Rohde, "An Exact Expression for Noise Resistance R_n of a Bipolar Transistor for the Hawkins Noise Model," *IEEE Microwave and Guided Wave Letters* , **3**(2), 35–37, February 1993.
15. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits.* New York: Wiley, 1986, p. 386.
16. D. Frohman-Bentchkowsky and A. S. Grove, "Conductance of MOS Transistors in Saturation," *IEEE Trans. Electron Devices* , **ED-16**, 108–113, January 1969.
17. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits* . New York: Wiley, 1986, p. 480.
18. "New LDMOST Technology Superb for RF Applications," *Philips Semiconductors World News,* Vol. 7, Issue 3, Article No. 29, http://www-us2.semiconductors.philips.com/publications/content/file_344.html.
19. Y. P. Tsividis, *Operation and Modeling of the MOS Transistor* . New York: McGraw-Hill, 1987, p. 136.
20. A. Materka and T. Kacprzak, "Computer Calculation of Large-Signal GaAs FET Amplifier Characteristics," *IEEE Trans. Microwave Theory Tech.* , **MTT-33**(2), 129–135, February 1985.
21. R. E. Anholt and S. E. Swirhun, "Experimental Investigation of the Temperature Dependence of GaAs FET Equivalent Circuits," *IEEE Trans. Electron Devices* , **ED-39**(9), 2029–2036, September 1992.
22. Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits* . New York: Wiley, 1986.
23. Paul W. Tuinenga, *SPICE: A Guide to Circuit Simulation and Analysis Using PSpice ®*. Englewood Cliffs, NJ: Prentice Hall, 1992.
24. Andrei Vladimirescu, *The SPICE Book* . New York: Wiley, 1994.

25. H. Statz, H. Haus, and R. A. Pucel, "Noise Characteristics of Gallium Arsenide Field-Effect Transistors," *IEEE Trans. Electron Devices,* **ED-21**, 549–562, September 1974.

26. V. Rizzoli, F. Mastri, and C. Cecchetti, "Computer-Aided Noise Analysis of MESFET and HEMT Mixers," *IEEE Trans. Microwave Theory Tech.* , **MTT-37**, 1401–1410, September 1989.

27. V. Rizzoli, F. Mastri, and D. Masotti, "General-Purpose Noise Analysis of Forced Nonlinear Microwave Circuits," *Military Microwave*, 1992.

28. Ulrich L. Rohde and Chao-Ren Chang, "Parameter Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits," *Microwave J.,* **36**(5), 222–239, May 1993.

29. Ulrich L. Rohde, "Parameters Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits Like Mixers and Oscillators," *23rd European Microwave Conference,* Madrid, Spain, September 6–9, 1993.

30. Ulrich L. Rohde, Chao-Ren Chang, and J. Gerber, "Design and Optimization of Low-Noise Oscillators Using Non-linear CAD Tools," presented at the *Frequency Control Symposium,* Boston, June 1–2, 1994.

31. Ulrich L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design* . New York: Wiley, 1997, Appendix B-3, "Bias Independent Noise Model."

FURTHER READING

George Anderson, "Moto Aims FETs at UHF," *Electron. Eng. Times*, 96, September 2, 1996.

Jean–Jacques Bouny, "Advantages of LDMOS in High Power Linear Amplification," *Microwave Eng. Eur.*, 37–40, April 1996.

Noël Boutin, "Develop SPICE High Frequency Small Signal Transistor Models from Data Sheets," *Appl. Microwave & Wireless*, 30–36, April 1999.

V. Danelon, F. Aniel, J. Benchimol, J. Mba, M. Riet, P. Crozat, G. Vernet, and R. Adde, "Parameters of InP-Based Double-Heterojunction Base Collector Self-Aligned Bipolar Transistors," *IEEE Microwave Guided Wave Lett.*

Suman Datta, Shen Shi, Kenneth P. Roenker, Marc M. Cahay, and William Stanchina, "Simulationand Design of InAlAs/InGaAs *pnp* Heterojunction Bipolar Transistors," *IEEE Trans. Electron Devices*, **ED-45**(8), 1634–1643, August 1998.

Russell J. Elias, Gordon C. Ma, and Laurence S. Golonka, "Statistical Optimization High Frequency LDMOS Devices Via Hyper-Fractionated Experimental Designs," *1996 IEEE/CPMT International Electronics Manufacturing Technology Symposium,* October 14–16, 1996, pp. 462–466.

Al Franceschino, "LDMOS Devices Enhance Cellular Amplifier Design," *Microwaves & RF*, **35**(10), 67–68, 70, 72, 74, October 1996.

GaAs IC Symposium Technical Digest 1995 , IEEE Gallium Arsenide Integrated Circuit Symposium, San Diego, CA, October 29 to November 1, 1995.

Hewlett-Packard, *Communications Components Designer's Catalog* , Hewlett-Packard, Jan Jose, CA, 1997.

ITT Capacitance Diodes, Tuner Diodes, Diode Switch, PIN Diodes, Basics and Applications , International Telephone and Telegraph Corporation, 1976.

Aron Kain, *Final Report for Bias Dependence Noise Modeling of Heterojunction Bipolar Transistors* , USAF SBIR Phase II (PIIN), F33615-95-C-1707, November 1997. Issued by USAF/AFMC/ASC, Wright Laboratory WL/AAKE BLD 7, 2530 C ST, Wright-Patterson AFB, OH 45433-7607.

X. McDade and X. Schiavone, "Switching Time Performance of Microwave PIN Diodes," *Microwave J.*, **17**, September 1974.

Danny Molezion, "A 1 watt Power Amplifier Design for 800–900 MHz Applications Using Discrete LDMOS FETs," *Wireless Technology '96 Conference & Exposition Conference Proceedings,* October 7–11, 1996, pp. 260–267.

Nadir Rahman, "A Survey of Existing Literature on Reverse Modeling of Field Effect Transistors," http://www-mtl.mit.edu/CAPAM/memos/95-3/95-3.html.

Jin-Ho Shin, Jiyoung Kim, Yuji Chugn, Joonwoo Lee, Youngsuk Suh, Kyu Hwan Ahn, and Bumman Kim, "Low-Frequency Noise Characterization of Self-Aligned AlGaAs/GaAs Heterojunction Bipolar Transistors with a Noise Corner Frequency Below 3 kHz," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, pp. 1604–1613, November 1998.

K. Y. Toh, P. K. Ko, and R. G. Meyer, "An Engineering Model for Short-Channel MOS Devices," *IEEE J. Solid-State Circuits*, **23**, 950–958, August 1988.

Robert J. Zavrel, Jr., "Silicon–Germanium and the World of RF Design," *RF Des.*, 46–50, June 1997.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

2

MODELS FOR ACTIVE DEVICES

Because we are dealing with amplifiers, both small-signal and power, oscillators, and active mixers, active devices will play a major role in circuit design. At our frequency range of interest (about 1000 MHz and above), distributed elements predominate, and even in device modeling we must consider some distributed effects such as packaging.

As we dive into providing functional explanations of the active devices used in practical circuits, there is some temptation not to provide enough information on semiconductor behavior. Because manufacturers of microwave products will use the services of a foundry or specify semiconductors from their suppliers, we will cover the device operation in depth. That said, for the beginner looking into designing circuits from high frequencies to low microwave frequencies, we will provide examples of a variety of devices unique to this frequency range. Some treatment of device physics will therefore be useful in providing a basic understanding of the nonlinear behavior of active devices and modeling in preparing for practical circuits using a computer-aided-design (CAD) tool. For CAD applications, we assume the user will have access to nonlinear CAD software, such as a high-performance SPICE program or a harmonic-balance program, such as Ansoft Serenade 8.0 or its derivatives. It is interesting that most semiconductor houses today still only offer very limited nonlinear data sets, while in the linear area a large body of useful data is made available, mostly on CDs. What CAD does for us—and this applies to SPICE and harmonic-balance programs alike—is predict the dc bias point of a device in the context of its surrounding circuitry. Configuring the simulation to drive the device into large-signal operation provides us with insight into the various forms of distortion and compression.

In evaluating active devices for RF circuit design, we have a few different technologies to consider. First, we will examine the semiconductor junction diode—a basic *pn* junction—and venture from there to the bipolar transistor. An investigation of metal-oxide semiconductor (MOS) devices, another important set of semiconductors for medium frequencies, will follow. We will also examine the gallium arsenide metal-gate Schottky field-effect transistor (GaAs MESFET) and members of its family. Since critical large-signal parameters for some devices may not be available from their manufacturers, we will examine

parameter-extraction techniques and problems. For various device types, we will provide insight into the noise properties of the device.

2-1 DIODES*

The diode model contains a nonlinear current source that follows the Shockley equation:

$$\text{Current} = \text{IS}\left(e^{V_j/\text{N}V_t} - 1\right) \tag{2-1}$$

where V_j = voltage across the junction
V_t = thermal voltage (= kT/q)

These values, with the model parameters IS and N, are used to model the current–voltage effects of the semiconductor junction. This does not include the nonideal operation of real diodes. For example, at low currents (less than 1 nA), other semiconductor processes that increase the flow of currents become noticeable.

By setting IS to different values, we can obtain the characteristics of other devices, such as a Schottky barrier diode, or a silicon diffused-junction diode. High-current effects are modeled, grossly, by including a series resistance that is intended to combine the effects of bulk resistance (the material on each side of the junction) and high-level injection. At high currents, the observed diode current stops following the Shockley form

$$I_{\text{forward}} = \text{IS} \times e^{V_j/\text{N}V_t} \tag{2-2}$$

and approaches a modified form

$$I_{\text{forward}} = \text{IS} \times e^{V_j/2\text{N}V_t} \tag{2-3}$$

2-1-1 Large-Signal Diode Model

Three diode models are used in the industry:

- Microwave diode model
- PIN diode model
- Enhanced SPICE diode model

Figure 2-1 shows the large-signal microwave diode model. Its keywords appear in Table 2-1. This model can also be used to simulate varactor diodes.

Table 2-2 lists SPICE parameters for a selection of Schottky mixer diodes by Alpha.

In most cases, the diode capacitance is modeled by a voltage-dependent capacitor, which is connected in parallel with the nonlinear current generator described previously, to represent the charge-storage effects of the junction. There are two components to this charge

- The reverse-voltage capacitive effect of the depletion region
- The forward-voltage charge represented by mobile carriers in the diode junction

*Portions of this chapter's diode coverage are based on material in Alpha Industries Semiconductor Application Reports 80800, 80200, and 80500. Used with permission.

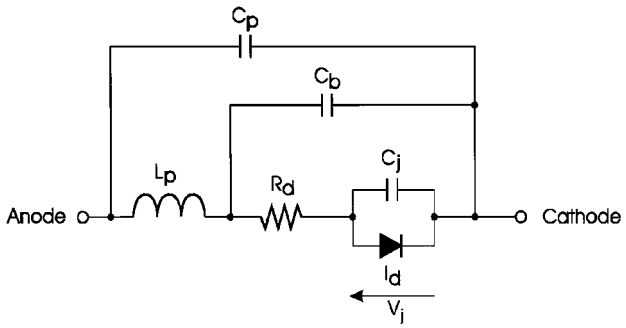


Figure 2-1 The large-signal microwave diode model. This model is temperature dependent.

Reverse-voltage capacitance follows the simple approximation that the depletion region (the area of the junction that is depleted of carriers) serves as the gap between the "plates" of a capacitor. This region varies in thickness, and therefore the capacitance varies with applied voltage. For a step (abrupt) junction, or linearly graded junction, the capacitance approximation is

Table 2-1 Large-signal microwave diode model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
JS	Saturation current	ampere	0
ALFA	Slope factor of conduction current	/volt	38.696
JB	Breakdown saturation current	ampere	10 mA
VB	Breakdown voltage	volt	$-\infty$
E	Power-law parameter of breakdown current	—	10.0
CT0	Zero-bias depletion capacitance	farad	0
FI	Built-in barrier potential	volt	0.8
GAMA	Capacitance power-law parameter	0.5	
GC1	Varactor capacitance polynomial coefficient 1	/volt	0.0
GC2	Varactor capacitance polynomial coefficient 2	/volt^2	0.0
GC3	Varactor capacitance polynomial coefficient 3	/volt^3	0.0
CD0	Zero-bias diffusion capacitance (*pn* diodes)	farad	0
AFAC	Slope factor of diffusion capacitance	/volt	38.696
R0	Bias-dependent part of series resistance in forward-bias condition	ohm	0
T	Intrinsic time constant of depletion layer for abrupt-junction diodes	second	0
KF	Flicker noise coefficient	—	0.0
AF	Flicker noise exponent	—	1.0
FCP	Flicker noise frequency shape factor	—	1.0
AREA	Area multiplier	—	1.0
	Extrinsic Model		
CP	Package parasitic capacitance	farad	0.0
CB	Beam-lead parasitic capacitance	farad	0.0
LP	Package parasitic inductance	henry	0.0

Table 2-2 Diode SPICE parameters

Parameter	Unit	SMS1546	SMS3922	SMS3923	SMS3924	SMS3926	SMS3927	SMS3928	SMS7621	SMS7630
I_S	A	3E–7	3E–8	5E–9	2E–11	2.5E–07	1.3E–09	9E–13	4E–8	5E–06
R_S	Ω	4	9	11	11	4	4	4	12	30
n	—	1.04	1.08	1.05	1.08	1.04	1.04	1.04	1.05	1.05
T_d	S	1E–11	8E–11	8E–11	8E–11	1E–11	1E–11	1E–11	1E–11	1E–11
C_{J0}	PF	0.38	0.9	0.93	1.6	0.42	0.39	0.39	0.1	0.14
m	—	0.36	0.26	0.24	0.4	0.32	0.37	0.42	0.35	0.4
E_G	EV	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
X_{ti}	—	2	2	2	2	2	2	2	2	2
F_C	—	0.5	0.5	0.5	0.5	0.5	0.5	0.5	5	0.5
B_J	V	3	20	46	100	2	3	4	3	2
I_{BV}	A	1E–5	1E–5	1E–5	1E–5	1.00E–05	1.00E–05	1.00E–05	1E–5	0.0001
V_J	—	0.51	0.65	0.15	0.84	0.495	0.595	0.800	0.51	0.34

$$\text{Capacitance} = \frac{C_{J0}}{\left(1 - V_j/\phi\right)^M} \tag{2-4}$$

where C_{J0} is the zero-bias value, ϕ (phi) is the junction barrier potential, and M is the grading coefficient that varies ($\frac{1}{2}$ is used for step junctions and $\frac{1}{3}$ is used for linearly graded junctions, and most junctions are somewhere in between).

There is often confusion about the barrier potential, which appears in the capacitance equation. From capacitance measurements, ϕ (model parameter VJ, and not to be confused with V_j in the equations) takes on a value of nearly 0.7 volt for regular (silicon) junction diodes and a range of 0.58–0.85 volt for various Schottky barrier diodes. This value is sometimes confused with the forward-current voltage drop of the diode or the energy gap of the material; it is *neither* of these.

Varying M generates a variety of reverse-bias capacitance characteristics. Inspection of the capacitance formula reveals that it predicts infinite capacitance for a forward bias, which is not the case for a real junction. Several depletion-capacitance formulas have been proposed that more correctly fit observed operation; however, SPICE uses a simple approach: For forward biases beyond some fraction (set by the parameter FC) of the value for ϕ, the capacitance is calculated as the linear extrapolation of the capacitance at the departure. This provides a continuous numerical result and does not affect circuit operation significantly because, for forward bias, the device capacitance is normally dominated by diffusion capacitance.

The diffusion charge (and therefore the capacitance) varies with forward current and is simply modeled as a transit time (model parameter TT) for the carriers to cross the diffusion region of the junction. The total charge is

$$\text{Diffusion charge} = \text{Device current} \times \text{Transit time} \tag{2-5}$$

and capacitance is the derivative, with respect to bias, of this:

$$\text{Diffusion capacitance} = \text{TT}\,\frac{\text{IS}}{\text{N}V_t}\,e^{V_j/\text{N}V_t} \tag{2-6}$$

Diffusion charge manifests itself as the *storage time* of a switching diode, which is the time required to discharge the diffusion charge in the junction, which must happen before the junction can be reverse-biased (switched off). Storage time is normally specified as the time to discharge the junction so that it is supporting only a fraction (typically 10%) of the initial reverse current. First, a forward current is supplied to the device to charge the junction. Then, as quickly as possible, a reverse current is supplied to the device. Internally, the junction is still forward-biased to a voltage nearly the same as before the switch in current; the junction is still conducting at the forward-current rate. This internal current adds to the external current as the total current discharging the junction. As the junction voltage decreases, the internal current falls off exponentially (according to the Shockley equation). This system is a relatively simple differential equation that can be solved to an explicit equation for the TT parameter (assuming complete discharge) as follows:

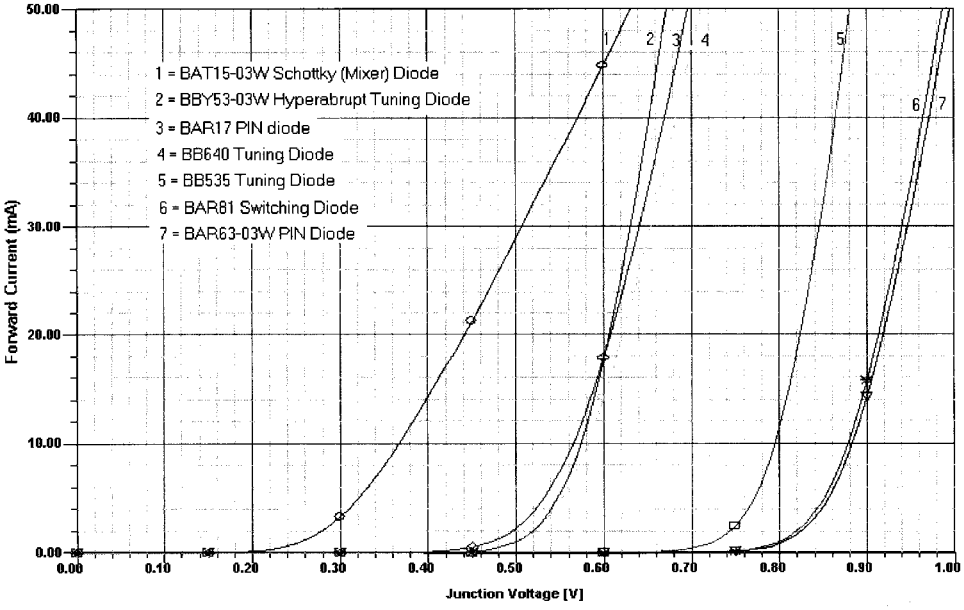


Figure 2-2 The dc *I–V* curves for seven diodes, showing the various barrier voltages that result from different doping profiles.

$$\text{Transit time} = \frac{\text{Storage time}}{\ln\left(\dfrac{I_F - I_R}{-I_R}\right)} \tag{2-7}$$

The diffusion charge dominates the reverse recovery characteristic of the diode. During the last part of the recovery, as the junction becomes reverse-biased, the depletion capacitance dominates. This causes the small tail at the end of the discharge cycle. Total capacitance is taken to be the sum of these capacitances: the depletion approximation dominates for reverse bias as the device current is small, and the diffusion proximity dominates for forward bias as the device current is large.

A special case for diode application is the switching diode, and its description and application will be part of a later chapter. Figure 2-2 shows the dc *I–V* curves, which indicate the different voltage potentials that are a result of the different doping profiles.

2-1-2 Mixer and Detector Diodes

Electrical Characteristics and Physics of Schottky Barriers. Schottky barrier diodes differ from junction diodes in that current flow involves only one type of carrier instead of both types. That is, in *n*-type Schottky barrier diodes the forward current consists of holes flowing from the *n*-type material into the metal.

Diode action results from a contact potential set up between the metal and the semiconductor, similar to the voltage between the two metals in a thermocouple. When metal is brought into contact with an *n*-type semiconductor (during fabrication of the chip), electronics diffuse out of the semiconductor into the metal, leaving a region under the contact that

has no free electrons ("depletion layer"). This region contains donor atoms that are positively charged (because each lost its excess electron), and this charge makes the semiconductor positive with respect to the metal. Diffusion continues until the semiconductor is so positive with respect to the metal that no more electrons can go into the metal. The internal voltage difference between the metal and the semiconductor is called the contact potential and is usually in the range of 0.3–0.8 volt for typical Schottky diodes. A cross section is shown in Figure 2-3.

When a positive voltage is applied to the metal, the internal voltage is reduced, and electrons can flow into the metal. The process is similar to thermionic emission of electrons from the hot cathode of a vacuum tube, except that the electrons are "escaping" into a metal instead of into a vacuum. Unlike the vacuum tube case, room temperature is "hot" enough for this to happen if enough voltage is applied. However, only those electronics whose thermal energy happens to be many times the average can escape, and these "hot electrons" account for all the forward current from the semiconductor into the metal.

One important thing to note is that there is no flow of minority carriers from the metal into the semiconductor and thus no neutral plasma of holes and electrons is formed. Therefore, if the forward voltage is removed, current stops "instantly," and reverse voltage can be established in a few picoseconds. There is no delay effect to charge storage as in junction diodes. This accounts for the exclusive use of barrier diodes in microwave mixers, where the diode must switch conductance states at microwave oscillator rates.

The current–voltage relationship for a barrier diode is described by the Richardson equation (which also applies to thermionic emission from a cathode). The derivation is given in many textbooks (e.g., Sze).

$$I = AA_{\mathrm{RC}}T^2 \exp\left(-\frac{q\phi_B}{kT}\right)\left[\exp\left(\frac{qV_J}{kT}\right) - M\right] \tag{2-8}$$

where A = area (cm^2)

A_{RC} = modified Richardson constant ($\mathrm{A}/K^2 \cdot \mathrm{cm}^2$)

k = Boltzmann's constant

T = absolute temperature (K)

ϕ_{B} = barrier height in volts

V_J = external voltage across the depletion layer (positive for external voltage) = $V - IR_S$

R_S = series resistance

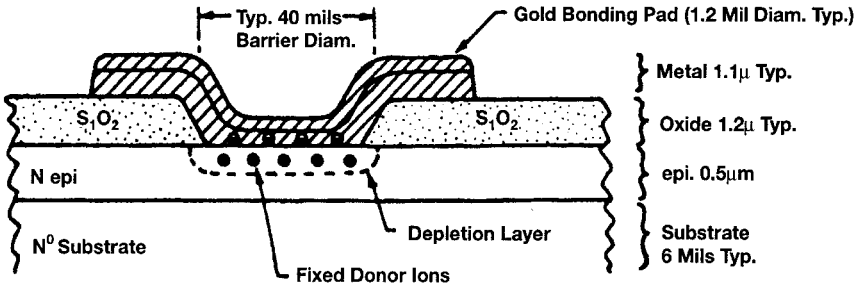


Figure 2-3 Schottky diode chip cross section.

M = avalanche multiplication factor
I = diode current in amperes (positive forward current)

The barrier height, ϕ_B, is typically a few tenths of a volt higher than the contact potential, ϕ_C (about 0.15 V higher than ϕ_C for silicon). This equation agrees well with experimental data for diodes without surface leakage, but it is difficult to use because A_{RC}, ϕ_B, and M are all dependent on applied voltage.

The major cause for variation in ϕ_B with voltage is the "image effect," in which the barrier height is lowered as the electric field near the metal is increased, especially at the edges. A better equation for circuit designers to use is one in which all parameters are independent of voltage and current. The simplest one that agrees reasonably well with Richardson's equation is

$$I = I_S\left[\exp\left(\frac{V_J}{0.028}\right) - 1 + \frac{K}{\left(1 - V_B/V\right)}\right] \tag{2-9}$$

where I_S = saturation current (a temperature-dependent quantity)
0.028 = nkT/q at room temperature (n = 1.08)
n = forward slope factor (derived from the variation of ϕ_B with forward voltage)
K = reverse slope factor (expressing the variation of ϕ_B with reverse voltage)
V_B = breakdown voltage (the voltage at which M = 1)

As before, V and I are considered positive for forward bias and negative for reverse bias.

Typical ranges for these parameters for microwave Schottky and point-contact mixer diodes are:

I_S = 10^{-12} to 10^{-5} A
n = 1.04–1.10
R_S = 2–20 Ω
K = 8–100
V_B = 2–20 V

The quantities I_S and 0.028 are strongly temperature dependent, while both R_S and V_B increase with temperature to a slight degree. R_S increases with current at high current levels (due to carrier velocity saturation) but is essentially independent of current at 10 mA and below for mixer diodes. Thus, for normal mixer and detector operation, R_S can be considered constant.

Agreement between Eqs. (2-8) and (2-9) is not perfect but Eq. (2-9) is much easier to use and is preferred by most circuit designers. A comparison of the two equations near zero bias gives the following relationship between zero-bias barrier height, ϕ_0, and saturation current:

$$I_S = AA_{RC}T^2\exp - \frac{q\phi_0}{kT}$$

$$\cong \left(\frac{10^7 A}{\text{cm}^2}\right)A\exp\left(-\frac{\phi_0}{0.026}\right) \quad \text{(for } n \text{ silicon at room temperature)} \tag{2-10}$$

Small-Signal Parameters. By combining Eqs. (2-9) and (2-10), the values of the parameters in Eq. (2-9) can be derived from a few simple measurements. Many specific equations can be derived, but the following are commonly used for production measurements:

$$R_S = \frac{V_{F10} - V_{F1} - 0.065}{0.009} \quad (\text{for } n = 1.08) \tag{2-11}$$

$$\phi_0 = \frac{V_{F1} - 0.001R_S + 0.28 + 0.12\log_{10} D}{1.08} \tag{2-12}$$

$$n = \frac{V_{F1} - V_{F0.1} - 0.0009R_S}{0.060} \tag{2-13}$$

$$K = \left(\frac{I_{R1}}{I_S} - 1\right)V_B \tag{2-14}$$

$$I_S = \exp\left(-\frac{V_{F1}}{0.028} + \frac{R_S}{28}\right) \quad (\text{in mA}) \tag{2-15}$$

where $V_{F0.1}$, V_{F1}, and V_{F10} are the forward voltages at 0.1, 1, and 10 mA, respectively, and I_{R1} is the reverse current at 1 volt. (The derivation of these equations requires that I_S be small compared to 0.1 mA.) The quantity D is the diameter of the metal–silicon contact in mils.

Measuring V_{F1} at 1 mA and 10 mA instead of some other current levels leads to the best accuracy for typical mixer diodes.

The total dynamic resistance for a forward-biased diode is given by

$$R_T = \frac{dV}{dI} = R_S + \frac{nkT}{q(I + I_S)} = R_S + R_B \tag{2-16}$$

and

$$R_B = \frac{28}{I + I_S} \quad \text{at room temperature (with } I \text{ and } I_S \text{ in mA, } n = 1.08) \tag{2-17}$$

This equation is also good at zero bias (unless K is very large or there is significant surface leakage). That is,

$$R_0 = R_S + \frac{28}{I_S} \tag{2-18}$$

For reverse voltages of a few volts, the dynamic resistance is dominated by the K term:

$$R_R = \text{Reverse resistance} = \frac{dV}{dI} \cong \frac{V_B}{KI_S} \tag{2-19}$$

For typical values of I_S, R_0 is larger than 5000 Ω and R_R is larger than 100 kΩ. For some zero-bias Schottky applications, it is desirable for R_0 to be made smaller than this.

The factors that determine R_S are (1) the thickness of the epitaxial layer, (2) the epitaxial layer doping level (N_D), (3) the barrier diameter, (4) the substrate resistivity ("spreading resistance"), (5) the contact resistances of the metals used for the barrier and the substrate contact, and (6) the resistance associated with the bonding wire or whisker. The barrier height is about 0.15 volt higher than the contact potential between the barrier metal and the semiconductor and is influenced by the method used to apply the metal, conditions at the edge of the junction, and the doping level. Saturation current depends on barrier height, junction area, and temperature; and the slope factors, n and K, depend on doping level, punchthrough voltage, and edge conditions.

Junction Capacitance. The capacitance of a Schottky barrier chip results mainly from two sources: the depletion layer under the metal–semiconductor contact and the capacitance of the oxide layer under the bonding pad (the so-called overlay capacitance). The bonding pad is required because the typical Schottky barrier diameter is so small that it is impractical to bond directly to the metal on the junction. If the semiconductor epitaxial layer is uniformly doped, the capacitance–voltage characteristic is similar to that of a textbook "abrupt-junction" diode.

$$C_J = \frac{\varepsilon_S \varepsilon_0 A'}{X_D} + C_0 \tag{2-20}$$

$$X_D = \sqrt{\frac{2\varepsilon_S \varepsilon_0 (\phi_C - V)}{qN}} \tag{2-21}$$

where ϕ = contact potential
C_0 = overlay (bonding pad) capacitance
ε_S = dielectric constant of the semiconductor ($\cong$ 12 for Si or GaAs)
N = doping level for the epitaxial layer
A' = effective contact area, including fringing corrections

In practical terms, the capacitance can be related to the 0-V barrier capacitance defined by

$$C_{B0} = \frac{\varepsilon_S \varepsilon_0 A'}{X_{D0}} \tag{2-22}$$

where

$$X_{D0} = \sqrt{\left(\frac{1.3 \times 10^{15}}{N_D}\right)\phi_C} \quad \text{(in } \mu\text{m)} \tag{2-23}$$

The resulting C–V relationship can be written

$$C_J = \frac{C_{B0}}{\sqrt{1 - V/\phi_C}} + C_0 \tag{2-24}$$

The contact potential, ϕ_C, is related to the barrier height as follows:

$$\phi_C = \phi_B - 0.026\left[1 + L_n\left(\frac{N_C}{N}\right)\right] \tag{2-25}$$

The theoretical meaning of these terms can be clarified by looking at Figure 2-4.

Parameter Trade-Offs

Barrier Height. The barrier height of a Schottky diode is important because it directly determines the forward voltage. In order to get a good noise figure, the LO drive voltage, V_L, must be large compared to V_T, which is essentially V_{F1}. Normally, it is best to have a low

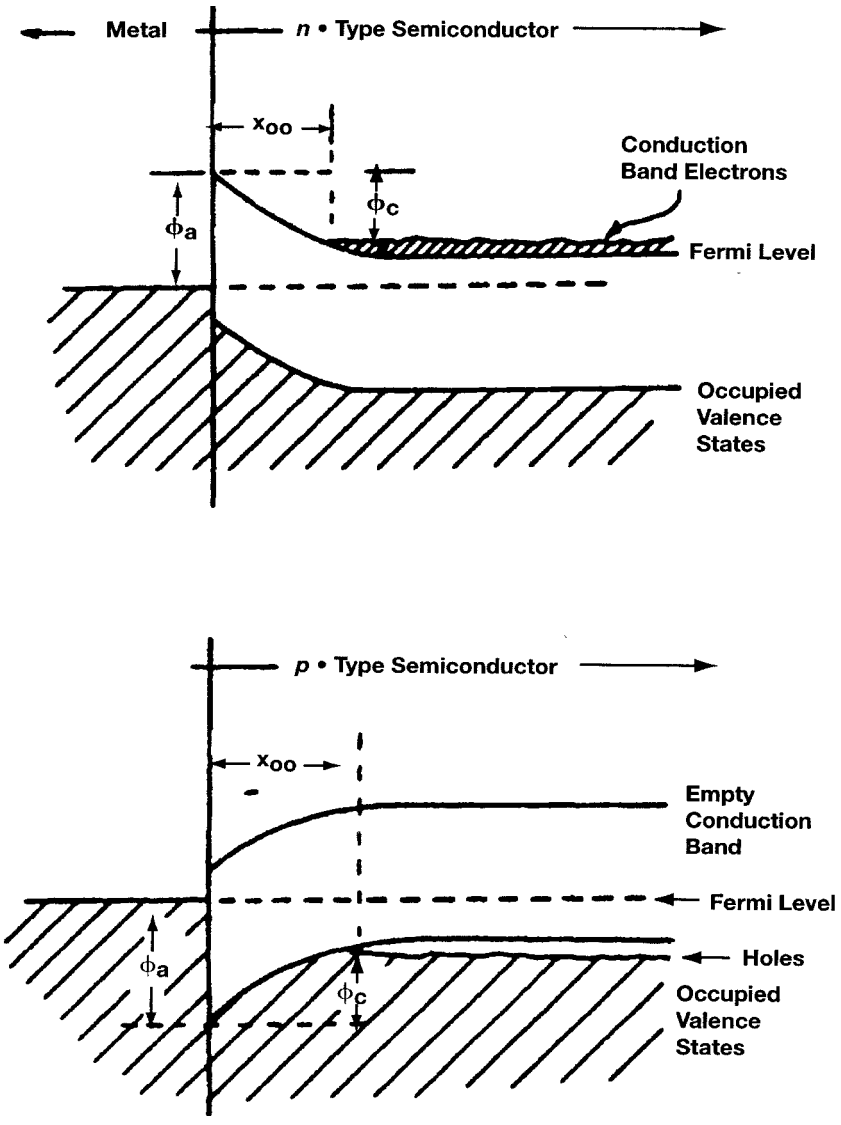


Figure 2-4 Schottky diode band diagrams.

Table 2-3 LO power and V_{F1} for various applications

Type	Typical V_{F1}	LO Power	Application
"Zero bias"	0.10–0.25	< 0.1 mW	Mainly for detectors
Low barrier	0.25–0.35	0.2–2 mW	Low-drive mixers
Medium barrier	0.35–0.50	0.5–10 mW	General purpose
High barrier	0.50–0.80	> 10 mW	High dynamic range

forward voltage (low V_{F1}), or low drive) diode, to reduce the amount of LO power needed. However, if high dynamic range is important, high LO power is needed, and the diode can have a higher V_F and should also have a high V_B (see Table 2-3).

Noise Figure Versus LO Power. At low LO drive levels, the noise figure is poor because of poor conversion loss, due to a too-low conduction angle. At very high LO drive levels, the noise figure again increases due to diode heating, excess noise, and reverse conduction.

If a high LO drive level is needed, for example, to get higher dynamic range, a high V_B (> 5 V) should be specified. However, Nature requires that you pay for this with higher R_S (lower f_C), so the noise figure will be degraded compared to what could be obtained with diodes designed for lower LO drive. Forward voltage and breakdown are basically independent parameters, but high breakdown is not needed or desirable unless high LO power is used.

Such a high-breakdown diode will have low reverse current (which is important only if the diode has to run hot).

Silicon Versus GaAs. Typical silicon Schottky diodes have cutoff frequencies in the 800–200-GHz range, which is good enough through the Ku band.

At the Ku band and above, or for image-enhanced mixers, higher f_C may be needed, which calls for the use of GaAs diodes. These have lower R_S due to higher mobility, which translates to cutoff frequencies in the 400–1000-GHz range.

However, if your IF is low, be careful; GaAs diodes have high 1/*f* noise. They also have high V_{F1}, so more LO power is required.

C_J Versus Frequency. There is quite a lot of latitude in choosing C_J. However, in general, the capacitive reactance should be a little lower than the transformed line impedance (Z_0). If Z_0 is not known, a good way to start is to use $X_C = 100\ \Omega$. Experience has shown that most practical mixers use an X_C near this value (a little higher in waveguide, and lower in 50-Ω systems). This translates to the following rule of thumb for choosing the junction capacitance of a diode for operation at frequency *f* (in GHz):

$$C_{J0} \approx \frac{100}{\omega}$$

$$\approx \frac{1.6}{f} \quad \text{(in pF)} \tag{2-26}$$

In order to evaluate possible tolerances, we show the range of forward current as a function of diode voltage (Figure 2-5), the junction capacitance as a function of the bias voltage

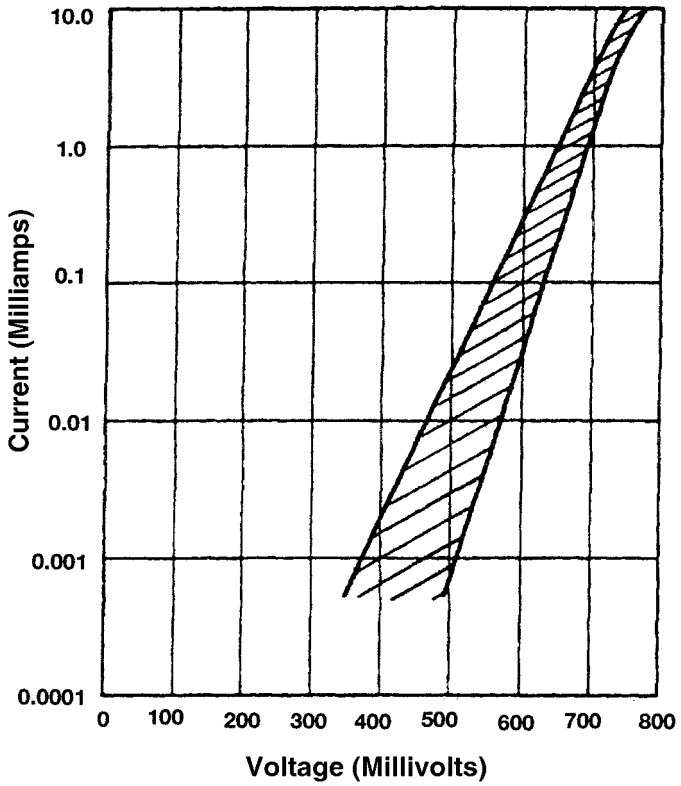


Figure 2-5 Forward dc characteristic curve range—current versus voltage.

(Figure 2-6), and finally some important RF parameters, such as noise figure and IF impedance, as a function of the local-oscillator drive (Figure 2-7).

Mixer Diodes. As an example of some of the parameters for mixer diodes, Table 2-4 gives data on some of the X-band mixer diodes. NF is measured at 9.375 GHz.

Linear Diode Model. Figure 2-8 shows the linear diode model. Its keywords appear in Table 2-5.

The circuit simulators, such as those supplied by Ansoft and Hewlett-Packard, provide a model library that has SPICE-type parameters for diodes (regular diodes, varactor diodes, and PIN diodes) as well as bipolar transistors and FETs, which will be discussed later.

2-1-3 PIN Diodes

Introduction. The PIN diode, in comparison with other microwave semiconductor devices, is fairly easy to understand. This makes possible the reduction of its complex behavior to relatively simple terms.

We do not attempt to describe the many possible microwave circuits in which PIN diodes are used. Rather, we attempt to explain the behavior of the diode in all aspects, giving the facts and some of the theory behind the facts. We offer the circuit designer the opportunity

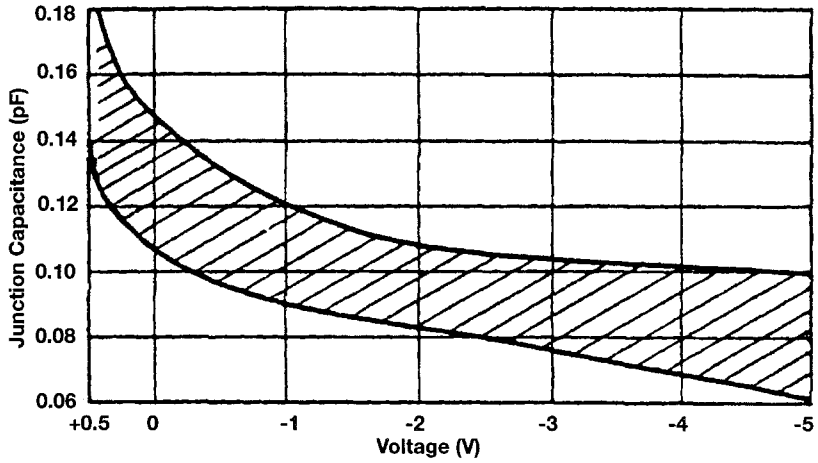


Figure 2-6 Junction capacitance range versus voltage.

to understand the PIN, so that he/she can understand its behavior in circuits. We assume the reader knows the circuit equations; to that knowledge we hope to add diode equations.

Most of the material presented consists of generalized data and explanations of the behavior of PIN diodes; we conclude with a brief description of circuit performance, test methods, and some hints on proper PIN specification writing.

The user can then evaluate the trade-offs involved in diode design and performance and be able to select the most nearly optimum diode from the wide range of diodes offered.

Large-Signal PIN Diode Model. Figure 2-9 shows the large-signal model for a PIN diode. Table 2-6 lists its keywords.

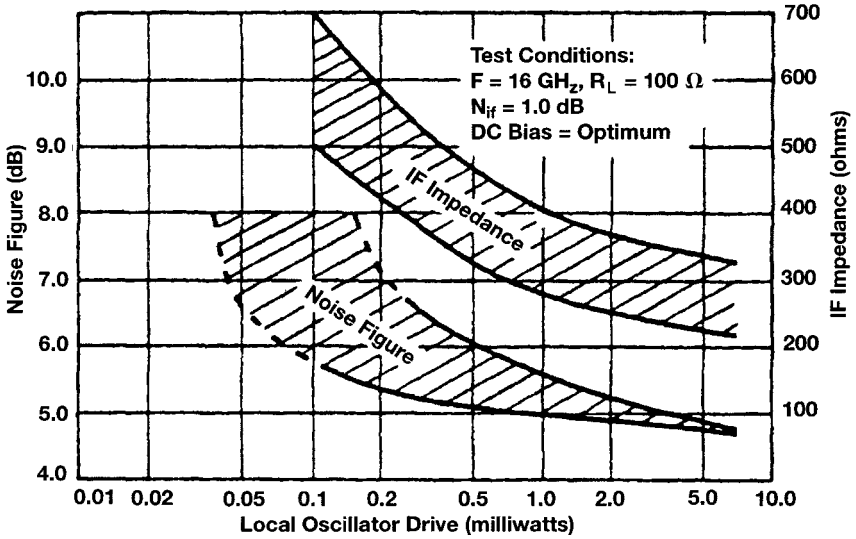


Figure 2-7 RF parameters versus local-oscillator drive level.

Table 2-4 X-band mixer diode data

Material	Barrier	Typical V_F (@1 mA)	Typical F_{C0} (GHz)	Typical R_S (Ω)	Typical C_{J0} (pF)	Maximum NF (dB)
n GaAs	High	0.70	1000	—	0.15	5.0[a]
n GaAs (BL)	High	0.70	500	—	0.15	6.0[a]
n GaAs (chip)	High	0.70	1000	—	0.15	5.3[a]
n Silicon (BL)	Low	0.28	150	6	0.20	6.5
n Silicon (quad)	Low	0.28	150	6	0.20	6.5
p Silicon (BL)	Low	0.28	150	12	0.20	6.5
n Silicon (BL)	High	0.60	100	8	0.20	6.5
n Silicon (quad)	High	0.60	100	8	0.20	6.5
n Silicon	Low	0.28	200	6	0.15	5.5
p Silicon	Low	0.28	200	18	0.14	6.0
p Silicon	Medium	0.40	150	150	0.12	6.5
n Silicon	Low	0.28	150	8	0.18	6.5
p Silicon	Low	0.28	150	12	0.18	6.5

[a]Specified for N_F = 1.0 dB.

Basic Theory: Variable Resistance. Intrinsic or "pure" silicon, as it can be grown in a laboratory, is an almost lossless dielectric. Some of is physical properties include:

Dielectric constant (relative)	12
Dielectric strength	400 V/mil (approximate)
Specific density	2.3
Specific heat	0.72 J/g · °C
Thermal conductivity	1.5 W/cm · °C
Resistivity	300,000 Ω/cm

Since a PIN diode is valuable essentially because it is a variable resistor, let us concentrate initially on the resistivity. Consider a volume comparable to a typical PIN diode chip, say, 20 mils in diameter and 2 mils thick. This chip has a dc resistance of about 0.75 MΩ. High

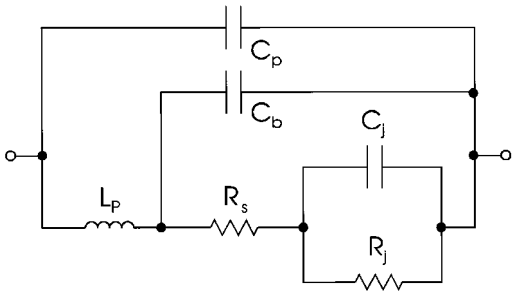


Figure 2-8 The linear diode model. This model is temperature dependent.

Table 2-5 Linear diode model keywords

Keyword	Description	Unit	Default
LP	Package inductance	henry	0.0
CB	Beam-lead capacitance	farad	0.0
CP	Package capacitance	farad	0.0
RS	Contact resistance	ohm	0.0
RJ	Junction resistance	ohm	
CJ	Junction capacitance	farad	

resistivity in any material indicates that most of the likely carriers of electric charge, electrons and holes, are tightly held in the crystal lattice and cannot "conduct."

In real life there are impurities, typically boron, that cannot be segregated out of the crystal. Such impurities contribute carriers, holes or electrons, that are not very tightly bound to the lattice and therefore lower the resistivity of the silicon.

Through various techniques we can adjust the level of impurities, called *dopants*, to produce resistivities ranging from 10 kΩ/cm (for good PIN diodes) to 0.001 Ω/cm (for substrates).

If the impurity adds "electrons" to the crystal, it is called a *donor*; if it adds holes it is called an *acceptor*. Boron adds holes, hence it is an acceptor, and the silicon plus boron combination is called *p*-type, or *positive*, because it has an excess of positive carriers. Phosphorus, on the other hand, is a *donor*, adding electrons, and the corresponding mix is *n*-type, or *negative*.

Notes on the PIN Diode Model

1. The PIN diode model is used to model a bias-dependent RF resistance for use in PIN diode circuits such as attenuators and switches. The resistance varies from RMAX to RS using the R function above. A typical R versus *I* characteristic is shown in Figure 2-10 with parameters IS = 5.96 nA, RS = 2.016, RMAX = 6500, K1 = 0.1272, K2 = 1.0, N = 2.077.
2. The transit-time parameter, TT, can also be used to approximately model a switching PIN diode's reverse-recovery time—a value often provided by diode manufacturers.
3. Diode breakdown can be modeled by specifying IBV and BV parameters.

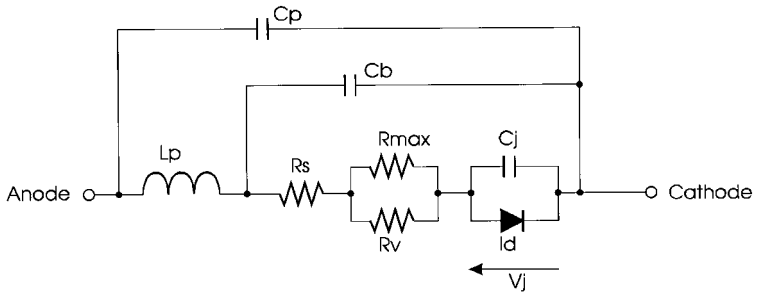


Figure 2-9 The large-signal PIN diode model. This model is temperature dependent.

Table 2-6 PIN diode model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
IS	Saturation current	ampere	1.0E–14
N	Emission coefficient		1.0
IBV	Magnitude of current at the reverse breakdown voltage	ampere	1.0E–10
BV	Magnitude of the reverse breakdown voltage	volt	∞
FC	Coefficient for forward-bias depletion capacitance		0.5
CJ0	Zero-bias *pn* junction capacitance	farad	0.0
VJ	Built-in junction potential	volt	1.0
M	*pn* Junction grading coefficient		0.5
GC1	Varactor capacitance polynomial coefficient 1	/volt	0.0
GC2	Varactor capacitance polynomial coefficient 2	/volt^2	0.0
GC3	Varactor capacitance polynomial coefficient 3	/volt^3	0.0
TT	Transit time	second	0.0
K1	Variable resistance coefficient	volt	0.0
K2	Variable resistance current exponent		1.0
RMAX	Maximum resistance of PIN intrinsic region	ohm	0.0
KF	Flicker noise coefficient		0.0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
AREA	Area multiplier		1.0
	Extrinsic Model		
RS	Series resistance (minimum resistance of PIN diode)	ohm	0.0
CP	Package parasitic capacitance	farad	0.0
CB	Beam-lead parasitic capacitance	farad	0.0
LP	Package parasitic inductance	henry	0.0

4. The reverse-bias capacitance characteristics can be more accurately modeled than the common expression derived from *pn* junction theory. The capacitance grading coefficient exponent can be expressed as a polynomial function of voltage by specifying values for GC1, GC2, and GC3.
5. The PIN diode model was derived from J. Walston, "SPICE Circuit Yields Recipe for PIN Diode," *Microwaves & RF*, November 1992.
6. Following Simon M. Sze, *Physics of Semiconductor Devices* (New York: John Wiley, 1982), the variable resistance may be modeled by setting

$$\text{K1} = \frac{3}{8} V_T \frac{W^2}{D_a \tau_a} \qquad \text{K2} = 1.0$$

where W is the width of the intrinsic region, D_a is the ambipolar diffusion coefficient, τ_a is the ambipolar lifetime, and V_T is the thermal voltage.

There are many concepts important to the physicist, but not to the diode user, that elaborate on the impact of impurities on the behavior of silicon. These can be studied in Watson [1]. The more carriers added, the lower the resistivity.

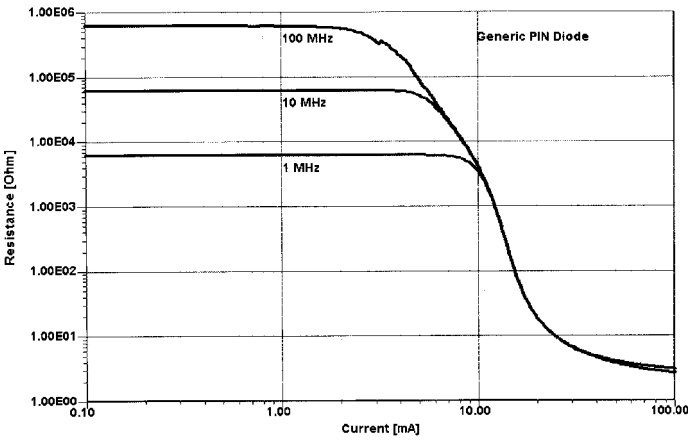


Figure 2-10 Simulated PIN diode resistance as a function of dc at 1, 10, and 100 MHz.

If one wished to vary the resistance of a given diode, in principle he/she could bring it into a semiconductor laboratory, add or subtract carriers as desired, and perhaps even make the process reversible. However, this is a slow, expensive, and impractical way to make a variable resistor; one would be better advised to take a wrench and a soldering iron and replace a component.

The PIN diode derives its value from the fact that the free-charge-carrier concentration in silicon, and hence its resistance, can be varied electronically by means of current from a simple bias supply. This can be done rapidly (in nanoseconds in some cases), reversibly, repeatably, and accurately. The thing that makes this possible is called a *junction*, the interface between the relatively pure silicon in the middle of the PIN diode (the I stands for *intrinsic*) and the heavily doped layers on either end, p^+ and n^+. The p^+ region is rich in holes; the n^+ region is rich in electrons. Both of these regions have low resistance. The I region is the variable resistive element in the diode (see Figure 2-11). In the absence of any external bias, internal effects within the crystal keep the charges fixed; the resistance of the I region is high.

When the p^+ region (anode) is biased positively with respect to the n^+ region (cathode), the interface potential "barrier" is overcome, and direct current flows in the form of holes

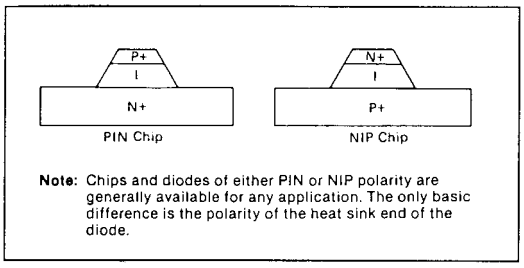


Figure 2-11 General outline of PIN diode construction.

streaming from p^+ toward n^+, with electrons moving in the opposite direction; we say that free carriers have been injected into the I region. The resistance of the I region becomes low.

The number of free carriers within the I region determines the resistivity of the region and thus the resistance of the diode.

Consider "one hole" and "one electron" drifting in opposite directions in the I region under the impetus of the applied field. Under certain conditions, imperfections in the silicon may cause these carriers to recombine. They are no longer available to constitute current or to lower the resistivity of the I region.

It can be shown that the amount of "recombination" between holes and electrons that continuously takes place in a semiconductor is governed by a property of the lattice called *lifetime*. In fact, lifetime is defined as the reciprocal of recombination rate.

Thus, $Q_S = Q_0 \exp(-t/T_L)$, where Q_S is the total amount of free charge "stored" in the I region, and T_L is the lifetime, or the mean time between recombination events. In the steady-state condition, the bias supply must deliver current to maintain a constant Q_S. The required current is

$$I_{\text{dc}} = \frac{Q_S d}{dt} = \frac{-Q_S}{T_L} \tag{2-27}$$

or $Q_S = I_{\text{dc}} T_L$, dropping the minus sign.

Ignoring some details that are not crucial to this section, we can now calculate the resistance of a given diode, of area A and thickness W. (W stands for *base width*, the width or thickness of the intrinsic layer.) The p^+ and n^+ regions have essentially zero resistance, as they are very heavily doped.

The resistivity of a given material is inversely proportional to the number of free carriers, N, and the mobility (not quite the same as velocity) of the carriers. Thus,

$$\rho = \frac{1}{q(\mu_n N + \mu_p P)} \tag{2-28}$$

where μ_n and μ_p are mobilities of electrons and holes, and N and P are numbers of electrons and holes. Simplifying, we find

$$\rho = \frac{C}{Q_S d} \tag{2-29}$$

where C is a collection of constants, and $Q_S d$ is the stored charge density (numbers per unit volume). For our piece of silicon, volume is WA, and

$$Q_S d = \frac{Q_S}{WA} \tag{2-30}$$

The resistivity is

$$\rho = \frac{CWA}{Q_S} \tag{2-31}$$

and the resistance is

$$R_S = \rho \frac{W}{A} = \frac{CW^2}{I_{\text{dc}} T_L} \tag{2-32}$$

This is a fundamental equation in PIN diode theory and design.

Rigorous analysis shows that

$$R = \frac{2kT/q}{I_f} \sinh\left(\frac{W}{2\sqrt{DT_L}} \tan^{-1}\left[\sinh\frac{W}{2\sqrt{DT_L}}\right]\right) \tag{2-33}$$

where k = Boltzmann's constant
T = temperature (in kelvins)
D = diffusion coefficient = $\mu kt/q$

For most PIN diodes, W/DT_L is less than unity, and the equation simplifies to the simple equation above.

Typical data on R_S as a function of bias current are shown in Figure 2-12. Many design choices are available, as the data indicate. Many combinations of W and T_L have been developed to satisfy the full range of applications.

Breakdown Voltage, Capacitance, Q Factor. The previous section on R_S explained how a PIN can become a low resistance, or a "short." This paragraph will describe the other state: a high impedance, or an "open." Clearly, the better PIN diode is the one that has the better on/off ratio at the frequency and power level of interest.

If we return to the undoped or intrinsic I region, we note that it is an almost lossless dielectric. As such, it has a dielectric strength of about 400 V/mil, and all PIN diodes have

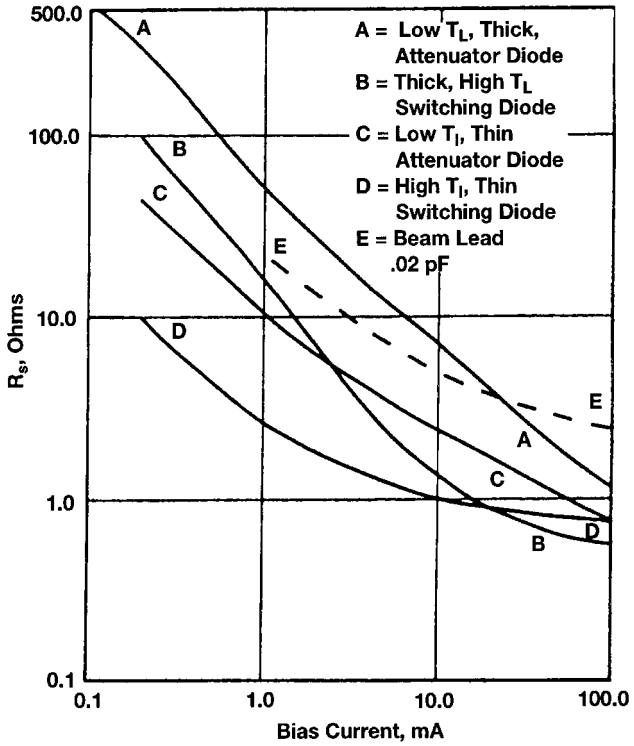


Figure 2-12 Typical series resistance as a function of bias (1 GHz).

a parameter called V_b, breakdown voltage, which is a direct measure of the width of the I region. Voltage in excess of this parameter results in a rapid increase in current flow (called avalanche current); see Figure 2-13. When the negative bias voltage is below the bulk breakdown of the I region, a few nanoamperes will be drawn. As V_b is approached, the leakage current increases often gradually, as is exaggerated in the curve. This current is primarily caused by less-than-perfect diode fabrication, although there is some contribution from temperature. Typically, the leakage current occurs at the periphery of the I region. For this reason, various *passivation* materials (silicon dioxide, silicon nitride, hard glass) are grown or deposited to protect and stabilize this surface and minimize leakage. These techniques have been well advanced over the years, and PIN diode reliability has improved as a result.

Most diodes are specified in terms of minimum V_b for a nominal leakage, usually 10 μA.

It will be noted later that RF voltage swings in excess of the rated V_b are permitted, for the mechanisms causing leakage current do not always respond at radiofrequencies. However, bulk breakdown is effectively instantaneous, and that voltage should never be exceeded.

The next characteristic of our "open" circuit is the capacitance. In simplest form, the capacitance of a PIN diode is determined by the area and width of the I region and the dielectric constant of silicon; however, we have discussed the fact that intrinsic material does contain some carriers and therefore has some conductivity. An *E* field could not exist unless all these carriers were swept out, or depleted.

Application of a reverse bias accomplishes this. At zero bias, the excess carriers on either side of the junction are separated, held apart, by "built-in" fields. This is the contact potential (about 0.5 V for silicon). If there are only a few excess carriers in the I region, this "potential" can separate the charges more easily. The junction "widens" in the sense that, starting at the p^+ and I interface, there is a region of no free carriers called the *depletion zone*. Beyond this depletion zone, the I region still contains the free charges with which it started. With the application of reverse bias, the depletion zone widens.

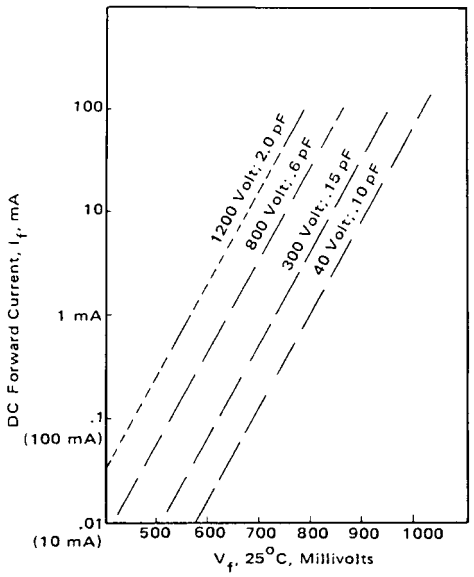


Figure 2-13 Current versus voltage for various PIN diodes.

Eventually, at a bias equal to a so-called *punchthrough* voltage (V_{PT}), the depletion zone fills the entire I region. At this voltage, the 1-MHz capacitance bottoms out and the diode Q reaches its maximum. Figure 2-14 illustrates the equivalent circuit of the I region before punchthrough.

Some very interesting facts can be derived from this model. Consider the undepleted region; this is a lossy dielectric consisting of a volume (area A, length ℓ) of silicon of permittivity 12 and resistivity ε. The capacitance is

$$12\frac{\varepsilon_0 A}{\ell} \tag{2-34}$$

and the admittance is

$$2\pi\frac{12\varepsilon_0 A}{\ell} \tag{2-35}$$

The resistance is $\varepsilon\ell/A$ and the conductance is $A/\varepsilon\ell$.

At very low frequencies, the undepleted zone looks like a pure resistor. At very high frequencies it looks like a lossy capacitor. The "crossover" frequency depends on the resistivity of the I-region material. For ε of 160 Ω/cm, the frequency is 1 GHz. Higher resistivity is generally used for PINs—say, 1000 Ω/cm, and the crossover frequency is 160 MHz.

Diode manufacturers measure junction capacitance at 1 MHz; clearly, what is measured is the depletion-zone capacitance.

If the I region thickness is W and the depletion is X_d, the undepleted region is $(W - X_d)$.

The capacitance of the depleted zone is, proportionally,

$$\frac{1}{X_d} \tag{2-36}$$

and of the undepleted zone,

$$\frac{1}{W - X_d} \tag{2-37}$$

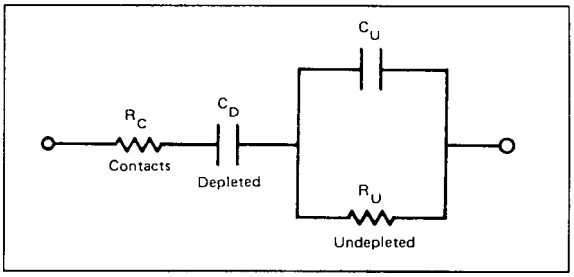


Figure 2-14 Equivalent circuit of I region before punchthrough.

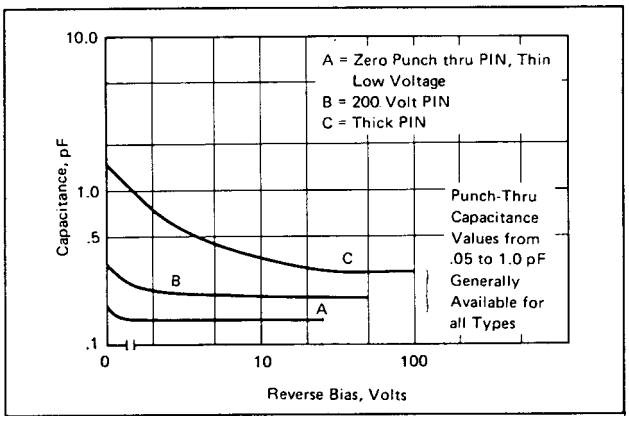


Figure 2-15 Typical 1-MHz capacitance.

The 1-MHz capacitance as a function of reverse bias is seen in Figure 2-15. The 1-MHz capacitance decreases with bias until punchthrough, where $X_d = W$. However, at microwave frequencies well above the crossover, the junction looks like two capacitors in series:

$$C_T = \frac{C_d C_u}{C_d + C_u} \propto \frac{1}{W} \tag{2-38}$$

That is, the microwave capacitance tends to be constant, independent of X_d and bias voltage.

However, since the undepleted zone is lossy, an increase in bias up to the punchthrough voltage reduces the loss.

At any given frequency, the equivalent network can now be drawn as Figure 2-16. R_v is now the equivalent series resistance of the undepleted region. Typical R_v data are shown in Figure 2-17. An alternate equivalent network is shown in Figure 2-18, and typical R shunt data are shown in Figure 2-19.

A good way to understand the effects of series resistance is to observe the insertion loss of a PIN chip shunt mounted in a 50-Ω line, as shown in Figure 2-20. An accepted way to include reverse loss in the figure of merit of a PIN diode is to write

$$Q = \frac{1}{2\pi + \sqrt{R_s R_v C}} \tag{2-39}$$

where R_s and R_v are measured under the expected forward- and reverse-bias conditions at the frequency of interest.

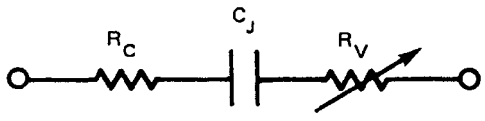


Figure 2-16 Simplified equivalent series circuit.

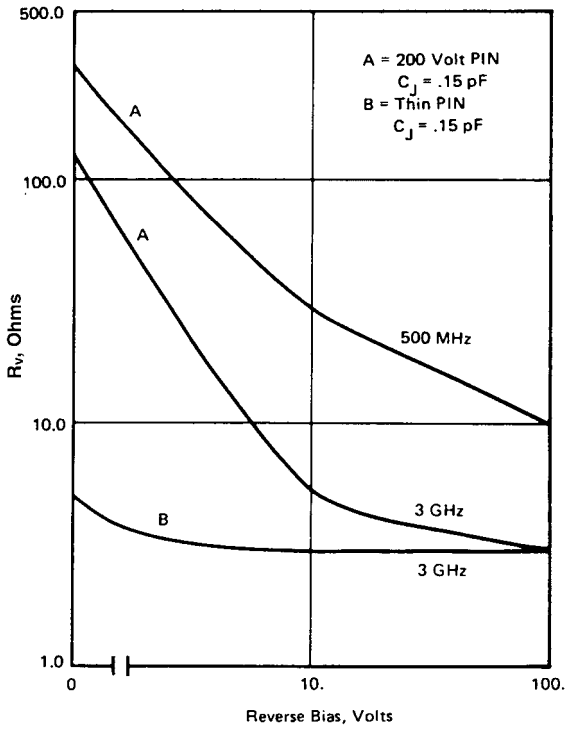


Figure 2-17 Reverse series resistance R_V.

The punchthrough voltage is a function of the resistivity and thickness of the I region. It is advisable to measure loss as a function of bias voltage and RF voltage to determine if the correct diode has been selected for your application. (*Note*: At frequencies below crossover, and diodes with thin I regions, the effective junction capacitance can increase substantially at low forward bias, on the order of 1–200 μA.)

Incidentally, if you are working with PIN or NIP chips that do not have an opaque covering, note that PIN diodes are photosensitive. Incident light causes photogeneration of carriers in the I region, increasing the chip's insertion loss.

PIN Diode Applications. If the intrinsic zone is thick (10–100 μm), we then have a high-reverse-voltage rectifier with a low forward-voltage drop at high current or, in other

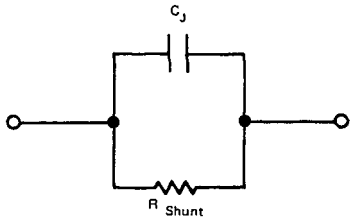


Figure 2-18 Simplified equivalent circuit, shunt.

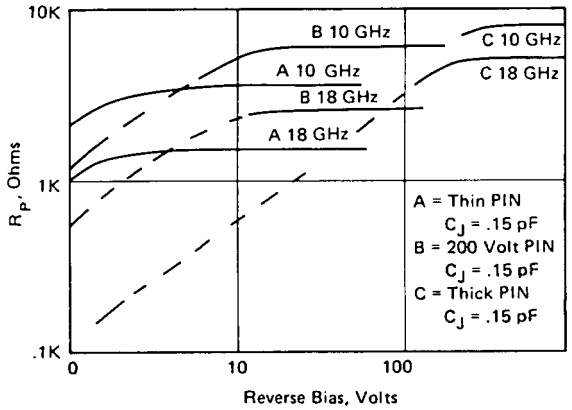


Figure 2-19 Reverse shunt resistance, R_p.

words, a highly efficient rectifier. The low forward voltage results from the fact that the conductivity of the I zone can be modulated by large amounts of charge carriers injected from the *p* and the *n* zones.

Another application of PIN diodes is the high-frequency (HF) field. Here, the fact is exploited that, due to the long carrier lifetime at frequencies beyond approximately 10 MHz, a rectifying effect will no longer occur and the PIN diode rather behaves like a real resistance, the magnitude of which depends on the forward direct current passed by the device and

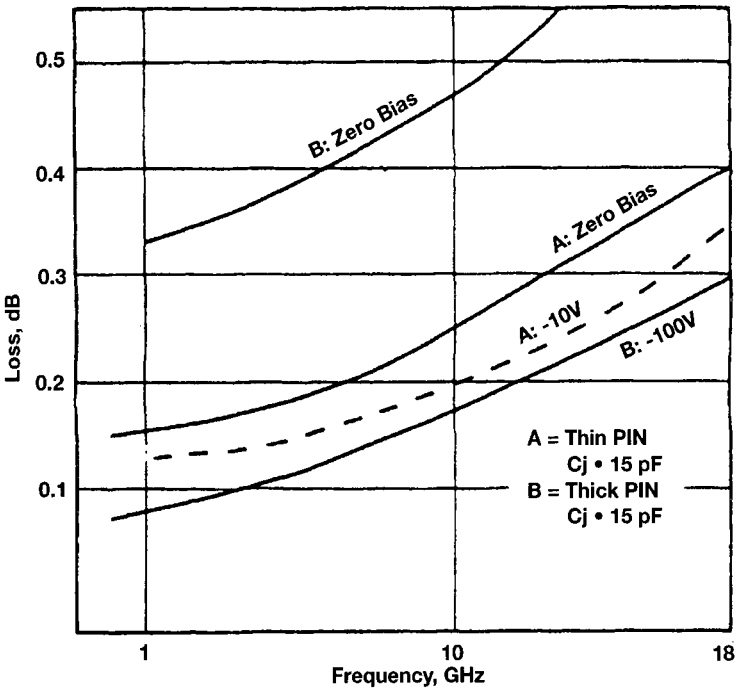


Figure 2-20 Insertion loss versus frequency.

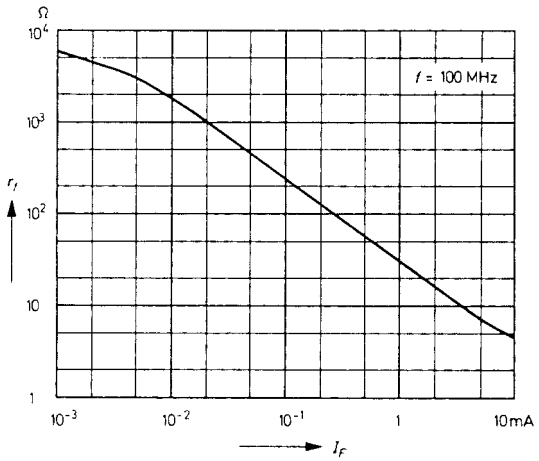


Figure 2-21 Forward resistance versus forward current.

produces an equal effect on both half waves of the HF signal. In view of this behavior, the PIN diode can be used as a switch or a variable resistor for HF signals. Thus it becomes possible, for example, to subject an HF signal to amplitude modulation by means of an AF-controlled PIN diode.

An important application of PIN diodes that has found favor in recent times is for dc-operated attenuators in TV tuners and antenna distribution amplifiers. Figure 2-21 shows the real HF forward resistance r_f as a function of the forward current I_f, measured at 100 MHz. Figures 2-22 and 2-23 show second-order IMD and cross-modulation for PIN diodes.

Applying the PIN in Amplitude Control of High-Frequency Signals. In conventional transistorized TV tuners, automatic gain control (AGC) is usually achieved by varying the emitter current of the input transistor. This method exploits the fact that the input transistor exhibits maximum gain at a certain level of emitter current and that this gain decreases when the emitter current is either increased or decreased relative to this point. Since modulation capacity and cross-modulation resistance grow with the emitter current, "upward" gain control has lately become the preferred solution, and according to it, gain reduction is achieved by an increase of the input transistor's emitter current. An input stage based on this design, therefore, is least resistant to cross-modulation when receiving a signal from a weak transmitter because it is operating at or below the point at which AGC action begins. This effect is particularly disadvantageous when the signal of a weak transmitter is to be received in the presence of a strong local transmitter. The unsatisfactory cross-modulation properties of such an input stage manifest themselves more and more frequently in the form of image perturbations as the TV reception band is crowded by an increasing number of transmitters. Another disadvantage is the variation in the transistor parameters that results from the control action. The variations affect antenna matching at the input and the response of the RF filter at the output.

To obviate these problems, methods being adopted today provide for the input stage to be equipped with cross-modulation-resistant transistors with fixed bias, gain control being provided by a variable attenuator preceding the transistor(s). The input-filter termination

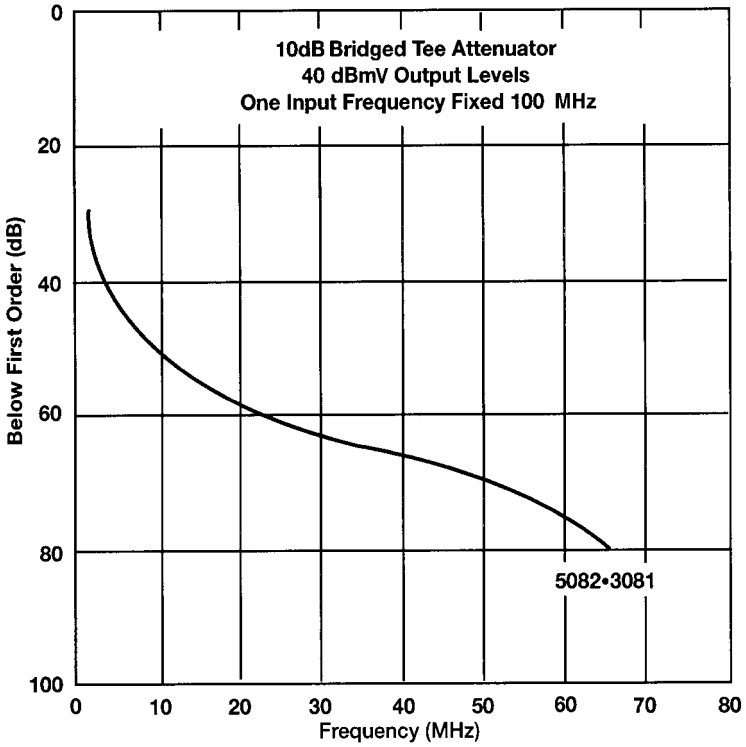


Figure 2-22 Second-order IMD in a PIN diode. Equation (1-48) can be used to determine the diode's IP_2 from the absolute values of its fundamental and IM_2 outputs.

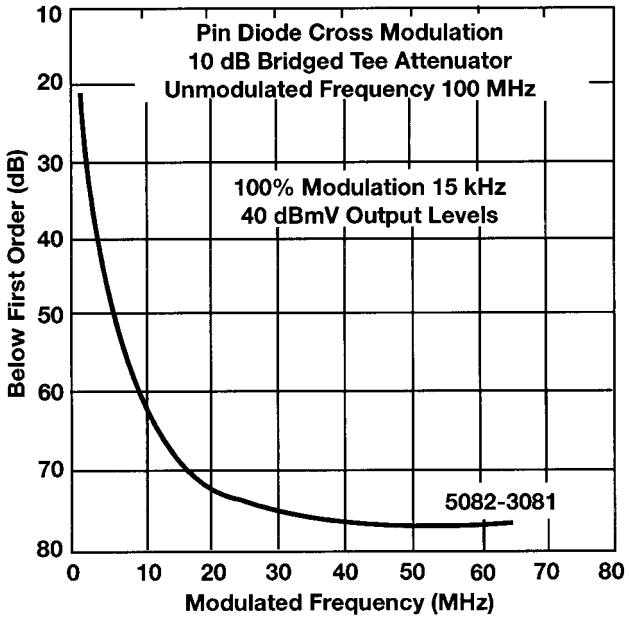


Figure 2-23 Cross-modulation in a PIN diode. Equation (2-61) relates cross-modulation level to IMD.

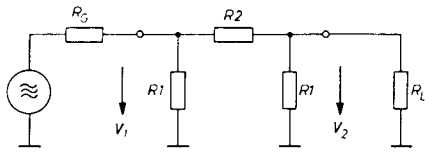


Figure 2-24 Basic circuit of a π-mode filter attenuator.

provided by this attenuator must be independent of the degree of attenuation to preserve the filter's response characteristics.

The first two requirements are met by the properties of the PIN diode and the latter requirement can be met by designing a PIN diode control in the form of a π network, a schematic of which is shown in Figure 2-24. The attenuation of a matched π network is expressed by the following equations:

$$\frac{a}{\text{dB}} = 20 \log \frac{Z + R_2}{R_2 - Z} \tag{2-40}$$

$$\frac{a}{\text{dB}} = 20 \log \left(\frac{R_1}{Z} + \sqrt{\frac{R_1^2}{Z^2} + 1} \right) \tag{2-41}$$

Figure 2-25 plots the relationship between the resistance levels of R_1 and R_2 and the input and output voltages of the π network. Figure 2-26 shows the circuit diagram of an implementation of the π network composed of PIN diodes. The resistance characteristic required according to Figure 2-25 is achieved approximately by varying the control current I_{co}.

If the control current I_{co} in Figure 2-26 equals zero, then—due to the fact that $R_1 = R_2$ and $R_3 = R_4$—the auxiliary voltage V_h causes forward currents of equal magnitude to flow through diodes D_1 and D_3. The voltage drops across R_1 and R_2 are of equal magnitude, so that no current passes through diode D_2. Therefore, the latter presents a high resistance, whereas D_1 and D_3 present a low resistance. Under these conditions, the network produces maximum attenuation. A control current I_{co} reduces the forward currents through D_1 and D_3 and allows a forward current to be passed by diode D_2. When it has attained its maximum value, diodes

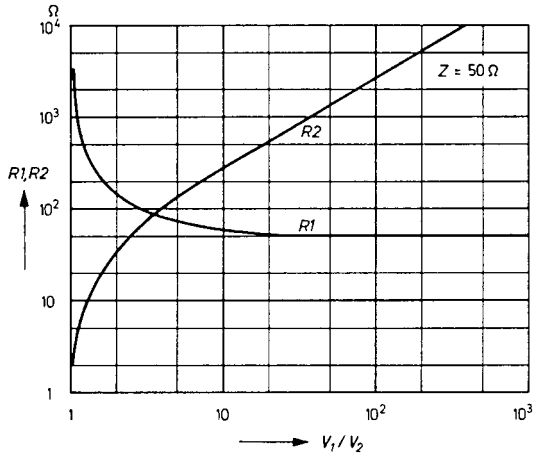


Figure 2-25 Values of R_1 and R_2 in Figure 2-24 versus relation of input and output voltage.

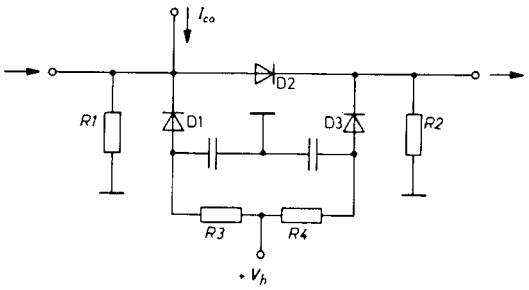


Figure 2-26 PIN diode π-mode attenuator.

D_1 and D_3 are blocked—that is, they are high-ohmic—and D_2 is low-ohmic. Under these conditions, the network produces its minimum attenuation.

If the maximum attenuation of the π network is to be increased, it is possible to make the two shunt diodes D_1 and D_3 more low-ohmic than the impedance Z, although this would reduce the reflection coefficient of attenuation. For example, if a reflection coefficient of about 7 dB is accepted over the entire control range, then a maximum attenuation of 25–30 dB could be achieved in the VHF range. The series inductances of the transverse diodes and the shunt capacitance of the series diode are responsible for this relatively unsatisfactory value. Improved attenuation is brought about with the integrated π network described in the next section.

Example: A PIN Diode π Network for TV Tuners. The since-discontinued TDA1053 is an example of an integrated PIN π network attenuator intended for TV-tuner use. It comprises three silicon planar PIN diodes connected to form a π network (Figure 2-27) and serves for the electronic amplitude control of the input signals of TV tuners and antenna distribution amplifiers in the 40–1000-MHz range. Its input and output impedances remain constant over the entire control range. This can also be achieved with discrete diodes.

The TDA1053 was normally supplied with vertical leads. The characteristics stated below apply to devices of this configuration.

MAXIMUM RATINGS OF INDIVIDUAL DIODES

Reverse voltage	V_R	30	V
Forward current at $T_{amb} = 25$ °C	I_F	50	mA
Junction temperature	T_j	125	°C
Storage temperature range	T_s	−55 to +125	°C

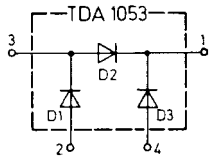


Figure 2-27 Internal circuitry of the TDA1053.

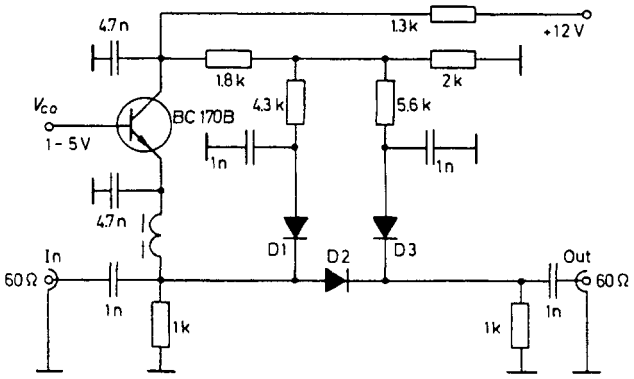


Figure 2-28 Test and application circuit.

MAXIMUM RATINGS OF THE π NETWORK

Ambient operating temperature range when operating according to Figure 2-28	T_{amb}	100	°C

CHARACTERISTICS OF INDIVIDUAL DIODES AT T_{amb} = 25 °C

Forward voltage at I_F = 50 mA	V_F	< 1.2	V
Forward current at V_R = 15 V	I_R	< 500	nA
Differential forward resistance:			
At I_F = 10 mA, f = 100 MHz	r_f	5	Ω
At I_F = 10 µA, f = 100 MHz	r_f	1.4	kΩ

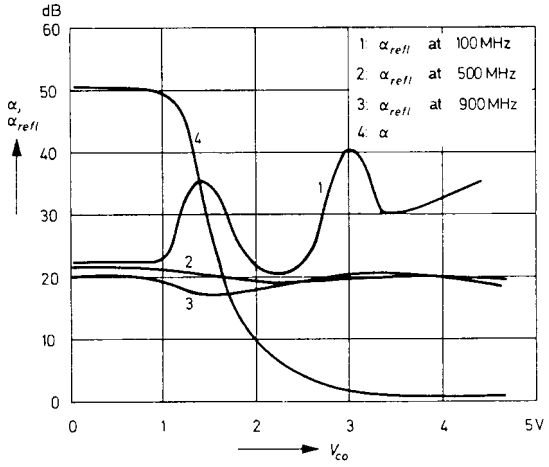


Figure 2-29 Attenuation and reflection attenuation versus control voltage.

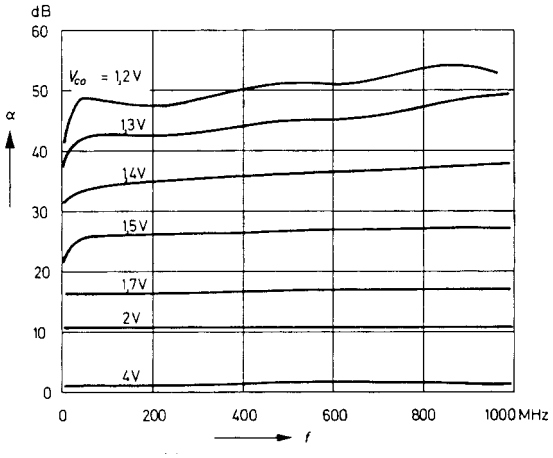


Figure 2-30 Attenuation versus frequency.

CHARACTERISTICS IN THE TEST CIRCUIT (FIGURE 2-28) AT T_{amb} = 25 °C

Voltage for 1% cross-modulation	V_{cr}	1	V
Attenuation in the 40–1000-MHz range:			
At V_{co} = 1.5 V (1–2 V)	a_{max}	45 (> 36)	dB
At V_{co} = 5 V (4–5 V)	a_{min}	1.5 (< 2)	dB
Reflection coefficient in the 40–1000-MHz range over the entire control range, depending on circuit design	a_{refl}	20 (> 16)	dB

These data reveal that the compact design of the three PIN diodes in a common 50B4 plastic package guarantees favorable values for minimum and maximum attenuation, as well as reflection attenuation. The test and application circuit shown in Figure 2-28 also comprises the transistor control-signal amplifier. The typical characteristic of the attenuation and the reflection attenuation for this circuit are shown in Figure 2-29, as a function of the control voltage V_{co}. Figure 2-30 shows the attenuation at different control voltages, as a function of frequency.

2-1-4 Tuning Diodes

Introduction. In recent years, continuous development of tuning diodes—also known as *varactors* or *varicaps*—together with increased commercial and military use has led to substantial improvement in Q, reproducibility, and reliability. Concurrently, new techniques for producing and controlling a hyperabrupt dopant profile in the semiconductor permit the capacitance–voltage law to be much faster than the classical square root or cube root behavior.

Current tuning diode materials include silicon and gallium arsenide; silicon is favored for low cost and lower Q applications from HF through microwave frequencies. Hyperabrupt varactors, also of silicon, are finding many applications in commercial television tuners, where their high tuning ratios, linear tuning, and low cost are needed. New developments include low-capacitance hyperabrupts for microwave and wireless applications.

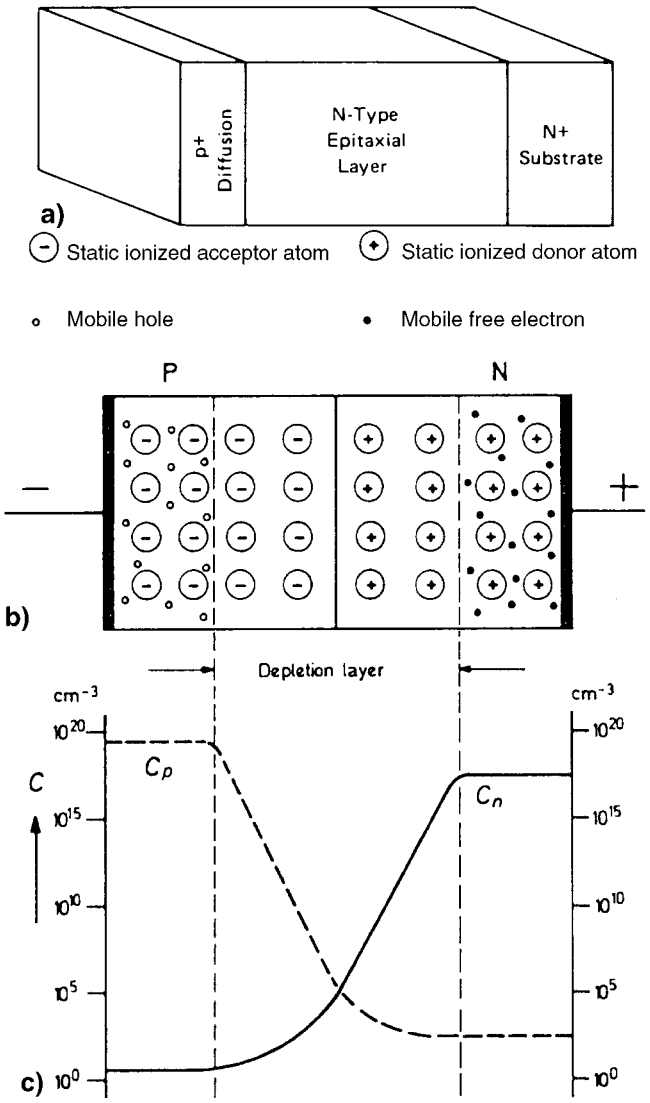


Figure 2-31 (a) Basic PIN structure; (b) cross section of a reverse-biased *pn* junction; and (c) density distribution of free charge carriers.

Gallium arsenide is used with high operating frequency and dictates the highest Q possible, as in parametric amplifiers and millimeter multipliers.

This section will acquaint the reader with tuning diodes: how they work, and what they can or cannot be expected to do in an electronic circuit. The basic properties of a tuning diode will be described in terms of the parameters that manufacturers use in characterizing them. The following topics will also be addressed:

- Capacitance ratio with respect to voltage and voltage breakdown
- Q as a function of design and operating conditions

- Stability—leakage current, temperature coefficient, and post-tuning drift
- Distortion products
- Packaging parasitics
- Applications—suggestions on how to specify a varactor

Tuning Diode Physics

Introduction. All junction diodes are made up of the same physical parts: a *pn* junction, a carefully controlled epitaxial layer, and a very-low-resistance substrate. These parts are shown in Figure 2-31.

No matter what type of junction device we are discussing—a tuning diode, a step-recovery diode, or a PIN diode—these parts are all present; the main difference between these devices is the resistivity and thickness of the epitaxial layer. Tuning diodes and multiplier diodes need epitaxial layers where both the resistivity and thickness are carefully controlled.

Abrupt Junction. An abrupt-junction diode is one in which the p^+ diffused region of the diode is much more highly doped than the epitaxial layer. Also, the high doping drops to the doping level of the epitaxial layer in a distance that is short compared to the epitaxial layer thickness, and the doping level of the epitaxial layer is constant over its thickness. This is shown in Figure 2-32, with the corresponding C–V curve shown in Figure 2-33. When these requirements are satisfied, the diode capacity, diode area, epitaxial layer doping level, and diode voltage are related by

$$\frac{C(V)}{A} = K\left(\frac{N}{V+\varphi}\right)^n \tag{2-42}$$

where $C(V)$ = capacitance of the diode at voltage V
A = area of the diode
N = doping level of the epitaxial layer

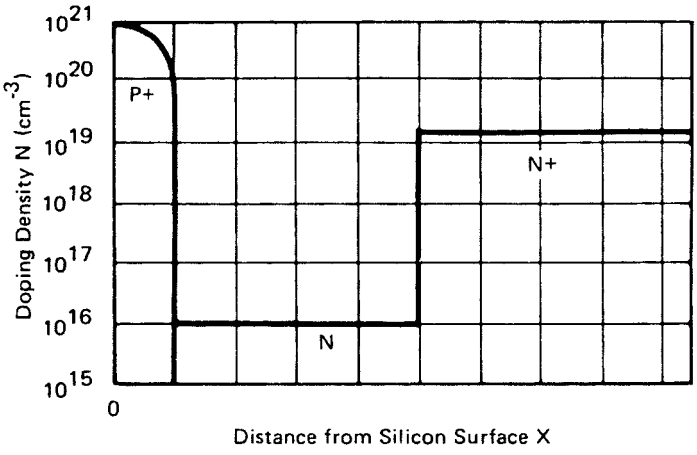


Figure 2-32 *N–X* abrupt junction.

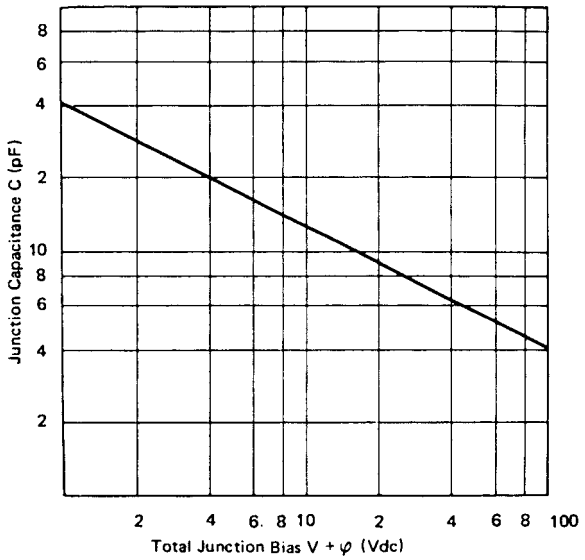


Figure 2-33 Capacitance versus total junction bias for abrupt-junction diode.

V = voltage applied to the diode
φ = built-in potential of the diode (0.6–0.8 V)
n = slope of diode C–V curve; $n \approx 0.5$ for an abrupt-junction diode
K = constant

As a consequence of the physical properties of a *pn* junction, a depletion layer is formed between the *p* and *n* regions whose width depends on the voltage applied to the diode. The capacitance of the diode is inversely proportional to the width of the depletion layer; that is, $C \sim 1/\ell_0$. In addition, the series resistance of the diode is proportional to the width of the undepleted epitaxial layer. Thus, as diode reverse bias is increased, the depletion layer increases, causing a decrease in capacitance and a decrease in series resistance. As the diode reverse bias is increased further, a point is reached where the electric field caused by the reverse bias reaches a critical level, and current through the diode increases rapidly; this is

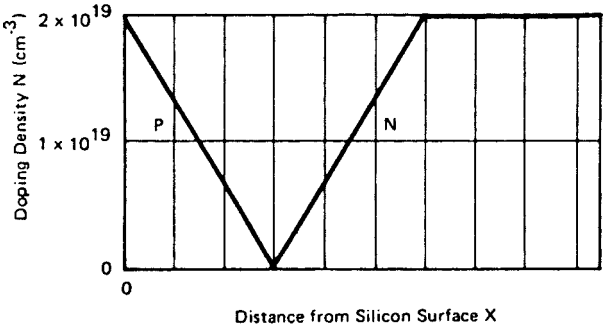


Figure 2-34 *N–X* linearly graded junction.

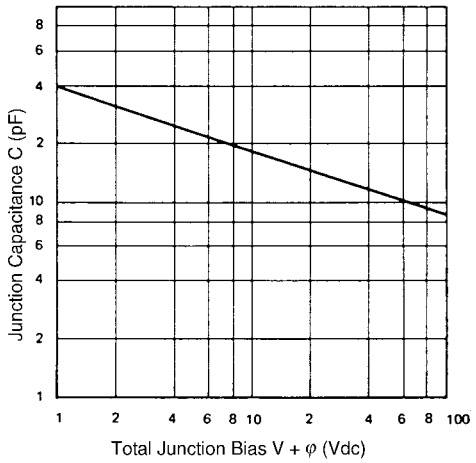


Figure 2-35 Capacitance versus total junction bias for linearly graded junction.

breakdown voltage of the diode. If, at the breakdown voltage, the epitaxial layer is not completely depleted, the diode will have excessive series resistance. Conversely, if the epitaxial layer is depleted before the breakdown voltage is reached, no further capacitance decrease occurs after the total depletion, and a condition called *punchthrough* occurs.

While, in the ideal case, voltage breakdown will occur just as the epitaxial layer is totally depleted, this seldom occurs in practice, and we generally have a condition of either punchthrough or excess series resistance.

Linearly Graded Junction. If, instead of the junction profile shown in Figure 2-32, we have a p^+ region and an n region whose doping levels increase linearly with distance from the pn junction as shown in Figure 2-34, with its corresponding C–V curve in Figure 2-35, then we have what is called a linearly graded junction diode. This diode follows Eq. (2-42) with the exception that the exponent n is equal to $\frac{1}{3}$. This means that, for a given voltage change, the linearly graded junction will have a smaller capacitance change than an abrupt-junction diode. Since, in most cases, the designer is looking for the maximum

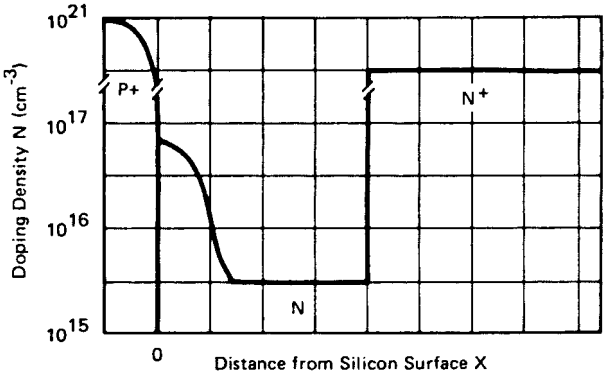


Figure 2-36 *N–X* hyperabrupt-junction diode.

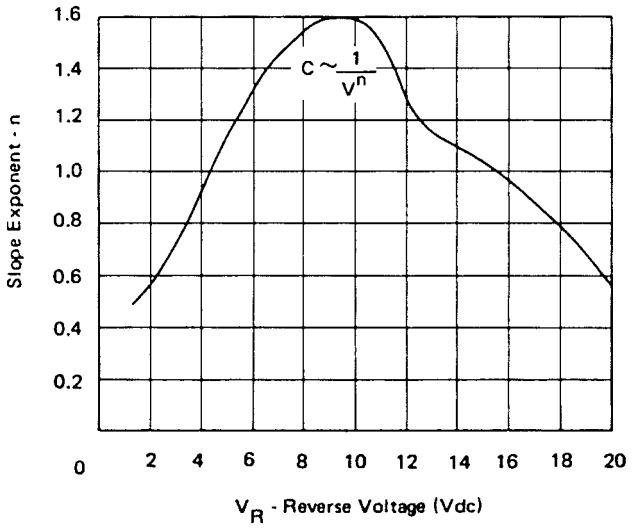


Figure 2-37 Typical *n* versus reverse voltage characteristic for hyperabrupt-junction diode.

capacitance change obtainable, the linearly graded junction is not used as a tuning diode. This structure found its greatest use several years ago as a "cube law" multiplier, but even this use has decreased as new structures have been developed.

Hyperabrupt Junction. The hyperabrupt-junction diode provides a greater capacitance change than the abrupt-junction diode for a given voltage change, as well as a linear frequency versus voltage characteristic over a limited voltage range. The structure of the hyperabrupt-junction diode is shown in Figure 2-36 and can be seen to be an abrupt-junction diode with an additional, increased doping level at the *pn* junction. This diode also follows Eq. (2-42) with the exception that n is not a function of voltage and is generally in the range of 0.5–2. A typical curve of n versus voltage is shown in Figure 2-37.

The C–V curve in a hyperabrupt-junction diode is shown in Figure 2-38 and is seen to start at a high value of capacitance per unit area at low bias (high epitaxial doping) and change to a lower value of capacitance per unit area (low epitaxial doping) at high bias. The details of the curve depend on details of the shape of the more highly doped region near the *pn* junction.

Unfortunately, with a hyperabrupt-junction diode, you must settle for a lower Q than an abrupt-junction diode with the same breakdown voltage and same capacitance at 4 V.

It should be noted that any diode that has an n value that exceeds 0.5 at any bias voltage is, by definition, a hyperabrupt-junction diode. Thus, the hyperabrupt-junction diode family can have an infinite number of different C–V curves. Since the abrupt-junction diode has a well-defined C–V curve, the capacitance value at one voltage is sufficient to define the C–V capacitance at any other voltage. This is not the case for the hyperabrupt-junction diode. In order to adequately define the C–V characteristics of a hyperabrupt-junction diode, two and sometimes three points on the curve must be specified.

Silicon Versus Gallium Arsenide. Everything mentioned so far applies to both silicon and gallium arsenide (GaAs) diodes. The main difference between silicon and GaAs from a

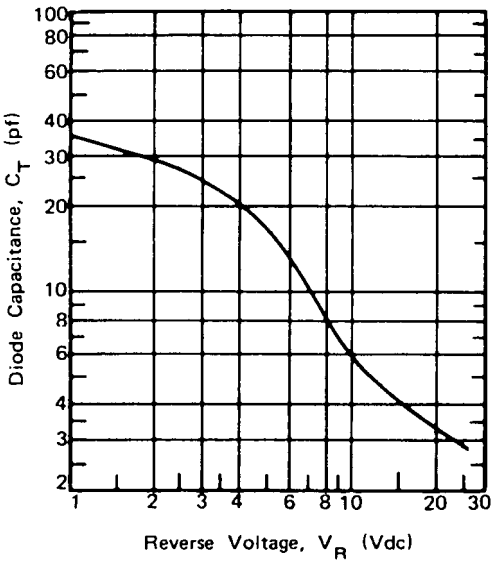


Figure 2-38 Capacitance versus junction bias for hyperabrupt-junction diode.

user's point of view is that higher Q can be obtained from GaAs devices. This is due to the lower resistivity of GaAs from a given doping level N. The resistivity of the epitaxial layer, or substrate, of a diode is given by

$$\rho = \frac{1}{Ne\mu} \tag{2-43}$$

where ρ = the resistivity
N = the doping level of the layer
e = the charge on an electron
μ = is the mobility of the charge carriers in the layer

Gallium arsenide has a mobility about four times that of silicon and, thus, a lower resistivity and higher Q for a given doping level N. Since diode capacitance is proportional to $\sqrt{N}$, independent of resistivity, a silicon diode and a GaAs diode of equal area and doping will have a capacitance difference proportional to the square root of the dielectric constant ratio. This gives the GaAs diode a 5% higher capacitance and is thus of little practical significance. The penalty paid for using GaAs is an unpassivated diode and a more expensive diode due to higher material and processing costs. If the higher Q of the GaAs device is not really needed, a substantial price saving will be obtained by using a silicon device.

Planar Versus Mesa Construction. The two basic construction techniques used to manufacture tuning diodes are planar and mesa; cross sections for both of these devices are shown in Figure 2-39. The planar process, which is the backbone of the integrated circuit industry, lends itself to large-volume production techniques and is the one used for the 1N series of tuning diodes. Mesa processing, on the other hand, requires more

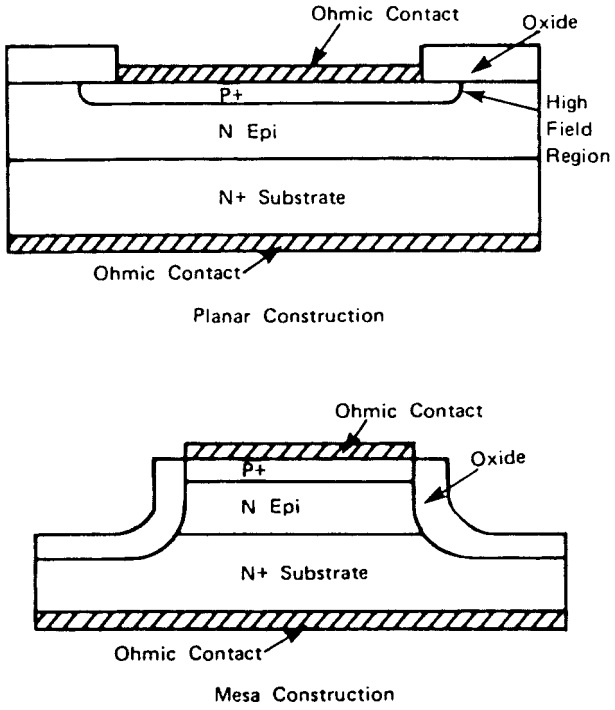


Figure 2-39 Cross sections of planar and mesa devices.

processing steps and is generally done on a wafer-by-wafer basis. This results in a more costly process and thus a more expensive diode. All microwave tuning diodes are of mesa design because of greatly higher Q. Due to the relatively small radius of curvature at the junction edge of a planar diode, the electric field in this area is greater than the electric field in the center (flat) portions of the junction. As a result, the breakdown voltage of the diode is determined by both the epitaxial resistivity and the radius of curvature of the junction edge. Thus, for a given breakdown voltage, a planar diode must use higher resistivity epitaxial material than a mesa diode, which has a completely flat junction. The end result is that the planar diode has a greater series resistance than a mesa diode for the same capacitance and breakdown voltage, and thus lower Q.

Capacitance

Capacitance Ratio. From the user's point of view, the capacitance ratio is simply the capacitance change available in the circuit. Thus, a user tuning from, say, –4 to –45 volts defines capacitance ratio as

$$R = \frac{C_T(-4)}{C_T(-45)} \tag{2-44}$$

where C_T includes C_J plus C_P plus C_F. The manufacturer, however, defines C_T as $C_J + C_P$.

To explore the significance of this difference, let's take two examples, one with a large C_J and another with a small C_J, in chip, package, and "typical" fringe situations. Both are 45-V tuning varactors.

Device	C_{J0}	C_{J45}	Ratio
A	0.6 pF	0.1 pF	6.0
B	15.0 pF	2.5 pF	6.0

Put both devices in a standard 023 package with C_P (strap and ceramic) of 0.18 pF:

Device	C_{T0}	C_{T45}	Ratio
C	0.78 pF	0.28 pF	2.75
D	15.18 pF	2.68 pF	5.67

Notice the drop in capacitance ratio, especially for the low-C_J diode. If we now add a typical 0.04 pF for external fringe capacitance, we get the following results:

Device	C_{T0}	C_{T45}	Ratio
C	0.82 pF	0.32 pF	2.56
D	15.22 pF	2.72 pF	5.6

The reduction in capacitance ratio, and thus circuit tuning capability, by the fringing fields is quite obvious and amounts to 7% in this example.

Because of the often stringent specifications on tuning ratio, it is mandatory that the manufacturer and customer clearly agree on the exact design of the holder used to measure the varactor in question.

Having described how to measure capacitance, it is relatively easy to describe the results. The section on diode physics described the various types of "laws," or C–V curves, and we will not repeat them here. Nonetheless, several important points must be covered.

The first is "available capacitance swing." The laws indicate a steadily decreasing capacitance with voltage, which indicates that the epitaxial region is widening and the electric field is increasing. (For an abrupt junction, since $C \propto 1/\sqrt{V}$, the depletion zone with W is increasing as $\sqrt{V}$, and the electric field V/W increases as $\sqrt{V}$.)

Two things can happen:

1. The junction width widens so that the entire intrinsic region is depleted. The capacitance bottoms out, resulting in voltage punchthrough.
2. The electric field exceeds the dielectric strength of silicon (or GaAs), and "solid-state discharge" or "avalanche" current is drawn.

The diode impedance drops, the varactor no longer "varacts," and circuit operation ceases. Moreover, if more than a few milliamperes of current is drawn, localized overheating may destroy the diode, resulting in breakdown voltage. All varactors are characterized for breakdown voltage—for example, 45 V minimum.

The theoretical tuning varactor is designed so that the punchthrough occurs at a voltage equal to the voltage breakdown of the diode. Logically, this means that in order to obtain

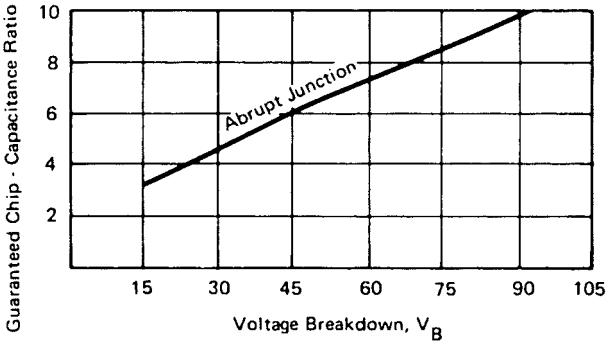


Figure 2-40 Capacitance ratio versus breakdown voltage.

greater tuning ratios, it is necessary to be able to increase the depletion-layer width without reaching punchthrough or breakdown. You must have a thicker epitaxial region to make this possible.

Figure 2-40 shows catalog ratio values, from zero bias to breakdown, as a function of breakdown voltage necessary. In the next section, on Q, we will discuss other elements in your choice of V_B.

To complete this section, we should mention that semiconductor processing control has been refined so well that capacitance tracking to within ±1% over the full range from zero to breakdown is now readily obtainable in production quantities.

Temperature Coefficient of Capacitance (T_{CC}). Unfortunately, since most datasheets give the value of T_{CC} at 4 V, it is sometimes assumed that this value applies at all bias voltages. This is not the case. Consider Eq. (2-45), a rewritten form of Eq. (2-42):

$$C(V) = \frac{C(O)}{(V + \varphi)^n} \tag{2-45}$$

Taking the derivative of this with respect to temperature T we have

$$\frac{dC(V)}{dT} = \frac{+nC(O)}{(v + \varphi)(v + \phi)^n} \frac{d\varphi}{dT} \tag{2-45a}$$

or, after substituting Eq. (2-45):

$$T_{CC} = \frac{1}{C(V)} \frac{dC(V)}{dT} = \frac{-n}{(V + \varphi)} \frac{d\varphi}{dT} \tag{2-46}$$

As a first approximation, we can say that $d\varphi/dT = -2.3$ mV/°C over the temperature range of interest.

From Eq. (2-46) we can draw the following conclusions:

1. The temperature coefficient is inversely proportional to the applied voltage.
2. The temperature coefficient is directly proportional to the diode slope, n.

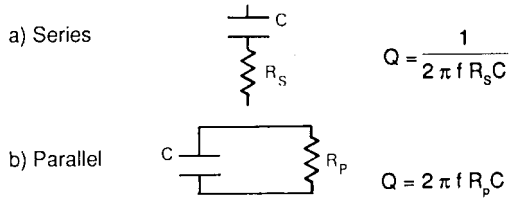
For an abrupt-junction diode that has a constant value of n (0.5), the temperature coefficient has a smooth curve in the form $K/(v + \varphi)$. However, in the case of hyperabrupt-junction diodes, n is a function of voltage, and the shape of the T_{CC} curve depends on the details of the $n(V)$ curve. A typical T_{CC} curve for an abrupt-junction diode is shown in Figure 2-41 and for an Alpha DKV6520 series hyperabrupt-junction diode in Figure 2-42. The inflection in the hyperabrupt T_{CC} is due to the fact that in this voltage range $n(V)$ is increasing faster than $1/V$, giving an increase in T_{CC}. It should also be noted, however, that over the range of the T_{CC} minimum the temperature coefficient is relatively constant, and operation in this area may be advantageous in some applications where a restricted tuning range can be used.

Q Factor or Diode Loss

Definitions. The classical definition of the Q of any device or circuit is

$$Q = \frac{2\pi \times \text{Energy stored}}{\text{Energy dissipated per cycle}} \tag{2-47}$$

For a capacitor, two formulations are possible:



Clearly, the two definitions must be equal at any frequency, which establishes

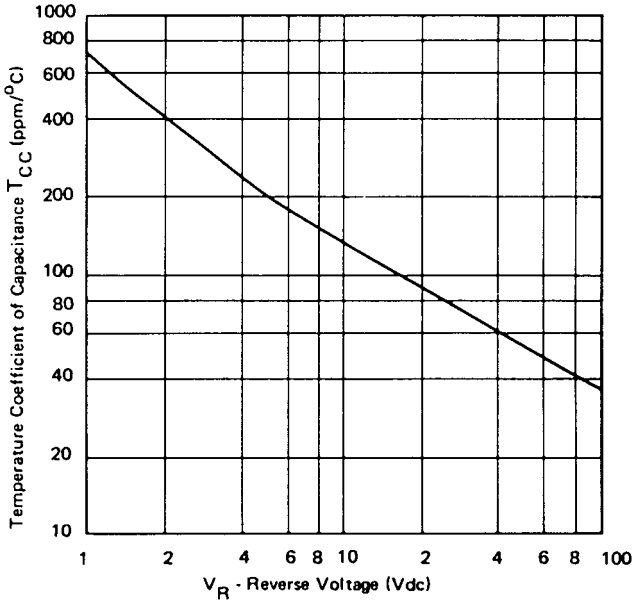


Figure 2-41 Temperature coefficient of capacitance versus tuning voltage for an abrupt-junction diode.

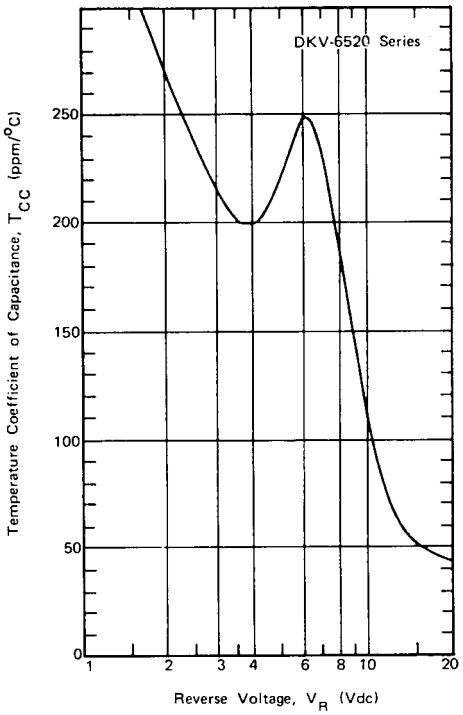


Figure 2-42 Temperature coefficient of capacitance versus tuning voltage ($T_A = 25$ °C) hyperabrupt-junction diode.

$$R_P = \frac{1}{(2\pi f)^2 C^2 R_S} \tag{2-48}$$

In the case of a high-Q tuning diode, the proper physical model is the series configuration, because the depleted region is almost perfectly pure capacitance, and the undepleted region, due to its relatively low resistivity, is almost a pure resistor in series with the capacitance. Furthermore, the contact resistances are also clearly in series. Q, then, for a tuning diode, is given by

$$Q_{(-v)} = \frac{1}{2\pi f_o R_{(-v)} C_{(-v)}} \tag{2-49}$$

where f_o is the operating frequency, $C_{(-v)}$ is the junction capacitance, and $R_{(-v)} = R(\text{epi}) + R_{C'}$, the sum of the resistance of the undepleted epitaxial layer and the fixed contact resistance.

Cutoff frequency, f_c, is defined as that frequency at which Q equals unity. Thus,

$$f_{c_{(-v)}} = \frac{1}{2\pi R_{(-v)} C_{(-v)}} \tag{2-50}$$

Historically, the tuning diode business has specified Q at 50 MHz, despite the fact that Q values of microwave diodes are so high that it is almost impossible to measure them at 50

Table 2-7 Parameters for 0.6-pF diode

Breakdown Voltage V_B	Resistivity ($\Omega \cdot$ cm)	Junction Diameter (mills)	Epitaxial Region (μm)	Depleted Epitaxial $V = -4$ V (μm)	Undepleted Epitaxial $V = -4$ V (μm)	R Undepleted (Ω)	R_{sp} (Ω)	Q_4
30	0.31	2.4	1.37	0.54	0.83	0.81	0.18	5200
45	0.52	2.8	2.25	0.73	1.52	1.86	0.15	2600
60	0.74	3.2	3.20	0.90	2.30	3.24	0.13	1500
90	1.25	3.7	5.27	1.21	4.06	7.17	0.10	700

MHz. Instead, as discussed below, Q is measured at microwave frequencies (e.g., 1–3 GHz) and related to 50 MHz by the relationship

$$Q_{(f1)} = Q_{(f2)} \frac{f_2}{f_1} \tag{2-51}$$

which derives quickly from the assumption that f_c is independent of the measuring frequency.

Since both junction capacitance and epitaxial layer resistance are functions of the applied bias, it is not possible to calculate Q as a function of bias from a measurement of capacitance alone. Catalog specifications typically show Q at –4 V, together with the capacitance at two or more voltages.

Relative to $Q_{(-4)}$, Q increases faster than the reduction in capacitance for bias greater than 4 V and, conversely, decreases faster for bias less than 4 V.

In the following section, we will discuss the diode design parameters that determine Q. Following this, we will describe some elementary Q measurement techniques.

Causes. In the discussion of device physics, the resistivity of the epitaxial region was discussed, together with its impact on punchthrough and breakdown. For example, Table 2-7 supplies typical resistivity and relative parameters of 0.6 pF ($C_{J(4)}$) diodes of different breakdowns. If we remember that if at any bias lower than breakdown the epitaxial region is not completely depleted, it follows that the undepleted portion presents a resistance in series with the pure capacitance of the depleted zone. The magnitude of this "undepleted" resistance is also shown in the table.

The entry R_{sp} ($R_{spreading}$) is the series resistance between the epitaxial region and the low-resistivity substrate. The calculations are for idealized cylindrical epitaxial regions of uniform resistivity, low-resistivity contact on the anode (top), and low-resistivity substrate on the cathode. This resistance is constant, independent of bias; also shown are epitaxial thickness region and width of the depletion zone at –4 V bias.

Note the substantial reduction in Q for high-voltage diodes caused by the increased epitaxial layer resistance; this is true for any value of capacitance or any type of junction. For greater voltage breakdown, the epitaxial region thickness must be increased, which requires an increase in epitaxial layer resistivity; the higher the resistivity of the undepleted zone, multiplied by the fact that it is much wider for high-voltage diodes, means the resistance increases substantially.

Consequently, a rule of thumb emerges: For maximum Q, never choose a tuning diode with a voltage breakdown in excess of what is needed for the necessary tuning range. If the

Table 2-8 ***Q*** **versus bias for 0.6-pF diode**

Breakdown Voltage V_B	C_{J0}	$Q(0)$	$C_J(-4)$	$Q(-4)$	$C_J(-10)$	Q (−10)	$C_J(-30)$	$Q(-30)$	$C_J(-45)$	$Q(-45)$	$C_J(-60)$	$Q(-60)$	$C_J(-90)$	Q(−90)
30	1.43	1700	0.6	5200	.4	10k	0.23	64k						
45	1.43	850	0.6	2600	.4	5k	0.23	20k	0.19	90k				
60	1.43	550	0.6	1500	.4	3k	0.23	9k	0.19	23k	0.17	170k		
90	1.43	270	0.6	700	.4	1.3k	0.23	3.5	0.19	6k	0.17	10k	0.14	220k

Table 2-9 Capacitance ratios (C_{J0}/C_JV_B)

Breakdown Voltage V_B	Optimum Ratio	Minimum Guaranteed Ratio	$Q_{(-4)}$ Typical
30	6.2	4.5	3000
45	7.5	6.0	2500
60	8.4	7.5	1400
90	10.2	8.7	650

required tuning range is an octave, requiring a 4:1 ratio, the selection of a 30-V diode will result in diode losses half those of a 60-V diode.

Table 2-8 lists capacitance and Q for each of these chips as a function of bias. Please remember that Q is calculated at 50 MHz.

Table 2-9 rewrites the data of Table 2-8 to show available capacitance ratios between zero bias and breakdown. The first column is the theoretical optimum, as tabulated. The second column is the typical catalog specification. The reduction in tuning ratio below theoretical optima is caused by nonideal junction fabrication. The junctions are never perfectly abrupt.

Although the tables and numbers given refer to abrupt-junction silicon devices, the principle applies without exception to all types of tuning diodes. For comparison, Table 2-10 lists available ratios and Q values for a number of different varactors. The high Q values for GaAs and the low values for hyperabrupts are apparent.

One last point: The Q values and series resistance refer to chips only. The effects of package parasitics will be discussed later, but it is important to consider circuit contact losses here. For low-capacitance diodes—for example, $C_{J4} = 0.6$ pF—the epitaxial region contributes a high value of resistance and dominates Q except at punchthrough. Diode contact losses are less significant.

Post-Tuning Drift. Post-tuning drift (PTD) is the change in oscillator frequency with time after the tuning voltage has stabilized. The minimization of PTD has assumed greater importance with the design of more sophisticated electronic countermeasure systems, where rapid, accurate frequency changes are required.

Table 2-10 Comparative Tuning Diodes

Type	C_{J0}	Breakdown Voltage V_B	$Q_{(-4)}$[a]	Ratio C_{J0}/C_JV_B[a]
Silicon abrupt	1.0	30	5,000	4.5
Silicon abrupt	2.5	30	4,600	4.5
Silicon abrupt	5.0	30	3,800	4.5
Gallium arsenide abrupt	1.0	25	10,000	3.6
Gallium arsenide abrupt	0.5	10	17,000	2.5
Silicon hyperabrupt	50.0	22	300	17.0
Silicon hyperabrupt	2.5	22	500	14

[a] Minimum guaranteed.

Post-tuning drift can be characterized as short-term and long-term. Short-term PTD occurs in the time range of tens of nanoseconds to a few seconds, while long-term PTD occurs in the time range of seconds to minutes, hours, or days.

Short-term PTD is mainly dependent on the thermal properties of the diode and is improved by high Q (low power loss) and flip chip construction. Long-term PTD depends on oxide stability and freedom of mobile charge in the oxide. It should be noted that actual oscillation frequency change may occur even with a perfect tuning diode because of variation with frequency in the power dissipated by the diode, changes in the diode heat-sink temperature, and frequency changes due to other circuit elements. Less than 0.01% short- and long-term PTD can be obtained.

Distortion Products. Inasmuch as nonlinear components generate harmonics and other distortion products, an understanding of this mechanism is of prime interest to the circuit designer. In some instances, the distortion products are the desired end result of the circuit design, as in frequency multipliers, where harmonics of the input signal frequency are the required output signal. For other applications, such as tuning-diode-tuned linear circuits, distortion products are extremely undesirable, and in some instances the end product specification may set a maximum limit on the distortion products allowed.

Cross-Modulation. Cross-modulation is the transfer of the modulation on one signal to another signal and is caused by third-order and higher odd-order nonlinearities in the behavior of the device. Rewriting Eq. (2-42) we have

$$C(V) = \frac{C_0}{(1 + V/\varphi)^n} \tag{2-52}$$

where C_0 = capacitance
V = applied voltage = $V_0 + v$
V_0 = dc applied voltage
v = ac applied voltage

Then, for a desired signal of

$$S_1 = v_1 \sin \omega_1 t \tag{2-53}$$

and a second, amplitude-modulated signal of

$$S_2 = v_2(1 + m \cos \omega_m t)\sin \omega_2 t \tag{2-54}$$

it can be shown that the cross-modulation, γ, defined by

$$\text{Output signal} \sim v_1 \sin \omega_1 t + \gamma \sin (\omega_1 \pm \omega_m)t \tag{2-55}$$

is found to be

$$\gamma = \frac{n(n+1)mv_2^2}{4(V_0+\varphi)^2} \tag{2-56}$$

From this equation it can be seen that cross-modulation is:

- Proportional to the square of the interfering signal
- Directly proportional to the interfering signal's modulation index, m
- Independent of the strength of the desired signal
- Independent of the frequencies of the desired and interfering signals (assuming that the nonlinearity to which both signals are subjected is sufficiently frequency-indiscriminate so this is the case)
- Present for all values of n; that is, no value of n gives zero cross-modulation

Solving Eq. (2-56) for the signal level v_2 required to produce cross-modulation of value γ, we have

$$v_2 = \frac{2(V_0+\phi)\gamma}{n(n+1)^m} \tag{2-57}$$

The interfering signal levels required to produce 1% cross-modulation from a 30%-modulated interfering signal applied to an abrupt-junction diode and a hyperabrupt-junction diode are shown in Figure 2-43.

From this figure, it can be seen that the hyperabrupt-junction diode is more susceptible to cross-modulation than the abrupt-junction diode in the region of maximum slope of the

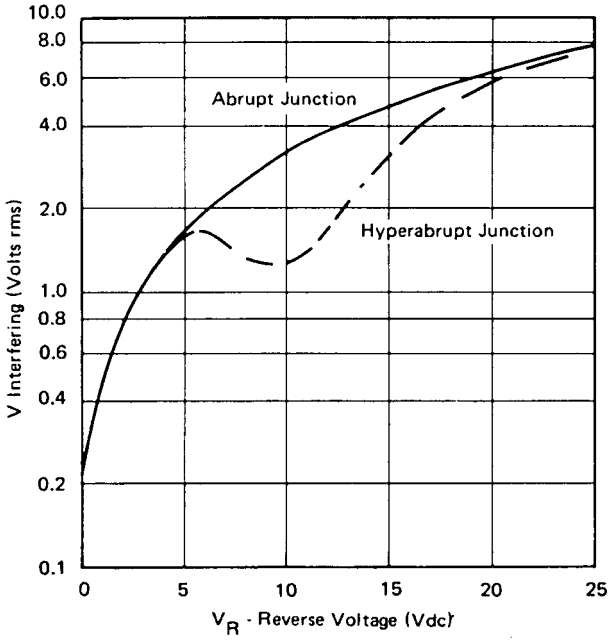


Figure 2-43 Interfering signal (30% amplitude modulated) level versus bias for 1% cross-modulation for abrupt-junction and hyperabrupt-junction diodes.

hyperabrupt-junction diode. For many applications, however, distortion products will be generated by other devices, such as transistors, at signal levels considerably below those given in Figure 2-43.

Intermodulation. Intermodulation is the production of undesired frequencies in the form

$$\sin(2\omega_1 t - \omega_2 t) \quad \text{and} \quad \sin(\omega_1 t - 2\omega_2 t) \tag{2-58}$$

from an input signal in the form of

$$v(\cos \omega_1 t + \cos \omega_2 t) \tag{2-59}$$

From an analysis similar to that done for cross-modulation, it can be shown that

$$\text{Intermodulation} = \frac{n(n+1)v^2}{8(V_0 + \varphi)^2} \tag{2-60}$$

or

$$\text{Cross-modulation} = (2m) \times \text{Intermodulation} \tag{2-61}$$

Harmonic Distortion. Harmonic distortion products are integral multiples of the signal frequencies and decrease in amplitude as the harmonic number decreases. Due to passband considerations and amplitude decrease with harmonic number, the second harmonic is the one of prime concern. Again, it can be shown that the second harmonic, v_2, of a signal of amplitude v_1 is

$$v_2 = \frac{n}{3(V_0 + \phi)} v_1^2 \tag{2-62}$$

Figure 2-44 shows the signal level required to produce 10% second-harmonic distortion in an abrupt-junction and a hyperabrupt-junction diode. Again, as in the case of cross-modulation, the hyperabrupt-junction diode is slightly worse than the abrupt-junction diode in the region of maximum slope of the hyperabrupt-junction diode.

Reduction of Distortion Products. In some cases, the signal levels applied to the diode generate distortion products larger than desirable for the circuit application. In this case, significant reduction in the distortion products can be achieved by using two diodes in a back-to-back configuration, as shown in Figure 2-45. Analysis shows that the fundamental signal components through the diodes are in phase and add, while some distortion products are out of phase and cancel, thus improving distortion performance.

Since the gradient of the electrical field produced in the depletion layer by a reverse bias applied to the device is proportional to the space charge density, the following equations can be written for the junction width W as a function of the reverse bias V_R:

For an Abrupt *pn* Junction (Alloyed Diodes)

$$W = {}_2\sqrt{2\frac{\varepsilon_r \varepsilon_0}{q}\left(\frac{1}{C_p} + \frac{1}{C_n}\right)(V_R + V_D)} \tag{2-63}$$

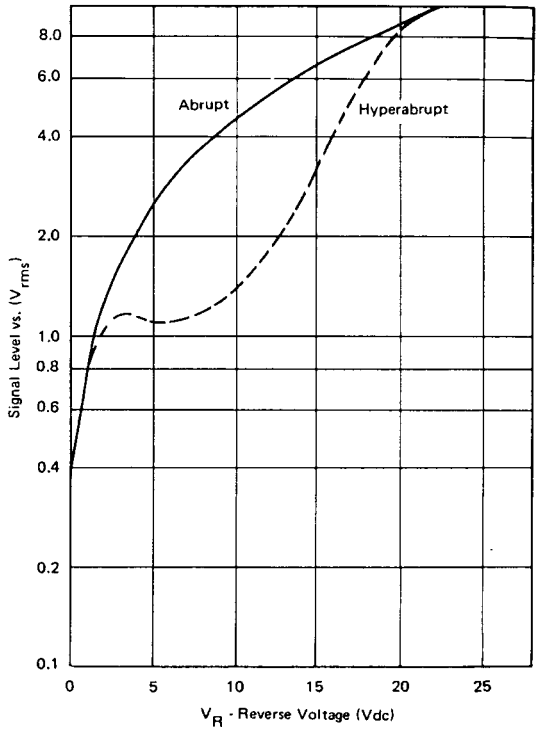


Figure 2-44 Signal level versus reverse voltage for 10% harmonic distortion.

For Linear *pn* Junctions (Single-Diffused Diodes, such as BA110–BA112)

$$W = \sqrt[3]{12\frac{\varepsilon_r \varepsilon_0}{aq}(V_R + V_D)} \tag{2-64}$$

where a is the impurity gradient within the depletion layer, ε_0 is the absolute dielectric constant ($8.85 \times 10^{-14}\ A_s/V_{cm}$, and $\varepsilon_r \approx 12$, the relative dielectric constant of silicon.

The junction capacitance, which is inversely proportional to the junction width, therefore varies in alloyed diodes with the square root, and in single-diffused diodes with the cube root of the externally applied reverse bias, and can be calculated from the general equation

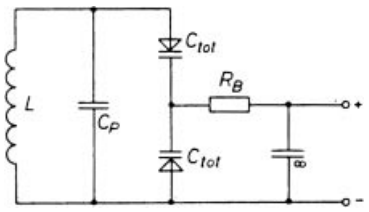


Figure 2-45 Back-to-back diodes. C_P is a fixed parallel capacitance, R_B is a bias decoupling resistor, and the capacitor marked ∞ is a low-impedance bypass.

$$C = \frac{\varepsilon_r \varepsilon_0 S}{W} \tag{2-65}$$

where S is the surface area of the *pn* junction. By way of approximation, we can also use the equation

$$C = \frac{K}{(V_R + V_D)^n} \tag{2-66}$$

where all constants and all parameters determined by the manufacturing process are contained in K. The exponent n is a measure of the slope of the capacitance/voltage characteristics and is 0.5 for alloyed diodes, 0.33 for single-diffused diodes, and (on average) 0.75 for tuner diodes with a hyperabrupt *pn* junction. Figure 2-46 shows the capacitance/voltage characteristics of an alloyed, a diffused, and a tuner diode.

Recently, an equation has been indicated which, although purely formal, describes the practical characteristics better than Eq. (2-66):

$$C = C_0 \left(\frac{A}{A + V_R} \right)^m \tag{2-67}$$

where C_0 is the capacitance at $V_R = 0$, and A is a constant whose dimension is a voltage. The exponent m is much less dependent on voltage than the exponent n in Eq. (2-66).

Equations (2-63) through (2-67) express the pure junction capacitance of the capacitance diode, but to this must still be added a constant capacitance, determined by structure parameters, in order to obtain the diode capacitance C_{tot}, which interests the user. With high inverse voltages—that is, low junction capacitance—a difference will therefore arise between the theoretical capacitance/voltage characteristic according to Eq. (2-66) and the practical characteristic, as shown in Figure 2-47.

The operating range of a capacitance diode or its useful capacitance ratio

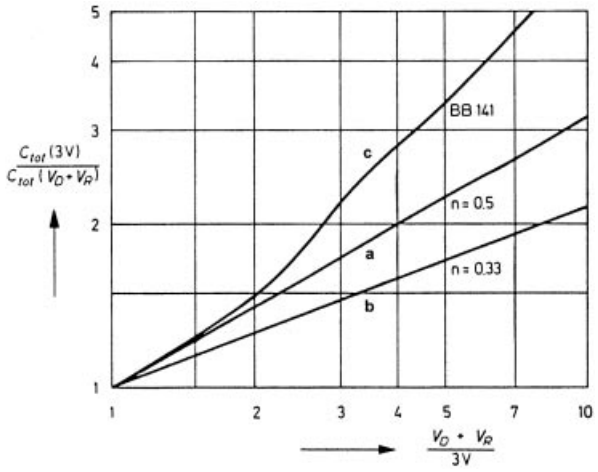


Figure 2-46 Capacitance/voltage characteristic for (a) an alloyed capacitance diode, n = 0.5; (b) a diffused capacitance diode, n = 0.33; and (c) a wide-range tuner diode (BB141).

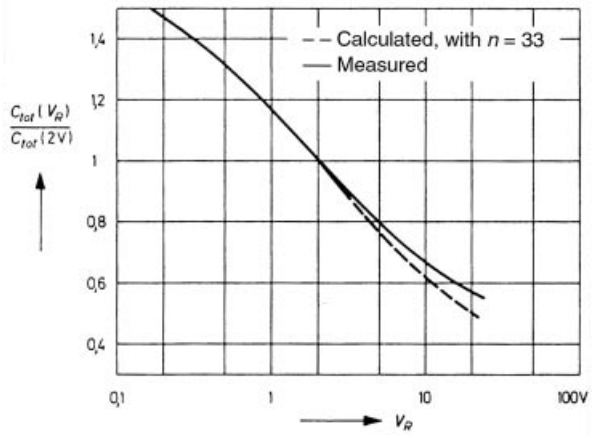


Figure 2-47 Capacitance/voltage characteristic of the BA110 diode.

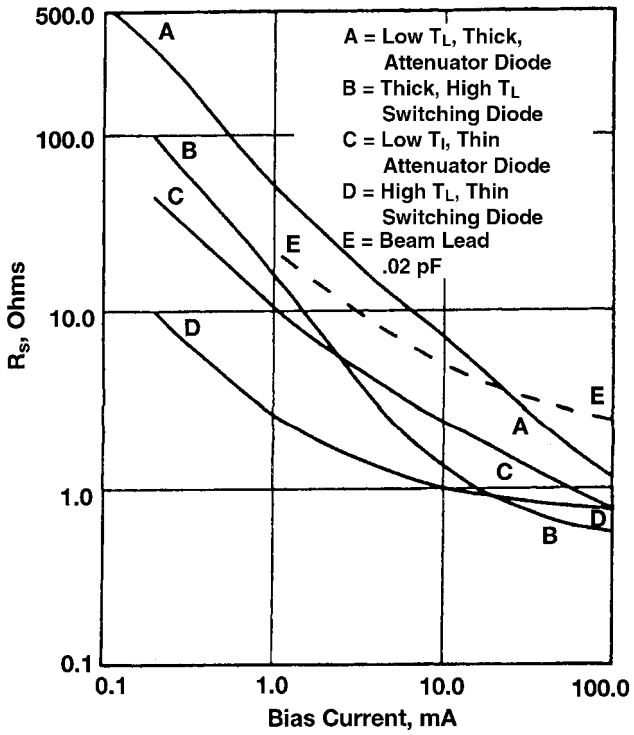


Figure 2-48 Basic current/voltage and capacitance/voltage characteristics.

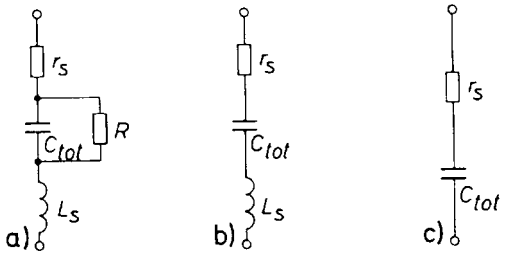


Figure 2-49 Equivalent circuits for capacitance diodes: (a) complete circuit, (b) simplified circuit, and (c) further simplification for low frequencies.

$$\frac{C_{\max}}{C_{\min}} = \frac{C_{\text{tot}}(V_{R\min})}{C_{\text{tot}}(V_{R\max})} \tag{2-68}$$

is limited by the fact that the diode must not be driven by the alternating voltage superimposed on the tuning voltage either into the forward mode or the breakdown mode. Otherwise, rectification, which would shift the diode's bias and considerably affect its figure of merit, would take place. Figure 2-48 plots the capacitance/voltage characteristics of a capacitance diode to clarify the relationship. The useful operating range lies between the voltages

$$V_{\min} > \hat{V} - V_F \quad \text{and} \quad V_{\max} < V_{(\text{BR})R} - \hat{V} \tag{2-69}$$

As has already been indicated, the exponent n for large-capacitance-ratio or tuning diodes used today for TV tuners is not constant but voltage-dependent and subject to manufacturing tolerances. This means that the capacitance/voltage characteristic of these diodes is likewise subject to manufacturing tolerances. Since, in a TV tuner, it is necessary for two or three circuits to be tuned uniformly, tuner diodes must be selected empirically for identical characteristics and supplied in equipment lots.

Electrical Properties of Tuning Diodes. In this section, the electrical properties of capacitance diodes are described, with reference to data published on the tuner diodes BB141 and BB142.

Equivalent Circuit. Since a capacitance diode is not an ideal capacitor, it is useful to introduce an equivalent circuit that can serve as a basis for discussing the diode's electrical properties. Figure 2-49 shows various types of equivalent circuits.

The complete equivalent circuit (Figure 2-49a), which conforms closely to physical conditions, comprises, in addition to the diode capacitance C_{tot}, the series resistance r_S and the series inductance L_S, and also the reverse resistance $R = dV_R/dI_R$. At higher frequencies, this resistance can usually be disregarded, so that the equivalent circuit is simplified in terms of Figure 2-49b—the configuration usually employed. In many cases, the series inductance can also be disregarded, in which case we obtain Figure 2-49c.

Capacitance. Figure 2-50 shows the capacitance/voltage characteristic borrowed from the datasheets of diodes BB141 and BB142. As this characteristic, like most characteristics published in datasheets, represents merely a typical curve, it would not convey sufficient

information for the user to dimension a tuner. Therefore, the datasheets contain additional data on capacitance and useful capacitance ratio:

Capacitance		BB141A (pF)	BB141B (pF)	BB142 (pF)
At $V_R = 1$ V	C_{tot}	16	19	17
At $V_R = 3$ V	C_{tot}	11	13	12
At $V_R = 25$ V	C_{tot}	2–2.35	2.25–2.65	2–3

$$\text{Useful capacitance ratio} = \frac{C_{tot}(3\text{ V})}{C_{tot}(25\text{ V})} = 4 \text{ to } 6$$

On the basis of the spread of the diode capacitance C_{tot} guaranteed for $V_R = 25$ V, it is possible to calculate the tuner with the aid of the spread guaranteed for the capacitance ratio.

Three further capacitance graphs are shown in Figures 2-51 through 2-53. Figure 2-51 shows a (typical) curve representing the normalized slope as a function of the reverse voltage, Figure 2-52 the normalized capacitance as a function of the junction temperature, and Figure 2-53 the temperature coefficient of the capacitance as a function of the reverse voltage.

The variation of diode capacitance with ambient temperature, as shown in Figures 2-52 and 2-53, is really only a function of the temperature dependence of the diffusion voltage [see Eq. (2-66)]. The diffusion voltage drops with rising temperatures by about 2 mV/°C , which means that the diode capacitance rises with temperature. The influence of the diffusion voltage, and thus the temperature coefficient of the capacitance, decreases as the reverse voltage rises. It is therefore advisable to run a capacitance diode with as high a reverse voltage as the required capacitance ratio permits. Compensation for the dependence of capacitance on temperature will be discussed later.

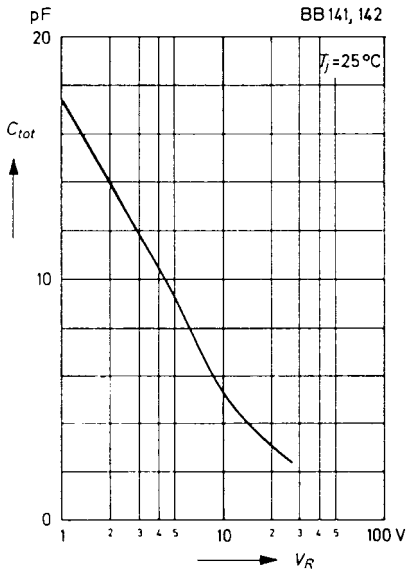


Figure 2-50 Capacitance/voltage characteristic of diodes BB141 and BB142.

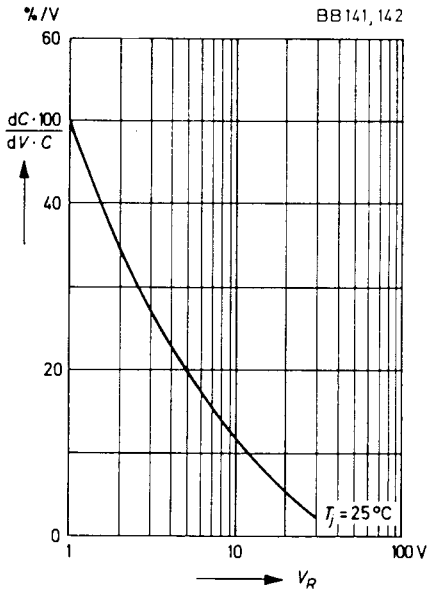


Figure 2-51 Slope (normalized) as a function of the reverse voltage.

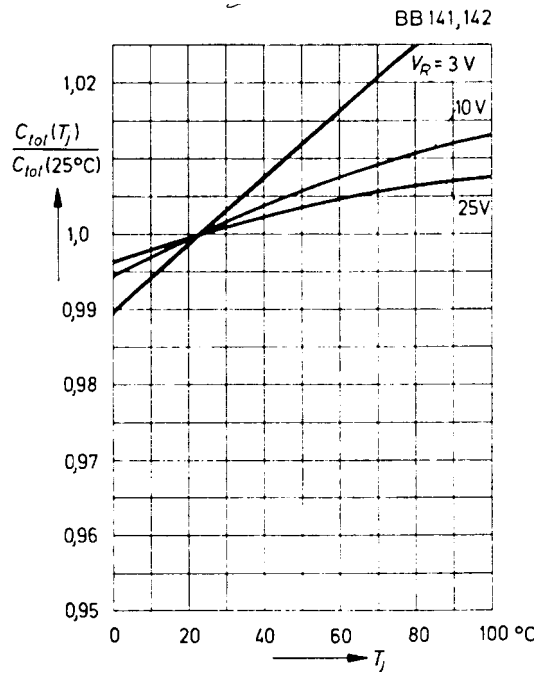


Figure 2-52 Capacitance (normalized) as a function of the junction temperature.

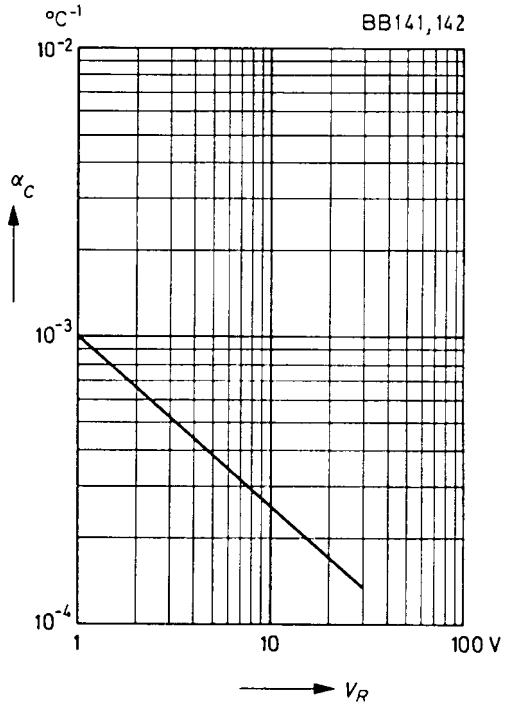


Figure 2-53 Temperature coefficient of capacitance as a function of the reverse voltage.

Series Resistance, Figure of Merit (Q). Since capacitance diodes are intended for employment in resonant circuits, indications are required as to the circuit attenuation that they produce. Usually, the series resistance or the figure of merit that can be calculated therefrom is indicated. With reference to the equivalent circuit shown in Figure 2-49a, the figure of merit (or Q factor) is calculated:

$$Q = \frac{1}{\omega C_{\text{tot}} r_S + 1/\omega C_{\text{tot}} R} \tag{2-70}$$

Figure 2-54 shows the theoretical (normalized) Q as a function of frequency. With the frequency range of 10–1000 MHz, which is essential for tuner diodes, the parallel resistance R caused by the reverse current (see Figure 2-49b) can be disregarded, and Eq. (2-70) can be simplified to read as follows:

$$Q = \frac{1}{\omega C_{\text{tot}} r_S} \tag{2-71}$$

In Figure 2-55, Q is plotted as a function of frequency for the tuner diodes BB141 and BB142. The datasheet indicates, in addition to this curve, the value of the series resistance at $f = 470$ MHz and $C_{\text{tot}} = 9$ pF:

	BB141A, BB141B	BB142
Series resistance (r_S)	0.6 (< 0.8)	0.9 (< 1.2)

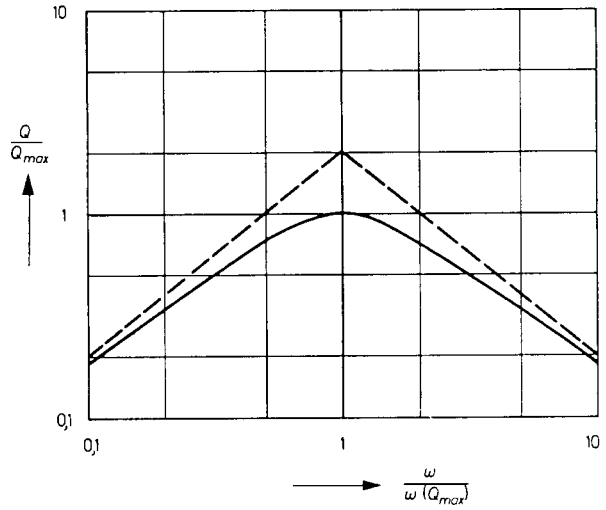


Figure 2-54 *Q* (normalized) as a function of frequency.

Since the diode capacitance and the series resistance decrease as the reverse voltage rises, Q is likewise dependent on the reverse voltage. It is smallest at a low reverse voltage, that is, at a low frequency. The series and parallel capacitances inevitable in a diode-tuned resonant circuit influence the effective Q of the tuning capacitance, as will be described later.

Series Inductance, Series Resonant Frequency, Cutoff Frequency for Q = 1. These three parameters define the behavior of the diode at high frequencies when the influence of

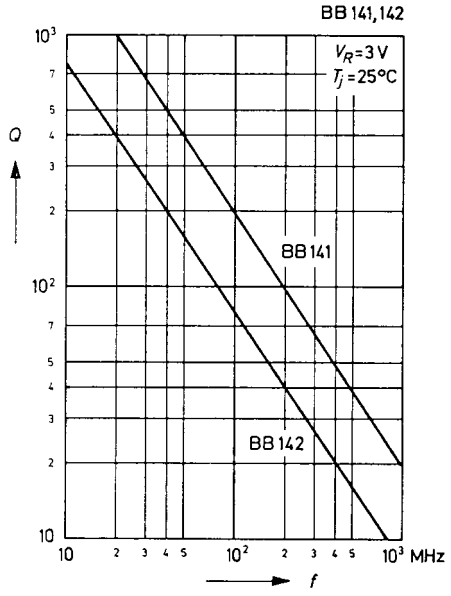


Figure 2-55 *Q* as a function of frequency.

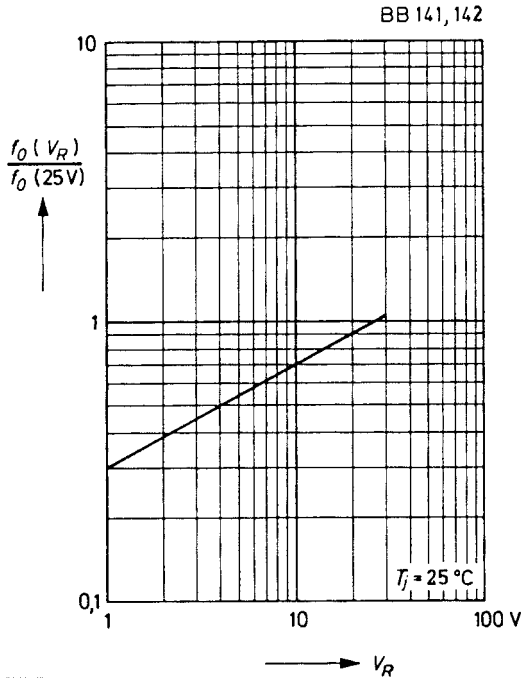


Figure 2-56 Series-resonant frequency (normalized) as a function of the reverse voltage.

the series inductance can no longer be neglected. The parameter on which the calculation is based is the series inductance L_S, which is determined by the connection leads. As indicated in Figure 2-49a, it is in series with the capacitance and, given a suitable frequency, forms a series-resonant circuit therewith. In that case

$$\omega L_S = \frac{1}{\omega C_{\text{tot}}} \tag{2-72}$$

and the series-resonant frequency is

$$f_0 = \frac{1}{2\pi\sqrt{C_{\text{tot}} L_S}} \tag{2-73}$$

Above this frequency, the impedance of the diode is inductive, but it depends on the reverse voltage. Figure 2-56 shows how the (normalized) series-resonant frequency depends on the reverse voltage for the tuner diodes BB141 and BB142.

The cutoff frequency for $Q = 1$ is defined as the frequency at which Q assumes the value of unity. In a rough calculation, the reactive impedance of the diode capacitance can be neglected and the frequency f_{Q1} obtained for the condition in which the inductive reactance equals the series resistance:

$$\omega L_S = r_S \tag{2-74}$$

$$f_{Q1} = \frac{r_S}{2\pi L_S} \tag{2-75}$$

For the tuner diodes BB141 and BB142, Figure 2-57 shows how the (normalized) cutoff frequency depends on the reverse voltage. The datasheets contain the following numerical information:

		BB141	BB142	
Series inductance measured 1.5 mm from the case	L_S	2.5	2.5	nH
Series-resonant frequency at $V_R = 25$ V	f_0	2	1.8	GHz
Cutoff frequency for $Q = 1$ at $V_R = 3$ V	f_{Q1}	24	16	GHz

Depending on the intended application, either the series-resonant frequency or the cutoff frequency for $Q = 1$ determines the maximum useful frequency range.

Leakage Current, Breakdown Voltage. These two parameters mark the dc performance of the diode when a reverse bias is applied to it. In the databook for the tuner diodes BB141 and BB142, the breakdown voltage is quoted as follows:

$$\text{Reverse breakdown voltage at } I_R = 100\ \mu\text{A}: \quad V_{(BR)R} > 30\ \text{V}$$

This means that the maximum reverse voltage that may be applied to the diode as a tuning voltage is 30 V.

The leakage current for the above-mentioned tuner diodes is guaranteed as follows:

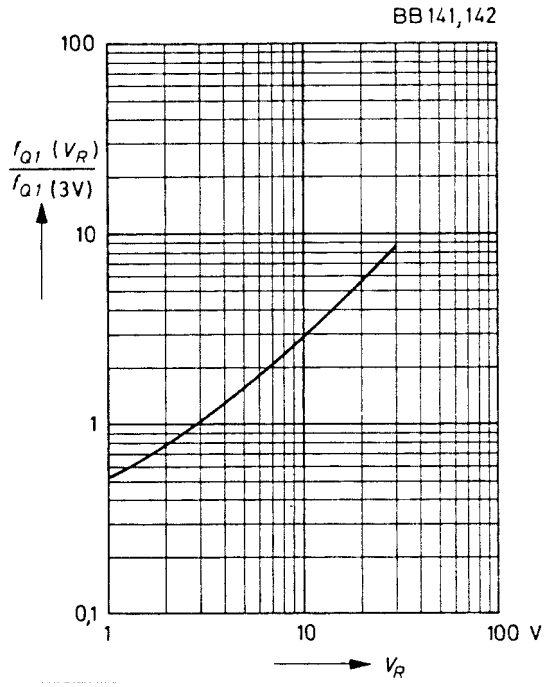


Figure 2-57 Cutoff frequency for $Q = 1$ (normalized) as a function of the reverse voltage.

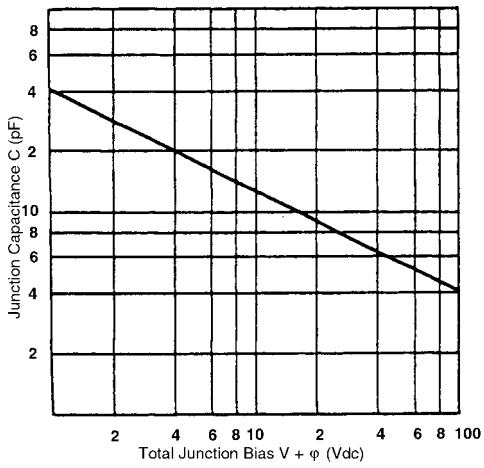


Figure 2-58 Leakage current as a function of reverse voltage.

$$\text{Leakage current at } V_R = 28 \text{ V:} \quad I_R < 50 \text{ nA}$$

Moreover, Figure 2-58 shows a graph that illustrates the leakage current as a function of the reverse voltage. Since this current is temperature dependent (as with every silicon diode)—it doubles with each temperature rise of about 10 °C—care should be taken when dimensioning the tuner circuit that the leakage current does not cause inadmissible voltage variations at increased ambient temperatures.

Matching of Tuner Diodes, Uniform Parameters. As has already been mentioned, the capacitance versus voltage characteristics of modern tuner diodes are subject to a certain scatter, so that it becomes necessary to test these diodes to obtain equipment lots empirically.

Diode-Tuned Resonant Circuits

Tuner Diode in the Parallel-Resonant Circuit. Figures 2-59, 2-60, and 2-61 illustrate three basic circuits for the tuning of parallel-resonant circuits by means of capacitance diodes.

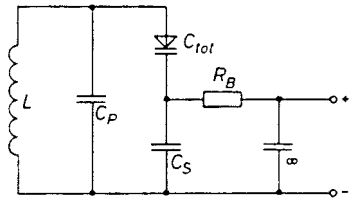


Figure 2-59 Parallel-resonant circuit with tuner diode, and bias resistor parallel to the series capacitor.

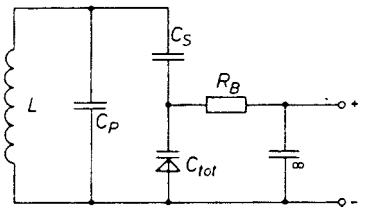


Figure 2-60 Parallel-resonant circuit with tuner diode, and bias resistor parallel to the diode.

In the circuit diagram of Figure 2-59, the tuning voltage is applied to the tuner diode via the input coil and the bias resistor R_B. Series-connected to the tuner diode is the series capacitor C_S, which completes the circuit for ac but isolates the cathode of the tuner diode from the coil and thus from the negative terminal of the tuning voltage. Moreover, a fixed parallel capacitance C_P is provided. The decoupling capacitor preceding the bias resistor is large enough for its value to be disregarded in the following discussion. Since for high-frequency purposes the biasing resistor is connected in parallel with the series capacitor, it is transformed into the circuit as an additional equivalent shunt resistance R_C. We have the equation

$$R_C = R_B\left(1 + \frac{C_S}{C_{\text{tot}}}\right)^2 \tag{2-76}$$

If in this equation the diode capacitance is substituted by the resonant circuit frequency ω, we obtain

$$R_C = R_B\left(\frac{\omega^2 LC_S}{1 - \omega^2 LC_P}\right)^2 \tag{2-77}$$

The resistive loss R_C caused by the bias resistor R_B is seen to be highly frequency dependent, and this may result in the bandwidth of the tuned circuit being independent of frequency if the capacitance of the series capacitor C_S is not chosen sufficiently high.

Figure 2-60 shows that the tuning voltage can also be applied directly and in parallel to the tuner diode. For the parallel loss resistance transformed into the circuit, we have the expression

$$R_C = R_B\left(\frac{1 + C_{\text{tot}}}{C_S}\right)^2 \tag{2-78}$$

and

$$R_C = R_B\left(\frac{\omega^2 LC_S}{\omega^2 L(C_S + C_P) - 1}\right)^2 \tag{2-79}$$

The influence of the bias resistor R_B in this case is larger than in the circuit of Figure 2-60, provided that

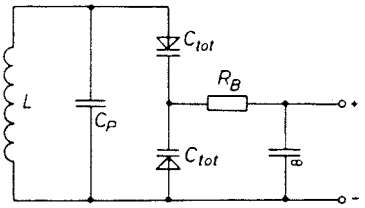


Figure 2-61 Parallel-resonant circuit with two tuner diodes.

$$C_S^2 > C_S(C_{\text{tot}} + C_P) + C_{\text{tot}}C_P \tag{2-80}$$

This is usually the case because the largest possible capacitance will be preferred for the series capacitor C_S, and the smallest for the shunt capacitance C_P. The circuit of Figure 2-59 is therefore normally preferred to that of Figure 2-60. An exception would be the case in which the resonant circuit is meant to be additionally damped by means of the bias resistor at higher frequencies.

In the circuit of Figure 2-61, the resonant circuit is tuned by two tuner diodes that are connected in parallel via the coil for tuning purpose, but series-connected in opposition for high-frequency signals. This arrangement has the advantage that the capacitance shift caused by the ac modulation (see section entitled "Modifying the Diode Capacitance by the Applied ac Voltage") takes effect in opposite directions in these diodes and therefore cancels itself. The bias resistor R_B, which applies the tuning voltage to the tuner diodes, is transformed into the circuit at a constant ratio throughout the whole tuning range. Given two identical, loss-free tuner diodes, we obtain the expression

$$R_C = 4R_B \tag{2-81}$$

Capacitances Connected in Parallel or Series with the Tuner Diode. Figures 2-59 and 2-60 show that a capacitor is usually in series with the tuner diode, in order to close the circuit for alternating current and, at the same time, to isolate one terminal of the tuner diode from the rest of the circuit with respect to direct current, so as to enable the tuning voltage to be applied to the diode. As far as possible, the value of the series capacitor C_S will be chosen such that the effective capacitance variation is not restricted. However, in some cases, as, for example, in the oscillator circuit of receivers whose intermediate frequency is of the order of magnitude of the reception frequency, this is not possible and the influence of the series capacitance will then have to be taken into account. By connecting the capacitor C_S, assumed to be lossless, in series with the diode capacitance C_{tot}, the tuning capacitance is reduced to the value

$$C^* = C_{\text{tot}} \frac{1}{1 + C_{\text{tot}}/C_S} \tag{2-82}$$

The Q of the effective tuning capacitance, taking into account the Q of the tuner diode, increases to

$$Q^* = Q\left(1 + \frac{C_{\text{tot}}}{C_S}\right) \tag{2-83}$$

The useful capacitance ratio is reduced to the value

$$\frac{C^*_{\max}}{C^*_{\min}} = \frac{C_{\max}}{C_{\min}} \frac{1 + C_{\min}/C_S}{1 + C_{\max}/C_S} \tag{2-84}$$

where $C_{\max}$ and $C_{\min}$ are the maximum and minimum capacitances of the tuner diode.

On the other hand, the advantage is gained that, due to capacitive potential division, the amplitude of the alternating voltage applied to the tuning diode is reduced to

$$\hat{v}^* = \hat{v} \frac{1}{1 + C_{\text{tot}}/C} \tag{2-85}$$

so that the lower value of the tuning voltage can be smaller, and this results in a higher maximum capacitance $C_{\max}$ of the tuner diode and a higher useful capacitance ratio. The influence exerted by the series capacitor, then, can actually be kept lower than Eq. (2-83) would suggest.

The parallel capacitance C_P that appears in Figures 2-59 to 2-61 is always present, since wiring capacitances are inevitable and every coil has its self-capacitance. By treating the capacitance C_P, assumed to be lossless, as a shunt capacitance, the total tuning capacitance rises value and, if C_S is assumed to be large enough to be disregarded, we obtain

$$C^* = C_{\text{tot}} \left(1 + \frac{C_P}{C_{\text{tot}}} \right) \tag{2-86}$$

The Q of the effective tuning capacitance, as derived from the Q of the tuner diode, is

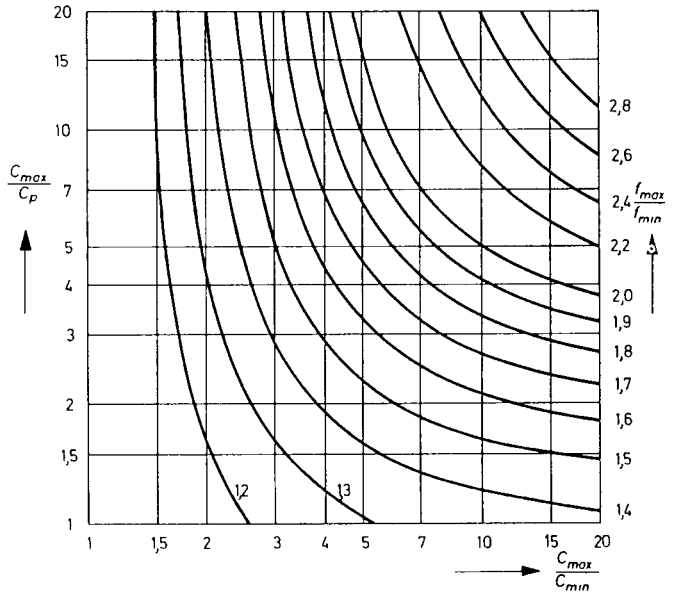


Figure 2-62 Diagram for determining the capacitance ratio and minimum capacitance.

$$Q^* = Q\left(1 + \frac{C_P}{C_{\text{tot}}}\right) \tag{2-87}$$

or, in other words, it rises with the magnitude of the parallel capacitance. The useful capacitance ratio is reduced:

$$\frac{C^*_{\text{max}}}{C^*_{\text{min}}} = \frac{C_{\text{max}}}{C_{\text{min}}} \frac{1 + C_P/C_{\text{max}}}{1 + C_P/C_{\text{max}}} \tag{2-88}$$

In view of the fact that even a comparatively small shunt capacitance reduces the capacitance ratio considerably, it is necessary to ensure low wiring and coil capacitances in the circuit-design stage.

TUNING RANGE. The frequency range over which a parallel-resonant circuit according to Figure 2-59 can be tuned by means of the tuner diode depends on the useful capacitance ratio of the diode and on the parallel and series capacitances present in the circuit.

The ratio can be found from

$$\frac{f_{\text{max}}}{f_{\text{min}}} = \sqrt{1 + \frac{\dfrac{C_{\text{max}}}{C_P\,(1 + C_{\text{max}}/C_S)}}{1 + \dfrac{C_{\text{max}}}{C_P\,(C_{\text{max}}/C_{\text{min}} + C_{\text{max}}/C_S)}}} \tag{2-89}$$

In many cases, the series capacitor can be chosen large enough for its effect to be negligible. In that case, Eq. (2-89) is simplified as follows:

$$\frac{f_{\text{max}}}{f_{\text{min}}} = \sqrt{\frac{1 + C_{\text{max}}/C_P}{1 + C_{\text{min}}/C_P}} \tag{2-89a}$$

From this equation, the diagram shown in Figure 2-62 is computed. With the aid of this diagram, the tuner diode parameters required for tuning a resonant circuit over a stipulated frequency range—that is, the maximum capacitance and the capacitance ratio—can be determined.

Whenever the series capacitance C_S cannot be disregarded, the effective capacitance ratio is reduced according to Eq. (2-84).

TRACKING. Some applications require the maintenance of a fixed frequency relationship between two or more tuned circuits as their tuning is simultaneously adjusted. Referred to as *tracking*, this technique requires narrow tolerances of capacitance versus tuning voltage. Minimizing tracking error requires special care if the tracking circuits must cover the same frequency span beginning at different start and end frequencies, as is necessary when simultaneously tuning oscillator and mixer/RF circuitry in a superheterodyne receiver. Then, tracking error must be minimized by means of series and shunt capacitances in accordance with methods known from variable capacitors. The frequency deviations that must be anticipated are summarized in the equation

$$\frac{df}{f} = -\frac{1}{2}\frac{dC_0}{C_0} - \frac{1}{2}\frac{d(L-L_0)}{L} - \frac{1}{2}\frac{dL_0}{L_0} + \frac{n}{2}\frac{dV_R}{V_R + V_D} \tag{2-90}$$

The spread of parameters dC_0/C_0 and dL_0/L_0 can only be compensated by varying the circuit inductances $d(L–L_0)/L$ or by varying the bias $dV_R/(V_R + V_D)$.

Modulating the Diode Capacitance by the Applied ac Voltage. In normal operation, the sum of the tuning voltage and the alternating signal voltage of the resonant circuits is applied to the tuner diode. The bias, and thus the capacitance, of the tuner diode therefore varies at the rhythm of the alternating voltage. Due the nonlinear character of the capacitance versus voltage curve, voltage distortions and capacitance shifts are inevitable, and these must be kept within adequate limits. This is done by maintaining the ac applied to the diode(s) at sufficiently low ac amplitude and by choosing an adequate minimum value for the tuning voltage. In the resonant circuit, a tuner diode is modulated predominantly by a current free from harmonics, according to the equation

$$i = \hat{i}\cos \omega t \tag{2-91}$$

The alternating voltage across the diode is

$$v = (V_R + V_D)\left[\left(1 + \frac{\hat{i}(1-n)}{\omega C_{\text{tot}} V_R}\sin \omega t\right)^{1/(1-n)} - 1\right] \tag{2-92}$$

An evaluation of this equation shows that especially the first harmonic makes its appearance. The capacitance shift caused by the alternating voltage superimposed on the tuning voltage is shown in Figure 2-63. However, the voltage distortion, and thus the capacitance shift, can be largely avoided if two tuner diodes are used, as in Figure 2-61.

COMPENSATING TEMPERATURE DEPENDENCE. As has already been mentioned, variations are caused in the capacitance of the diode mainly by the dependence of the diffusion voltage on

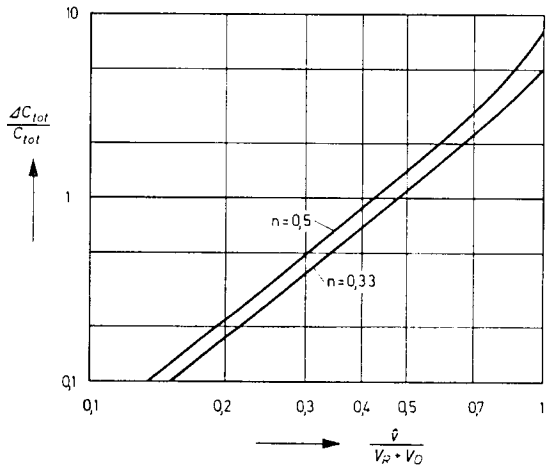


Figure 2-63 Capacitance increase as a function of the ac voltage drop across the tuner diode.

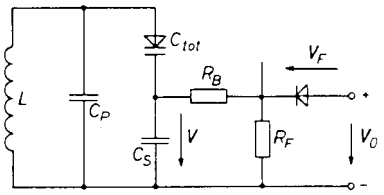


Figure 2-64 Temperature-compensation circuit with diode.

temperature. In a diode-tuned resonant circuit, the temperature coefficient of the resonant frequency, therefore, depends on the tuning voltage and thus on the resonant frequency. It is therefore impossible to compensate the temperature dependence of the resonant frequency by means of temperature-dependent capacitors—the method usually adopted for mechanically tuned circuits.

To achieve satisfactory compensation, the tuning voltage should be increased by an amount equal to that by which the diffusion voltage of the diode is reduced with rising temperature; that is, by approximately 2 mV/°C. This can be done easily by connecting a forward-biased silicon diode in series with the source of the tuning voltage, as shown in Figure 2-64. Since the forward voltage of the diode varies by –2 mV/°C with ambient temperature, the voltage

$$V = V_0 + V_D - V_F$$

which determines the bias of the tuner diode, and therefore also the diode capacitance, is almost temperature independent. In practice, it is advisable to feed this diode with additional current via the resistor R_F to keep its differential forward resistance sufficiently low and thus to prevent capacitive noise voltages from leaking into the tuning circuit.

To prevent the load placed by the resistor R_F on the source of the tuning voltage from fluctuating in a manner depending on the tuning voltage, a circuit such as shown in Figure 2-65 may be used in which the emitter diode of the transistor effects the compensation, and the only load placed on the source of the tuning voltage is the base current of this transistor.

Dynamic Stability. The resonant frequency of a diode-tuned RF circuit follows the equation

$$f = f_0 \sqrt{\left(\frac{V_R + V_D}{V_{R0} + V_D}\right)^n} \tag{2-93}$$

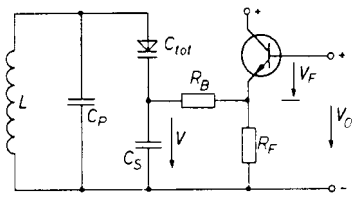


Figure 2-65 Temperature-compensation circuit with transistor.

The frequency f_0 is set with the aid of the circuit inductance when the tuning voltage equals V_{R0}. Since the signal amplitude in the circuit depends on the field strength of the received transmitter and since the resonant frequency of the circuit depends on the amplitude, undesirable feedback between the tuning voltage and the received alternating voltage may occur at the tuner diode. A disturbance that has a similar effect is the microphonic effect, which can sometimes be observed with mechanically variable capacitors. In an investigation of the feedback in diode-tuned resonant circuits, the following assumptions were made:

- The impedance of the antenna (or of the signal generator) is negligible.
- The full alternating voltage is applied to the diode (whereas, in practice, the series capacitor brings about a potential division).
- The capacitance versus voltage characteristic is an exponential function.
- The Q of this circuit is independent of the amplitude of the alternating voltage.

As a result of a high alternating voltage at the resonant circuit, three undesirable effects—harmonic generation, frequency shift, and cross-modulation—are encountered.

HARMONIC GENERATION. The nonlinearity of the capacitance versus voltage characteristic produces harmonics. However, the selectivity of the receiver reduces the effects of these harmonics to noise level.

FREQUENCY SHIFT. When a sinusoidal voltage v is applied to the tuner diode of the RF circuit, the variation of the diode capacitance does not follow the sine law, but the equation

$$\frac{C_\omega}{C} = \left(1 + \frac{v}{V_R + V_D} \sin \omega t\right)^{-n} \tag{2-94}$$

This results in a change of the resonant frequency of the circuit. The inductive slope of the resonant curve steepens and may even turn back. This could lead to bistable behavior of the resonant circuit. The change of frequency takes place as if the tuning voltage, and thus the resonant frequency, has decreased. The order of magnitude of this change of resonant frequency was determined with the aid of a computer from Fourier analysis of

$$\frac{\Delta f}{f} = 1 + \frac{1}{2\pi} \int_0^{2\pi} \left(1 + \frac{v}{V_R + V_D} \sin \omega t\right)^{n/2} d(\omega t) \tag{2-95}$$

Figure 2-66 shows the admissible alternating voltage at the diode, computed for a 2% frequency shift, as a function of the tuning voltage. The measured values differ by not more than 10% from the computed values. Since the Q of the circuits is about 50 for the VHF as well as the UHF range, a 2% frequency shift is permissible.

CROSS-MODULATION. Cross-modulation is a disturbance caused by nonlinearities of the characteristics of the components concerned. The modulation of an undesired amplitude-modulated signal is here transferred to the carrier of the amplitude-modulated intelligence signal. Cross-modulation is virtually independent of the amplitude of the intelligence signal.

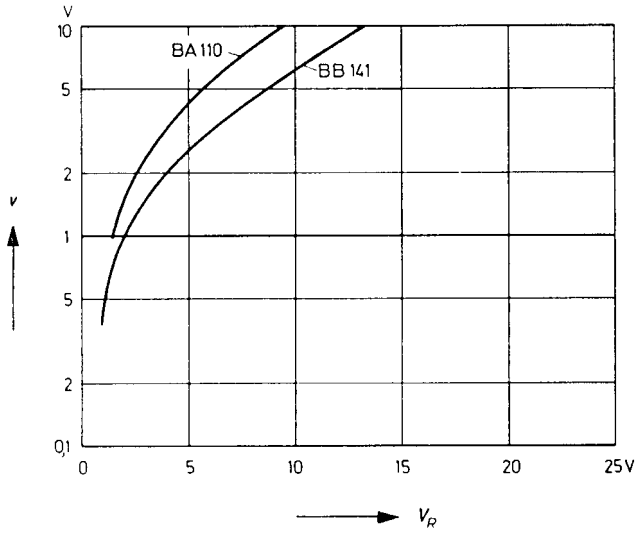


Figure 2-66 Admissible alternating voltage for a 2% frequency shift as a function of tuning voltage.

The cross-modulation factor for tuner diodes is

$$k = \frac{v^2}{(V_R + V_D)^2} m_v \frac{n(n+1)}{2} \tag{2-96}$$

Resolving this equation in terms of v, we obtain

$$v = \frac{V_R + V_D}{n} \sqrt{\frac{2kn}{m_v (n+1)}} \tag{2-97}$$

The usual definition of the cross-modulation percentage for measurement purposes is as follows: A 100% modulated noise signal ($m_v = 1$) of amplitude v is received simultaneously with the intelligence signal in the diode-tuned RF circuit. If a 1% modulation of the signal carrier by the noise signal takes place, this amounts to a cross-modulation factor $k = 0.01$. We thus obtain for tuner diodes

$$V = 0.14 \frac{V_R + V_D}{n} \sqrt{\frac{n}{n+1}} \tag{2-98}$$

Figure 2-67 gives the graphs for this equation. In practical operation, FM and VHF/UHF tuners equipped with the tuner diodes BB121 or BB141 exhibit merely the cross-modulation caused by the BJTs or FETs employed. For bipolar transistors, $v \approx 15$–100 mV; for FETs, 100–300 mV.

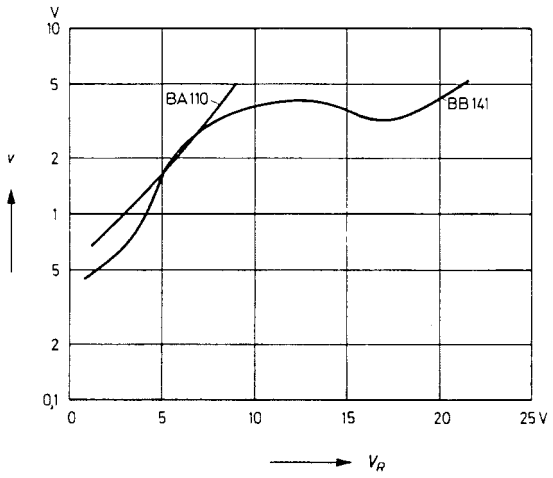


Figure 2-67 Graphical representation of Eq. (2-98).

Static Stability. To ensure trouble-free operation, the static parameters of the tuner diodes, especially the leakage current, must be taken into account. When differentiating the expression for the tuning voltage

$$V_R = V_B - I_R R_B \tag{2-99}$$

we obtain for a known leakage current I_R the tuning voltage variation

$$dV_R = dV_B - R_B I_R \left(\frac{dR_B}{R_B} + \frac{dI_R}{I_R} + dV_B \right) \tag{2-100}$$

This leads to a change in the resonant frequency

$$\frac{df}{f} = \frac{n}{2(V_R + V_D)} \left[dV_D - R_B I_R \left(\frac{dR_B}{R_B} + \frac{dI_R}{I_R} \right) + dV_B \right] \tag{2-101}$$

The resonant frequency is more stable the smaller the exponent n and the higher the tuning voltage V_R. The influence of the temperature dependence of the diffusion voltage V_D and its compensation have already been discussed. If the series bias resistance R_B is not too high, the effect of its variation ($dR_B/R_{B)}$ and the effect of the temperature dependence of the leakage current—doubling for every 10 °C temperature rise—may be neglected.

Generating the Tuning Voltage. The supply voltage V_B in Eq. (2-101) exerts an influence in many respects. The temperature coefficient of the voltage source—usually, for economic reasons, only a Zener diode is provided for stabilization—affects conditions not only with ambient temperature variations, but also when the Zener diode heats up after the set is switched on, or when its operating current varies. Because external temperature compensation of the operating voltage of the Zener diode cannot cover all detrimental effects and, moreover, increases the series bias resistance R_B, as well as any variations the latter may

undergo, the use of a temperature-compensated Zener diode, such as the 1N4065A (nominal Zener voltage, 33 V; temperature coefficient, 0.002 °C/W), is recommended.

Diode Switches. The diode switches described here differ somewhat from the switching diodes used in computer and pulse technology. Whereas in the case of switching diodes the signal itself triggers the switching operation—current does or does not pass through the switching diode depending on the signal level—diode switches allow an alternating current to be switched on or off by means of a direct voltage or a direct current. Figure 2-68 shows an example of the use of diode switches.

Diode Switch Technology. Special manufacturing techniques are needed to attain the low junction capacitance, low differential forward resistance r_F, and the low-inductance structure required for the special applications envisaged. Two conflicting requirements have to be met: On the one hand, a high-resistivity basic material should be used and the diode area should be kept as small as possible to minimize the junction capacitance for a given reverse voltage; on the other hand, the bulk resistance should be kept low. The size of the junction area, therefore, can represent only a compromise between these two requirements. To obtain a low bulk resistance R_B despite the use of high-resistivity silicon around the junction, the diodes are manufactured by the epitaxial process, characterized by a thin high-resistivity layer on a low-resistivity substrate. If the forward current I_F is sufficiently high, then the specific resistance of the epitaxial layer does not affect the bulk resistance, because under these conditions the resistivity of this layer is considerably reduced by being flooded with carriers.

The junction impedance Z_j is somewhat more critical (Figure 2-69). It is composed of the junction capacitance C_S, which does not depend on current, the current-dependent diffusion capacitance

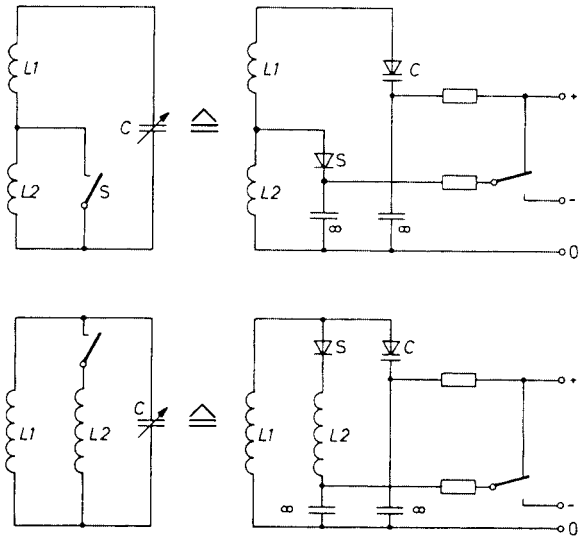


Figure 2-68 Comparison of mechanically and electronically tuned and switched resonant circuits.

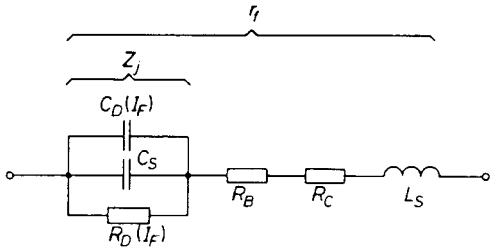


Figure 2-69 Equivalent circuit of diode switches in the frequency range from 1 MHz to 1 GHz.

$$C_D = \frac{\mathfrak{t}_p I_F}{V_T} \tag{2-102}$$

and the diffusion resistance

$$R_D = \frac{V_T}{I_F} \tag{2-103}$$

where $\mathfrak{t}_p$ is the lifetime of the minority charge carriers and

$$V_T = \frac{kT}{q} = 26 \text{ mV} \tag{2-104}$$

is the voltage equivalent of thermal energy. The resistive component of the impedance Z_j is calculated as follows:

$$\text{Re}(Z_j) = \frac{R_D}{1 + \omega^2 R_D^2 (C_S + C_D)^2} \tag{2-105}$$

As can be seen from the following approximations, this resistance becomes rather dependent on current at higher frequencies:

$$\text{For } I_F \to 0 \quad \text{Re}(Z_j) \approx \frac{I_F}{\omega^2 C_S^2 V_T} \tag{2-106}$$

$$\text{For } I_F \to \infty \quad \text{Re}(Z_j) \approx \frac{I_F}{\omega t_p I_F} \tag{2-107}$$

In Figure 2-70 the differential forward resistance is plotted as a function of the forward current. The only parameter by which the differential forward resistance can be reduced for a given current and a given frequency according to Eq. (2-107) is the carrier lifetime. Steps will have to be taken, therefore, to make the carrier lifetime as long as possible.

Because diode switches are designed for applications up to the gigahertz range, a geometry resulting in a low inherent capacitance and inductance had to be used. Furthermore, the device had to be assembled in such a way that virtually no additional contact resistance R_C was introduced. For this reason, a double-plug construction (hard-glass pressure contacts)

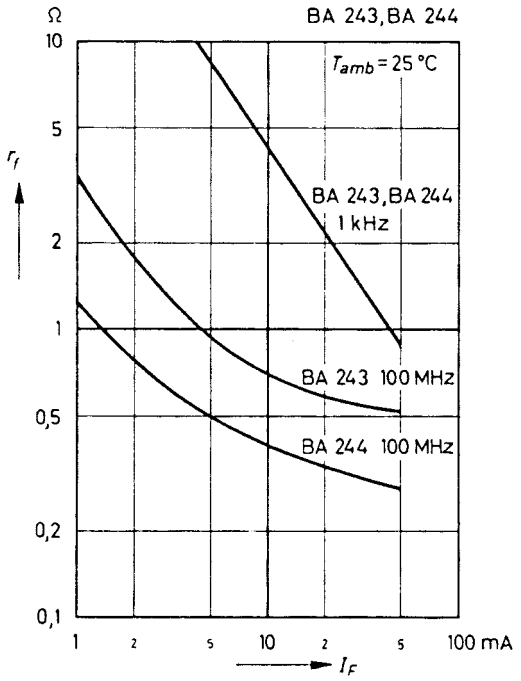


Figure 2-70 Differential forward resistance as a function of forward current.

was chosen to avoid the S- or U-bends customary with other encapsulations. The system is pressure-clamped between two stud-shaped enlargements of the connecting leads, the necessary high pressure being produced as the glass shrinks during cooling. This construction results in an encapsulation of short length (only 4 mm) and low series inductance, the advantages of which can be fully exploited because soldered connections to the leads can be made directly at the glass body.

Diode Switch Data. The silicon planar diode switches BA243 and BA244 are designed for electronic band selection in tuners in the frequency range from 10 to 1000 MHz. The differential forward resistance remains constant and is very low over a wide frequency and current range. The diode capacitance is likewise small and independent of voltage through a wide range. The devices are characterized by low-inductance geometry. The BA243 is intended for the VHF range, and the BA244 for the UHF range. See Table 2-11.

Resonant Circuits Incorporating Diode Switches. Figure 2-71 illustrates the function of the diode switch with reference to a parallel-resonant circuit whose resonant frequency can be switched by short-circuiting part of the circuit inductance. Such a switching procedure may be necessary for a VHF tuner, for example, to switch from band I to band III, because the frequencies of these bands are too far apart.

Capacitor C_K represents the capacitance of the resonant circuit. The symbol ∞ at capacitor C is meant to indicate that the value of this capacitor should be chosen to be much higher than the circuit capacitance. C completes the ac circuit when the inductor L_2 is shorted and

Table 2-11 BA243/244 specifications

Maximum Ratings			
Reverse voltage	V_R	20	V
Forward current at $T_A = 25$ °C	I_F	100	mA
Junction temperature	T_j	150	°C
Storage temperature range	T_s	–55 to +150	°C
Characteristics at $T_A = 25$ °C			
Forward voltage at $I_F = 100$ mA	V_F	<1	V
Leakage current at $V_R = 15$ V	I_R	< 100	nA
Dynamic forward resistance at $f = 50$–1000 MHz, $I_F = 10$ mA			
BA243	r_F	0.7 (< 1)	Ω
BA244	r_F	0.4 (< 0.5)	Ω
Relative variation of dynamic forward resistance with the variation of forward current in the range of $I_F + 2$–40 mA	$\frac{\Delta r_F 100}{r_F \Delta I_F}$	5	%/mA
Capacitance at $V_R = 15$ V, $f = 1$ MHz	C_{tot}	1.3 (< 2)	pF
Relative variation of capacitance with the variation of reverse voltage in the range of $V_R = 7$–20 V, $f = 100$ MHz	$\frac{\Delta C_{tot} 100}{C_{tot} \Delta V_R}$	1	%/V
Series inductance across case	L_S	2.5	nH

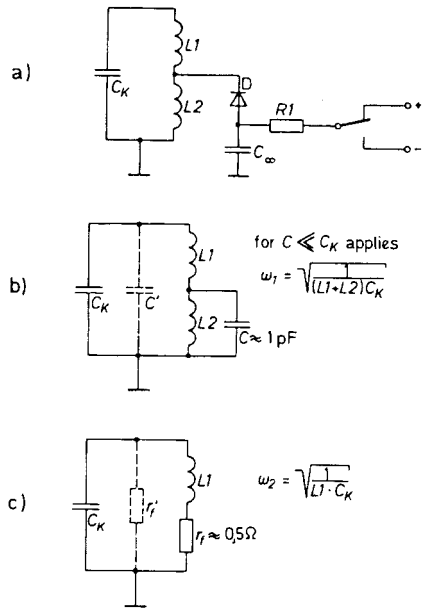


Figure 2-71 (a) Diode switch as bandswitch, shorting a partial inductance; (b) equivalent circuit with blocked diode switch; (c) equivalent circuit with conducting diode switch.

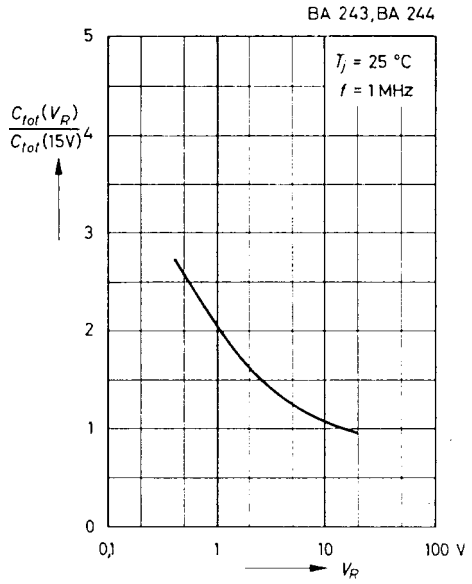


Figure 2-72 Capacitance as a function of the reverse voltage.

enables the switching or reverse voltage to be applied at a point of virtually zero RF potential. The resistor R_1 limits the diode current in the forward direction.

At a reverse voltage of $V_R = 15$ V, for example, the diode switch represents a 1-pF capacitor, as shown in Figure 2-71b. This represents parasitic capacitance C' in the circuit, which reduces the effective capacitance variation of the resonant circuit in the case of capacitive tuning. The capacitance C' is small enough to be negligible in most cases relative to the wiring and winding capacitances. Figure 2-72 shows the relative capacitance of the blocked-diode switches, as a function of the reverse voltage.

At a forward current of, for example, 10 mA, the diode has a differential resistance of approximately 0.5 Ω (see Figure 2-70). The effect of this resistance r_F on the resonant circuit is illustrated in Figure 2-71c. Conversion of a series loss resistance r_F into a parallel loss resistance r'_F is based on the following considerations: The equivalent parallel resistance r'_F affects the tuned circuit in the same way as the series resistance r_F; that is, the Q of the circuit should be the same in both cases. If one assumes that $\omega L >> r_F$, the equivalent of parallel resistance can be expressed as follows:

$$Q = \frac{\omega L_1}{r_F} = \frac{r'_F}{\omega L_1} \tag{2-108}$$

$$r'_f = \frac{(\omega L_1)^2}{r_F} \tag{2-109}$$

An evaluation of this equation shows that a differential resistance $r_F \approx 0.5$ Ω can be tolerated in most applications as regards amplification and selectivity.

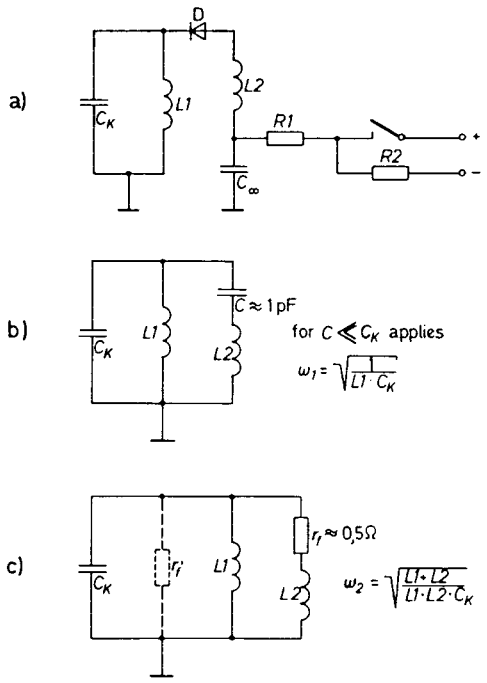


Figure 2-73 (a) Diode switch as a bandswitch, with an inductance in parallel; (b) equivalent circuit with blocked-diode switch; (c) equivalent circuit with conducting diode switch.

Instead of altering the resonant frequency of a tuned circuit by short-circuiting part of its inductance, the same result can be obtained by connecting a second inductance in parallel. The use of the diode switch for this purpose is illustrated in Figure 2-73. This differs from the one shown in Figure 2-71 in that the circuit inductance consists of either L_1 or of L_1 and L_2 in parallel. By applying a permanent negative bias to the diode via a high-value resistor R_2, it is possible to dispense with the use of one switch contact. Figure 2-73b shows the equivalent circuit for the diode under reverse-bias conditions (corresponding to a low resonant frequency), and Figure 2-73c shows the equivalent circuit under forward-bias conditions (corresponding to a higher resonant frequency).

In radio and television receivers, it usually is possible to derive the necessary reverse bias for the diode by rectification of the local oscillator signal, since the power required for blocking the diodes is extremely small. Note that under reverse-bias conditions the inductance L_2 and the capacitance C form a series-tuned acceptor circuit. However, this capacitance being extremely small, any undesirable effects due to its presence can normally be avoided. The same diode switching arrangements described for single-tuned circuits can also be employed in multisection bandpass and broadband filters.

In the following paragraph, a practical example for the use of diode switches will be given, and it will be seen that it is possible, in principle, to replace all of the mechanical RF switch contacts normally employed in receivers by diode switches. In cases where application of the diode dc potentials to a "cold" point in the circuit is not feasible, it may be necessary to apply them via *LC* or *RC* isolating networks.

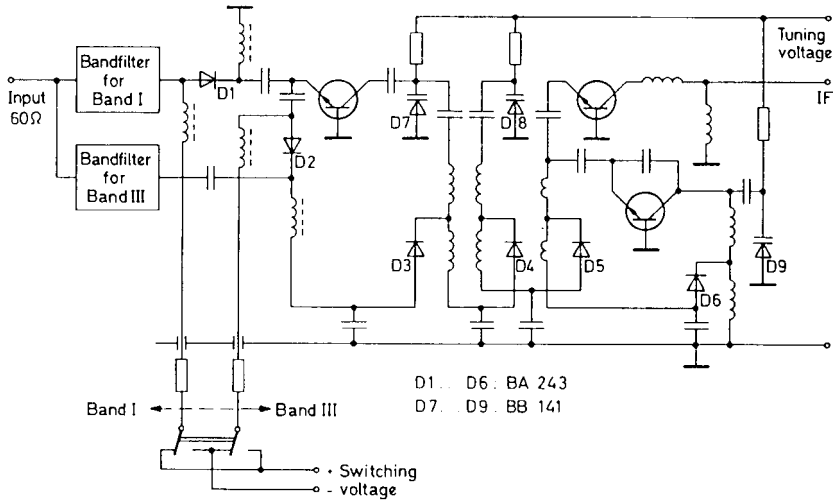


Figure 2-74 Schematic of a VHF television tuner with electronic band selection.

Use of the Diode Switch in a Television Receiver. Figure 2-74 shows the schematic of an electronically tuned and switched band I (48–62 MHz)/band III (175–224 MHz) TV tuner that requires switching of the input filter, bandpass filter, and oscillator circuit. These switching functions are performed by BA243 diode switches. As is customary in most European countries, in this example a broadband input filter between the antenna and the first RF transistor is used rather than a selective front-end circuit. In view of the need to correct noise and power matching over a relatively large frequency range, a switchable filter is employed in most tuners, switching being effected in this case by means of diodes D_1 and D_2.

Diodes D_3–D_6 are used for band-filter switching, for switching the coupling to the mixer transistor, and for switching the oscillator frequency. To minimize the bias-current requirements (each diode requires a bias current of approximately 10 mA), all diodes that are to be activated at the same time are connected in series. This arrangement requires only a few additional RF chokes, which are easily provided because of the high frequencies involved. The direct current required for the diode switches could be obtained by simple rectification, for example, from the heater circuit of the television receiver CRT, and the diode reverse bias could be obtained by rectification of the local oscillator signal, in which case the minute additional power requirement would have to be allowed for in the oscillator design.

A technically sophisticated design for a television tuner takes the form of a combined VHF/UHF unit incorporating electronically switched circuits in which several elements would be common. Such an all-band tuner can be assembled from only three transistors and a combined VHF/UHF band filter in $\lambda/2$ or $\lambda/4$ technology for UHF reception. The VHF circuits are short-circuited when the UHF position is selected.

Multistandard TV receivers require a large number of switch contacts in the tuner, the IF amplifier, the video amplifier, and the time base. The use of diode switches is particularly advantageous in this situation because it renders the physical layout of the equipment virtually independent of the location of the controls and obviates the need for long RF cable runs, which may be prone to interference.

2-2 BIPOLAR TRANSISTORS

2-2-1 Transistor Structure Types

All current RF silicon transistors are of the bipolar *npn* planar epitaxial type. Briefly, the significance of each of these terms is as follows.

Bipolar. In its broadest sense, it is the basic structure shown schematically in Figure 2-75, that is, the familiar three-semiconductor-region structure. *Bipolar* specifically means that the charge carriers of both negative (electron) and positive (hole) polarities are involved in the transistor action. In way of contrast, unipolar types include the junction-gate and insulated-gate field-effect transistors, which are basically one- or two-semiconductor-region structures in which carriers of a single polarity dominate.

npn. This abbreviation is for *negative–positive–negative* and identifies the regions of the structure as to polarity of the dominant or majority carrier in each region. The other polarity type is *pnp*. (See Figure 2-75.)

Silicon. Silicon (Si) is one of two elements from the fourth column of the periodic table that are in widespread use for transistor fabrication [the other is germanium (Ge)]. Other materials used include the compound gallium arsenide (GaAs). Although silicon is in predominant use because it results in the most favorable compromise among high-frequency, high-temperature, high-reliability, and ease-of-use attributes of the usable semiconductor material, silicon–germanium (SiGe) technology is increasingly important because it affords high-frequency performance unachievable with silicon alone.

Planar. The term planar denotes that both emitter–base and base–collector junctions of the transistor intersect the device surface in a common plane (hence, a better term might be *coplanar*). However, the real significance of the so-called planar structure is that the technique of diffusing dopants through an oxide mask, used in fabricating such a structure, results in junctions being formed beneath a protective oxide layer. These protected junctions are less prone to the surface problems sometimes associated with other types of structures, such as the mesa.

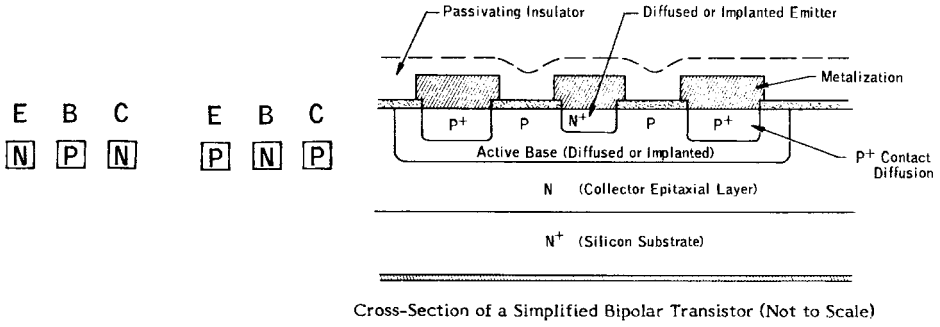


Figure 2-75 Transistor structure schematic.

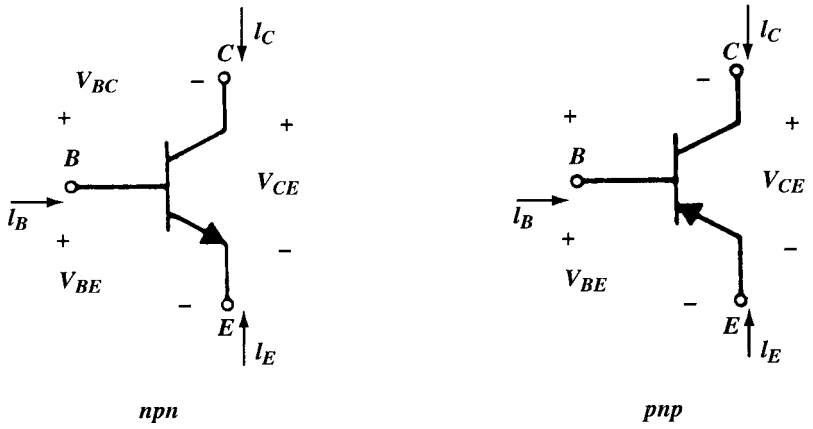


Figure 2-76 Bipolar transistor sign convention.

Epitaxial. This term, as it is commonly used, is actually a shortening of the term *epitaxial-collector*. That is, the collector region of the transistor is formed by the epitaxial technique, rather than by diffusion, which is commonly used to form the base and emitter regions. The epitaxial layer is formed by condensing a single-crystal film of semiconductor material on a wafer or substrate that is usually of the same material. Thus, an epitaxial (collector) transistor is one in which the collector region is formed on a low-resistivity substrate. Subsequently, the base and emitter regions are diffused into the "epi" layer. The epitaxial technique lends itself to precise tailoring of collector-region thickness and resistivity with consequent improved device performance and uniformity.

2-2-2 Large-Signal Behavior of Bipolar Transistors*

In this section, the large-signal or dc behavior of bipolar transistors is considered. Large-signal models are developed for the calculation of total currents and voltages in transistor circuits, and such effects as breakdown-voltage limitations, which are usually not included in models, are also considered. Second-order effects, such as current-gain variation with collector current and Early voltage, can be important in many circuits and are treated in detail.

The sign conventions used for bipolar transistor currents and voltages are shown in Figure 2-76. All bias currents for both *npn* and *pnp* transistors are assumed positive going into the device.

Electrical and Performance Characteristics and Specifications. Electrical characteristics may be described as uniquely defined, measurable electrical transistor properties that are not a function of the measuring circuit or apparatus (except insofar as standard terminations and measurement accuracy are concerned). Performance or operating characteristics are also electrical properties but they are, in general, not unique because their values depend on the measuring circuit (in particular, source and load impedance, which may be arbitrary). As might be expected, the terms are often used somewhat loosely (and sometimes

* Portions of this section, Sections 2-2-3 through 2-2-5, and Sections 2-3-3 through 2-3-6 are based on Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed., ©1993 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

interchangeably), especially in cases where only subtle differences are involved. The terms are generally used on transistor datasheets to segregate (for emphasis) under performance or operating characteristics those properties most directly applicable to the primary intended application.

Performance (Operating) Characteristics. Of the numerous performance characteristics that can be specified for high-frequency transistors, perhaps the most fundamental and pertinent characteristics are:

- Power gain and noise figure, for small-signal applications
- Power gain, power output, and efficiency, for large-signal applications

All of these characteristics are, of course, functions of frequency, bias temperature, and so on, and to completely characterize a transistor over its full frequency, bias, and temperature ranges would be prohibitively costly. Consequently, characterization data are given only for restricted ranges of these variables. These data should portray sufficiently the capabilities of a particular device for its intended applications. In the case of maximum ratings, some applications may require additional characterization by the user or through applications assistance from the manufacturer.

POWER GAIN

G_{max}. Of the various definitions for the measurement of power flow in an active two-port device, such as a transistor, two are unique enough to allow specification without recourse to specifying the complete measuring circuit in detail. One of these definitions is termed *maximum available gain* (G_{max}) and is the power gain obtained when the input and output ports are simultaneously conjugately matched to source and load impedances, respectively. Implicit in this definition is the assumption that the two-port is unconditionally stable; that is, that no combination of input/output tuning can result in increasing gain to the point of oscillation.

$|S_{21}|^2$. The other unique power gain is the gain realized when the transistor is inserted between a source and load with identical impedances (in practice, usually $50 + j0\ \Omega$). This particular insertion of transducer gain happens to coincide with the usual definition of the two-port forward scattering parameter, S_{21}. More precisely, it is equal to the magnitude squared of this parameter and is therefore often identified by the symbol $|S_{21}|^2$. For wideband applications, $|S_{21}|^2$ is important because wideband terminations "not too different from 50 Ω" are more easily realized than are wideband transforming networks that provide the matching required for G_{max}.

NOISE FIGURE. A common measure of the noise generated by an active two-port device—noise that sets a lower limit on amplifier sensitivity—is noise factor (F). This is defined as

$$F = \frac{\text{Input signal-to-noise ratio}}{\text{Output signal-to-noise ratio}} \tag{2-110}$$

or more generally as

$$F = \frac{\text{Total output noise power}}{\text{Output power due to source resistance}} \tag{2-111}$$

At high frequencies, spot noise factor or noise factor for a small fractional bandwidth (say, 1%) is used and is usually expressed as *noise figure* (*NF*), in decibels:

$$NF = 10 \log F \tag{2-112}$$

As already discussed, noise figure is a function of source impedance (as well as functions of frequency, bias, etc.) and hence, there is an infinity of noise figures associated with a given device corresponding to the infinity of possible impedances that may be presented to the device input. The only unique one, in the sense that it does not involve arbitrary source impedances, is $NF_{\min}$, the minimum noise figure obtained (at a given bias and frequency) when the input is tuned to optimize this parameter. This is the noise figure value commonly given on datasheets.

In practical amplifiers that involve more than one stage, *overall noise factor* (F_O) is given by

$$F_O = F_1 + \frac{F_2 - 1}{G_1} + \cdots + \frac{F_n - 1}{G_{(n-1)}} \tag{2-113}$$

where n = number of stages
G_n = gain of the nth stage
F_n = noise factor of the nth stage

This expression emphasizes the important fact that for low-noise amplifiers, the first stage must be designed for the lowest noise figure and highest gain possible. [Note that the noise contribution of the second stage is divided (reduced) by the gain of the first stage.] Since the optimum source impedance and bias currents for optimum gain and noise figure do not often coincide, very careful circuit design is required to minimize overall noise figure. For more accuracy, we recommend the noise correlation approach.

POWER OUTPUT. This characteristic is important for both amplifier and oscillator transistors. In both cases, it is extremely circuit sensitive. For amplifiers, maximum useful output is often limited to that power output level at which gain has compressed 1 dB, an indicator of the upper limit of the device's linearity range. For oscillators, it is merely a quantitative measure of RF power for given dc input power.

EFFICIENCY. In the most general sense, this characteristic expresses as a percentage the ratio of RF power output to the total circuit input power, both dc and RF. That expression is efficiency, η, defined as

$$\eta = \frac{P_O}{P_i + P_{\text{dc}}} \times 100 \tag{2-114}$$

where P_O = RF output power
P_i = RF input power
P_{dc} = total dc power input

Since for oscillator transistors there is no RF power input, and for amplifier transistors the maximum input RF power is calculable from the power gain and power output specifications,

the inclusion of P_i in efficiency is redundant. Moreover, since the major portion of the dc power is dissipated by the transistor collector, a more restricted definition of efficiency is pertinent. This parameter, termed *collector efficiency* (η_C), is given by

$$\eta_C = \frac{P_O}{P_{CC}} \times 100 \tag{2-115}$$

where $P_{CC} = V_{CC} \times I_{CC}$
V_{CC} = collector supply voltage
I_{CC} = collector supply current

Electrical Characteristics. Electrical characteristics may be conveniently classified into two main types, dc and ac.

dc CHARACTERISTICS. The importance of dc characteristics of high-frequency transistors lies primarily in biasing and reliability considerations. However, certain dc characteristics are also directly related to high-frequency performance. For example, high-frequency noise figure is affected by dc current gain. The dc characteristics discussed here are those usually found on high-frequency transistor datasheets.

$V_{(BR)CBO}$, I_{CBO}. These two parameters serve to characterize the reverse-biased collector–base *pn* junction and are defined as follows (with the aid of Figure 2-77a). The collector–base breakdown voltage, $V_{(BR)CBO}$, identifies the voltage at which collector current tends to increase without limit, usually due to the high electric field developed across the junction. This voltage sets a limit on the maximum transistor operating voltage and, as mentioned before under maximum ratings, usually is the basis for the collector–base maximum voltage rating. $V_{(BR)CBO}$ should be specified at a value of $I_C = I_{C1}$ in the figure, which is within the avalanche (or high slope) region of the device's reverse characteristic. Typical values of I_{C1} are in the 1–10-μA region for high-frequency transistors.

To further define the quality of the reverse *I–V* characteristic, a specification is usually placed on the collector cutoff current, I_{CBO}, which is measured at some value of collector–base voltage less than $V_{(BR)CBO}$. For a good-quality silicon junction ("sharp" instead of soft; see Figure 2-77a), I_{CBO} is in the nanoampere range.

$V_{(BR)EBO}$, I_{EBO}. These two parameters characterize the reverse-biased emitter–base *pn* junction in an analogous manner to the collector–base junction parameters $V_{(BR)CBO}$ and I_{CBO}, given above, and are shown in Figure 2-77b. No further discussion of these parameters will be given here.

$V_{(BR)CEO}$, I_{CEO}. The collector–emitter breakdown voltage and cutoff current are somewhat more complex in nature than the collector–base and emitter–base parameters. In the latter two, only a single *pn* diode is involved. In the collector–emitter case, two diodes are involved. Moreover, each is influenced by the other through transistor action, since the reverse current of the collector–base diode flows through the emitter–base junction as forward current. Thus, the collector–base reverse current is amplified by the dc current gain of the transistor, resulting in:

- I_{CEO} being greater than I_{CBO} (for a given voltage)
- Typically, the familiar negative-resistance region in the *I–V* characteristic as shown in Figure 2-77c.

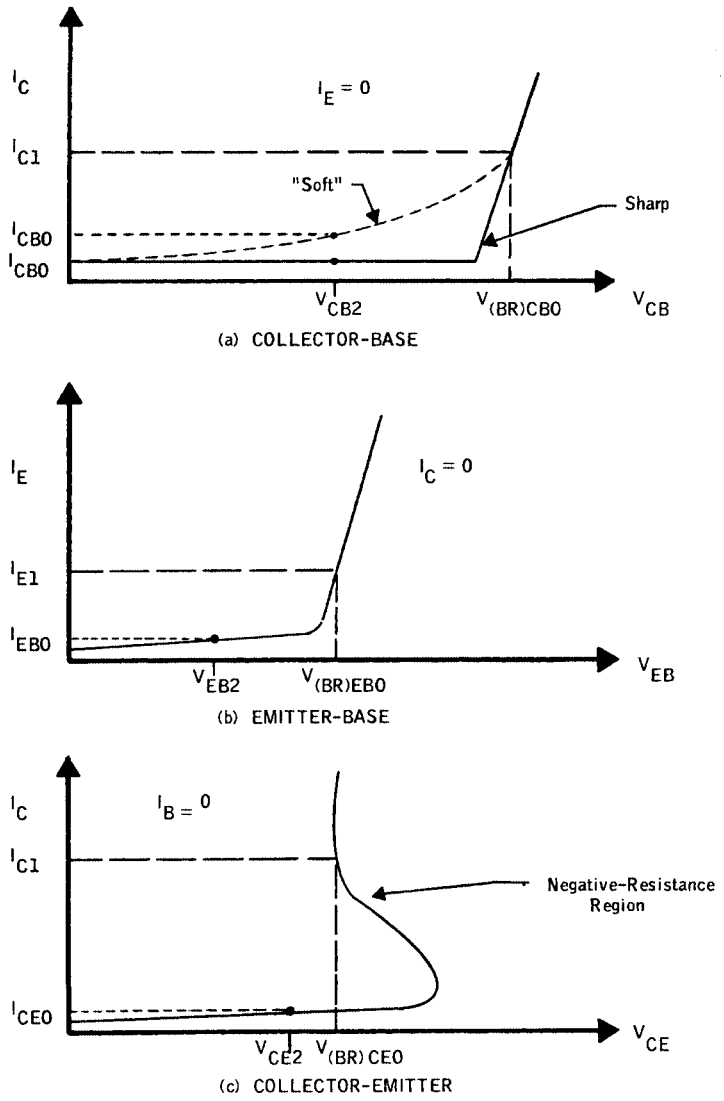


Figure 2-77 Transistor reverse *I–V* characteristics.

Consequently, $V_{(BR)CEO}$ is typically specified at collector currents one to three orders of magnitude higher than in the case of $V_{(BR)CBO}$ and $V_{(BR)EBO}$ to establish the minimum value of this characteristic.

h_{fe}. This parameter is simply the dc common-emitter current gain; that is, the ratio of collector current to base current at some specified collector voltage and current.

ac CHARACTERISTICS. Of the numerous ac characteristics defined for transistors, only relatively few are commonly used in characterizing high-frequency transistors. Some of the more pertinent parameters are briefly covered here.

S Parameters. By far the most useful and conveniently measured set of two-port parameters for transistor high-frequency (roughly 100 MHz and above) characterization is

the *S*-parameter or scattering-matrix set. These parameters completely and uniquely define the small-signal gain and input/output immittance properties of any linear "black box." (By definition, a transistor or any active device is linear under small-signal conditions.) However, these parameters reveal nothing (except possibly indirectly and approximately) about large-signal or noise behavior. Simply interpreted (more general definitions and other interpretations abound in the technical literature), the *S* parameters are merely insertion gains, forward and reverse, and reflection coefficients, input and output, with driven and nondriven ports both terminated in equal impedances, usually $50 + j0\ \Omega$. Such an interpretation tends to make *S* parameters very attractive, once some familiarity is gained, at high (especially microwave) frequencies, since the power flow or gain and reflection-coefficient concepts are more intuitively meaningful than voltage and current conceptual schemes. It should also be mentioned that *S*-parameters can be converted through straightforward matrix transformations to other two-port parameter sets, such as *h*-, *y*-, and *z*-parameters.

Proceeding with more specific definitions, the *S*-parameters are defined analytically by

$$b_1 = S_{11}a_1 + S_{12}a_2 \tag{2-116}$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \tag{2-117}$$

or, in matrix form,

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{2-118}$$

where, referring to Figure 2-78

a_1 = (incoming power at Port 1)$^{1/2}$
b_1 = (outgoing power at Port 1)$^{1/2}$
a_2 = (incoming power at Port 2)$^{1/2}$
b_2 = (outgoing power at Port 2)$^{1/2}$
E_1, E_2 = electrical stimuli at Port 1, Port 2

From the figure and defining linear equations, for $E_2 = 0$, then $a_2 = 0$, and (skipping through numerous rigorous steps)

$$S_{11} = \frac{b_1}{a_1}$$

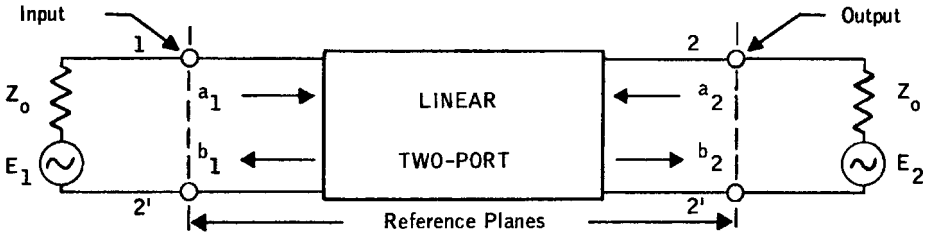


Figure 2-78 Two-port *S*-parameter definition.

$$= \left(\frac{\text{Outgoing input power}}{\text{incoming input power}}\right)^{1/2}$$

$$= \frac{\text{Reflected voltage}}{\text{Incident voltage}}$$

$$= \text{Input reflection coefficient} \tag{2-119}$$

$$S_{21} = \frac{b_2}{a_1}$$

$$= \left(\frac{\text{Outgoing output power}}{\text{incoming input power}}\right)^{1/2}$$

$$= \left(\frac{\text{Output power}}{\text{Available input power}}\right)^{1/2}$$

$$= (\text{Forward transducer gain})^{1/2} \tag{2-120}$$

or, more precisely in the case of S_{21}:

$$\text{Forward transducer gain} = G_{\text{TF}} = |S_{21}|^2 \tag{2-121}$$

$$Z_i = Z_o \tag{2-122}$$

Similarly, at Port 2 for $E_1 = 0$, then $a_1 = 0$, and

$$S_{22} = \frac{b_2}{a_2} = \text{Output reflection coefficient} \tag{2-123}$$

$$S_{12} = \frac{b_1}{a_2} = (\text{Reverse transducer gain})^{1/2} \tag{2-124}$$

$$G_{\text{TR}} = |S_{12}|^2 \tag{2-125}$$

Since many measurement systems display *S*-parameter magnitudes in decibels, the following relationships are particularly useful:

$$|S_{11}|_{\text{dB}} = 10\log|S_{11}|^2$$

$$= 20\log|S_{11}| \tag{2-126}$$

$$|S_{22}|_{\text{dB}} = 20\log|S_{22}| \tag{2-127}$$

$$|S_{21}|_{\text{dB}} = 10\log|S_{21}|^2$$

$$= 20 \log|S_{21}|$$

$$= 10 \log|G_{TF}| = |G_{TF}|_{dB} \tag{2-128}$$

$$|S_{12}|_{dB} = 10 \log|S_{12}|^2$$

$$= 20 \log|S_{12}|$$

$$= 10 \log|G_{TR}| = |G_{TR}|_{dB} \tag{2-129}$$

Figure 2-79 shows an example of the use of *S* parameters in transistor stage design.

Transition Frequency. One of the better known, but perhaps least understood, figures of merit for high-frequency transistors is the so-called transition frequency, f_T. Part of the misunderstanding that appears to exist is due to the use of the misleading (but common, for

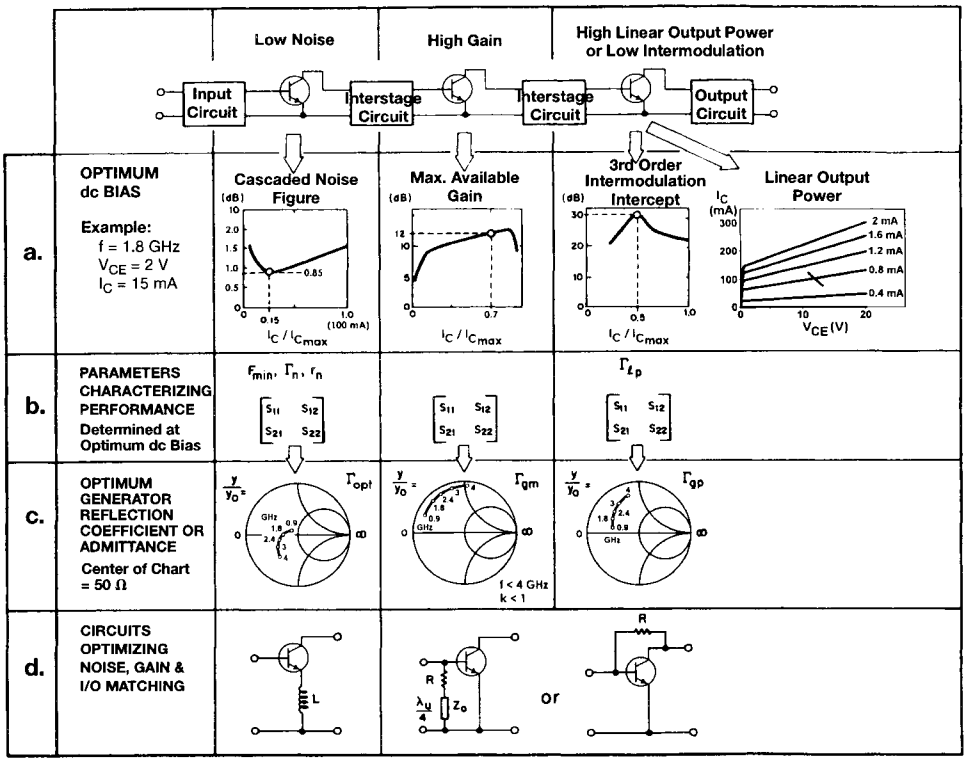


Figure 2-79 Key parameters in applying a BJT in low-noise front-end, high-gain, and linear-power stages. (a) The BJT characteristics that lead to optimum dc bias for a low-noise front end; (b) parameters characterizing device performance; (c) Γ_{opt} for the first stage, and S_{11} and S_{22} for the output stage as simulated by CAD software. The example is based on the Siemens BJTs BFP420 (low-noise stage), BFP450 (high-gain stage), and BFG235 (output stage). Circuits to optimize stage noise, gain, and I/O matching are shown in (d). (Presentation based on Figure 14 in Charles A. Liechti, "Microwave Field-Effect Transistors—1976," *IEEE Trans. Microwave Theory Tech.,* vol. MTT-24, pp. 279–300, June 1976.)

historical reasons) term *short-circuit gain–bandwidth product* for this parameter. By definition, f_T is that characteristic frequency described by the equation

$$f_T = h_{fe} \times f_{\text{meas}} \tag{2-130}$$

where h_{fe} is the magnitude of the small-signal common-emitter short-circuit current gain, and f_{meas} is the frequency of measurement, chosen such that

$$2 \le h_{fe} \le \frac{h_{fe0}}{2} \tag{2-131}$$

where h_{fe0} is the low-frequency value of h_{fe}.

To varying degrees of approximation, depending on the transistor type, f_T is the frequency at which h_{fe} approximates unity. It is not, in general, the frequency at which h_{fe} is precisely equal to unity. To clarify these points further, consider the plot of h_{fe} against frequency sketched in Figure 2-80.

At low frequencies, $f << f_B$, h_{fe} is constant and equal to h_{fe0}.

At f_B, h_{fe} has decreased to 0.707 h_{fe0}; that is, f_B is the 3-dB cutoff frequency for common-emitter short-circuit current gain, h_{fe}.

For frequencies such that

$$2f_B < f < f_T \tag{2-132}$$

h_{fe} varies inversely proportional to frequency. That is, the f_T defining relationship holds:

$$h_{fe} \times f_{\text{meas}} = \text{constant} = f_T \tag{2-133}$$

At frequencies approaching f_T, other parameters, especially package parasitics, can cause $|h_{fe}|$ to depart significantly from this $1/f$ variation. Therefore, the frequency, f_1, at which $|h_{fe}|$ actually equals unity can be somewhat different from f_T.

Applying these frequency-gain characteristics to the common-emitter wideband, low-pass amplifiers gives rise to the terminology of f_T being a "gain–bandwidth product." However,

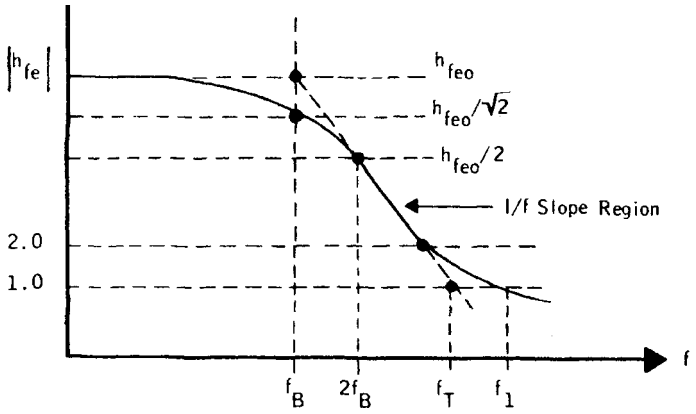


Figure 2-80 $|h_{fe}|$ frequency characteristics.

this is an optimistic approximation at best, since the product of low-frequency circuit gain and the 3-dB cutoff frequency is reduced from f_T by an amount that depends on circuit impedances.

The real significance of f_T lies in the fact that it is a measure of certain internal transistor parameters that do, in fact, affect high-frequency performance: for example, gain (though not in the convenient quantitative manner implied by the term "gain–bandwidth product"). In particular, good high-frequency noise performance requires that f_T be high. Thus, f_T is included on transistor datasheets primarily as a figure of merit, not as a parameter to be used directly in design.

Collector–Base Time Constant, $r_b'C_c$. This is an internal device parameter that relates only indirectly to high-frequency performance. It is primarily a measure of internal feedback within the transistor. It also relates to transistor high-frequency impedance. As the name says (in symbols), it is a measure of transistor base resistance and collector capacitance in combination; however, except for certain low-frequency transistors, it cannot be considered the simple two-element lumped *RC* time constant implied by the terminology. (In high-frequency transistors, base resistance and collector impedance must be considered to be distributed when considered in detail). As a figure of merit, it is included on transistor datasheets to indicate how well base resistance and collector impedance have been minimized. It also allows the estimation of certain gain properties of the transistor (see the $f_{\max}$ parameter, which we'll describe shortly).

Collector–Base Capacitance, C_{cb}. This parameter is simply the total collector–base *pn* junction capacitance measured at a low frequency (typically, 1 MHz), where it can be considered a single lumped element. For high-frequency transistors it is, of course, desirable that C_{cb} be small from bandwidth and stability considerations as well as from gain considerations alone.

Maximum Frequency of Oscillation, $f_{\max}$. This is another figure-of-merit parameter, as opposed to measurable parameters directly usable in the application of transistors. Its importance stems from the following approximate relationships (which will not be derived here):

$$f_{\max} \approx \left(\frac{f_T}{8\pi r_b' C_C} \right)^{1/2} \tag{2-134}$$

$$G_{\max} \approx \left(\frac{f_{\max}}{f_{\text{oper}}} \right)^2 \tag{2-135}$$

These expressions illustrate in a quantitative way the importance and the interrelationship between high f_T and low $r_b'C_C$ insofar as high-frequency gain is concerned. However, since they are approximations and since their derivation involves several assumptions that are not always valid, they must be interpreted with caution. For example, the expression for $G_{\max}$ is obviously not applicable at low frequencies, since as $f \to 0$, $G_{\max} \to \infty$, according to this expression. As a rule of thumb, the $G_{\max}$ expression is a reasonable approximation for frequencies such that

$$5 > \frac{f_{\max}}{f_{\text{oper}}} > 1 \tag{2-136}$$

For accurate analysis of transistor gain and stability, a complete set of two-port parameters must be employed in exact expressions, such as those from which the approximations shown above were derived.

Datasheet Transistor Specifications. The following is a reprint of a Siemens microwave transistor datasheet. This 25-GHz f_T transistor will soon be replaced by a 40-GHz family, based on a silicon–germanium design. Key parameters describing the transistor are:

- The dc parameters, such as maximum current and voltage and dissipation
- Noise figure as a function of generator impedance, bias point, and frequency
- Frequency-dependent behavior, typically expressed in *S* parameters, and also as a function of bias point and frequency
- Large-signal performance to be calculated based on a nonlinear model, such as the Gummel–Poon

The transistor description typically gives details about packaging, as the Gummel–Poon model is intrinsic (without parasitics). Not only does the package determine the parasitics, it also affects the device's ability to transfer heat to its heat sink. The following figures, reproduced with permission from the datasheet for the BFP420 transistor by Siemens (now Infineon Technologies), are a typical representation of the dc-I_C curves, *S* parameters, and noise parameters of a normal, packaged microwave transistor. Besides its dc parameters and capacitance, there are other critical parameters. A BJT's noise figure is a strong function of its dc current gain, and one can determine a 3-dB cutoff frequency referred to as f_T, which shows the gain roll-off as a function of collector current and frequency. Another important measure or figure of merit is $f_{\max}$, the maximum frequency of oscillation:

$$f_{\max} = \sqrt{\frac{|Y_{21}|}{16\pi R_{bb} C_{be} C_{bc}}} \tag{2-137}$$

Figure 2-81 shows a photograph of the die for the BFP420.

Since the intercept point is an important figure of merit, Figure 2-82 shows intercept point as a function of bias point for the BFR93AW BJT. We wish more manufacturers would supply this information.

2-2-3 Large-Signal Transistors in the Forward-Active Region

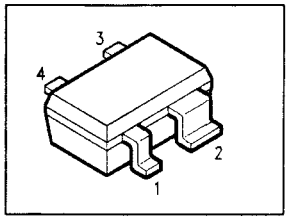
We saw in the diode case that its description is physics-based and it is, therefore, logical to apply the same modeling for the transistors. The first complete modeling for transistors was described in the Gummel–Poon model, but many engineers use the somewhat more simpler Ebers–Moll model. By combining two diodes, the parasitic elements, and the base separating resistor, we can show the "intrinsic" nonlinear model for the bipolar transistor.

Many equations that have been published contain some approximations, but for exact nonlinear analysis, we have to resort to the somewhat more complex or complete equations as shown in Appendix A for the bipolar model.

Recently, some other models have been proposed by researchers, such as the MEXTRAM models (Philips), but the trouble with these models is that while they may give quite good results for RF applications, it is virtually impossible to obtain parameters for them. So far the only industry-standard model is the Gummel–Poon, and to the authors' knowledge there

SIEMENS **SIEGET®25 BFP420**

NPN Silicon RF Transistor



- For High Gain Low Noise Amplifiers
- For Oscillators up to 10 GHz
- Noise Figure F = 1.05 dB at 1.8 GHz
  Outstanding G_{ms} = 20 dB at 1.8 GHz
- Transition Frequency f_T = 25 GHz
- Gold metallization for high reliability
- **SIEGET®25-Line**
  Siemens **G**rounded **E**mitter **T**ransistor-
  25 GHz f_T-**Line**

ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering Code (8-mm taped)	Pin Configuration 1	2	3	4	Package [1]
BFP420	AMs	Q62702-F1591	B	E	C	E	SOT343

Maximum Ratings

Parameter	Symbol		Unit
Collector-emitter voltage	V_{CEO}	4.5	V
Collector-base voltage	V_{CBO}	15	V
Emitter-base voltage	V_{EBO}	1.5	V
Collector current	I_C	35	mA
Base current	I_B	3	mA
Total power dissipation, Ts ≤ 107°C [2)3)]	P_{tot}	160	mW
Junction temperature	T_j	150	°C
Ambient temperature range	T_A	-65...+150°C	°C
Storage temperature range	T_{stg}	-65...+150°C	°C

Thermal Resistance

Junction-soldering point [2)]	$R_{th\ JS}$	270	K/W

1) For detailed information see chapter Package

2) T_S is measured on the emitter lead at the soldering point to the pcb.

3) P_{tot} due to Maximum Ratings.
At typical Ts ≤ 80°C: P_{tot} = 250 mW due to thermical characteristics.

SIEMENS **SIEGET®25 BFP420**

Electrical Characteristics

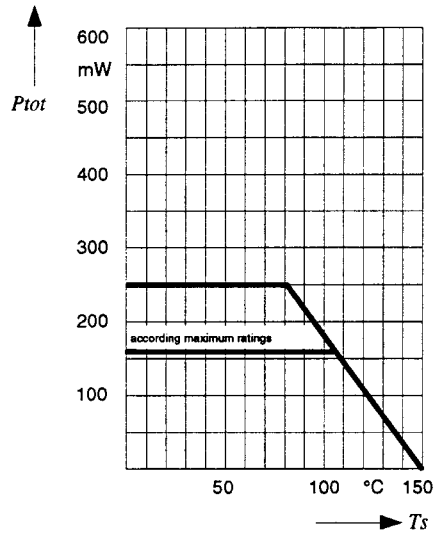
at T_A = 25 °C, unless otherwise specified.

Parameter	Symbol	Value			Unit
		min.	typ.	max.	
DC Characteristics					
Collector-emitter breakdown voltage I_C = 1 mA	$V_{(BR)CEO}$	4.5	5	6.5	V
Collector-cutoff current V_{CB} = 5 V, I_E = 0	I_{CBO}	-	-	200	nA
Emitter base cuttoff current V_{EB} = 1.5 V, I_C = 0	I_{EBO}	-	-	35	µA
DC current gain I_C = 20 mA, V_{CE} = 4 V	h_{FE}	50	80	150	
AC Characteristics					
Transition frequency I_C = 30 mA, V_{CE} = 3 V, f = 2 GHz	f_T	20	25	-	GHz
Collector-base capacitance V_{CB} = 2 V, V_{BE} = v_{be} = 0, f = 1 MHz	C_{cb}	-	0.15	0.24	pF
Collector-emitter capacitance V_{CE} = 2 V, V_{BE} = v_{be} = 0, f = 1 MHz	C_{ce}	-	0.41	-	pF
Emitter-base capacitance V_{EB} = 0.5 V, V_{CB} = v_{cb} = 0, f = 1 MHz	C_{eb}	-	0.55	-	pF
Noise figure I_C = 5 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}	F	-	1.05	1.4	dB
Power gain I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt} , Z_L = Z_{Lopt}	G_{ms} [1]	-	20	-	dB
Insertion power gain I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_L = 50Ω	$\lvert S_{21}\rvert^2$	14	17	-	dB
Third order intercept point at output I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}, Z_L = Z_{Lopt}	IP_3	-	22	-	dBm
1dB Compression point I_C = 20 mA, V_{CE} = 2 V, f = 1.8 GHz, Z_S = Z_{Sopt}, Z_L = Z_{Lopt}	P_{-1dB}	-	12	-	dBm

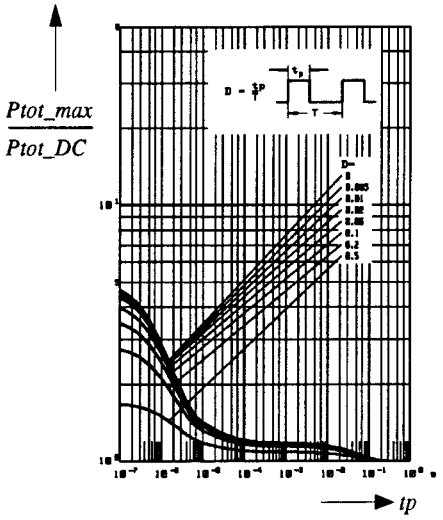
1) $G_{ms} = \left|\frac{S21}{S12}\right|$

SIEMENS **SIEGET®25 BFP420**

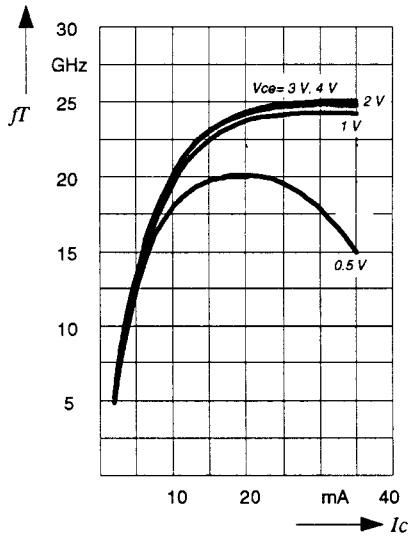
Total Power Dissipation
versus Soldering Point Temperature



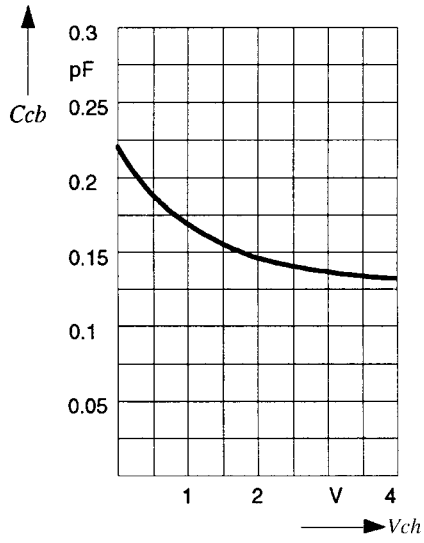
Permissible Pulse Power Dissipation
versus On-Time (V_{CE0max} = 4.5 V)



Transition Frequency
versus Collector Current
f = 2 GHz



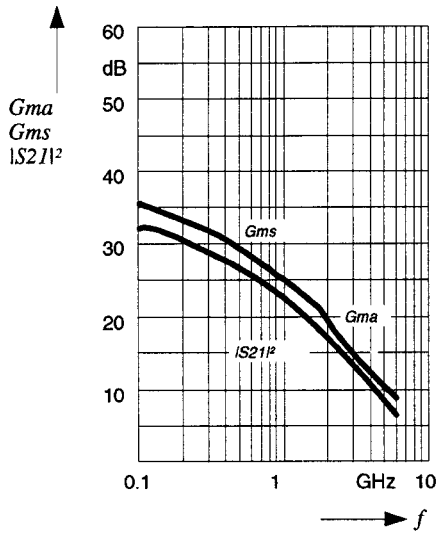
Collector-base Capacitance
versus Collector-base Voltage
V_{BE} = 0 V, f = 1MHz



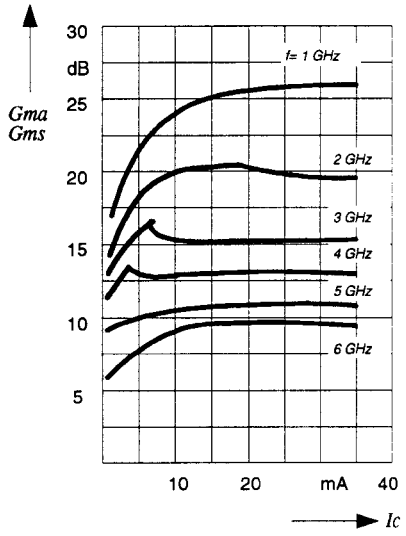
SIEMENS

SIEGET®25 BFP420

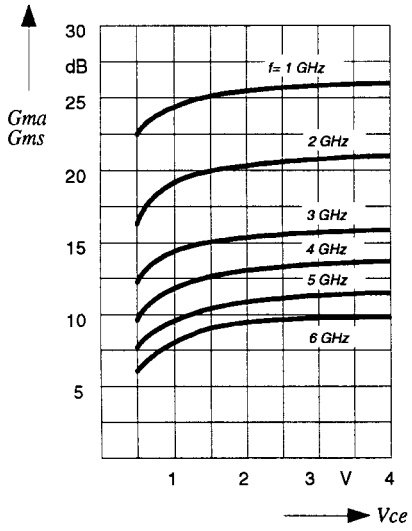
Power Gain
versus Frequency
V_{CE} = 2 V , I_C = 20 mA



Power Gain
versus Collector Current
V_{CE} = 2 V

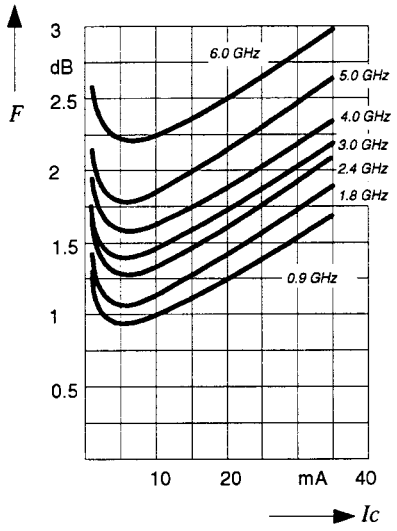


Power Gain
versus Collector Voltage
I_C = 20 mA

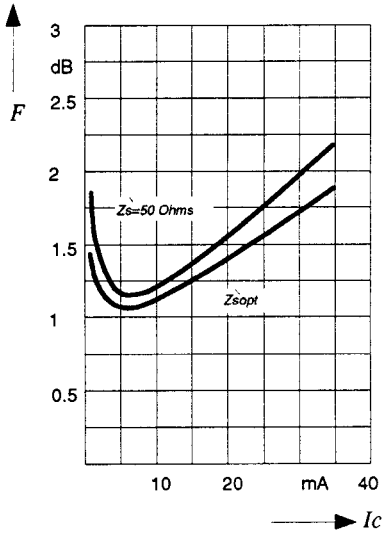


SIEMENS SIEGET®25 BFP420

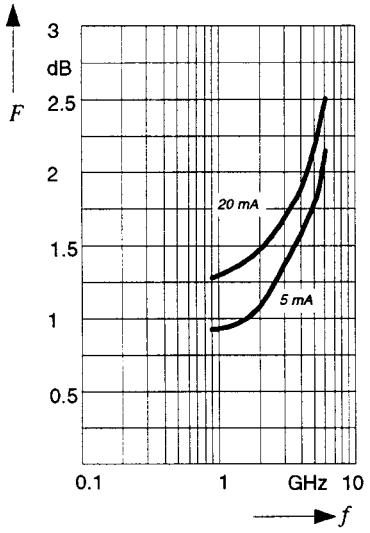
Noise Figure
versus Collector Current
V_{CE} = 2 V, Z_S = Z_{Sopt}



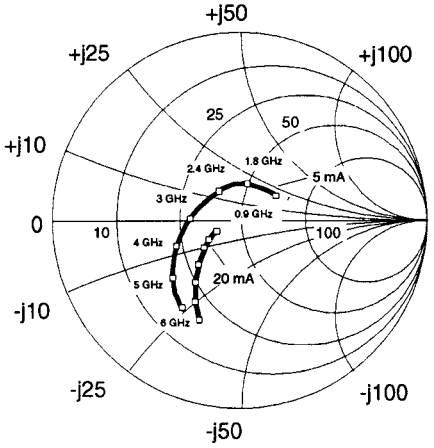
Noise Figure
versus Collector Current
VCE = 2 V, f = 1.8 GHz



Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA,
Z_S = Z_{Sopt}



Source Impedance for min.
Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA



SIEMENS **SIEGET®25 BFP420**

Common Emitter S-Parameters

f	S_{11}		S_{21}		S_{12}		S_{22}	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
V_{CE} = 2 V, I_C = 20 mA								
0.01	0.452	-2.3	37.62	178.3	0.0011	94.4	0.956	-0.6
0.1	0.447	-25.1	36.30	164.7	0.0068	82.5	0.941	-12.4
0.5	0.386	-101.1	23.41	121.0	0.0262	61.7	0.632	-47.2
1.0	0.378	-146.2	13.99	96.0	0.0395	57.8	0.395	-63.9
2.0	0.405	173.5	7.18	70.8	0.0664	54.0	0.222	-87.3
3.0	0.446	149.4	4.77	52.6	0.0949	47.1	0.133	-111.3
4.0	0.501	130.0	3.52	36.8	0.1206	38.5	0.133	-158.5
6.0	0.599	104.8	2.27	8.2	0.1646	18.9	0.196	142.0
8.0	0.700	78.5	1.51	-20.8	0.1800	-2.4	0.289	99.3
9.0	0.758	67.6	1.25	-34.4	0.1820	-13.0	0.379	84.1
10.0	0.800	62.0	1.04	-43.5	0.1800	-19.3	0.465	76.6
V_{CE} = 2 V, I_C = 5 mA								
0.01	0.790	-1.0	15.14	179.2	0.0012	83.4	0.988	-0.7
0.1	0.786	-11.6	14.98	171.8	0.0092	84.1	0.982	-6.5
0.5	0.702	-55.7	12.86	140.1	0.0398	62.8	0.857	-29.8
1.0	0.589	-99.1	9.63	112.6	0.0603	46.5	0.647	-48.6
2.0	0.507	-156.0	5.60	79.4	0.0798	34.6	0.401	-70.3
3.0	0.511	168.5	3.84	57.1	0.0957	29.8	0.267	-84.2
4.0	0.549	142.0	2.87	38.5	0.1121	25.1	0.207	-110.5
5.0	0.604	123.9	2.26	22.1	0.1285	19.4	0.150	-137.3
6.0	0.633	110.0	1.86	6.7	0.1442	13.1	0.173	-169.8

Common Emitter Noise Parameters

f	F_{min} 1)	G_a 1)	Γ_{opt}		R_N	r_n	$F_{50\Omega}$ 2)	$\lvert S_{21}\rvert^2$ 2)
GHz	dB	dB	MAG	ANG	Ω	-	dB	dB
V_{CE} = 2 V, I_C = 5 mA								
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	-175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	-157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	-142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	-123.0	15.0	0.30	3.30	5.3

1) Input matched for minimum noise figure, output for maximum gain 2) Z_S=Z_L=50 Ω

SIEMENS **SIEGET®25 BFP420**

SPICE Parameters:

Transistor Chip Data T502 (Berkeley-SPICE 2G.6 Syntax):

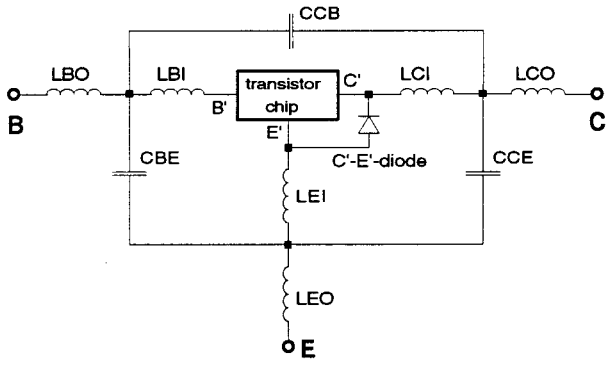
IS =	0.20045	fA	BF =	72.534	-	NF =	1.2432	-
VAF =	28.383	V	IKF =	0.48731	A	ISE =	19.049	fA
NE =	2.0518	-	BR =	7.8287	-	NR =	1.3325	-
VAR =	19.705	V	IKR =	0.69141	A	ISC =	0.019237	fA
NC =	1.1724	-	RB =	3.4849	OHM	IRB =	0.72983	mA
RBM =	8.5757	OHM	RE =	0.31111	OHM	RC =	0.10105	OHM
CJE =	1.8063	fF	VJE =	0.8051	V	MJE =	0.46576	-
TF =	6.7661	ps	XTF =	0.42199	-	VTF =	0.23794	V
ITF =	1.0	mA	PTF =	0	deg	CJC =	234.53	fF
VJC =	0.81969	V	MJC =	0.30232	-	XCJC =	0.3	-
TR =	2.3249	ns	CJS =	0	fF	VJS =	0.75	V
MJS =	0	-	XTB =	0	-	EG =	1.11	eV
XTI =	3.0	-	FC =	0.73234	-	Tnom =	300	K

C'-E'- Diode Data (Berkeley-SPICE 2G.6 Syntax):

IS =	3.5	fA	N =	1.02	-	RS =	10	Ω

All parameters are ready to use, no scaling is necessary. Caution: In these extracted parameters, RBM > the operating RB.

Package Equivalent Circuit:



SOT343-3:

LBI = 0.47 nH
LBO = 0.53 nH
LEI = 0.23 nH
LEO = 0.05 nH
LCI = 0.56 nH
LCO = 0.58 nH
CBE = 136 fF
CCE = 134 fF
CCB = 6.9 fF

valid up to 6 GHz

The SOT343 package has two emitter leads. To avoid high complexity of the package equivalent circuit, both leads are combined in one electrical connection.

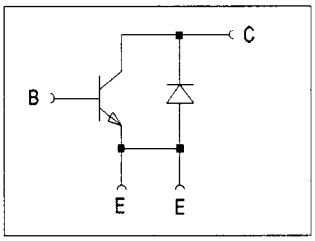
Extracted on behalf of SIEMENS Small Signal Semiconductors by:
Institut für Mobil- und Satellitenfunktechnik (IMST)
© 1996 SIEMENS AG

For non-linear simulation:

- Use transistor chip parameters in Berkeley SPICE 2G.6 syntax for all simulators.
- If you need simulation of the reverse characteristics, add the diode with the C'-E'-diode data between collector and emitter.
- Simulation of the package is not necessary for frequencies < 100 MHz.
  For higher frequencies add the wiring of the package equivalent circuit around the non-linear transistor and diode model.

Note:

- This transistor is constructed in a common emitter configuration. This feature causes an additional, reverse biased diode between emitter and collector, which does not effect normal operation.



Transistor Schematic Diagram

The common emitter configuration shows the following advantages:

- Higher gain because of lower emitter inductance.
- Power is dissipated via the grounded emitter leads, because the chip is mounted on the copper emitter leadframe.

Please note, that the broadest lead is the emitter lead.

- The AC-Characteristics are verified by random sampling.

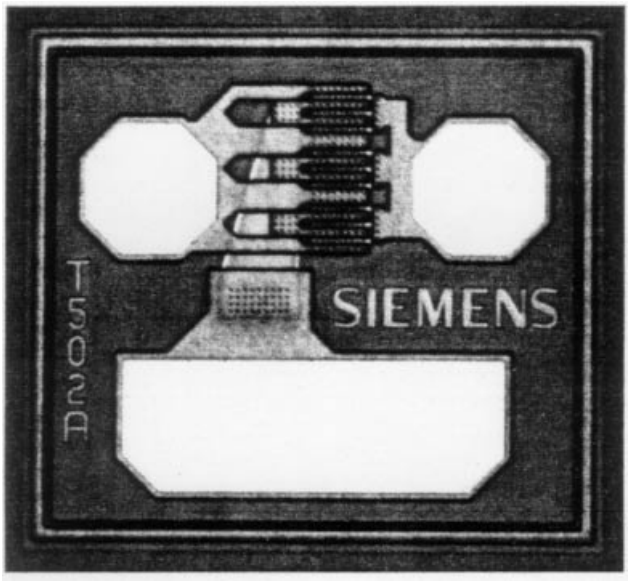


Figure 2-81 Photograph of the BFP420 die (310 × 290 μm). The bonding pad for the base is at the upper right; for the collector at the upper left. The large, lower pad connects to the emitter.

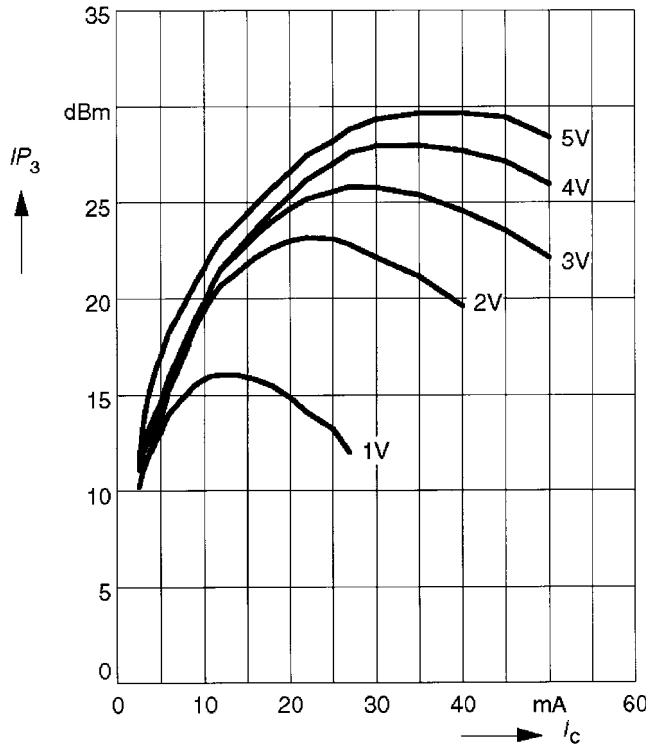


Figure 2-82 $IP_{3,out}$ versus I_C at five values of V_{CE} for the BRF93AW BJT.

are no accepted or affordable parameter extractor programs available, not even to consider the cost of hardware for such extractions. Hewlett-Packard recently published a system with a hardware cost in the vicinity of $1 million.

A typical *npn* planar bipolar transistor structure is shown in Figure 2-83a, where collector, base, and emitter are labeled *C*, *B*, and *E*, respectively. The impurity doping density in the base and the emitter of such a transistor is not constant but varies with distance from the top surface. However, many of the characteristics of such a device can be predicted by analyzing

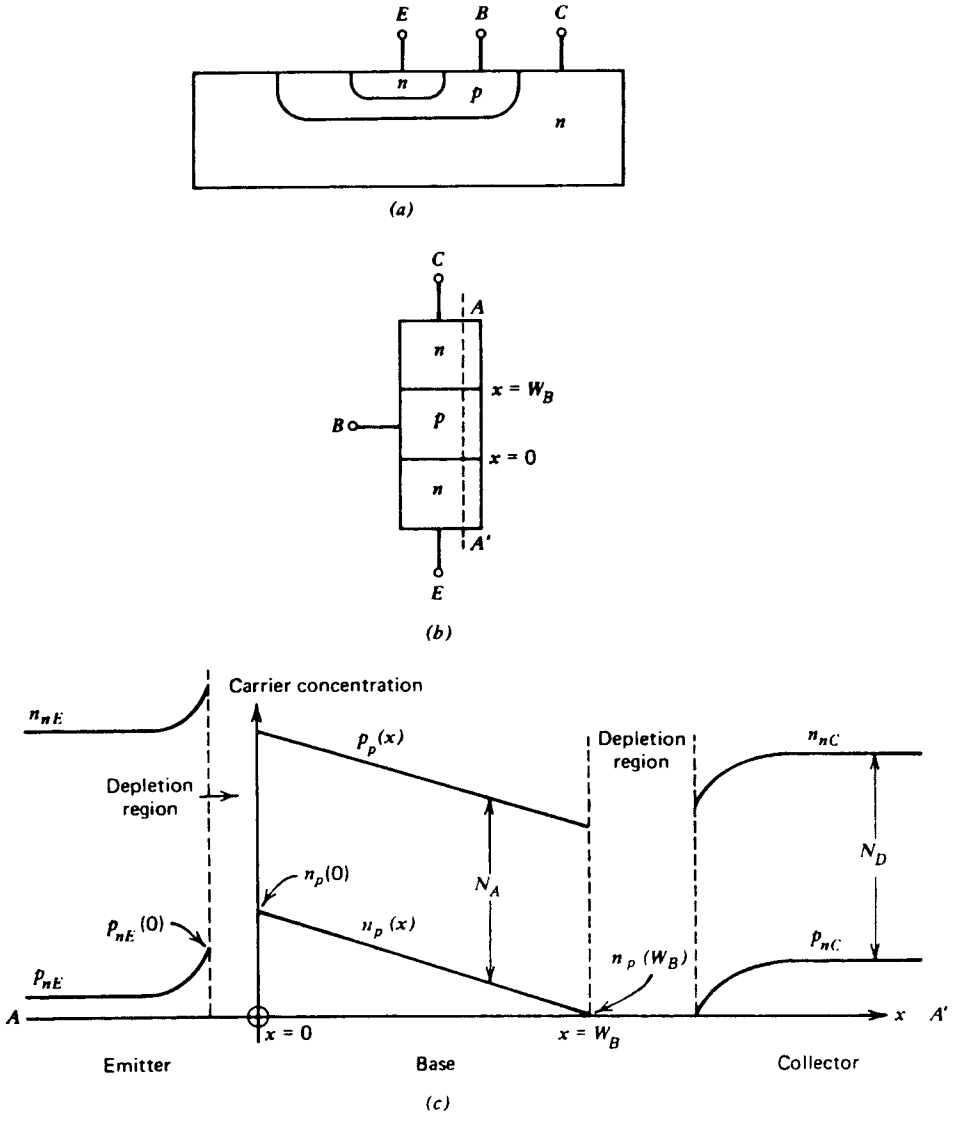


Figure 2-83 (a) Cross section of a typical *npn* planar bipolar transistor structure; (b) idealized transistor structure; (*c*) carrier concentrations along the cross section *AA*′ of the transistor in (b). Uniform doping densities are assumed. (Not to scale)

the idealized transistor structure shown in Figure 2-83b. In this structure, the base and emitter doping densities are assumed constant, and this is sometimes called a "uniform-base" transistor. Where possible in the following analyses, the equations for the uniform-base analysis are expressed in a form that applies also to nonuniform-base transistors.

A cross section AA' is taken through the device of Figure 2-83b and carrier concentrations along this section are plotted in Figure 2-83c. Hole concentrations are denoted by p and electron concentrations by n with subscripts p or n representing p-type or n-type regions. The n-type emitter and collector regions are distinguished by subscripts E and C, respectively. The carrier concentrations shown in Figure 2-83c apply to a device in the *forward-active region*. That is, the base–emitter junction is forward-biased and the base–collector junction is reverse-biased. The minority-carrier concentrations in the base at the edges of the depletion regions can be calculated from a Boltzmann approximation to the Fermi–Dirac distribution function to give [2]

$$n_p(0) = n_{po} \exp \frac{V_{BE}}{V_T} \tag{2-138}$$

$$n_p(W_B) = n_{po} \exp \frac{V_{BC}}{V_T} \simeq 0 \tag{2-139}$$

where W_B is the width of the base from the base–emitter junction depletion layer edge of the base–collector depletion layer edge and n_{po} is the equilibrium concentration of electrons in the base. Note that V_{BC} is negative for an *npn* transistor in the forward-active region and thus $n_p(W_B)$ is very small. Low-level injection conditions are assumed in the derivation of Eqs. (2-138) and (2-139). This means that the minority-carrier concentrations are always assumed much smaller than the majority-carrier concentrations.

If *recombination* of holes and electrons in the base is small, it can be shown that the minority-carrier concentration $n_p(x)$ in the base varies *linearly* with distance [3]. Thus, a straight line can be drawn joining the concentrations at $x = 0$ and $x = W_B$ in Figure 2-83c.

For charge neutrality in the base, it is necessary that

$$N_A + n_p(x) = p_p(x) \tag{2-140}$$

and thus

$$p_p(x) - n_p(x) = N_A \tag{2-141}$$

where $p_p(x)$ is the hole concentration in the base and N_A is the base doping density that is assumed constant. Equation (2-141) indicates that the hole and electron concentrations are separated by a constant amount and thus $p_p(x)$ also varies linearly with distance.

Collector current is produced by minority-carrier electrons in the base diffusing in the direction of the concentration gradient and being swept across the collector–base depletion region by the field existing there. The diffusion current density due to electrons in the base is

$$J_n = qD_n \frac{dn_p(x)}{dx} \tag{2-142}$$

where D_n is the diffusion constant for electrons. From Figure 2-83c,

$$J_n = -qD_n \frac{n_p(0)}{W_B} \tag{2-143}$$

If I_C is the collector current and is taken as positive flowing *into* the collector, it follows from Eq. (2-143) that

$$I_C = qA D_n \frac{n_p(0)}{W_B} \tag{2-144}$$

where A is the cross-sectional area of the emitter. Substitution of Eq. (2-138) into Eq. (2-144) gives

$$I_C = \frac{qAD_n n_{po}}{W_B} \exp \frac{V_{BE}}{V_T} \tag{2-145}$$

$$= I_S \exp \frac{V_{BE}}{V_T} \tag{2-146}$$

where

$$I_S = \frac{qA D_n n_{po}}{W_B} \tag{2-147}$$

and I_S is a constant used to describe the transfer characteristic of the transistor in the forward-active region. Equation (2-147) can be expressed in terms of the base doping density by noting that [4]

$$n_{po} = \frac{n_i^2}{N_A} \tag{2-148}$$

and substitution of Eq. (2-148) in Eq. (2-147) gives

$$I_S = \frac{qAD_n n_i^2}{W_B N_A} = \frac{qA\overline{D}_n n_i^2}{Q_B} \tag{2-149}$$

where $Q_B = W_B N_A$ is the number of doping atoms in the base per unit area of the emitter and n_i is the intrinsic carrier concentration in silicon. In this form Eq. (2-149) applies to both uniform- and nonuniform-base transistors and D_n has been replaced by $\overline{D}_n$, which is an average effective value of the electron diffusion constant in the base. This is necessary for nonuniform-base devices because the diffusion constant is a function of impurity concentration. Typical values of I_S as given by Eq. (2-149) are 10^{-14} to 10^{-16} A.

Equation (2-146) gives the collector current as a function of base–emitter voltage. The base current I_B is also an important parameter and, at moderate current levels, consists of two major components. One of these (I_{B1}) represents recombination of holes and electrons in the base and is proportional to the minority-carrier charge Q_e in the base. From Figure 2-83c, the minority-carrier charge in the base is

$$Q_e = \frac{1}{2} n_p(0) W_B q A \tag{2-150}$$

and we have

$$I_{B1} = \frac{Q_e}{\tau_b} = \frac{1}{2} \frac{n_p(0) W_B q A}{\tau_b} \tag{2-151}$$

where τ_b is the minority-carrier lifetime in the base. I_{B1} represents a flow of majority holes from the base lead into the base region. Substitution of Eq. (2-138) into Eq. (2-151) gives

$$I_{B1} = \frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} \exp \frac{V_{BE}}{V_T} \tag{2-152}$$

The second major component of base current (usually the dominant one in the integrated-circuit *npn* devices) is due to injection of holes from the base into the emitter. This current component depends on the gradient of minority-carrier holes into the emitter and is [5]

$$I_{B2} = \frac{q A D_p}{L_p} p_{nE}(0) \tag{2-153}$$

where D_p is the diffusion constant for holes and L_p is the diffusion length (assumed small) for holes in the emitter. $p_{nE}(0)$ is the concentration of holes in the emitter at the edge of the depletion region and is

$$P_{nE}(0) = p_{nEo} \exp \frac{V_{BE}}{V_T} \tag{2-154}$$

If N_D is the donor atom concentration in the emitter (assumed constant) then

$$p_{nEo} \simeq \frac{n_i^2}{N_D} \tag{2-155}$$

The emitter is deliberately doped much more heavily than the base, making N_D large and p_{nEo} small, so that the base-current component, I_{B2}, is minimized.

Substitution of Eqs. (2-154) and (2-155) into Eq. (2-153) gives

$$I_{B2} = \frac{q A D_p}{L_p} \frac{n_i^2}{N_D} \exp \frac{V_{BE}}{V_T} \tag{2-156}$$

The total base current, I_B, is the sum of I_{B1} and I_{B2}:

$$I_B = I_{B1} + I_{B2} = \left(\frac{1}{2} \frac{n_{po} W_B q A}{\tau_b} + \frac{q A D_p}{L_p} \frac{n_i^2}{N_D} \right) \exp \frac{V_{BE}}{V_T} \tag{2-157}$$

Although this equation was derived assuming uniform base and emitter doping, it gives the correct functional dependence of I_B on device parameters for practical double-diffused

nonuniform-base devices. Second-order components of I_B, which are important at low current levels, are considered later.

Since I_C in Eq. (2-146) and I_B into Eq. (2-147) are both proportional to $\exp(V_{BE}/V_T)$ in this analysis, the base current can be expressed in terms of collector current as

$$I_B = \frac{I_C}{\beta_F} \tag{2-158}$$

where β_F is the forward current gain. An expression for β_F can be calculated by substituting Eqs. (2-145) and (2-157) into Eq. (2-158) to give

$$\beta_F = \frac{qAD_n n_{po}/W_B}{\frac{1}{2}\frac{n_{po}W_B qA}{\tau_b} + \frac{qAD_p n_i^2}{L_p N_D}} = \frac{1}{\frac{W_B^2}{2\tau_b D_n} + \frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \tag{2-159}$$

where Eq. (2-148) has been substituted for n_{po}. Equation (2-159) shows that β_F is maximized by minimizing the base width W_B and maximizing the ratio of emitter to base doping densities N_D/N_A. Typical values of β_F for *npn* transistors in integrated circuits are 50–500, whereas lateral *pnp* transistors have values of 10–100. Finally, the emitter current is

$$I_E = -(I_C + I_B) = -\left(I_C + \frac{I_C}{\beta_F}\right) = -\frac{I_C}{\alpha_F} \tag{2-160}$$

where

$$\alpha_F = \frac{\beta_F}{1+\beta_F} \tag{2-161}$$

The value of α_F can be expressed in terms of device parameters by substituting Eq. (2-159) into Eq. (2-161) to obtain

$$\alpha_F = \frac{1}{1+\frac{1}{\beta_F}} = \frac{1}{1+\frac{W_B^2}{2\tau_b D_n} + \frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \simeq \alpha_T \gamma \tag{2-162}$$

where

$$\alpha_T = \frac{1}{1+\frac{W_B^2}{2\tau_b D_n}} \tag{2-163}$$

$$\gamma = \frac{1}{1+\frac{D_p}{D_n}\frac{W_B}{L_p}\frac{N_A}{N_D}} \tag{2-164}$$

The validity of Eq. (2-162) depends on $W_B^2/2\tau_b D_n << 1$ and $(D_p/D_n)(W_B/L_p)(N_A/N_D) << 1$, and this is always true if β_F is large [see Eq. (2-159)]. The term γ in Eq. (2-162) is

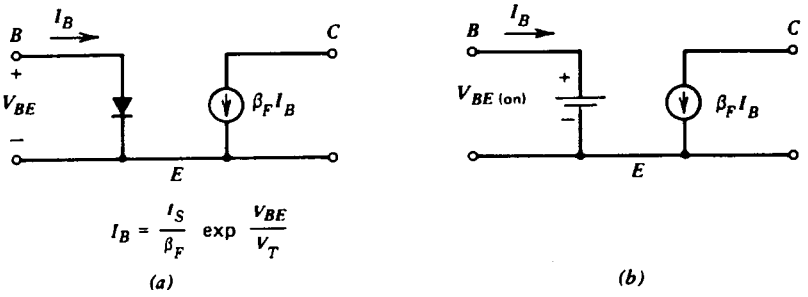


Figure 2-84 Large-signal models of *npn* transistors for use in bias calculations. (a) Circuit incorporating an input diode. (b) Simplified circuit with an input voltage source.

called the *emitter injection efficiency* and is equal to the ratio of the electron current (*npn* transistor) injected into the base from the emitter to the total hole and electron current crossing the base–emitter junction. Ideally $\gamma \to 1$, and this is achieved by making N_D/N_A large and W_B small. In that case very little reverse injection occurs from base to emitter.

The term α_T in Eq. (2-162) is called the *base transport factor* and represents the fraction of carriers injected into the base (from the emitter) that reach the collector. Ideally $\alpha_T \to 1$, and this is achieved by making W_B small. It is evident from the above development that fabrication changes that cause α_T and γ to approach unity also maximize the value of β_F of the transistor.

The results derived above allow formulation of a large-signal model of the transistor suitable for bias-circuit calculations with devices in the forward-active region. One such circuit is shown in Figure 2-84 and consists of a base–emitter diode to model Eq. (2-157) and a controlled collector-current generator to model Eq. (2-158). Note that the collector voltage ideally has no influence on the collector current and the collector node acts as a high-impedance current source. A simpler version of this equivalent circuit, which is often useful, is shown in Figure 2-84b, where the input diode has been replaced by a battery with a value $V_{BE(\text{on})}$, which is usually 0.6–0.7 V. This represents the fact that in the forward-active region the base–emitter voltage varies very little because of the steep slope of the exponential characteristic. In some circuits the temperature coefficient of $V_{BE(\text{on})}$ is important and a typical value for this is –2 mV/°C. The equivalent circuits of Figure 2-84 apply for *npn* transistors. For *pnp* devices the corresponding equivalent circuits are shown in Figure 2-85.

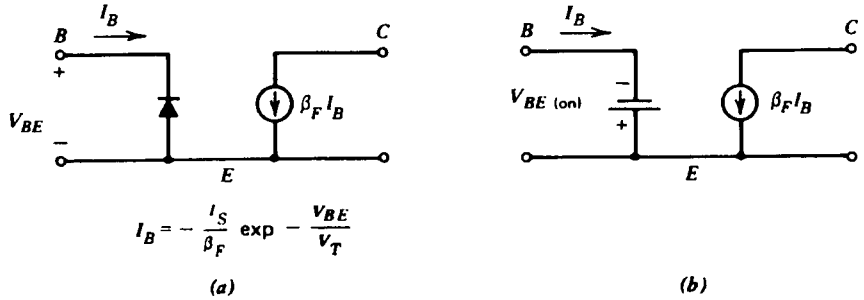


Figure 2-85 Large-signal models of *pnp* transistors corresponding to the circuits of Figure 2-84.

2-2-4 Effects of Collector Voltage on Large-Signal Characteristics in the Forward-Active Region

In the analysis of the previous section, the collector–base junction was assumed reverse-biased and ideally had no effect on the collector currents. This is a useful approximation for first-order calculations, but it is not strictly true in practice. There are occasions where the influence of collector voltage on collector current is important, and this will now be investigated.

The collector voltage has a dramatic effect on the collector current in two regions of device operation. These are the saturation (V_{CE} approaches zero) and breakdown (V_{CE} very large) regions that will be considered later. For values of collector–emitter voltage V_{CE} between these extremes, the collector current increases slowly as V_{CE} increases. The reason for this can be seen from Figure 2-86, which is a sketch of the minority-carrier concentration in the base of the transistor. Consider the effect of changes in V_{CE} on the carrier concentration for constant V_{BE}. Since V_{BE} is constant, the change in V_{CB} equals the change in V_{CE}, and this causes an increase in the collector–base depletion-layer width as shown. The change in the base width of the transistor, ΔW_B, equals the change in depletion-layer width and causes an increase ΔI_C in the collector current.

From Eqs. (2-146) and (2-149) we have

$$I_C = \frac{qA\bar{D}_n n_i^2}{Q_B} \exp \frac{V_{BE}}{V_T} \tag{2-165}$$

Differentiation of Eq. (2-165) yields

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{qA\bar{D}_n n_i^2}{Q_B^2} \left(\exp \frac{V_{BE}}{V_T} \right) \frac{dQ_B}{dV_{CE}} \tag{2-166}$$

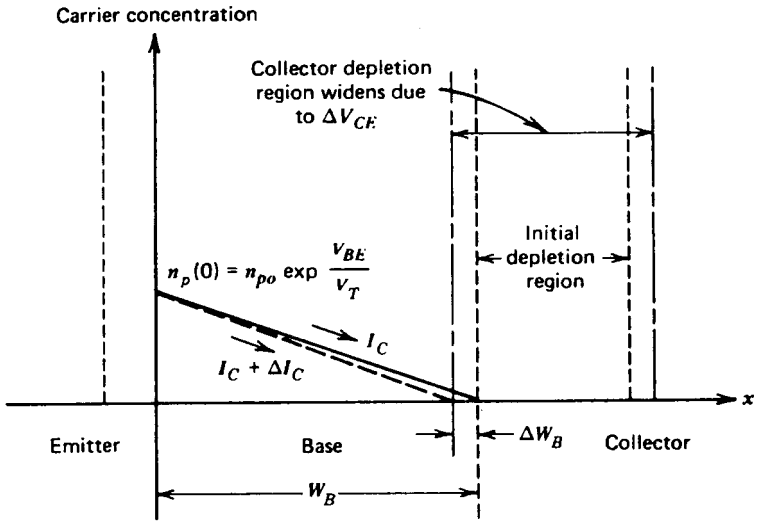


Figure 2-86 Effect of increases in V_{CE} on the collector depletion region and base width of a bipolar transistor.

and substitution of Eq. (2-165) in Eq. (2-166) gives

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{I_C}{Q_B}\frac{dQ_B}{dV_{CE}} \tag{2-167}$$

For a uniform-base transistor $Q_B = W_B N_A$ and Eq. (2-167) becomes

$$\frac{\partial I_C}{\partial V_{CE}} = -\frac{I_C}{W_B}\frac{dW_B}{dV_{CE}} \tag{2-168}$$

Note that since the base width *decreases* as V_{CE} increases, dW_B/dV_{CE} in Eq. (2-168) is negative and thus $\partial I_C/\partial V_{CE}$ is positive. dW_B/dV_{CE} is a function of the bias value of V_{CE}, but the variation is typically small for a reverse-biased junction and dW_B/dV_{CE} is often assumed constant. The resulting predictions agree adequately with experimental results.

Equation (2-168) shows that $\partial I_C/\partial V_{CE}$ is proportional to the collector-bias current and inversely proportional to the transistor base width. Thus narrow-base transistors show a greater dependence of I_C on V_{CE} in the forward-active region. The dependence of $\partial I_C/\partial V_{CE}$ on I_C results in typical transistor output characteristics as shown in Figure 2-87. In accordance with the assumptions made in the foregoing analysis, these characteristics are shown for constant values of V_{BE}. However, in most integrated-circuit transistors the base current is dependent only on V_{BE} and not on V_{CE}, and thus constant-base-current characteristics can often be used in the following calculation. The reason for this is that the base current is usually dominated by the I_{B2} component of Eq. (2-156), which has no dependence on V_{CE}. Extrapolation of the characteristics of Figure 2-87 back to the V_{CE} axis gives an intercept V_A called the Early voltage, where

$$V_A = \frac{I_C}{\dfrac{\partial I_C}{\partial V_{CE}}} \tag{2-169}$$

Substitution of Eq. (2-168) in Eq. (2-169) gives

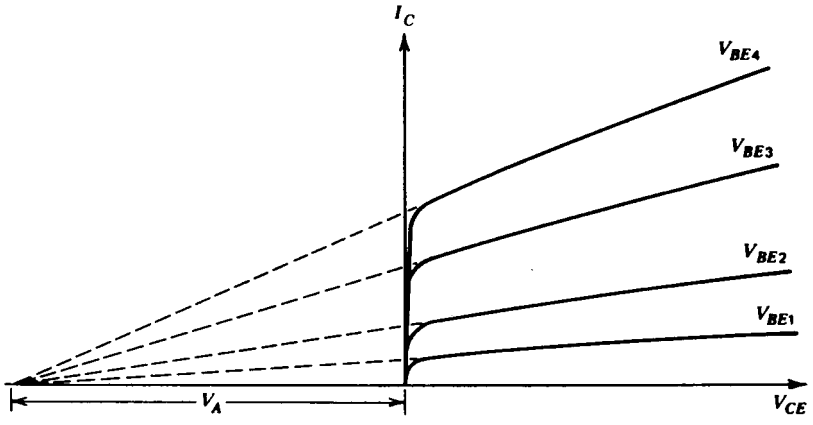


Figure 2-87 Bipolar transistor output characteristics showing the Early voltage, V_A.

$$V_A = -W_B \frac{dV_{CE}}{dW_B} \tag{2-170}$$

which is a constant, independent of I_C. Thus all the characteristics extrapolate to the same point on the V_{CE} axis. The variation of I_C with V_{CE} is called the Early effect and V_A is a common model parameter for circuit-analysis computer programs. Typical values of V_A for integrated-circuit transistors are 15–100 V. The inclusion of Early effect in dc bias calculations is usually limited to computer analysis because of the complexity introduced into the calculation. However, the influence of the Early effect is often dominant in small-signal calculations for high-gain circuits and this point will be considered later.

Finally, the influence of the Early effect on the transistor large-signal characteristics in the forward-active region can be represented approximately by modifying Eq. (2-146) to

$$I_C = I_S \left(1 + \frac{V_{BE}}{V_A}\right) \exp \frac{V_{BE}}{V_T} \tag{2-171}$$

This is a common means of representing the device output characteristics for computer simulation.

2-2-5 Saturation and Inverse Active Regions

Saturation is a region of device operation that is usually avoided in analog circuits because the transistor gain is very low in this region. Saturation is much more commonly encountered in digital circuits, where it provides a well-specified output voltage that represents a logic state. In saturation, both emitter–base and collector–base junctions are forward biased. Consequently, the collector–emitter voltage V_{CE} is quite small and usually in the range 0.05–0.3 V. The carrier concentrations in a saturated *npn* transistor with uniform base doping are shown in Figure 2-88. The minority-carrier concentration in the base of the edge of the depletion region is again given by Eq. (2-139) as

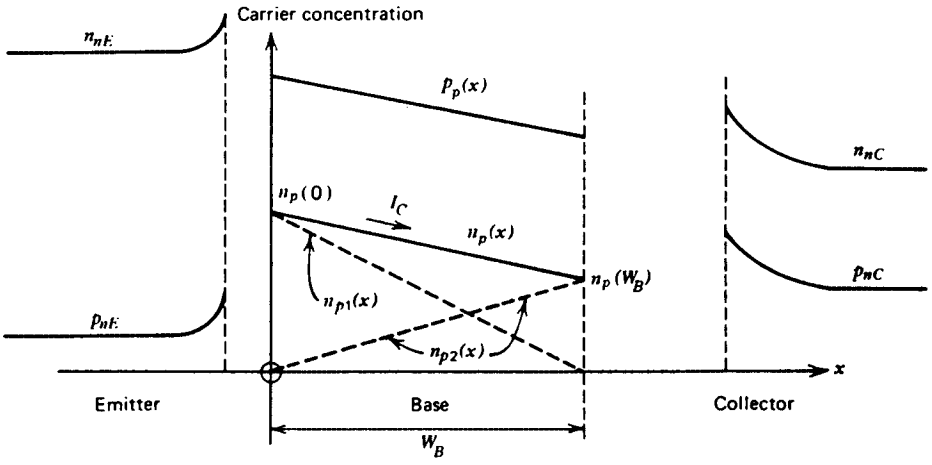


Figure 2-88 Carrier concentrations in a saturated *npn* transistor. (Not to scale)

$$n_p(W_B) = n_{po} \exp \frac{V_{BC}}{V_T} \tag{2-172}$$

but since V_{BC} is now positive, the value of $n_p(W_B)$ is no longer negligible. Consequently, changes in V_{CE} with V_{BE} held constant (which cause equal changes in V_{BC}) directly affect $n_p(W_B)$. Since the collector current is proportional to the slope of the minority-carrier concentration in the base [see Eq. (2-142)], it is also proportional to $[n_p(0) - (n_p(W_B)]$ from Figure 2-88. Thus, changes in $n_p(W_B)$ directly affect the collector current, and the collector node of the transistor appears to have a *low impedance*. As V_{CE} is decreased in saturation with V_{BE} held constant, V_{BC} increases as does $n_p(W_B)$ from Eq. (2-172). Thus, from Figure 2-88 the collector current decreases because the slope of the carrier concentration decreases. This gives rise to the saturation region of the $I_C - V_{CE}$ characteristic shown in Figure 2-89. The slope of the $I_C - V_{CE}$ characteristic in this region is largely determined by the resistance in series with the collector lead due to the finite resistivity of the *n*-type collector material. A useful model for the transistor in this region is shown in Figure 2-90

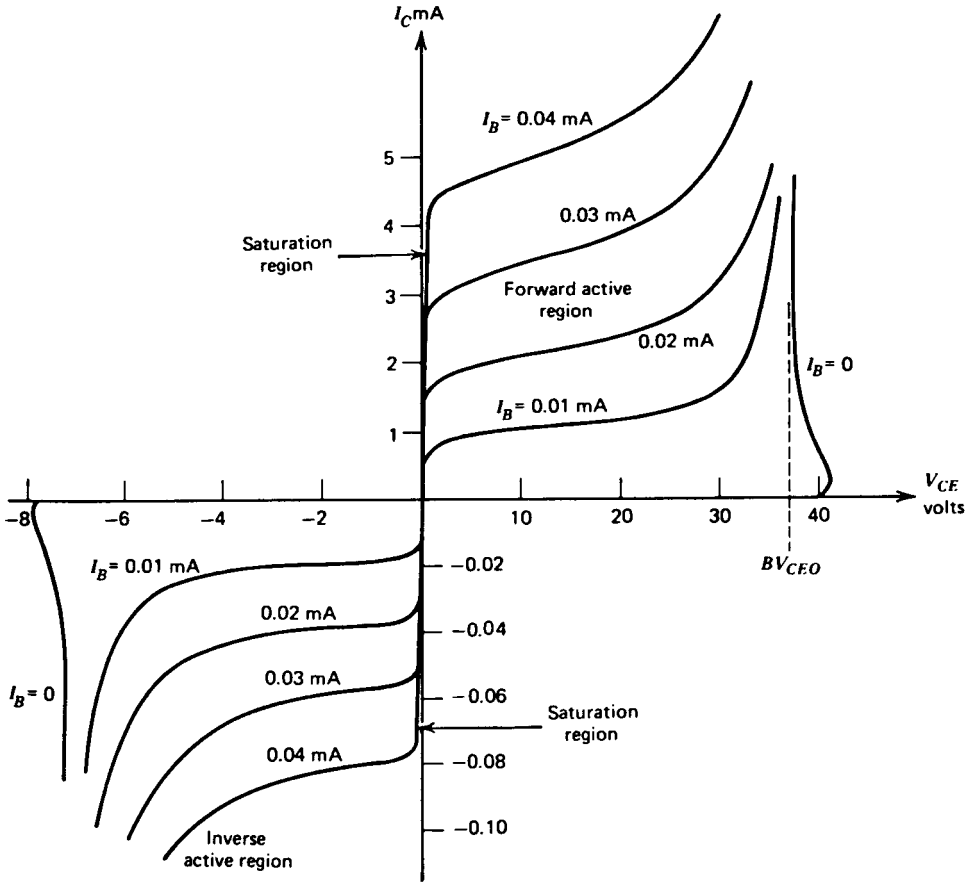


Figure 2-89 Typical $I_C - V_{CE}$ characteristics for an *npn* bipolar transistor. Note the different scales for positive and negative currents and voltages.

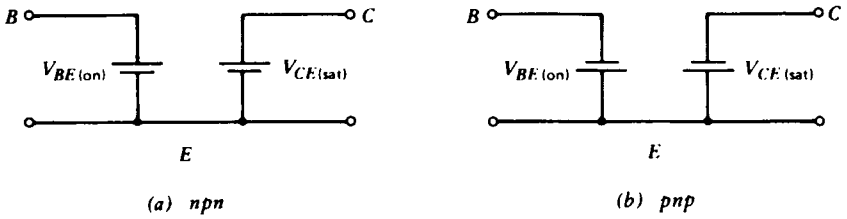


Figure 2-90 Large-signal models for bipolar transistors in the saturation region.

and consists of a fixed voltage source to represent $V_{BE(on)}$, and a fixed voltage source to represent the collector–emitter voltage $V_{CE(sat)}$. A more accurate but more complex model includes a resistor in series with the collector. This resistor can have a value ranging from 20 to 500 Ω, depending on the device structure.

An additional aspect of transistor behavior in the saturation region is apparent from Figure 2-88. For a given collector current, there is now a much larger amount of stored charge in the base than there is in the forward-active region. Thus the base-current contribution represented by Eq. (2-152) will be larger in saturation. In addition, since the collector–base junction is now forward biased, there is a new base-current component due to injection of carriers from the base to the collector. These two effects result in a base current I_B in saturation, which is larger than in the forward-active region for a given collector current I_C. Ratio I_C/I_B in saturation is often referred to as the *forced* β and is always less than β_F. As the forced β is made lower with respect to β_F, the device is said to be more *heavily saturated*.

The minority-carrier concentration in saturation, shown in Figure 2-88, is a straight line joining the two end points, assuming that recombination is small. This can be represented by a linear superposition of the two dotted distributions as shown. The justification for this is that the terminal currents depend *linearly* on the concentrations $n_p(0)$ and $n_p(W_B)$. This picture of device-carrier concentrations can be used to derive some general equations describing transistor behavior. Each of the distributions in Figure 2-88 is considered separately and the two contributions are combined. The *emitter* current that would result from $n_{p1}(x)$ above is given by the classical diode equation

$$I_{EF} = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) \tag{2-173}$$

where I_{ES} is a constant that is often referred to as the "saturation current" of the junction (no connection with the transistor saturation described above). Equation (2-173) predicts that the junction current is given by $I_{EF} \simeq I_{ES}$ with a reverse-bias voltage applied. However, in practice Eq. (2-173) is applicable only in the forward-bias region, since second-order effects dominate under reverse-bias conditions and typically result in a junction current several orders of magnitude larger than I_{ES}. The junction current that flows under reverse-bias conditions is often called the "leakage current" of the junction.

Returning to Figure 2-88, we can describe the *collector* current resulting from $n_{p2}(x)$ alone as

$$I_{CR} = -I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-174}$$

where I_{CS} is a constant. The total collector current I_C is given by I_{CR} plus the fraction of I_{EF} that reaches the collector (allowing for recombination and reverse emitter injection). Thus,

$$I_C = \alpha_F I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) - I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-175}$$

where α_F has been defined previously by Eq. (2-162). Similarly, the total emitter current is composed of I_{EF} plus the fraction of I_{CR} that reaches the emitter with the transistor acting in an inverted mode. Thus,

$$I_E = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + \alpha_R I_{CS}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-176}$$

where α_R is the ratio of emitter to collector current with the transistor operating *inverted* (i.e., with the collector–base junction forward biased and emitting carriers into the base and the emitter–base junction reverse biased and collecting carriers). Typical values of α_R are 0.5–0.8. An inverse current gain β_R is also defined,

$$\beta_R = \frac{\alpha_R}{1 - \alpha_R} \tag{2-177}$$

and has typical values 1–5. This is the current gain of the transistor when operated inverted and is much lower than β_F because the device geometry and doping densities are designed to maximize β_F. The inverse-active region of device operation occurs for V_{CE} negative in an *npn* transistor and is shown in Figure 2-89. In order to display these characteristics adequately in the same figure as the forward-active region, the negative voltage and current scales have been expanded. The inverse-active mode of operation is rarely encountered in analog circuits.

Equations (2-175) and (2-176) describe *npn* transistor operation in the saturation region when V_{BE} and V_{BC} are both positive, and also in the forward-active and inverse-active regions. These equations are from the *Ebers–Moll* equations. In the forward-active region, they degenerate into a form similar to that of Eqs. (2-146), (2-158), and (2-160) derived earlier. This can be shown by putting V_{BE} positive and V_{BC} negative in Eqs. (2-175) and (2-176) to obtain

$$I_C = \alpha_F I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + I_{CS} \tag{2-178}$$

$$I_E = -I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) - \alpha_R I_{CS} \tag{2-179}$$

Equation (2-178) is similar in form to Eq. (2-146) except that leakage currents that were previously neglected have now been included. This minor difference is significant only at high temperatures or very low operating currents. Comparison of Eq. (2-178) with Eq. (2-146) allows us to identify $I_S = \alpha_F I_{ES}$ and it can be shown [6] in general that

$$\alpha_F I_{ES} = \alpha_R I_{CS} = I_S \tag{2-180}$$

where this expression represents a reciprocity condition. Using Eq. (2-180) in Eqs. (2-175) and (2-176) allows the Ebers–Moll equations to be expressed in the general form

$$I_C = I_S\left(\exp\frac{V_{BE}}{V_T} - 1\right) - \frac{I_S}{\alpha_R}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-181}$$

$$I_E = -\frac{I_S}{\alpha_F}\left(\exp\frac{V_{BE}}{V_T} - 1\right) + I_S\left(\exp\frac{V_{BC}}{V_T} - 1\right) \tag{2-182}$$

This form is often used for computer representation of transistor large-signal behavior.

The effect of leakage currents mentioned above can be further illustrated as follows. In the forward-active region, we have, from Eq. (2-179),

$$I_{ES}\left(\exp\frac{V_{BE}}{V_T} - 1\right) = -I_E - \alpha_R I_{CS} \tag{2-183}$$

Substitution of Eq. (2-183) in Eq. (2-178) gives

$$I_C = -\alpha_F I_E + I_{CO} \tag{2-184}$$

where

$$I_{CO} = I_{CS}(1 - \alpha_R\alpha_F) \tag{2-185}$$

and I_{CO} is the collector–base leakage current with the emitter open. Although I_{CO} is given theoretically by Eq. (2-185), in practice, surface leakage effects dominate when the collector–base junction is reverse biased and I_{CO} is typically several orders of magnitude larger than the value given by Eq. (2-185). However, Eq. (2-184) is still valid if the appropriate measured value for I_{CO} is used. Typical values of I_{CO} are 10^{-10} to 10^{-12} A at 25 °C, and the magnitude doubles about every 8 °C. As a consequence, these leakage terms can become very significant at high temperatures. For example, consider the base current I_B. This is

$$I_B = -(I_C + I_E) \tag{2-186}$$

If I_E is calculated from Eq. (2-184) and substituted in Eq. (2-186), the result is

$$I_B = \frac{1 - \alpha_F}{\alpha_F} I_C - \frac{I_{CO}}{\alpha_F} \tag{2-187}$$

But from Eq. (2-161)

$$\beta_f = \frac{\alpha_F}{1 - \alpha_F} \tag{2-188}$$

and use of Eq. (2-188) in Eq. (2-187) gives

$$I_B = \frac{I_C}{\beta_F} - \frac{I_{CO}}{\alpha_F} \tag{2-189}$$

Since the two terms in Eq. (2-189) have opposite signs, the effect of I_{CO} is to *decrease* the magnitude of the external base current at a given value of collector current.

2-2-6 Small-Signal Models of Bipolar Transistors

Figure 2-91 shows the small-signal equivalent BJT model. It consists of an intrinsic model and a package model. Table 2-12 lists its keywords.

In order to better understand the difference between the various technologies, we show Table 2-13, which allows the reader to compare their performances.

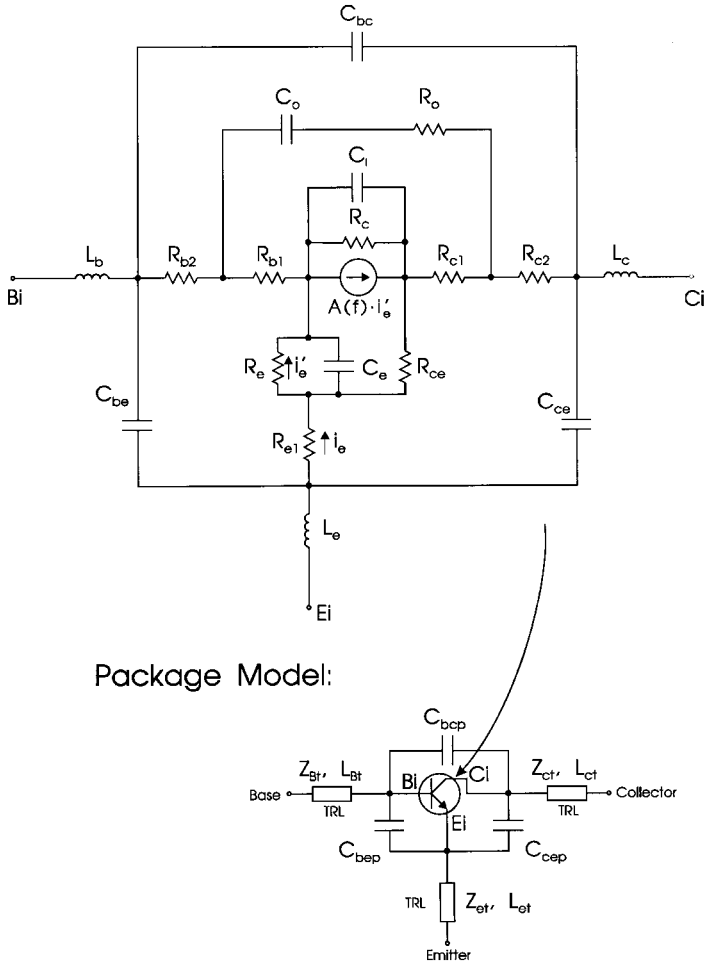


Figure 2-91 Linear BJT model.

Table 2-12 Small-signal equivalent BJT model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
A	Ratio of I_C to I_E at dc		0.0
RE	Emitter resistance	ohm	0.0
F	Current generator roll-off frequency	hertz	∞
T	Time delay	second	0.0
CE	Emitter capacitance	farad	0.0
CI	Collector capacitance	farad	0.0
RCE	Collector–emitter resistance	ohm	∞
RC	Collector resistance	ohm	∞
RO	Extrinsic base–collector resistance	ohm	0.0
CO	Extrinsic base–collector capacitance	farad	0.0
RB1	Intrinsic base resistance (R_{bb})	ohm	0.0
RC1	Parasitic collector resistance	ohm	0.0
RE1	Parasitic emitter resistance	ohm	0.0
RB2	Parasitic base resistance	ohm	0.0
RC2	Parasitic collector resistance	ohm	0.0
CBE	Base-to-emitter package capacitance	farad	0.0
CBC	Base-to-collector package capacitance	farad	0.0
CCE	Collector-to-emitter package capacitance	farad	0.0
LB	Base-lead inductance	henry	0.0
LC	Collector-lead inductance	henry	0.0
LE	Emitter-lead inductance	henry	0.0
TJ	Chip temperature	Kelvin	298
NFAC	Noise factor proportional to drive		1.0
FC	Flicker noise (1/*f* noise) corner frequency	hertz	50
	Package Model		
CBCP	Base-to-collector package capacitance	farad	0.0
CBEP	Base-to-emitter package capacitance	farad	0.0
CCEP	Collector-to-emitter package capacitance	farad	0.0
ZBT	Base transmission line impedance	ohm	50
ZCT	Collector transmission line impedance	ohm	50
ZET	Emitter transmission line impedance	ohm	50
LBT	Base transmission line length at $\varepsilon_r = 1$	meter	0.0
LCT	Collector transmission line length at $\varepsilon_r = 1$	meter	0.0
LET	Emitter transmission line length at $\varepsilon_r = 1$	meter	0.0

Notes:

1. $A \equiv \alpha = I_c/I_e$; β = dc current gain = $\alpha(1 - \alpha)$.
2. The bipolar current gain in this model is described by

$$\mathrm{A} = \mathrm{A}(0) = \frac{e^{-j\omega\mathrm{T}}}{1 + j\,(f/F)}$$

where $\omega = 2\,\pi f$ and f is frequency.

3. The current source is controlled by the current through R_e. The current generator has a cutoff frequency with respect to the total emitter current, I_E:

$$\mathrm{F} = g_m/2\pi C_e$$

where $g_m = 1/R_e$. This frequency becomes infinite for the default value for C_e (0.0). The parammeter F specifies the frequency roll-off for the current generator with respect to the current through R_e. Effectively, this frequency parameter may be used to model additional delays in the device.

Table 2-13 Overview of current BJT technologies including BiCMOS

Parameter	MOSAIC 3	MOSAIC 5	MOSAIC 5SE	RFBICMOS1	TFSOI
Emitter Size	$0.95 \times 3.2\ \mu m^2$	$0.3 \times 0.9\ \mu m^2$	$0.3 \times 0.9\ \mu m^2$	$0.4 \times 1.3\ \mu m^2$	$0.1 \times 1.2\ \mu m^2$
β	150	150	90	80	40–65
F_t (GHz)	14@ $V_{ce} = 1$ V	14@ $V_{ce} = 1$ V	27@ $V_{ce} = 1$ V	16@ $V_{ce} = 1$ V	8@ $V_{ce} = 2$ V
C_{je} (fF)	19.9	4.3	3.8	5.1	1.5
C_{jc} (fF)	25.1	2.8	2.8	3.7	0.34
C_{js} (fF)	49.6	4.3	3.6	7.4	0.95
R_e (Ω)	22	36	50	34	72
R_b (Ω)	462	578	583	462	n/a
BV_{ceo} (V)	6.5	7.0	4.8	6.0	>3.5
BV_{cbo} (V)	20	20	15	18	>10
BV_{ebo} (V)	5.2	4.2	3.8	6.0	4.3
V_A (V)	30	48	20	>25	25–40
Minimum T_{pd}	75 ps @ 1.5 mA	35 ps @ 800 μA	36 ps @ 800 μA	43 ps @ 300 μA	52 ps @ 25 μA

Parameter	TFSOI Ring Device	RFBICMOS3 Analog	RFBICMOS3 Digital	BiGCMOS3 Analog	BiGCMOS3 Digital
Emitter Size	$0.1 \times 7.2\ \mu m^2$	$0.25 \times 0.75\ \mu m^2$	$0.25 \times 0.75\ \mu m^2$	$0.2 \times 0.6\ \mu m^2$	$0.2 \times 0.6\ \mu m^2$
β	40–65	100	150	100	100
F_t (GHz)	8 @ $V_{ce} = 2$ V	15 @ $V_{ce} = 2$ V	25 @ $V_{ce} = 2$ V	15 @ $V_{ce} = 1$ V	35 @ $V_{ce} = 1$ V
C_{je} (fF)	3.4	1.8	1.8	1.7	1.7
C_{jc} (fF)	2.8	1.1	1.1	0.2	0.3
C_{js} (fF)	2.4	12.0	12.0	4	4
R_e (Ω)	26	140	140	113	113
R_b (Ω)	790	n/a	n/a	10 kΩ/sq.	15 kΩ/sq.
BV_{ceo} (V)	>3.5	6.5	4	6.5	4.8
BV_{cbo} (V)	>10	15	15	20	15
BV_{ebo} (V)	4.3	4	4	4	4
V_A (V)	30–40	40	25	40	20
Minimum T_{pd}	75 ps @ 100 μA	n/a	n/a	35 ps @ 50 μA	22 ps @ 200 μA

Parameter	AT&T BEST2[a]	Fujitsu POSET	HP HP-25	HP ISOSAT II[b]	Hitachi SMI
Emitter Size	$1.0 \times 2.0\ \mu m^2$	$0.1 \times 1.7\ \mu m^2$	$0.4 \times 20\ \mu m^2$	$0.6 \times ?\ \mu m^2$	$0.2 \times 2.0\ \mu m^2$
β	100	110	75	100	500
F_t (GHz)	23 @ $V_{ce} = 3$ V	38 @ $V_{ce} = 3$ V	28.5 @ $V_{ce} = 3$ V	14/F_{max} > 20 GHz	82@ $V_{ce} = 2$ V
C_{je} (fF)	5.0	3.5	42	n/a	2.9
C_{jc} (fF)	5.0	2.2	37	n/a	4.8
C_{js} (fF)	8.0	n/a	123	n/a	4.1
R_e (Ω)	25	25	n/a	n/a	30
R_b (Ω)	10 kΩ/sq.	960	100	n/a	240
BV_{ceo} (V)	5.5	5.4	5.2	15	2.5
BV_{cbo} (V)	11	18.2	15	25	6.4
BV_{ebo} (V)	8.0	4.5	>5	2.0	6.2
V_A (V)	24	n/a	n/a	25	8

(continued)

Table 2-13 (Continued)

Parameter	AT&T BEST2[a]	Fujitsu POSET	HP HP-25	HP ISOSAT II[b]	Hitachi SMI
Minimum T_{pd}	37 ps @ 2.1 mA 48 ps @ 600 µA	21.5 ps @ 320 µA	n/a	n/a	16 ps @ $I_g = 1.3$

Parameter	IBM SiGe[c]	Maxim GST-1[d]	Maxim GST-2[e]	Mitsubishi[f]	National ABIC V
Emitter Size	0.5×1.0 µm^2	1.2×2.6 µm^2	0.7×1.5 µm^2	0.5×5.7 µm^2	0.5×0.5 µm^2
β	100	75	190	90	85
F_t (GHz)	41 @ $V_{ce} = 1$ V	12.5	26 @ $V_{ce} = 4$ V	20	13 @ $V_{ce} = 2$ V
C_{je} (fF)	5.2	9	4	n/a	1.9
C_{jc} (fF)	3.4	15	7	9	3.7
C_{js} (fF)	3.6	11	9	n/a	9.6
R_e (Ω)	n/a	n/a	80	n/a	n/a
R_b (Ω)	300	600	710	155	383
BV_{ceo} (V)	3.3	>5	>4	6.2	8.5
BV_{cbo} (V)	9.5	>12	>14	n/a	20
BV_{ebo} (V)	n/a	>4	>4	n/a	7.5
V_A (V)	n/a	20	15	n/a	n/a
Minimum T_{pd}	n/a	n/a	n/a	n/a	38 ps @ 1.8 mA

Parameter	NEC BSA	NEC SiGe[g]	NEC NESAT III	Philips QBiC2	Philips PRET[g]
Emitter	0.4×7.6 µm^2	0.2×1.6 µm^2	0.6×40 µm^2	0.5×1.2 µm^2	0.2×0.9 µm^2
β	n/a	120	80	100	50
F_t (GHz)	43	51	20 @ $V_{ce} = 3$ V	17	14
C_{je} (fF)	20	n/a	120	2.4	1.5
C_{jc} (fF)	16	1.7	140	3.5	1.0
C_{js} (fF)	n/a	n/a	160	6.5	5.0
R_e (Ω)	n/a	80	n/a	120	120
R_b (Ω)	43	220	55	700 Ω/sq.	1 kΩ/sq.
BV_{ceo} (V)	3.3	2.7	6	4.5	7
BV_{cbo} (V)	n/a	n/a	17	13	12.5
BV_{ebo} (V)	2	2.2	4.5	5.5	5.5
V_A (V)	n/a	n/a	n/a	15	20
Minimum T_{pd}	29 ps @ 300 µA	19 ps @ 300 µA	n/a	n/a	n/a

Parameter	Philips SiGe[g]	Siemens B6HF	Siemens B6HF Plus	Siemens SiGe[g]	SONY[h]
Emitter Size	0.4×1.4 µm^2	0.8×3.3 (drawn) µm^2	0.25×2.7 µm^2	0.27×2.5 µm^2	$0.2 \times ?$ µm^2
β	100	>60	115	130	100
F_t (GHz)	45 @ $V_{ce} = 1.5$ V	25 @ $V_{cb} = 0$ V	50 @ $V_{cb} = 2$ V	61 @ $V_{cb} = 1$ V	30
C_{je} (fF)	9.1	11	8.5	7.5	n/a
C_{jc} (fF)	4.6	6.5	8.0	5.5	n/a
C_{js} (fF)	10.8	31	26.5	13	n/a

(continued)

Table 2-13 Overview of current BJT technologies including BiCMOS (Continued)

Parameter	Phillips SiGe[g]	Siemens B6HF	Siemens B6HF Plus	Siemens SiGe[g]	SONY[h]
R_e (Ω)	n/a	n/a	n/a	n/a	n/a
R_b (Ω)	n/a	165	14 kΩ/sq.	8 kΩ/sq.	n/a
BV_{ceo} (V)	3.2	>3.5	7	2.5	4.0
BV_{cbo} (V)	11	n/a	10.8	9.9	18
BV_{ebo} (V)	3.5	n/a	2.0	3.0	3.5
V_A (V)	70–100	n/a	20	130	10
Minimum T_{pd}	14 ps @ 6 mA	25 ps @ 1.2 mA (CML)	16 ps @ 2.0 mA (CML)	11 ps @ 1.9 mA (CML)	21 ps @ I_g = 3.5 mA (CML)

Parameter	STM HSB[i]	STM HSB2A	iSTM HSB3[g,j]	Toshiba	Toshiba Epi Base[g]
Emitter Size	$0.35 \times ? \ \mu m^2$	$0.35 \times ? \ \mu m^2$	$0.2 \times ? \ \mu m^2$	$0.4 \times 5.0 \ \mu m^2$	$0.6 \times 5.0 \ \mu m^2$
β	>50	>50	>50	60	105
F_t (GHz)	n/a	20	~50	22 @ V_{ce} = 2 V	52 @ V_{ce} = 3 V
C_{je} (fF)	n/a	n/a	n/a	9.0	16.1
C_{jc} (fF)	n/a	n/a	n/a	3.6	14.1
C_{js} (fF)	n/a	n/a	n/a	11	n/a
R_e (Ω)	n/a	n/a	n/a	14	22.7
R_b (Ω)	10 kΩ/sq.	n/a	n/a	90	205
BV_{ceo} (V)	4.5	n/a	n/a	6.2	4.1
BV_{cbo} (V)	n/a	n/a	n/a	14.8	9.5
BV_{ebo} (V)	n/a	n/a	n/a	4.2	n/a
V_A (V)	n/a	>30	n/a	n/a	26.5
Minimum T_{pd}	n/a	30 ps	n/a	19 ps @ 1.2 mA	n/a

[a] 1.0 µm Design Rule; NF = 3.25 dB, G_A = 12 dB at I_C = 2 mA, freq. = 2 GHz for $16 \times 1 \times 2 \ \mu m^2$; Resistors: 150 Ω/sq. (mono), 300 Ω/sq. (mono), 500 Ω/sq. (poly); Layers of metal, Schottky diodes, linear capacitor (silicided n^+ poly/n^+Si); Gate length = 0.68 µm (*n*-ch)/0.8 µm (*p*-ch), Gate Oxide = 12.5 nm (3 V)/15 nm (5 V), V_{th} = 0.65 V (*n*-ch)/–0.9 V (*p*-ch), Gm (µS/µm) = 204 (*n*-ch)/65 (*p*-ch).

[b] 0.6 µm Design Rule; Gold metallization, ILD: polyimide; Inductors: $Q > 10$ for $L \leq 10$ nH; Capacitors: $Q > 25$ for $C \leq 20$ pF; Schottky diodes (f_c = 100 GHz), Zener diodes, MIS capacitor, thin film resistors, and LPNP, 10 mask process.

[c] Analog devices using this process.

[d] LPNP F_t = 80 MHz; three-layer metallization; Schottky diodes (R_sC_j ~ 4 ps); Implanted resistors: 200 Ω/sq.; Poly resistors: 400/2000 Ω/sq.; NiCr thin film resistors (50 Ω/sq.) with laser trim option; MOS capacitor.

[e] LPNP F_t = 85 MHz; two-layer metallization; Schottky diodes (R_sC_j ~ 0.8 ps); Implanted resistors: 170/600/1200 Ω/sq.; NiCr thin film resistors (50 Ω/sq.) with laser trim option; MOS capacitor.

[f] L_n = 0.8 µm, T_{ox} = 18 nm, V_{th} = 0.76 V (nmos), V_{th} = –0.78 V (pmos), 2poly-MIM capacitor: 2.7 fF/µm^2, Schottkys, inductors.

[g] Not known if this process is in production.

[h] 0.8 µm Design Rules; 2.5 fF/µm^2 MIS capacitor; 600 MHz PJFET; 12L; F_t = 4 GHz VPNP; poly resistors 80/400/2000 Ω/sq.

[i] Targeted for 900-MHz GSM radio front-end blocks, a CT2 cordless single-chip radio, and DECT (digital enhanced cordless telephones).

[j] Targeted for multimode 900-MHz GSM, 1.8-GHz PCN and 1.9-GHz PCS based systems. Company is also evaluating integrating SiGe in HSB3.

2-3 FIELD-EFFECT TRANSISTORS

Both diodes and *pnp/npn* transistors work on the injection principle, which means the base–emitter junction generates either free electrons or other carriers, which are being "collected" by the collector connection and the base connection controls it. While this is a fairly complex process, its physics is well understood, and its latest spin-offs, heterojunction BJTs and silicon–germanium transistors, seem to offer a bright future.

In opposition to this, the field-effect transistor (FET) controls the electronic conduction in a solid by an electric field, and this concept actually predates the invention of the bipolar transistor. J. E. Lilienfeld filed for a patent on such a device in 1925 (U.S. Patent No. 1,745,175), and W. Shockley presented a comprehensive theory of the FET in 1952. It took until 1960 before the first commercially available FETs came on the market.

Several types of semiconductor materials have been used for making FETs: silicon, germanium, gallium arsenide, and others have been used. By far the most widely used is silicon, followed now by gallium arsenide.

The FET is a class of electronic semiconductor device in which the conduction of a "channel" between source (*S*) and drain (*D*) terminals is controlled by an electric field impressed on the channel via a gate (*G*) terminal. The conduction channel may utilize *n*-type carriers (electrons) or *p*-type carriers (holes). The electric field that controls the channel conduction may be introduced via a *pn* junction [for a "junction" FET (JFET)], a metal plate separated from the semiconductor channel by an oxide dielectric [for a metal-oxide semiconductor (MOS) FET], or a combination of the two methods. The polarity of the controlling electrical field is a function of the type of carriers in the channel. A FET "family tree" is shown in Figure 2-92.

For today's microwave and RF integrated circuits, the JFET with its high-frequency problems (low cutoff frequency) and large variation in parameters, such as transconductance and pinchoff voltage, has fallen out of favor and is mainly used in discrete circuits such as two-way radios or other RF applications that tolerate these wide variations and temperature dependence. The most-used FET types are MOS for small-signal applications, LDMOS (lateral diffused MOS) for power applications, and GaAs MESFET for lower-power applications (low-voltage applications) as well as high-power applications (recent advertisements by NEC have shown output power up to 100 W). Additional information on the FET family is shown in Figure 2-93.

All members of the FET family are high-impedance-input devices. Figure 2-94 shows the key operating parameters for MOSFETs and MESFETs in low-noise stages, high-gain stages, and linearized power stages. The linearization also improves the matching. This figure is similar to its counterpart in Section 2-2-2.

The following several pages present a typical example of a datasheet* for a JFET (U310) and the chip (NZA) on which it is based, followed by an example of a device with a much higher dynamic range (the CP640 family). The CP640 family and devices similar to them are being replaced by low-power LDMOS, which are not yet well-documented. At least one company, InterFET, is devoted entirely to producing discrete JFETs and JFET-related hybrid and small-scale-integration IC products.

* Reproduced with permission of Vishay-Siliconix.

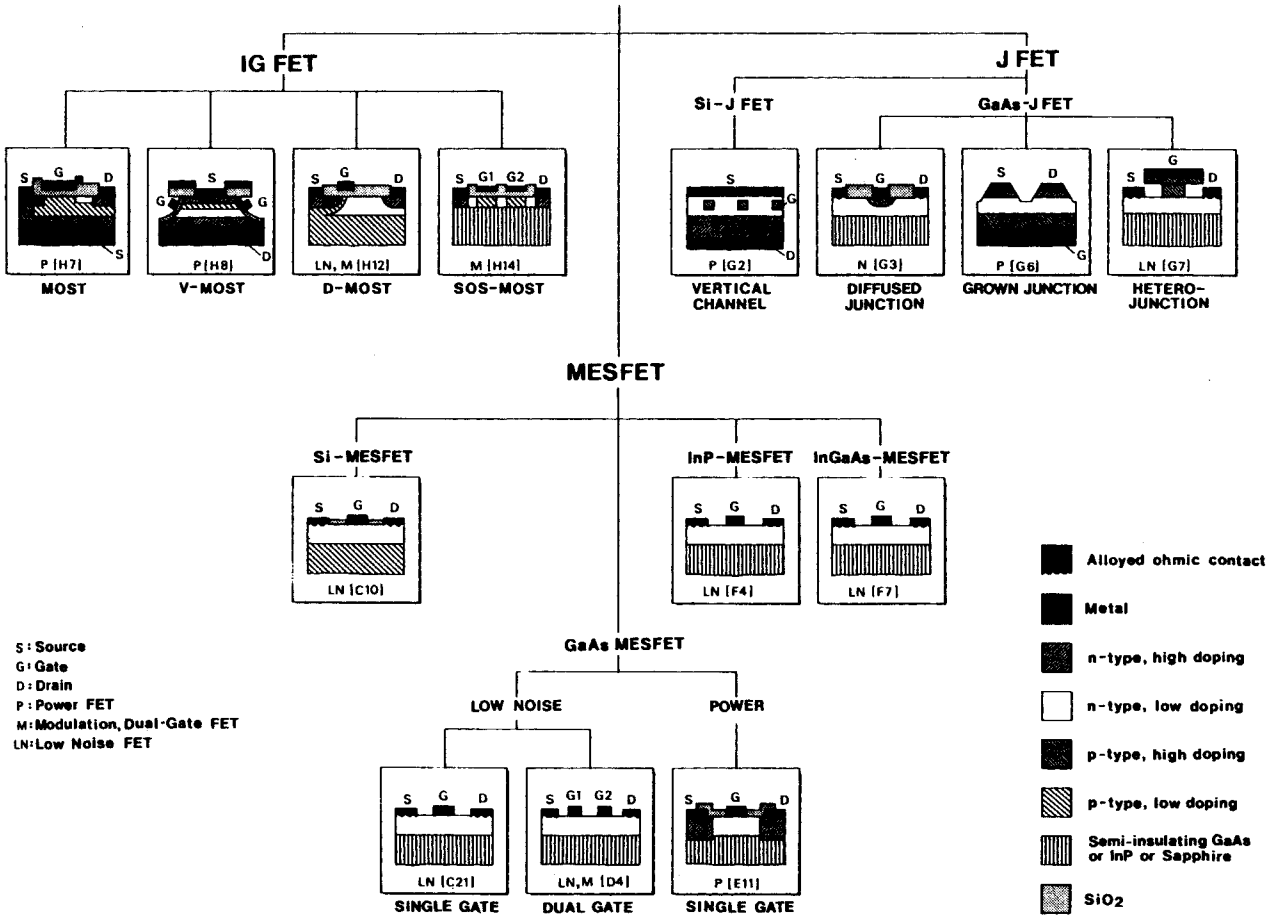


Figure 2-92 "Family tree" of microwave FET types showing cross sections of their structures [7]. Copyright ©1976 IEEE.

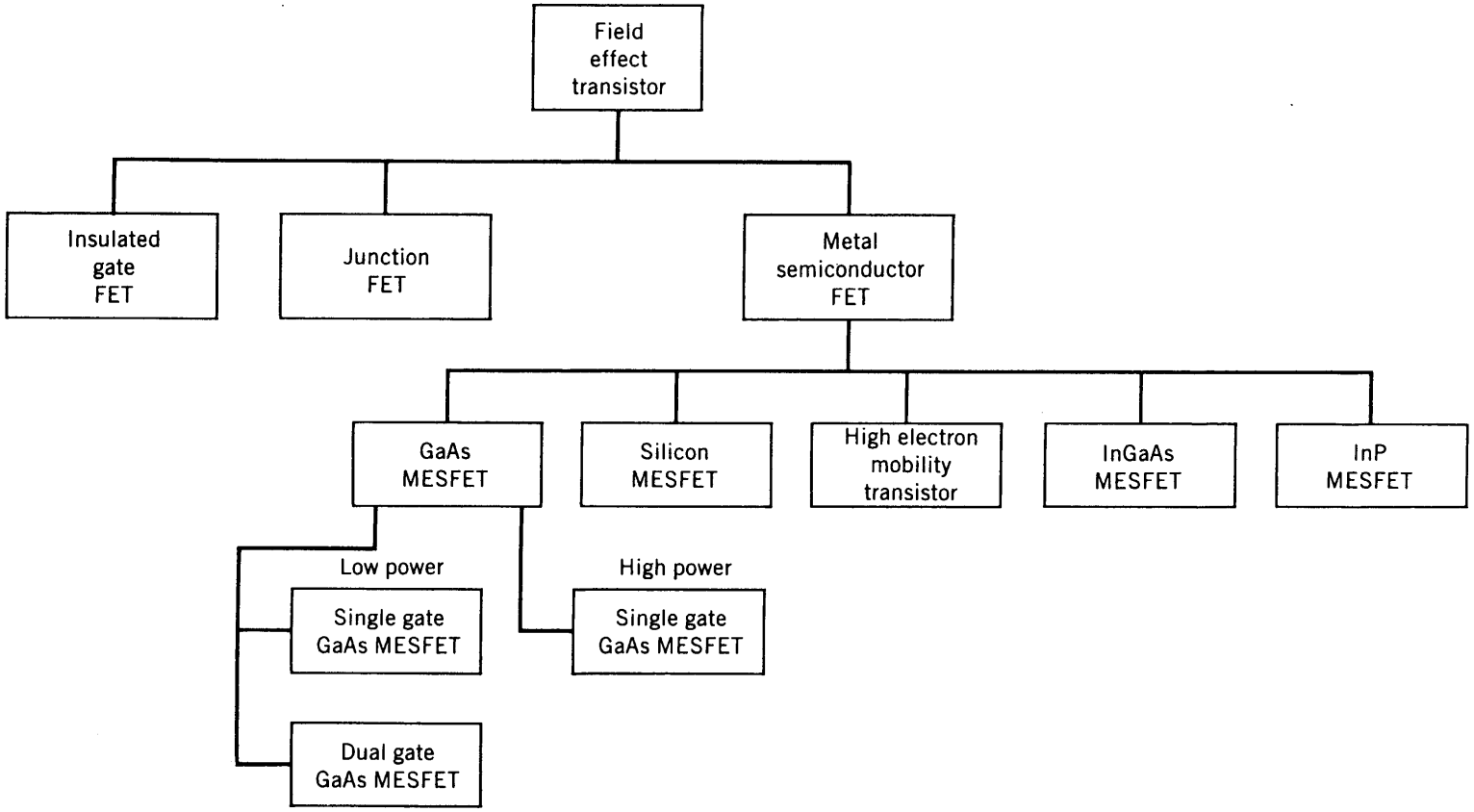


Figure 2-93 Another FET family tree [8].

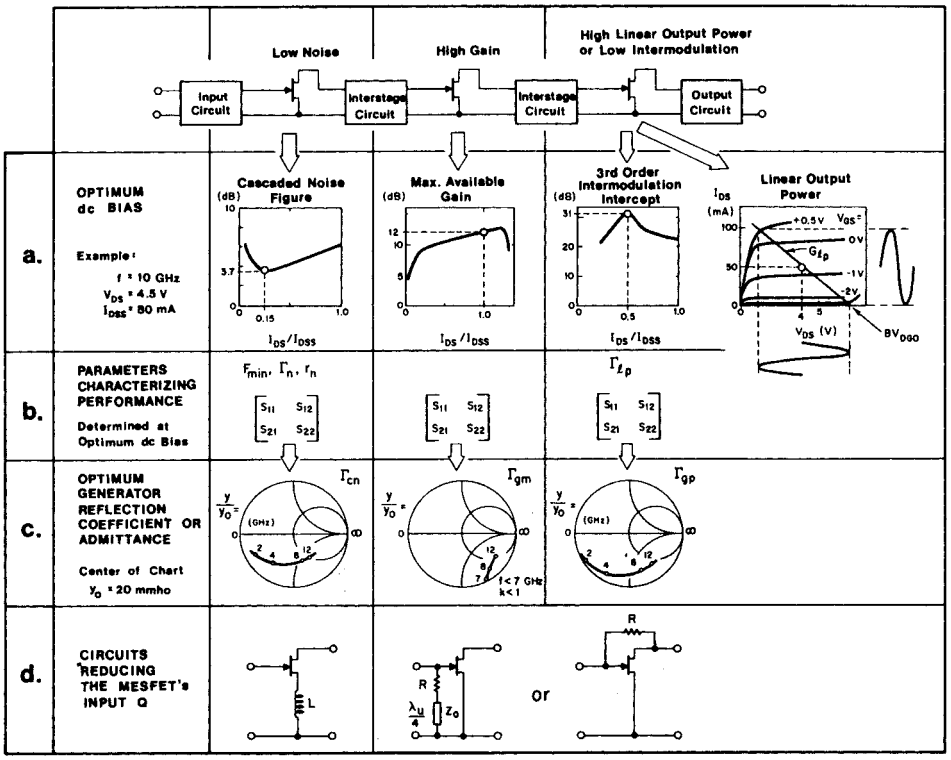


Figure 2-94 Key parameters in operating a MESFET in a low-noise front-end stage, a high-gain intermediate stage, and a linear power stage, showing (a) the device characteristics that lead to optimum dc bias for each service, (b) the parameters that characterize the device performance, (c) the optimum generator reflection coefficient that must be synthesized by each circuit, and (d) circuit arrangements that reduce the transistor's input *Q* or stabilize it at low frequencies. Although this example is based on an HP MESFET with a gate length of 1 μm and a gate width of 500 μm, this principle applies to all members of the FET family [7]. Copyright ©1976 IEEE.

n-channel JFETs *designed for . . .*

Siliconix

Performance Curves NZA See Section 4

- **VHF Amplifiers**
- **Front End High Sensitivity Amplifiers**
- **Oscillators**
- **Mixers**

BENEFITS

- Industry Standard
- High Power Gain
  16 dB at 105 MHz, Common-Gate
  11 dB at 450 MHz, Common-Gate
- Low Noise
  2.7 dB Noise Figure at 450 MHz
- Wide Dynamic Range
  Greater than 100 dB
- 75 Ω Input Match Common Gate

ABSOLUTE MAXIMUM RATINGS (25°C)

Gate-Drain or Gate-Source Voltage . . . . . -25 V
Gate Current . . . . . 20 mA
Total Power Dissipation at T_A = 25°C . . . . . 500 mW
Power Derating to 150°C . . . . . 4.0 mW/°C
Storage Temperature Range . . . . . -65 to +200°C
Lead Temperature
(1/16" from case for 10 seconds) . . . . . 300°C

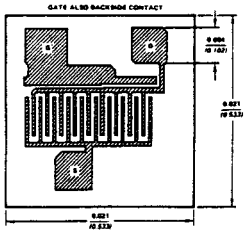
TO-52
See Section 6

ELECTRICAL CHARACTERISTICS (25°C unless otherwise noted)

		Characteristic		U308 Min	U308 Typ	U308 Max	U309 Min	U309 Typ	U309 Max	U310 Min	U310 Typ	U310 Max	Unit	Test Conditions	
1	STATIC	I_{GSS}	Gate Reverse Current			-150			-150			-150	pA	V_{GS} = -15 V,	
2						-150			-150			-150	nA	V_{GS} = 0	T_A = 125°C
3		BV_{GSS}	Gate-Source Breakdown Voltage	-25			-25			-25			V	I_G = -1 μA, V_{DS} = 0	
4		$V_{GS(off)}$	Gate-Source Cutoff Voltage	-1.0		-6.0	-1.0		-4.0	-2.5		-6.0		V_{DS} = 10 V, I_D = 1 nA	
5		I_{DSS}	Saturation Drain Current (Note 1)	12		60	12		30	24		60	mA	V_{DS} = 10 V, V_{GS} = 0	
6		$V_{GS(f)}$	Gate-Source Forward Voltage			1.0			1.0			1.0	V	I_G = 10 mA, V_{DS} = 0	
7	DYNAMIC	g_{fg}	Common-Gate Forward Transconductance (Note 1)	10	17		10	17		10	17		mmho	V_{DS} = 10 V, I_D = 10 mA	f = 1 kHz
8		g_{og}	Common-Gate Output Conductance			250			250			250	μmho		
9		C_{gd}	Drain-Gate Capacitance			2.5			2.5			2.5	pF	V_{GS} = -10 V, V_{DS} = 10 V	f = 1 MHz
10		C_{gs}	Gate-Source Capacitance			5.0			5.0			5.0			
11		$\overline{e_n}$	Equivalent Short Circuit Input Noise Voltage		10			10			10		$\frac{nV}{\sqrt{Hz}}$	V_{DS} = 10 V, I_D = 10 mA	f = 100 Hz
12	HI FREQ	g_{fg}	Common-Gate Forward Transconductance		15			15			15		mmho	V_{DS} = 10 V, I_D = 10 mA	f = 105 MHz
13					14			14			14				f = 450 MHz
14		g_{og}	Common-Gate Output Conductance		0.18			0.18			0.18				f = 105 MHz
15					0.32			0.32			0.32				f = 450 MHz
16		G_{pg}	Common-Gate Power Gain (Note 2)	14	16		14	16		14	16		dB		f = 105 MHz
17				10	11		10	11		10	11				f = 450 MHz
18		NF	Noise Figure		1.5	2.0		1.5	2.0		1.5	2.0			f = 105 MHz
19					2.7	3.5		2.7	3.5		2.7	3.5			f = 450 MHz

NOTES:
1. Pulse test duration = 2 ms.
2. Gain (G_{pg}) measured at optimum input noise match.

NZA



n-channel JFET
designed for . . .

- VHF/UHF Amplifiers
- Front End High Sensitivity Amplifiers
- Oscillators
- Mixers

BENEFITS

- Industry Standard
- High Power Gain
  16 dB at 100 MHz, Common Gate
  11 dB at 450 MHz, Common Gate

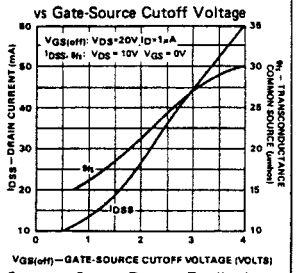
Siliconix

BENEFITS

- Low Noise
  3 dB Noise Figure at 450 MHz
- Wide Dynamic Range
  Greater Than 100 dB
- 75 Ohm Input Match Common Gate

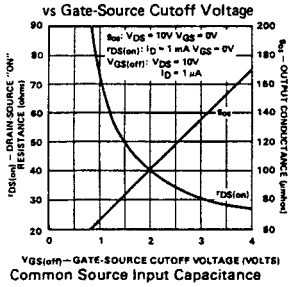
TYPE	PACKAGE	PRINCIPAL DEVICES
Single	TO-52	U308-10
Single	TO-72	U311
Single	TO-92	J308-10
Dual	TO-99	U430-1
Single	Chip	J308CHP-10CHP, U308CHP-10CHP, U311CHP
Dual	Chip	U430CHP-1CHP

Drain Current & Transconductance vs Gate-Source Cutoff Voltage

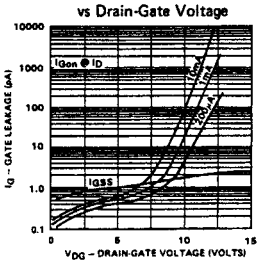


PERFORMANCE CURVES (25°C unless otherwise noted)

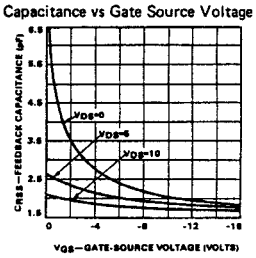
On Resistance & Output Conductance vs Gate-Source Cutoff Voltage



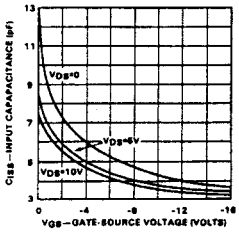
Gate Operating Current vs Drain-Gate Voltage



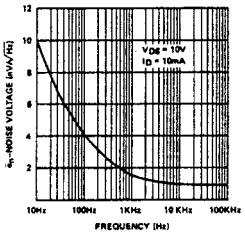
Common Source Reverse Feedback Capacitance vs Gate Source Voltage



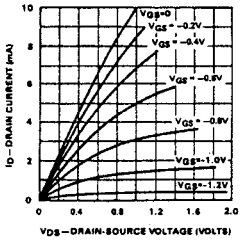
Common Source Input Capacitance vs Gate-Source Voltage



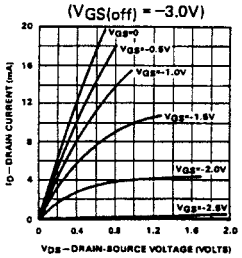
Noise Voltage vs Frequency



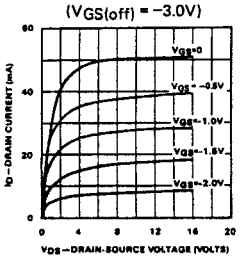
Output Characteristic ($V_{GS(off)} = -1.7V$)



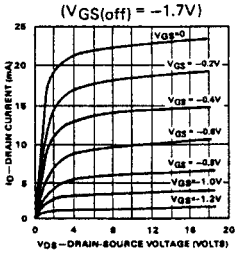
Output Characteristic ($V_{GS(off)} = -3.0V$)



Output Characteristic ($V_{GS(off)} = -3.0V$)



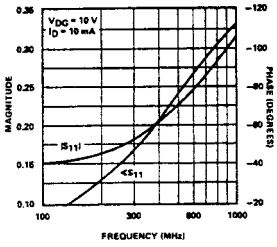
Output Characteristic ($V_{GS(off)} = -1.7V$)



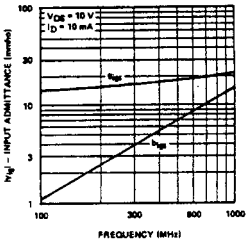
NZA

PERFORMANCE CURVES (Con't) (25°C unless otherwise noted)

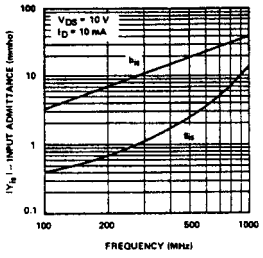
Forward Reflection Coefficient Common Gate



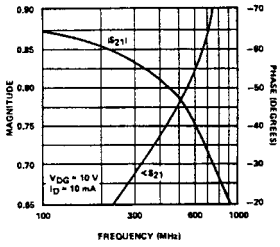
Input Admittance Common Gate



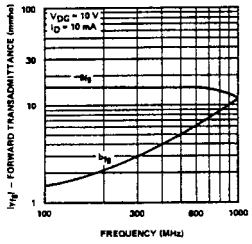
Input Admittance Common Source



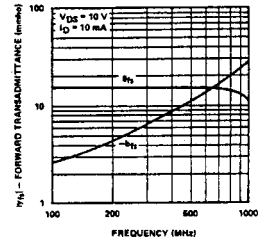
Forward Transmission Coefficient Common Gate



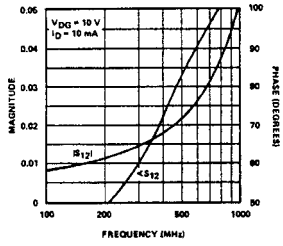
Forward Transfer Admittance Common Gate



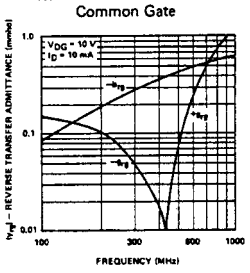
Forward Transfer Admittance Common Source



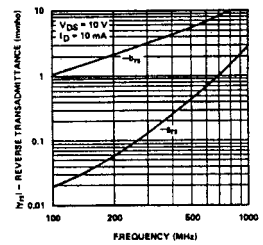
Reverse Transmission Coefficient Common Gate



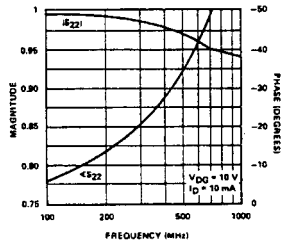
Reverse Transfer Admittance Common Gate



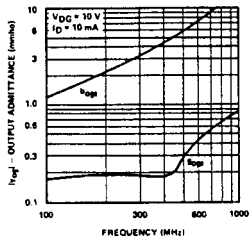
Reverse Transfer Admittance Common Source



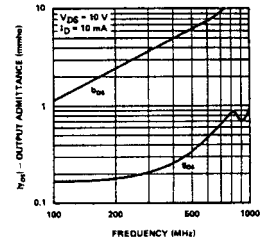
Reverse Reflection Coefficient Common Gate



Output Admittance Common Gate



Output Admittance Common Source



BROADBAND RF FET
SILICON EPITAXIAL JUNCTION N-CHANNEL FIELD EFFECT TRANSISTOR

HIGH DYNAMIC RANGE HF AND VHF AMPLIFIER
FOR USE IN COMMON GATE CONFIGURATION

- USABLE TO OVER 300 MHz
- 50 Ohm VSWR < 1.5:1 0.5-50 MHz (FIG. 1)
- LOW NOISE FIGURE — 2.2 dB TYPICAL @ 50 MHz
- INPUT Z CONSTANT 0.5-50 MHz
- HIGH IM INTERCEPT POINT — > + 40 dBm
- HIGH TRANSCONDUCTANCE — 100,000 μmhos (TYP.)
- 1 dB COMPRESSION POINT > + 20 dBm
- DYNAMIC RANGE > 140 dB (TO 1 dB COMPRESSION)
- HIGH VOLTAGE—TO 50 V.

CP640
CP664
CP665
CP666

ELECTRICAL DATA ABSOLUTE MAXIMUM RATINGS

PARAMETER	SYMBOL	CP 640	CP 664	CP 665	CP 666	UNITS
Drain to Source Voltage	BVDSO	20	30	40	50	Volts
Drain to Gate Voltage	BVDGO	20	30	40	50	Volts
Gate to Source Voltage	BVGSO	−15	−20	−20	−20	Volts
Peak Drain Current	ID	1.2	1.2	1.2	1.2	Amps
Power Dissipation 25 °C CASE	PD	8.0	8.0	8.0	8.0	Watts
Derating Factor (slope)	DF	22	22	22	22	°C/W
Junction Temp.(Oper. & Store)	TJ	−55 °C to +200 °C				

TYPICAL TWO TONE 3rd ORDER IM PRODUCTS — CIRCUIT FIGURE 1

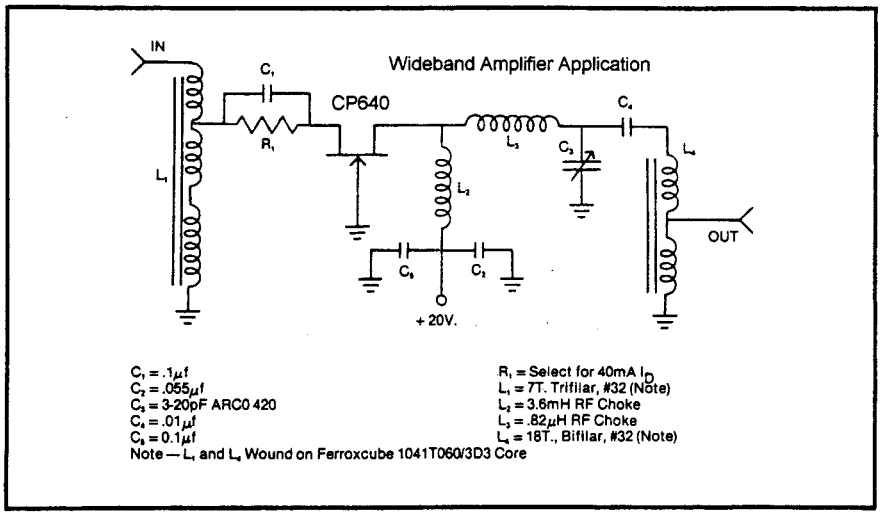
Tones at 3MHz/5MHz

Signal Level EMF	3rd Order Product
1 Volt	-44 dB
0.3 Volt	-75 dB
0.25 Volt (0dBM)	-80 dB

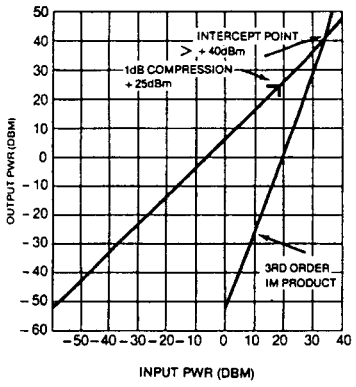
ELECTRICAL CHARACTERISTICS: T_{CASE} = 25 °C (UNLESS OTHERWISE STATED)

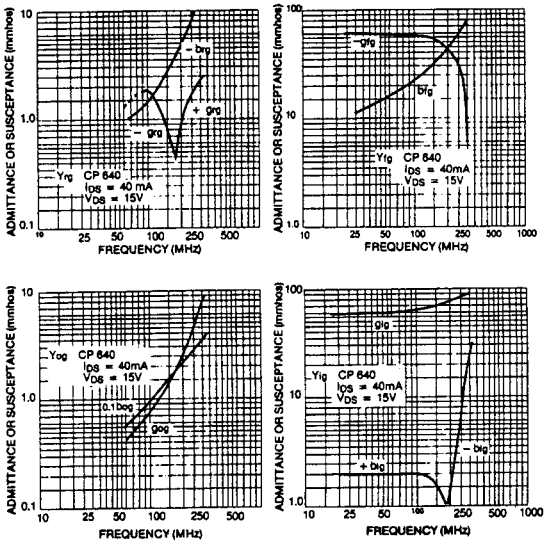
PARAMETERS	CONDITIONS		SYMBOL	Min.	Typ.	Max.	UNITS
Gate Leakage Current	V_{GS} = 15V, V_{DS} = 0	25 °C	I_{GSS}		5	100	nA
		150 °C	I_{GSS}			10	μA
Operating Transconductance	V_{DS} = 15V, I_{DS} = 40 mA		g_{fg}	40	60	80	mmho
Zero Bias Transconductance	V_{DS} = 15V, V_{GS} = 0(1)		g_{fg}	75	100	200	mmho
Gate-Source Cut-Off Voltage	V_{DS} = 5V, I_{DS} = 1.0 mA		$V_{GS(off)}$	2	5	10	Volts
Zero Bias Drain Current	V_{DS} = 15V, V_{GS} = 0 (1)		I_{DSS}	100	400	800	mA
Gate to Source Cap.	V_{GS} = −20V		C_{GS}		15	20	pf
Gate to Drain Cap.	V_{GD} = −20V		C_{GD}		15	20	pf
Power Gain	I_{DS} = 40mA, f = 50MHz, Fig. 1		Gpg	8	8.5	9.5	dB
Noise Figure	I_{DS} = 40mA, f = 30MHz, Fig. 1		N.F.		2.2	3.0	dB
Voltage Standing Wave Ratio	f = 0.5-50MHz, 50 Ω Source, Fig. 1		VSWR			1.5:1	
Common Gate Input Conductance	f = 0.5-50MHz, V_{DS} = 15, I_D = 40mA		g_{igs}		60		mmho
Common Gate Output Conductance	f = −50MHz, V_{DS} = 15, I_D = 40mA		g_{ogs}		0.4		mmho

(1)Pulse Measurement 1% Duty Cycle 10 mS Max.



TYPICAL INTERCEPT AND COMPRESSION POINT





2-3-1 Large-Signal Behavior of JFETs

A section view of a junction field-effect transistor (JFET) is shown in Figure 2-95. This structure contains, between the source and the drain contacts, an *n*-type "channel" embedded in a *p*-type silicon substrate. If it is assumed that the *pn* junction forms a barrier to current flow, then it can be seen that channel conduction is a function of the channel width, length, and thickness and of the density and mobility of the carriers. In this structure, current can flow equally well in either direction through the channel; that is, the drain can be positive or negative with respect to the source.

Since we have found that the SPICE models for JFETs are really incomplete and give poor answers, we will only briefly touch here on the JFET's performance. The modified Materka model, one of the best GaAsFET models, turns out to be best suited for JFETs because both transistors act similarly in the dc area. Both have a gate-to-source diode that becomes conductive above 0.7 V at the input and they are also quite similar in other respects. Their equivalent circuits are also quite similar, as parameter extractions for both approaches have shown.

In general, in a three-terminal device, the drain current I_D is a function of two variables: V_{DS} and V_{GS}. This function is best represented by families of characteristic curves, as shown in Figure 2-96. These curves are for the "common source" configuration, with the drain as the output and the gate as the input. They reveal that for this device, if V_{DS} is greater than about 2 V (but less than the drain breakdown), I_D is primarily determined by the gate voltage V_{GS}. Under these conditions it is valid for small signals to characterize the FET by a single transfer characteristic curve, commonly called the "forward transconductance curve," such as the upper curve shown in Figure 2-96b. Some of the relationships between the forward transfer curve and the output characteristic curves of Figure 2-96a will be examined. The value of V_{GS} that reduces I_D to approximately zero is the gate–source cutoff voltage, $V_{GS(\text{off})}$. With reference to the output curve, Figure 2-96a, note that the drain current at $V_{GS} = 0$ tends to become saturated at a drain voltage approximately equal in magnitude to $-V_{GS(\text{off})}$. This drain voltage is often referred to as the pinchoff voltage V_P; however, in this section pinchoff voltage is used interchangeably with gate–source cutoff voltage. V_P will have the same meaning as $V_{GS(\text{off})}$. The symbol I_{DSS} is commonly used to indicate the value of saturated drain current at $V_{GS} = 0$.

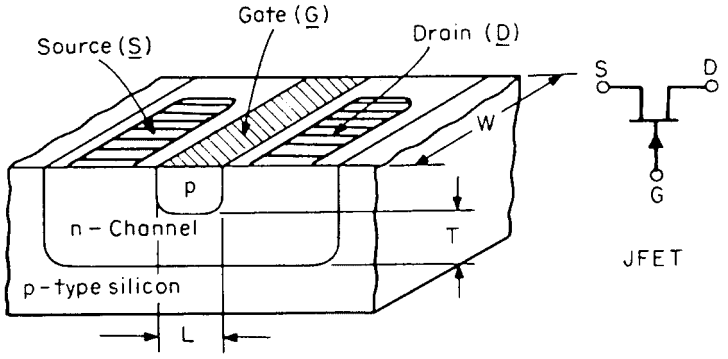


Figure 2-95 Junction field-effect transistor [9].

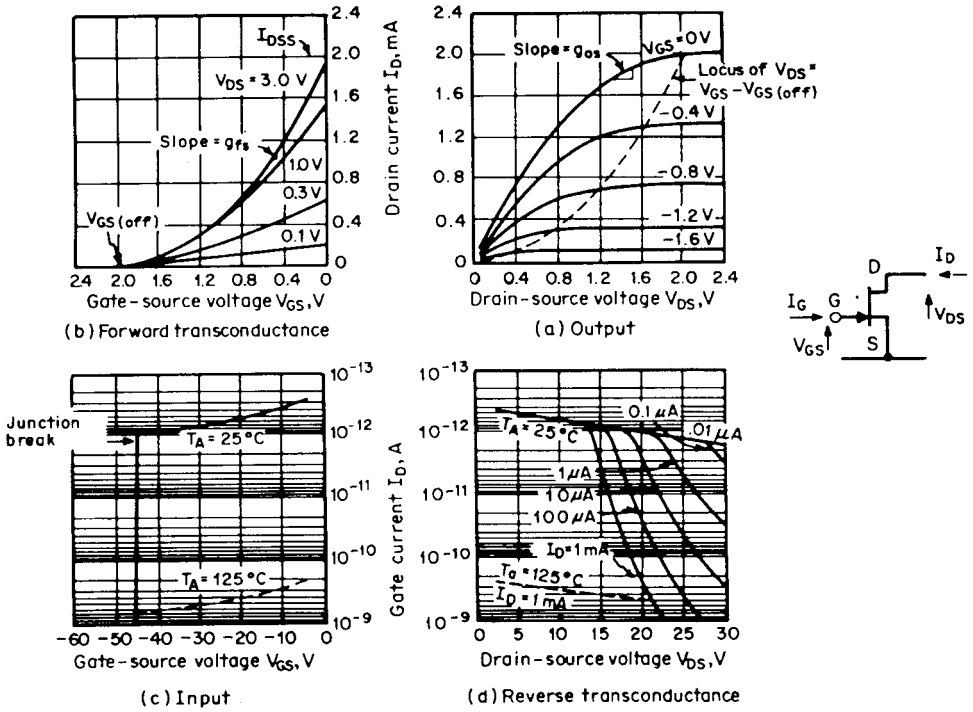


Figure 2-96 Static characteristics of an *n*-channel JFET.

The forward transconductance characteristic of Figure 2-96b can be approximated by a power law relation expressed as

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right)^n \tag{2-190}$$

if $V_{DS} \geq -V_{GS(\text{off})}$. By differentiation the small-signal transconductance g_{fs} is given by

$$g_{fs} = \frac{dI_D}{dV_{GS}} = -n\frac{I_{DSS}}{V_{GS(\text{off})}}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right)^{n-1} \tag{2-191}$$

Some texts indicate a value of $\frac{3}{2}$ for n; however, experimental measurements on a number of n-channel JFET geometries indicate that the exponent n is close to 2, which is the value derived in an approximate treatment by R. D. Middlebrook [10].

A useful relationship between g_{fso}, I_{DSS}, and $V_{GS(\text{off})}$ is derived from the ratio of Eqs. (2-190) and (2-191):

$$\frac{g_{gs}}{I_D} = n(V_{GS} - V_{GS(\text{off})})^{-1} \tag{2-192}$$

At $V_{GS} = 0$, $I_D = I_{DSS}$ and $g_{fs} = g_{fso}$. Using 2 as the value of the constant n leads to

$$g_{fso} = -2\frac{I_{DSS}}{V_{GS(\text{off})}} \tag{2-193}$$

For n-channel FETs I_{DSS} is positive and $V_{GS(\text{off})}$ is negative; for p-channel FETs I_{DSS} is negative and $V_{GS(\text{off})}$ is positive; thus g_{fs} is a positive quantity for both p- and n-channel FETs.

Equation (2-190) indicates that for $V_{GS} = V_{GS(\text{off})}$, $I_D = 0$. In a real device this does not happen. Starting from zero, as the gate voltage is made more negative, the drain current decreases until it reaches a very low value equal to the drain-leakage current. At this value, the source current will consist of source–gate leakage. Any further increase in the magnitude of the negative gate voltage will result in an increase in I_D leakage. For the small-signal device illustrated, this minimum I_D is on the order of 2×10^{-13} A.

For some types of applications of the FET, it is helpful to understand the characteristics at very low values of V_{DS}, such as those shown in Figure 2-97. A "very low value" is one that is small compared to the magnitude of $V_{GS} - V_{GS(\text{off})}$. In this region, V_{DS} is small enough to have little effect on channel thickness, so that the I_D/V_{DS} slope is nearly linear. Since the slope is a function of V_{GS}, the FET can be utilized as a voltage-controlled resistor. The conductance slope ($\Delta I_D/V_{DS}$) at $V_{DS} = 0$ is approximately a linear function of $V_{GS} - V_{GS(\text{off})}$.

If g_{ds} at $V_{GS} = 0$ is given the term g_{dso}, then

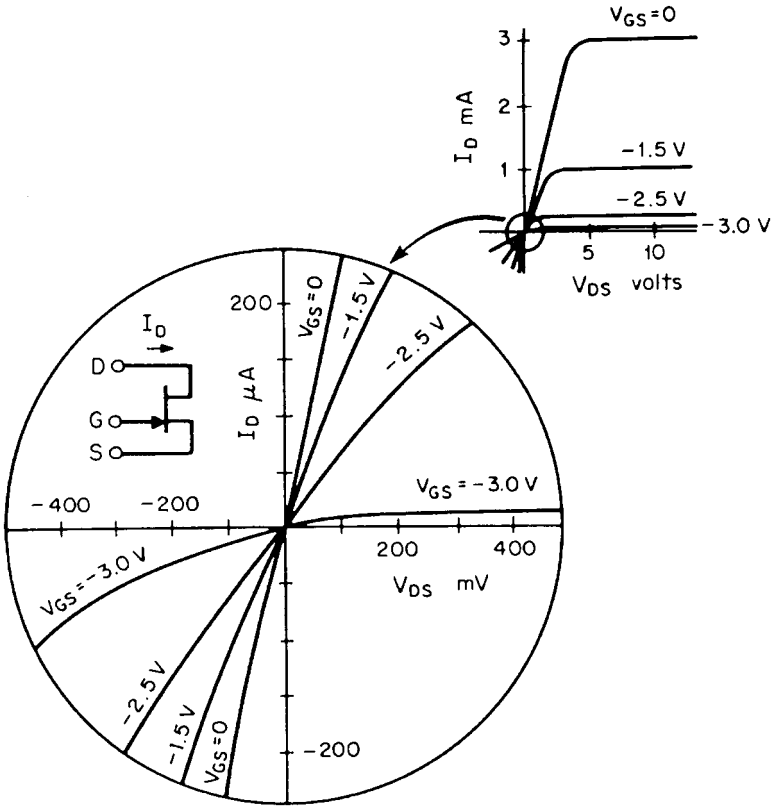


Figure 2-97 Enlargement of n-channel output characteristic around $V_{DS} = 0$.

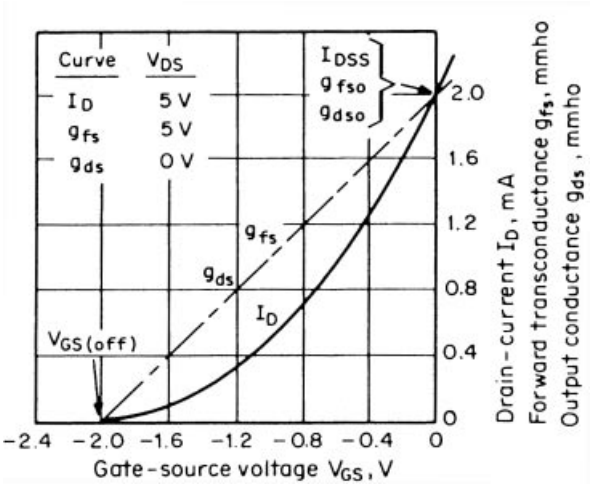


Figure 2-98 JFET I_D, g_{fs}, and g_{ds}.

$$g_{ds} = g_{dso}\left(1 - \frac{V_{GS}}{V_{GS(\text{off})}}\right) \tag{2-194}$$

with $V_{DS} = 0$. A plot of this characteristic is shown in Figure 2-98, along with g_{fs} and I_D characteristics.

The relationship between g_{dso}, $V_{GS(\text{off})}$, and I_{DSS} is given by the equation

$$g_{dso} = -2\frac{I_{DSS}}{V_{GS(\text{off})}} \tag{2-195}$$

where I_{DSS} and $V_{GS(\text{off})}$ are as indicated in Figure 2-96. It is important to note that Eqs. (2-194) and (2-195) and the g_{ds} curve of Figure 2-98 are valid only for the case where V_{DS} is very small compared to V_P. Drain–source conductance at higher values of V_{DS} will be discussed later. FETs designed to be used as voltage-controlled resistors typically have a high $V_{GS(\text{off})}$ because the $V_{DS}/(V_{GS} - V_{GS(\text{off})})$ ratio should be low to keep distortion low. The actual large-signal JFET model as used in most simulators is shown in Figure 2-99, including a list of its intrinsic keywords (Table 2-14). Again, we had very little luck using this model as published by several SPICE CAD software manufacturers as offered. While we cannot judge on the quality of the parameters that came with the software, it was not possible to closely match even optimized sets of parameters to this large-signal model and obtain acceptable results. However, since the modified Materka model worked very well for this, we recommend that this popular JFET model not be used at frequencies above 10 MHz.

2-3-2 Small-Signal Behavior of JFETs

Figure 2-100 shows the small-signal equivalent transistor model we would recommend. It consists of an empirical FET model and a package model. Table 2-15 lists its keywords. Its advantage is that its built-in noise model is very accurate, even for JFETs and metal semiconductor FETs.

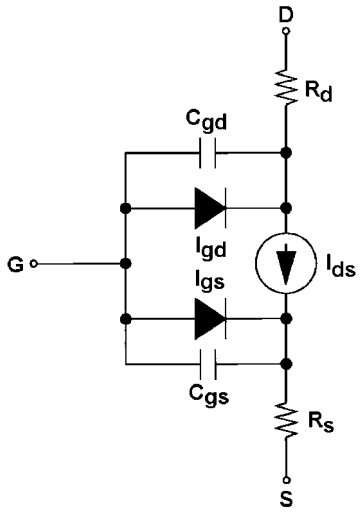


Figure 2-99 The nonlinear JFET model. Table 2-14 lists its intrinsic and extrinsic parameters. In our opinion, this model is limited to frequencies below 10 MHz; above this frequency, the modified Materka model should replace it.

Table 2-14 Large-signal JFET model keywords

Keyword	Description	Unit	Default
	Intrinsic Model		
NJF/PJF	Channel-type selection (NJF = *n*-channel, PJF = *p*-channel)		NJF
VT0	Threshold voltage	volt	–2.0
BETA	Transconductance coefficient	ampere/volt^2	1.0E–4
LAMB	Channel-length modulation	/volt	0.0
IS	Gate-junction saturation current	ampere	1.0E–14
PB	Gate-junction potential	volt	1.0
FC	Forward-bias depletion capacitance coefficient		0.5
CGS	Zero-bias gate–source junction capacitance	farad	0.0
CGD	Zero-bias gate–drain junction capacitance	farad	0.0
T	Channel transit-time delay	second	0.0
KF	Flicker noise coefficient		0.0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
SN	Switch to turn device shot noise on (1) or off (0)		1
AREA	Area multiplier		1.0
NOIS	Reference label to a set of noise data		
NAME	Required user-specified name up to 8 characters		
	Extrinsic Model		
RD	Drain ohmic resistance	ohm	0.0
RS	Source ohmic resistance	ohm	0.0

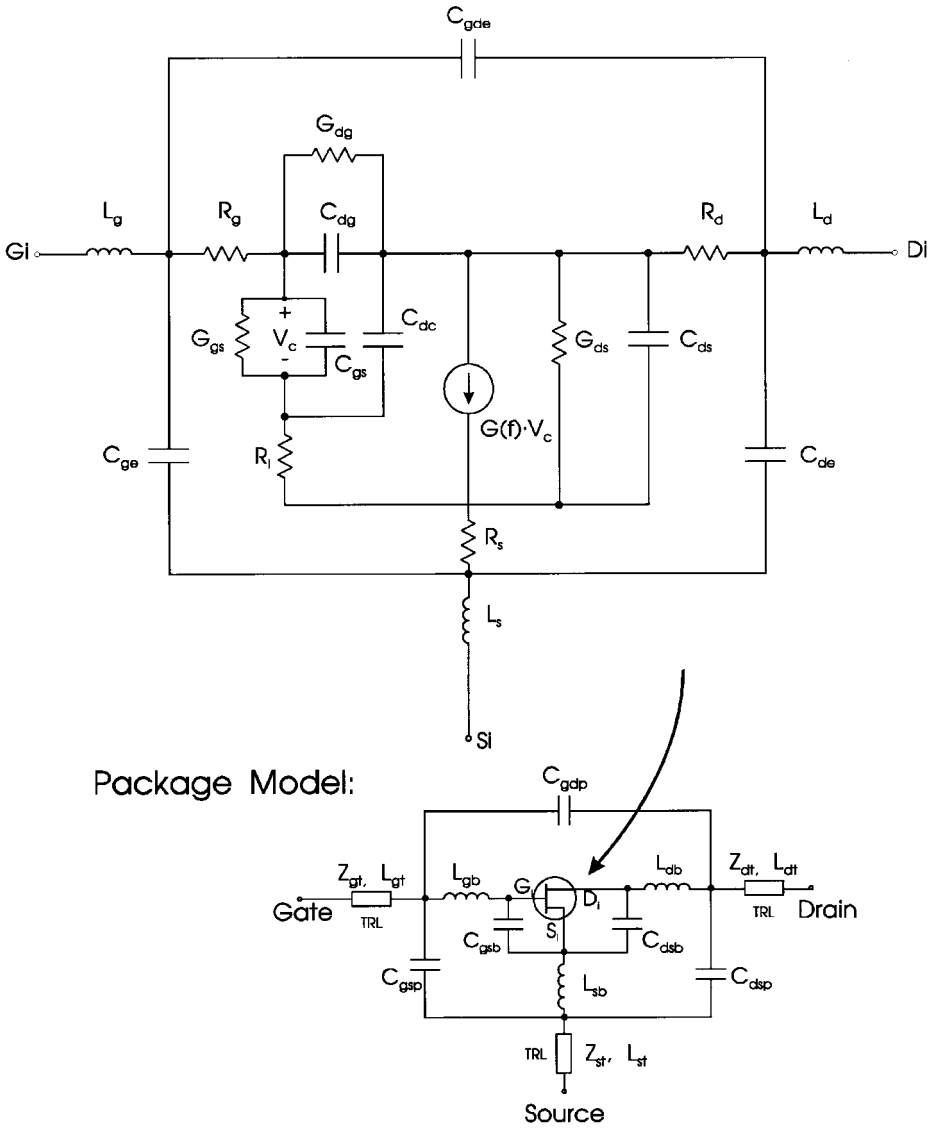


Figure 2-100 Linear FET model.

Electrical noise generated within a JFET is usually represented by equivalent noise sources, $\overline{e}_n$ and $\overline{i}_n$. Both noise voltage $\overline{e}_n$ and noise current $\overline{i}_n$ are frequency dependent and have the characteristics shown in Figure 2-101.

An equivalent noise circuit is shown in Figure 2-102. Above the frequency f_1, $\overline{e}_n$ is approximately given by

$$\overline{e}_n \cong \left(4KTB\,\frac{0.67}{g_{fs}}\right)^{1/2} \tag{2-196}$$

Table 2-15 Small-signal FET model keywords

Keyword	Description	Unit	Default
	Intrinsic model		
G	Transconductance at dc, G_0[a]	/ohm	
CGS	Gate–source capacitance	farad	
F	3-dB roll-off frequency	hertz	∞
T	Time delay	second	0.0
TDS	Drain–source time delay	second	0.0
GGS	Gate–source conductance	/ohm	0.0
CDG	Drain–gate capacitance	farad	0.0
CDC	Dipole layer capacitance	farad	0.0
CDS	Drain–source capacitance	farad	0.0
GDS	Drain–source conductance	/ohm	0.0
RI	Channel resistance	ohm	0.0
RG	Gate resistance	ohm	0.0
RD	Drain resistance	ohm	0.0
RS	Source resistance	ohm	0.0
CGE	External gate capacitance	farad	0.0
CDE	External drain capacitance	farad	0.0
LG	Gate-lead inductance	henry	0.0
LD	Drain-lead inductance	henry	0.0
LS	Source-lead inductance	henry	0.0
CGDE	External gate–drain capacitance	farad	0.0
GDG	Gate–drain conductance	/ohm	0.0
TJ	Chip temperature	kelvin	298
	Package Parasitics		
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate to source package capacitance	farad	0.0
CDSP	Drain to source package capacitance	farad	0.0
CGDP	Gate to drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for at $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for $\varepsilon_r = 1$	meter	0.0
FC	Corner frequency of flicker (1/*f*) noise[b]	hertz	10 MHz
FCP	Shape factor of the 1/*f* noise response		1.0
label	User-defined term that refers to temperature coefficient		

[a]The transconductance of this model may be approximately described by $g_m = G\,\frac{e^{-j\omega T}}{1 + j(f/F)}$ where $\omega = 2\pi f$ and f is frequency.
[b]The flicker noise frequency dependence is given by $1/(f/F_c)^{\mathrm{FCP}}$.

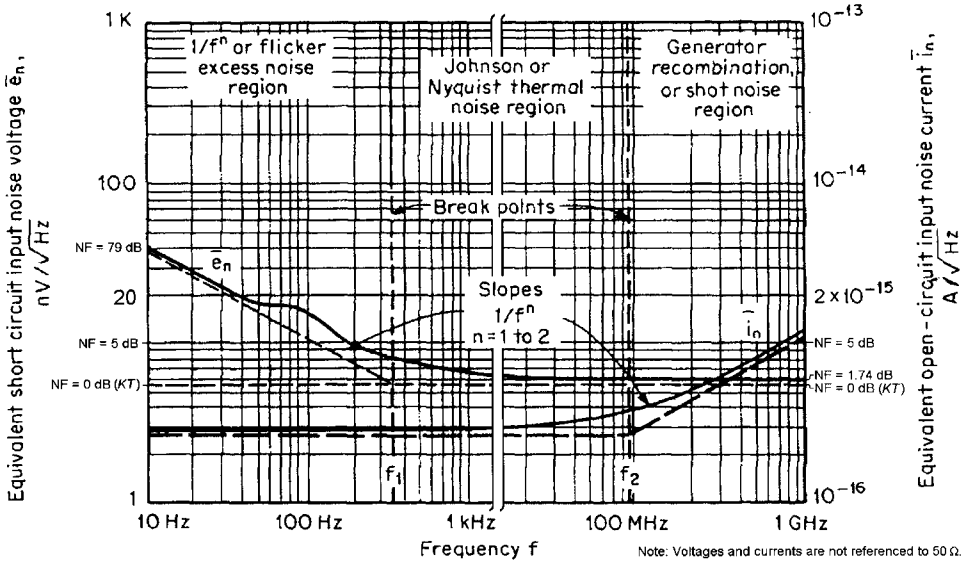


Figure 2-101 JFET noise characteristics.

where $K = 1.374 \times 10^{-23}$ J/K

T = absolute temperature in kelvins (273 K = 0 °C)

B = frequency range in hertz

g_{fs} = transconductance of FET

With the input short-circuited, the noise voltage across the load R_L resulting from the FET is

$$\text{Output noise voltage} = \bar{e}_n A_V \tag{2-197}$$

Below the frequency f_1, $\bar{e}_n$ increases proportional to $1/f^n$ and is expressed as

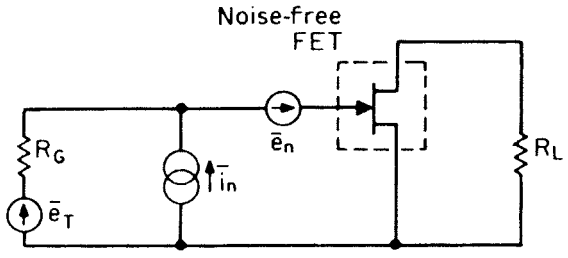


Figure 2-102 Equivalent noise circuit of the FET.

$$\bar{e}_n = \left[4KTB \left(\frac{0.67}{g_{fs}} \right) \left(1 + \frac{f_1}{f^n} \right)^{1/2} \right] \tag{2-198}$$

The low-frequency corner frequency f_1 for JFETs is typically in the 100-Hz to 1-kHz range, and the exponent n is usually between 1 and 2. As indicated by the equations, $\bar{e}_n$ is inversely proportional to the square root of g_{fs}.

The equivalent input noise current $\bar{i}_n$ is caused by the current in the gate-to-channel junction. Its approximate value below f_2 is

$$\bar{i}_n = (2qI_G B)^{1/2} \tag{2-199}$$

where $q = 1.602 \times 10^{-19}$
B = frequency range in hertz
I_G = dc gate current

This expression is fairly accurate when I_G is the result of the active device conductance. Typically, $\bar{i}_n$ will be lower than the calculated value because part of I_G is due to conductance across the device package.

At higher frequencies (above f_2),

$$\bar{i}_n = \left(\frac{4KTB}{R_P} \right)^{1/2} \tag{2-200}$$

where R_P is the real part of gate input impedance. R_P, in terms of Y_{11}, can range from several tens of megohms (at audiofrequencies) to 1 kΩ or less (at VHF/UHF frequencies). The high-frequency corner f_2 is typically in the range of 5–50 kHz and more than 100 MHz for high-frequency devices.

Another form of noise is known as "popcorn" or burst noise, the causes of which have not been completely identified. It shows up as a random short-duration step-function change in drain current, equivalent to an input gate–source voltage change of a few tenths of a microvolt. It is not unlike big bubbles on the surface of boiling water.

As far as the noise is concerned, we can also state that the Materka built-in proprietary noise model is significantly more accurate than the one mentioned above [11–14]. This model is implemented, with minimal documentation of its inner workings because of its proprietary nature, in the linear portion of the circuit simulator in Ansoft's Serenade Design Environment. In many cases, the manufacturer supplies only limited data, so it is most convenient that after SPICE-type parameter extraction is done by using the Scout program, one even gets good noise prediction for JFETs using the modified Materka model.

2-3-3 Large-Signal Behavior of MOSFETs

Metal-oxide-semiconductor field-effect transistors (MOSFETs) are important components in contemporary analog integrated circuits. Whereas initial applications were centered on all-MOS processes, combined bipolar and MOS processes give the designer the best of both worlds. A major advantage of MOS processes for realizing analog functions is that complex, dense digital functions can be realized on the same chip.

Transfer Characteristics of MOS Devices. A cross section of a typical enhancement-mode *n*-channel MOS transistor (NMOS) is shown in Figure 2-103a. Heavily doped *n*-type source and drain regions are fabricated in a *p*-type substrate (often called the body). A thick layer of silicon dioxide is grown over the substrate material and a conductive gate material (metal or polycrystalline silicon) covers this oxide between source and drain. The operation of the device is very similar to that of a JFET in that the gate–source voltage is used to modify the conductance of the region under the gate. This allows the gate voltage to control the current flowing between source and drain, giving gain in analog circuits and switching characteristics in digital circuits.

The enhancement-mode NMOS device of Figure 2-103a shows significant conduction between source and drain only when an *n*-type channel exists under the gate, and this is the origin of the "*n*-channel" designation. The term "enhancement mode" refers to the fact that no conduction occurs for $V_{GS} = 0$, and thus the channel must be "enhanced" to cause conduction. MOS devices can be made equally well by using an *n*-type substrate with a *p*-type conducting channel. Such devices are called enhancement-mode *p*-channel MOS transistors (PMOS). In technologies employing one or the other of these device types, the circuit symbol of Figure 2-103b is commonly used for either one. In complementary MOS technology (CMOS) both device types are present and the circuit symbols of Figure 2-103c will be used to distinguish them. In PMOS and NMOS technologies, the substrate is common to all devices, invariably connected to a dc power supply voltage, and usually not shown on the circuit diagram. In CMOS technology, however, devices of one type or another are fabricated in individual, separate isolation regions, which may or may not be connected to a power supply voltage. If these isolation regions *are* connected to the appropriate power supply, the symbols of Figure 2-103c will be used and the substrate connection will not be shown. If the individual isolation regions are connected elsewhere, however, the devices will be represented by the symbols of Figure 2-103d, where the substrate is labeled *B*. Finally, in NMOS technology an additional device type called a "depletion-mode" device is usually available. This is a conducting channel implanted between the source and drain so that conduction occurs for $V_{GS} = 0$. This device has characteristics that are almost identical to that of a JFET and will be represented by the symbol of Figure 2-103e.

The derivation of the transfer characteristics of the enhancement-mode NMOS device of Figure 2-103a begins by noting that with $V_{GS} = 0$, the source and drain regions are separated by back-to-back *pn* junctions. These junctions are formed between the *n*-type source and drain regions and the *p*-type substrate, resulting in an extremely high resistance (about 10^{12} Ω) between drain and source when the device is off.

Now consider substrate, source, and drain grounded and a positive voltage V_{GS} applied to the gate as shown in Figure 2-104. The gate and substrate then form the plates of a capacitor with the SiO_2 as a dielectric. Positive charge accumulates on the gate and the negative charge in the substrate. Initially, the negative charge in the *p*-type substrate is manifested by creation of a *depletion region* and resulting exclusion of holes under the gate. This is shown in Figure 2-104. The depletion-layer width *X* under the oxide is

$$X = \left(\frac{2\varepsilon\phi}{qN_A} \right)^{1/2} \tag{2-201}$$

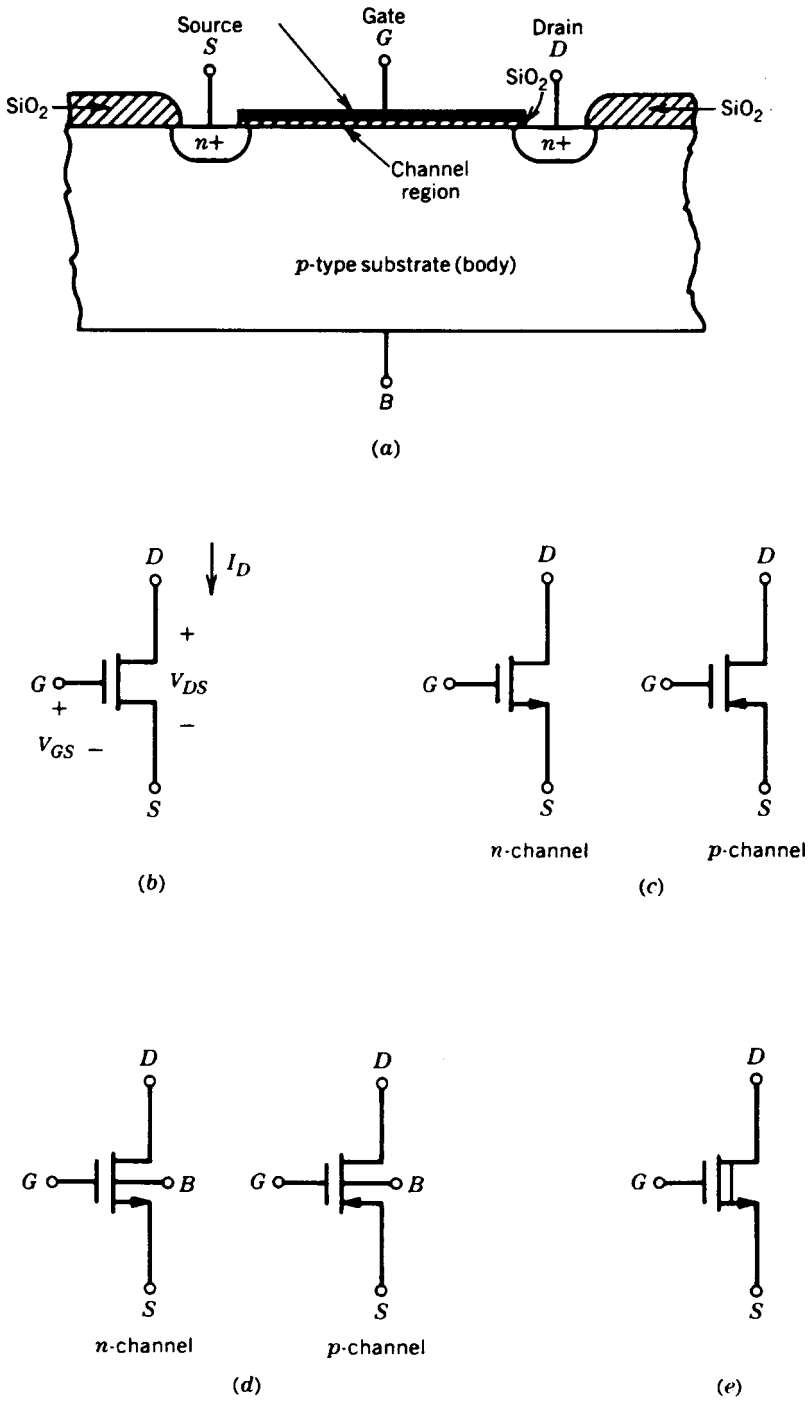


Figure 2-103 (a) Typical enhancement-mode NMOS structure. (b) Enhancement-mode NMOS or PMOS circuit symbol when one device type only is present. (c) NMOS and PMOS symbols used in CMOS circuits. (d) NMOS and PMOS symbols used when the substrate connection is nonstandard. (e) Depletion MOS device symbol.

where ϕ is the potential in the depletion layer at the oxide–silicon interface, N_A atoms/cm^3 is the doping density (assumed constant) of the p-type substrate, and ε is the permittivity of the silicon. The charge per unit area in this depletion region is

$$Q = qN_A X = \sqrt{2qN_A \varepsilon \phi} \tag{2-202}$$

When the potential in the silicon reaches a critical value equal to twice the Fermi level, $\phi_f \simeq 0.3$ V, a phenomenon known as "inversion" occurs [15]. Further increases in gate voltage produce no further changes in the depletion-layer width but instead a thin layer of electronics is induced in the depletion layer directly under the oxide. This produces a continuous n-type region with the source and drain regions and is the conducting channel between source and drain. This channel can then be modulated by increases or decreases in the gate voltage. In the presence of an inversion layer, and with no substrate bias, the depletion region contains a fixed charge

$$Q_{b0} = \sqrt{2qN_A \varepsilon 2\phi_f} \tag{2-203}$$

If a substrate bias voltage V_{SB} (source positive for n-channel devices) is applied between source and substrate, the potential required to produce inversion becomes $(2\phi_f + V_{SB})$ and the charge stored in the depletion region in general is

$$Q_b = \sqrt{2qN_A \varepsilon (2\phi_f + V_{SB})} \tag{2-204}$$

The gate voltage V_{GS}, required to produce an inversion layer, is called the threshold voltage V_t and can now be calculated. This voltage consists of several components. First, a voltage $[2\phi_f + (Q_b/C_{ox})]$ is required to sustain the depletion-layer charge Q_b, where C_{ox} is the gate oxide capacitance per unit area. Second, a work-function difference ϕ_{ms} exists between the

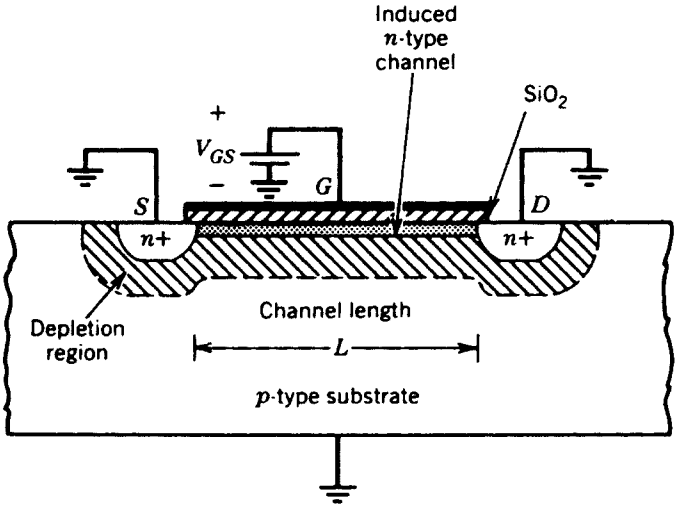


Figure 2-104 Idealized NMOS device cross section with positive V_{GS} applied, showing depletion regions and the induced channel.

gate metal and the silicon. Third, there is always charge density Q_{ss} (positive) in the oxide at the silicon interface. This is caused by crystal discontinuities at the Si–SiO_2 interface and must be compensated by a gate voltage contribution of $-Q_{ss}/C_{ox}$. Thus we have a threshold voltage

$$V_t = \phi_{ms} + 2\phi_f + \frac{Q_b}{C_{ox}} - \frac{Q_{ss}}{C_{ox}}$$

$$= \phi_{ms} + 2\phi_f + \frac{Q_{b0}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} + \frac{Q_b - Q_{b0}}{C_{ox}} \tag{2-205}$$

$$= V_{t0} + \gamma\left(\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f}\right) \tag{2-206}$$

where Eqs. (2-203) and (2-204) have been used and V_{t0} is the threshold voltage with $V_{SB} = 0$. The parameter γ is defined as

$$\gamma = \frac{1}{C_{ox}}\sqrt{2q\varepsilon N_A} \tag{2-207}$$

and

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \tag{2-208}$$

where ε_{ox} and t_{ox} are the permittivity and thickness of the oxide, respectively. A typical value of γ is 0.5 $V^{1/2}$ and $C_{ox} = 3.5 \times 10^{-4}$ pF/μ^2 for $t_{ox} = 0.1$ μm.

In practice, the value of V_{t0} is usually adjusted in processing by implanting additional impurities into the channel region. Extra *p*-type impurities are implanted in the channel to make $V_{t0} \simeq 0.5$–1.5 V for *n*-channel enhancement devices. By implanting *n*-type impurities in the channel region, a conducting channel can be formed, even for $V_{GS} = 0$. This MOS transistor is called a depletion device, with typical values of V_{t0} in the range –1 to –4 V. If Q_i is the charge density per unit area due to the implant, then the threshold voltage given by Eq. (2-205) is shifted by approximately Q_i/C_{ox}.

The preceding equations can now be used to calculate the large-signal characteristics of an NMOS transistor. For purposes of analysis the source is assumed grounded and bias voltages V_{GS}, V_{DS}, and V_{SB} are applied as shown in Figure 2-105. If V_{GS} is greater than V_t, a conducting channel exists and V_{DS} causes a larger reverse bias from drain to substrate than exists from source to substrate, and thus a wider depletion region exists at the drain. However, for simplicity, we assume that the voltage drop along the channel itself is small so that the depletion layer is constant along the channel.

At a distance *y* along the channel the voltage with respect to the source is $V(y)$ and the gate-to-channel voltage at that point is $V_{GS} - V(y)$. We assume this voltage exceeds the threshold voltage V_t, and thus the induced electron charge per unit area in the channel is

$$Q_I(y) = C_{ox}[V_{GS} - V(y) - V_t] \tag{2-209}$$

The resistance *dR* of a length *dy* of the channel is

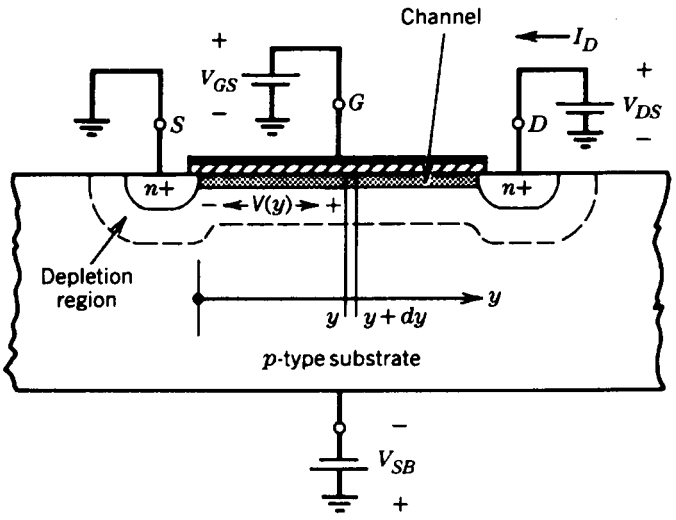


Figure 2-105 NMOS device with bias voltages applied.

$$dR = \frac{dy}{W\mu_n Q_I(y)} \tag{2-210}$$

where W is the width of the device, perpendicular to the plane of Figure 2-105, and μ_n is the average electron mobility of the channel.

The voltage drop dV along the length of the channel dy is

$$dV = I_D\, dR = \frac{I_D}{W\mu_n Q_I(y)}\, dy \tag{2-211}$$

If L is the total channel length, then substitution of Eq. (2-209) into Eq. (2-211) and integration gives

$$\int_0^L I_D\, dy = \int_0^{V_{DS}} W\mu_n C_{\text{ox}}(V_{GS} - V - V_t)\, dV \tag{2-212}$$

This results in

$$I_D = \frac{k'}{2}\frac{W}{L}[2(V_{GS} - V_t)V_{DS} - V_{DS}^2] \tag{2-213}$$

where

$$k' = \mu_n C_{\text{ox}} = \frac{\mu_n \varepsilon_{\text{ox}}}{t_{\text{ox}}} \tag{2-214}$$

Equation (2-213) is important and describes the I–V characteristics of an MOS transistor, assuming a continuous induced channel. A typical value of k' for $t_{\text{ox}} = 0.1$ μm is about 20 μA/V^2 for an n-channel device.

As the value of V_{DS} is increased, the induced conducting channel narrows at the drain end and Eq. (2-209) indicates that Q_I at the drain end approaches zero as V_{DS} approaches $(V_{GS} - V_t)$. This results in the same pinchoff phenomenon as occurs in a JFET, and further increases in V_{DS} produce little change in I_D. Equation (2-213) is thus no longer valid if V_{DS} is greater than $(V_{GS} - V_t)$. The value of I_D in this region is obtained by substituting $V_{DS} = (V_{GS} - V_t)$ in Eq. (2-213), giving

$$I_D = \frac{k'}{2}\frac{W}{L}(V_{GS} - V_t)^2 \tag{2-215}$$

for an MOS transistor in the pinchoff region. As in the case of a JFET, the drain current in the pinchoff region varies slightly as the drain voltage is varied. This is due to the presence of a depletion region between the physical pinchoff point in the channel at the drain end and the drain region itself. If this depletion-layer width is X_d, then the *effective* channel length is given by

$$L_{\text{eff}} = L - X_d \tag{2-216}$$

If L_{eff} is used in place of L in Eq. (2-215), we obtain a more accurate formula for the pinchoff region:

$$I_D = \frac{k'}{2}\frac{W}{L_{\text{eff}}}(V_{GS} - V_t)^2 \tag{2-217}$$

The fact that X_d (and thus L_{eff}) is a function of the drain–source voltage results in a variation of I_D and V_{DS} in the pinchoff region. Using Eqs. (2-216) and (2-217), we obtain

$$\frac{\partial I_D}{\partial V_{DS}} = -\frac{k'}{2}\frac{W}{L_{\text{eff}}^2}(V_{GS} - V_t)^2\frac{dL_{\text{eff}}}{dV_{DS}} \tag{2-218}$$

and thus

$$\frac{\partial I_D}{\partial V_{DS}} = \frac{I_D}{L_{\text{eff}}}\frac{dX_d}{dV_{DS}} \tag{2-219}$$

This equation is analogous to Eq. (2-168) for bipolar transistors. Following a similar procedure, a voltage analogous to the Early voltage can be defined as

$$V_A = \frac{I_D}{\partial I_D/\partial V_{DS}} \tag{2-220}$$

and thus

$$V_A = L_{\text{eff}}\left(\frac{dX_d}{dV_{DS}}\right)^{-1} \tag{2-221}$$

For MOS transistors, the most widely used parameter for the characterization of output resistance is

$$\lambda = \frac{1}{V_A} \tag{2-222}$$

As in the bipolar case, the large-signal properties of the transistor can be approximated by assuming that λ and V_A are constants, independent of bias conditions. Thus we can formulate a better approximation to the I–V characteristics as

$$I_D = \frac{k'}{2}\frac{W}{L}(V_{GS} - V_t)^2 (1 + \lambda V_{DS}) \tag{2-223}$$

In practical MOS transistors variation of X_d with voltage is complicated by the fact that the field distribution in the drain depletion region is not one dimensional. A vertical component in the field distribution is introduced by the potential difference between the gate and channel and the gate and drain. As a result, the calculation of λ from the device structure is quite different [16], and it is usually necessary to develop effective values of λ from experimental data. The parameter λ is a linear function of effective channel length and is an increasing function of the doping level in the channel. Typical values of λ are in the range 0.05–0.005 V^{-1}. Note that, in the case of a JFET, the pinchoff region for MOS devices is often called the *saturation* region.

A plot of I_D versus V_{DS} for an NMOS transistor is shown in Figure 2-106. Below pinchoff, the device behaves as a nonlinear voltage-controlled resistor, which is often called the *ohmic* or *triode* region. Above pinchoff, the device approximates a voltage-controlled current source. Note that for depletion MOS devices, V_t is negative and I_D is finite, even for $V_{GS} = 0$. For PMOS devices, all polarities of voltage and current are reversed.

The results as previously derived can be used to form a large-signal model of an MOS transistor. The model topology in the pinchoff region is the same as that for the JFET, but using Eq. (2-215) for the controlled-current generator.

MOS Device Voltage Limitations. The voltage limitations of MOS transistors depend on the gate length L. For small values of L (less than about 10 μm), the drain-depletion region

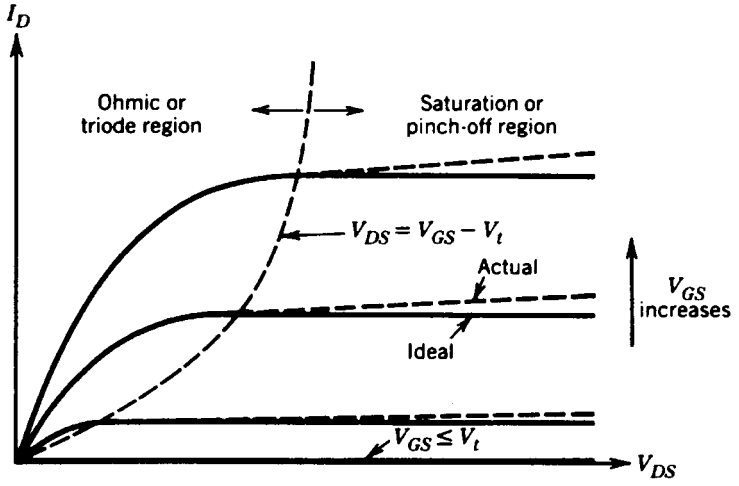


Figure 2-106 NMOS device characteristics.

exerts an appreciable influence on the channel. This causes I_D to rise with increasing V_{DS} in a similar fashion to the bipolar transistor curves of Figure 2-89.

For long channel lengths, the drain-depletion region has little effect on the channel and the I_D-versus-V_{DS} curves follow closely the ideal curves of Figure 2-106. Eventually the drain–substrate *pn*-junction breakdown voltage is exceeded and a sharp breakdown characteristic is obtained, which is similar to the JFET characteristic.

In addition to V_{DS} limitations, MOS devices must also be protected against excessive gate voltages. Typical gate oxides break down with about 25–50 V applied from gate to channel, and this process is destructive to the transistor.

2-3-4 Small-Signal Model of the MOS Transistor in Saturation

The preceding large-signal equations can now be used to derive the small-signal model of the MOS transistor in the saturation or pinchoff region. The important equations are (2-220) and (2-206). Note from Eq. (2-206) that the source–substrate voltage V_{BS} affects V_t, and thus I_D. This is due to the influence of the substrate acting as a second gate and is called *body effect*. As a consequence, I_D is a function of both V_{GS} and V_{BS}, and we require *two* transconductance generators in the small-signal model as shown in Figure 2-107. Variations in the voltage v_{bs} from source to body cause current $g_{mb}v_{bs}$ to flow from drain to source. Note that the body (or substrate) of an NMOS integrated circuit is usually connected to the most negative supply voltage and is thus an ac ground. However, the source connection can have a significant ac voltage impressed on it. Parasitic resistances due to the channel contact regions should be included in series with the source and drain of the model but are usually neglected in hand calculations. These resistances have an inverse dependence on channel width W and have typical values of 50–100 Ω for devices with W of about 1 μm.

The parameters of Figure 2-107 can be determined from Eq. (2-223) by differentiating.

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = k' \frac{W}{L} (V_{GS} - V_t)(1 + \lambda V_{DS}) \tag{2-224}$$

If $\lambda V_{DS} << 1$, this is often approximated as

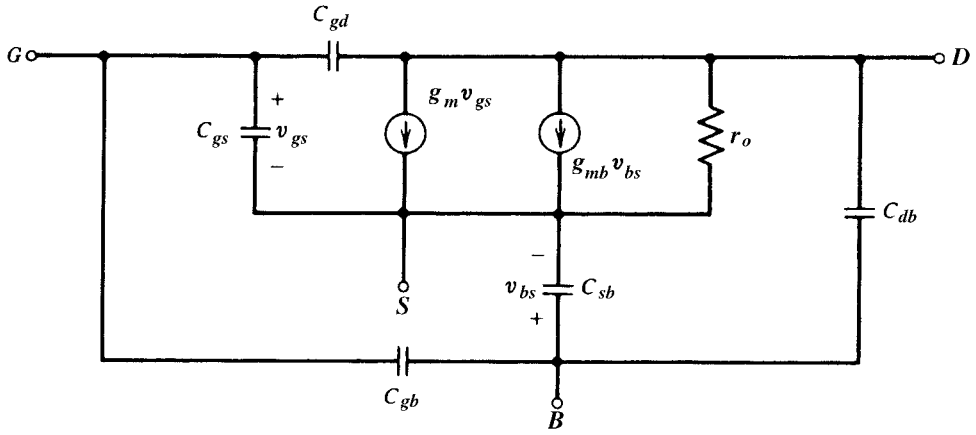


Figure 2-107 Small-signal MOS transistor equivalent circuit.

$$g_m = k' \frac{W}{L} (V_{GS} - V_t) = \sqrt{2k' \frac{W}{L} I_D} \tag{2-225}$$

Like the JFET and unlike the bipolar transistor, the transconductance of the MOSFET depends on both bias current and the *W*/*L* ratio (also on the oxide thickness via *k*′). Similarly,

$$g_{mb} = \frac{\partial I_D}{\partial V_{BS}} = -k' \frac{W}{L} (V_{GS} - V_t) (1 + \lambda V_{DS}) \frac{\partial V_t}{\partial V_{BS}} \tag{2-226}$$

From Eq. (2-206)

$$\frac{\partial V_t}{\partial V_{BS}} = -\frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} = -\chi \tag{2-227}$$

where this equation defines a factor χ. This parameter is equal to the rate of change of threshold voltage with body bias voltage. One obtains

$$\chi = \frac{C_{js}}{C_{\text{ox}}} \tag{2-228}$$

where C_{js} is the capacitance per unit area of the depletion region under the channel, assuming a one-sided step junction with a built-in potential $\psi_0 = 2\phi_f$.

Substitution of Eq. (2-227) into Eq. (2-226) gives

$$g_{mb} = \frac{\gamma k' (W/L) (V_{GS} - V_t) (1 + \lambda V_{DS})}{2\sqrt{2\phi_f + V_{SB}}} \tag{2-229}$$

Again, if $\lambda V_{DS} << 1$, we have

$$g_{mb} = \frac{\gamma \sqrt{k' (W/L) I_D}}{\sqrt{2(2\phi_f + V_{SB})}} \tag{2-230}$$

An important quantity is the ratio g_{mb}/g_m, and from Eqs. (2-224) and (2-229) we find

$$\frac{g_{mb}}{g_m} = \frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} = \chi \tag{2-231}$$

The factor χ is typically in the range 0.1–0.3.

Finally, the small-signal output resistance can be obtained directly from Eq. (2-218):

$$r_o = \left(\frac{\partial I_D}{\partial V_{DS}} \right)^{-1} = \frac{L_{\text{eff}}}{I_D} \left(\frac{dX_d}{dV_{DS}} \right)^{-1} \tag{2-232}$$

and using Eqs. (2-221) and (2-222), we find

$$r_o = \frac{1}{\lambda I_D} = \frac{V_A}{I_D} \tag{2-233}$$

Small-signal model capacitances are also shown in Figure 2-107. Of these, only the gate–source capacitance C_{gs} is intrinsic to the device operation in the saturation region. Capacitances C_{sb} and C_{db} are parasitic depletion-region capacitances between the substrate and the source and drain regions, respectively. These can be expressed as follows:

$$C_{sb} = \frac{C_{sb0}}{(1 + V_{SB}/\psi_0)^{1/2}} \tag{2-234}$$

$$C_{db} = \frac{C_{db0}}{(1 + V_{DB}/\psi_0)^{1/2}} \tag{2-235}$$

These capacitances are proportional to the gate and source region areas (including sidewalls), and C_{sb} also includes depletion-region capacitance from the induced channel in the body.

Capacitance C_{gb} between gate and substrate models parasitic oxide capacitance between the gate contact material and the substrate outside the active-device area. This is a constant capacitance and models coupling between polysilicon and metal interconnects and the underlying substrate. In fact, parasitic capacitance of this type underlies all polysilicon and metal traces on the chip and should be taken into account when simulating and calculating high-frequency circuit and device performance. Typical values depend on oxide thicknesses and range from about 0.04 to 0.15 fF per square micron of interconnect, with fringing effects becoming important for narrow lines (several microns or less in width).

Capacitances C_{gs} and C_{gd} exist from gate to source and drain, respectively. If C_{ox} is the oxide capacitance per unit area from gate to channel then the total capacitance under the gate is $C_{\text{ox}}WL$. This capacitance is intrinsic to the device operation and models the gate control of the channel conductance. In the ohmic region of device operation, this capacitance is split equally between source and drain so that $C_{gs} = C_{gd} = 1/2C_{\text{ox}}WL$. However, in the saturation region, the channel is very narrow at the drain end and the drain voltage exerts little influence on either the channel or the gate charge. As a consequence, the intrinsic portion of C_{gd} is essentially zero in the saturation region and C_{gd} then consists of a constant parasitic oxide-capacitance contribution due to gate overlap of the drain region. This is on the order of 1–10 fF for small devices.

In order to calculate the corresponding value of C_{gs} in the saturation region, we must calculate the total charge Q_T stored in the channel. This can be obtained by integrating Eq. (2-209) to obtain

$$Q_T = WC_{\text{ox}} \int_0^L [V_{GS} - V(y) - V_t]dy \tag{2-236}$$

Substituting for dy/dV from Eq. (2-211) into Eq. (2-236), we find

$$Q_T = \frac{W^2 C_{\text{ox}}^2 \mu_n}{I_D} \int_0^{V_{GS}-V_t} (V_{GS} - V - V_t)^2 dV \tag{2-237}$$

where the limit $y = L$ corresponds to $V = (V_{GS} - V_t)$ in saturation. Solution of Eq. (2-237) and use of Eqs. (2-214) and (2-215) gives

$$Q_T = \frac{2}{3} WLC_{\text{ox}}(V_{GS} - V_t) \tag{2-238}$$

and thus

$$C_{gs} = \frac{\partial Q_T}{\partial V_{GS}} = \frac{2}{3} WLC_{\text{ox}} \tag{2-239}$$

In addition, there is a contribution to C_{gs} from the constant parasitic oxide capacitance due to gate overlap of the source region.

The f_T of the MOSFET is given by

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{gs} + C_{gd} + C_{gb}} \tag{2-240}$$

The dependence of MOSFET f_T on device and process parameters can be seen by assuming that the intrinsic device capacitance G_{gs} dominates. Thus from Eq. (2-240) we have

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{gs}} \tag{2-241}$$

Substituting in Eq. (2-241) for g_m from Eq. (2-225) and C_{gs} from Eq. (2-239) we find for a MOSFET

$$f_T = 1.5 \frac{\mu_n}{2\pi L^2} (V_{GS} - V_t) \tag{2-242}$$

It is interesting to compare this with the intrinsic f_T of a bipolar transistor when parasitic depletion-layer capacitance is neglected. From

$$f_T = \frac{1}{2\pi\tau_F} \tag{2-243}$$

and substituting for τ_f and using the Einstein relationship $D_n/\mu_n = kT/q = V_T$, we find for a bipolar transistor

$$f_T = 2 \frac{\mu_n}{2\pi W_B^2} V_T \tag{2-244}$$

The similarity in form between Eqs. (2-242) and (2-244) is striking. In both cases the intrinsic device f_T increases as the inverse square of the critical device dimension across which carriers are in transit. The voltage V_t = 26 mV is fixed for a bipolar transistor, but MOSFET f_T can be increased by operating at high values of $(V_{GS} - V_t)$. Note that the base width W_B in a bipolar transistor is a vertical dimension determined by diffusions or implants and can typically be made much smaller than the channel length L of a MOSFET, which depends on surface geometry and photolithographic processes. Thus bipolar transistors generally have higher f_T than MOSFETs made with comparable processing. Finally, Eq. (2-242) was derived assuming that the MOSFET square law is valid. However, as discussed in Section 2-3-5, submicron MOSFETs depart significantly from square-law characteristics, and we find that for such devices f_T increases as L^{-1} rather than L^{-2}.

2-3-5 Short-Channel Effects in FETs

The evolution of integrated-circuit processing techniques has led to continuing reductions in both the horizontal and vertical dimensions of the active devices (the minimum allowed dimension of passive devices has also increased). This trend is driven primarily by economics in that more devices and circuits can be processed at one time on a given wafer. A second benefit has been that the frequency capability of the active devices continues to increase, as intrinsic f_T values increase with smaller dimensions while parasitic capacitances decrease.

Vertical dimensions such as the base width of a bipolar transistor in production processes may now be on the order of 0.05 μm or less, whereas horizontal dimensions such as bipolar emitter width or FET gate length may be significantly less than 1 μm. Even at these very small dimensions, the large-signal and small-signal models of bipolar transistors given in previous sections remain valid. However, significant short-channel effects become important in FETs of all types at channel lengths on the order of 1 μm or less and require modifications to the FET models given previously. The primary effect is to modify the classical FET square-law transfer characteristic in the saturation region to make the device more closely approach an ideal linear transfer function. Note, however, that even in processes with submicron capability, many of the FETs in a given analog circuit may be deliberately chosen to be larger than the minimum size and may be well approximated by the square-law model.

The most important short-channel effect in FETs is due to velocity saturation of carriers in the channel [17]. At low electric field values, the linear relation between carrier velocity and field implied by Eqs. (2-210) and (2-211) is valid. At high fields, however, the carrier velocities approach the thermal velocities and subsequently the carrier velocities increase more slowly with increasing field. This is illustrated in Figure 2-108, which shows typical measured electron drift velocity v_d versus tangential electric field strength magnitude $\mathscr{E}$ in an NMOS surface channel. Note that at low field values the velocity is proportional to the field, while at high fields the velocity approaches a constant value called the scattering-limited velocity v_{scl}. A first-order analytical approximation to this curve is

$$v_d = \frac{\mu_n \mathscr{E}}{1 + \mathscr{E}/\mathscr{E}_c} \tag{2-245}$$

where $\mathscr{E} \simeq 1.5 \times 10^6$ V/m and $\mu_n \simeq 0.07$ m^2/V · s is the low-field mobility. Equation (2-245) is also plotted in Figure 2-108. From Eq. (2-245), as $\mathscr{E} \to \infty$ we have $v_{\text{scl}} = \mu_n \mathscr{E}_c$. At the critical field value $\mathscr{E}_c$, the carrier velocity is a factor of 2 less than the low-field formula would predict. In a device with a channel length $L = 1$ μm, we need a voltage drop of only 1.5 V along the channel to have an average field equal to $\mathscr{E}_c$, and this condition is readily achieved in small MOSFETs. Similar results are found for PMOS devices.

As an example of the effects of velocity saturation on FET characteristics, we consider the example of the MOSFET. In the analysis of "Transfer Characteristics of MOS Devices" in Section 2-3-3, we now use the more general expression

$$I_D = WQ_I(y)v_d(y) \tag{2-246}$$

Substituting Eq. (2-245) into Eq. (2-246) and using

$$\mathscr{E} = \frac{dV}{dy} \tag{2-247}$$

for the magnitude of the field, we find that

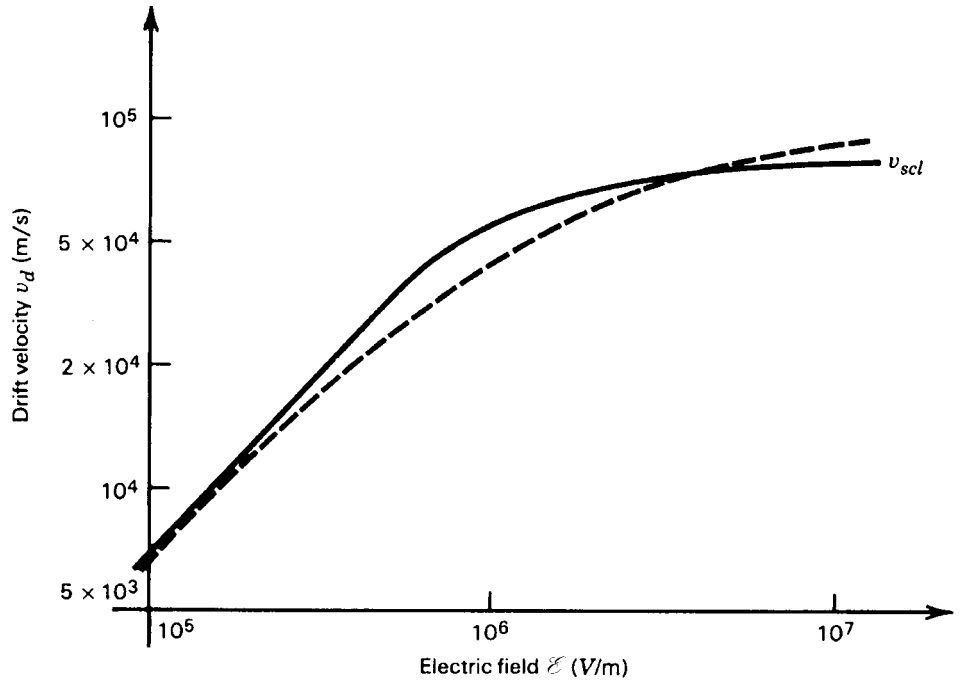


Figure 2-108 Typical measured electron drift velocity v_d versus tangent electric field $\mathscr{E}$ in an MOS surface channel (solid line). Also shown (dotted line) is the analytical approximation of Eq. (2.245) with $\mathscr{E}_c = 1.5 \times 10^6$ V/m and $\mu_n = 0.07$ m^2/V · s.

$$I_D\left(1 + \frac{1}{\mathscr{E}_c}\frac{dV}{dy}\right) = WQ_I(y)\mu_n\frac{dV}{dy} \tag{2-248}$$

Note that as $\mathscr{E}_c \to \infty$ and velocity saturation becomes negligible, Eq. (2-248) approaches the original equation (2-211).

Integrating Eq. (2-248) along the channel we obtain

$$\int_0^L I_D\left(1 + \frac{1}{\mathscr{E}_c}\frac{dV}{dy}\right)dy = \int_0^{V_{DS}} WQ_I(y)\mu_n dV \tag{2-249}$$

and thus

$$I_D = \frac{\mu_n C_{ox}}{2\left(1 + \dfrac{1}{\mathscr{E}_c}\dfrac{V_{DS}}{L}\right)}\frac{W}{L}[2(V_{GS} - V_t)V_{DS} - V_{DS}^2] \tag{2-250}$$

The quantity V_{DS}/L in Eq. (2-250) can be interpreted as the average field in the channel. If this is comparable to $\mathscr{E}_c$, the drain current for a given V_{DS} is less than the simple expression (2-213) would predict.

Equation (2-250) is valid in the triode region. The MOSFET transfer function in saturation can be obtained by using $V_{DS} = (V_{GS} - V_t)$ in Eq. (2-250) to obtain

$$I_D = \frac{k'}{2[1 + \theta(V_{GS} - V_t)]} \frac{W}{L} (V_{GS} - V_t)^2 \tag{2-251}$$

where $\theta = 1/L\mathscr{E}_c$ and has the dimension volts^{-1}. For $L = 1$ μm, a typical value is $\theta \approx 0.7\ \text{V}^{-1}$. Note that in the presence of velocity saturation effects, the device enters the saturation region for $V_{DS} < (V_{GS} - V_t)$. However, Eq. (2-251) still gives a good estimate of the saturation current.

Thus far, we have only considered the effects of the tangential field due to the V_{DS} along the channel when examining velocity saturation effects. However, there exists a normal field originating from the gate voltage that also inhibits channel carrier mobility. Since the normal field depends on the value of V_{GS}, we find that an empirical modification to θ in Eq. (2-251) can adequately model this effect. In practice, θ is determined by a best fit to measured device characteristics.

Returning to Eq. (2-251) we note that for very short channel lengths, θ becomes large and Eq. (2-251) reduces to

$$I_D \propto (V_{GS} - V_t) \tag{2-252}$$

Thus, the FET characteristics tend toward a *linear* transfer function as the channel length becomes very small (less than 1 μm).

Equation (2-251) has a simple circuit representation. Consider the circuit of Figure 2-109, where an ideal square-law MOSFET has a resistance R_{SX} in series with the source of the FET. Assume

$$I_D = \frac{\mu C_{\text{ox}}}{2} \frac{W}{L} (V'_{GS} - V_t)^2 \tag{2-253}$$

Now

$$V_{GS} = V'_{GS} + I_D R_{SX} \tag{2-254}$$

and substituting Eq. (2-254) into Eq. (2-253) we find that

$$I_D = \frac{\mu C_{\text{ox}}}{2} \frac{W}{L} (V_{GS} - I_D R_{SX} - V_t)^2 \tag{2-255}$$

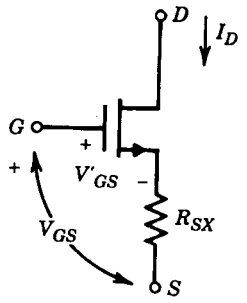


Figure 2-109 Model of velocity saturation of a MOSFET by addition of a series source resistance to an ideal square-law device.

Rearranging Eq. (2-255) we find

$$I_D = \frac{\mu C_{\text{ox}}}{2\left[1 + \mu C_{\text{ox}} \frac{W}{L} R_{SX}(V_{GS} - V_t)\right]} \frac{W}{L} (V_{GS} - V_t)^2 \tag{2-256}$$

This has exactly the same form as Eq. (2-251) if we identify

$$\theta = \mu C_{\text{ox}} \frac{W}{L} R_{SX} \tag{2-257}$$

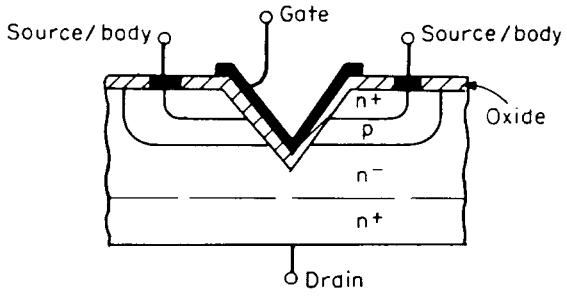
Substituting $\theta = 1/L\mathscr{E}_c$ into Eq. (2-257) we have

$$R_{SX} = \frac{1}{\mathscr{E}_c} \frac{1}{\mu C_{\text{ox}}} \frac{1}{W} \tag{2-258}$$

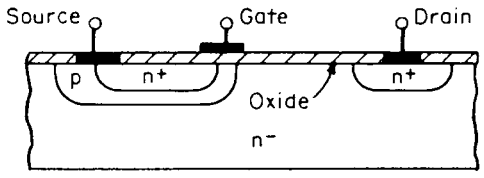
Thus, the influence of velocity saturation on the large-signal characteristics of a FET can be modeled to first order by a resistor R_{SX} in series with the source of an ideal square-law device. Note that R_{SX} varies inversely with W, as does the intrinsic physical series resistance due to source and drain contact regions. Typically, R_{SX} is larger than the physical series resistance. For $W = 2\ \mu\text{m}$, $\mu C_{\text{ox}} = 40\ \mu\text{A}/V^2$, and $\mathscr{E}_c = 1.5 \times 10^6$ V/m, we find $R_{SX} = 8\ \text{k}\Omega$.

The foregoing analysis has developed a modified large-signal mode for the MOSFET including velocity saturation effects. Small-signal MOSFET modeling for small devices can still be done using the equivalent circuit of Figure 2-107 if the values of g_m and g_{mb} are modified to account for the effects of velocity saturation using Eq. (2-251).

Figure 2-110 shows a cross section of VMOS (vertical MOS) and DMOS (double-diffused MOS) transistors. There is also a TMOS device—a Motorola variant of VMOS—and the



(a) VMOS (vertical channel MOS)



(b) DMOS (double-diffused MOS)

Figure 2-110 Cross section of (a) VMOS and (b) DMOS FETs.

now very popular LDMOS (lateral diffused MOS) power transistor (Figures 2-111 and 2-112). LDMOS FETs operate from a single supply (approximately 8 V and up) and are available for output powers surpassing 100 W at 2 GHz. The feedback capacitance is much lower than VMOS, TMOS, and earlier models, significantly improving circuit stability.

Following is a listing of Motorola high-power LDMOS transistors capable of operation up to 2 GHz:

RF High-Power LDMOS Transistors by Motorola
2.0 GHz, V_{DD} = 26 V, Class AB

Device	P_{out} Output Power (watts)	P_{in} Input Power Typical (watts)	G_{ps} (Typ.)/Freq. (dB/MHz)	IMDs (–dBc)	η Efficiency, Typ. %	θ_{JC} (°C/W)	Package/Style
MRF281S	4	0.2	13.6/2000	–30	40	8.75	458/1
MRF281Z	4	0.2	13.6/2000	–30	40	8.75	458A/1
MRF6525-5	5	0.4	11.5/2000	—	40	15	458A/1
MRF6525-10	10	1.0	10/2000	—	40	6.0	458A/1
MRF282S	10	0.5	13/2000	–30	35	2.9	458/1
MRF282Z	10	0.5	13/2000	–30	35	2.9	458A/1
MRF284	30	2.4	11.5/2000	–30	33	2.0	360B/1
MRF284S	30	2.4	11.5/2000	–30	33	2.0	360C/1
MRF286	60	4.75	11.4/2000	–30	33	.73	465/1
MRF286S	60	4.75	11.4/2000	–30	33	.73	465A/1

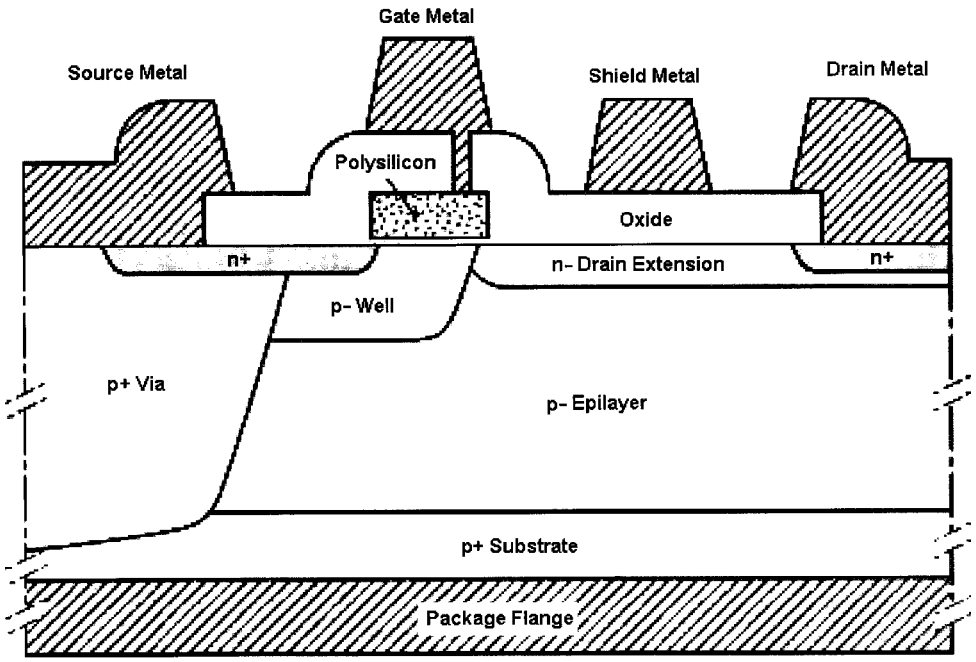


Figure 2-111 Cross section of a Philips LDMOS FET [18].

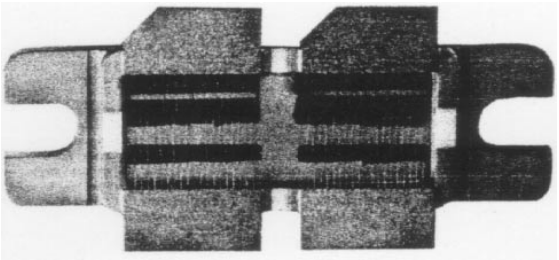


Figure 2-112 Photo of the underside of an LDMOS FET.

2-3-6 Small-Signal Models of MOSFETs

Subthreshold Conduction in MOSFETs. The MOSFET analysis of Section 2-3-3 focused on the normal region of operation where this is a well-defined conducting channel under the gate. Changes in the gate voltage are assumed to cause changes in the channel charge only, and not in the depletion region below. However, for gate voltages less than the extrapolated threshold voltage V_t, the applied gate potential affects both the depletion-region charge and the channel charge, which is then very small but not zero. The device can thus conduct finite (but small) current for $V_{GS} < V_t$, so that Eq. (2-215) is not valid in this region. The electrons in the n^+ source region of an NMOS transistor can overcome the potential barrier to the p-type substrate and enter the channel region. This process is very similar to the turn-on of a bipolar transistor, and in fact the MOSFET characteristics in this subthreshold region (also called weak inversion) are very similar to those of a bipolar transistor. Analysis [19] shows that in the subthreshold region of the MOSFET characteristics can be defined by the equation

$$I_D = k_x \frac{W}{L} e^{V_{GS}/nV_T}(1 - e^{-V_{DS}/V_T}) \tag{2-259}$$

where k_x depends on process parameters and $n \simeq 1.5$. The value of n is not equal to unity in this case (as it is for the bipolar transistor) because the applied voltage V_{GS} appears partially at the silicon surface and partially across the depletion layer.

To illustrate this effect, we show measured NMOS characteristics plotted on three different scales in Figure 2-113. In Figure 2-113a we show the transfer characteristic in the forward-active region plotted on linear scales. For this device W = 20 μm and L = 20 μm and short-channel effects are negligible. The same data are plotted in Figure 2-113b as $\sqrt{I_D}$ versus V_{GS}. The resulting straight line shows that the device characteristic is close to an ideal square law. Plots like this are commonly used to obtain V_t by extrapolation (0.7 V in this case) and also k' from the slope of the curve (54 μA/V^2 in this case). Note that near the threshold voltage, the curve deviates from the straight line representing the square law. This is the subthreshold region. The data are plotted a third time in Figure 2-113c on log–linear scales. The straight line obtained for $V_{GS} < V_t$ fits Eq. (2-259) with n = 1.5. At currents below 10^{-12} A, the effect of leakage currents becomes evident.

The major application of subthreshold operation is in very low power applications at relatively low signal frequencies. The limitation to low signal frequencies occurs because the MOSFET f_T becomes very small. Since the device capacitances at very low bias currents

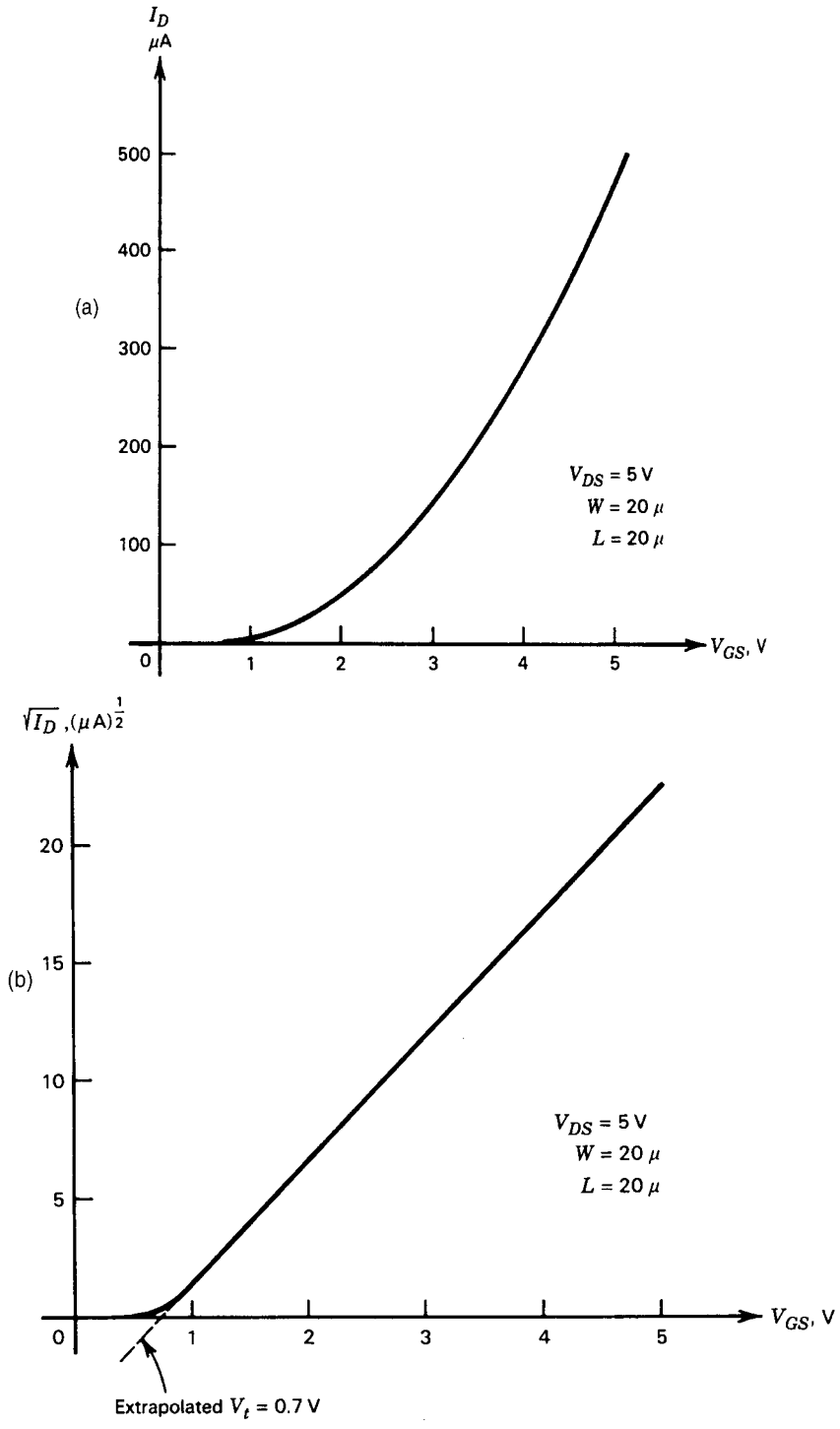


Figure 2-113 (a) Measured NMOS transfer characteristic in the forward-active region. (b) Data from part (a) plotted as $\sqrt{I_D}$ versus V_{GS} showing the square-law characteristic. (c) Data from part (a) plotted on log–linear scales showing the exponential characteristic of the subthreshold region.

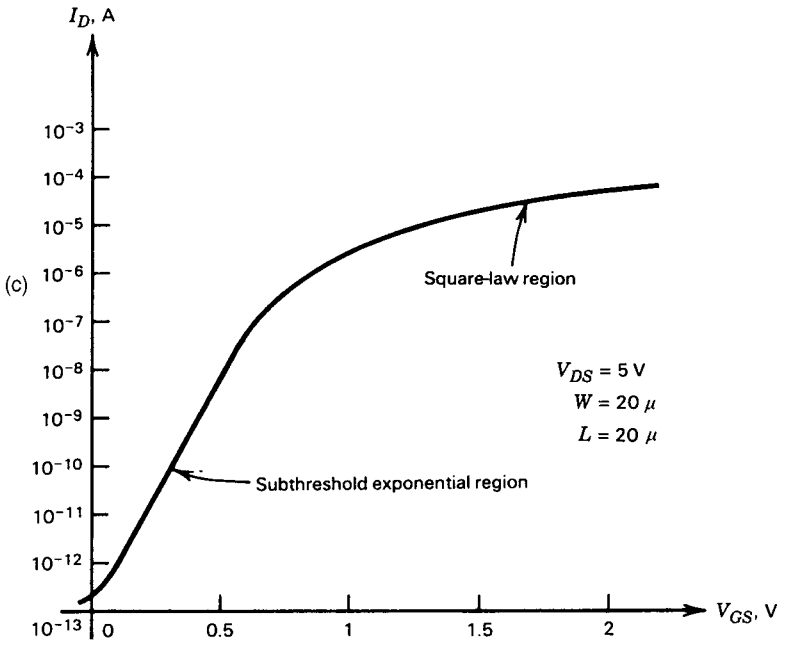


Figure 2-113. (*Continued*)

are essentially fixed, and the small-signal g_m calculated from Eq. (2-259) becomes proportional to I_D, we see that the value of f_T becomes very small at very low values of I_D.

Substrate Flow in MOSFETs. As the reverse-bias voltages on the device are increased, carriers traversing the depletion regions gain sufficient energy to create new hole–electron pairs in lattice collisions by a process known as *impact ionization.* Eventually, at high enough bias voltages, the process results in large avalanche currents. For collector–base bias voltages well below the breakdown value, a small enhanced current flow may occur across the collector–base junction due to this process with little apparent effect on the device characteristics.

Impact ionization also occurs in MOSFETs but has a significantly different effect on the device characteristics. This is because the channel electrons (for the case of NMOS) create hole–electron pairs in lattice collisions in the drain depletion region, and some of the resulting holes then flow to the substrate, creating a substrate current. (The electrons created

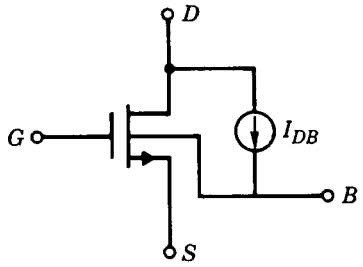


Figure 2-114 Representation of impact ionization in a MOSFET by a drain–substrate current generator.

in the process flow out the drain terminal.) The carriers created by impact ionization are thus not confined within the device as they are in a bipolar transistor. The effect of this phenomenon can be modeled by inclusion of a controlled current generator I_{DB} from drain to substrate, as shown in Figure 2-114 for an NMOS device. The magnitude of this substrate current depends on the voltage across the drain depletion region (which determines the energy of the ionizing channel electrons), and also on the drain current (which is the rate at which channel electrons enter the depletion region). It has been found empirically [20] that the current I_{DB} can be expressed as

$$I_{DB} = K_1(V_{DS} - V_{DS,\text{sat}})I_D e^{-[K_2/(V_{DS}-V_{DS,\text{sat}})]} \tag{2-260}$$

where K_1 and K_2 are process-dependent parameters and $V_{DS,\text{sat}}$ is the value of V_{DS} where the drain characteristics enter the saturation region. Typical values for NMOS devices are $K_1 = 5\ \text{V}^{-1}$ and $K_2 = 30$ V. The effect is generally much less significant in PMOS devices because the holes carrying the charge in the channel are much less efficient in creating hole–electron pairs than are energetic electrons.

The major impact of this phenomenon on circuit performance is that a parasitic resistance from drain to substrate now exists. Since the substrate of an NMOS device in a *p*-substrate process is an ac ground (the common substrate terminal must always be connected to the most negative supply voltage in the circuit), the parasitic resistance shunts the drain to ground and can be a limiting factor in many circuit designs. Differentiating Eq. (2-260) we find for the drain–substrate small-signal conductance

$$g_{db} = \frac{\partial I_{DB}}{\partial V_D}$$

$$= K_2 \frac{I_{DB}}{(V_{DS} - V_{DS,\text{sat}})^2} \tag{2-261}$$

where the gate and source are assumed held at fixed potentials.

In IC applications for RF/UHF circuits, many not-yet-solved problems remain:

1. The flicker corner frequency for MOS devices is somewhere between 100 kHz and 1 MHz.
2. Because of backplane problems in ICs, such as the substrate capacitance, also relative to the source, the performance of an IC-MOS/CMOS is not as good as expected if these effects are not considered.
3. Discrete devices, such as *n*-channel MOS devices, which are also available as dual-gate MOSFETs, can work well up to 1200 MHz. Such performance is not yet shown in silicon-based RFICs.
4. For RF power devices, such as LDMOS and others mentioned, promising results are only shown for discrete devices. There are no equivalent ICs on the market yet.

It will be interesting to see how this technology will work itself upward in frequency. Its main advantage is the BiCMOS process, which allows the mixing of analog and digital circuits nicely. Heterojunction transistors, such as SiGe, are inherently much more linear

and, for the same performance, require less current than MOS technology. On the other hand, we have not seen large-scale production SiGe-based ICs.

In the following five pages, we reproduce the datasheet for the Siemens (now Infineon Technologies) BF999 single-gate MOSFET.* Above 500 MHz, a dual-gate (also referred to as tetrode) MOSFET, such as the BF998, is preferred over a single-gate device. The main reasons for this are:

1. Internally, it is equivalent to two single-gate devices connected in a cascade arrangement, with the output device operating in the grounded-gate configuration, which entirely avoids the Miller effect and therefore remains stable at higher frequencies, operating up to 1.2 GHz.
2. The second gate can be used for AGC. As the Gate 2 voltage changes, the gain varies heavily.
3. By applying a local oscillator signal to Gate 2, this configuration can be used as a mixer.

The major drawbacks of these devices are their wide tolerances and temperature sensitivity. They follow to some degree the CMOS design but seem to have better performance.

The BF995, also called a tetrode, is in reality a dual-gate MOSFET specified to only 300 MHz; however, we are reproducing its datasheet* because the one more applicable for wireless circuits, the BF1005, does not have such a complete set of data.

To demonstrate the capabilities of low-voltage LDMOS transistors, we show the MRF5003 datasheet.† This device is intended for low-voltage applications down to 7.5 V. To the best of our knowledge, this is the lowest-voltage LDMOS transistor on the market. Other members of this family, which produce 5 W (MRF5005) and 15 W (MRF5015), begin to require higher supply voltages. Just to make sure it is not forgotten, the advantage of LDMOS is, besides lower-voltage operation, much-reduced feedback capacitance. LDMOS devices are now available that operate above 1 GHz and produce 60 W. There is now some competition with the recently introduced NEC GaAsFET, which goes up to 100 W in a similar frequency range.

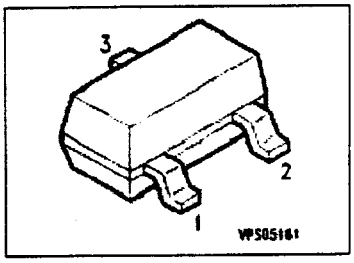
*Reproduced with permission.
†Copyright of Motorola, used by permission.

SIEMENS

Silicon N Channel MOSFET Triode

BF 999

- For high-frequency stages up to 300 MHz, preferably in FM applications



Type	Marking	Ordering Code (tape and reel)	Pin Configuration 1	2	3	Package
BF 999	LB	Q62702-F1132	G	D	S	SOT-23

Maximum Ratings

Parameter	Symbol	Values	Unit
Drain-source voltage	V_{DS}	20	V
Drain current	I_D	30	mA
Gate-source peak current	$\pm I_{GSM}$	10	
Total power dissipation, $T_A \leq 60$ °C	P_{tot}	200	mW
Storage temperature range	T_{stg}	– 55 ... + 150	°C
Channel temperature	T_{ch}	150	

Thermal Resistance

Junction - ambient [1]	$R_{th\,JA}$	≤ 450	K/W

[1] Package mounted on alumina 15 mm × 16.7 mm × 0.7 mm.

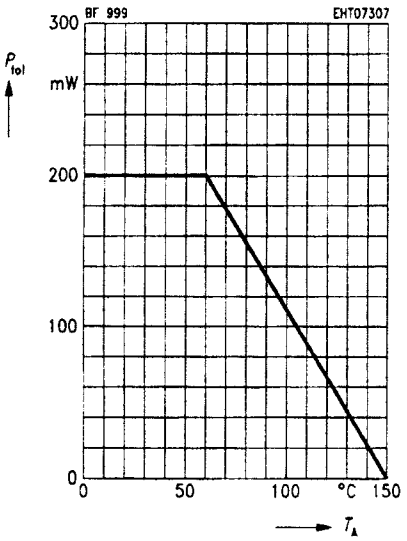
SIEMENS **BF 999**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

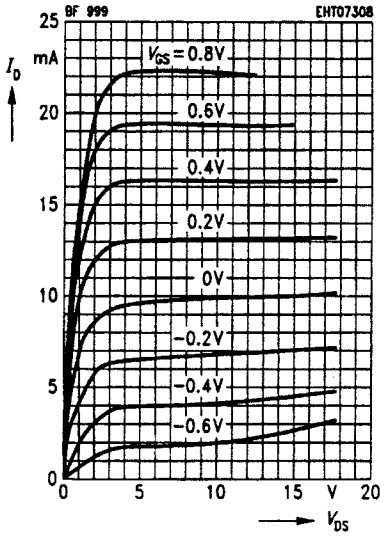
Parameter	Symbol	Values			Unit
		min.	typ.	max.	
DC Characteristics					
Drain-source breakdown voltage I_D = 10 µA, – V_{GS} = 4 V	$V_{(BR)\,DS}$	20	–	–	V
Gate-source breakdown voltage ± I_{GS} = 10 mA, V_{DS} = 0	± $V_{(BR)\,GSS}$	6.5	–	12	
Gate-source leakage current ± V_{GS} = 5 V, V_{DS} = 0	± I_{GSS}	–	–	50	nA
Drain current V_{DS} = 10 V, V_{GS} = 0	I_{DSS}	5	–	18	mA
Gate-source pinch-off voltage V_{DS} = 10 V, I_D = 20 µA	– $V_{GS\,(p)}$	–	–	2.5	V
AC Characteristics					
Forward transconductance V_{DS} = 10 V, I_D = 10 mA, f = 1 kHz	g_{fs}	14	16	–	mS
Gate input capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{gss}	–	2.5	–	pF
Reverse transfer capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{dg}	–	25	–	fF
Output capacitance V_{DS} = 10 V, I_D = 10 mA, f = 1 MHz	C_{dss}	–	1	–	pF
Power gain (test circuit) V_{DS} = 10 V, I_D = 10 mA, f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS	G_p	–	25	–	dB
Noise figure (test circuit) V_{DS} = 10 V, I_D = 10 mA, f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS	F	–	1	–	

SIEMENS **BF 999**

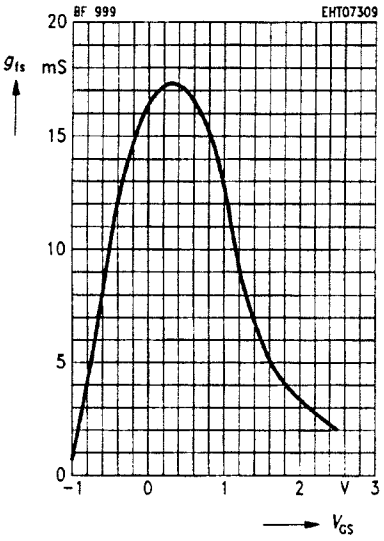
Total power dissipation $P_{tot} = f(T_A)$



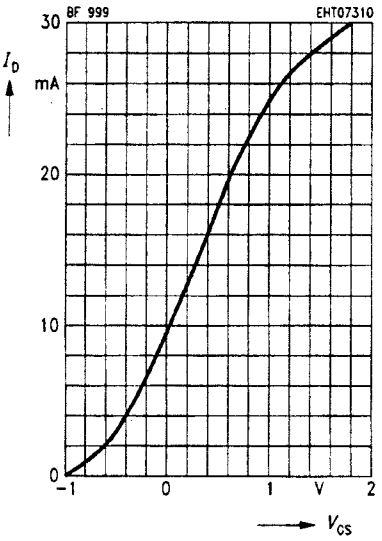
Output characteristics $I_D = f(V_{DS})$



Gate transconductance $g_{fs} = f(V_{GS})$
V_{DS} = 10 V, I_{DSS} = 10 mA, f = 1 kHz

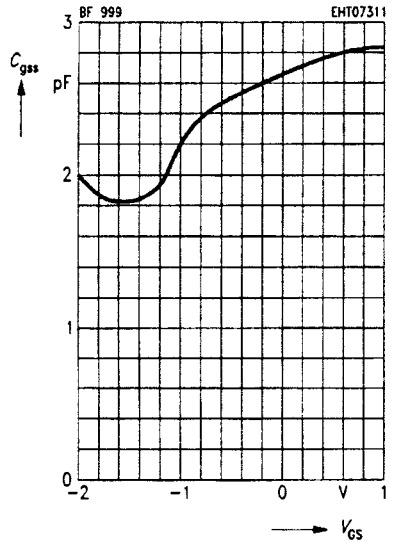


Drain current $I_D = f(V_{GS})$
V_{DS} = 10 V

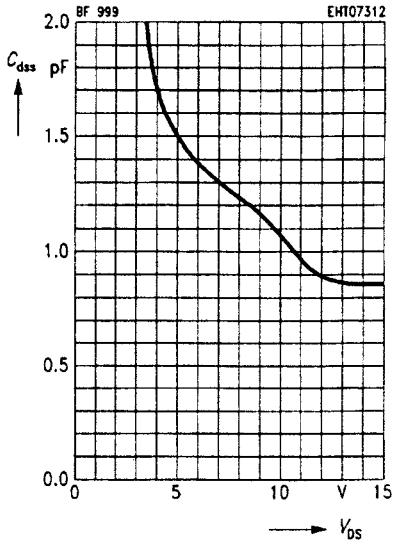


SIEMENS **BF 999**

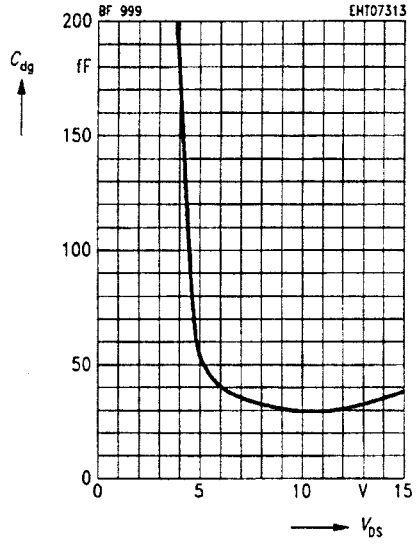
Gate input capacitance $C_{gss} = f(V_{GS})$
V_{DS} = 10 V, I_{DSS} = 10 mA, f = 1 MHz



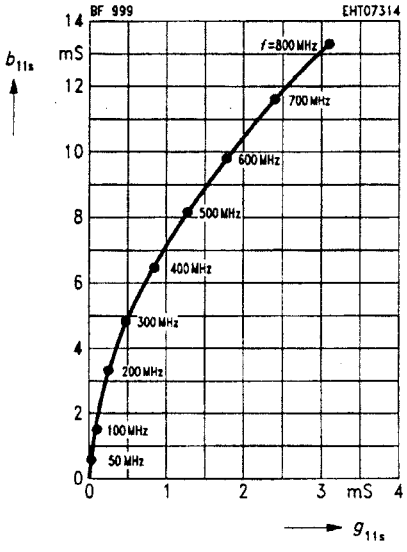
Output capacitance $C_{dss} = f(V_{DS})$
V_{GS} = 0, I_{DSS} = 10 mA, f = 1 MHz



Reverse transfer capacitance
$C_{dg} = f(V_{DS})$
I_{DSS} = 10 mA, f = 1 MHz, V_{GS} = 0

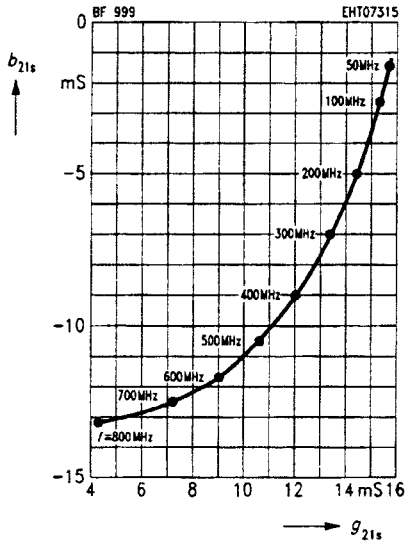


Gate input admittance y_{11s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)

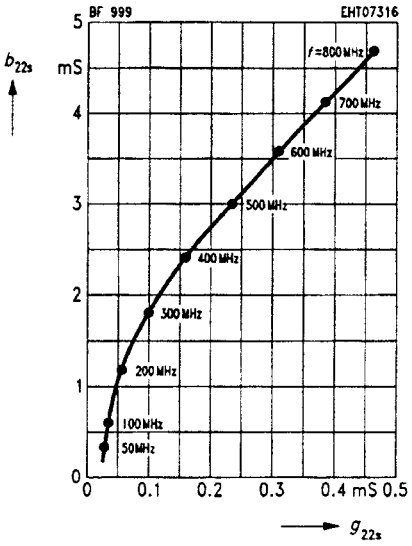


SIEMENS **BF 999**

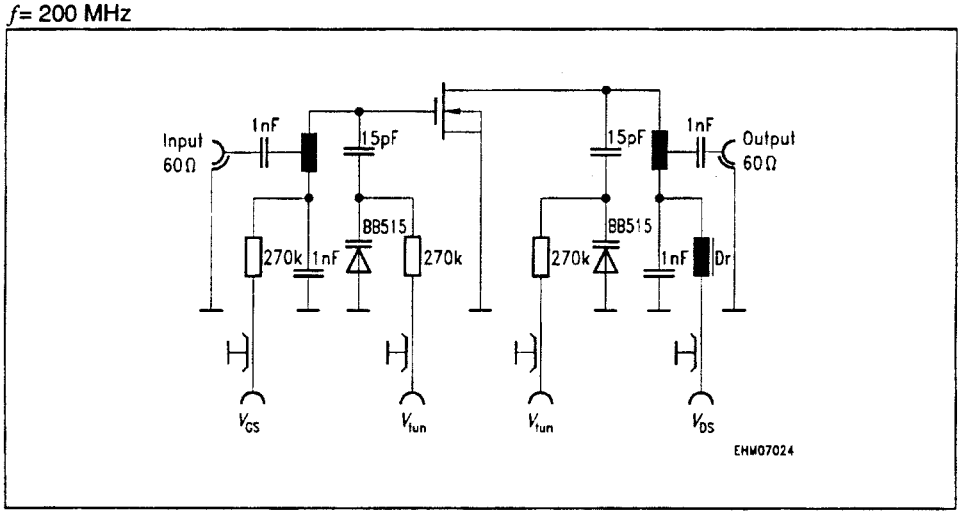
Gate forward transfer admittance y_{21s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)



Output admittance y_{22s}
V_{DS} = 10 V, V_{GS} = 0,
I_{DSS} = 10 mA, (common-source)



Test circuit for power gain and noise figure
f= 200 MHz

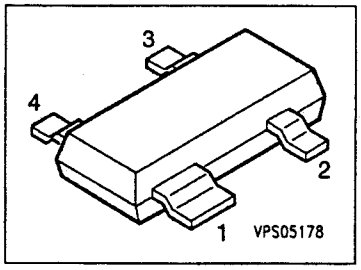


SIEMENS

Silicon N Channel MOSFET Tetrode

BF 995

- For input and mixer stages in FM and VHF TV tuners



Type	Marking	Ordering Code (tape and reel)	Pin Configuration 1	2	3	4	Package
BF 995	MB	Q62702-F936	S	D	G_2	G_1	SOT-143

Maximum Ratings

Parameter	Symbol	Values	Unit
Drain-source voltage	V_{DS}	20	V
Drain current	I_D	30	mA
Gate 1/gate 2 peak source current	$\pm I_{G1/2SM}$	10	
Total power dissipation, T_S < 76 °C	P_{tot}	200	mW
Storage temperature range	T_{stg}	– 55 ... + 150	°C
Channel temperature	T_{ch}	150	

Thermal Resistance

Junction - soldering point	$R_{th\,JS}$	< 370	K/W

SIEMENS **BF 995**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

Parameter	Symbol	Values			Unit
		min.	typ.	max.	
DC Characteristics					
Drain-source breakdown voltage I_D = 10 µA, – V_{G1S} = – V_{G2S} = 4 V	$V_{(BR)\,DS}$	20	–	–	V
Gate 1 source breakdown voltage ± I_{G1S} = 10 mA, V_{G2S} = V_{DS} = 0	± $V_{(BR)\,G1SS}$	8.5	–	14	
Gate 2 source breakdown voltage ± I_{G2S} = 10 mA, V_{G1S} = V_{DS} = 0	± $V_{(BR)\,G2SS}$	8.5	–	14	
Gate 1 source leakage current ± V_{G1S} = 5 V, V_{G2S} = V_{DS} = 0	± I_{G1SS}	–	–	50	nA
Gate 2 source leakage current ± V_{G2S} = 5 V, V_{G1S} = V_{DS} = 0	± I_{G2SS}	–	–	50	
Drain current V_{DS} = 15 V, V_{G1S} = 0, V_{G2S} = 4 V	I_{DSS}	4	–	20	mA
Gate 1 source pinch-off voltage V_{DS} = 15 V, V_{G2S} = 4 V, I_D = 20 µA	– $V_{G1S\,(p)}$	–	–	2.5	V
Gate 2 source pinch-off voltage V_{DS} = 15 V, V_{G1S} = 0, I_D = 20 µA	– $V_{G2S\,(p)}$	–	–	2.0	

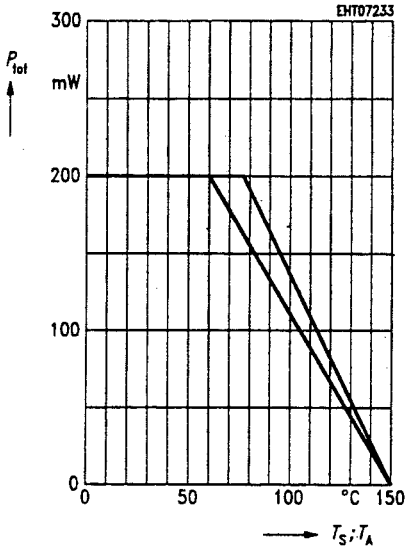
SIEMENS **BF 995**

Electrical Characteristics
at T_A = 25 °C, unless otherwise specified.

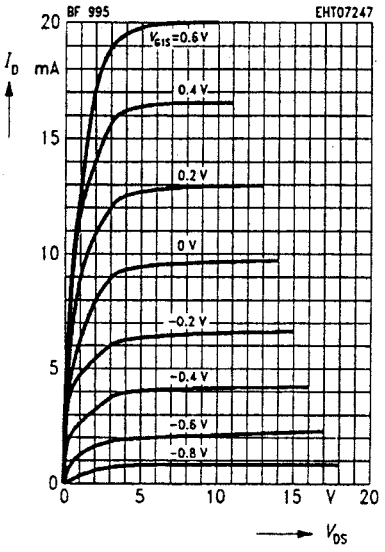
Parameter	Symbol	Values			Unit
		min.	typ.	max.	
AC Characteristics					
Forward transconductance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 kHz	g_{fs}	12	17	–	mS
Gate 1 input capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{g1ss}	–	3.6	–	pF
Gate 2 input capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{g2ss}	–	1.6	–	
Feedback capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{dg1}	–	25	–	fF
Output capacitance V_{DS} = 15 V, I_D = 10 mA, V_{G2S} = 4 V, f = 1 MHz	C_{dss}	–	1.6	–	pF
Power gain V_{DS} = 15 V, I_D = 10 mA f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS 2 Δf = 12 MHz (see test circuit 1)	G_{ps}	–	23	–	dB
Noise figure V_{DS} = 15 V, I_D = 10 mA f = 200 MHz, G_G = 2 mS, G_L = 0.5 mS (see test circuit 1)	F	–	1.1	–	
Gain control range V_{DS} = 15 V, V_{G2S} = 4 ... – 2 V, f = 200 MHz (see test circuit 1)	ΔG_{ps}	–	50	–	
Mixer gain (additive) V_{DS} = 15 V, V_{G2S} = 6 V, R_S = 220 Ω f = 200 MHz, f_F = 36 MHz 2 Δf_F = 5 MHz, V_{osc} = 0.5 V (see test circuit 2)	G_{psc}	–	16	–	
Mixer gain (multiplicative) V_{DS} = 15 V, V_{G1S} = 1.7 V, V_{G2S} = 2.5 V R_S = 220 Ω, f = 200 MHz, f_F = 36 MHz 2 Δf_F = 5 MHz, V_{osc} = 2 V (see test circuit 3)	G_{psc}	–	18	–	

SIEMENS **BF 995**

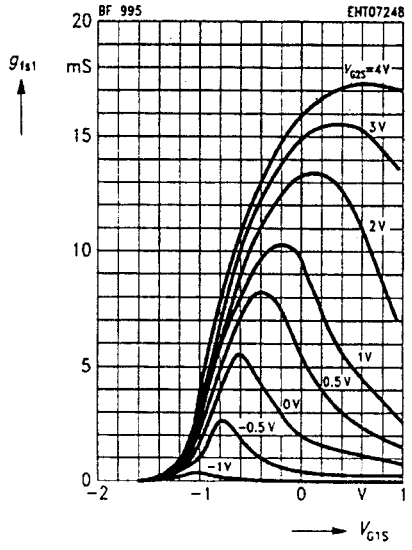
Total power dissipation $P_{tot} = f(T_A)$



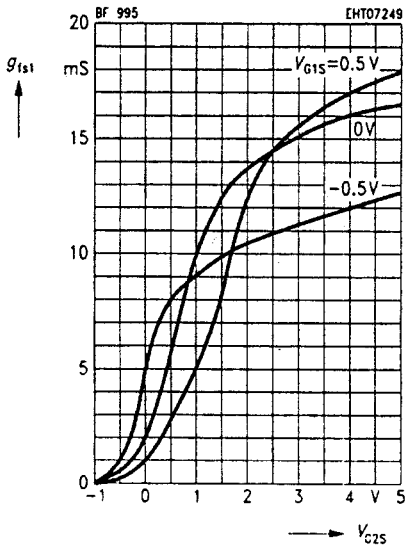
Output characteristics $I_D = f(V_{DS})$
V_{G2S} = 4 V



Gate 1 forward transconductance $g_{fs1} = f(V_{G1S})$
V_{DS} = 15 V, I_{DSS} = 10 mA, f = 1 kHz

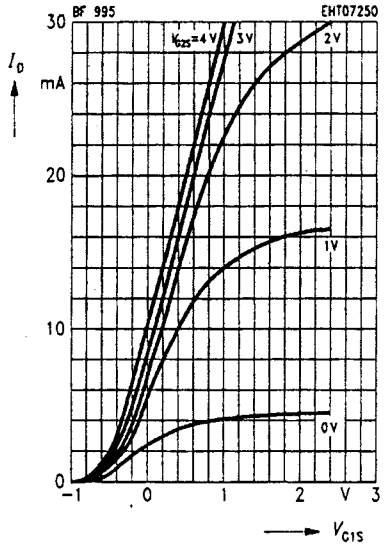


Gate 1 forward transconductance $g_{fs1} = f(V_{G2S})$
V_{DS} = 15 V, I_{DSS} = 10 mA, f = 1 kHz

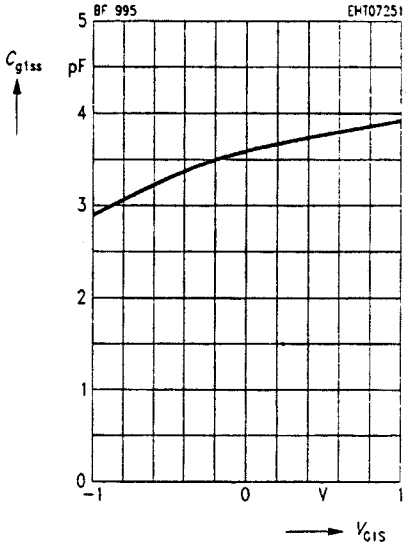


SIEMENS **BF 995**

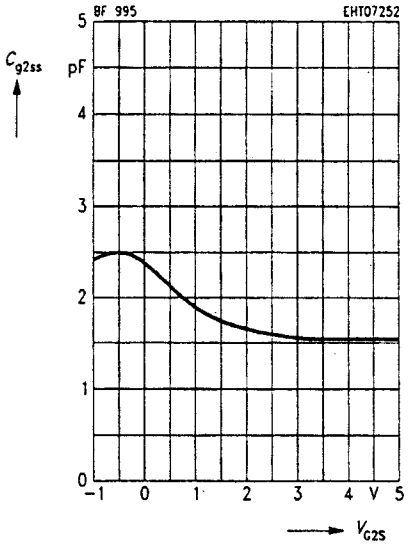
Drain current $I_D = f(V_{G1S})$
V_{DS} = 15 V



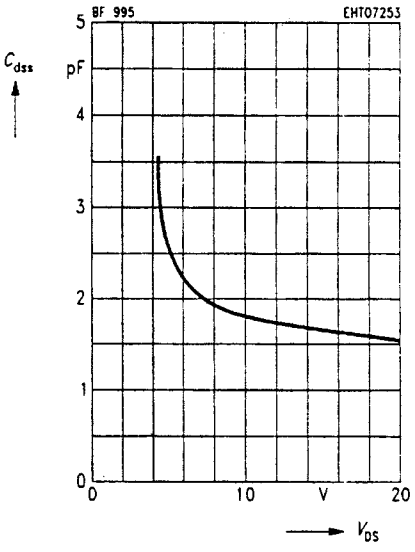
Gate 1 input capacitance $C_{g1ss} = f(V_{G1S})$
V_{G2S} = 4 V, V_{DS} = 15 V
I_{DSS} = 10 mA, f = 1 MHz



Gate 2 input capacitance $C_{g2ss} = f(V_{G2S})$
V_{G1S} = 0 V, V_{DS} = 15 V
I_{DSS} = 10 mA, f = 1 MHz

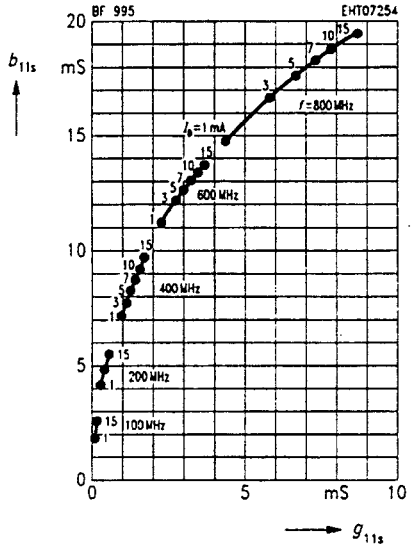


Output capacitance $C_{dss} = f(V_{DS})$
V_{G1S} = 0 V, V_{G2S} = 4 V
I_{DSS} = 10 mA, f = 1 MHz

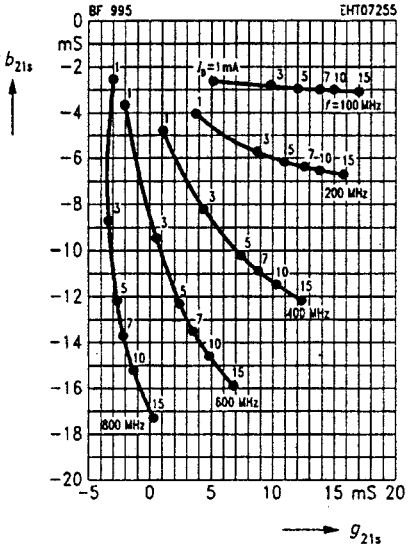


SIEMENS **BF 995**

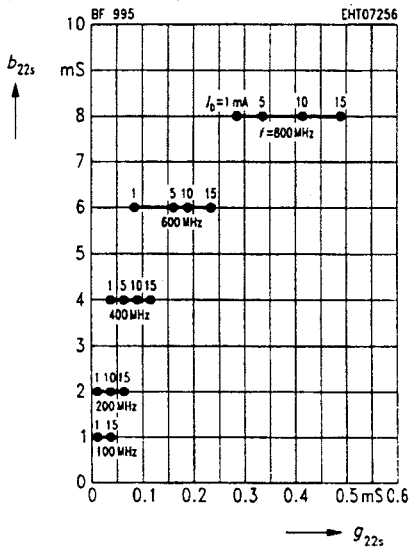
Gate 1 input admittance y_{11s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)



Gate 1 forward transfer admittance y_{21s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)

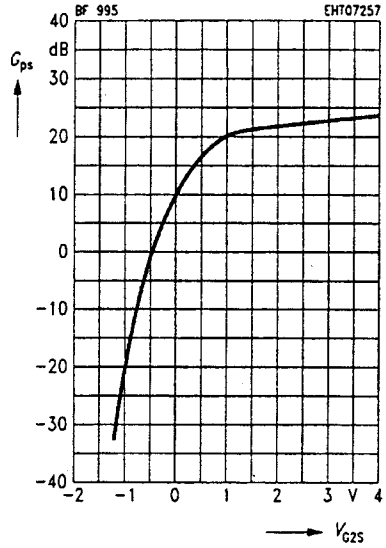


Output admittance y_{22s}
V_{DS} = 15 V, V_{G2S} = 4 V
(common source)

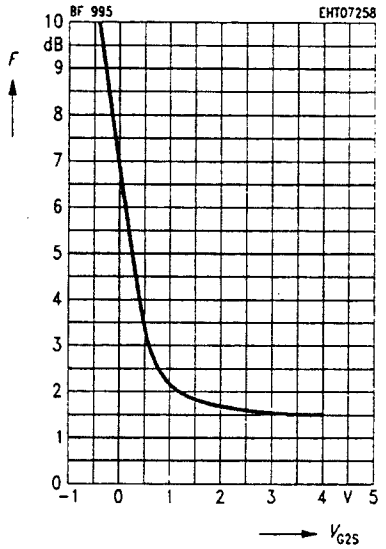


SIEMENS **BF 995**

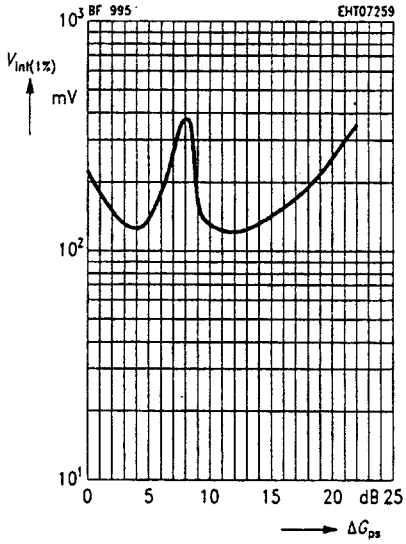
Power gain $G_{ps} = f(V_{G2S})$
V_{DS} = 15 V, V_{G1S} = 0 V, I_{DSS} = 10 mA
f = 200 MHz (see test circuit 1)



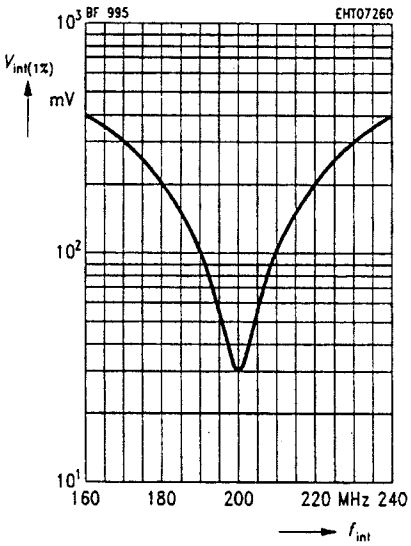
Noise figure $F = f(V_{G2S})$
V_{DS} = 15 V, V_{G1S} = 0 V, I_{DSS} = 10 mA
f = 200 MHz (see test circuit 1)



Interference voltage for 1% cross modulation $V_{int\,(1\%)} = f(\Delta G_{ps})$[1)]
V_{DS} = 15 V, V_{G1S} = 0, f = 200 MHz
f_{int} = 221 MHz (see test circuit 1)



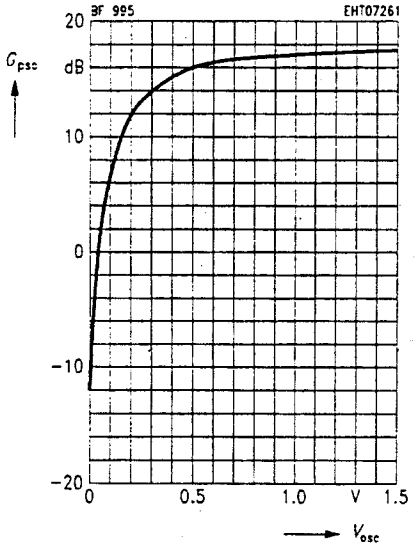
Interference voltage for 1% cross modulation $V_{int\,(1\%)} = f(f_{int})$[1)]
V_{DS} = 15 V, V_{G2S} = 4 V, V_{G1S} = 0
f = 200 MHz (see test circuit 1)



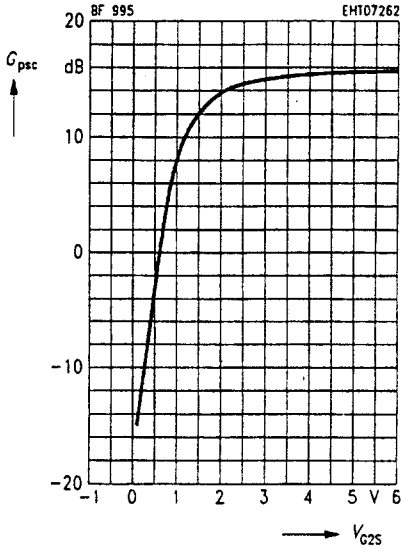
1) For footnote refer to the last page of this data sheet.

SIEMENS **BF 995**

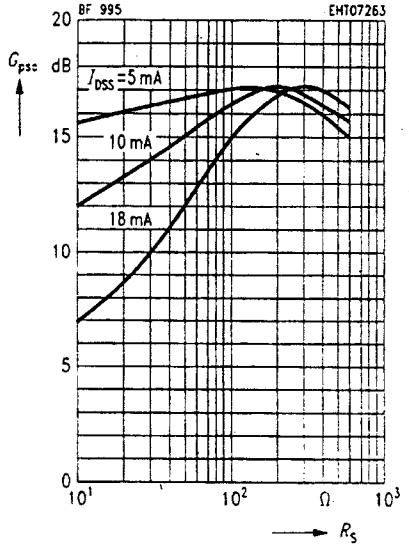
Mixer gain (additive) $G_{psc} = f(V_{osc})$
V_D = 15 V, V_{G1S} = 0, V_{G2S} = 6 V
R_S = 220 Ω, I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)



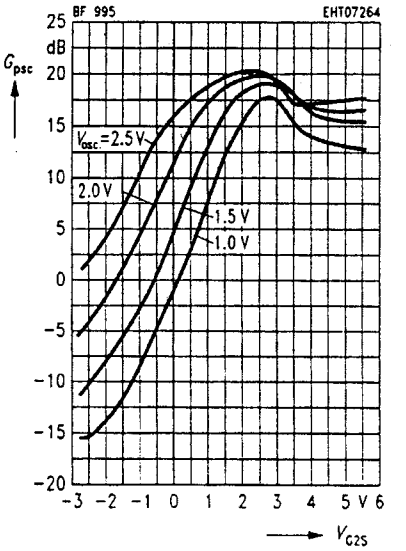
Mixer gain (additive) $G_{psc} = f(V_{G2S})$
V_D = 15 V, V_{G1S} = 0, R_S = 220 Ω
V_{osc} = 0.5 V, I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)



Mixer gain (additive) $G_{psc} = f(R_S)$
V_D = 15 V, V_{G1S} = 0, V_{G2S} = 6 V
V_{osc} = 0.5 V, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 2)

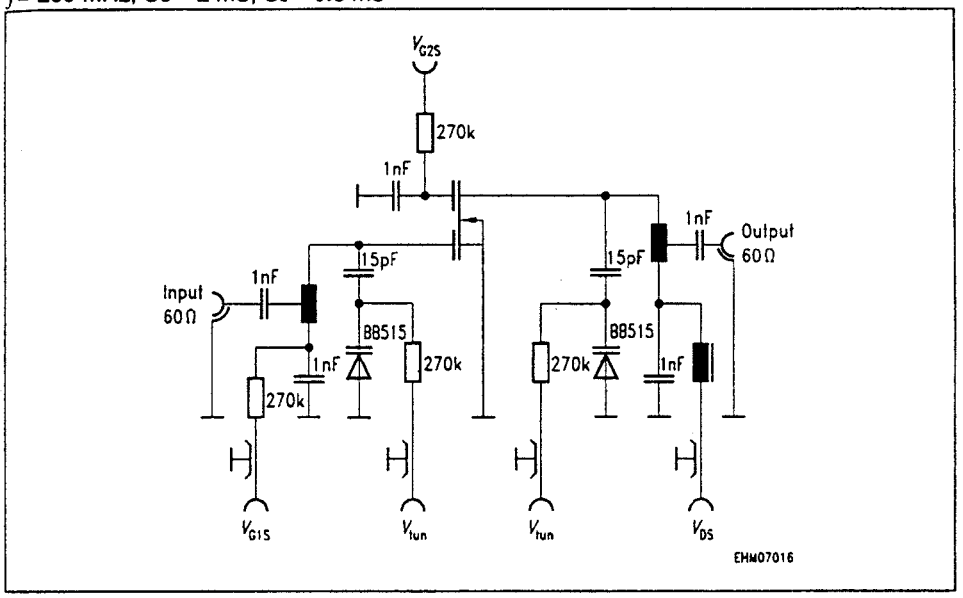


Mixer gain (multiplicative) $G_{psc} = f(V_{G2S})$
V_D = 15 V, V_{G1S} = 1.7 V, R_S = 200 Ω
I_{DSS} = 10 mA, f = 200 MHz
f_{IF} = 36 MHz (see test circuit 3)

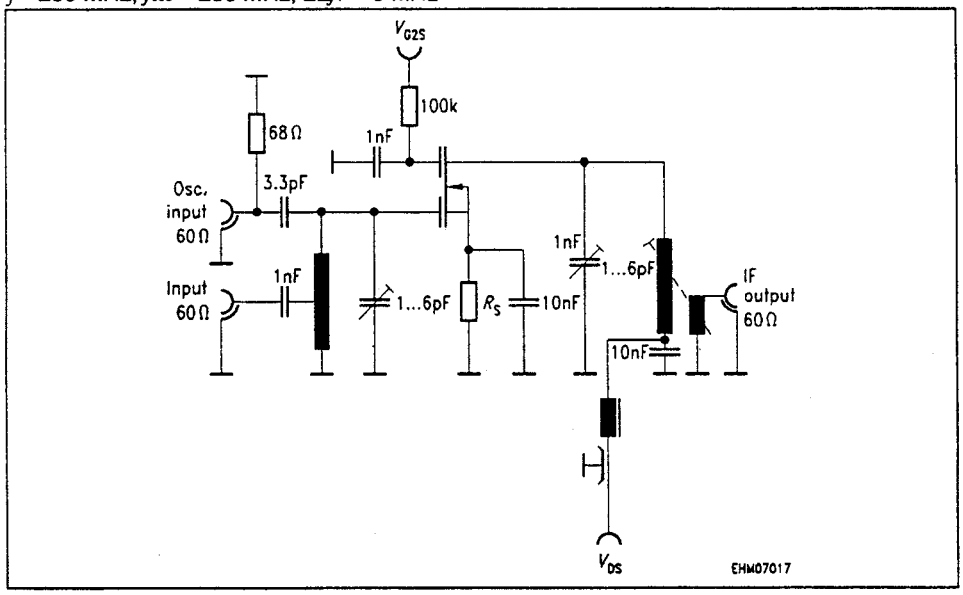


SIEMENS **BF 995**

Test circuit 1 for power gain, noise figure and cross modulation
f= 200 MHz, G_S = 2 mS, G_L = 0.5 mS

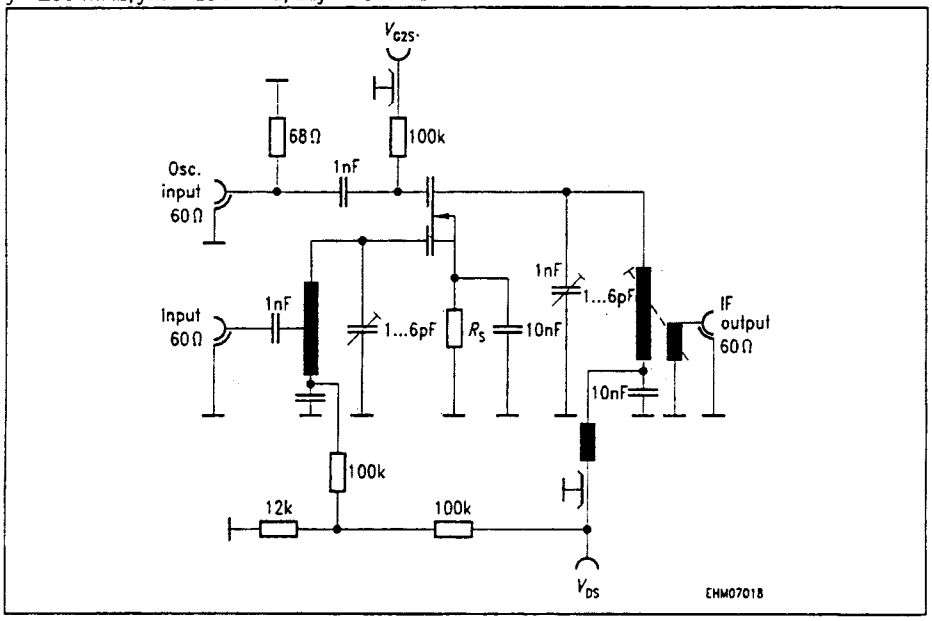


Test circuit 2 for mixer gain (additive)
f= 200 MHz, f_{osc} = 236 MHz, $2\Delta f_F$ = 5 MHz



SIEMENS **BF 995**

Test circuit 3 for mixer gain (multiplicative)
f = 200 MHz, f_{osc} = 236 MHz, $2\Delta f_F$ = 5 MHz



1) $V_{int\,(1\%)}$ is the rms value of half the emf (terminal voltage at matching) of a 100 % sine modulated TV carrier at an internal generator resistance of 60 Ω, causing 1 % amplitude modulation on the active carrier.

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF5003/D

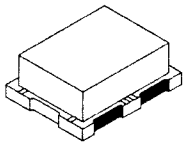
The RF MOSFET Line
RF Power Field Effect Transistor
N–Channel Enhancement–Mode

MRF5003

3.0 W, 7.5 V, 512 MHz
N–CHANNEL
BROADBAND
RF POWER FET

The MRF5003 is designed for broadband commercial and industrial applications at frequencies to 520 MHz. The high gain and broadband performance of this device makes it ideal for large–signal, common source amplifier applications in 7.5 Volt and 12.5 Volt mobile, portable, and base station FM equipment.

- Guaranteed Performance at 512 MHz, 7.5 Volts
  Output Power = 3.0 Watts
  Power Gain = 9.5 dB
  Efficiency = 45%
- Characterized with Series Equivalent Large–Signal Impedance Parameters
- S–Parameter Characterization at High Bias Levels
- Excellent Thermal Stability
- All Gold Metal for Ultra Reliability
- Capable of Handling 20:1 VSWR, @ 15.5 Vdc, 512 MHz, 2.0 dB Overdrive
- Suitable for 12.5 Volt Applications
- True Surface Mount Package
- Available in Tape and Reel by Adding R1 Suffix to Part Number.
  R1 Suffix = 500 Units per 16 mm, 7 inch Reel.
- Circuit board photomaster available upon request by contacting
  RF Tactical Marketing in Phoenix, AZ.



CASE 430–01, STYLE 2

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Drain–Source Voltage	V_{DSS}	36	Vdc
Drain–Gate Voltage (R_{GS} = 1.0 Meg Ohm)	V_{DGR}	36	Vdc
Gate–Source Voltage	V_{GS}	±20	Vdc
Drain Current — Continuous	I_D	1.7	Adc
Total Device Dissipation @ T_C = 25°C Derate above 25°C	P_D	12.5 0.07	Watts W/°C
Storage Temperature Range	T_{stg}	–65 to +150	°C
Operating Junction Temperature	T_J	200	°C

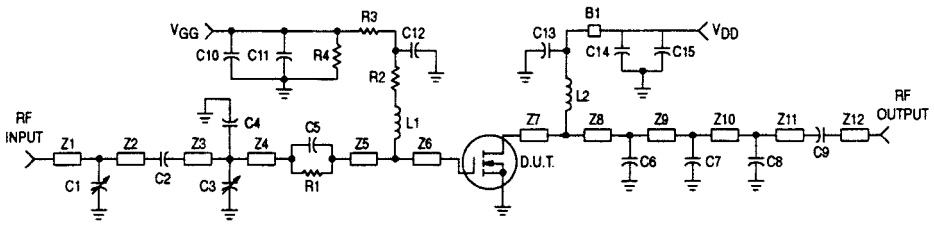
THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	14	°C/W

NOTE – **CAUTION** – MOS devices are susceptible to damage from electrostatic charge. Reasonable precautions in handling and packaging MOS devices should be observed.

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted.)

Characteristic		Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS						
Drain–Source Breakdown Voltage (V_{GS} = 0, I_D = 2.5 mAdc)		$V_{(BR)DSS}$	36	—	—	Vdc
Zero Gate Voltage Drain Current (V_{DS} = 15 Vdc, V_{GS} = 0)		I_{DSS}	—	—	1.0	mAdc
Gate–Source Leakage Current (V_{GS} = 20 Vdc, V_{DS} = 0)		I_{GSS}	—	—	1.0	μAdc
ON CHARACTERISTICS						
Gate Threshold Voltage (V_{DS} = 10 Vdc, I_D = 5.0 mAdc)		$V_{GS(th)}$	1.25	2.25	3.5	Vdc
Drain–Source On–Voltage (V_{GS} = 10 Vdc, I_D = 0.5 Adc)		$V_{DS(on)}$	—	—	0.375	Vdc
Forward Transconductance (V_{DS} = 10 Vdc, I_D = 0.5 Adc)		g_{fs}	0.6	—	—	mho
DYNAMIC CHARACTERISTICS						
Input Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{iss}	—	16.5	—	pF
Output Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{oss}	—	37	—	pF
Reverse Transfer Capacitance (V_{DS} = 12.5 Vdc, V_{GS} = 0, f = 1.0 MHz)		C_{rss}	3.5	4.4	5.4	pF
FUNCTIONAL TESTS (In Motorola Test Fixture)						
Common–Source Amplifier Power Gain (V_{DD} = 7.5 Vdc, P_{out} = 3.0 W, I_{DQ} = 50 mA)	f = 512 MHz f = 175 MHz	G_{ps}	 9.5 —	 10.5 15	 — —	dB
Drain Efficiency (V_{DD} = 7.5 Vdc, P_{out} = 3.0 W, I_{DQ} = 50 mA)	f = 512 MHz f = 175 MHz	h	 45 —	 50 55	 — —	%



C1, C3, C7, C8	0 to 20 pF Johanson	Z1	0.350″ x 0.08″ Microstrip
C2, C9	56 pF, 100 mil Chip	Z2	0.190″ x 0.08″ Microstrip
C4	10 pF, 100 mil Chip	Z3	0.800″ x 0.08″ Microstrip
C5	47 pF, Miniature Clamped Mica Capacitor	Z4	0.380″ x 0.08″ Microstrip
C6	22 pF, 100 mil Chip	Z5	0.150″ x 0.08″ Microstrip
C10, C15	10 µF, 50 V, Electrolytic	Z6	0.285″ x 0.08″ Microstrip
C11, C14	0.1 µF, Capacitor	Z7	0.340″ x 0.08″ Microstrip
C12	1000 pF, 100 mil Chip	Z8	0.070″ x 0.08″ Microstrip
C13	160 pF, 100 mil Chip	Z9	0.280″ x 0.08″ Microstrip
R1	35 Ω, 1/4 W Carbon	Z10	0.840″ x 0.08″ Microstrip
R2	30 Ω, 0.1 W Chip	Z11	0.180″ x 0.08″ Microstrip
R3	1.0 kΩ, 0.1 W Chip	Z12	0.600″ x 0.08″ Microstrip
R4	1.0 MΩ, 1/4 W Carbon	L1	7 Turns, 0.076″ ID, #24 AWG Enamel
B1	Fair Rite Products Short Ferrite Bead (2743021446)	L2	5 Turns, 0.126″ ID, #20 AWG Enamel
Board — Glass Teflon®, 31 mils		Input/Output Connectors — Type N	

Note: Plated ceramic part locators (0.1″ x 0.15″) soldered onto Z6 and Z7.

Figure 1. 512 MHz Narrowband Test Circuit

TYPICAL CHARACTERISTICS

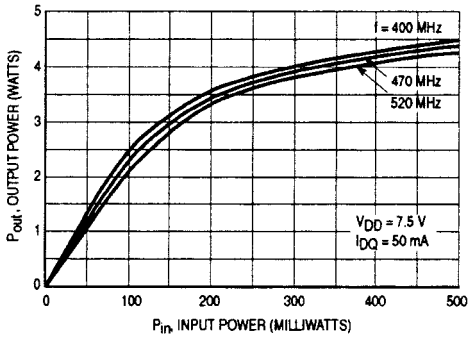


Figure 2. Output Power versus Input Power

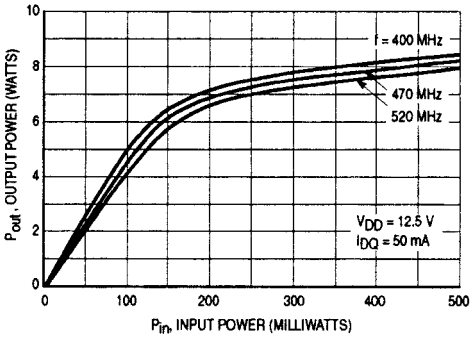


Figure 3. Output Power versus Input Power

TYPICAL CHARACTERISTICS

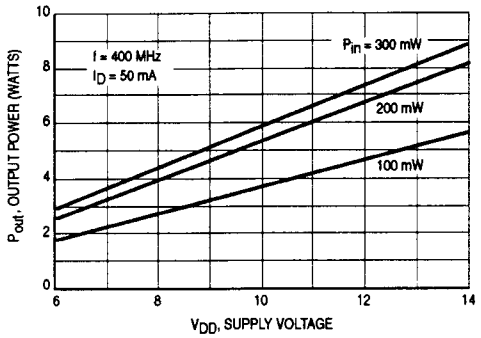


Figure 4. Output Power versus Supply Voltage

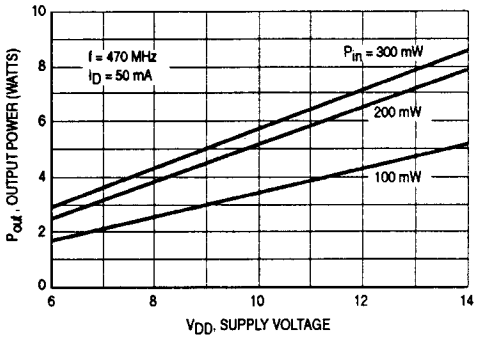


Figure 5. Output Power versus Supply Voltage

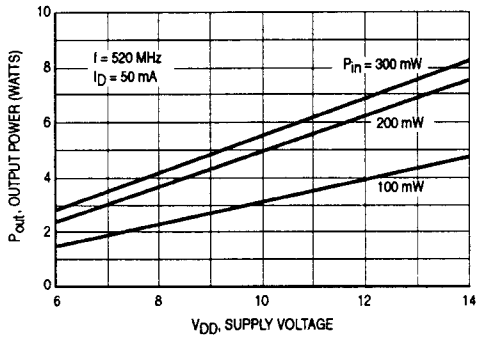


Figure 6. Output Power versus Supply Voltage

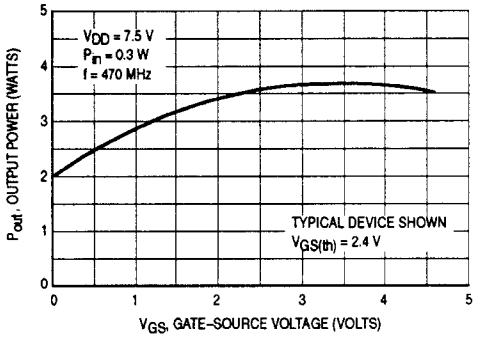


Figure 7. Output Power versus Gate Voltage

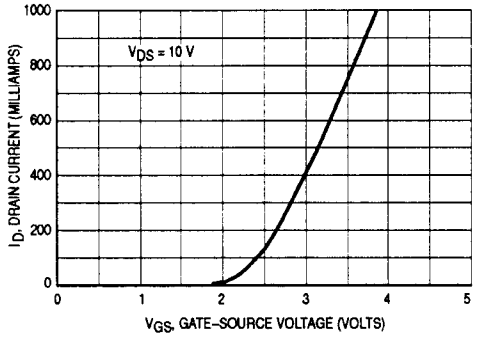


Figure 8. Drain Current versus Gate Voltage (Typical Device Shown)

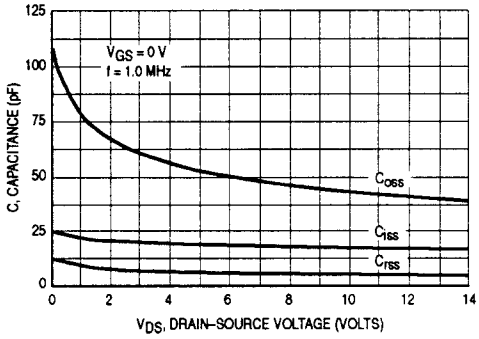


Figure 9. Capacitance versus Voltage

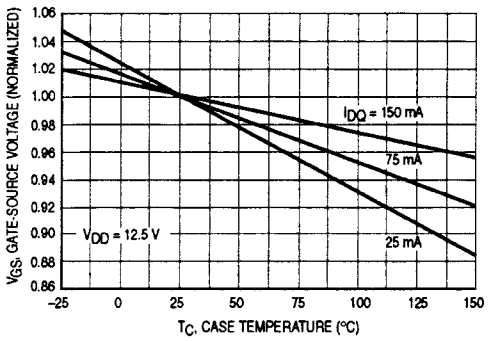


Figure 10. Gate–Source Voltage versus Case Temperature

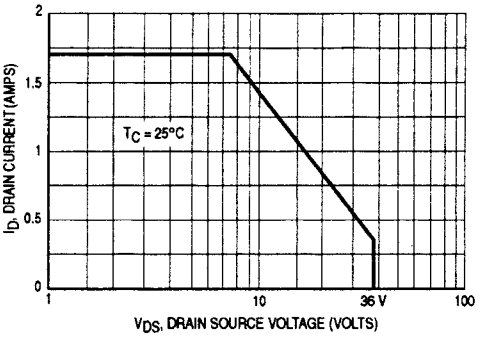
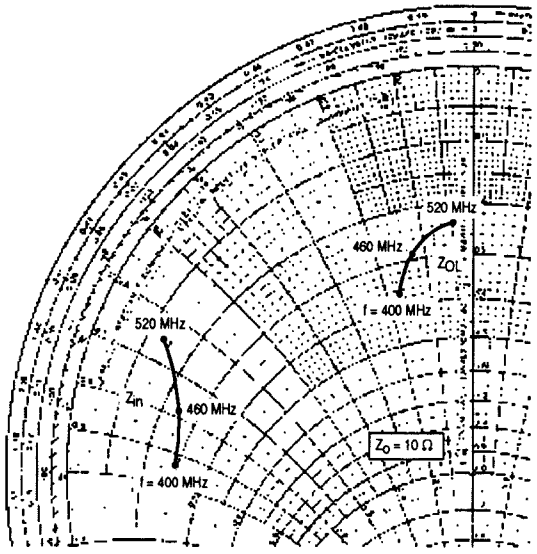


Figure 11. Maximum Rated Forward Biased Safe Operating Area



V_{DD} = 7.5 V, I_{DQ} = 50 mA, P_{out} = 3.0 W

f MHz	Z_{in} Ohms	Z_{OL}^* Ohms
400	2.8 – j9.2	3.6 – j1.7
430	2.7 – j8.5	3.3 – j1.5
460	2.5 – j7.8	2.7 – j1.1
490	2.0 – j7.2	2.5 – j0.8
520	1.3 – j6.5	2.4 – j0.5

Z_{in} = Conjugate of source impedance with parallel 35 Ω resistor and 47 pF capacitor in series with gate.

Z_{OL}^* = Conjugate of the load impedance at given output power, voltage, frequency, and η_D > 50%.

Note: Z_{ol}^* was chosen based on tradeoffs between gain, drain efficiency, and device stability.

Figure 12. Series Equivalent Input and Output Impedance

Table 1. Common Source Scattering Parameters (V_{DS} = 10 V)

I_D = 50 mA

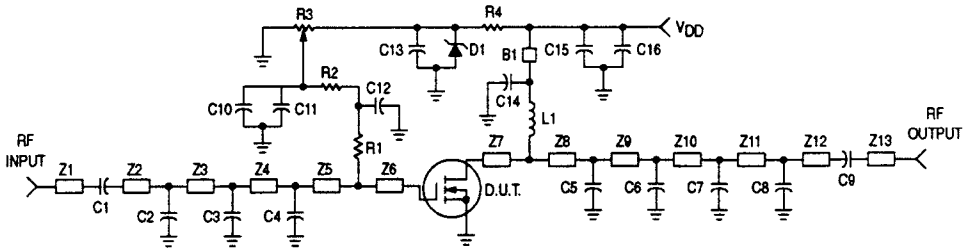
f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.69	−90	10.8	117	0.07	29	0.74	−119
100	0.58	−120	6.0	96	0.08	10	0.78	−146
200	0.58	−139	3.0	75	0.08	−7	0.81	−161
300	0.64	−147	1.9	61	0.07	−16	0.84	−166
400	0.70	−152	1.3	50	0.06	−21	0.86	−169
500	0.75	−157	0.99	41	0.05	−24	0.88	−172
700	0.82	−165	0.61	28	0.03	−15	0.92	−176
850	0.86	−171	0.45	21	0.02	13	0.94	−179
1000	0.89	−176	0.34	16	0.02	47	0.95	178

I_D = 500 mA

f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.76	−124	15.0	109	0.04	23	0.76	−151
100	0.72	−150	7.9	94	0.04	12	0.81	−165
200	0.72	−163	4.0	80	0.04	6	0.83	−172
300	0.73	−168	2.6	71	0.04	5	0.84	−175
400	0.75	−171	1.9	62	0.04	7	0.85	−176
500	0.77	−173	1.5	55	0.03	12	0.86	−178
700	0.81	−177	0.97	42	0.03	29	0.89	−180
850	0.84	−180	0.75	35	0.03	44	0.90	178
1000	0.86	177	0.60	29	0.04	55	0.92	176

I_D = 1.0 A

f	S_{11}		S_{21}		S_{12}		S_{22}	
MHz	$\|S_{11}\|$	$\angle\phi$	$\|S_{21}\|$	$\angle\phi$	$\|S_{12}\|$	$\angle\phi$	$\|S_{22}\|$	$\angle\phi$
50	0.80	−125	14.6	110	0.04	23	0.75	−155
100	0.76	−150	7.8	95	0.04	10	0.81	−167
200	0.76	−164	3.9	81	0.04	1	0.83	−173
300	0.77	−169	2.6	71	0.04	−3	0.84	−175
400	0.79	−172	1.9	63	0.03	−5	0.85	−176
500	0.80	−174	1.4	56	0.03	−5	0.86	−177
700	0.83	−178	0.95	43	0.03	−1	0.88	−179
850	0.85	179	0.73	35	0.02	9	0.90	179
1000	0.87	177	0.58	28	0.02	22	0.91	178



C1, C9	100 pF 100 mil Chip
C2	16 pF, 100 mil Chip
C3	24 pF, 100 mil Chip
C4	68 pF, 100 mil Chip
C5	51 pF, 100 mil Chip
C6	39 pF, 100 mil Chip
C7	6.2 pF, 100 mil Chip
C8	9.1 pF, 100 mil Chip
C10, C15	39000 pF, 100 mil Chip
C11, C16	10 µF, 50 V Electrolytic
C12	10000 pF, 100 mil Chip
B1	Fair Rite Products Short Ferrite Bead (2743021446)
C13	0.1 µF, 100 mil Chip
C14	160 pF, 100 mil Chip
R1	43 Ω, 0.1 W Chip Resistor
R2	1000 Ω, 0.1 W Chip Resistor
R3	10 kΩ Potentiometer
R4	3000 Ω, 0.1 W Chip Resistor
L1	5 Turns, 0.126″ ID, #20 AWG Enamel
Z1 to Z13	See Photomaster
D1	1N4734 Motorola Zener
Board — G10, 1/32″	
Input/Output Connectors — SMA	

Figure 13. Schematic of Broadband Demonstration Amplifier

PERFORMANCE CHARACTERISTICS OF BROADBAND DEMONSTRATION AMPLIFIER

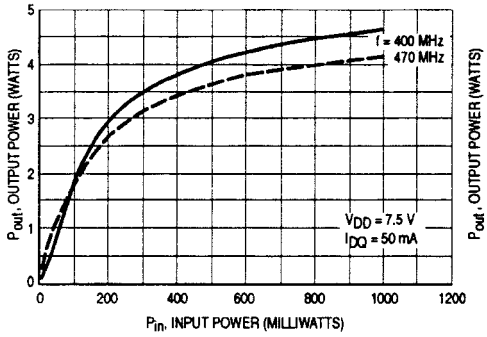


Figure 14. Output Power versus Input Power

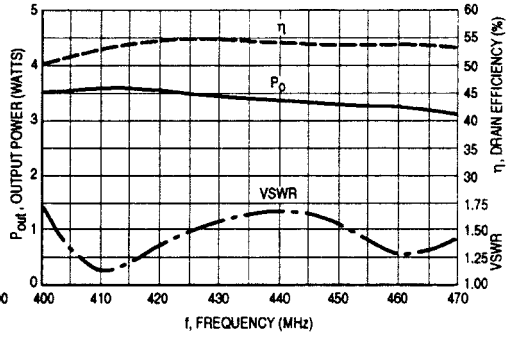


Figure 15. Output Power, Drain Efficiency and VSWR versus Frequency

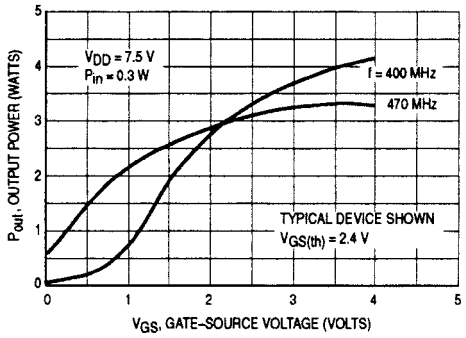


Figure 16. Output Power versus Gate Voltage

DESIGN CONSIDERATIONS

The MRF5003 is a common–source, RF power, N–Channel enhancement mode, Metal–Oxide Semiconductor Field–Effect Transistor (MOSFET). Motorola RF MOSFETs feature a vertical structure with a planar design. Motorola Application Note AN211A, "FETs in Theory and Practice", is suggested reading for those not familiar with the construction and characteristics of FETs.

This surface mount packaged device was designed primarily for VHF and UHF power amplifier applications. Manufacturability is improved by utilizing the tape and reel capability for fully automated pick and placement of parts.

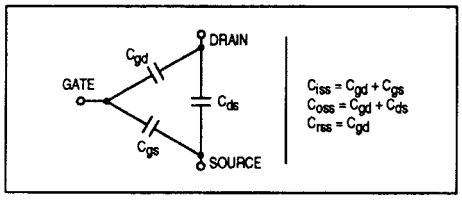
The major advantages of RF power MOSFETs include high gain, simple bias systems, relative immunity from thermal runaway, and the ability to withstand severely mismatched loads without suffering damage.

MOSFET CAPACITANCES

The physical structure of a MOSFET results in capacitors between all three terminals. The metal oxide gate structure determines the capacitors from gate–to–drain (C_{gd}), and gate–to–source (C_{gs}). The PN junction formed during fabrication of the RF MOSFET results in a junction capacitance from drain–to–source (C_{ds}). These capacitances are characterized as input (C_{iss}), output (C_{oss}) and reverse transfer (C_{rss}) capacitances on data sheets. The relationships between the inter–terminal capacitances and those given on data sheets are shown below. The C_{iss} can be specified in two ways:

1. Drain shorted to source and positive voltage at the gate.
2. Positive voltage of the drain in respect to source and zero volts at the gate.

In the latter case, the numbers are lower. However, neither method represents the actual operating conditions in RF applications.



DRAIN CHARACTERISTICS

One critical figure of merit for a FET is its static resistance in the full–on condition. This on–resistance, $R_{DS(on)}$, occurs in the linear region of the output characteristic and is specified at a specific gate–source voltage and drain current. The drain–source voltage under these conditions is termed $V_{DS(on)}$. For MOSFETs, $V_{DS(on)}$ has a positive temperature coefficient at high temperatures because it contributes to the power dissipation within the device.

GATE CHARACTERISTICS

The gate of the RF MOSFET is a polysilicon material, and is electrically isolated from the source by a layer of oxide. The input resistance is very high — on the order of $10^9\ \Omega$ — resulting in a leakage current of a few nanoamperes.

Gate control is achieved by applying a positive voltage to the gate greater than the gate–to–source threshold voltage, $V_{GS(th)}$.

Gate Voltage Rating — Never exceed the gate voltage rating. Exceeding the rated V_{GS} can result in permanent damage to the oxide layer in the gate region.

Gate Termination — The gates of these devices are essentially capacitors. Circuits that leave the gate open–circuited or floating should be avoided. These conditions can result in turn–on of the devices due to voltage build–up on the input capacitor due to leakage currents or pickup.

Gate Protection — These devices do not have an internal monolithic zener diode from gate–to–source. If gate protection is required, an external zener diode is recommended with appropriate RF decoupling.

Using a resistor to keep the gate–to–source impedance low also helps dampen transients and serves another important function. Voltage transients on the drain can be coupled to the gate through the parasitic gate–drain capacitance. If the gate–to–source impedance and the rate of voltage change on the drain are both high, then the signal coupled to the gate may be large enough to exceed the gate–threshold voltage and turn the device on.

DC BIAS

Since the MRF5003 is an enhancement mode FET, drain current flows only when the gate is at a higher potential than the source. See Figure 8 for a typical plot of drain current versus gate voltage. RF power FETs operate optimally with a quiescent drain current (I_{DQ}), whose value is application dependent. The MRF5003 was characterized at I_{DQ} = 50 mA, which is the suggested value of bias current for typical applications. For special applications such as linear amplification, I_{DQ} may have to be selected to optimize the critical parameters.

The gate is a dc open circuit and draws no current. Therefore, the gate bias circuit may generally be just a simple resistive divider network. Some special applications may require a more elaborate bias system.

GAIN CONTROL

Power output of the MRF5003 may be controlled from its rated value down to zero (negative gain) with a low power dc control signal, thus facilitating applications such as manual gain control, ALC/AGC and modulation systems. Figure 16 is an example of output power variation with gate–source bias voltage. This characteristic is very dependent on frequency and load line.

MOUNTING

The specified maximum thermal resistance of 14°C/W assumes a majority of the 0.100″ x 0.200″ source contact on the back side of the package is in good contact with an appropriate heat sink. In the test fixture shown in Figure 1, the device is clamped directly to a copper pedestal. In the demonstration amplifier, the device was mounted on top of the G10 circuit board and heat removal was accomplished through several solder filled plated through holes. As with all RF power devices, the goal of the thermal design should be to minimize the temperature at the back side of the package.

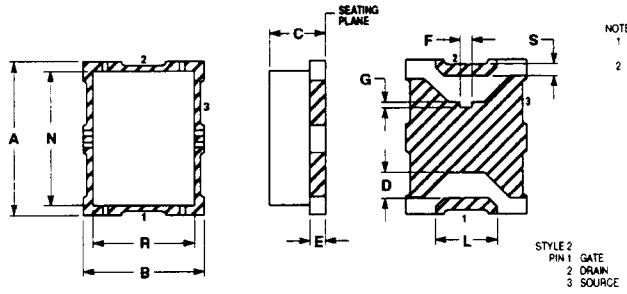
AMPLIFIER DESIGN

Impedance matching networks similar to those used with bipolar transistors are suitable for the MRF5003. For examples see Motorola Application Note AN721, "Impedance Matching Networks Applied to RF Power Transistors." Both small–signal S–parameters and large–signal impedances are provided. While the S–parameters will not produce an exact design solution for high power operation, they do yield a good first approximation. This is an additional advantage of RF power MOSFETs.

Since RF power MOSFETs are triode devices, they are not unilateral. This coupled with the very high gain of the MRF5003 yield a device capable of self oscillation. Stability may be achieved by techniques such as drain loading, input shunt resistive loading, or output to input feedback. Different stabilizing techniques were applied to the test fixture and demonstration amplifiers. The RF test fixture implements a parallel resistor and capacitor in series with the gate while the demonstration amplifier utilizes a 43 Ω shunt resistor from gate to ground. Both circuits have a load line selected for a higher efficiency, lower gain, and more stable operating region.

Two port stability analysis with the MRF5003 S–parameters provides a useful tool for selection of loading or feedback circuitry to assure stable operation. See Motorola Application Note AN215A, "RF Small–Signal Design Using Two–Port Parameters," for a discussion of two port network theory and stability.

PACKAGE DIMENSIONS



NOTES
1 DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2 CONTROLLING DIMENSION: INCH.

	INCHES		MILLIMETERS	
DIM	MIN	MAX	MIN	MAX
A	0.260	0.270	6.60	6.86
B	0.200	0.210	5.08	5.33
C	0.090	0.104	2.29	2.64
D	0.040	0.050	1.02	1.27
E	0.022	0.028	0.56	0.71
F	0.015	0.025	0.38	0.64
G	0.005	0.015	0.13	0.38
L	0.100	0.110	2.54	2.79
N	0.226	0.236	5.74	5.99
R	0.166	0.176	4.22	4.47
S	0.025	0.035	0.64	0.89

CASE 430–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters can and do vary in different applications. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and Ⓜ are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

2-3-7 GaAs MESFETs

Introduction. More than twenty years ago, for RF applications, the JFET was the dominant FET. As mentioned earlier, its major disadvantage is the fact that the f_T cutoff frequency, defined as

$$f_T = g_m / 2\pi C_{GS} \tag{2-262}$$

limits the operating range of this semiconductor device to 1 GHz at most. Also, the JFET's temperature sensitivity and loose tolerances in transconductance and other RF-related values made it difficult to use. Finally, no RFICs made use of this technology. While at the moment silicon-based MOS technology probably is the most cost-effective and is produced in the largest quantities, it does not rival the metal-semiconductor FET designed on GaAs technology. Figure 2-115 shows the structure of the silicon MOSFET, silicon JFET, and the GaAs MESFET.

Large-Signal Behavior of GaAs MESFETs. The mathematics for the large-signal behavior of the GaAsFET is basically quite similar to that for the JFET; however, the computations for JFET channel current, diode currents, and capacitance are much simpler than for the GaAsFET. Temperature effects are embedded in all the equations for all the transistors mentioned so far. In our opinion, the modified Materka model used by Ansoft and a large number of users is the most complete one; however, the following models are supported by most CAD tools:

- Chalmers (Angelov)
- Curtice–Ettenberg cubic
- Curtice quadratic
- IAF (Berroth)
- ITT PFET and TFET
- Modified Materka–Kacprzak
- Physics-based MESFET
- Raytheon (Statz)
- TriQuint (TOM-1, TOM-2, and TOM-3)
- With some restrictions, metal-insulator-semiconductor FETs (MISFETs), modulation-doped FETs (MODFETs), and high-electron-mobility transistors (HEMTs)

The main advantages of GaAsFET technology derive from the fact that it uses a metal-semiconductor junction with a barrier voltage of 0.8 V, its input capacitance is typically less than 0.2 pF, and the reverse feedback capacitance is less than 0.02 pF, or roughly 10% of the input capacitance. As a result of this, $f_{\max}$ is approximately five times higher than f_T. Table 2-16 shows a comparison of silicon BJT and GaAsFET technologies. While $f_{\max}$ for the bipolar transistor is

$$f_{\max} \simeq \sqrt{\frac{f_t}{8\pi r'_{bb} C_C}} \tag{2-263}$$

the $f_{\max}$ determination for the GaAsFET is given by

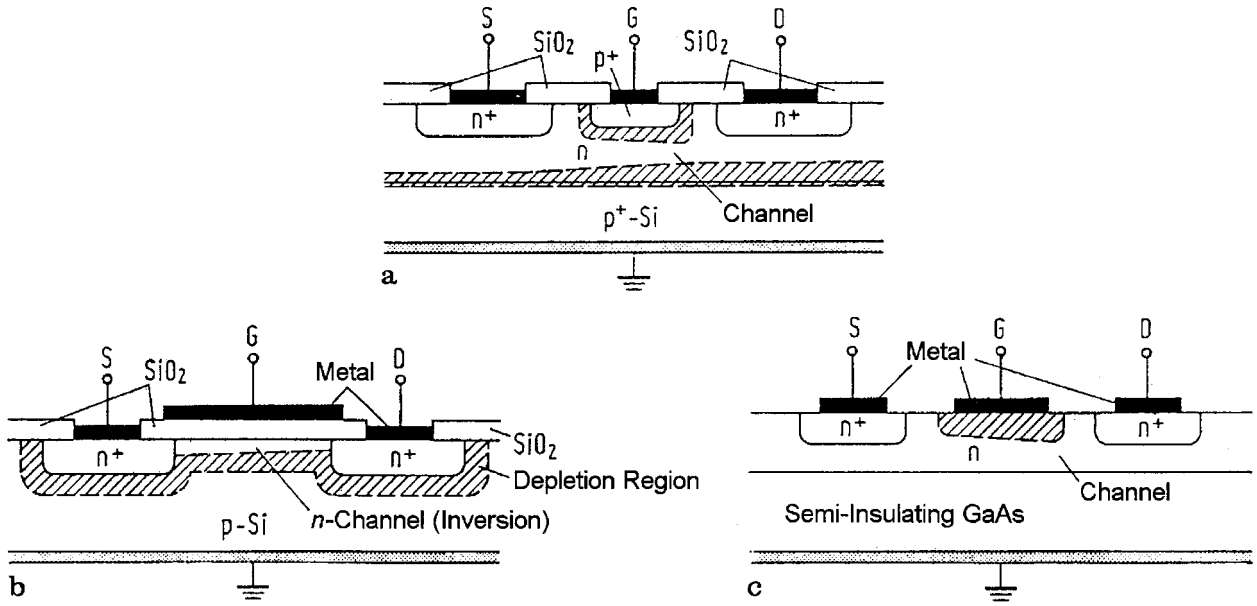


Figure 2-115 This picture is provided to differentiate between the structure of the three technologies. (a) The JFET and (b) the MOSFET are based on silicon; (c) the GaAsFET is based on semi-insulating GaAs. While the MOS transistor has silicon dioxide (SiO_2) as gate insulation, both the JFET and GaAs MESFET have a diode from gate to source, with a barrier voltage of 0.65 V for silicon and 0.8 V for GaAs.

Table 2-16 Comparison of Si BJT, SiGe HBT, and GaAsFET technologies

	Si Bipolar	SiGe HBT	GaAs FET
f_T	25 GHz	100 GHz	22 GHz
$f_{\max}$	40 GHz	200 GHz	110 GHz
Features	Low cost, low 1/*f* noise (5 kHz)	Low 1/*f* noise, very low distortion, high cost, IC only	Highest flexibility, lowest NF_0, well established

$$f_{\max} = \frac{f_T}{2}\sqrt{\frac{R_0}{R_i + R_s + R_g}} \tag{2-264}$$

As a sample calculation,

$$f_{\max}\ \frac{21.9\text{ GHz}}{2}\sqrt{\frac{450}{1 + 1.5 + 2}} = 110\text{ GHz} \tag{2-265}$$

The three major drawbacks of the GaAsFET are:

1. Much higher flicker corner frequency (somewhere between 10 and 100 MHz), probably due to a lack of surface passivation.
2. Much higher output conductance. This tends to load down any circuit connected to the drain. On the other hand, since the transconductance is quite high for even low currents, these devices have very high gains at low frequencies, which can make them quite unstable. In the saturated mode it is not uncommon to find a drain–source resistance of 100–500 Ω, while BJTs and JFETs offer values of several kilohms and higher.
3. Because of the very high flicker corner frequency (from 10 to 100 MHz), MESFETs are really not useful for low-noise mixers and oscillators, and unless there are no devices available in the frequency range above 30 GHz, they should be avoided for these applications.

As to the MESFET's construction and dc properties, Figure 2-116 shows a MESFET's cross section and dc *I*–*V* characteristics.

It was outlined in the beginning that while the GaAsFET is a close relative to the JFET and MOSFET, its actual behavior was found to be described best by a set of analytic equations. The first such model was the one by Curtice in the form of quadratic and cubic models, but it does not have enough derivatives to give enough insight into subtle things such as third- and higher-order intermodulation distortion and accurate harmonic generation. Other researchers have addressed various areas, but we still find that the Materka model has the best approximation, especially using the two-transistor approach outlined later [20, 21]. The large-signal topology for all FETs consists of an intrinsic model with some extrinsic parameters, further complicated by the package, as shown by Figures 2-117 and 2-118. Table 2-17 lists their keywords. The actual intrinsic model and its parameter definition depend on the particular model and since designs using GaAsFET will always be done using CAD tools, we will not go into any details of the equations but will list them. They are not dissimilar from the JFET and MOSFET equations.

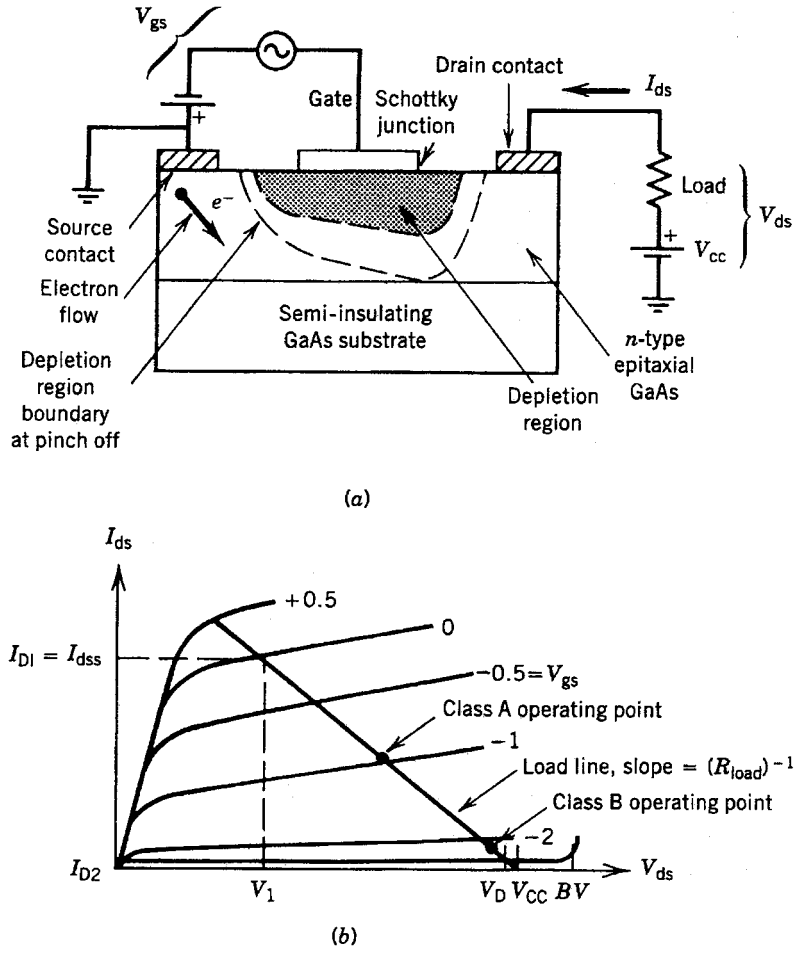


Figure 2-116 (a) Cross section and bias circuit and (b) dc *I–V* curve, including ac load line, for a MESFET.

The Modified Materka-Kacprzak Model Figure 2-119 shows the intrinsic model of the Materka FET. Table 2-18 lists its keywords.

Large-Signal Equations

Device Equations

V_{gsi} = Intrinsic gate–source voltage
V_{dsi} = Intrinsic drain–source voltage
V_1 = Voltage across C_{GS} and R_i
V_{gdi} = Intrinsic gate–drain voltage
$V_T = kT_J/q$ (thermal voltage)
k = Boltzmann's constant
q = Electron charge
T_J = Analysis temperature (kelvins)

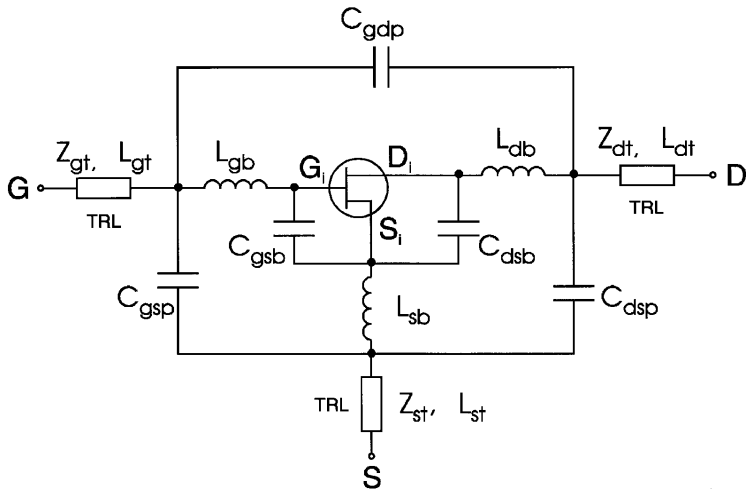


Figure 2-117 MESFET extrinsic model.

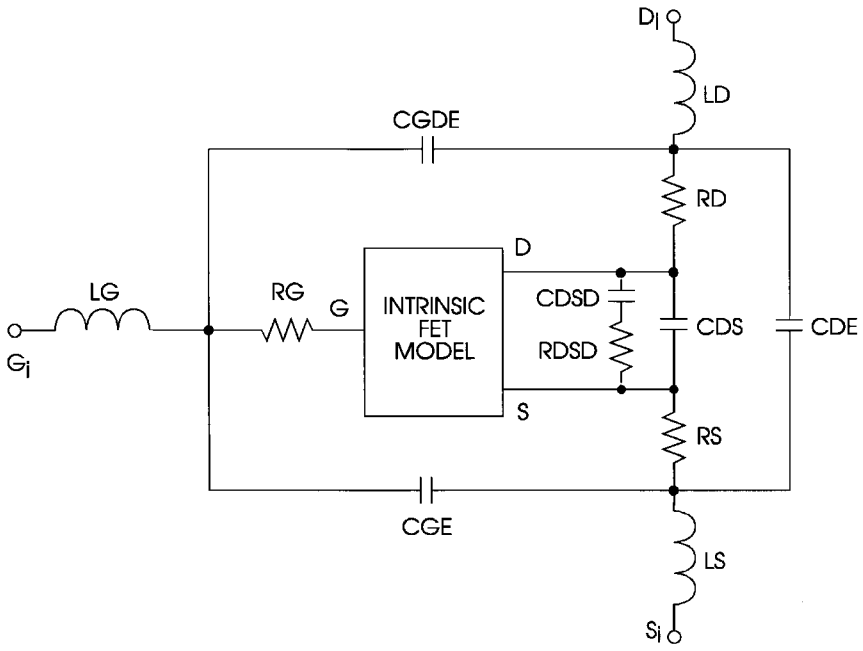


Figure 2-118 MESFET package model.

Table 2-17 Large-signal MESFET model keywords

Keyword	Description	Unit	Default
	Extrinsic Parameters		
RG	Gate bulk and ohmic resistance	ohm	0.0
RD	Drain bulk and ohmic resistance	ohm	0.0
RS	Source bulk and ohmic resistance	ohm	0.0
LG	Gate-lead inductance (metallization)	henry	0.0
LD	Drain-lead inductance (metallization)	henry	0.0
LS	Source-lead inductance (via)	henry	0.0
CDS	Drain–source capacitance	farad	0.0
CDSD	Low-frequency trapping capacitor	farad	0.0
RDSD	Channel trapping resistance	ohm	∞
CGE	Gate–source electrode capacitance	farad	0.0
CDE	Drain–source electrode capacitance	farad	0.0
CGDE	Gate–drain electrode capacitance	farad	0.0
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate-to-source package capacitance	farad	0.0
CDSP	Drain-to-source package capacitance	farad	0.0
CGDP	Gate-to-drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for $\varepsilon_r = 1$	meter	0.0

Channel Current

$$I_{ds} = \mathrm{IDSS}\left(1 + \mathrm{SS}\,\frac{V_{dsi}}{\mathrm{IDSS}}\right)\left(1 - \frac{V_{gsi}(t-T)}{\mathrm{VP0} + \mathrm{GAMA}V_{dsi}}\right)^{(\mathrm{E}+\mathrm{KE}V_{gsi}(t-T))}$$

$$\times \tanh\left(\frac{\mathrm{SL}V_{dsi}}{\mathrm{IDSS}(1 - \mathrm{KG}V_{gsi}(t-T))}\right)$$

Diode

$$I_{gd} = I_{gdc} - \begin{cases} \mathrm{IB0}\,\exp[-\mathrm{AFAB}(V_{gdi} + \mathrm{VBC})] \\ \dfrac{\mathrm{GMAX}}{4}\,\{\tanh[\mathrm{K1D}(V_{gsi} - \mathrm{K2D})] - 1\}\,[V_{gdi} + \mathrm{VBC} - \sqrt{(V_{gdi} + \mathrm{VBC})^2 + \mathrm{K3D}}] \end{cases}$$

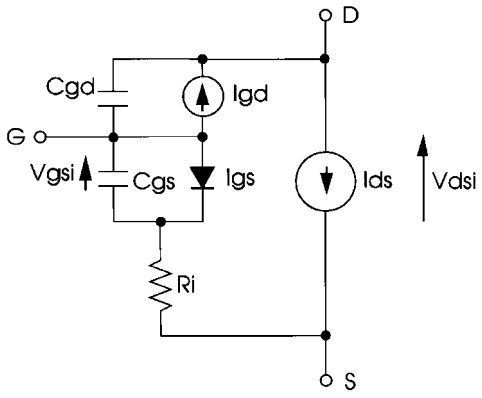


Figure 2-119 Intrinsic model of the modified Materka–Kacprzak MESFET.

where

$$I_{gdc} = \text{IG0}[\exp(\text{AFAB}V_{gdi}) - 1]$$

Channel Resistance

$$R_i = \begin{cases} \text{R10}(1 - \text{KR}V_{gsi}), & \text{KR}V_{gsi} < 1.0 \\ 0, & \text{KR}V_{gsi} \geq 1.0 \end{cases}$$

Capacitance Model

$$C_{gs} = \text{CGS0}\frac{F1F2}{\sqrt{1 - \dfrac{V_{\text{new}}}{\text{VBI}}}} + \text{CGD0}\ F3$$

$$C_{gd} = \text{CGS0}\frac{F1F3}{\sqrt{1 - \dfrac{V_{\text{new}}}{\text{VBI}}}} + \text{CGD0}\ F2$$

where

$$F1 = \frac{1}{2}\left(1 + \frac{V_{\text{eff}} - V_T}{\sqrt{(V_{\text{eff}} - V_T)^2 + \delta^2}}\right)$$

$$F2 = \frac{1}{2}\left(1 + \frac{V_{gsi} - V_{gdi}}{\sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}}\right)$$

Table 2-18 Materka–Kacprzak MESFET model keywords

Keyword	Description	Unit	Default
	Area, Noise, and Name		
AREA	Area multiplier		1.0
KFN	Flicker noise coefficient (Materka model only)[a]		0
AF	Flicker noise exponent		1.0
FCP	Flicker noise frequency shape factor		1.0
	Channel Current Model		
IDSS	Drain saturation current for $V_{GS} = 0$	ampere	0.1
VP0	Pinchoff voltage for $V_{DS} = 0$	volt	–2.0
GAMA	Voltage slope parameter of pinchoff voltage	/volt	0.0
E	Constant part of power-law parameter		2.0
KE	Dependence of power law on V_{GS}	/volt	0.0
SL	Slope of the $V_{GS} = 0$ drain characteristic in the linear region	ampere/volt	0.15
KG	Drain dependence on V_{GS} in the linear region	/volt	0.0
SS	Slope of the drain characteristic in the saturated region	ampere/volt	0.0
T	Channel transit-time delay	second	0.0
IG0	Diode saturation current	ampere	0
AFAG	Slope factor of forward diode current	/volt	38.696
IB0	Breakdown saturation current	ampere	0
AFAB	Slope factor of breakdown current	/volt	0
VBC	Breakdown voltage	volt	∞
GMAX	Breakdown conductance	ampere/volt	0
K1D	Fitting parameter	/volt	0
K2D	Fitting parameter	volt	0
K3D	Fitting parameter	$volt^2$	0
R10	Intrinsic channel resistance for $V_{GS} = 0$	ohm	0.0
KR	Slope factor of intrinsic channel resistance	/volt	0.0
	Materka Capacitance Model		
C10	Gate–source Schottky barrier capacitance for $V_{GS} = 0$	farad	0.0
K1	Slope parameter of gate–source capacitance	/volt	1.25
MGS	Gate–source grading coefficient		0.5
C1S	Constant parasitic component of gate–source capacitance	farad	0.0
CF0	Gate–drain feedback capacitance for $V_{GD} = 0$	farad	0.0
KF	Slope parameter of gate–drain feedback capacitance	/volt	1.25
MGD	Gate–drain grading coefficient		0.5
FCC	Forward-bias depletion capacitance coefficient		0.8

[a]The flicker noise parameter of the Materka model is KFN so as not to conflict with the KF parameter in the capacitance model.

$$F3 = \frac{1}{2}\left(1 - \frac{V_{gsi} - V_{gdi}}{\sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}}\right)$$

$$V_{\text{new}} = \begin{cases} A1, & A1 < \text{VMAX} \\ \text{VMAX}, & A1 \geq \text{VMAX} \end{cases}$$

$$A1 = \frac{1}{2}\left(V_{\text{eff}} + V_T + \sqrt{(V_{\text{eff}} - V_T)^2 + \delta^2}\right)$$

$$V_{\text{eff}} = \frac{1}{2}\left(V_{gsi} + V_{gdi} + \sqrt{(V_{gsi} - V_{gdi})^2 + (1/\text{ALFA})^2}\right)$$

$$V_T = \text{VP0} + \text{GAMA}\, V_{dsi}$$

$$\delta = 0.2$$

Some of the modifications to the Materka model have been done by Ansoft under various Department of Defense contracts, and by Raytheon and Texas Instruments under similar contracts, with Ansoft being a subcontractor.

The most relevant equation is really the channel current. Its derivatives are largely responsible for the accuracy of the intermodulation distortion, power-added efficiency, and, of course, its dc *I–V* curves.

Enhancement/Depletion FETs To make the designer's life more difficult, it turns out that there are two types of GaAsFETs:

1. *Depletion FETs (DFETs).* Most similar to the JFET; here V_G must be negative to control the device. They are the most commonly produced and are the FET type most referred to in this book.

On one hand:

- They require a negative gate voltage with respect to the source.
- Self-bias allows operation from a single supply voltage.

On the other hand:

- For low-voltage operation, a negative voltage generator may be required.
- Supply voltage must be doubled to accommodate full-swing operation.

2. *Enhancement FETs (EFETs).* Most similar to the MOSFET; here V_G must be positive to bring life to the device. Practically speaking, EFETs are used mostly in integrated circuits; they're typically not available in discrete, packaged form.

On one hand:

- They need only positive supply for biasing.
- They provide higher g_m/mA (for same device width)—5.1 mS versus 3.9 mS at 8 mA.

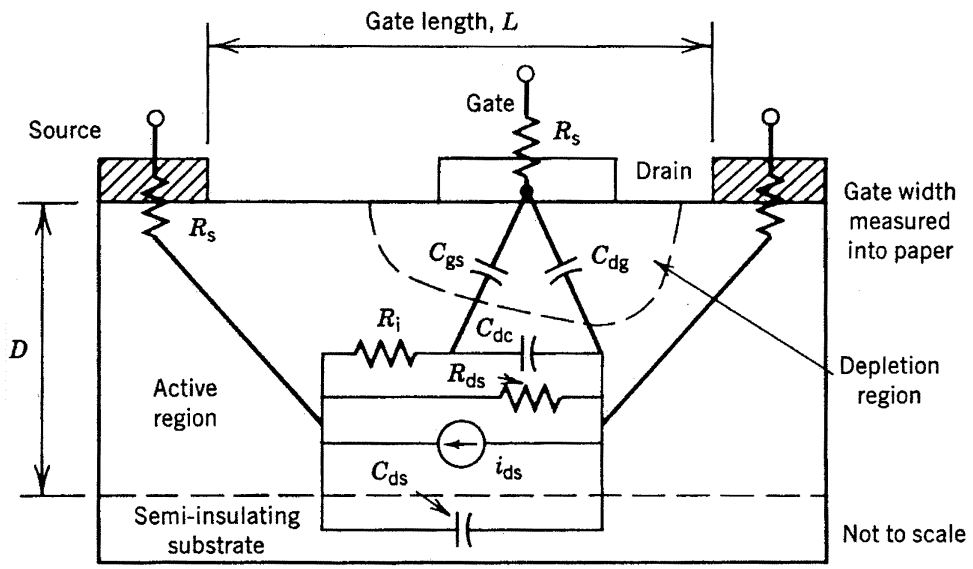


Figure 2-120 Location of lumped-element components for a MESFET.

- They are good for low-power LNAs, giving slightly better NF that DFETs, an NF of 1 dB at 1 GHz, and $I_{dd} < 10$ mA.

On the other hand:

- They have a very limited gate bias range (V_{GS} between 0.15 and 0.7 V).
- The gate conduction degrades NF and input impedance.
- The gate capacitance is higher than that of DFETs.
- The linearity is not as good as that of DFETs.

With today's technologies, all GaAs devices are *n*-channel; we have not seen any *p*-channels yet. More information about biasing will be given in the next chapter.

Figure 2-120 is a lumped-element, two-part equivalent circuit of a MESFET showing the location of lumped-element components.

2-3-8 Small-Signal GaAs MESFET Model

Figure 2-121 shows the applicable linear equivalent circuit for a MESFET and Table 2-19 lists its keywords. As with the MOS transistors, there is a gallium-arsenide dual-gate MOSFET available that is used mostly in special circuits, such as preamplifiers and mixers, whose IF has to be significantly higher than the flicker corner frequency, for example, higher than 200 MHz.

The following two datasheets* show typical GaAsFETs, both low-noise and medium-power. Such devices are, of course, available from a large variety of manufacturers—to name at least two, Siemens and the Nippon Electric Company, which is represented in the United States by California Eastern Laboratories.

*Reproduced with permission.

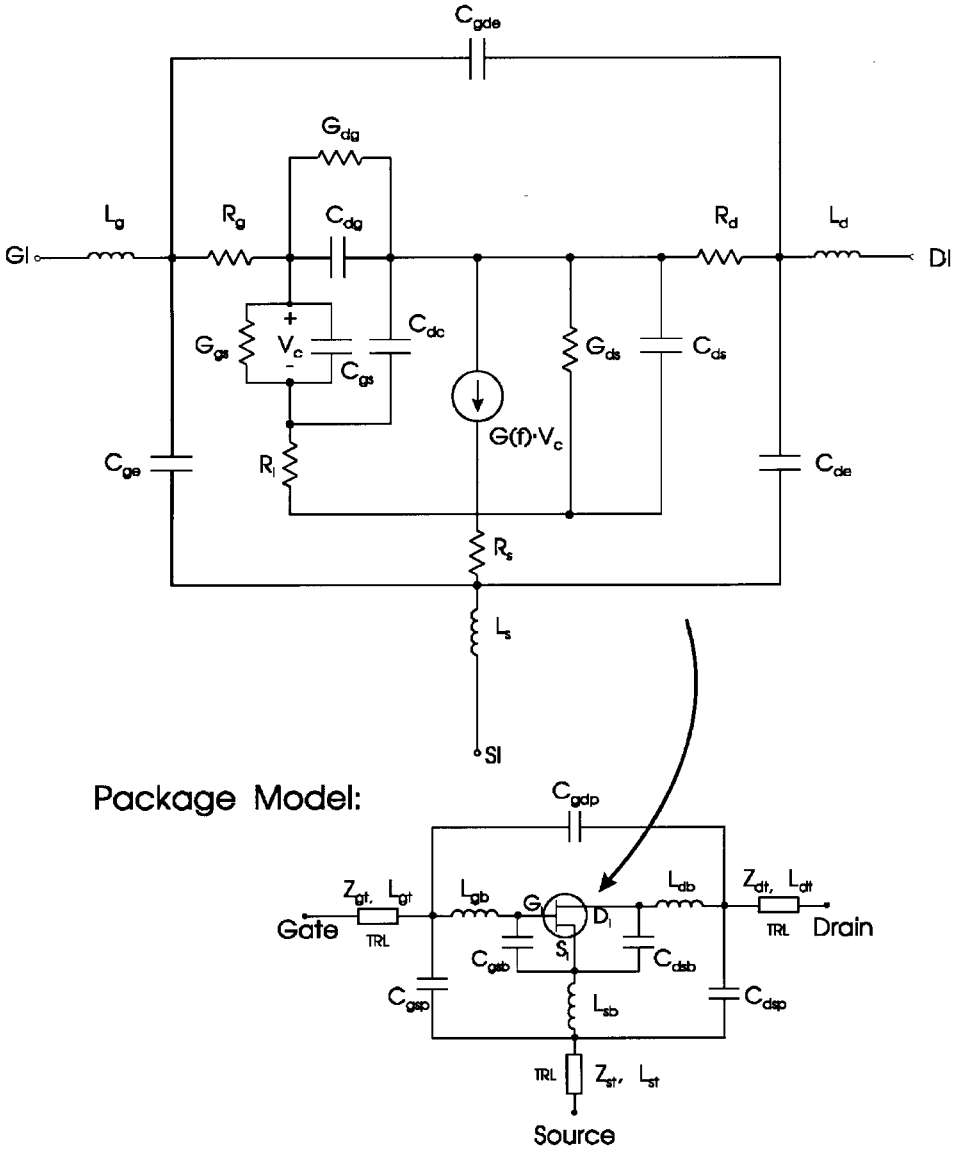


Figure 2-121 Small-signal model of a MESFET.

Table 2-19 Small-signal MESFET model keywords

Keyword	Description	Unit	Default
G	Transconductance at dc, G_0[a]	/ohm	
CGS	Gate–source capacitance	farad	
F	3-dB Roll-off frequency	hertz	∞
T	Time delay	second	0.0
TDS	Drain–source time delay	second	0.0
GGS	Gate–source conductance	/ohm	0.0
CDG	Drain–gate capacitance	farad	0.0
CDC	Dipole-layer capacitance	farad	0.0
CDS	Drain–source capacitance	farad	0.0
GDS	Drain–source conductance	/ohm	0.0
RI	Channel resistance	ohm	0.0
RG	Gate resistance	ohm	0.0
RD	Drain resistance	ohm	0.0
RS	Source resistance	ohm	0.0
CGE	External gate capacitance	farad	0.0
CDE	External drain capacitance	farad	0.0
LG	Gate-lead inductance	henry	0.0
LD	Drain-lead inductance	henry	0.0
LS	Source-lead inductance	henry	0.0
CGDE	External gate–drain capacitance	farad	0.0
GDG	Gate–drain conductance	/ohm	0.0
TJ	Chip temperature	kelvin	298
	Package Parasitics		
LGB	Gate wirebond inductance	henry	0.0
LDB	Drain wirebond inductance	henry	0.0
LSB	Source wirebond inductance	henry	0.0
CGSB	Gate bondpad to source capacitance	farad	0.0
CDSB	Drain bondpad to source capacitance	farad	0.0
CGSP	Gate-to-source package capacitance	farad	0.0
CDSP	Drain-to-source package capacitance	farad	0.0
CGDP	Gate-to-drain package capacitance	farad	0.0
ZGT	Gate transmission line impedance	ohm	50
ZDT	Drain transmission line impedance	ohm	50
ZST	Source transmission line impedance	ohm	50
LGT	Gate transmission line length for at $\varepsilon_r = 1$	meter	0.0
LDT	Drain transmission line length for at $\varepsilon_r = 1$	meter	0.0
LST	Source transmission line length for at $\varepsilon_r = 1$	meter	0.0
FC	Corner frequency of flicker (1/*f*) noise[b]	hertz	10 MHz
FCP	Shape factor of the 1/*f* noise response	1.0	
label	User-defined term that refers to temperature coefficient		

[a]The transconductance of this model may approximately be described by $g_m = G\dfrac{e^{-j\omega T}}{1 = j(f/F)}$ where $\omega = 2\pi f$ and f is frequency.

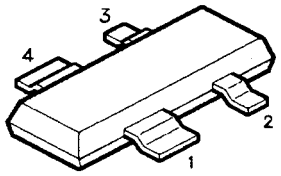
[b]The flicker noise frequency dependence is given by $1/(f/F_c)^{\mathrm{FCP}}$

SIEMENS AlGaAs / InGaAs HEMT CFY 77

Datasheet

Features

* Very low noise
* Very high gain
* For low noise front end amplifiers up to 20 GHz
* For DBS down converters



ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering code (taped)	Package 1)
CFY77-08	HG	Q62702-F1549	MW-4
CFY77-10	HH	Q62702-F1559	MW-4

Maximum ratings	Symbol		Unit
Drain-source voltage	V_{DS}	3.5	V
Drain-gate voltage	V_{DG}	4.5	V
Gate-source voltage	V_{GS}	-3.0	V
Drain current	I_D	60	mA
Channel temperature	T_{Ch}	150	°C
Storage temperature range	T_{stg}	-65...+150	°C
Total power dissipation ($T_s \leq 51$°C) 2)	P_{tot}	180	mW
Thermal resistance			
Channel-soldering point source	R_{thChS}	550	K/W

1) Dimensions see chapter Package Outlines
2) T_S: Temperature measured at soldering point

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Electrical characteristics at T_A = 25°C

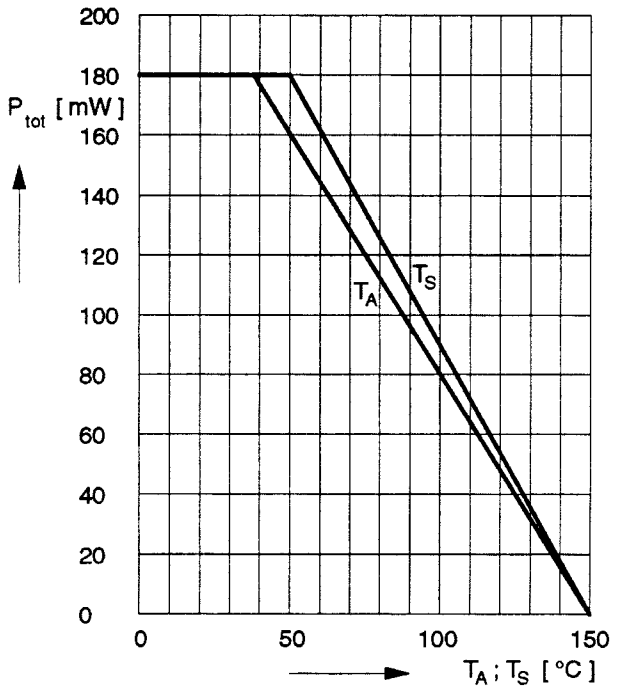
unless otherwise specified

Characteristics	Symbol	min	typ	max	Unit
Drain-source saturation current V_{DS} = 2 V V_{GS} = 0 V	I_{DSS}	15	30	60	mA
Pinch-off voltage V_{DS} = 2 V I_D = 1 mA	$V_{GS(P)}$	-2	-0.7	-0.2	V
Gate leakage current V_{DS} = 2 V I_D = 15 mA	I_G	-	0.05	2	µA
Transconductance V_{DS} = 2 V I_D = 15 mA	g_m	50	65	-	mS
Noise figure V_{DS} = 2 V I_D = 15 mA f = 12 GHz	F				dB
CFY77-08 CFY77-10		- -	0.7 0.9	0.8 1	
Associated gain V_{DS} = 2 V I_D = 15 mA f = 12 GHz	G_a				dB
CFY77-08 CFY77-10		10 9.5	10.5 10	- -	

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Total Power Dissipation $P_{tot} = f(T_s;T_A)$

Package mounted on alumina



Typical Common Source Noise Parameters

I_D = 15 mA U_{DS} = 2.0 V Z_0 = 50 Ω

f	F_{min}	G_a	Γ_{opt}		R_n	r_n	N	$F_{50\Omega}$
GHz	dB	dB	MAG	ANG	Ω			dB
2	0.36	19.4	0.79	27	13.7	0.274	0.03	1.2
4	0.44	15.9	0.72	60	10.1	0.202	0.04	1.1
6	0.51	13.9	0.63	92	5.85	0.117	0.05	1.05
8	0.58	12.4	0.56	134	2.35	0.047	0.06	1.0
10	0.65	11.2	0.52	180	1.1	0.022	0.07	1.0
12	0.72	10.4	0.54	-135	2.9	0.058	0.08	1.1
14	0.80	9.7	0.59	-108	7.15	0.143	0.10	1.5

SIEMENS AlGaAs / InGaAs HEMT **CFY 77**

Typical Common Source S-Parameters

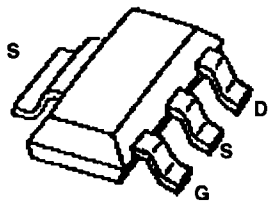
I_D = 15 mA U_D = 2.0 V Z_o = 50 Ω

	S11		S21		S12		S22	
GHz	Mag	Ang	Mag	Ang	Mag	Ang	Mag	Ang
1	0.98	-22.8	5.55	159.6	0.030	87.5	0.633	-16.3
2	0.94	-46.1	5.40	139.3	0.053	57.8	0.60	-32.5
3	0.88	-68.4	5.09	120.1	0.074	44.9	0.54	-48.0
4	0.82	-90.6	4.77	101.2	0.089	30.7	0.48	-63.3
5	0.77	-110.8	4.45	84.0	0.101	18.1	0.42	-77.5
6	0.72	-131.4	4.16	67.3	0.112	7.9	0.35	-92.6
7	0.66	-153.6	3.88	50.2	0.119	-3.3	0.28	-110.8
8	0.63	-175.2	3.58	34.5	0.122	-12.7	0.22	-132.0
9	0.62	164.4	3.29	18.9	0.120	-22.0	0.16	-157.3
10	0.62	145.0	3.01	4.0	0.119	-29.5	0.14	177.3
11	0.64	128.3	2.76	-10.3	0.119	-37.4	0.15	136.2
12	0.64	113.1	2.51	-23.5	0.114	-44.0	0.18	115.4
13	0.66	101.3	2.32	-35.7	0.114	-47.3	0.23	100.9
14	0.67	89.4	2.18	-48.2	0.116	-53.1	0.25	91.0
15	0.69	73.6	2.06	-62.4	0.116	-58.6	0.28	75.4
16	0.73	59.2	1.85	-75.9	0.115	-65.8	0.36	57.1
17	0.76	51.7	1.65	-86.5	0.112	-69.4	0.39	53.1
18	0.78	45.4	1.56	-96.7	0.115	-72.3	0.42	43.8
19	0.77	36.2	1.51	-108.6	0.121	-76.7	0.44	38.8

SIEMENS GaAs FET **CLY 15**

Datasheet

* Power amplifier for mobile phones
* For frequencies from 400 MHz to 2.5 GHz
* Operating voltage range: 2.7 to 6 V
* P_{out} at V_D=3V, f=1.8 GHz typ. 31.5 dBm
* Efficiency better 50%



ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering code (taped)	Package 1)
CLY 15	CLY 15	Q62702-L99	SOT 223

Maximum ratings	Symbol		Unit
Drain-source voltage	V_{DS}	9	V
Drain-gate voltage	V_{DG}	12	V
Gate-source voltage	V_{GS}	-6	V
Drain current	I_D	5	A
Channel temperature	T_{Ch}	150	°C
Storage temperature	T_{stg}	-55...+150	°C
Total power dissipation (Ts ≤ 80°C) Ts: Temperature at soldering point	P_{tot}	4.7	W
Thermal resistance Channel-soldering point (GND)	R_{thChS}	< 15	K/W

1) Dimensions see chapter Package Outlines

SIEMENS GaAs FET **CLY 15**

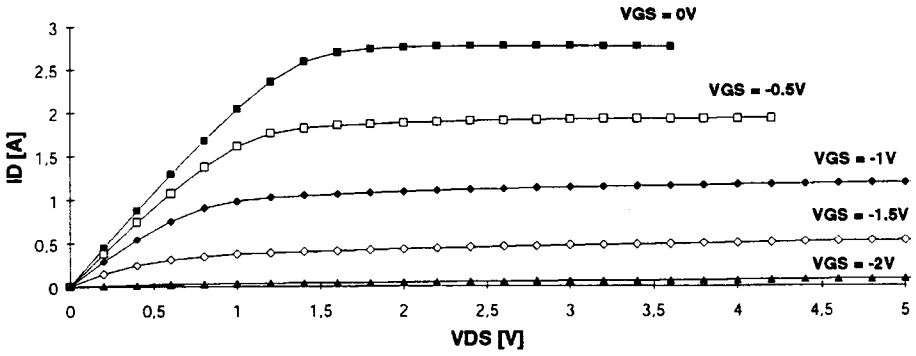
Electrical characteristics (T_A = 25°C , unless otherwise specified)

Characteristics	Symbol	min	typ	max	Unit
Drain-source saturation current *) V_{DS} = 3V V_{GS} = 0V	I_{DSS}	2.4	3.2	4.8	A
Cut-off current V_{DS} = 3V V_{GS} = -3.8V	I_D	-	-	400	µA
Gate cut-off current V_{DS} = 3V V_{GS} = -3.8V	I_G	-	20	70	µA
Pinch-off Voltage VDS=3V ID=400µA	$V_{GS(p)}$	-3.8	-2.8	-1.8	V
Small Signal Gain *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 5dBm	G	-	6	-	dB
Output Power *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 29dBm	P_o	32	32.5	-	dBm
Output Power *) V_{DS} = 5V I_D = 1.4A f = 1.8GHz P_{in} = 30 dBm	P_o	34.5	35	-	dBm
1dB-Compression Point *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz	P_{1dB}	-	31.5	-	dBm
1dB-Compression Point *) V_{DS} = 5V I_D = 1.4A f = 1.8GHz	P_{1dB}	-	34.5	-	dBm
Power Added Efficiency *) V_{DS} = 3V I_D = 1.4A f = 1.8GHz P_{in} = 29dBm	η_D	45	50	-	%

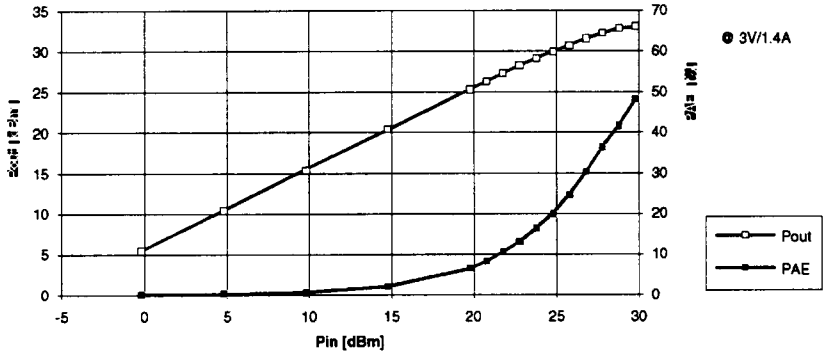
*) pulsed measurement; duty cycle 1:10; t_{on} = 1ms, power matching conditions.

SIEMENS GaAs FET **CLY 15**

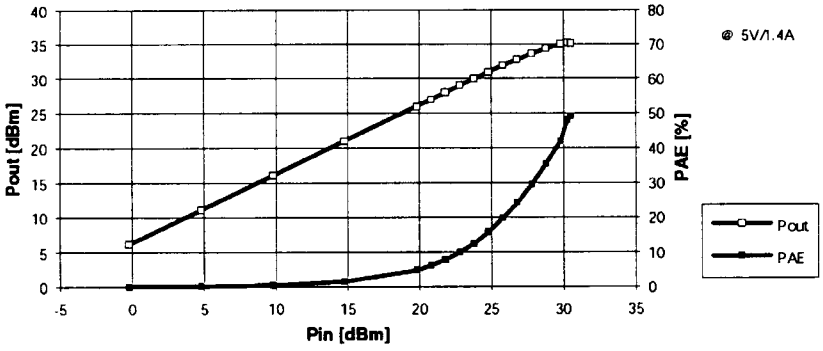
Output Characteristics



Power Characteristics



Power Characteristics



SIEMENS GaAs FET **CLY 15**

typ. Common Source S-Parameter

V_{DS} = 3V I_D=1.4A Z_O=50Ω

f	S11		S21		S12		S22	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
0.200	0.91	-150.9	5.69	99.7	0.01	48.5	0.90	176.0
0.250	0.91	-160.6	4.63	93.5	0.02	45.9	0.87	173.8
0.300	0.90	-167.9	3.89	88.7	0.02	46.8	0.88	171.8
0.350	0.90	-173.7	3.34	84.6	0.02	47.4	0.87	170.8
0.400	0.90	-178.7	2.92	80.9	0.02	47.5	0.87	168.8
0.450	0.90	176.9	2.60	77.5	0.02	47.6	0.87	167.3
0.500	0.90	173.0	2.34	74.6	0.02	48.1	0.87	165.8
0.550	0.90	169.5	2.12	71.4	0.02	47.7	0.87	164.2
0.600	0.90	166.1	1.95	68.7	0.03	47.0	0.87	162.8
0.650	0.90	163.1	1.79	66.1	0.03	47.1	0.87	161.2
0.700	0.90	160.0	1.66	63.5	0.03	46.6	0.87	159.7
0.750	0.90	157.2	1.54	60.9	0.03	45.6	0.87	158.3
0.800	0.90	154.6	1.45	58.6	0.03	45.0	0.87	156.9
0.850	0.90	152.0	1.36	56.1	0.03	43.9	0.87	155.6
0.900	0.90	149.3	1.28	53.8	0.04	43.0	0.87	154.0
0.950	0.90	146.9	1.21	51.5	0.04	41.9	0.87	152.6
1.000	0.90	144.5	1.15	49.0	0.04	41.0	0.87	151.3
1.200	0.91	135.2	0.95	40.3	0.05	36.1	0.87	145.8
1.400	0.91	126.7	0.81	31.8	0.05	31.9	0.88	140.1
1.600	0.92	118.5	0.70	23.8	0.06	26.1	0.88	134.7
1.800	0.92	110.6	0.61	16.3	0.06	20.8	0.88	129.7
2.000	0.93	103.2	0.55	8.7	0.06	15.6	0.89	124.3
2.200	0.93	96.3	0.49	2.1	0.07	10.4	0.88	119.1
2.400	0.93	89.3	0.44	-4.1	0.07	5.2	0.90	114.4
2.600	0.94	82.8	0.40	-10.0	0.07	0.2	0.90	109.3
2.800	0.94	77.0	0.37	-14.9	0.07	-4.2	0.90	104.5
3.000	0.94	71.3	0.34	-19.6	0.08	-9.7	0.91	99.8
3.200	0.93	66.0	0.32	-23.4	0.08	-15.0	0.92	95.1
3.400	0.92	61.6	0.31	-26.8	0.08	-19.4	0.93	90.8
3.600	0.91	57.3	0.30	-29.7	0.07	-23.7	0.92	87.0
3.800	0.90	53.1	0.31	-33.1	0.07	-28.1	0.93	83.1
4.000	0.89	49.2	0.32	-38.1	0.07	-31.9	0.93	79.8
4.200	0.86	46.4	0.34	-44.9	0.07	-35.4	0.92	76.4
4.400	0.83	44.7	0.36	-55.4	0.07	-37.5	0.92	73.4
4.600	0.89	44.2	0.07	-36.2	0.07	-38.1	0.92	71.0
4.800	0.83	43.7	0.34	-80.6	0.07	-39.4	0.92	68.2
5.000	0.85	42.2	0.30	-92.1	0.07	-40.3	0.92	65.2
5.200	0.88	39.4	0.27	-100.8	0.07	-42.5	0.92	62.2
5.400	0.89	36.5	0.24	-107.8	0.07	-45.0	0.92	58.7
5.600	0.90	33.1	0.22	-113.6	0.07	-48.6	0.92	55.7
5.800	0.91	29.6	0.19	-118.9	0.07	-51.4	0.92	52.1
6.000	0.92	26.4	0.18	-124.4	0.07	-54.4	0.92	48.0

SIEMENS GaAs FET **CLY 15**

typ. Common Source S-Parameter
V_{DS} = 5V I_D=1.4A Z_O=50Ω

f	S11		S21		S12		S22	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
0.200	0.90	-151.1	7.61	98.8	0.01	46.5	0.84	176.7
0.250	0.89	-160.6	6.18	92.4	0.02	43.4	0.82	174.7
0.300	0.89	-167.8	5.19	87.5	0.02	46.5	0.82	172.9
0.350	0.89	-173.7	4.45	83.3	0.02	46.0	0.82	171.8
0.400	0.88	-178.7	3.90	79.4	0.02	46.5	0.82	169.7
0.450	0.89	177.0	3.47	75.9	0.02	47.3	0.82	168.3
0.500	0.88	173.2	3.11	72.8	0.02	47.9	0.82	166.7
0.550	0.88	169.6	2.82	69.5	0.02	47.8	0.82	165.5
0.600	0.89	166.4	2.59	66.6	0.03	47.4	0.82	163.9
0.650	0.88	163.1	2.38	63.9	0.03	47.4	0.82	162.6
0.700	0.89	160.3	2.20	61.1	0.03	46.5	0.82	161.0
0.750	0.89	157.5	2.05	58.4	0.03	45.6	0.82	159.6
0.800	0.89	154.9	1.91	55.9	0.03	45.3	0.82	158.0
0.850	0.89	152.1	1.79	53.2	0.03	44.8	0.82	156.8
0.900	0.89	149.7	1.69	50.7	0.03	43.9	0.82	155.4
0.950	0.89	147.1	1.59	48.4	0.04	42.7	0.82	154.1
1.000	0.89	144.7	1.51	45.7	0.04	42.0	0.82	152.8
1.200	0.89	135.5	1.24	36.2	0.04	37.8	0.83	147.3
1.400	0.90	127.1	1.04	27.1	0.05	32.2	0.83	141.9
1.600	0.91	119.1	0.90	18.3	0.05	27.4	0.84	136.8
1.800	0.92	111.1	0.78	10.1	0.06	22.5	0.84	131.6
2.000	0.92	103.7	0.68	2.1	0.06	18.2	0.85	126.3
2.200	0.93	96.6	0.61	-5.1	0.06	12.5	0.86	121.3
2.400	0.93	89.8	0.54	-12.0	0.07	7.3	0.86	116.1
2.600	0.93	83.2	0.48	-18.6	0.07	2.6	0.87	111.4
2.800	0.93	77.3	0.43	-24.0	0.07	-2.6	0.88	106.3
3.000	0.93	71.8	0.39	-29.3	0.07	-7.2	0.89	101.8
3.200	0.92	66.6	0.37	-33.5	0.07	-12.1	0.90	97.2
3.400	0.92	61.8	0.34	-37.2	0.07	-16.8	0.91	92.5
3.600	0.91	57.9	0.32	-40.5	0.07	-21.1	0.91	88.8
3.800	0.90	54.1	0.32	-43.9	0.07	-24.9	0.92	85.1
4.000	0.88	50.5	0.31	-48.3	0.07	-27.6	0.92	81.4
4.200	0.87	47.8	0.32	-53.8	0.07	-31.5	0.92	78.1
4.400	0.86	45.7	0.32	-60.9	0.07	-33.4	0.92	74.9
4.600	0.85	43.4	0.32	-68.9	0.07	-35.4	0.92	72.2
4.800	0.85	42.3	0.31	-77.5	0.07	-37.2	0.92	69.3
5.000	0.86	40.3	0.30	-86.7	0.07	-39.3	0.92	66.2
5.200	0.87	37.7	0.28	-94.5	0.07	-41.6	0.92	63.1
5.400	0.88	35.2	0.26	-101.8	0.07	-44.0	0.92	59.6
5.600	0.89	32.1	0.24	-108.4	0.08	-48.5	0.92	56.6
5.800	0.90	29.0	0.22	-114.5	0.08	-50.3	0.92	53.0
6.000	0.90	26.0	0.21	-121.1	0.08	-54.0	0.93	49.0

2-4 PARAMETER EXTRACTION OF ACTIVE DEVICES

2-4-1 Introduction

We have already seen that all linear models that are used on the market are really derived from large-signal models, and it is our point that the quality of the model depends mostly on the quality of the parameters used to describe the model.

As far as modeling is concerned there are two options. The first option is a physics-based model, which can be applied to items such as diodes and bipolar transistors, including its most advanced versions, such as HBTs, SiGe versions, and other future derivatives. The bipolar transistor is essentially a combination of two diodes whereby the base–emitter junction generates the electrons or holes (depending on how the reader is accustomed to viewing the process) and then transfers those to the collector. The emitter, as its name states, emits the charged particles and the collector collects them, minus some current losses, which are expressed in the current gain of the device. A current gain (β) of 100 means that the collector receives 1% less than the emitter emits.

The second modeling option involves the construction of an analytical mathematic equivalent for the device modeled. Earlier in this chapter, we went through a detailed physics-based derivation about the inner workings of these devices, explaining how, for a given dc bias point, one can simplify them to a small-signal equivalent circuit. This is possible since we assume that the RF current and voltages are less than 1% of the dc equivalent currents and voltages. Reducing a device's complex nonlinear equivalent circuit to a quasistatic linear equivalent circuit allows its manufacturer to generate and publish sets of *S* parameters for it. These parameters are bias, voltage, and temperature dependent. At higher frequencies, parasitics play an enormous role in determining a device's RF performance.

A device operating at signal levels that are more than 1% of its dc equivalent currents and voltages must be evaluated in terms of large-signal performance. Most of the large-signal-equivalent circuits for nonbipolar devices, such as FETs (MOS, silicon JFETs, and members of the GaAsFET family), are analytic approximations that essentially bear no physical insight into the inner workings of the transistors. This fact is particularly painful because to look at high-order intermodulation distortion, one needs to have third and fourth derivatives of mathematically continuous equations. The SPICE approach cannot easily be translated into modern harmonic-balance simulators because such cop-outs as IF THEN ELSE statements in programming provide everything else but a continuous model. In the case of our own work, we used complicated curve-fitting equations to obtain analytic equations and their derivatives to accurately provide a large-signal model that is valid over a wide range of dc and RF.

2-4-2 Typical SPICE Parameters

Since we are about to evaluate bipolar microwave transistors, junction FETs, MOSFETs (model level 3), and GaAsFETs, Tables 2-20, 2-21, 2-22, and 2-23 list typical parameters for the devices we have used. These parameters can be obtained by the Scout program with the appropriate measurements. The BSIM model for MOSFETs, applicable for submicron technology transistors, requires an enormous level of parameter extraction and has not fully been validated for the LDMOS-type transistors currently favored for RF and microwave applications.

Table 2-20 BFR193W BJT

IS = 2.738E – 16	BF = 125	NF = 0.95341
VAF = 24	IKF = 0.26949	ISE = 1.0627E – 14
NE = 1.935	BR = 14.267	NR = 1.4289
VAR = 3.8742	IKR = 0.037925	ISC = 3.7409E – 17
NC = 0.94371	RB = 15	IRB = 0.00091763
RBM = 1.8368	RE1 = 0.76534	RC2 = 0.11938
CJE = 1.1824E – 15	VJE = 0.70276	MJE = 0.48654
TF = 1.8828E – 11	XTF = 0.69477	VTF = 0.8
ITF = 0.00096893	PTF = 0	CJC = 9.3503E – 13
VJC = 1.1828	MJC = 0.30002	XCJC = 0.053563
TR = 1.0037E – 09	VJS = 0.75	MJS = 0
XTB = 0	EG = 1.11	XTI = 3
FCC = 0.72063	LB = 0.57E – 9	LC = 0.00E – 9
LE = 0.43E – 9	CBCP = 0.101E – 12	CCEP = 0.175E – 12
CBEP = 0.061E – 12	VCMX = 10 V	

The meaning and significance of the various parameters is best explored in a book on SPICE or semiconductor physics [22–24].

2-4-3 Noise Modeling

Diode Noise Model. The noise model for the diodes (Figure 2-122 and Table 2-24) consists of two contributions: the shot noise and the flicker noise. The shot noise is computed automatically and does not require any parameters. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications).
2. Using bias-dependent flicker noise coefficients (specifying KF and AF at multiple bias points).

The noise generators in the diode noise model are the series parasitic resistance, R_S, and the intrinsic junction. Figure 2-122 illustrates the intrinsic junction noise generator. Let Δf be the bandwidth (usually normalized to 1 Hz). The intrinsic noise generator has a mean-square value of

$$\langle i_{Dn}^2 \rangle = 2qI_D\,\Delta f + \mathrm{KF}\,\frac{I_D^{\mathrm{AF}}}{f^{\mathrm{FCP}}}\,\Delta f \tag{2-266}$$

Notes on the Diode Noise Model:

1. Shot noise is always present unless the SN parameter is set to zero. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.

Table 2-21 2SK125 JFET[a]

IDSS	= 0.5250E – 01	VP0	= –0.3111E + 01	GAMA	= –0.1867E – 01	E	= 0.1520E + 01
KE	= –0.3856E – 03	SL	= 0.2818E – 01	KG	= –0.2398E + 00	T	= 0.0000E + 00
SS	= 0.7448E – 04	IG0	= 0.2000E – 14	AFAG	= 0.3846E + 02	IB0	= 0.1000E – 04
AFAB	= 0.3800E + 02	VBC	= 0.3000E + 02	R10	= 0.1711E + 02	KR	= 0.0000E + 00
C10	= 0.6609E – 11	K1	= 0.1675E + 01	C1S	= 0.6818E – 33	CF0	= 0.7261E – 11
KF	= 0.1156E + 01	RG	= 0.5000E + 00	RD	= 0.1542E + 01	RS	= 0.1333E + 01
LG	= 0.6098E – 09	LD	= 0.5159E – 08	LS	= 0.1482E – 08	CDS	= 0.4813E – 16
CGE	= 0.1590E – 11	CDE	= 0.3394E – 26	CGSP	= 0.8282E – 13	CDSP	= 0.4832E – 12
ZGT	= 0.5000E + 02	LGT	= 0.4712E – 01	ZDT	= 0.5000E + 02	LDT	= 0.3998E – 01
CGDP	= 0.3653E – 12	ZST	= 0.5000E + 02	LST	= 0.1495E – 01	CGDE	= 0.3831E – 12
CGSB	= 0.3120E – 13	CDSB	= 0.5896E – 12	VDMX	= 10		

[a]The currently implemented model for the junction FET is too primitive for serious RF applications. We have therefore taken the approach (liberty) of using the Materka parameter extraction approach for silicon junction FETs. This has resulted in unparalleled high-quality parameters; in particular, the knee voltage behavior has significantly improved, as well as the overall frequency response.

Table 2.22 GaAs MESFET

IDSS = 0.1077E+00	VP0 = –0.1800E+01	GAMA = –0.5741E-01	E = 0.1290E+01
KE = –0.1155E-01	SL = 0.1652E+00	KG = –0.1782E+00	T = 0.0000E+00
SS = –0.1208E+01	IG0 = 0.2130E-11	AFAG = 0.2740E+02	IB0 = 0.5680E-09
AFAB = 0.1826E+01	VBC = 0.9000E+01	R10 = 0.8382E+01	KR = 0.6359E+00
C10 = 0.5964E-12	K1 = 0.1296E+01	C1S = 0.0000E+00	CF0 = 0.6110E-13
KF = 0.9775E+00	RG = 0.1996E+01	RD = 0.1296E+01	RS = 0.1234E+01
CDS = 0.7852E-13	CDSD = 0.1000E-07	RDSD = 0.1581E+03	CGE = 0.1609E-12
CDE = 0.8674E-13	VDMX = 8		

Table 2.23 The 1 μm × 750 μm level 3 LDMOS FET

CBD = 0.863E-12	CGD0 = 166E-12	CGS0 = 246E-12	GAMA = 0.211
IS = 6.53E-16	KAPA = 0.809	MJ = 0.536	NSUB = 1E+15
PB = 0.71	PBSW = 0.71	PHI = 0.579	RD = 39
RS = 0.1	THET = 0.588	TOX = 4E-8	U0 = 835
VMAX = 3.38E5	VT0 = 2.78	XQC = 0.41	

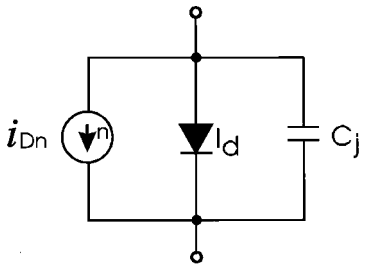


Figure 2-122 Equivalent noise circuit for a diode chip.

Table 2.24 Diode noise model keywords

Keyword	Description	Unit	Default
ID	Required bias current for the data point	Ampere	
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner fequency	hertz	

2. If the value of KF is specified as zero, then the flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.
3. The corner frequency noise model uses the system noise floor to internally compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the diode parameters and *kT*.
4. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

BJT Noise Model. The noise model for the Gummel–Poon BJT model consists of two contributions: shot noise and flicker noise. The shot noise is computed automatically and does not require any parameters. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications) (see Table 2-25).
2. Using bias-dependent flicker noise coefficients (specifying KF and AF at multiple bias points).

Option 1: *Specifying the Bias-Independent Flicker Noise Coefficient*. This option involves the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model.

1. Shot noise is always present unless it is turned off. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying The Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows a bias-dependent flicker noise coefficient (i.e., KF and AF vary with the bias point).

Notes on the BJT Noise Model:

Table 2.25 BJT noise model keywords

Keyword	Description	Unit	Default
IB	Required base bias current for the data point	ampere	
VCE	Required collector–emitter voltage for the data point	volt	
VBS	Base–substrate voltage required for LPNP type when four nodes are used	volt	
VCS	Collector–substrate voltage required for NPN or PNP type when four nodes are used	volt	
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

1. KF, AF, and FC can be specified as bias dependent. If only one set of noise data is specified, the corresponding bias point is not meaningful because all parameters are considered constant over all bias values. However, the bias point is needed for the program to identify the data as bipolar noise data.
2. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
3. This noise model of course considers the actual operating temperature, which must be supplied to the model.

Figure 2-123 shows the BJT noise model. Let Δf be the bandwidth (usually normalized to a 1-Hz bandwidth). The noise generators introduced in the intrinsic device are shown below and have mean-square values of

$$\langle i_{bn}^2 \rangle = 2qI_B \,\Delta f + \text{KF}\,\frac{I_B^{\text{AF}}}{f^{\text{FCP}}}\,\Delta f \tag{2-267}$$

$$\langle i_{cn}^2 \rangle = 2qI_C \Delta f \tag{2-268}$$

$$\langle i_{R_{bb}}^2 \rangle = \frac{4kT}{R_{bb}}\,\Delta f \tag{2-269}$$

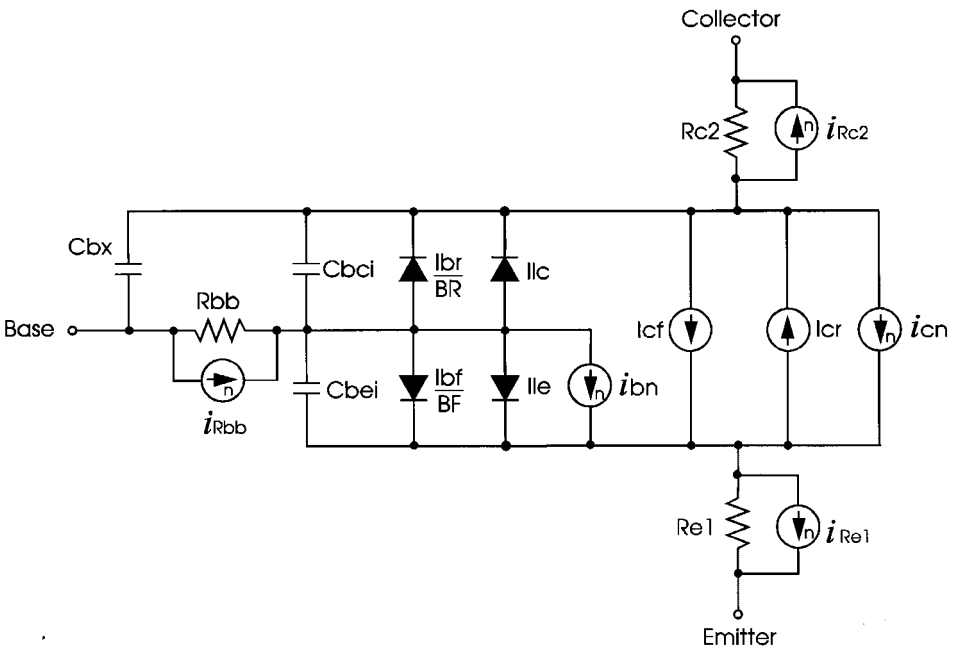


Figure 2-123 BJT noise model (not showing extrinsic parasitics). Current sources with *n* are noise sources.

$$\langle i_{R_{e1}}^2 \rangle = \frac{4kT}{R_{e1}} \Delta f \tag{2-270}$$

$$\langle i_{R_{c2}}^2 \rangle = \frac{4kT}{R_{c2}} \Delta f \tag{2-271}$$

$$I_B = \frac{I_{bf}}{BF} + I_{le} \tag{2-272}$$

$$I_C = I_{cf} - I_{cr} \tag{2-273}$$

JFET and MESFET Noise Models. The noise model for the FETs consists of two contributions: the shot noise and the flicker noise. There are two options to specify noise in the FET model:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (Table 2-26) to determine the flicker noise (this option is usually sufficient for most applications). The shot noise will be automatically computed using the SPICE equation.
2. Using bias-dependent flicker noise coefficients through a reference in the DATA block (specifying KF and AF at multiple bias points) and specifying the four noise parameters [$F_{\min}$, Γ_{opt} (magnitude and phase), and R_n] at multiple bias points.

Option 1: *Specifying the Enhanced SPICE Noise Model.* Option 1 is the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model. The drain noise model has the form

$$\langle |I_{dn}|^2 \rangle = 4K_B T \frac{2g_m}{3} \Delta f + \text{KF} \frac{|I_D|^{\text{AF}}}{f^{\text{FCP}}} \Delta f \tag{2-274}$$

Table 2.26 FET noise model keywords

Keyword	Description	Unit	Default
FN	Noise data measurement frequency	hertz	1.0 GHz
VGS	Required gate–source voltage for the data point	volt	
VDS	Required drain–source voltage for the data point	volt	
FMIN	Required minimum noise figure in dB at FN		
MGO	Required magnitude of optimum noise reflection coefficient at FN		
PGO	Required phase of optimum noise reflection coefficient at FN		
RN	Required normalized noise resistance at FN		
KF	Flicker noise coefficient		0.0
AF	Bias exponent of the flicker noise model		1.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

where the shot noise is derived from g_m and the flicker noise is proportional to KF and the drain channel current, I_D, and inversely proportional to frequency. The AF and FCP parameters tailor the flicker noise dependence on bias and frequency, respectively.

1. Shot noise is always present unless it is turned off. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, then flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying the Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows the specification of the complex bias-dependent nature of the shot noise and flicker noise. At high frequencies, the equivalent noise sources are correlated (the SPICE noise model does not account for this correlation). The complete evaluation of the shot noise sources can be determined from the four noise parameters. Since these are functions of bias, they can be specified over the (V_{GS}, V_{DS}) bias plane.

Additionally, a bias-dependent flicker noise coefficient (i.e., KF and AF vary with current) can be specified.

The MESFET noise model uses the four measured noise data [$F_{\min}$, Γ_{opt} (magnitude and phase), and R_n] at one frequency and multiple arbitrary bias points. The program uses these data and the FET model parameters to de-embed the noise data to an intrinsic noise model. The intrinsic model is accurate at all frequencies and, therefore, can predict the noise performance at all frequencies given data at just one frequency point. Built-in bias-dependent characteristics are used if multi-bias noise data are not provided.

Notes on the FET Noise Model

1. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
2. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

Noise in a MESFET is produced by sources intrinsic to the device. The same approach, but with different flicker corner frequencies, is highly applicable to JFETs and MOSFETs. For more detail as to simulation, see the Element library book for the active-device portion of Ansoft's Serenade Design Environment product. The equivalent noisy circuit of an intrinsic FET is represented in Figure 2-124.

The intrinsic FET is internally represented as a noiseless nonlinear two-port with one equivalent noise current connected across the gate–source terminal, and one across the drain–source terminal. The correlations of the gate and drain noise current sources are

$$\langle |I_{gn}|^2 \rangle = 4K_B T \, \Delta f \frac{\omega^2 C_{gs}^2}{g_m} R \tag{2-275}$$

$$\langle |I_{gn}|^2 \rangle = 4K_B T \, \Delta f \, g_m P \tag{2-276}$$

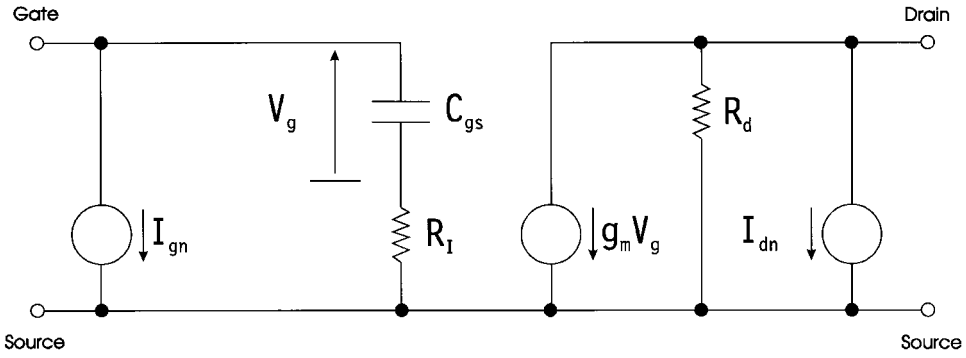


Figure 2-124 Equivalent noise circuit of an intrinsic FET device.

$$\langle I_{gn} I_{dn}^{*} \rangle = 4K_B T\, \Delta f\, j\omega C_{gs} \sqrt{PR}\, C \tag{2-277}$$

The correlation matrix of the noise current sources is

$$C_{dc}(\omega) = \frac{2}{\pi} K_B T\, d\omega \begin{bmatrix} \dfrac{\omega^2 C_{gs}^2}{g_m} R & -j\omega C_{gs}\sqrt{PR}\, C \\ j\omega C_{gs}\sqrt{PR}\, C & g_m P \end{bmatrix} \tag{2-278}$$

The gate and drain noise parameters R and P and the correlation coefficient C are related to the physical noise sources acting in the channel and are functions of the device structure and bias noise parameters. By defining measured noise parameters, $F_{\min}$, R_n, and Γ_{opt}, and using a noise de-embedding procedure, the parameters R, P, and C and the intrinsic noise correlation matrix of a FET device as functions of device bias are determined by the program.

In addition to the noise sources shown above, the flicker (1/f) noise can also be modeled by means of a noise current source connected in parallel with the intrinsic drain port. The flicker noise component in a narrow band, Δf, is expressed in the form

$$\langle |I_f|^2 \rangle = \mathrm{Q}\, \Delta f \frac{|I_D|^{\mathrm{AF}}}{f^{\mathrm{FCP}}} \tag{2-279}$$

where I_D is the instantaneous value of the channel current, and Q, AF, and FCP are empirical parameters. In most practical cases, AF and FCP are directly obtained from measurements (typically, AF = 2 and FCP = 1), while Q is not. In Ansoft's Serenade Design Environment, Q is either provided directly using KF or is computed by providing the flicker corner frequency (FC). FC is the frequency at which the flicker noise equals the shot/diffusion noise. The corner frequency is defined by the equation

$$\mathrm{Q} \frac{|I_D|^{\mathrm{AF}}}{f_{\mathrm{C}}^{\mathrm{FCP}}} = g_m P \tag{2-280}$$

Given the corner frequency FC and the measurement bias point V_{gs} and V_{ds}, the program automatically computes I_D, g_m, and P, and finally Q.

More information on FET noise modeling can be found in the literature [25–31].

MOSFET Noise Model. The MOSFET noise model (Figure 2-125 and Table 2-27) consists of two contributions: the shot noise and the flicker noise. The shot noise is computed automatically and does not require any parameters. It can be turned off by specifying SN = 0. The flicker noise can be specified in two ways:

1. Using the enhanced SPICE noise model by specifying KF, AF, and FCP in the model parameter list (this option is usually sufficient for most applications).

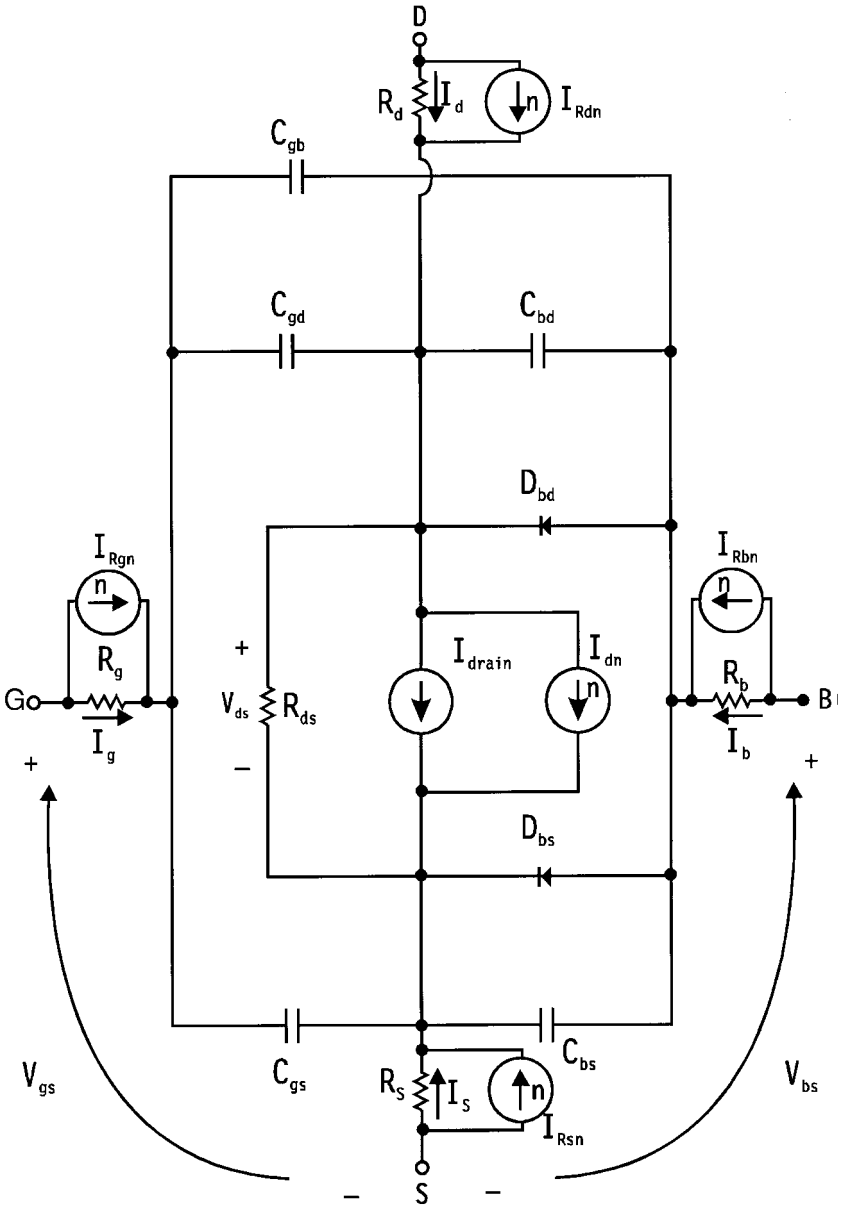


Figure 2-125 Equivalent noise circuit of an intrinsic MOSFET device.

Table 2.27 MOSFET noise model keywords

Keyword	Description	Unit	Default
VGS	Required gate–source bias for the data point	volt	
VDS	Required drain–source for the data point	volt	
VBS	Required drain–bulk for the data point	volt	
KF	Flicker noise coefficient		1.0E–13
AFT	Bias exponent of the flicker noise model		2.0
FCP	Frequency exponent of the flicker noise model		1.0
FC	Flicker noise corner frequency	hertz	

2. Using bias-dependent flicker noise coefficients through a reference (specifying KF and AF at multiple bias points).

Option 1: *Specifying the Enhanced SPICE Noise Model.* This option is the straightforward specification of KF, AF, and FCP that are constant with bias, as in the SPICE noise model (the flicker noise is considered bias dependent).

1. Shot noise is always present unless the SN parameter is set to zero. Turning noise off is useful for comparing the total circuit noise that is generated by the nonlinear devices and that generated by the linear circuit components.
2. If the value of KF is specified as zero, then the flicker noise will not be contributed by the device and only shot noise is considered in the intrinsic model.

Option 2: *Specifying the Bias-Dependent Flicker Noise Coefficient or Flicker Corner Frequency.* This option allows a bias-dependent flicker noise coefficient (i.e., KF and AF vary with drain current). The MOSFET noise model data are given and referenced by a model parameter.

Notes on the MOSFET Noise Model

1. KF, AF, and FC can be specified as bias dependent. If only one set of noise data is specified, the corresponding bias point is not meaningful because all parameters are considered constant over all bias values. However, the bias point is needed for the program to identify the data as MOSFET noise data.
2. The corner frequency noise model option uses the system noise floor to compute the flicker noise coefficient, KF. The system noise floor is computed by the program using the transistor parameters and kT.
3. This noise model, of course, considers the actual operating temperature, which must be supplied to the model.

Let Δf be the bandwidth (normalized to 1 Hz). The noise generators introduced in the intrinsic device are shown below and have mean-square values of

$$\langle i_{dn}^2 \rangle = \frac{8kTg_m}{3}\Delta f + \mathrm{KF}\frac{I_D^{\mathrm{AF}}}{f^{\mathrm{FCP}}}\Delta f$$

$$\langle i_{Rgn}^2 \rangle = 4\frac{kT}{R_g}\Delta f$$

$$\langle i_{Rdn}^2 \rangle = 4\frac{kT}{R_d}\Delta f$$

$$\langle i_{Rsn}^2 \rangle = 4\frac{kT}{R_s}\Delta f$$

$$\langle i_{Rbn}^2 \rangle = 4\frac{kT}{R_b}\Delta f$$

We include this MOSFET noise model (used for quite awhile) for completeness. At the moment, we do not know which MOSFET noise model the industry will settle on in the future.

Modern CAD tools, such as Ansoft's Serenade product, use these models to generate quite accurate noise data based on a good linear equivalent model. Internally Serenade uses the noise-correlation-matrix method (first introduced by Russer). The same noise-correlation techniques apply to BJTs, all FETs except MOS, and HBTs (such as SiGe devices).

2-4-4 Scalable Device Models

Since diodes and transistors are scalable, here are guidelines for how to scale them:

Microwave Diode

$$I_D = Area \times I_D$$
$$C_j = Area \times C_j$$
$$R_D = R_D/\text{Area}$$

PIN Diode

$$I_D = \text{Area} \times I_D$$
$$C_j = \text{Area} \times C_j$$
$$R_S = R_S/\text{Area}$$
$$R_{\max} = R_{\max}/\text{Area}$$

Bipolar

$$I_{bf} = \text{Area} \times I_{bf} \qquad I_{br} = \text{Area} \times I_{br}$$
$$I_{le} = \text{Area} \times I_{le} \qquad I_{lc} = \text{Area} \times I_{lc}$$
$$I_{cf} = \text{Area} \times I_{cf} \qquad I_{cr} = \text{Area} \times I_{cr}$$
$$C_{bc} = \text{Area} \times C_{bc} \qquad C_{be} = \text{Area} \times C_{be}$$
$$C_{bx} = \text{Area} \times C_{bx} \qquad R_{bb} = R_{bb}/\text{Area}$$
$$\text{RB2} = \text{RB2/Area} \qquad \text{RC2} = \text{RC2/Area}$$
$$\text{RE1} = \text{RE1/Area} \qquad I_{jss} = I_{jss} \times \text{Area}$$
$$C_{jss} = C_{jss} \times \text{Area}$$

Materka FET

$$\begin{aligned}
\text{IGSS} &= \text{IGSS} \times \text{Area} \\
\text{CGS0} &= \text{CGS0} \times \text{Area} \\
\text{CGS1} &= \text{CGS1} \times \text{Area} \\
\text{CDVC} &= \text{CDVC} \times \text{Area} \\
\text{CDVS} &= \text{CDVS} \times \text{Area} \\
R_G &= R_G \times \text{Area}/(\text{Number of fingers} \times 2) \\
R_D &= R_D/\text{Area} \\
R_S &= R_S/\text{Area}
\end{aligned}$$

Generating a Databank for Parameter Extraction. Given the fact that one can measure with an automated system (Rohde & Schwarz, Hewlett-Packard) the bias-dependent *S* parameters of any active device, bipolar and FET, and furthermore, that Hewlett-Packard makes a computer-controllable dc *I–V* curve tracer, we can generate a huge databank that allows us to have corresponding values for both RF and dc operating points. To validate this, one needs to run a large-signal simulator, such as Ansoft's Serenade Design Environment, or its equivalent from other manufacturers, enter the large-signal parameters, and generate a set of *S*-parameters that correspond to the dc values for which we have measurements. The agreement between measured and predicted points for both RF and dc values is a good measure to evaluate the accuracy of the parameter extraction and the simulator's performance. We cannot stress often enough, however, the fact that while most simulator models are quite good, the weakest point of the link is the parameter extraction program.

Currently, the best parameter extraction for bipolar transistors and FETs, excluding the BSIM model, is implemented in the Scout program by Ansoft, based on modified techniques found in the literature. These techniques are based on the fact that present-day nonlinear modeling of microwave devices, especially FETs, is not exactly adequate to describe all effects found within these devices. In the case of FETs, the most troublesome parameter is R_{DS}, which is really more bias dependent than practically all existing models take into consideration. Techniques used in published extraction methods extract nonlinear models independently for each section of a device model, and then assume that the proper response will be generated when all parts are put together. For instance, many extraction techniques independently fit the *I–V* data of the section describing the *I–V* curves and then fit the *C–V* data to the equivalent small-signal models. This results in an approximation to the bias-dependent response of the complete model.

In the technique used in the Scout program, model parameters are extracted simultaneously so the effects between model sections are accounted for and the device is treated as a whole. The most critical area happens to be in the linear region, typically around the pinchoff area, and for very small supply voltages. The Scout techniques use measured data taken at a wide variety of bias ranges and frequencies to define the model response. The possible solution region for the complete model becomes well identified and uniqueness of solution is vastly improved. The Scout program is interactive in nature and displays simultaneously measured versus predicted data. This allows for fine-tuning of the parameters.

Over the years, the industry has adapted a large number of models. Those supported by the Scout parameter extraction program include:

- *npn* Gummel–Poon BJT
- *pnp* Gummel–Poon BJT

- *npn* heterojunction BJT
- *pnp* heterojunction BJT
- General-purpose *n*-channel JFET
- General-purpose *p*-channel JFET
- Chalmers (Angelov) MESFET/HEMT
- Curtice–Ettenberg cubic MESFET/HEMT
- Curtice quadratic MESFET/HEMT
- IAF (Berroth) MESFET/HEMT
- Modified Materka–Kacprzak MESFET/HEMT
- Raytheon (Statz) MESFET/HEMT
- TriQuint (TOM-1 and TOM-2) MESFET/HEMT
- Ansoft physics-based MESFET

The Ansoft physics-based MESFET model is unique and politically sensitive because its users are essentially required to enter all device fabrication parameters to get the appropriate answer, and most, if not all, companies will stay away from handing out such company proprietary recipes. The validation of this model was done using manufacturing data for transistors that were developed under government contract and therefore accessible for the team members. Compact Software/Ansoft over the last ten years has been involved in a large number of government-funded research and development programs, and practically all modern modeling implementations and their validation were done in conjunction with the largest semiconductor manufacturers.

As far as further validation is concerned, besides our looking at the matching of dc *I–V* curves and *S* parameters, important parameters such as 1-dB compression point, power-added efficiency, and harmonic content are important criteria to evaluate the quality of such an undertaking. As far as hardware is concerned, one needs to have a network analyzer, such as the HP8510 or its equivalent. Its selection is determined by the highest cutoff frequency where measurements are necessary. A dc feed capability, up to several amperes, through the measuring ports, and appropriate power supplies are essential, as is a dc *I–V* curve tracer (HP4145). The built-in various optimizers of the program establish the best possible match between the selected parameters and the measured results.

The area of proper modeling has always been fascinating because the active-device model can make or break first-pass success. Figure 2-126 shows the doubler gain comparison of a model supplied by NEC (the device manufacturer, parameter extraction done with Scout, and parameter extraction supplied from the *Microwave Engineering Europe* (*MEE*) magazine in its CAD review of May 1994). It is somewhat unclear where the *MEE* model came from. The actual task given by *MEE* was to simulate a frequency doubler against measured data. This type of simulation required two areas of high precision, one being the model and the other being the electromagnetic modeling of the discontinuities of the circuit—in this case, a stub element. The circuit itself was very simple. The manufacturer-supplied data gave the poorest results (we believe the parameter extraction came from a third party not quite up to speed in this area). The results that came from the Scout parameter extraction and the third one that *MEE* supplied were fairly close, yet the Scout solution, including the electromagnetic simulation part, gave the best answer. More simulators have since appeared on the market, and it would be interesting to revisit this topic. Figure 2-127 shows a physical layout of the doubler circuit.

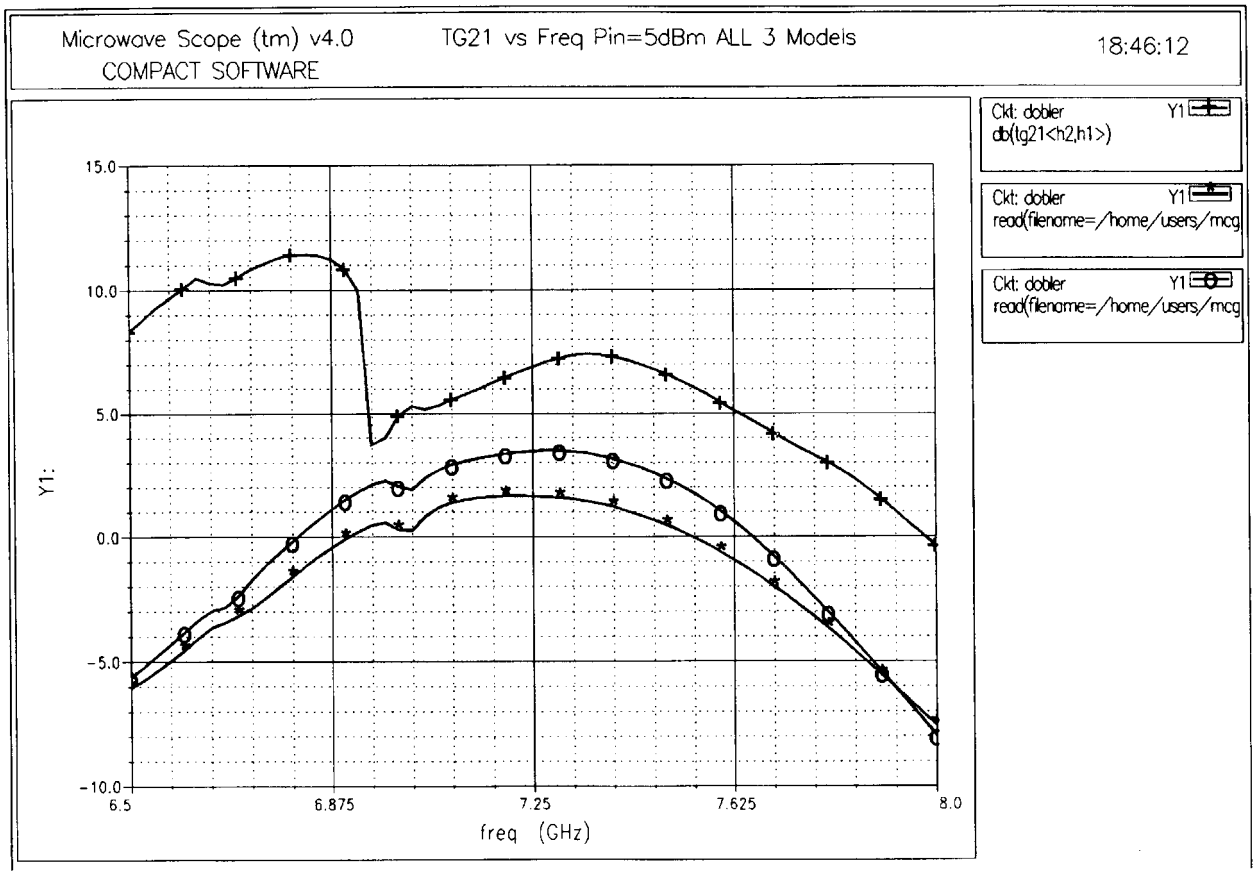


Figure 2-126 Gain comparison of a MESFET doubler based on the three device models used: NEC (+), Ansoft (O), and *Microwave Engineering Europe* (*).

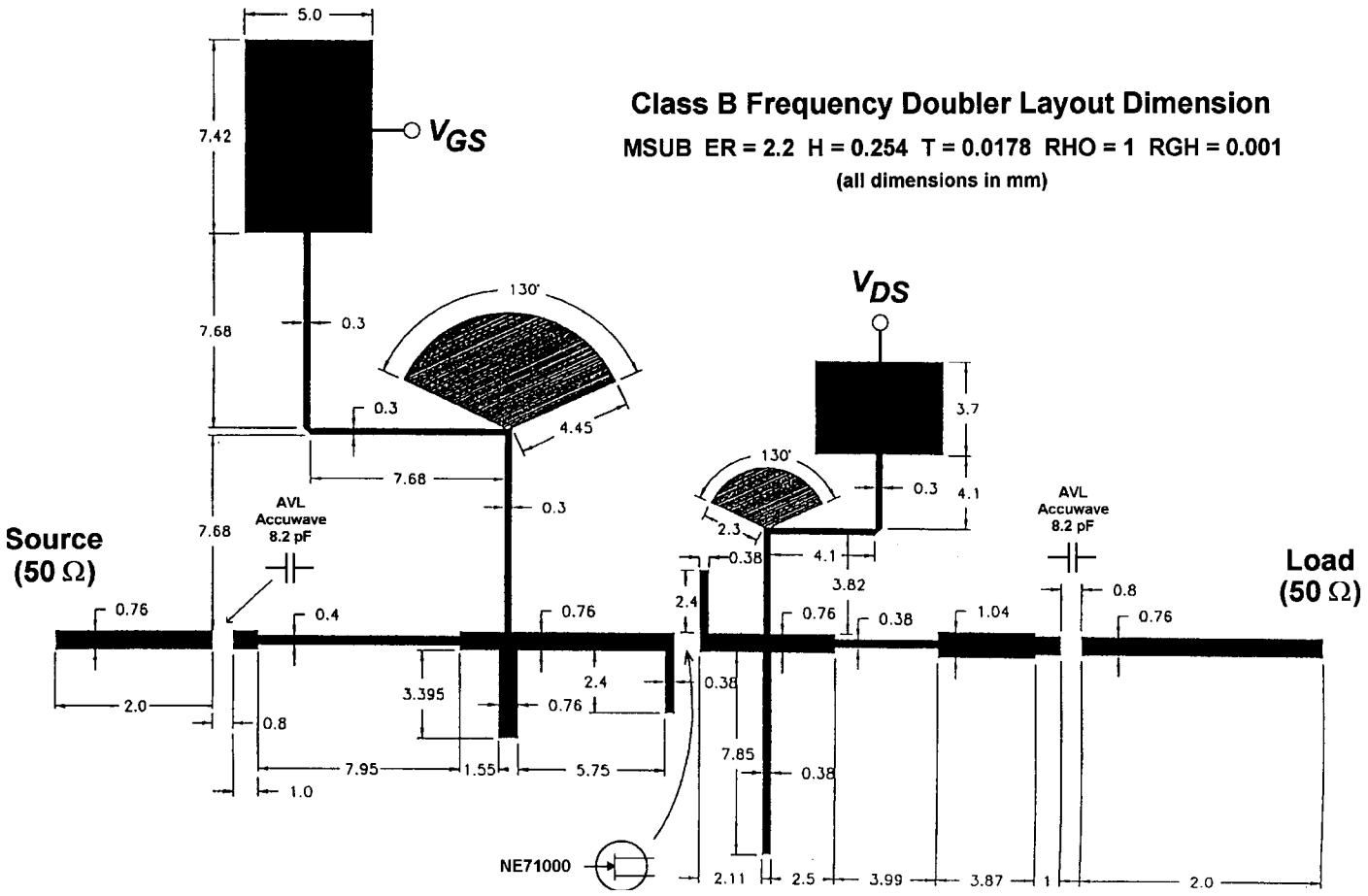


Figure 2-127 Layout of the MESFET doubler.

Test Setup. The measuring equipment required for the actual extraction includes:

- Semiconductor parameter analyzer (HP4145)
- Vector network analyzer (ZVR, HP8510, W360)
- Bias supplies and meters
- Power RF source, power meters, spectrum analyzer, and so on

The required software tools include:

- GPIB controller: Compact NETCOM (network analyzer control program)
- Parameter extraction performed by Scout
- Small-signal and harmonic-balance validation: Serenade nonlinear analysis (Microwave Harmonica) or its equivalent

After we have assembled all these tools, the device extraction process can begin. It will follow the flow as shown in Figure 2-128. Depending on the numbers for the dc/ac measurement, this process can take up to two hours. The Scout program, with its built-in optimizer, can take another two hours. It is therefore reasonable to assume a thorough parameter extraction of a device will take about one day.

Parameter Extraction. As we have outlined, the parameter extraction program Scout, which is optimized for high-frequency and millimeter-wave application, is a Windows-based parameter extraction and optimization program. It requires a network-communication program to acquire dc and *S* parameter measurement data. It supports interactive device fitting and optimization and supports the various models listed earlier. For ease of use, it has an attractive graphical user interface with multiple windows, as can be seen from Figure 2-129. Hewlett-Packard offers more general RF parameter extraction programs, but we did not have access to those. We also expect that some other companies will develop similar tools. An interesting program named HarPE, developed by Professor John Bandler in Canada, may have been withdrawn from the market after his company merger with HP. The program was acclaimed to provide quite good results using measured harmonics. One of the unfortunate deficiencies of most BJT parameter extraction programs known to us is the fact that (as one

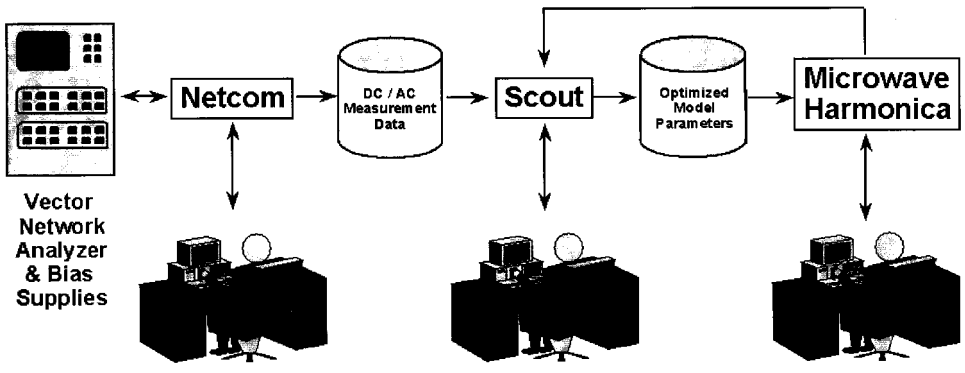


Figure 2-128 Flowchart of the parameter extraction process.

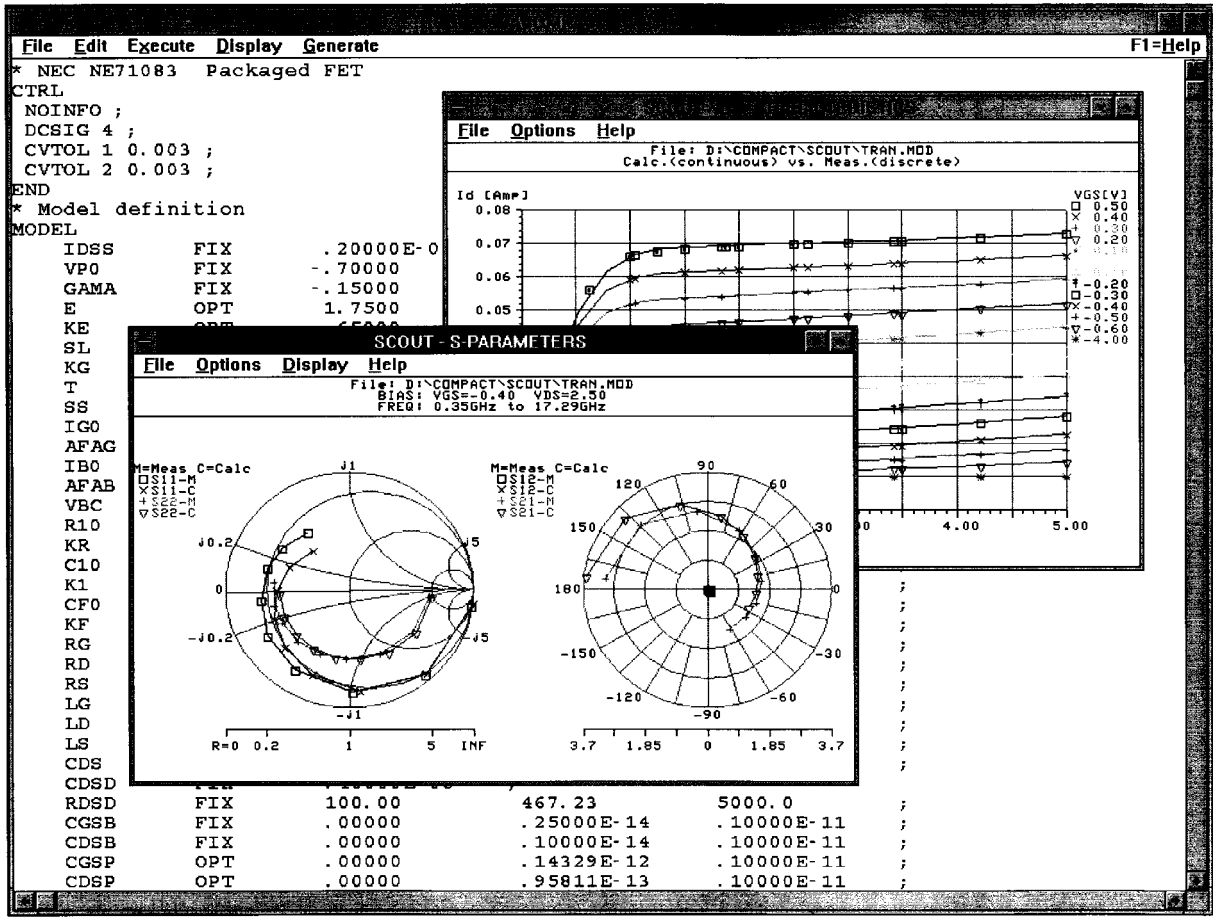


Figure 2-129 The Scout user interface.

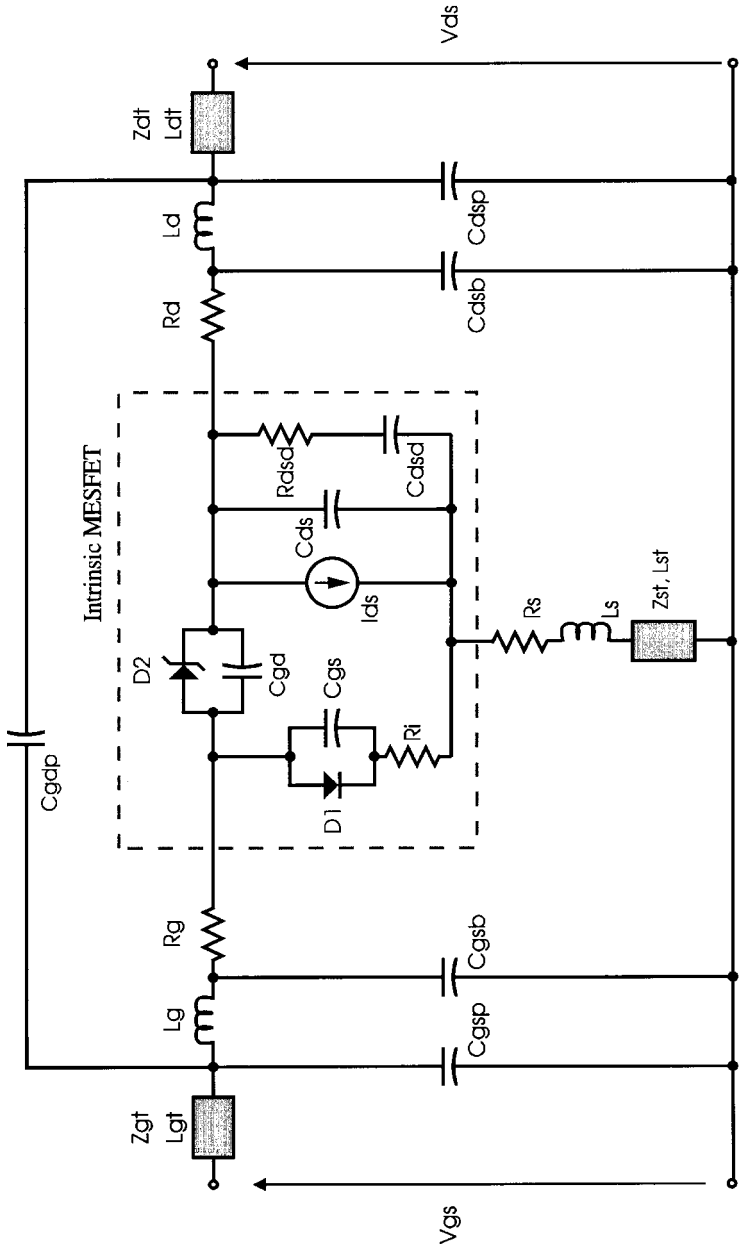


Figure 2-130 Intrinsic model and complete chip/package model.

example) the sensitivity to determine the actual base-spreading resistor for bipolar transistors is poor. The only useful workaround is considering the noise figure of the device, given the fact that the midrange noise factor of any bipolar transistor can be calculated from

$$F = 1 + \frac{R'_{bb}}{R_g} + \frac{0.5(26\ \text{mV}/I_C)}{R_g}$$

(approximate equation at medium frequencies) (2-281)

By solving the equation for R'_{bb}, since we know that R_g, the generator impedance, is 50 Ω, and the emitter diffusion resistance (26 mV/I_C) can be determined from the bias point, we can already eliminate the need to extract R'_{bb} from the HP test equipment. By using HP's program, errors of a factor up to 10 are not uncommon because of the low sensitivity for this parameter. Another headache in the case of the BJT is the determination of the excess phase (PTF) and forward and reverse transit time (TF and TR, respectively). The lack of quality modeling can best be found by examining the measured versus predicted values of S_{22}. Again, practically all the parameter extraction programs we have seen so far suffer from this deficiency.

MESFETs. In the case of a member of the MESFET family, a model as shown in Figure 2-130 is used. It consists of an intrinsic model and a complete chip/package parasitic model. In the case of the MESFET, to obtain first-class results, the following measurements are necessary:

- Dc *I*–*V* measurements
- Dc diode measurements
- Cold-FET and forward-biased-gate *S* parameters
- *S*-parameter measurements over bias
- Harmonic power measurements
- *Z*-Parameter Equations for Figure 2-131

In order to simplify the modeling, it is useful to do cold-FET [Figure 2-131 and Eqs. (2-282), (2-283), and (2-284)] and pinched-FET [Figure 2-132 and Eqs. (2-285), (2-286), and (2-287)] measurements to derive some of the intrinsic or extrinsic values.

$$Z_{11} = R_s + R_g + R_{gs} + \alpha_g R_{ch} + j\omega(L_g + L_s) \tag{2-282}$$

$$Z_{12} = Z_{21} = R_s + \alpha R_{ch} + j\omega L_s \tag{2-283}$$

$$Z_{22} = R_s + R_d + 2\alpha R_{ch} + j\omega(L_s + L_d) \tag{2-284}$$

Capacitance Equations for Figure 2-132

$$C_b = -\frac{\text{Im}(Y_{12})}{\omega} \tag{2-285}$$

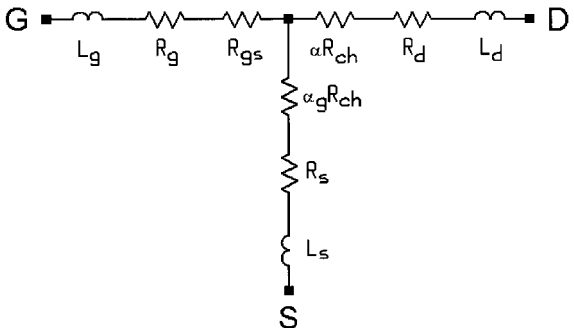


Figure 2-131 Forward-biased gate FET model for cold-FET (V_{DS} = 0, V_{GS} 0) measurements. This approach allows direct extraction of parasitic resistances and inductances.

$$C_{pg} = \frac{\mathrm{Im}(Y_{11})}{\omega} - 2C_b \tag{2-286}$$

$$C_{pd} + C_{ds} = \frac{\mathrm{Im}(Y_{22})}{\omega} - C_b \tag{2-287}$$

For the highest accuracy, the following extraction strategies are recommended:

- Pre-extraction using basic dc and *S* parameters—for example, C_{GS} from S_{11}
- Six extraction steps to determine parameters:
  Basic dc model
  Advanced dc model
  Basic ac model
  Advanced ac model
  Package determination
  Diode characterization
- Large-signal validation
  Power compression and bias shift
  Harmonic power comparison
  Intermodulation characterization

As an example, we are going to use the popular and well-characterized NE71000, which is not a particularly modern device, but has been used for various research projects. Figure

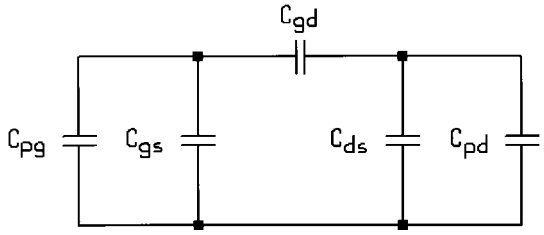


Figure 2-132 Pinched-FET (V_{DS} = 0, V_{GS} *VP*) model. This approach allows direct extraction of parasitic gate and drain capacitances.

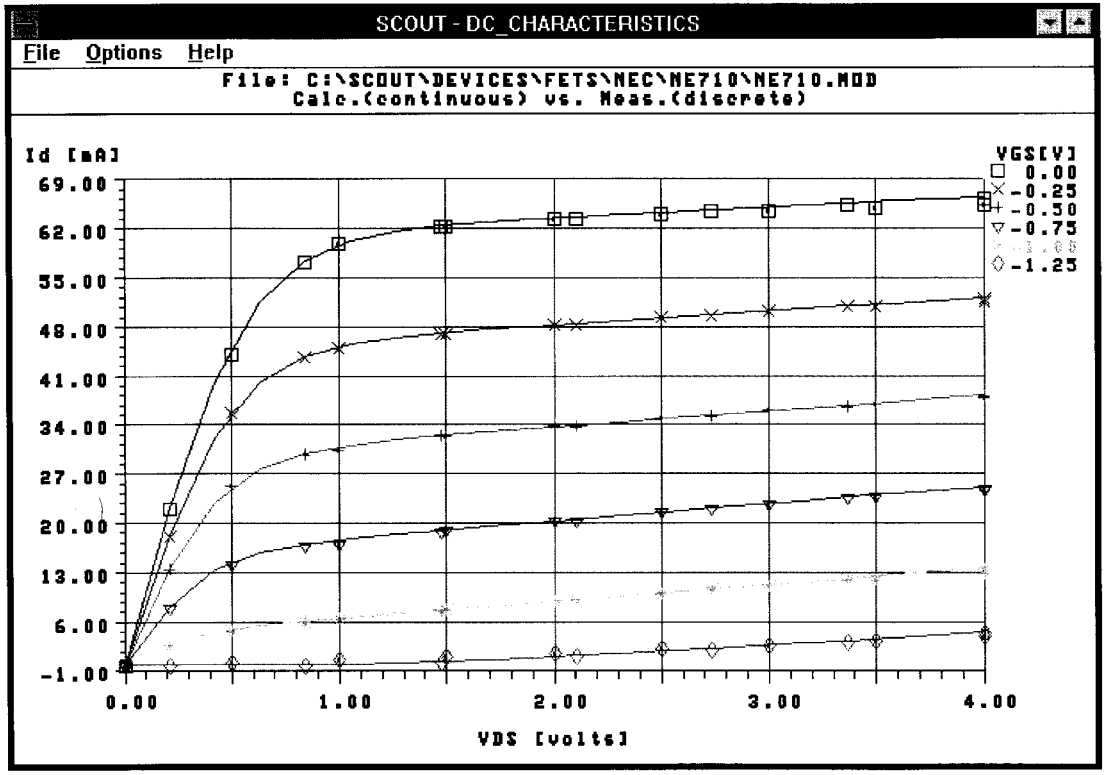


Figure 2-133 Measured NE71000 dc *I–V* curves versus those simulated using the Ansoft Materka model.

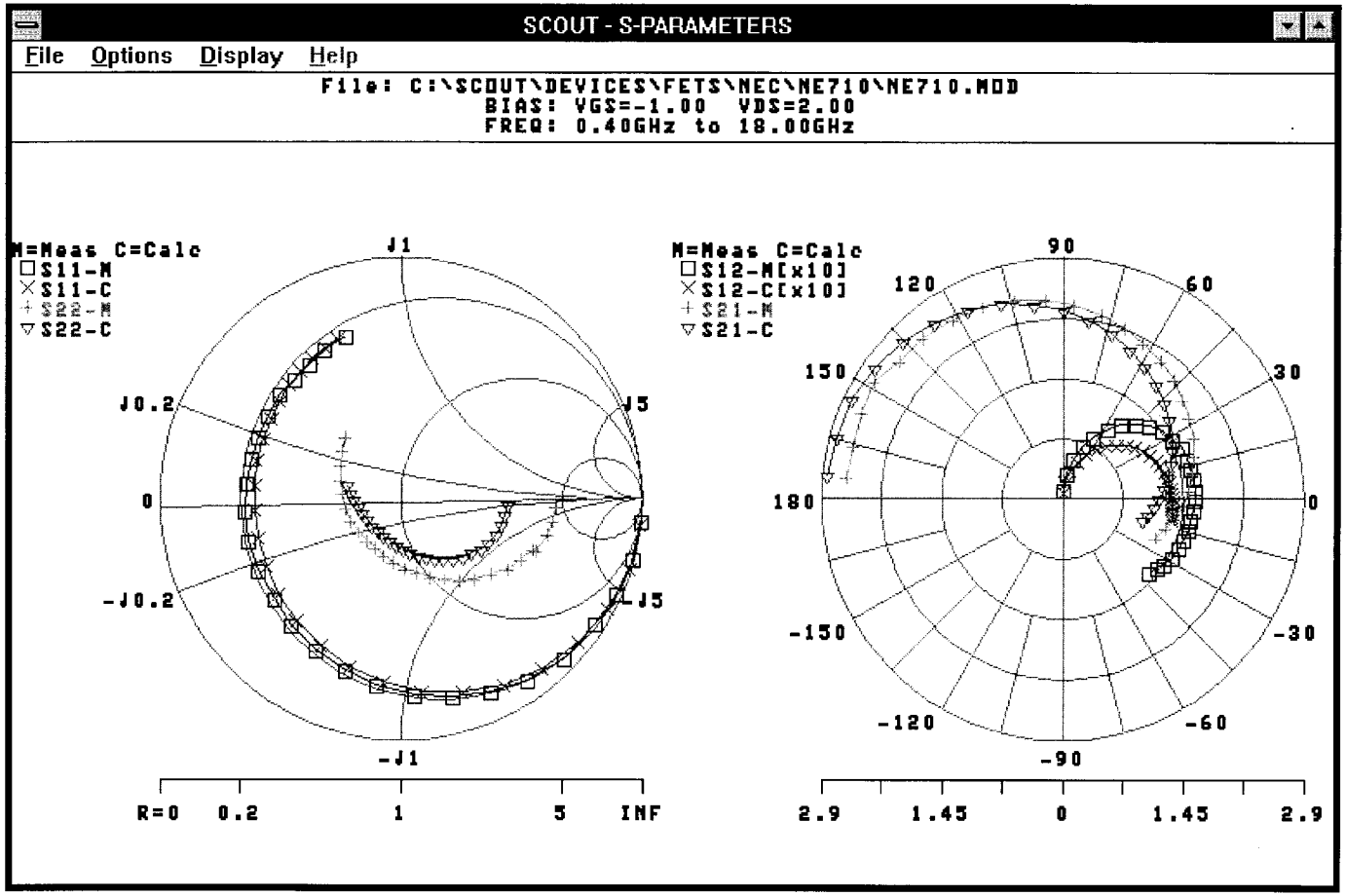


Figure 2-134 Measured versus calculated NE71000 *S* parameters for $V_{GS} = -1.00$ and $V_{DS} = 2.00$.

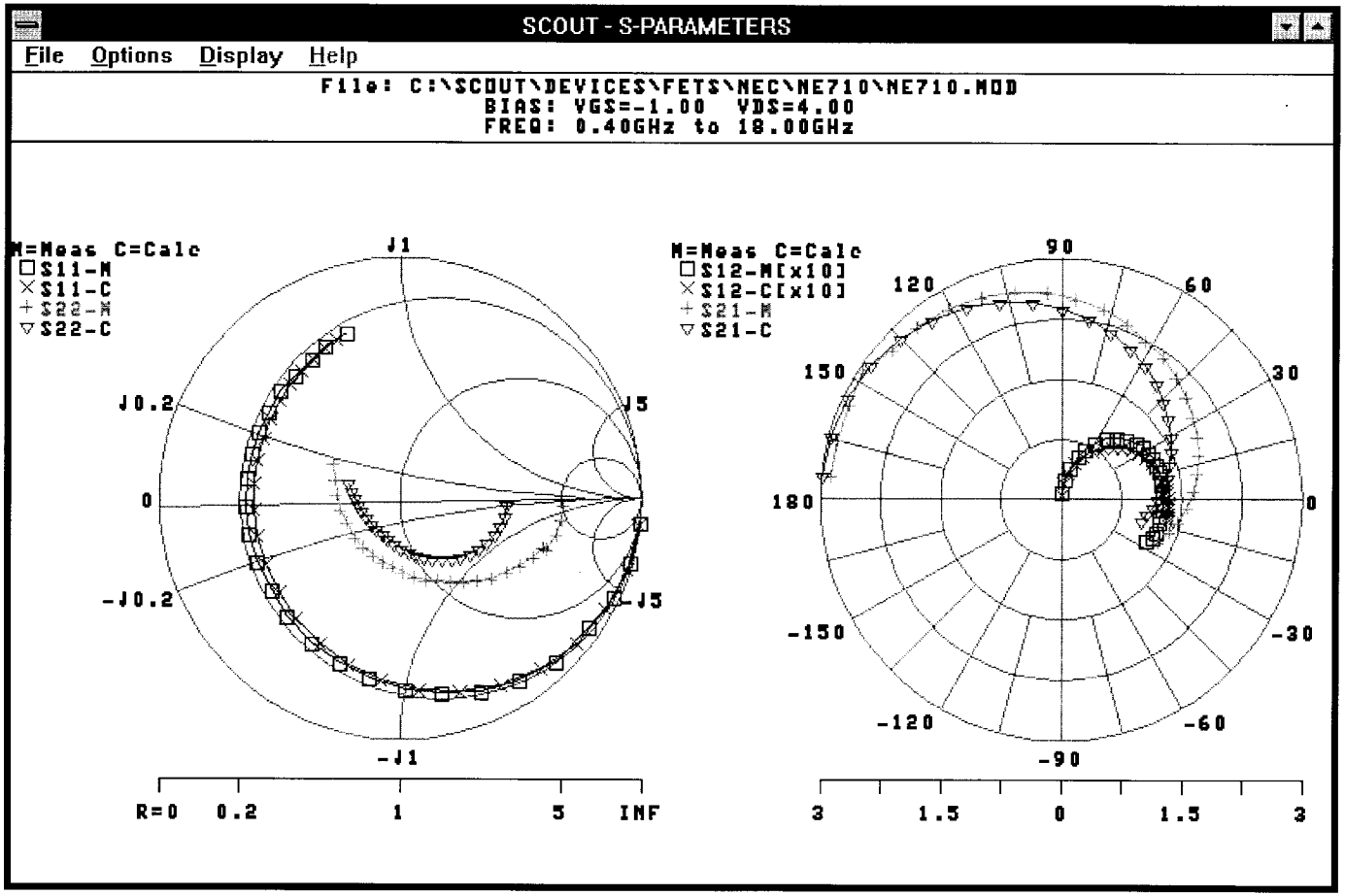


Figure 2-135 Measured versus calculated NE71000 *S* parameters for V_{GS} = –1.00 and V_{DS} = 4.00.

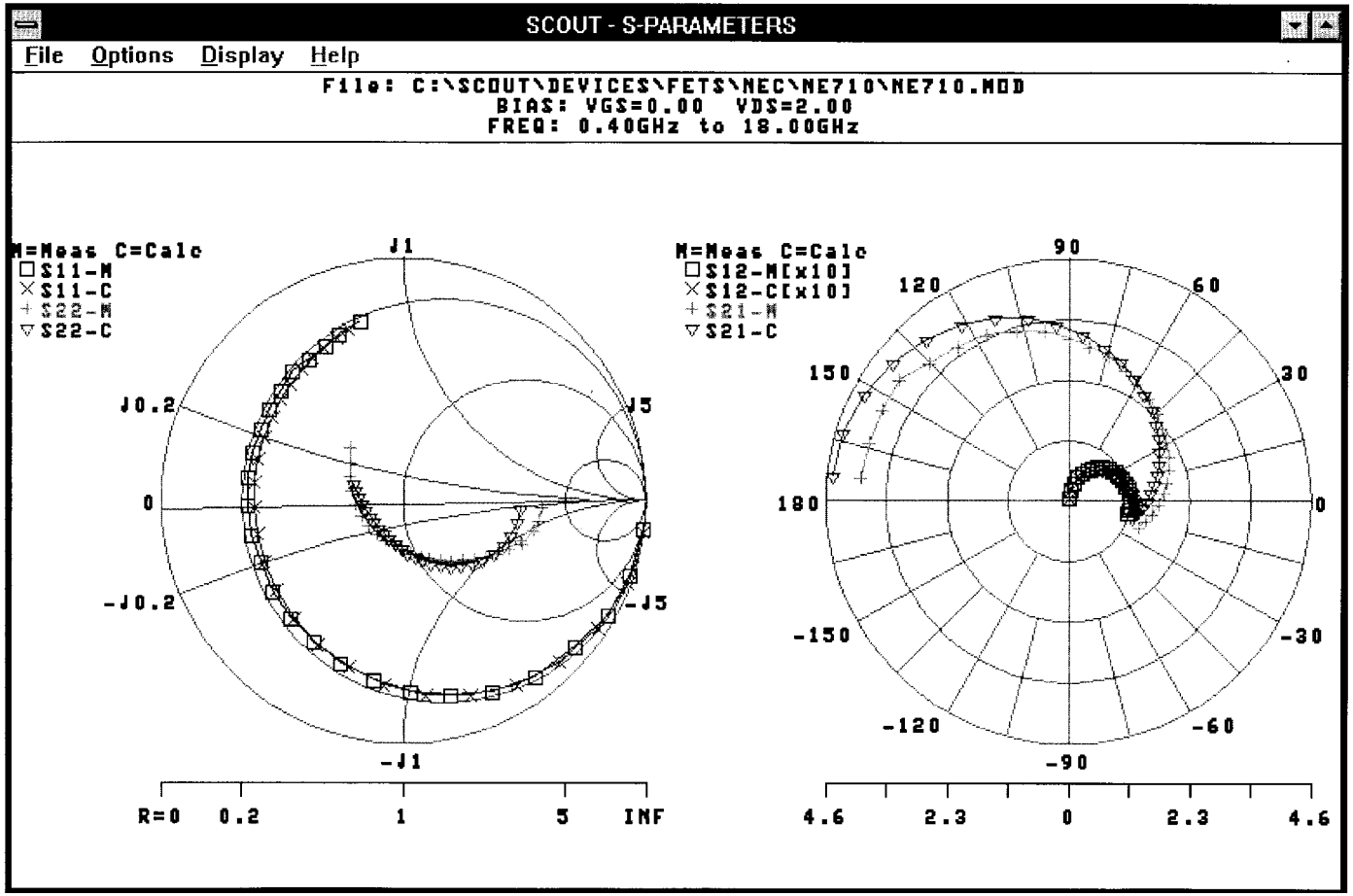


Figure 2-136 Measured versus calculated NE71000 *S* parameters for V_{GS} = 0.00 and V_{DS} = 2.00.

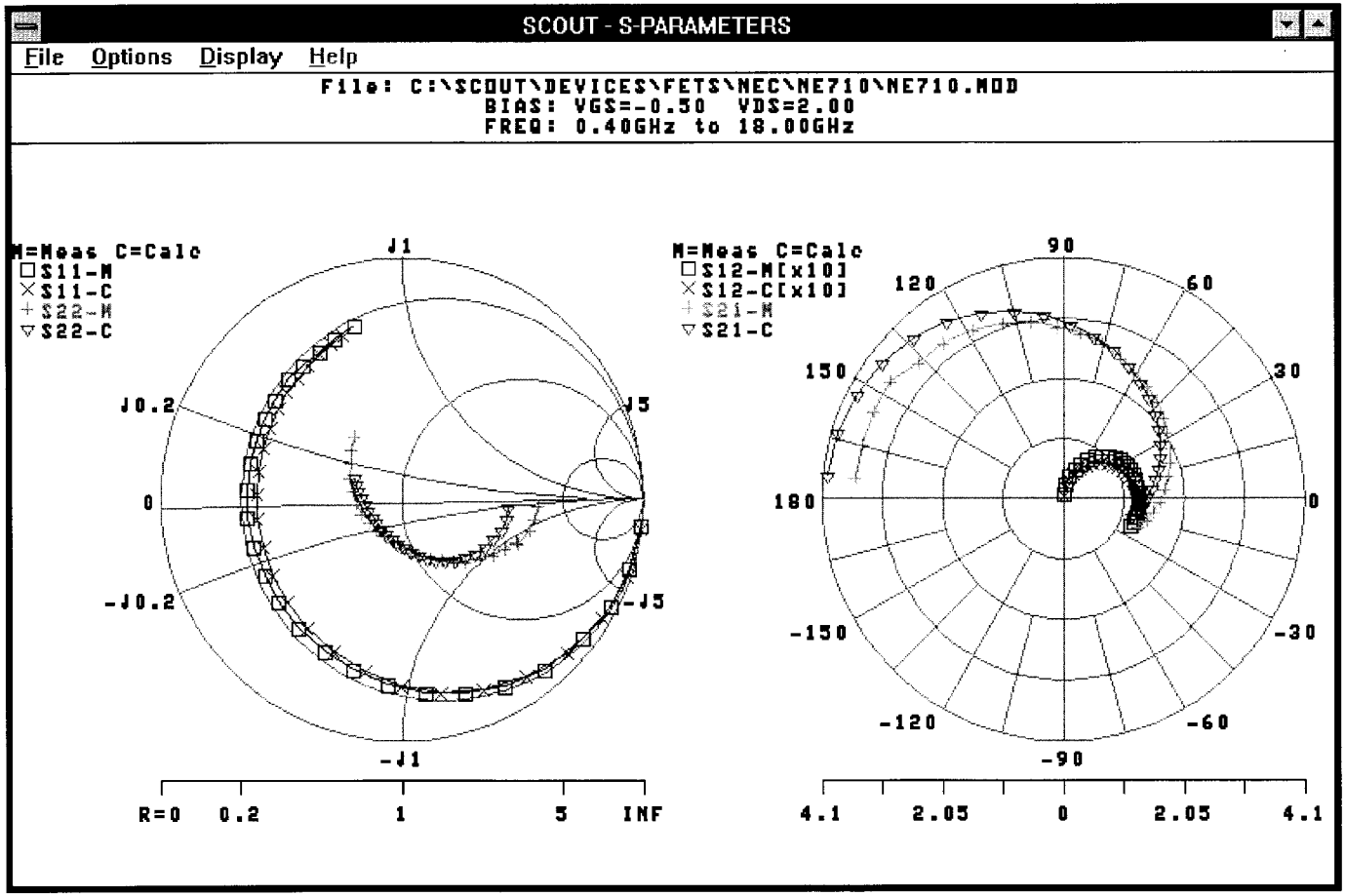


Figure 2-137 Measured versus calculated NE71000 *S* parameters for V_{GS} = –0.50 and V_{DS} = 2.00.

2-133 shows measured versus simulated dc I–V curves for the NE71000. The agreement between measured and predicted dc I–V curves is quite good. In a similar fashion, we obtain the S-parameter fits over bias. They are shown in Figures 2-134 through 2-137.

As mentioned above, 1-dB compression, power saturation, and drain current are functions of the drive. This particular model point is not particularly impressive. We believe this has to do with the original Materka model.

Having brought up the subject of models, Figure 2-138 compares measured data with that obtained with the Angelov, cubic, quadratic, TriQuint, Raytheon, and Materka models. The figure shows that there is no perfect model yet. To put this in perspective, in Figure 2-139 we show the dc I–V error bars for these six models. This graph shows that the least errors for the dc measurement versus modeling are obtained with the Curtice cubic model, followed by the modified Materka model (implemented in Serenade) and TriQuint models. The difference in S parameter matching for RF application is significantly more relevant for the designer.

Things become much more interesting if we compare the S-parameter errors between the Materka shown in Figure 2-140, the Curtice cubic model (Figure 2-141), and the TOM model (Figure 2-142). Improved modeling regarding the R_{DS} is being developed, and we will discuss this at the end of this chapter.

Since the trend is to go to low-voltage power RF operation, several issues must be battled:

- Models are most inaccurate at low V_{DS} and near pinchoff.
- Modeling of low-voltage operation is adversely impacted.
- Thermal effects must be considered, but older models do not account for this.

A Case Study. As a case study, we will consider the Oki KGF1608 power MESFET (0.5 μm × 28 mm). Figures 2-143 and 2-144 show the dc I–V curves and S parameters, respectively. The resolution of the S parameters, based on the scale used, is not particularly impressive.

Figure 2-145 shows the dc I–V curves, with load line, for the KGF1608; the self-heating effect of the transistor can be seen in the droop of the top two curves. Figures 2-146 and 2-147 show drain current and output power versus drive, respectively. The test conditions for these three figures were f = 850 MHz, V_{DS} = 3.4 V, and V_{DS} = 2.025 V.

2-4-5 Conclusions

From all plots of dc and ac error contours, dc I–V curves, and dc error bar charts we can draw the following conclusions.

Dc errors

- *In the Saturation Region*. All the models considered work very well, with a relative error of less than 15%. This means that all models are suitable for typical applications where the device operates in the saturation region.
- *In the Linear Region.* Though the errors in the saturation region are similar for different models, the errors in the linear region are quite different. The Materka, Curtice cubic, and TriQuint models behave better than other models in the linear region. The relative errors of these three models are less than 10%, while other models have more than 60% relative errors at some points.

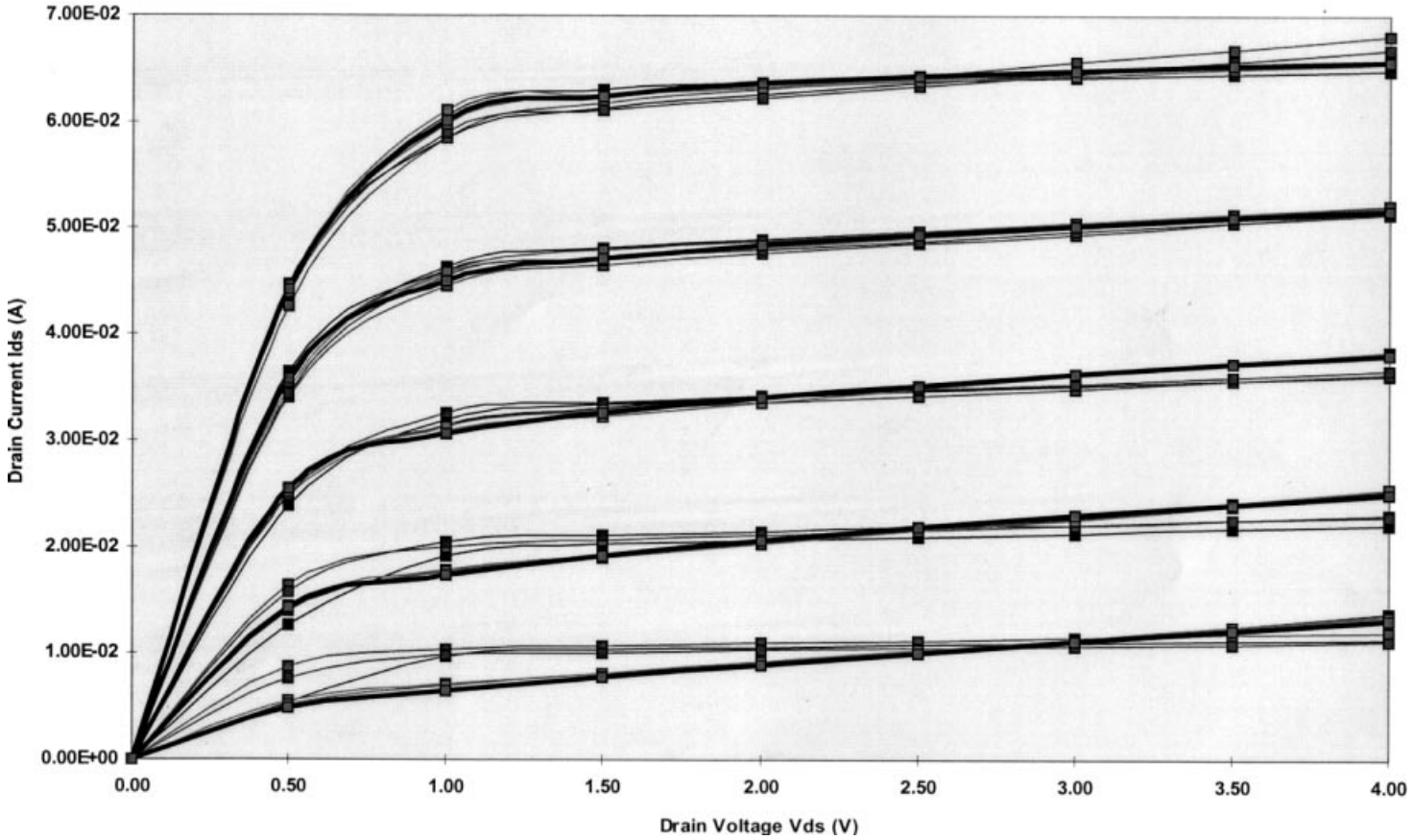


Figure 2-138 Comparison of dc models: Chalmers (Angelov), Curtice cubic, Curtice quadratic, Materka, Raytheon, and TOM (TriQuint).

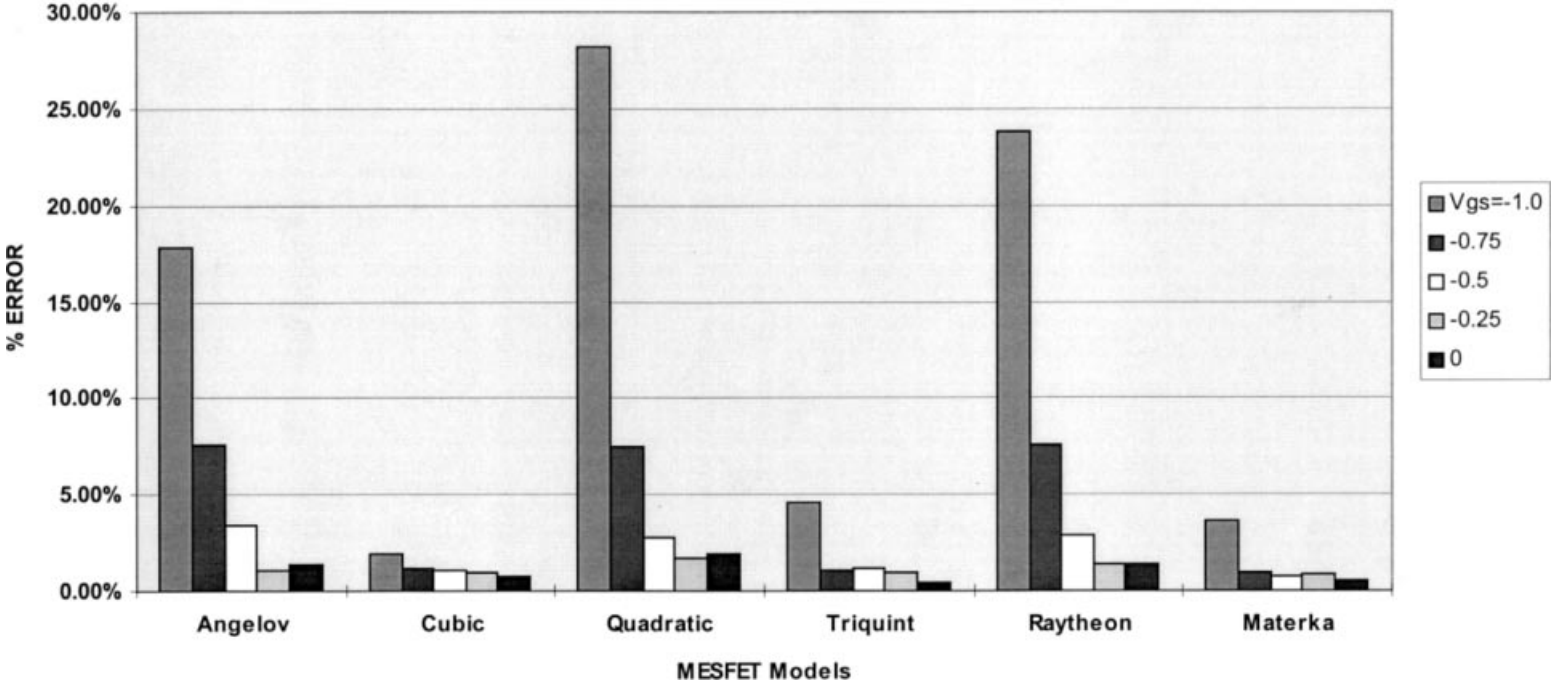


Figure 2-139 Dc *I–V* error bars: Chalmers (Angelov), Curtice cubic, Curtice quadratic, Materka, Raytheon, and TOM.

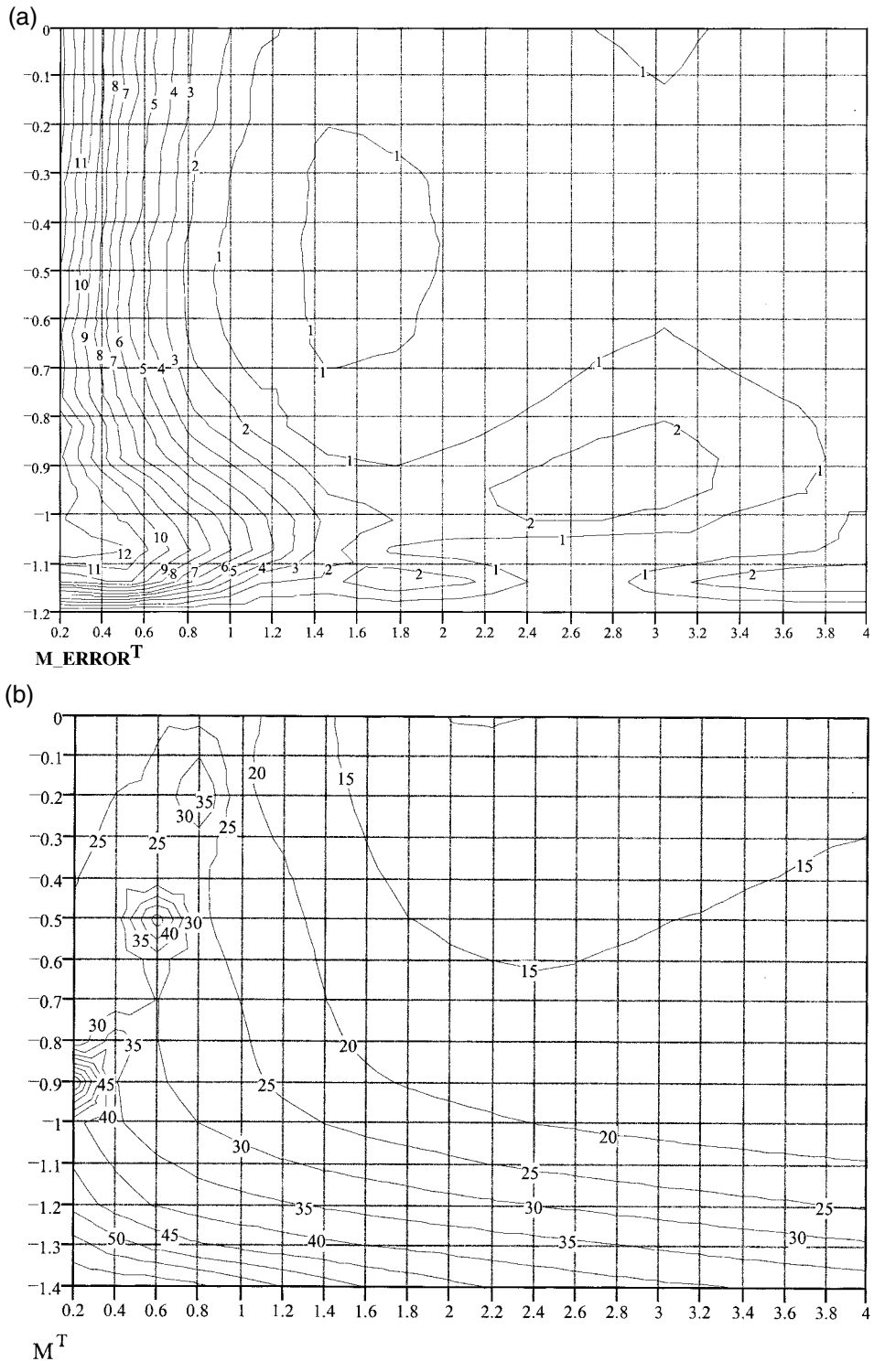


Figure 2-140 NE71000, Materka model, % error contours for (a) dc *I–V* and (b) *S* parameters.

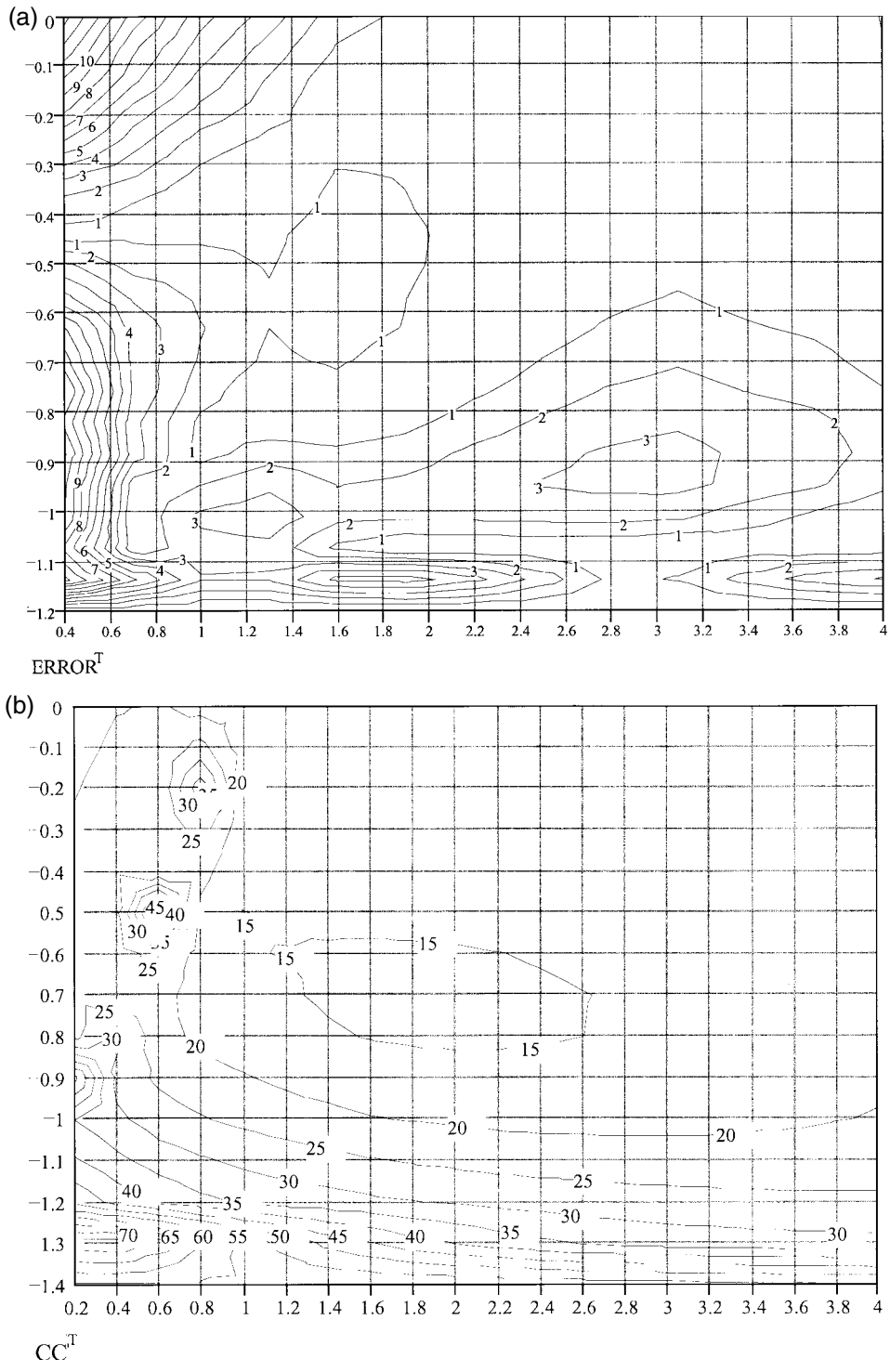


Figure 2-141 NE71000, Curtice cubic model, % error contours for (a) dc *I–V* and (b) *S* parameters.

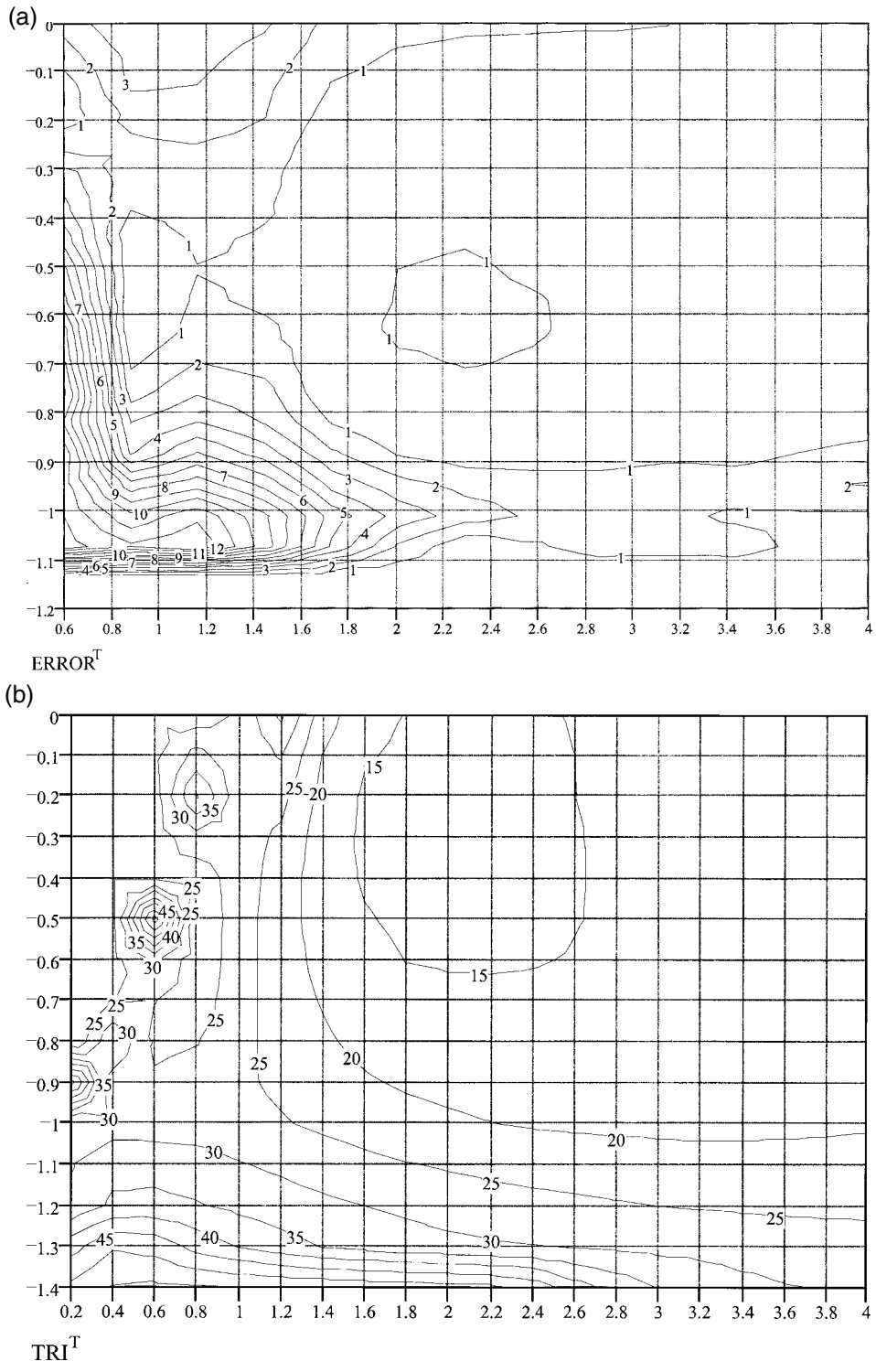


Figure 2-142 NE71000, TOM model, % error contours for (a) dc *I–V* and (b) *S* parameters.

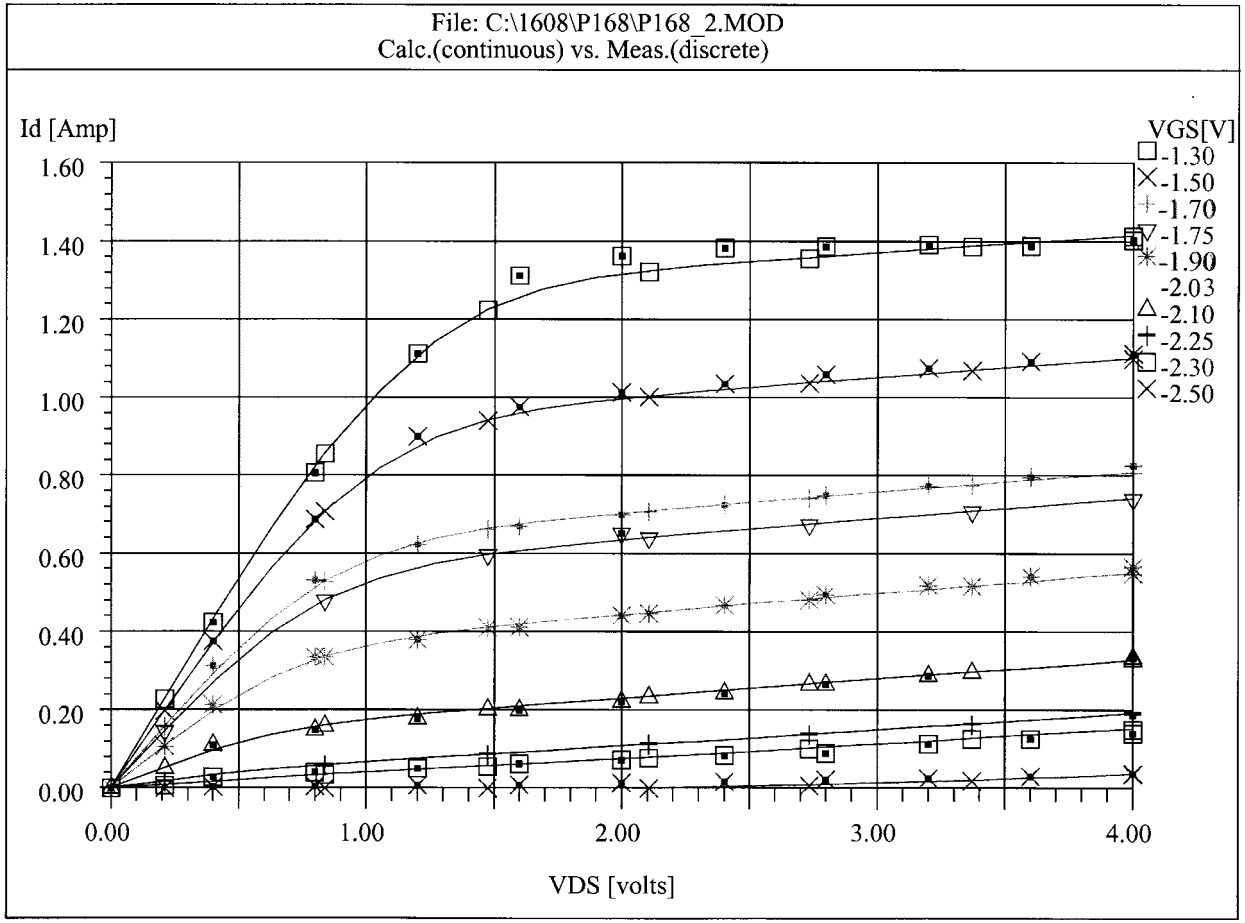


Figure 2-143 Calculated versus measured dc *I–V* curves for the Oki KGF1608 package MESFET.

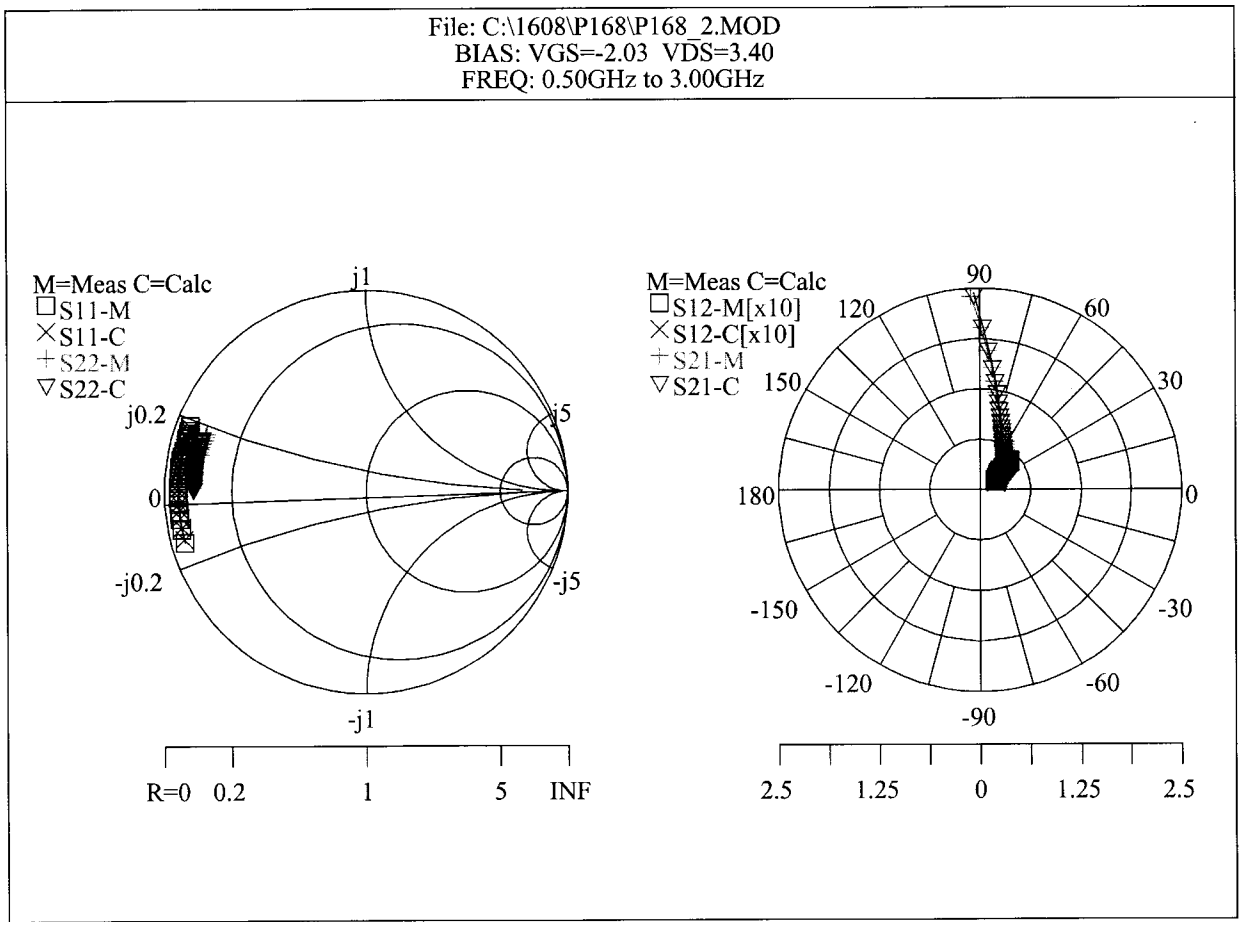


Figure 2-144 Calculated versus measured *S* parameters for the Oki KGF1608 package MESFET.

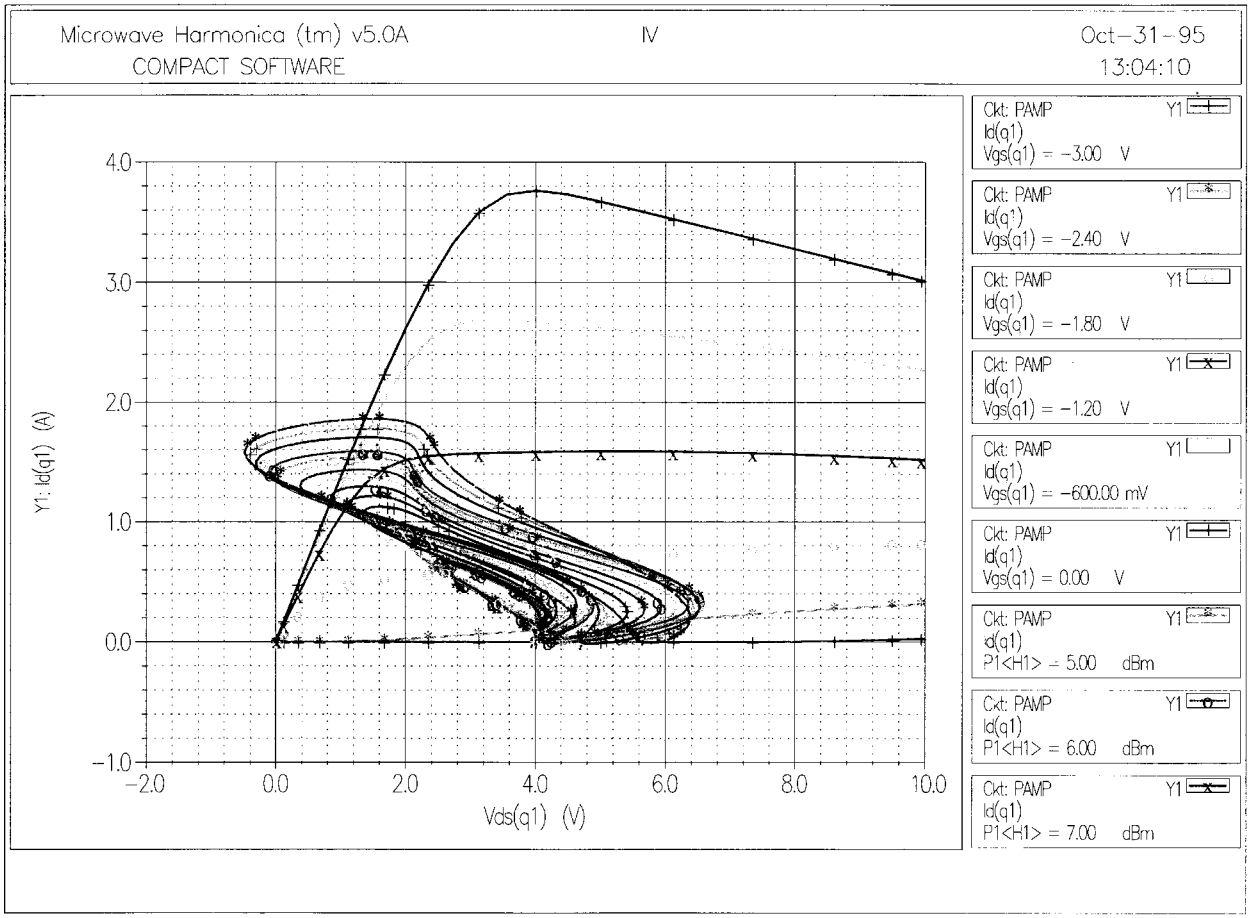


Figure 2-145 Load line and dc *I–V* curves for the KGF1608 MESFET.

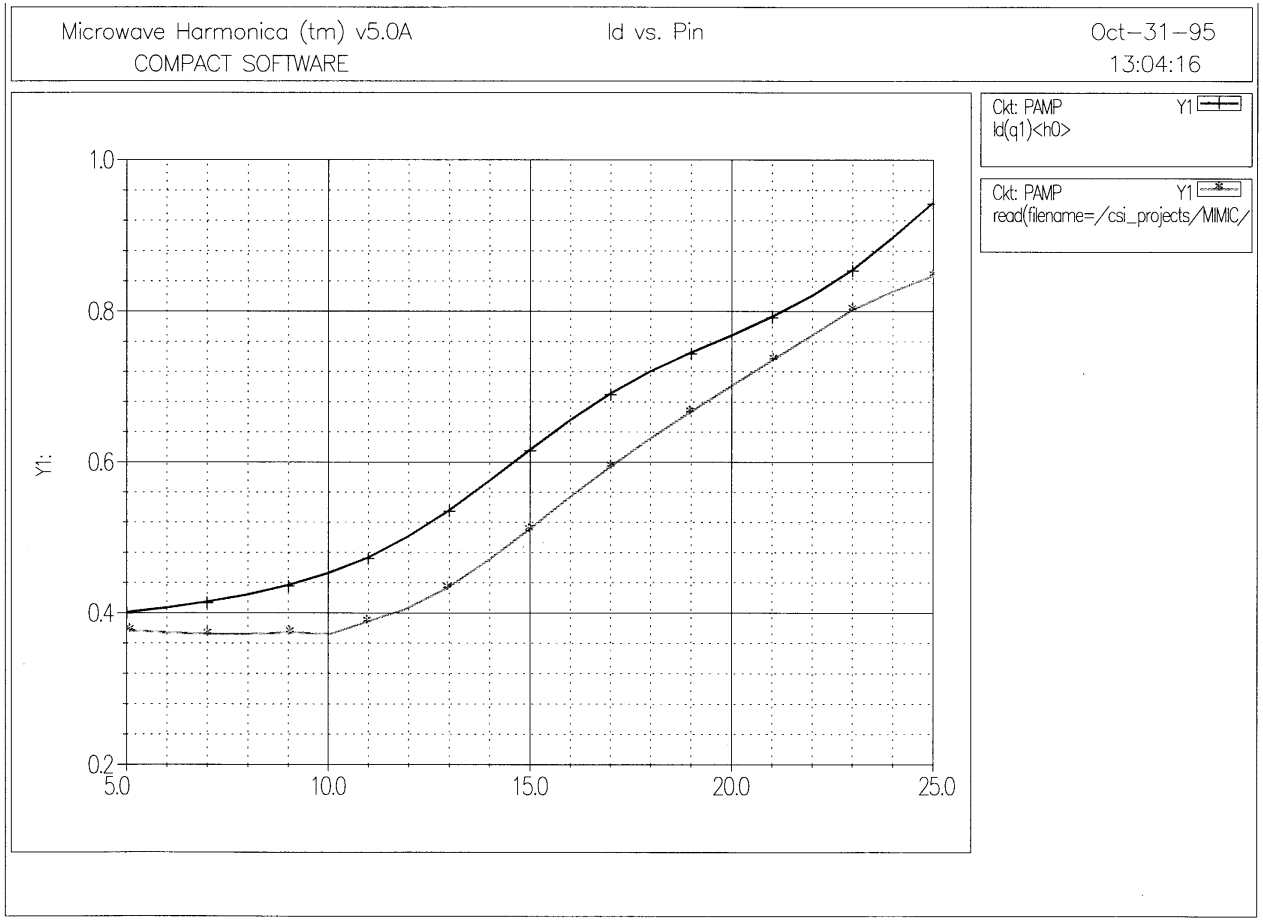


Figure 2-146 Measured versus calculated drain current for the KGF1608 MESFET.

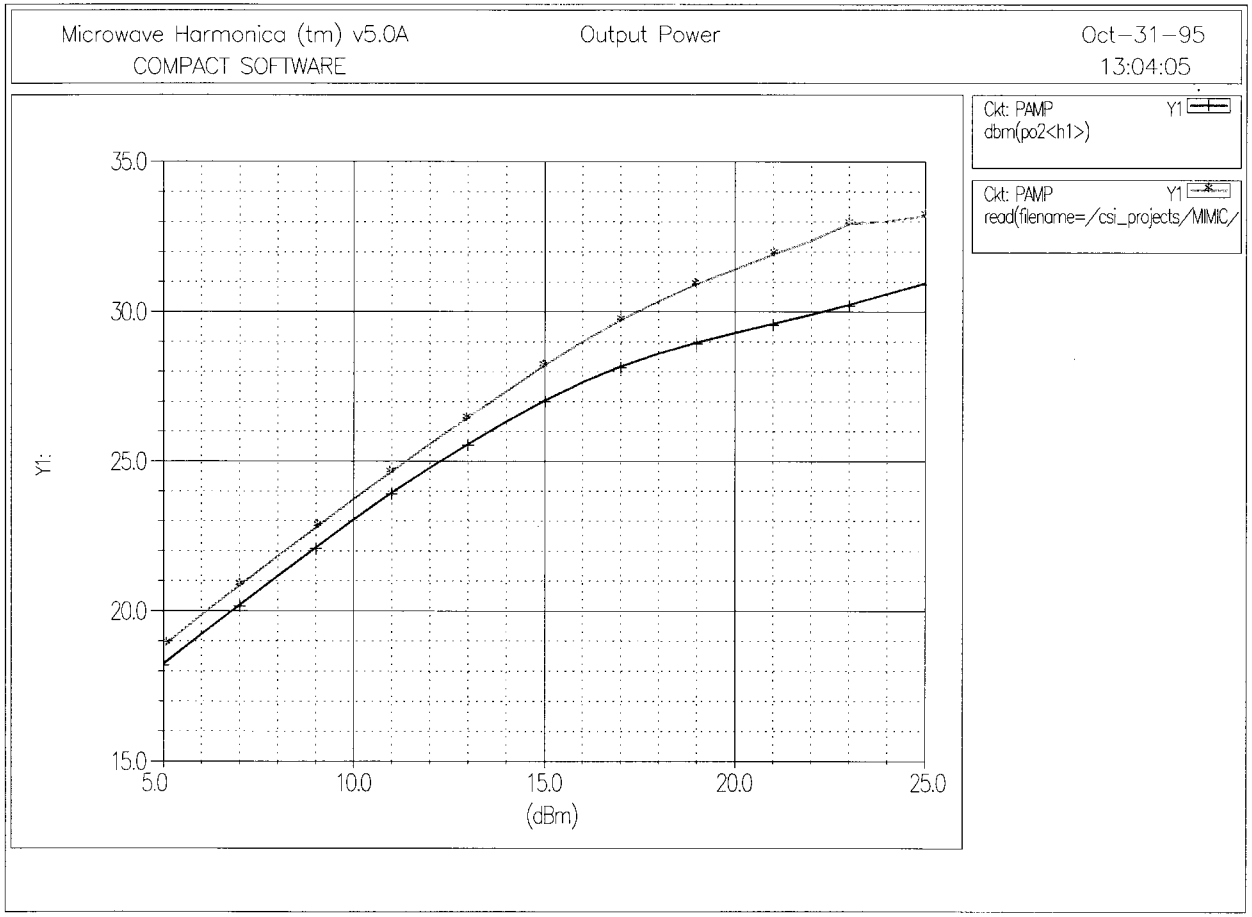


Figure 2-147 Measured versus calculated output power for the KGF1608 MESFET.

ac Errors. From the ac error contour plots we can see that the Chalmers (Angelov) model emerges as the most accurate for ac small-signal operation. The comparison also shows that all models considered are good for ac operations biased in the saturation region with a typical relative error of less than 15%. However, the ac relative errors are much higher in the linear region, particularly for the Curtice quadratic model, which exhibits more than 80% error at some points.

Notes. The above conclusions are based on the matching between the simulated responses of different models and the measured data from a typical NE71000 device; the models may behave differently for a different device. It is not possible to draw any broad-based conclusions as to what model is the best. In order to select the most suitable model for your application, you must consider all the factors, such as bias range, frequency range, and small-signal or large-signal operation, according to your experiences in device modeling.

2-4-6 Device Libraries

All these measurements lead to a nonlinear device library. Figures 2-148 and 2-149 shows a typical datasheet for a device in the library with measured and modeled data. The previous high-power example does not always lead to acceptable results as the models are far from being fully developed. We also show the simulation results using the physics-based model for FETs.

Physics-Based MESFET Modeling. We mentioned a physics-based model earlier. Figure 2-150 shows the measured versus predicted dc *I–V*, I/O power, and *S* parameters for an Alpha Industries MBE MESFET developed under a government contract. The simulation does not handle the trapped energy, which explains the negative slope on the third trace of the output dc *I–V* curve.

The process for bipolar transistors is quite similar and less tricky than the ones just described. Good examples can be found in the user manual for Compact Scout.

2-4-7 A Novel Approach for Simulation at Low Voltage and Near Pinchoff Voltage

As mentioned earlier, the most critical zone is the area of operating the FET both close to the pinchoff voltage and at low operating voltages. A crucial reason for the inadequate results is the modeling of the drain–source resistance as a function of bias point. This is particularly important in calculating the power-added efficiency. When doing your modeling, there is always a temptation to look at both high currents and operating voltages. Figure 2-151 shows a reasonable agreement between measured and predicted dc *I–V* performance. However, the very moment the transistor is modeled close to the linear region (Figure 2-152), things change drastically and the results from 0 to 6 mA have little resemblance to those from 0 to 50 mA. The modified Materka model (Figures 2-153 and 2-154) gives a significantly better match than the Angelov model.

In order to overcome the modeling inaccuracy for the FETs, Thomas Meier of Siemens AG in Munich has come up with the concept of combining two transistors in parallel, whereby the second one is not really connected for dc (Figure 2-155). Figure 2-156 compares measured and simulated R_{DS} using this new approach.

NE67300

Manufacturer : **NEC**
Model Keyword : **NE - 67300**
Device Type : **Mesfet**
Model Type : **Materka**
Package Type : **Chip**

Electrical Characteristics:
Idss : 40mA
Vp : 1.1V
Gain : 8.5dB @ 18GHz
P1db : N/A
Fmin : N/A

Model Data Range:
Vgs : -1.2V to 0.0V
Vds : 0.0V to 4.0V
Freq: 0.4GHz to 18GHz

Verification :
Type : Power sweep, 50Ω
Vgs : -0.6V Vds : 3.0V Ids : 30mA Freq: 6.0GHz

Comments : The S-parameters that appear in the figure were measured and simulated with wire bonds present. The length of each wire bond is approximately 30 mils. To obtain the S-parameters for the manufacture's data sheets, the wire bond inductances would be removed from the equivalent circuit.

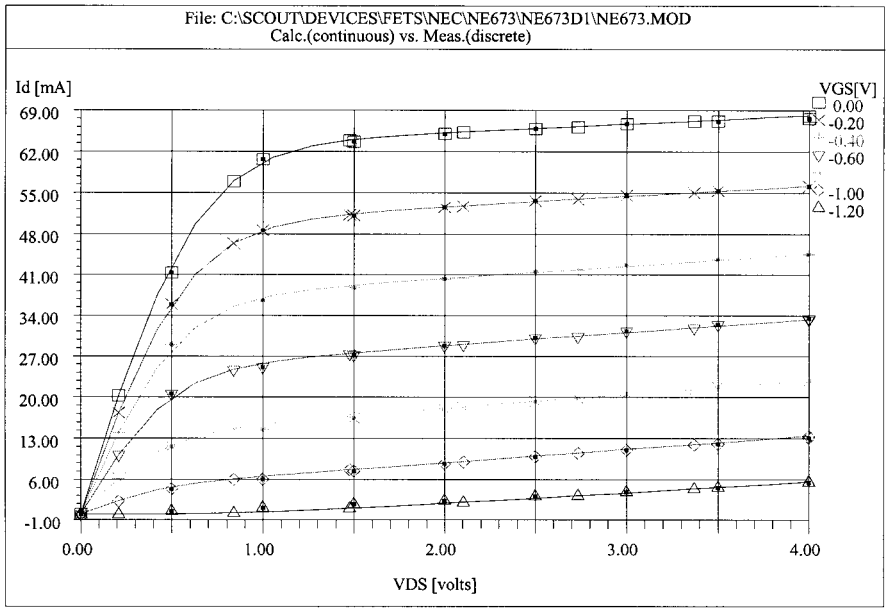
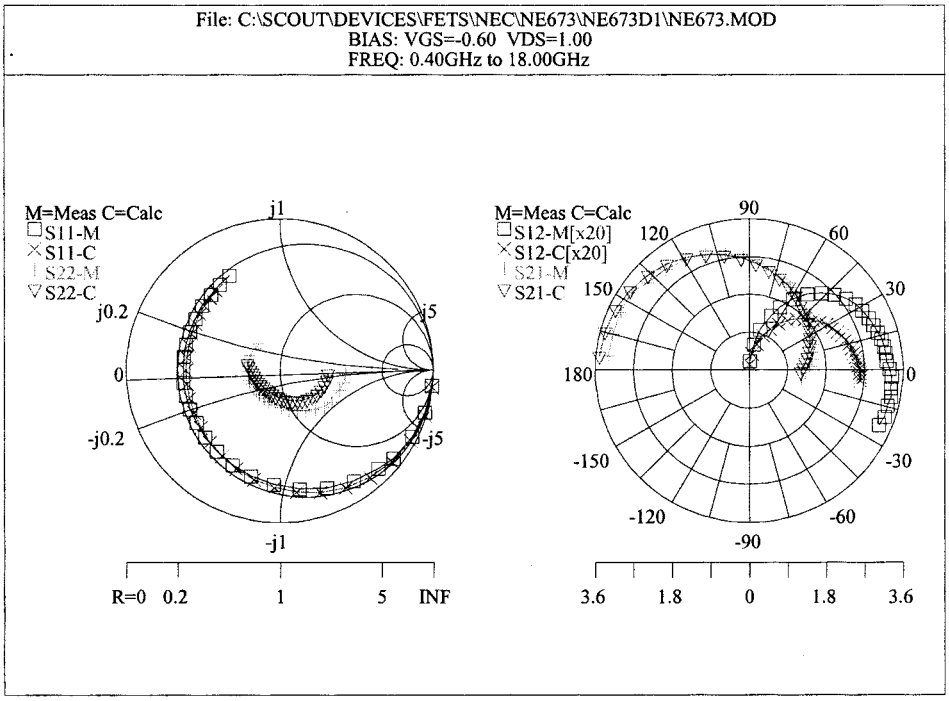


Figure 2-148 Page 1 of a nonlinear device library datasheet for the NE67300.



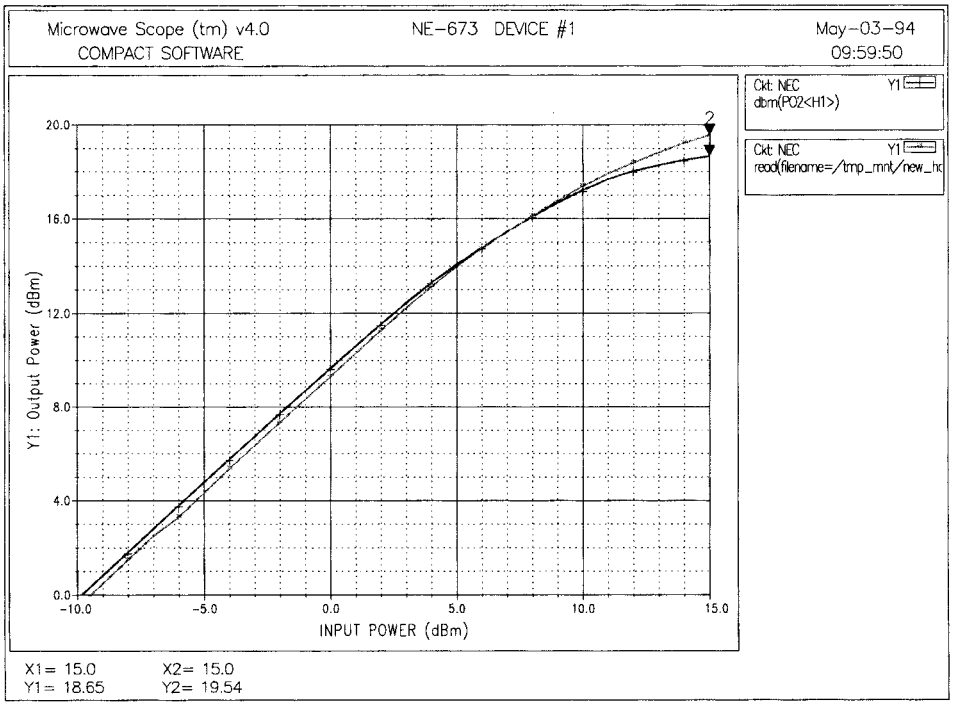


Figure 2-149 Page 2 of a nonlinear device library datasheet for the NE67300.

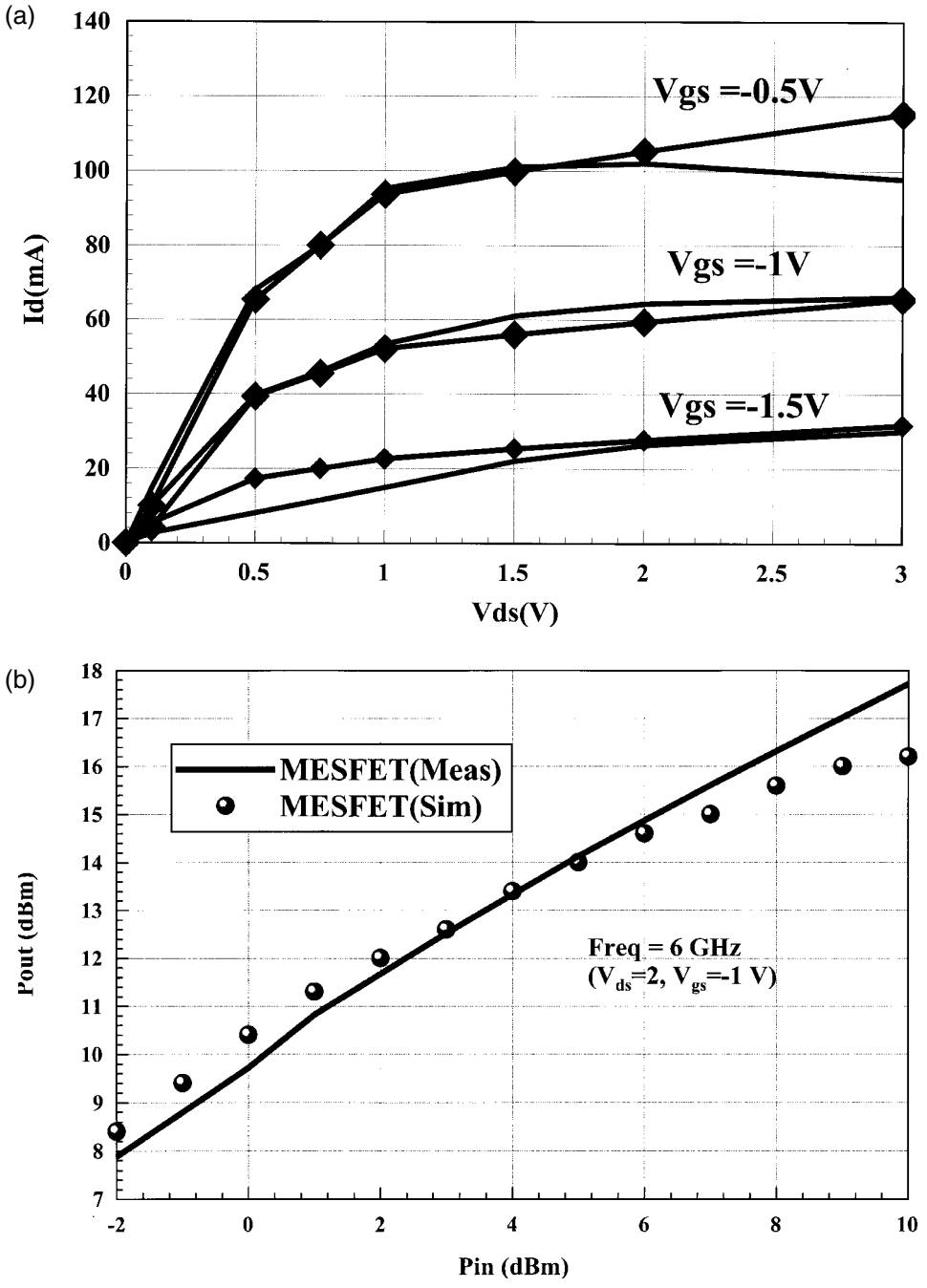


Figure 2-150 Comparison of measured data with PHYSFET and TOM models for an Alpha Industries MBE MESFET (0.25 μm × 400 μm) at $V_{GS} = -1$, $V_{DS} = 2$.

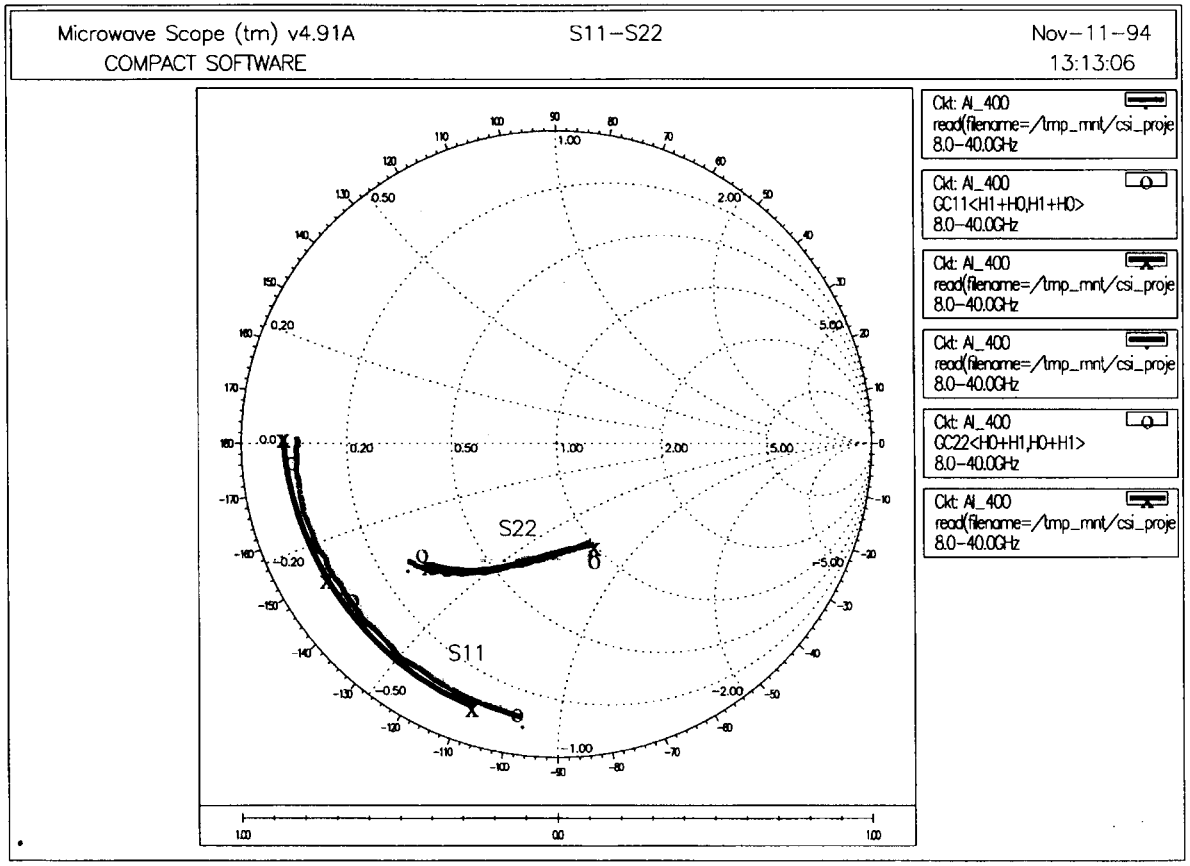


Figure 2-150 (*Continued*)

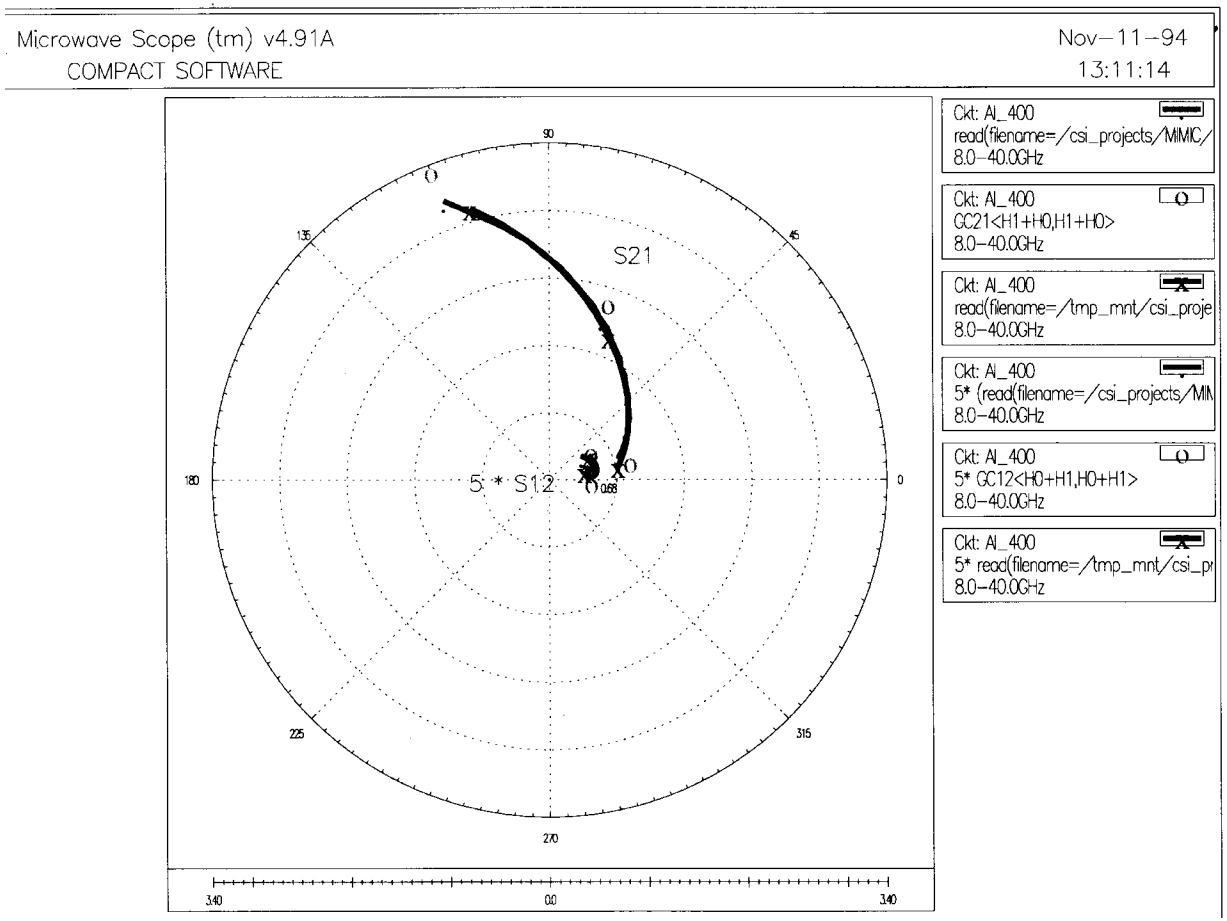


Figure 2-150 (*Continued*)

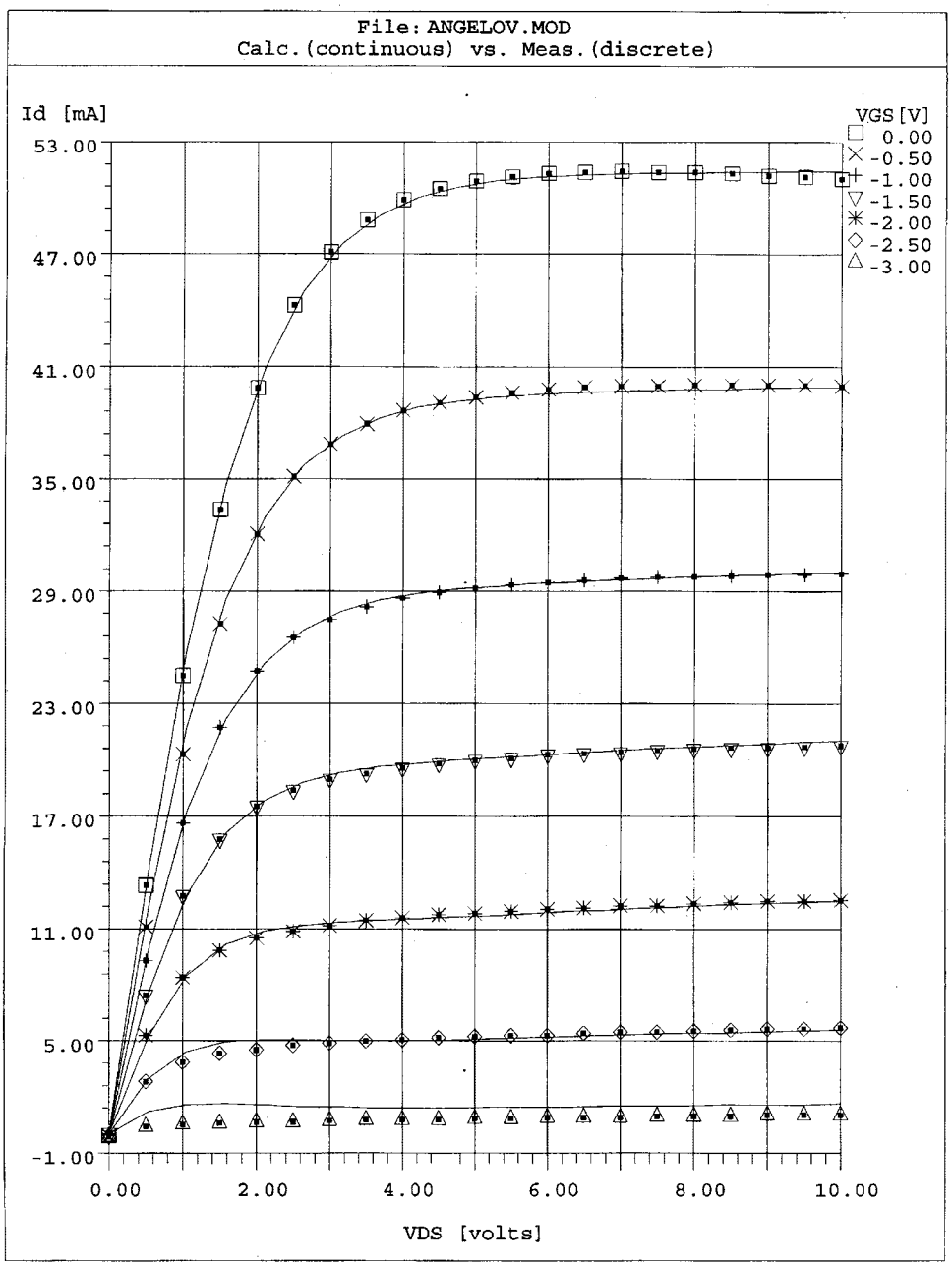


Figure 2-151 High-I_D dc I–V curves for the Angelov FET model.

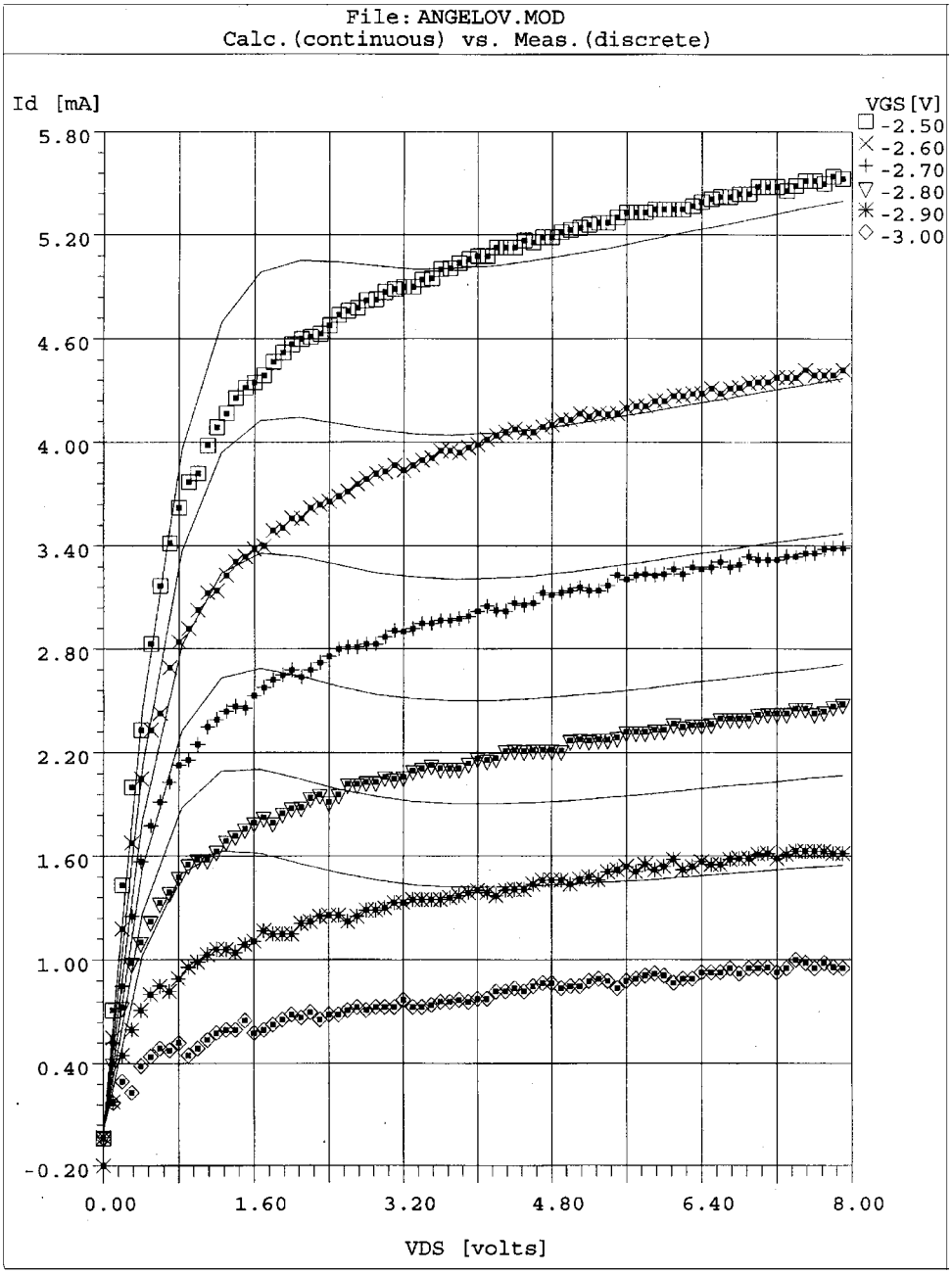


Figure 2-152 Low-I_D dc I–V curves for the Angelov FET model.

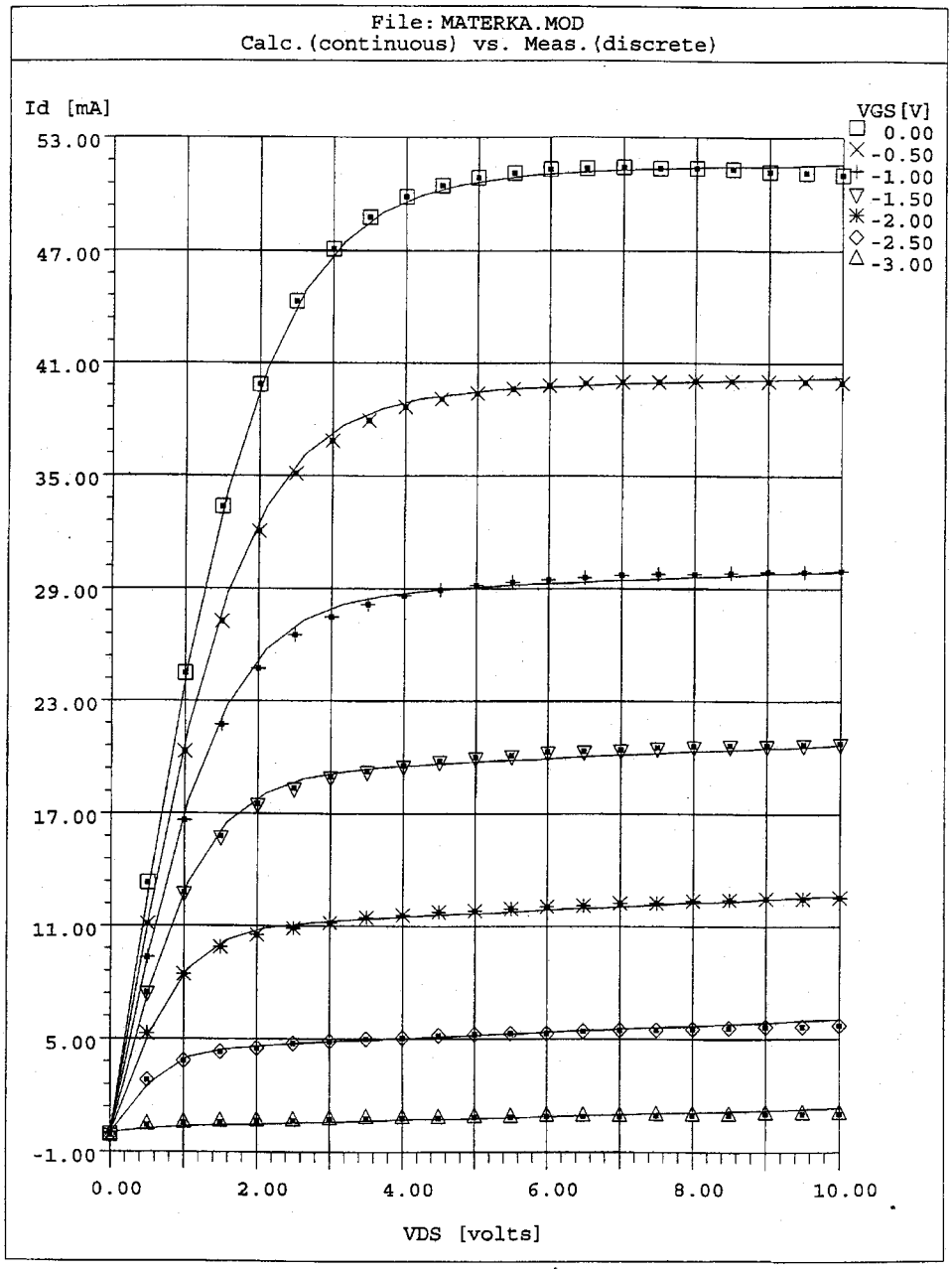


Figure 2-153 High-I_D dc I–V curves for the modified Materka FET model.

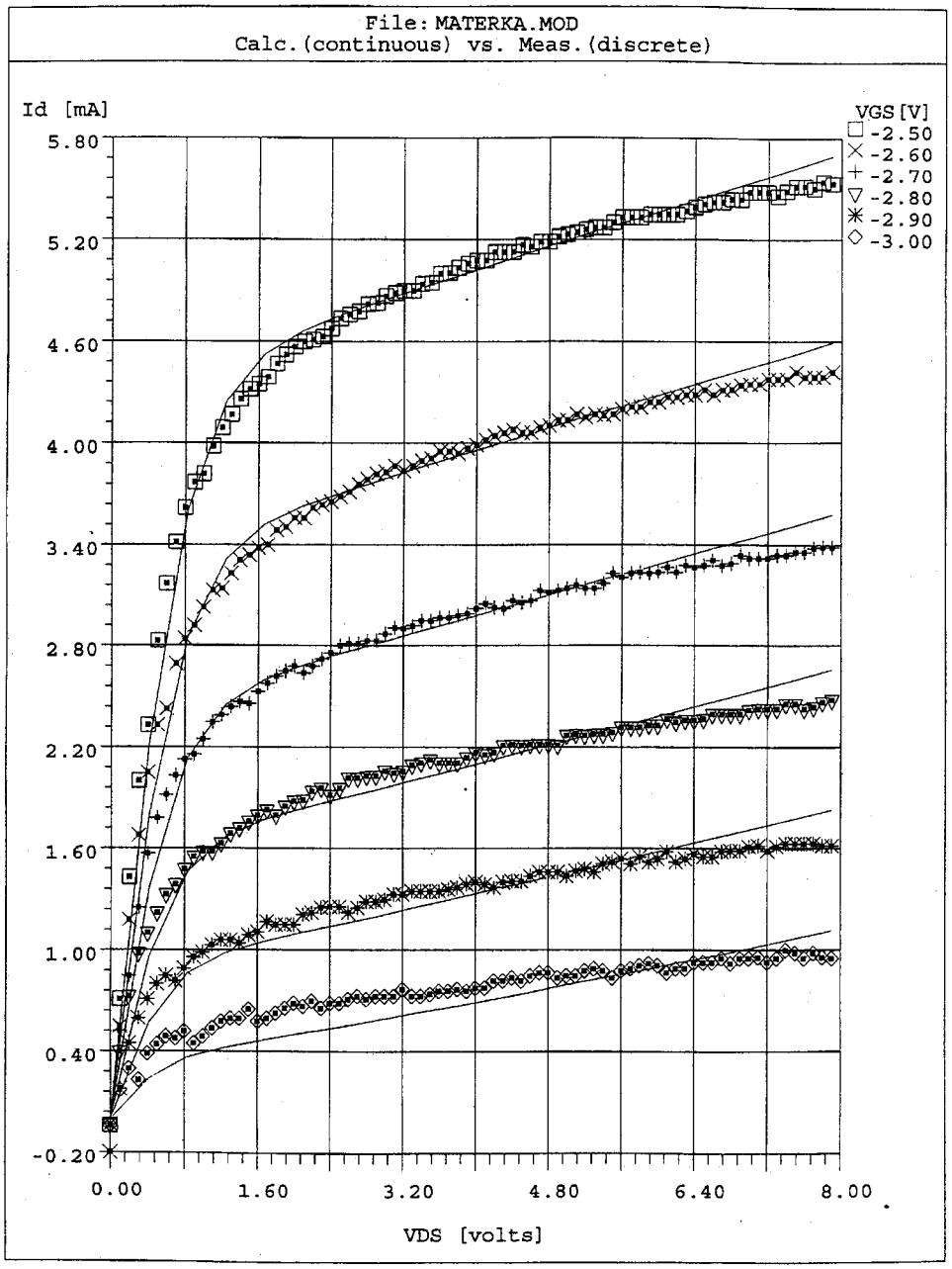


Figure 2-154 Low-I_D dc I–V curves for the modified Materka FET model.

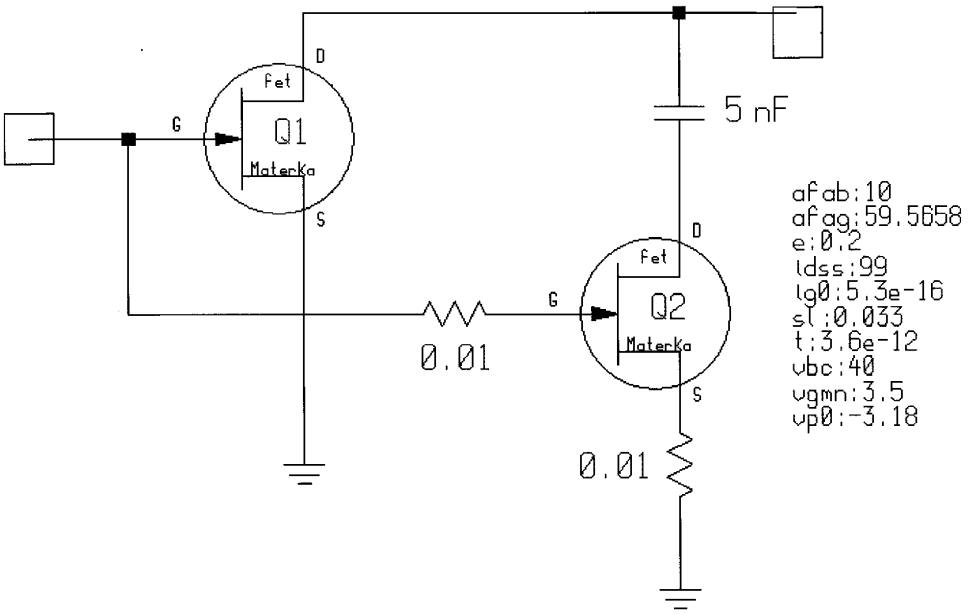


Figure 2-155 By putting two transistors in parallel, with the second one having no dc connection at the drain, one can simulate a dynamic R_{DS} quite well. Regardless of parameter values used for Q1, the parameters listed are necessary to get an accurate R_{DS} result.

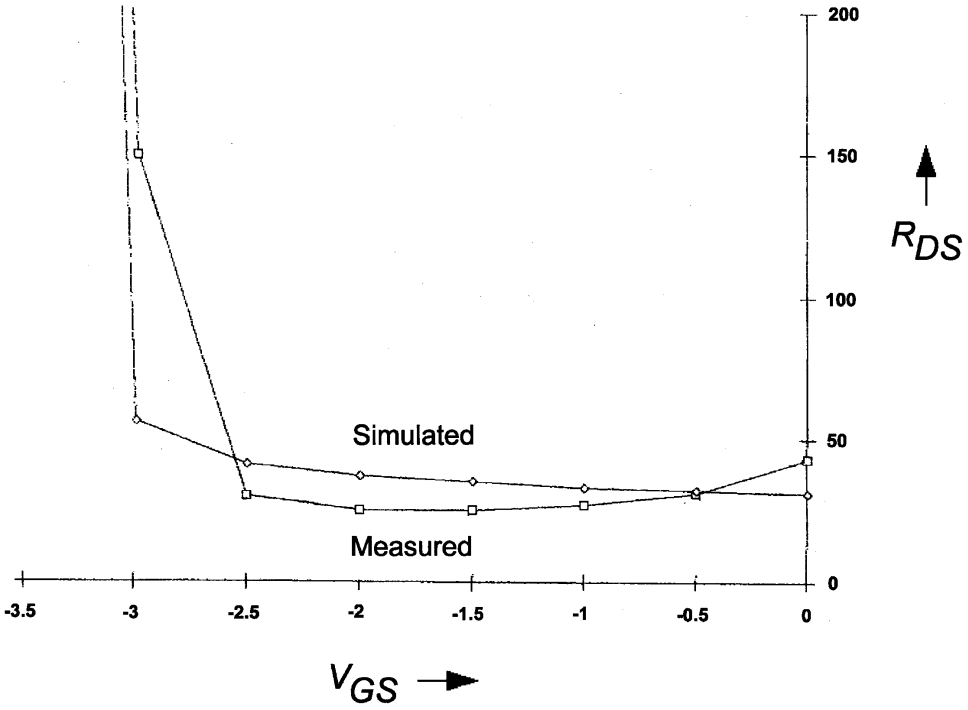


Figure 2-156 Measured versus simulated R_{DS} for the paralleled MESFET approach.

One might ask the question: "Why not fix the existing models?" It turns out that in order to get the best possible convergence and speed, one has to use analytic derivatives for voltage and current; at least the first and second derivatives, if not the third, are required. The programming of such a model becomes so difficult, however, and the calculation of its derivative is so highly likely to have subtle errors, that it turns out that the parallel configuration shown in Figure 2-155 is the best possible compromise for fast computation and accurate results.

A similar approach for bipolar transistors has been discussed by Philips with their MEXTRAM model. However, we have not seen any parameter extraction for this model and also we have not seen any industry acceptance for it. While, in principle, it will cover all derivatives of all members of the BJT family, its topology consists out of three discrete transistors connected and we doubt very much that a reliable parameter extraction program/mathematically continuous model for nonlinear simulation is supplied by the university that developed this jointly with Philips. At the time of writing this book, Philips has not yet provided us with the code for both, and until this happens, we are not in a position to qualify this any further.

2-4-8 Example: Improving the BFR193W Model

We have already mentioned the difficulty encountering the proper base-spreading resistor values and others by using standard parameter extraction programs. Table 2-28 is a copy from the Siemens semiconductor book showing Gummel–Poon Berkeley SPICE syntax parameters for the BFR193W BJT.

The first thing that will happen is that the observant reader will notice that the minimum possible base-spreading resistor, RBM (valid at large current densities), is larger than the actual base-spreading resistor, RB, "claimed to be measured" at 1 Ω (see Table 2-28).

Another headache is caused by the measured noise performance of the same device at an operating point of 10 mA. Using the above-mentioned simplified equation

Table 2.28. BFR193W transistor chip data: SPICE parameters (Gummel–Poon model, Berkeley–SPICE 2G.6 syntax)

IS	=	0.2738	fA	BF	=	125	—	NF	=	0.95341	—
VAF	=	24	V	IKF	=	0.26949	A	ISE	=	10.627	fA
NE	=	1.935	—	BR	=	14.267	—	NR	=	1.4289	—
VAR	=	3.8742	V	IKR	=	0.037925	A	ISC	=	0.037409	fA
NC	=	0.94371	—	RB	=	1	Ω	IRB	=	0.91763	mA
RBM	=	1.8368	Ω	RE	=	0.76534	Ω	RC	=	0.11938	Ω
CJE	=	1.1824	fF	VJE	=	0.70276	V	MJE	=	0.48654	—
TF	=	18.828	ps	XTF	=	0.69477	—	VTF	=	0.8	V
ITF	=	0.96893	mA	PTF	=	0	deg	CJC	=	935.03	fF
VJC	=	1.1828	V	MJC	=	0.30002	—	XCJC	=	0.053563	—
TR	=	1.0037	ns	CJS	=	1	fF	VJS	=	0.75	V
MJS	=	0	—	XTB	=	0	—	EG	=	1.1	eV
XTI	=	3	—	FC	=	0.72063	—	TNOM	=	300	K

All parameters are ready to use, no scaling is necessary. Extracted on behalf of SIEMENS Small Signal Semiconductors by Institut für Mobil-und Satellitenfunktechnik (IMST).

$$F = 1 + \frac{R'_{bb}}{R_g} + \frac{0.5(26\ \text{mV}/I_C)}{R_g}$$

(approximate equation at medium frequencies) (2-288)

we obtain the following results:

1. Inserting 1 Ω (value for RB provided by IMST) and R_B = 15 Ω (measured by the authors), here are the interesting results:

$$F = 1 + \frac{1}{50} + \frac{2.6}{100} = 1.046 \tag{2-289}$$

as noise factor or, as noise figure, F (dB) = 0.2 dB (IMST).

2. Or using the measured 15 Ω, one obtains

$$f = 1 + \frac{15}{50} + \frac{2.6}{100} = 1.326 \tag{2-290}$$

as noise factor or, as noise figure, F (dB) = 1.22.

The Siemens datasheet claims a noise figure of 1.3 dB, and therefore the measured versus calculated data differ by 0.08 dB. Since our noise figure was calculated well below f_T, which is seen from the fact that there are no frequency-dependent terms in the noise figure calculation, it is not unusual that the 900-MHz noise figure is 1.3 dB or slightly higher. This simple example shows how easy it is to provide measurement data that are inconsistent with the laws of physics. It is therefore highly recommended to add the noise figure as an additional parameter in the parameter extraction to reduce the number of variables—specifically, those for which the standard extraction has little sensitivity.

The collector–emitter junction capacitance is another important parameter that frequently gets measured incorrectly. The way around this is to remember the definition

$$f_{\max} = \sqrt{\frac{f_T}{8\pi R'_{bb} C_c}} \tag{2-291}$$

f_T can be measured from the 3-dB point of the emitter current gain (β) and R'_{bb} was just computed above, so the equation can be solved for C_c, as there is a relationship between measured $g_{\max}$ and $f_{\max}$. There are several correlations between the parameters of the equivalent circuit as published in different places, and therefore one can restrict the degree of freedom within the optimization process and obtain much better results. This approach is applicable to all types of transistors, bipolar and FET, and can be used to improve the accuracy of parameters that not only affect a transistor's RF performance but also its noise performance.

REFERENCES

1. H. A. Watson, *Microwave Semiconductor Devices and Their Circuit Applications* . New York: McGraw-Hill, 1969, Chap. 2.
2. A. S. Grove, *op. cit.,* Chap. 4.
3. A. S. Grove, *op. cit.,* Chap. 7.
4. P. E. Gray, D. DeWitt, A. R. Boothroyd, and J. G. Gibbons, *Physical Electronics and Circuit Models of Transistors* . New York: Wiley, 1964, p. 10.
5. P. E. Gray et al., *op. cit.,* p. 129.
6. P. E. Gray et al., *op. cit.,* p. 180.
7. Charles A. Liechti, "Microwave Field-Effect Transistors—1976," *IEEE Trans. Microwave Theory Tech.* , **MTT-24**, 279–300, June 1976.
8. Edward A. Wolff and Roger Kaul, *Microwave Engineering and Systems Applications* . New York: Wiley, 1988, p. 395.
9. Arthur D. Evans, ed., *Designing with Field-Effect Transistors* . New York: McGraw-Hill, 1981.
10. R. D. Middlebrook, "A Simple Derivation of Field-Effect Transistor Characteristics," *Proc. IEEE*, **50**, 1146–1147, 1963.
11. Tzu-Hwa and Craig P. Snapp, "Low Noise Microwave Bipolar Transistor with Sub-Half Micrometer Emitter Width," *IEEE Trans. Electron Devices* , **ED-25**, 723–730, June 1978.
12. Ulrich L. Rohde, "Designing a Match Low Noise Amplifier Using CAD Tools," *Microwave J.* , **20**(10) 154–160, October 1986.
13. R. A. Pucel, W. Strubel, R. Hallgren, and U. L. Rohde, "A General Noise De-embedding Procedure for Packaged Two-Port Linear Active Devices," *IEEE Trans. Microwave Theory Tech.* , **40**(11), 2013–2024, November 1992.
14. R. A. Pucel and U. L. Rohde, "An Exact Expression for Noise Resistance R_n of a Bipolar Transistor for the Hawkins Noise Model," *IEEE Microwave and Guided Wave Letters* , **3**(2), 35–37, February 1993.
15. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits.* New York: Wiley, 1986, p. 386.
16. D. Frohman-Bentchkowsky and A. S. Grove, "Conductance of MOS Transistors in Saturation," *IEEE Trans. Electron Devices* , **ED-16**, 108–113, January 1969.
17. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits* . New York: Wiley, 1986, p. 480.
18. "New LDMOST Technology Superb for RF Applications," *Philips Semiconductors World News,* Vol. 7, Issue 3, Article No. 29, http://www-us2.semiconductors.philips.com/publications/content/file_344.html.
19. Y. P. Tsividis, *Operation and Modeling of the MOS Transistor* . New York: McGraw-Hill, 1987, p. 136.
20. A. Materka and T. Kacprzak, "Computer Calculation of Large-Signal GaAs FET Amplifier Characteristics," *IEEE Trans. Microwave Theory Tech.* , **MTT-33**(2), 129–135, February 1985.
21. R. E. Anholt and S. E. Swirhun, "Experimental Investigation of the Temperature Dependence of GaAs FET Equivalent Circuits," *IEEE Trans. Electron Devices* , **ED-39**(9), 2029–2036, September 1992.
22. Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits* . New York: Wiley, 1986.
23. Paul W. Tuinenga, *SPICE: A Guide to Circuit Simulation and Analysis Using PSpice ®*. Englewood Cliffs, NJ: Prentice Hall, 1992.
24. Andrei Vladimirescu, *The SPICE Book* . New York: Wiley, 1994.

25. H. Statz, H. Haus, and R. A. Pucel, "Noise Characteristics of Gallium Arsenide Field-Effect Transistors," *IEEE Trans. Electron Devices,* **ED-21**, 549–562, September 1974.

26. V. Rizzoli, F. Mastri, and C. Cecchetti, "Computer-Aided Noise Analysis of MESFET and HEMT Mixers," *IEEE Trans. Microwave Theory Tech.* , **MTT-37**, 1401–1410, September 1989.

27. V. Rizzoli, F. Mastri, and D. Masotti, "General-Purpose Noise Analysis of Forced Nonlinear Microwave Circuits," *Military Microwave*, 1992.

28. Ulrich L. Rohde and Chao-Ren Chang, "Parameter Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits," *Microwave J.,* **36**(5), 222–239, May 1993.

29. Ulrich L. Rohde, "Parameters Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits Like Mixers and Oscillators," *23rd European Microwave Conference,* Madrid, Spain, September 6–9, 1993.

30. Ulrich L. Rohde, Chao-Ren Chang, and J. Gerber, "Design and Optimization of Low-Noise Oscillators Using Non-linear CAD Tools," presented at the *Frequency Control Symposium,* Boston, June 1–2, 1994.

31. Ulrich L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design* . New York: Wiley, 1997, Appendix B-3, "Bias Independent Noise Model."

FURTHER READING

George Anderson, "Moto Aims FETs at UHF," *Electron. Eng. Times*, 96, September 2, 1996.

Jean–Jacques Bouny, "Advantages of LDMOS in High Power Linear Amplification," *Microwave Eng. Eur.*, 37–40, April 1996.

Noël Boutin, "Develop SPICE High Frequency Small Signal Transistor Models from Data Sheets," *Appl. Microwave & Wireless*, 30–36, April 1999.

V. Danelon, F. Aniel, J. Benchimol, J. Mba, M. Riet, P. Crozat, G. Vernet, and R. Adde, "Parameters of InP-Based Double-Heterojunction Base Collector Self-Aligned Bipolar Transistors," *IEEE Microwave Guided Wave Lett.*

Suman Datta, Shen Shi, Kenneth P. Roenker, Marc M. Cahay, and William Stanchina, "Simulationand Design of InAlAs/InGaAs *pnp* Heterojunction Bipolar Transistors," *IEEE Trans. Electron Devices*, **ED-45**(8), 1634–1643, August 1998.

Russell J. Elias, Gordon C. Ma, and Laurence S. Golonka, "Statistical Optimization High Frequency LDMOS Devices Via Hyper-Fractionated Experimental Designs," *1996 IEEE/CPMT International Electronics Manufacturing Technology Symposium,* October 14–16, 1996, pp. 462–466.

Al Franceschino, "LDMOS Devices Enhance Cellular Amplifier Design," *Microwaves & RF*, **35**(10), 67–68, 70, 72, 74, October 1996.

GaAs IC Symposium Technical Digest 1995 , IEEE Gallium Arsenide Integrated Circuit Symposium, San Diego, CA, October 29 to November 1, 1995.

Hewlett-Packard, *Communications Components Designer's Catalog* , Hewlett-Packard, Jan Jose, CA, 1997.

ITT Capacitance Diodes, Tuner Diodes, Diode Switch, PIN Diodes, Basics and Applications , International Telephone and Telegraph Corporation, 1976.

Aron Kain, *Final Report for Bias Dependence Noise Modeling of Heterojunction Bipolar Transistors* , USAF SBIR Phase II (PIIN), F33615-95-C-1707, November 1997. Issued by USAF/AFMC/ASC, Wright Laboratory WL/AAKE BLD 7, 2530 C ST, Wright-Patterson AFB, OH 45433-7607.

X. McDade and X. Schiavone, "Switching Time Performance of Microwave PIN Diodes," *Microwave J.*, **17**, September 1974.

Danny Molezion, "A 1 watt Power Amplifier Design for 800–900 MHz Applications Using Discrete LDMOS FETs," *Wireless Technology '96 Conference & Exposition Conference Proceedings,* October 7–11, 1996, pp. 260–267.

Nadir Rahman, "A Survey of Existing Literature on Reverse Modeling of Field Effect Transistors," http://www-mtl.mit.edu/CAPAM/memos/95-3/95-3.html.

Jin-Ho Shin, Jiyoung Kim, Yuji Chugn, Joonwoo Lee, Youngsuk Suh, Kyu Hwan Ahn, and Bumman Kim, "Low-Frequency Noise Characterization of Self-Aligned AlGaAs/GaAs Heterojunction Bipolar Transistors with a Noise Corner Frequency Below 3 kHz," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, pp. 1604–1613, November 1998.

K. Y. Toh, P. K. Ko, and R. G. Meyer, "An Engineering Model for Short-Channel MOS Devices," *IEEE J. Solid-State Circuits*, **23**, 950–958, August 1988.

Robert J. Zavrel, Jr., "Silicon–Germanium and the World of RF Design," *RF Des.*, 46–50, June 1997.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

3

AMPLIFIER DESIGN WITH BJTs AND FETs

3-1 PROPERTIES OF AMPLIFIERS

3-1-1 Introduction

The goal of this chapter is to move from the details on the semiconductor devices themselves to our first practical application. One of the more interesting features of amplifiers is that they can easily become what will be covered in Chapter 5: *oscillators*. This is due to the high gain of the devices, unaccounted for parasitic elements, and other design flaws. The amplifiers we need will fall into three categories:

- Low-noise amplifiers
- High-gain amplifiers
- Medium- to high-power amplifiers

See Figure 3-1. The low-noise amplifier always operates in Class A, typically at 15–20% of its maximum useful current. The high-gain amplifier can operate in Class A, as well as Class B (mostly push–pull). The higher dc current for the same device results in a higher noise figure and more gain, and ultimately more output power. Class C operating mode is really reserved for either FM transmissions or constant-carrier modes like CW. Some of the modern digital modulation types are sensitive to phase distortion rather than amplitude changes and because of the resulting output spectrum designers have stayed away from Class C operation.

Some of these amplifiers, such as bipolar versions, can be dc-coupled with very few difficulties; others, like those in the FET families, will cause more headaches. An interesting example, although much too high in frequency for the purpose of this book, is shown in Figures 3-2, 3-3, 3-4, and 3-5. It shows a three-stage amplifier that has noise feedback for the first two stages and resistive feedback for the output stage. These are the three types of amplifiers, regardless of technology, we will be evaluating. The achievement of accurate noise analysis of this circuit, reported in Rohde et al. [1], marked the first time a complete

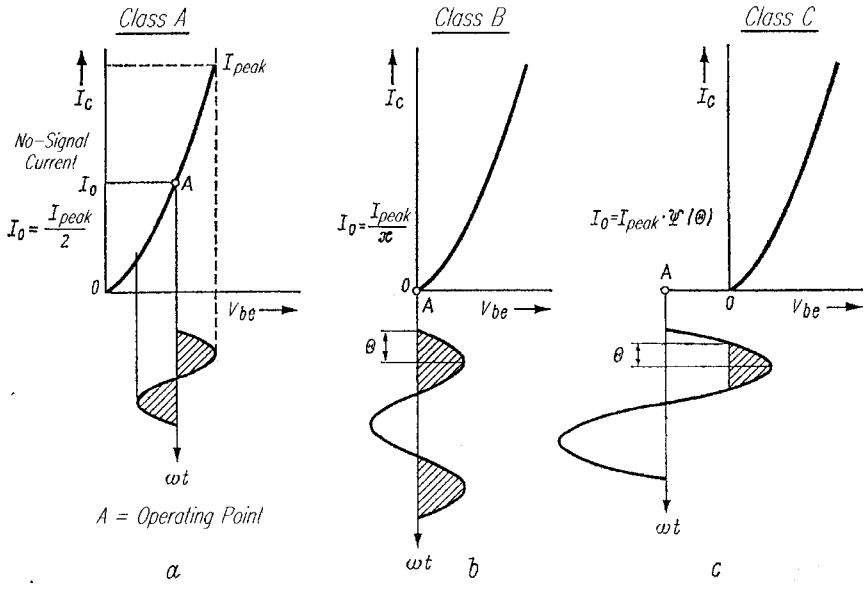


Figure 3-1 Definition of Class A, B, and C operation and resulting bias, including conduction angle (Θ, to be discussed in Section 3-2-2). The transfer characteristic can be either quadratic or exponential. This results in different distortion but does not change the basic operating mode.

linear noise model with essentially no frequency limitation had been developed. Its accuracy depends solely on the accuracy of the element values of the linear equivalent circuit, which can be obtained from measured *S*-parameters by an optimization process. Needless to say, this technique is also applicable at lower frequencies, such as 1.5–3 GHz, and its deviation from measured values has never been worse than 0.2 dB or, relative to the noise figure, 10% expressed in dB.

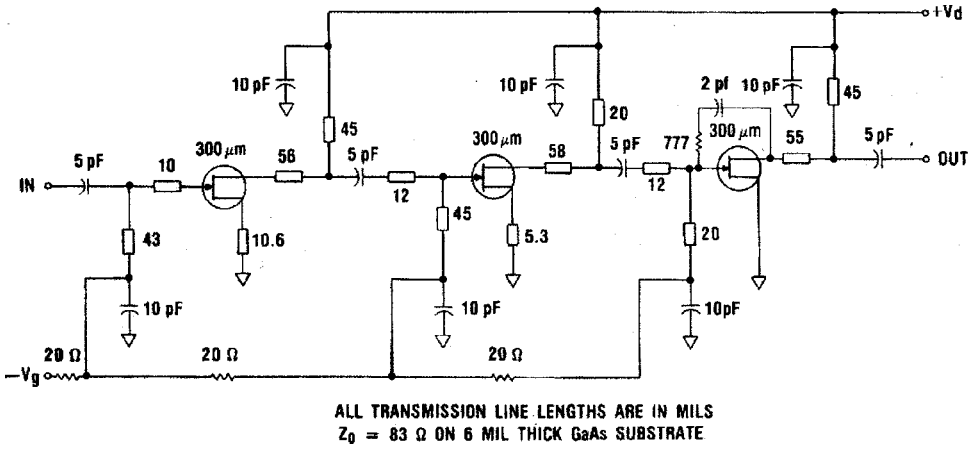


Figure 3-2 Schematic of the X-band GaAs monolithic low-noise amplifier (Texas Instruments EG8021). For reasons of linearity, all three stages operate in Class A.

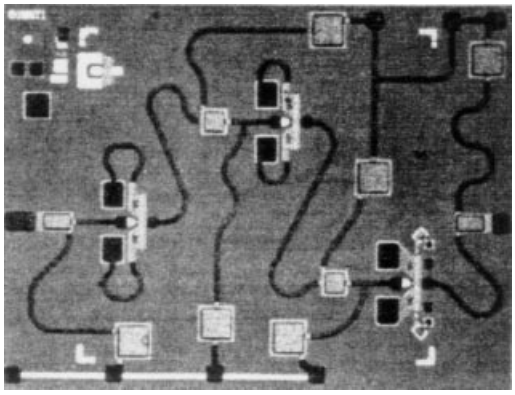


Figure 3-3 Photograph of the EG8021 monolithic amplifier chip. The area pictured is 0.09 inch by 0.12 inch in size.

The results of Figure 3-4 were obtained by using the linear FET model as supplied by Texas Instruments. It becomes somewhat obvious that a match for S_{11} of –60 dB is not likely. We then replaced the linear transistor with the SPICE-type nonlinear model, resulting in the responses shown in Figure 3-5. S_{11} now looks much more realistic in the center of the range, but at the same time exhibits a tendency toward instability at frequencies around 8 GHz. The actual circuit, when measured, did not exhibit this potential instability; on the other hand,

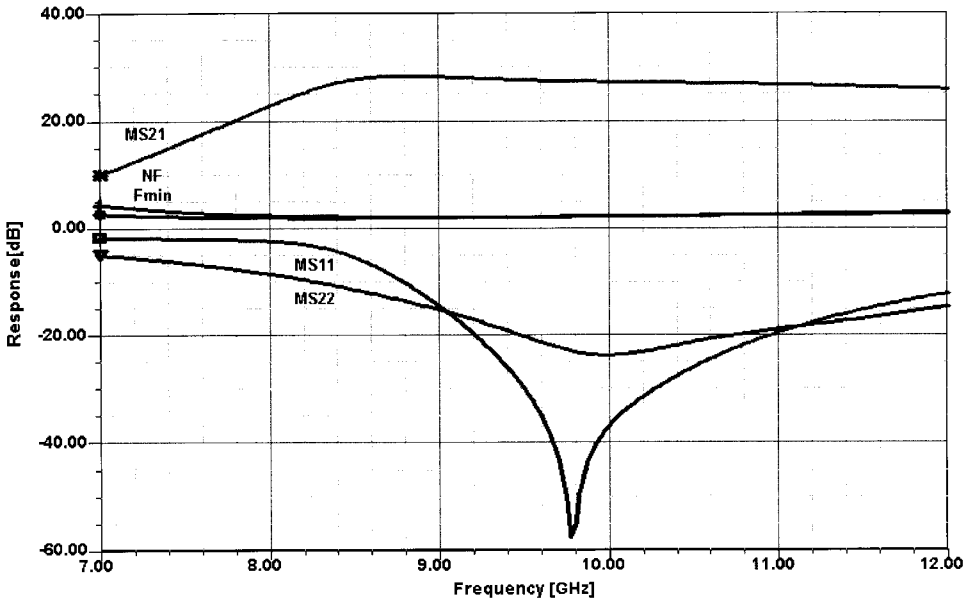


Figure 3-4 Simulated F_{min}, noise figure, S_{11}, S_{21}, and S_{22} responses for the three-stage GaAsFET amplifier using the TI linear FET model. The values at 10 GHz are: F_{min}, 2.20 dB; NF, 2.21 dB; S_{11}, –36.8 dB; S_{21}, 27.2 dB; and S_{22}, –23.7 dB.

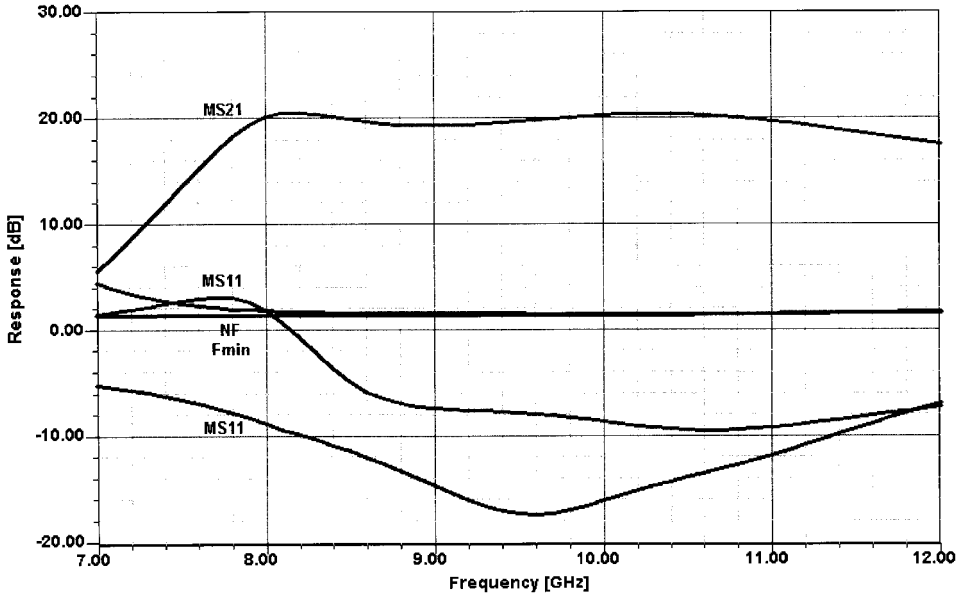


Figure 3-5 Simulated F_{min}, noise figure, S_{11}, S_{21}, and S_{22} responses for the three-stage GaAsFET amplifier using the nonlinear FET model. The values at 10 GHz are: F_{min}, 1.53 dB; NF, 1.65 dB; S_{11}, –8.5 dB; S_{21}, 20.3 dB; and S_{22}, –15.9 dB.

the noise figure and gain agreed quite well. The discrepancies between the results with the linear and nonlinear models emphasize the importance of accurate device modeling.

Figure 3-6 shows a wideband amplifier, specifically a distributed amplifier, that covers 1–20 GHz. Figure 3-7 shows its simulated frequency-dependent gain, matching, and noise performance. The example given here as the upper frequency for wireless systems is rapidly expanding and the technology to accommodate this is already in place. Modern CAD tools can handle such high frequencies accurately, which means one can trust them at the significantly lower wireless frequencies of 500 MHz to 3 GHz. Above 1500 MHz, one really should model using distributed elements. We will come to this later in this chapter.

Input/output and interstage matching will become another challenging task, specifically if one requires sufficient bandwidth. As circuits become more complex and as the operating frequency increases, passive components play an increasingly important role; the use of

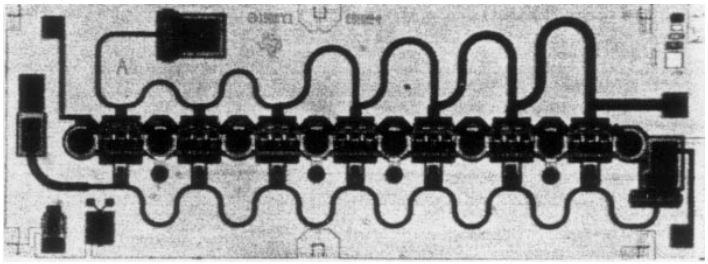


Figure 3-6 Photo of the 1–20-GHz distributed amplifier.

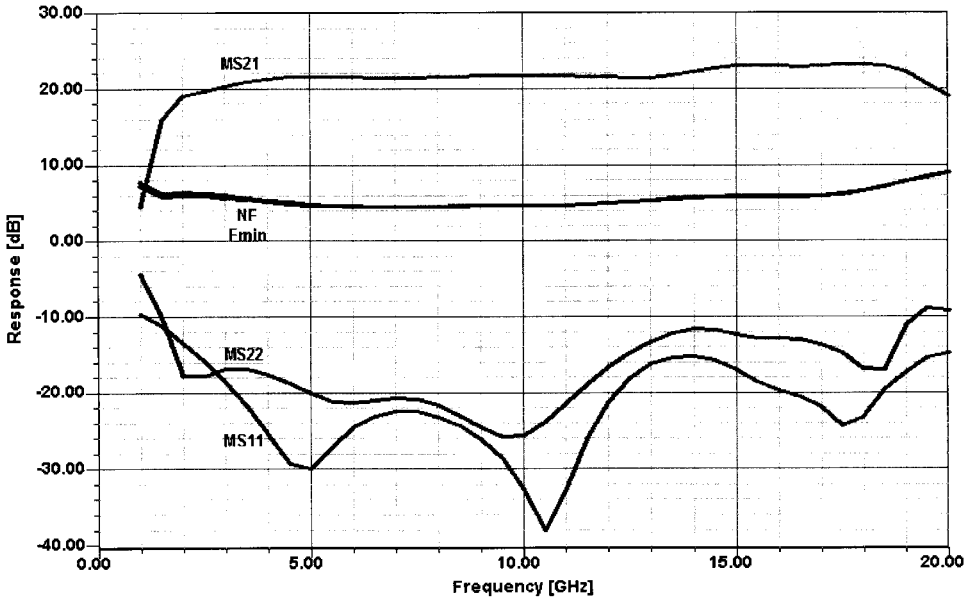


Figure 3-7 Simulated F_{min}, NF, S_{11}, S_{21}, and S_{22} responses. At the center of the range, the following numbers were found: F_{min}, 4.80 dB; NF, 4.80 dB; S_{11}, –32.8 dB; S_{21}, 21.8 dB; and S_{11}, –25.2 dB.

distributed elements is practically a necessity above 1500 MHz. For this reason, we begin this chapter by covering the application of transistors as amplifiers, following this with a discussion of the surrounding external circuitry, such as transmission lines, inductors, capacitors, and resistors. Monolithic ICs on silicon are a totally different ball game, and neither of the authors has sufficient expertise in this area. For this reason, we refer readers who will concentrate on MMIC design to the literature mentioned at the end of this chapter.

The goal of the applications presented here is to address, where possible, high-performance wireless stages that are less subject to constraints of space, power consumption, and cost, than portable/mobile applications. This is consistent with our preface, in which we have clearly stated our goal. The designers of handsets typically cannot consider the full variety of circuits we will present here. These types of stages (low noise, high gain and power) frequently will be used with filtering, and we will give some examples of input filters, including a tracking high-performance design.

A key topic besides the frequently mentioned gain is the whole noise issue. We will give the reader a complete introduction to this important area, followed by a similar introduction to the gain and matching of transistors. *Caution*: The majority of applications published go through the exercise of using *S*-parameters manually to design low-noise and small-signal amplifiers. While we have been tempted to repeat the information found in several textbooks, we feel that today's engineers will have access to CAD tools one way or another. It is highly unlikely that readers will undertake the inefficient process of going through all the design steps manually (a procedure that can take days); hopefully, engineers will resort to a linear/nonlinear CAD tool to do so. Only for the purpose of summarizing the necessary steps will we provide the relevant information regarding gain, stability, and related topics, demonstrating their usefulness or possible misuse with a modern CAD tool—in our case,

the Serenade Design Environment product from Ansoft. At this time, several additional companies beside Ansoft and Hewlett-Packard have emerged, but it is dangerous to list all of them as they come and go.

3-1-2 Gain

In practical terms, when we talk about gain we mean the output power relative to the input power. The industry standard now uses 50 Ω; however, at the point at which we have to match high-impedance devices such as monolithic crystal filters where the impedance jumps, things are not so obvious. The transformation equation that allows us to transform from one impedance to another is

$$m = \sqrt{\frac{R_2}{R_1}} \tag{3-1}$$

This correction factor, m, is required to do the transformation from R_1 (50 Ω) to R_2 (e.g., 1.2 kΩ). It needs to be understood that while the power gain remains the same, the voltage gain requires this correction factor. Since the power gain is expressed as 10 log A, A being the loop gain, 20 log A is the voltage gain, both expressed in dB. The reason for 10 versus 20 is that the power gain is always the voltage squared, which accounts for the 2. As a reminder of these relationships we present the following:

Distinction: Power Ratios–Voltage Ratios

$$n \text{ dB} = 10 \log_{10}(P_2/P_1) \quad \xrightarrow{PP = V^2/Z} \quad n = 10 \log_{10}\left(\frac{V_2^2/Z_2}{V_1^2/Z_1}\right)\text{dB}$$

$$Z_2 = Z_1 \rightarrow n \text{ dB} = 10 \log_{10}(V_2^2/V_1^2) = 10 \log_{10}(V_2/V_1)^2$$

$$\Longrightarrow \boldsymbol{n \text{ dB} = 20 \log_{10}(V_2/V_1) \text{ dB}}$$

The gain is typically defined under the condition of real resistive source and load terminations, which means that all the available power is being supplied to the amplifier, and likewise through proper matching at the output all the available power is fed to the termination. This gain figure is typically measured at a single spot frequency, but since most amplifiers are not totally flat, one needs to specify the gain tolerances or gain flatness. A good number is x dB ± 0.5 dB gain variation. In the case of a narrowband amplifier, the gain typically is shaped like a tuned circuit or like a critically coupled bandpass filter frequency response, while in the case of a real wideband application, the desired gain has some kind of ripple, which is due to the various compensation components in the circuit. An extreme case of a wideband amplifier is a 6–18-GHz single-stage amplifier that uses some very clever matching. This is shown in the following example. Figure 3-8 shows the circuit diagram, and Figure 3-9 shows its frequency-dependent performance.

While this is a cute approach for generating a single-stage amplifier over such a wide bandwidth, in real life it is not acceptable. The reason for this is that neither S_{11} nor S_{22} meets any useful specifications over the entire range. Typically, S_{11}/S_{22} values are only a few dB, while in reality more than 10 dB is required. We have chosen this example to show that gain

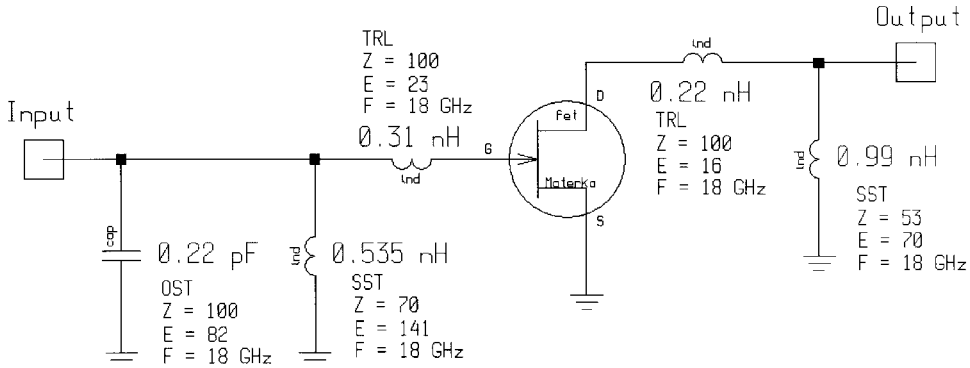


Figure 3-8 A 6–18-GHz FET amplifier.

and matching, depending on the circuit topology, are connected. We get similar results for the noise figure, whereby the best achievable noise figure ($F_{\min}$) and the actual 50-Ω noise figure are only a few dB apart; however, its frequency dependence in a wideband stage is certainly not acceptable.

Another parameter associated with gain is the stability factor K of the circuit. A value of $K > 1$ is mandatory to guarantee stability together with the S_{11} and S_{22} magnitudes being less than 1. Figure 3-10 plots the K stability factor, which for absolute stability must be >1, and maximum stable gain (MSG) as well as $G_{\max}$. It turns out that MSG is a highly artificial

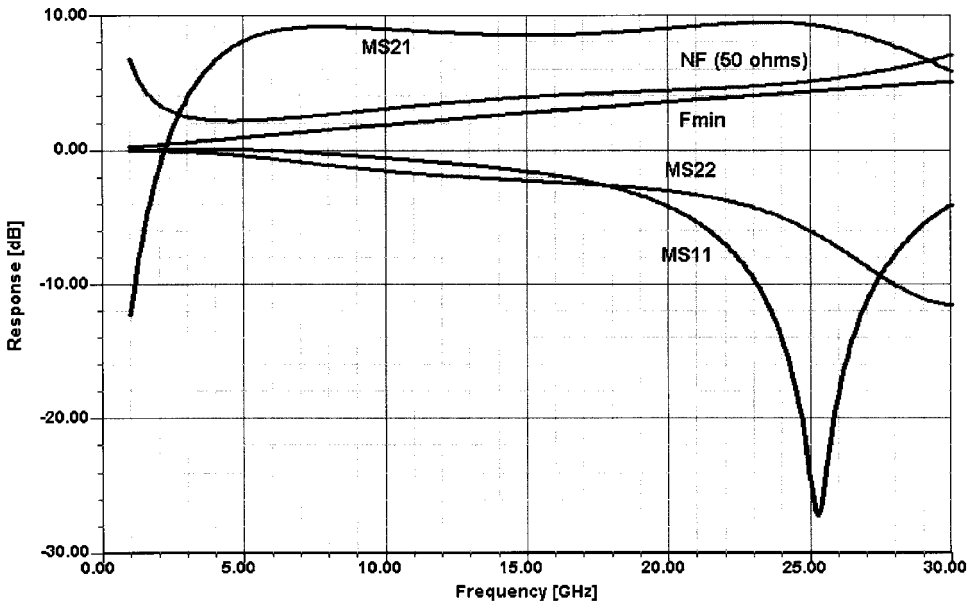


Figure 3-9 Frequency-dependent gain, matching, and noise performance of the amplifier shown in Figure 3-8.

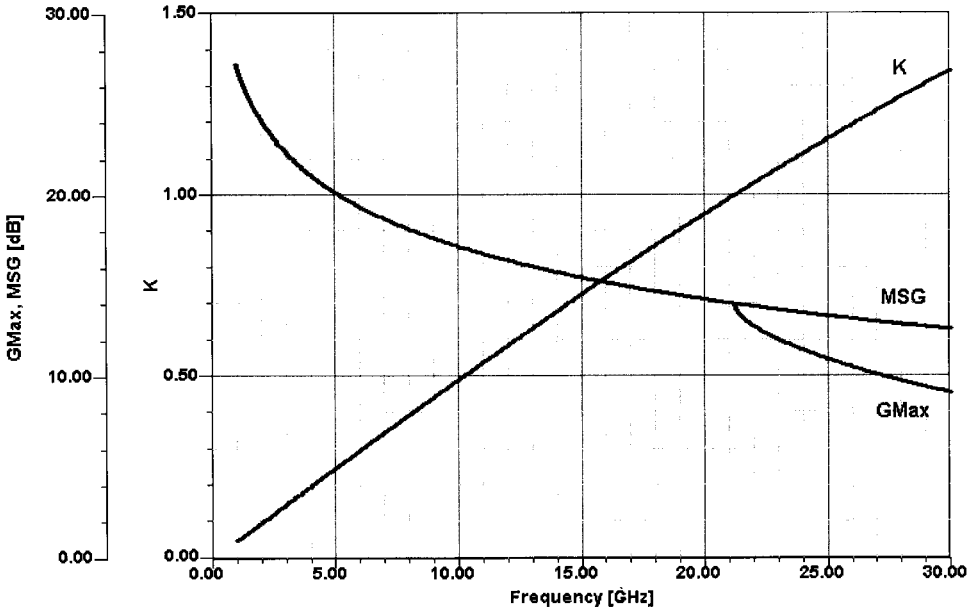


Figure 3-10 G_{max}, maximum stable gain (MSG), and K performance of the amplifier in Figure 3-8.

definition of stability that implies that input and output are terminated with pure resistances and then conjugately matched. There are four ways of achieving this resistive loading, either in the emitter as a shunt resistor between gate/drain or base/collector, or in parallel to the input and output to ground. Needless to say, this is not a real operating mode because it deteriorates both noise and gain. In our case, the entity G_{max} is a function of S_{12} of the circuit, not just the transistor, and depends highly on S_{12}. If we bring the system close to oscillation, then G_{max} will increase at the expense of bandwidth. Unfortunately, some of these microwave definitions of gain (see Table 3-1) have become widely used in industry without always being useful. Evaluating these equations for S_{12} is not legitimate because other quantities like K become infinite, and the entire system of equations falls apart. The power gain with input and output conjugately matched, or the transducer power gain for arbitrary source and load matching, still give more insight into the system.

In the above example, we specifically stated that the input and output VSWR was totally unacceptable, and this was due to a one-stage design. Figure 3-11 shows an amplifier chip built around two transistors with standard feedback techniques applied; Figures 3-12 and 3-13 show its frequency-dependent gain, matching, stability, and noise performance. This amplifier was one of the early examples of GaAsFET wideband amplifiers, but it is interesting because while it has wild (and wide) gain swings, it exhibits a fairly constant input and output impedance or reflection coefficient. A further evaluation of the circuit shows that not only is the circuit unconditionally stable, but that the maximum stable gain has an extremely high value based on essentially minimal S_{12}, the K factor is always significantly higher than 1, and G_{max} follows, to some degree, the actual gain response of the amplifier. (Ultimately, designers gave up on simple *RCL* feedback amplifiers for such a wide bandwidth, replacing them with distributed amplifiers previously called traveling wave amplifiers.) The problem of gain and stability starts at audiofrequency amplifiers, where the

Table 3-1 Nine power gains

Transducer power gain in 50-Ω system	$G_T = \lvert S_{21}\rvert^2$
Transducer power gain for arbitrary Γ_G and Γ_L	$G_T = \dfrac{(1-\lvert\Gamma_G\rvert^2)\,\lvert S_{21}\rvert^2\,(1-\lvert\Gamma_L\rvert^2)}{\lvert(1-S_{11}\Gamma_G)\,(1-S_{22}\Gamma_L)-S_{12}S_{21}\Gamma_G\Gamma_L\rvert^2}$
Unilateral transducer power gain	$G_{\mathrm{TU}} = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_G\rvert^2)\,(1-\lvert\Gamma_L\rvert^2)}{\lvert 1-S_{11}\Gamma_G\rvert^2\lvert 1-S_{22}\Gamma_L\rvert^2}$
Power gain with input conjugate matched	$G = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_L\rvert^2)}{\lvert 1-S_{22}\Gamma_L\rvert^2(1-\lvert S'_{11}\rvert^2)} = \dfrac{\lvert S_{21}\rvert^2}{1-\lvert S_{11}\rvert^2}$ (for $\Gamma_L = 0$)
Available power gain with output conjugate matched	$G_A = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_G\rvert^2)}{\lvert 1-S_{11}\Gamma_G\rvert^2(1-\lvert S'_{22}\rvert^2)} = \dfrac{\lvert S_{21}\rvert^2}{1-\lvert S_{22}\rvert^2}$ (for $\Gamma_G = 0$)
Maximum available power gain[a]	$G_{ma} = \left\lvert\dfrac{S_{21}}{S_{12}}\right\rvert\,(k-\sqrt{k^2-1})$
Maximum unilateral transducer power gain	$G_{\mathrm{TUmax}} = \dfrac{\lvert S_{21}\rvert^2}{(1-\lvert S_{11}\rvert^2)(1-\lvert S_{22}\rvert^2)}$
Maximum stable power gain	$G_{ms} = \dfrac{\lvert S_{21}\rvert}{\lvert S_{12}\rvert}$
Unilateral power gain	$U = \dfrac{1/2\lvert S_{21}/S_{12}-1\rvert^2}{k\lvert S_{21}/S_{12}\rvert-\mathrm{Re}(S_{21}/S_{12})}$

[a]Definition for $G_{\max}$: If $k > 1$, it is called G_{ma}, and if $k < 1$, it is called G_{ms}.

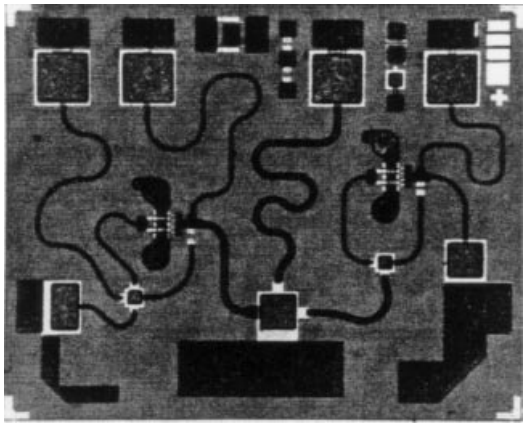


Figure 3-11 A two-stage, wideband feedback amplifier.

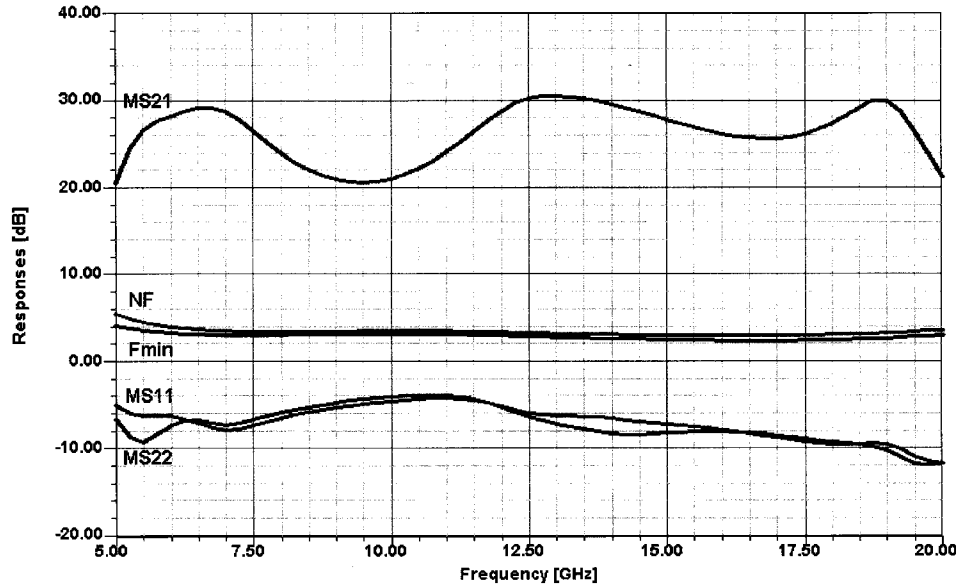


Figure 3-12 Frequency-dependent gain, matching, and noise performance for the amplifier in Figure 3-11.

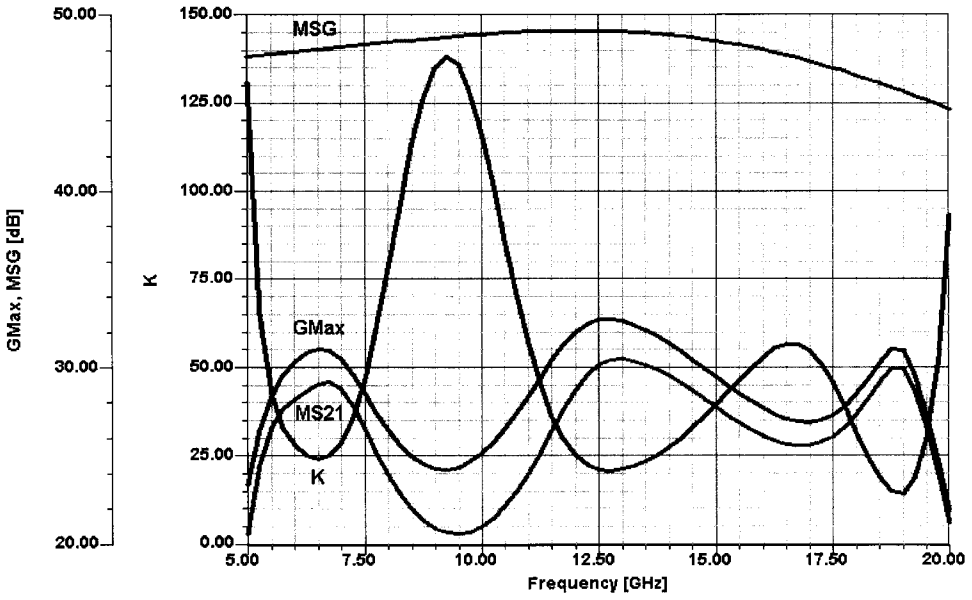


Figure 3-13 Frequency-dependent G_{max}, K, maximum stable gain, and gain performance for the amplifier in Figure 3-11.

motorboating phenomenon has been known for years, and the problem moves up in frequency with more modern technology and is largely dependent not only on the device and the circuit, but also on the layout. The motorboating effect has to do with ground loops and the fact that if the printed circuit traces from the power supply connection to the output power stage present too high a resistance, then at high currents the voltage will drop to a point below useful operation and then, as the volume is lowered, become stable again. In this case the amplitude of the signal through the amplifier controls the stability of the output stage. Needless to say, this most unpleasant effect applies to high-power stages for all output transistors, where we use Class B or C operation, meaning that the dc operating current as a function of output does not remain constant.

In terms of gain, there is also an important gain characteristic called *differential gain* and *differential phase*. These expressions were first used in television circuits, where linearity in both areas is an absolute must. The amplitude of the signal determines the black and white contrast, and the phase determines the color. As an example, transmitting an image of one of our presidents playing golf in California to the New York area is done by many microwave hops. Any multipath reception, such as that caused by reflection from airplanes, may have little effect on the signal amplitude, but the resulting phase shift can transform a green lawn to a red lawn if a total shift of 180° occurs. Given a particular operating point, if the 1-dB compression point of multiple amplifiers in the chain is approached or exceeded, their differential gain, which is defined by the IEEE as "the difference between (a) the ratio of the output of a small high-frequency sine-wave signal at two different levels of a low-frequency signal on which it is superimposed, and (b) unity," may increase. This mode of saturation charges all the capacitances of the amplifier transistors, causing huge phase shifts that result in the known problem of changing colors. The same applies, of course, in test equipment amplifiers, in which such changes cannot be tolerated. A somewhat similar, but not quite the same, problem is spectral regrowth [evaluated in terms of adjacent-channel power ratio (ACPR)], which is also due to nonlinearities. The question of linearity has already been covered in detail in Chapter 1. Needless to say, any overdrive condition also changes an amplifier's input and output SWR.

3-1-3 Noise Figure (NF)*

Introduction. In this section we look at the best way to describe noise in active devices (black boxes). Even when a two-port is linear, the output waveform will differ from the input, because of the failure to transmit all spectral components with equal gain (or attenuation) and delay. By careful design of the two-port, or by limitation of the bandwidth of the input waveform, such distortions can largely be avoided. However, noise generated within the two-port can still change the waveform of the output signal. In a linear passive two-port, noise arises only from the losses in the two-port; thermodynamic considerations indicate that such losses result in the random changes that we call noise. When the two-port contains active devices, such as transistors, there are other noise mechanisms present. A very important consideration in a system is the amount of noise that it adds to the transmitted signal. This is often judged by the ratio of the output signal power to the output noise power

*Section 3-1-3 is reprinted from Chapter 2 of George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*, copyright 1990 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

(S/N). The ratio of signal plus noise power to noise power [$(S + N)/N$] is generally easier to measure and approaches S/N when the signal is large.

In the evaluation of a two-port it is important to know the amount of noise added to a signal passing through it. An important parameter for expressing this characteristic is the noise factor. The signal energy coming from a generator or antenna is amplified or attenuated in passing from the input to the output of a two-port, as is the noise that accompanies the input signal energy. A system generally includes a cascade of two-port networks that constitute one overall two-port that amplifies the signal to a high-enough power level for its intended use. The *noise factor* of a system is defined as the ratio of signal-to-noise ratios available at input and output:

$$F = \frac{(S/N)_{\text{input}}}{(S/N)_{\text{output}}} \geq 1 \tag{3-2}$$

The noise figure (or factor) of a receiver is an easily measured quantity that describes the signal-to-noise ratio reduction of that receiver.

When this ratio of powers is converted to decibels, it is generally referred to as the *noise figure* rather than noise factor. Various conventions are used to distinguish the symbols used for noise factor and noise figure. Here we use F to represent the noise factor and NF to represent the noise figure, although the terms are usually used interchangeably.

For an amplifier with the power gain G, the noise factor can be rearranged as

$$F = \frac{S_i/N_i}{GS_i/G(N_i + N_a)} \tag{3-3}$$

when N_a is the additional noise power added by the amplifier referred to the input. This can be computed to be

$$F = 1 + N_a/N_i \tag{3-4}$$

The noise factor is often replaced by the noise figure (NF), which is defined in decibels as

$$\text{NF} = 10 \log_{10} F \tag{3-5}$$

In applications like satellite receivers, the noise factor becomes such a small number that it is inconvenient to work with. Many people have adopted the use of an equivalent noise temperature for a circuit to remedy this situation. Since the thermal noise power available from a resistor at temperature T_e is

$$N = kT_eB \tag{3-6}$$

where k is Boltzmann's constant (1.38×10^{-23} J/K), T_e is the effective temperature in kelvins, and B is the bandwidth in hertz. The equation above may be used to associate an effective noise temperature with circuits containing more than just thermal noise sources. This allows Eq. (3-4) to be written as

$$F = 1 + \frac{kT_eB}{kT_0B} = 1 + \frac{T_e}{T_0} \tag{3-7}$$

where T_e is the effective noise temperature of the circuit and T_0 is the temperature of the generator resistor in kelvins. The noise temperature T_e now characterizes our circuit noise contribution and can directly be related to the noise factor.

Assuming a reference noise temperature of 290 K (–273 + 290 = 17 °C), let us determine the noise temperature of the system with a noise factor of 2.6 (4.15 dB):

$$T_e = (2.6 - 1)\ (290) = 464 \text{ K}$$

This temperature T_e should not be confused with the environmental operating temperature T_0. It is quite common to operate low-noise amplifiers with T_e below 100 K at an ambient temperature of 290 K.

Signal-to-Noise Ratio. Let us consider the signal-to-noise ratio of power delivered from a generator to a load as shown in Figure 3-14. The signal power delivered to the input is given by

$$S_{\text{in}} = P_{\text{in}} = \frac{E_g^2 \text{Re}(Z_{\text{in}})}{|Z_g + Z_{\text{in}}|^2} \tag{3-8}$$

where E_g is the rms voltage of the input signal supplied to the system, and the noise power supplied to the input is expressed by

$$N_{\text{in}} = \frac{\overline{v_n^2} \text{Re}(Z_{\text{in}})}{|Z_g + Z_{\text{in}}|^2} \tag{3-9}$$

where the noise power at the input is provided by the noise energy of the real part of Z_g. The input impedance Z of the system in the form $Z = R_{\text{in}} + jX_{\text{in}}$ is assumed to be complex.

The Johnson noise of a resistor [here $\text{Re}(Z_g)$], given by the mean-square voltage

$$\overline{v_n^2} = 4kTRB \tag{3-10}$$

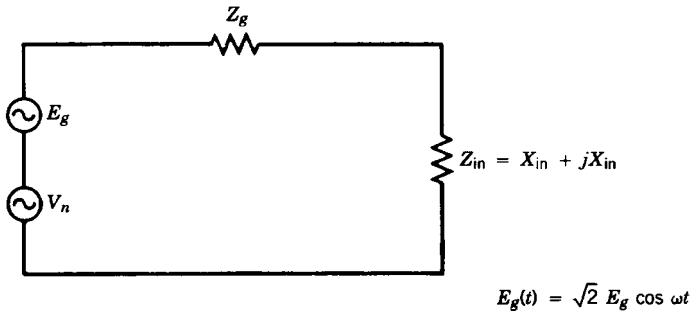


Figure 3-14 Combination of signal and noise voltages supplied to a complex termination.

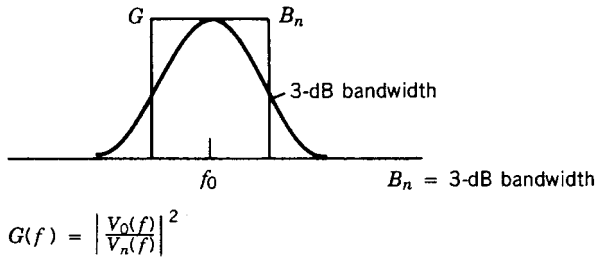
where k (Boltzmann's constant) = 1.38×10^{-23} J/K, T is the absolute temperature of the resistor, and B is the bandwidth, is sufficiently small that the resistive component of impedance does not change. The available signal power from the generator has a lower limit, even if the signal is attenuated by the highest possible attenuation. The generator resistor acts as a Johnson noise generator, its power being

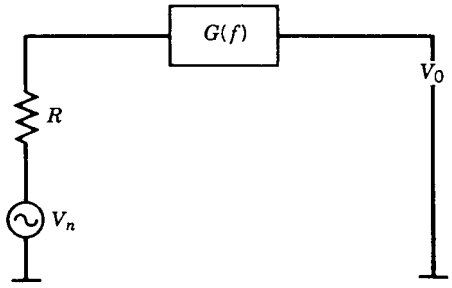
$$P_A = \frac{4kTRB}{4R} = kTB \tag{3-11}$$

where k is Boltzmann's constant, T is the absolute temperature, and B is the bandwidth. This power is the maximum available output power.

For an ambient temperature of 290 K, $kT = 4 \times 10^{-21}$ W/Hz. This expression is also given as kT = –204 dBW/Hz = –174 dBm/Hz = –114 dBm/MHz. We can combine Eqs. (3-8) to (3-10) to obtain

$$\left(\frac{S}{N}\right)_{\text{in}} = \frac{E_g^2}{4kT\text{Re}(Z_g)B} \tag{3-12}$$





$$V_0^2 = \int_0^\infty 4kT_0RG(f)\,dF$$
$$= 4kT_0R \int_0^\infty G(f)\,dF$$
$$B_n = \frac{1}{G} \int_0^\infty G(f)\,dF$$
B_n = noise bandwidth

Figure 3-15 Graphical and mathematical explanations of the noise bandwidth from a comparison of the Gaussian-shaped bandwidth to the rectangular filter response.

This is the value of *S/N* contributed by the generator, which does not include the noise generated by the load, in this case $\mathrm{Re}(Z_{\mathrm{in}})$, which would need to be included in the measurement of the total *S/N* across the input impedance.

A critical parameter is the noise bandwidth, B_n, which is defined as the equivalent bandwidth, as shown in Figure 3-15. For reasons of group delay correction, most practical filters have round rather than sharp corners. The noise figure measurements shown later can be used to determine the "integrated" bandwidth, which is B_n.

An active system such as a combination of amplifiers and mixers will add noise to the input signals, and the noise factor that describes this is defined as the *S/N* ratio at the input to the *S/N* ratio at the output, which is always greater than unity [2]. In practice, a certain minimum signal-to-noise ratio is required for operation. For example, in a communication system such a minimum is required for intelligible transmission, either voice or data. For high-performance TV reception, to provide a noise-free picture to the eye, a typical requirement is for a 60-dB *S/N*. In the case of a TV system, a large dynamic range is required, as well as a very large bandwidth to reproduce all colors truthfully and all shades from high-intensity white to black. Good systems will have a bandwidth of 8 MHz or more.

Noise Figure Measurements. Some of the noise equations are based on mathematical models and physics. To understand some of these expressions, it is useful to look at a practical system, with amplifiers, that has to be evaluated.

Let us look at Figure 3-16, which consists of a signal generator, the system or device under test (DUT), and a selective receiver with a built-in root-mean-square (rms) voltmeter to determine the signal and noise voltages. It is necessary that the system have enough gain so that the noise voltage supplied by the generator will be indicated [3].

If we assume that our selective receiver is a video noise meter calibrated in rms voltage levels, we can perform two measurements. With an input termination connected to the TV system (typically, 75 Ω for cable TV, 50 Ω for satellite TV), the noise receiver/meter will read a value for proper termination that can easily be calculated. Since one-half of the mean-square noise voltage appears across the input,

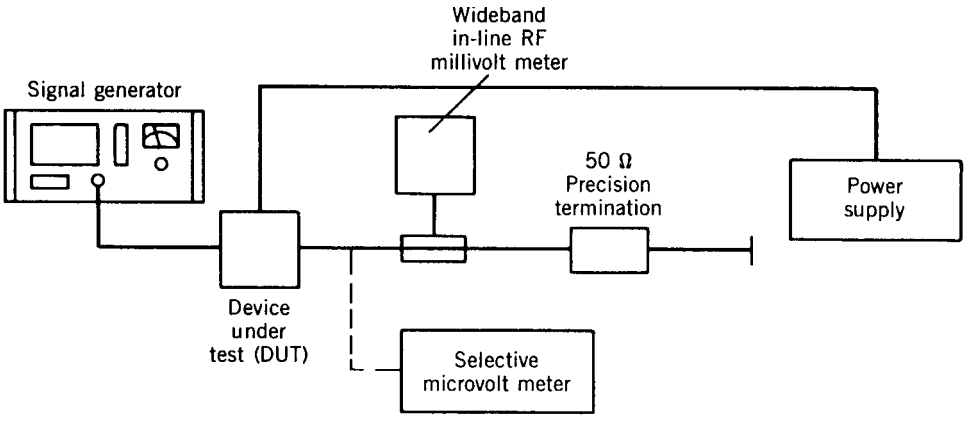


Figure 3-16 Test setup to measure signal-to-noise ratio.

$$v_{\text{in}} = \frac{v_n}{2} = \frac{\sqrt{4kTRB}}{2} \tag{3-13}$$

With $B = 10$ MHz, $T = 290$ K (T is always expressed in absolute temperature; $T_0 = -273$ °C), and $k = 1.38 \times 10^{-23}$ J/K, then for $R = 75\ \Omega$,

$$v_{\text{in}} = \frac{v_n}{2} = 1.73\ \mu\text{V} \tag{3-14}$$

where the rms noise voltage has been referred to the input port. We can verify this with our first measurement.

Now we increase the input voltage of the signal generator to a value that indicates a 60-dB *S/N* ratio at the output port. This should be about

$$E_g = \frac{v_n}{2}\sqrt{F} \times 1000 = 1.73\sqrt{F} \quad \text{mV}$$

where F is the noise factor of the receiver. For a receiver noise factor of 10, we would obtain $E_g = 5.48$ mV (rms value). If the noise energy equivalent to a noise factor of F is assumed, we need $\sqrt{F}$ times more voltage. For a 60-dB ratio, this means that $E_g = 1000 \times (v_n/2) \times \sqrt{F}$.

As they are done here, over a power range of 60 dB, the measurements can be performed over such a wide range only if special equipment is available. In cases where the internal detector of a piece of communication equipment is used, the signal-to-noise ratio measurements are performed over much smaller power ranges.

Let us assume that for the above-mentioned case ($F = 10$) we find a *S/N* ratio of 10 dB at the output for an input signal of 5.47 μV. By rewriting Eq. (3-7) as

$$E_g = \frac{v_n}{2}\sqrt{F} = \sqrt{KTRBF} \tag{3-15}$$

with F being the noise factor, we can solve for F with

$$F = \frac{P_s}{P_n} = \frac{E_g^2/R}{kTB} \tag{3-16}$$

While the input power from the thermal energy of the input termination resistor was $kTB = 4 \times 10^{-4}$ W, the input power required for the 10-dB *S/N* ratio was

$$P_s = \frac{(5.47 \times 10^{-6})^2}{75} = 3.98 \times 10^{-13}\ \text{W} \tag{3-17}$$

The noise factor is defined as the ratio P_s/P_n:

$$F = \frac{3.98 \times 10^{-13}}{4 \times 10^{-14}} = 10 \tag{3-18}$$

which is the proof.

This method is used more frequently at the 3-dB point, or double the input power if the dynamic range of the detector is small or only a linear indicator is available. Because of hum and other pickup, this is not an easy measurement. Using a signal generator is very expensive because in a laboratory or production environment a wide frequency range requires several generators.

Another method is the use of a wideband noise generator. Modern gas discharge diodes or avalanche diodes are available that provide essentially white noise energy over a large frequency range. These microwave diodes typically have an output of 30 dB above kT when switched on and kT when switched off. To provide good matching at microwave frequencies, a 15-dB attenuator is cascaded. This means that the noise power of the source in the ON condition is about 15 dB above kT.

In the early 1960s, low-cost noise figure test equipment was built around vacuum diodes whose operating range was limited to 1200 MHz due to the resonance effects of the structure. Today, automatic noise gain analyzers use calibrated solid-state noise sources up to 26.5 GHz. It appears that the upper frequency limit has to do with matching, and the lower-frequency limit with $1/f$ noise.

Noisy Two-Port Description. Based on the convention by Rothe and Dalke [4], any linear two-port can be in the form shown in Figure 3-17. This general case of a noisy two-port can be redrawn showing noise sources at the input and at the output. Figure 3-17b shows this in admittance form and Figure 3-17c in impedance form. The internal noise sources are assumed to produce very small currents and voltages, and we assume that linear two-port equations are valid. The internal noise contributions have been expressed by using external noise sources:

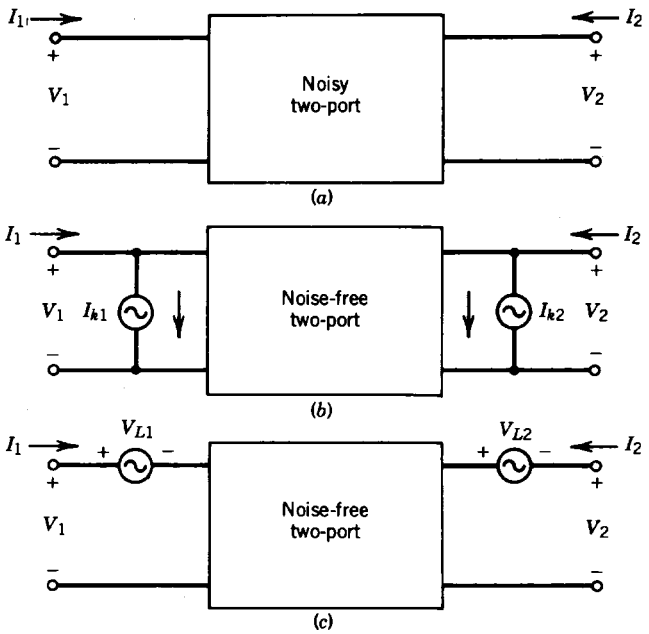


Figure 3-17 Noisy linear two-ports: (a) general form, (b) admittance form, and (c) impedance form.

$$I_1 = y_{11}V_1 + y_{12}V_2 + I_{K1}$$
$$I_2 = y_{21}V_1 + y_{22}V_2 + I_{K2} \tag{3-19}$$

$$V_1 = z_{11}I_1 + z_{12}I_2 + V_{L1}$$
$$V_2 = z_{21}I_1 + z_{22}I_2 + V_{L2} \tag{3-20}$$

where the external noise sources are I_{K1}, I_{K2}, V_{L1}, and V_{L2}.

Since we want to describe our noisy circuit in terms of the noise figure, the *ABCD*-matrix description will be more convenient since it refers both noise sources to the input of the two-port [5]. This representation is given below (note the change in direction of I_2):

$$V_1 = AV_2 + BI_2 + V_A$$
$$I_1 = CV_2 + DI_2 + I_A \tag{3-21}$$

where V_A and I_A are the external noise sources.

It is important to remember that all of these matrix representations are interrelated. For example, the noise sources for the *ABCD*-matrix description can be obtained from the *z*-matrix representation shown in Eq. (3-20). This transformation is

$$V_A = -\frac{I_{K2}}{y_{21}} = V_{L1} - \frac{V_{L2}z_{11}}{z_{21}} \tag{3-22}$$

$$I_A = I_{K1} - \frac{I_{K2}y_{11}}{y_{21}} = -\frac{V_{L2}}{z_{21}} \tag{3-23}$$

The *ABCD* representation is particularly useful because it allows us to define a noise temperature for the two-port referenced to the input. The two-port itself (shown in Figure 3-18) is assumed to be noise-free.

In the past, *z* and *y* parameters have been used, but in microwave applications it has become common to use *S*-parameter definitions. This is shown in Figure 3-19. The previous equations can be rewritten in their new form using *S*-parameters:

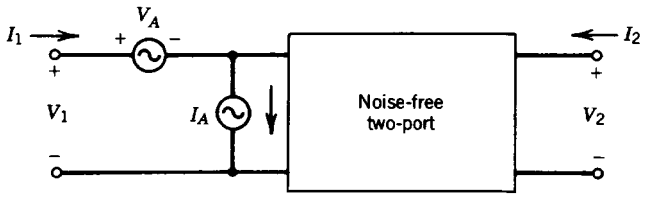


Figure 3-18 Chain-matrix form of linear noisy two-ports.

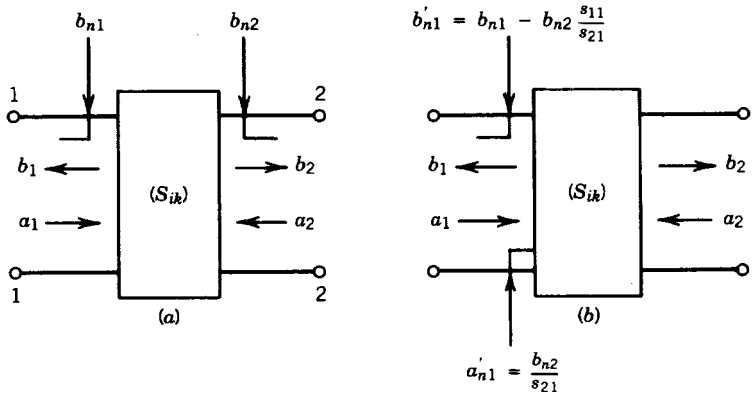


Figure 3-19 *S*-parameter form of linear noisy two-ports.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_{n1} \\ b_{n2} \end{bmatrix} \tag{3-24}$$

There are different physical origins for the various sources of noise. Typically, thermal noise is generated by resistances and loss in the circuit or transistor, whereas shot noise is generated by current flowing through semiconductor junctions and vacuum tubes. Since these many sources of noise are represented by only two noise sources at the device input, the two equivalent noise sources are often a complicated combination of the circuit internal noise sources. Often, some fraction of V_A and I_A is related to the same noise source. This means that V_A and I_A are not independent in general. Before we can use V_A and I_A to calculate the noise figure of the two-port, we must calculate the correlation between the V_A and I_A shown in Figure 3-18.

The noise source V_A represents all the device noise referred to the input when the generator impedance is zero; that is, when the input is short-circuited. The noise source I_A represents all the device noise referred to the input when the generator admittance is zero; that is, the input is open-circuited.

The correlation of these two noise sources considerably complicates analysis. By defining a correlation admittance, we can simplify the mathematics and get some physical intuition for the relationship between noise figure and generator admittance. Since some fraction of I_A will be correlated with V_A, we split I_A into correlated and uncorrelated parts as follows:

$$I_A = I_n + I_u \tag{3-25}$$

I_u is the part of I_A uncorrelated with V_A. Since I_n is correlated with V_A, we can say that I_n is proportional to V_A and the constant of proportionality is the correlation admittance:

$$I_n = Y_{\text{cor}} V_A \tag{3-26}$$

This leads us to

$$I_A = Y_{\text{cor}} V_A + I_u \tag{3-27}$$

The following derivation of noise figure will use the correlation admittance. Y_{cor} is not a physical component located somewhere in the circuit. Y_{cor} is a complex number derived by correlating the random variables I_A and V_A. To calculate Y_{cor}, we multiply each side of Eq. (3-27) by V_A^* and average the result. This gives

$$\overline{V_A^* I_A} = Y_{\text{cor}} \overline{V_A^2} \tag{3-28}$$

where the I_u term averaged to zero since it was uncorrelated with V_A. The correlation admittance is thus given by

$$Y_{\text{cor}} = \frac{\overline{V_A^* I_A}}{V_A^2} \tag{3-29}$$

Often, people use the term "correlation coefficient." This normalized quantity is defined as

$$c = \frac{\overline{V_A^* I_A}}{\sqrt{\overline{V_A^2} I_A^2}} = Y_{\text{cor}} \sqrt{\frac{\overline{V_A^2}}{I_A^2}} \tag{3-30}$$

Note that the dual of this admittance description is the impedance description. Thus, the impedance representation has the same equations as above with Y replaced by Z, I replaced by V, and V replaced by I.

V_A and I_A represent internal noise sources in the form of a voltage source acting in series with the input voltage and a source of current flowing in parallel with the input current. This representation conveniently leads to the four noise parameters needed to describe the noise performance of the two-port. Again using the Nyquist formula, the open-circuit voltage of a resistor at the temperature T is

$$\overline{V_A^2} = 4kTRB \tag{3-31}$$

This voltage is a mean-square fluctuation (or spectral density). It is the method used to calculate the noise identity. We could also define a noise equivalent resistance for a noise voltage as

$$R_n = \frac{\overline{V_A^2}}{4kTB} \tag{3-32}$$

The resistor R_n is not a physical resistor but can be used to simulate different portions of the noise equivalent circuit.

In a similar manner a mean-square current fluctuation can be represented in terms of an equivalent noise conductance G_n, which is defined by

$$G_n = \frac{\overline{I_A^2}}{4kTB} \tag{3-33}$$

and

$$G_u = \frac{\overline{I_u^2}}{4kTB} \tag{3-34}$$

for the case of the uncorrelated noise component. The input generator to the two-port has a similar contribution:

$$G_G = \frac{I_G^2}{4kTB} \tag{3-35}$$

with Y_G being the generator admittance and G_G being the real part. With the definition of F above, we can write

$$F = 1 + \left| \frac{I_A + Y_G V_A}{I_G} \right|^2 \tag{3-36}$$

The use of the voltage V_A and the current I_A has allowed us to combine all the effects of the internal noise sources.

We can use the previously defined [Eq. (3-29)] correlation admittance, $Y_{\text{cor}} = G_{\text{cor}} + jB_{\text{cor}}$, to simplify Eq. (3-36). First, we determine the total noise current:

$$\overline{I_A^2} = 4kT(|Y_{\text{cor}}|^2 R_n + G_u)B \tag{3-37}$$

where R_n and G_u are defined in Eqs. (3-32) and (3-34). The noise factor can now be determined:

$$F = 1 + \frac{G_u}{G_g} + \frac{R_n}{G_g}[(G_G + G_{\text{cor}})^2 + (B_G + B_{\text{cor}})^2] \tag{3-38}$$

$$F = 1 + \frac{R_u}{R_g} + \frac{G_n}{R_g}[(R_G + R_{\text{cor}})^2 + (X_G + X_{\text{cor}})^2] \tag{3-39}$$

The noise factor is a function of various elements, and the optimum impedance for the best noise figure can be determined by minimizing F with respect to generator reactance and resistance. This gives

$$R_{0n} = \sqrt{\frac{R_u}{G_n} + R_{\text{cor}}^2} \tag{3-40}$$

$$X_{0n} = -X_{\text{cor}} \tag{3-41}$$

and

$$F_{\min} = 1 + 2G_n R_{\text{cor}} + 2\sqrt{R_u G_n + (G_n R_{\text{cor}})^2} \tag{3-42}$$

(To distinguish between optimum noise and optimum power, we have introduced the convention 0*n* instead of the more familiar abbreviation *opt*.) At this point we see that the optimum condition for minimum noise figure is not a conjugate power match at the input port. We can explain this by recognizing that the noise source V_A and I_A represent all the

two-port noise, not just the thermal noise of the input port. We should observe that the optimum generator susceptance, $-X_{corr}$, will minimize the noise contribution of the two noise generators.

In rearranging for conversion to S parameters, we write

$$F = F_{\min} + \frac{g_n}{R_G}|Z_G - Z_{0n}|^2 \tag{3-43}$$

$$F = F_{\min} + \frac{R_n}{G_G}|Y_G - Y_{0n}|^2 \tag{3-44}$$

From the definition of the reflection coefficient,

$$\Gamma_G = \frac{Y_0 - Y_G}{Y_0 + Y_G} \tag{3-45}$$

and with

$$g_G = \frac{G_G}{Y_0} \tag{3-46}$$

$$r_n = \frac{R_n}{Z_0} \tag{3-47}$$

the normalized equivalent noise resistance,

$$F = F_{\min} + \frac{4r_n|\Gamma_G - \Gamma_{0n}|^2}{g_G(1 - |\Gamma|^2)|1 + \Gamma_{0n}|^2} \tag{3-48}$$

$$r_n = (F_{50} - F_{\min})\frac{|1 + \Gamma_{0n}|^2}{4|\Gamma_{0n}|^2} \tag{3-49}$$

$$\Gamma_{0n} = \frac{Z_{0n} - Z_0}{Z_{0n} + Z_0} \tag{3-50}$$

The noise performance of any linear two-port can now be determined if the values of the four noise parameters, $F_{\min}$, r_n (= $R_n/50$), and Γ_{0n} (magnitude and phase), are known.

Figure 3-20 shows the noise factor of a high-frequency transistor as a function of B_g for G_g = constant and as a function of G_g for $B_g = B_{opt}$.

Noise Figure of Cascaded Networks. In a system with many circuits connected in cascade (Figure 3-21), we must consider the contributions of the various circuits. In considering the equivalent noise resistor R_n in series with the input circuit,

$$F = \frac{R_G + R_n}{R_G} \tag{3-51}$$

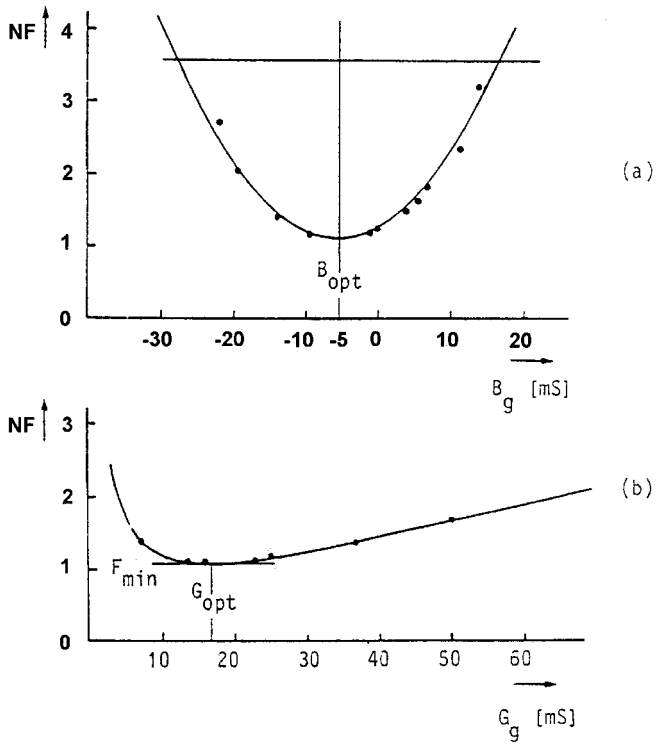


Figure 3-20 Noise factor in high-frequency BJTs for $f = 600$ MHz: (a) as a function of B_g for G_g = constant and (b) as a function of G_g for $B_g = B_{opt}$.

$$F = 1 + \frac{R_n}{R_G} \tag{3-52}$$

The excess noise added by the circuit is R_n/R_G.

In considering two cascaded circuits a and b, by definition the available noise at the output of b is

$$N_{ab} = F_{ab} G_{ab} kTB \tag{3-53}$$

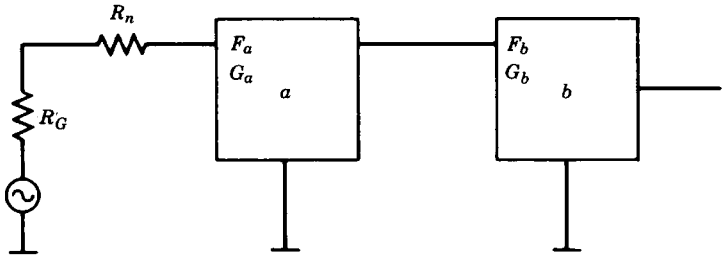


Figure 3-21 Cascaded noisy two-ports with the noise factors F_a and F_b and the gain figures G_a and G_b.

with B the equivalent noise bandwidth in which the noise is measured. The total available gain G is the product of the individual available gains, so

$$N_{ab} = F_{ab} G_a G_b kTB \tag{3-54}$$

The available noise from network a at the output of network b is

$$N_{a/b} = N_a G_b = F_a G_a G_b kTB \tag{3-55}$$

The available noise added by network b (its excess noise) is

$$N_{b/b} = (F_b - 1) G_b kTB \tag{3-56}$$

The total available noise N_{ab} is the sum of the available noise contributed by the two networks:

$$\begin{aligned} N_{ab} &= N_{a/b} + N_{b/b} = F_a G_a G_b kTB + (F_b - 1) G_b kTB \\ &= \left(F_a + \frac{F_b - 1}{G_a} \right) G_a G_b kTB \end{aligned} \tag{3-57}$$

$$F_{ab} = F_a + \frac{F_b - 1}{G_a} \tag{3-58}$$

For any number of circuits, this can be extended to

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \cdots \tag{3-59}$$

When considering a long chain of cascaded amplifiers, there will be a minimum noise figure achievable for this chain. This is a figure of merit and was proposed by Haus and Adler [6]. It is calculated by rearranging Eq. (3-59):

$$(F_{\text{tot}})_{\min} = (F_{\min} - 1) + \frac{F_{\min} - 1}{G_A} + \frac{F_{\min} - 1}{G_A^2} + \cdots + 1 \tag{3-60}$$

where $F_{\min}$ is the minimum noise figure for each stage and G_A is the available power gain of the identical stages. Using

$$\frac{1}{1 - X} = 1 + X + X^2 + \cdots \tag{3-61}$$

we find a quantity $(F_{\text{tot}} - 1)$, which is defined as the noise measure M. The minimum noise measure

$$(F_{\text{tot}})_{\min} - 1 = \frac{F_{\min} - 1}{1 - 1/G_A} = M_{\min} \tag{3-62}$$

refers to the noise of an infinite chain of optimally tuned, low-noise stages, so it represents a lower limit on the noise of an amplifier.

The minimum noise measure $M_{\min}$ is an invariant parameter and is not affected by feedback. It is somewhat similar to a gain–bandwidth product, in its use as a system invariant. The minimum noise measure is achieved when the amplifier is tuned for the available power gain and $\Gamma_G = \Gamma_{0n}$, given by Eq. (3-50).

Influence of External Parasitic Elements. Mounting an active two-port such as a transistor usually adds stray capacitance and lead inductance to the device, as shown in Figure 3-22. These external components consisting of transmission lines and parasitic reactances modify the noise parameters and the gain. Some researchers have published the results of these parasitic effects and have made manual computations or used some limited computer programs.

In a paper by Fukui [7], an attempt was made to determine the necessary equations, but the formulas are too involved even for pocket calculators. A more generic study by Iversen [8] is also very involved because of the various matrix manipulations and is more suitable for a computer. Besser's paper [9] in the *IEEE MTT-S* in 1975 and Vendelin's paper [10] in the same issue have shown for the first time some practical results using computers and even optimization methods, using an early version of COMPACT. The intention of these investigations was to find feedback that modifies the device noise and scattering parameters such that a noise match could also provide a low-input VSWR. It can be seen from these discussions that some feedback, besides resulting in some gain reduction, may improve the noise matching at the input for a limited frequency range.

A more recent paper by Suter [11] based on a report by Hartmann and Strutt [12] has given a simple transformation starting from the *S*-parameters and the noise parameters from common-source (or common-emitter) measurements. The noise parameters for the "packaged" device are calculated. This means that the parameters for the "new" device, including

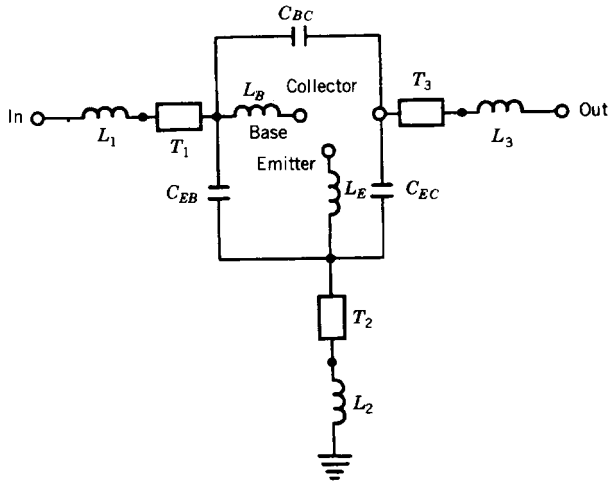
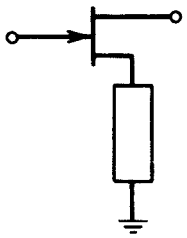


Figure 3-22 Equivalent circuit of the transistor package.

Table 3-2 Transformation matrix parameters

Series Feedback



$$Z_S = R_S + jX_S = Z_S$$

$$[n] = \begin{bmatrix} n_{11} = 1 & n_{12} = Z_0 \dfrac{S_{21}M - S'_{21}N}{S_{21}C'_1 + S'_{21}C_1}\ \Omega \\ n_{21} = 0 & n_{22} = \dfrac{S_{21}C'_1}{S_{21}C'_1 + S'_{21}C_1} \end{bmatrix}$$

where

$$S'_{11} = S'_{22} = \frac{-1}{1 + 2Z_S}$$

$$S'_{12} = S'_{21} = \frac{2Z_S}{1 + 2Z_S}$$

$$M = (1 + S'_{11})(1 - S'_{22}) + S'_{12}S'_{21}$$

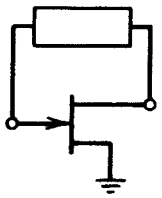
$$N = (1 + S_{11})(1 - S_{22}) + S_{12}S_{21}$$

$$C_1 = (1 - S_{11})(1 - S_{22}) - S_{12}S_{21}$$

$$C'_1 = (1 + S'_{11})(1 - S'_{22}) - S'_{12}S'_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

Shunt Feedback



$$Z_P = R_P + jX_P = Z_P$$

$$[n] = \begin{bmatrix} n_{11} = \dfrac{S_{21}C'_2}{S_{21}C'_2 + S'_{21}C_2} & n_{12} = 0\ \Omega \\ n_{21} = \dfrac{1}{Z_0}\dfrac{S_{21}P - S'_{21}Q}{S_{21}C'_2 + S'_{21}C_2}\,S & n_{22} = 1 \end{bmatrix}$$

where

$$S'_{11} = S'_{22} = \frac{Z_P}{2 + Z_P}$$

(*continued*)

Table 3-2 (*Continued*)

$$S'_{12} = S'_{21} = \frac{2}{2 + Z_P}$$

$$P = (1 - S'_{11})(1 + S'_{22}) + S'_{12}S'_{21}$$

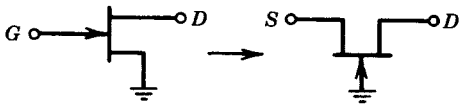
$$Q = (1 - S_{11})(1 + S_{22}) + S_{12}S_{21}$$

$$C_2 = (1 + S_{11})(1 + S_{22}) - S_{12}S_{21}$$

$$C'_2 = (1 + S'_{11})(1 + S'_{22}) - S'_{12}S'_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

Common Gate



$$[n] = \begin{bmatrix} n_{11} = \dfrac{2S_{21}}{-2S_{21} + C_4} & n_{12} = 0\ \Omega \\ n_{21} = \dfrac{1}{Z_0}\dfrac{C_3C_4 - 4S_{12}S_{21}}{V(-2S_{21} + C_4)}S & n_{22} = -1 \end{bmatrix}$$

where

$$V = (1 + S_{11})(1 + S_{22}) - S_{12}S_{21}$$

$$C_3 = (1 - S_{11})(1 + S_{22}) + S_{12}S_{21}$$

$$C_4 = (1 + S_{11})(1 - S_{22}) + S_{12}S_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \text{common-source } S \text{ parameters}$$

the common-gate (or common-base) case, are calculated. The equations are device dependent. They are valid for any active two-port.

A transformation matrix, n, may be used to combine the noise sources of the various circuit configurations. The transformation matrix parameters are given in Table 3-2 for series feedback, shunt feedback, and the common-gate (base) case, which will be important for oscillator analysis. The transformation matrix gives the new noise parameters as follows:

$$R'_n = R_n|n_{11} + n_{12}Y_{\text{cor}}|^2 + G_n|n_{12}|^2 \tag{3-63}$$

$$G'_n = \frac{G_nR_n}{R'_n}|n_{11}n_{22} - n_{12}n_{21}|^2 \tag{3-64}$$

$$Y'_{\text{cor}} = \frac{R_n}{R'_n}(n_{21} + n_{22}Y_{\text{cor}})(n^*_{11} + n^*_{12}Y^*_{\text{cor}}) + \frac{G_n}{R'_n}n_{22}n^*_{12} \tag{3-65}$$

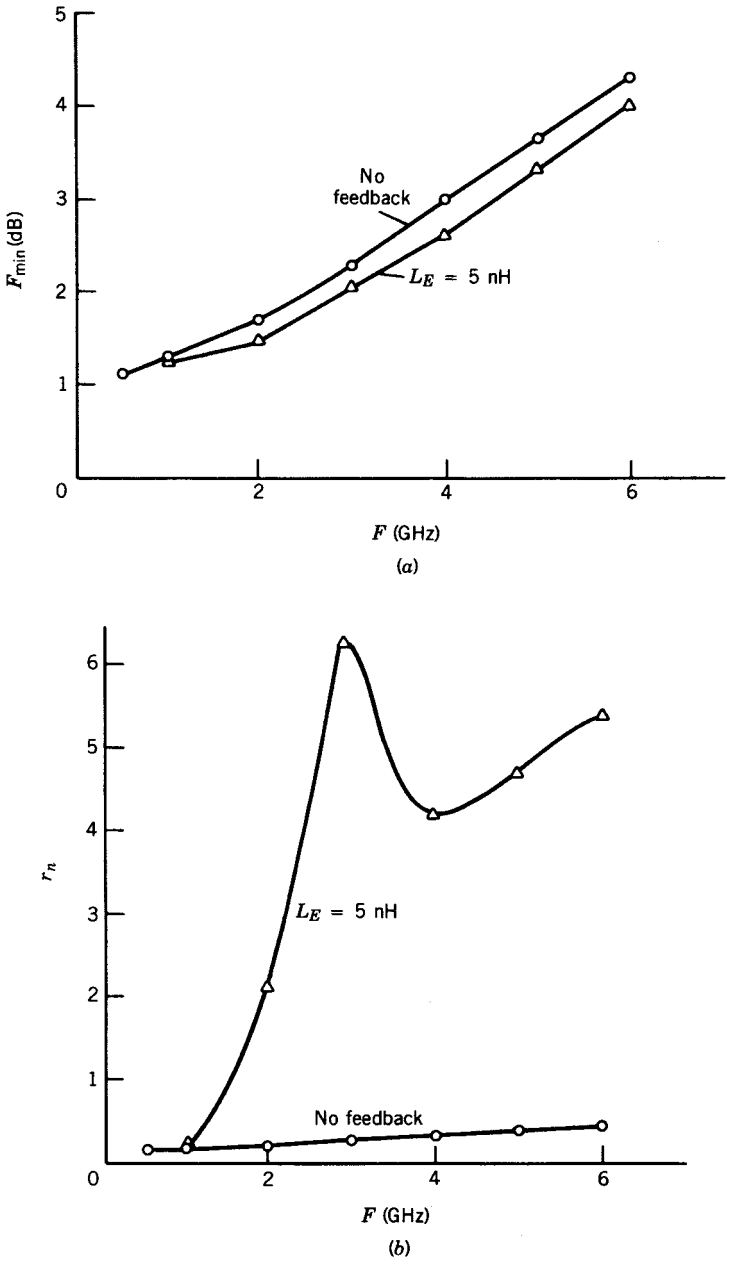


Figure 3-23 Noise parameters versus feedback for the AT-41435 silicon bipolar transistor: (a) F_{min} versus frequency and feedback; (b) r_n versus frequency and feedback; (c) Γ_{0n} versus frequency and feedback.

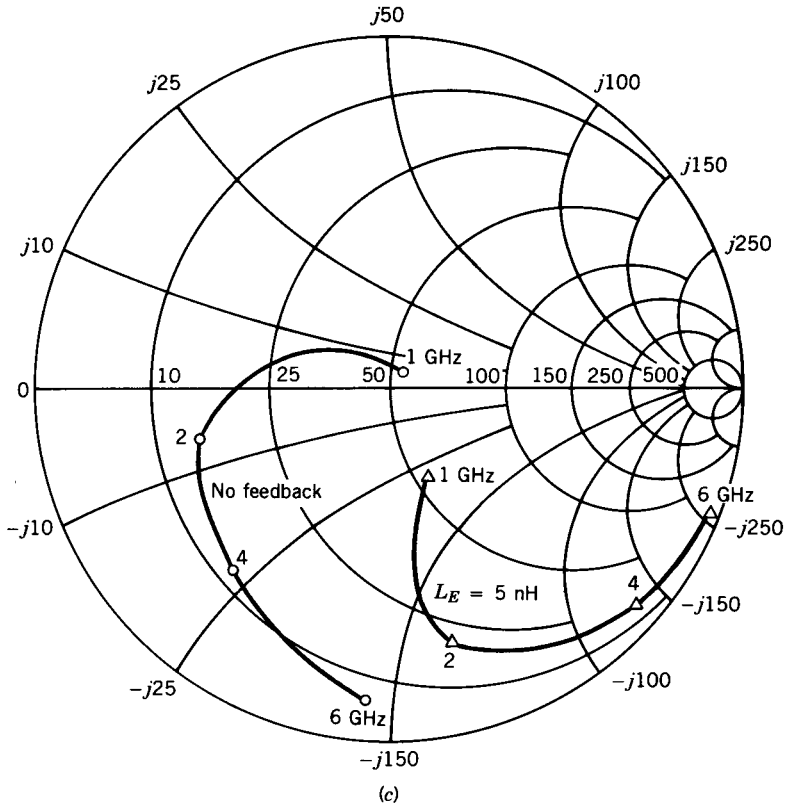


Figure 3-23 (*Continued*)

A final transformation to the more common noise-parameter format given by Eq. (3-44) is still needed [10]:

$$F_{\min} = 1 + 2R_n'(G_{\text{cor}}' + G_{0n}') \tag{3-66}$$

$$R_n = R_n' \tag{3-67}$$

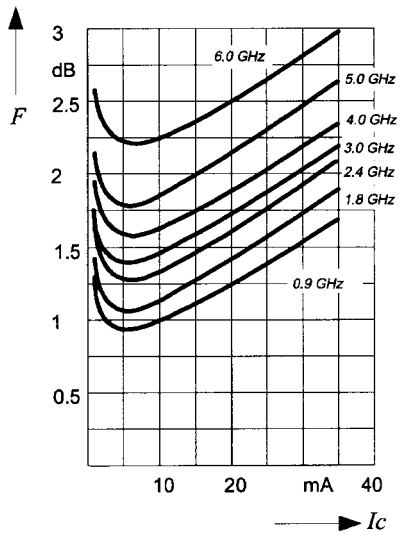
$$G_{0n} = \sqrt{\frac{G_n'}{R_n'} + G_{\text{cor}}'^2} \tag{3-68}$$

$$B_{0n} = -B_{\text{cor}}' \tag{3-69}$$

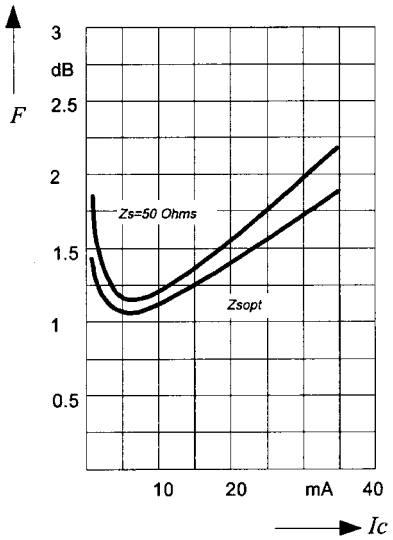
Figure 3-23 shows the noise figure as a function of external feedback for a low-noise microwave bipolar transistor, the AT-41435.

Bias-Dependent Noise Parameters. Since we have elaborated on the effect of parasitics, it is now useful to point out that noise is also a strong function of bias point and frequency. Figure 3-24 shows the noise figure expressed in dB as a function of different dc currents and frequencies for a BFP420 BJT by Siemens. A change in the collector voltage will also affect the noise parameters, not quite as strongly as shown here, but still significantly. The reason

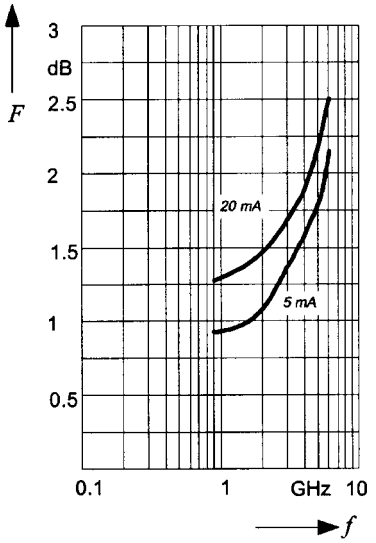
Noise Figure
versus Collector Current
V_{CE} = 2 V, $Z_S = Z_{Sopt}$



Noise Figure
versus Collector Current
VCE = 2 V, f = 1.8 GHz



Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA,
$Z_S = Z_{Sopt}$



Source Impedance for min.
Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA

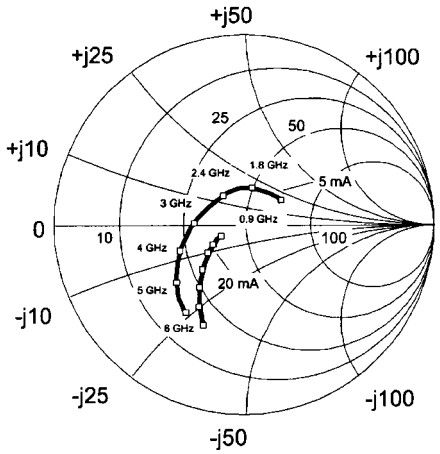


Figure 3-24 Noise figure versus collector current, noise figure versus frequency, and source impedance for minimum noise figure versus frequency for the Siemens BFP420 transistor.

Table 3-3 BFP420 common-emitter noise parameters

f	$F_{\min}{}^a$	$G_a{}^a$	Γ_{opt}		R_N	r_n	$F_{50\Omega}{}^b$	$\lvert S_{21} \rvert^{2b}$
(GHz)	dB	(dB)	(Magnitude)	(Angle)	(Ω)		(dB)	(dB)
V_{CE} = 2 V, I_C = 5 mA								
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	–175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	–157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	–142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	–123.0	15.0	0.30	3.30	5.3

[a]Input matched for minimum noise figure, output for maximum gain.
[b]$Z_S = Z_L = 50\ \Omega$.

for the turning point of the noise figure has to do with the fact that for lower currents the f_T cutoff frequency has not reached its peak yet, and therefore the noise increases. The increase of the noise at higher currents is due to the Schottky noise calculated from the equation

$$P_n = 2Iq \tag{3-70}$$

where P_n is the noise power, I is the saturation, collector, or drain current, and q is the charge of an electron (1.60×10^{-19} coulomb).

Table 3-3 shows the common-emitter noise parameters for the BFP420.

Noise Circles. From the section entitled "Influence of External Parasitic Elements," we see that the noise factor is a strong function of the generator admittance (or impedance) presented to the input terminals of the noisy two-port. Noise tuning is used to change the values of the input admittance to obtain the best noise performance. There is a range of values of input reflection coefficient over which the noise figure is constant. In plotting these points of constant noise figure, we obtain the so-called noise circles, which can be drawn on the Smith chart Γ_G plane [13]. Starting with the noise equation [see Eq. (3-44)] for a 50-Ω generator impedance, we find that

$$F_{50} = F_{\min} + 4r_n \frac{|\Gamma_{0n}|^2}{|1 + \Gamma_{0n}|^2} \tag{3-71}$$

We want to find the position of the reflection coefficient on the Smith chart, as in the case of the gain circles, for which F is constant. First we rearrange Eq. (3-67) to read

$$r_n = (F_{50} - F_{\min}) \frac{|1 + \Gamma_{0n}|^2}{4|\Gamma_{0n}|^2} \tag{3-72}$$

By introducing

$$N_i = \frac{F_i - F_{\min}}{4r_n} |1 + \Gamma_{0n}|^2 \tag{3-73}$$

we can find an expression for a circle of constant noise figure as introduced by Rothe and Dahlke [4] (see also Fukui [13]). The center for the noise circle is

$$C_i = \frac{|\Gamma_{0n}|}{1 + N_i} \tag{3-74}$$

and the radius is

$$r_i = \frac{\sqrt{N_i^2 + N_i(1 - |\Gamma_{0n}|^2)}}{1 + N_i} \tag{3-75}$$

with the definition of N used previously. However, if we consider only the minimum noise figure for a given device, we will not obtain the minimum noise figure for the multistage amplifier system. This was explained when the noise measure was introduced. (See Figure 3-25.) Therefore, a better way to design the amplifier would be to use circles at constant noise measure instead of circles of constant noise figure. This was recently done by Poole

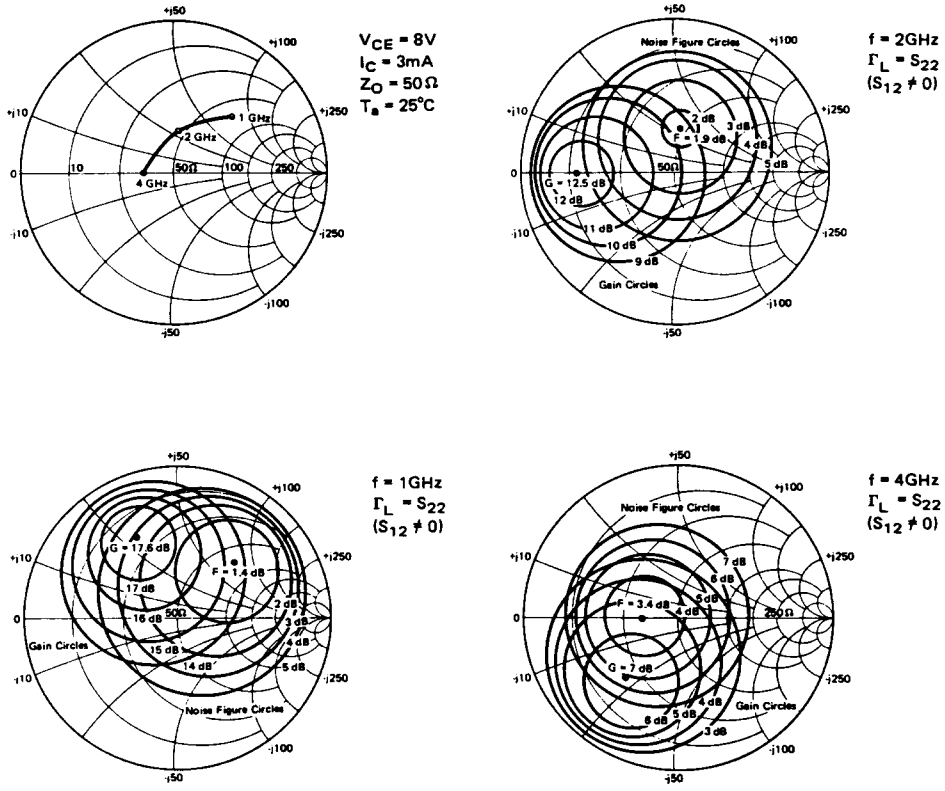


Figure 3-25 Typical noise figure circles and gain circles.

and Paul [14]. They derived the expressions for the noise measure circles as a function of S parameters, noise parameters, and Γ_G, using

$$G_A = \frac{|S_{21}|^2(1 - |\Gamma_G|^2)}{(1 - |S_{22}|^2) + |\Gamma_G|^2 (|S_{11}|^2 - |\Delta|^2) - 2\,\mathrm{Re}(\Gamma_G C_1)} \tag{3-76}$$

where

$$C_1 = S_{11} - S_{22}^* \Delta$$

$$\Delta = S_{11}S_{22} - S_{12}S_{11}$$

In terms of the generator reflection coefficient, the noise measure can be expressed as

$$M = [(F_{\min} - 1)(1 - |\Gamma_G|^2)|S_{21}|^2|1 + \Gamma_{0n}|^2 + 4r_n|S_{21}|^2|\Gamma_G - \Gamma_{0n}|^2]$$

$$\times \left\{|1 + \Gamma_{0n}|^2[|S_{21}|^2(1 - |\Gamma_G|^2) - (1 - |S_{22}|^2) - |\Gamma_G|^2(|S_{11}|^2 - |\Delta|^2) + 2\,\mathrm{Re}(\Gamma_G C_1)]\right\}^{-1} \tag{3-77}$$

Equation (3-77) can be shown to represent circles in the source reflection coefficient plane described by the following:

$$|\Gamma_G|^2 + |\Gamma_m|^2 - \Gamma_G\Gamma_m^* - \Gamma_G^*\Gamma_m = r_m^2 \tag{3-78}$$

The centers and radii of the constant-noise-measure circles are given by

$$C_m = \frac{M|1 + \Gamma_{0n}|^2 C_1^* + 4r_n|S_{21}|^2\Gamma_{0n}}{M|1 + \Gamma_{0n}|^2 P + |S_{21}|^2(4r_n - W)} \tag{3-79}$$

$$r_m = \frac{\sqrt{M^2M_a + MM_b + M_c}}{M|1 + \Gamma_{0n}|^2 P + |S_{21}|^2(4r_n - W)} \tag{3-80}$$

where

$$P = |S_{21}|^2 + |S_{11}|^2 - |\Delta|^2$$

$$Q = |S_{21}|^2 + |S_{22}|^2 - 1$$

$$W = |1 + \Gamma_{0n}|^2(F_{\min} - 1)$$

$$M_a = |1 + \Gamma_{0n}|^4(PQ + |C_1|^2)$$

$$M_b = |1 + \Gamma_{0n}|^2|S_{21}|^2[8r_n\,\mathrm{Re}(\Gamma_{0n}C_1)$$

$$- (4r_n|\Gamma_{0n}|^2 + W)P - (W - 4r_n)Q]$$

$$M_c = |S_{21}|^4[W - 4r_n(1 - |\Gamma_{0n}|^2)]$$

The value of the minimum noise measure can be found by considering the noise measure circle of zero radius; that is, set r_m equal to zero in Eq. (3-80). This results in

$$M_{\min} = \frac{-M_b \pm \sqrt{M_b^2 - 4M_a M_c}}{2M_a} \tag{3-81}$$

Equation (3-81) yields the same value of $M_{\min}$ as would have been obtained by using the immittance parameter equation given by Fukui [13] and can, therefore, be considered as the reflection coefficient plane analog of Fukui's expression. The elimination of the need for the parameter R_{eg}, however, results in a considerable simplification as compared with the earlier approach [13].

The value of the minimum noise measure is taken as the smallest nonnegative value of $M_{\min}$ given by Eq. (3-81). The source reflection coefficient that results in the minimum noise measure can now be obtained by employing Eq. (3-79):

$$\Gamma_{0m} = \frac{M_{\min}|1 + \Gamma_{0n}|^2 C_1^* + 4r_n|S_{21}|^2\Gamma_{0n}}{M_{\min}|1 + \Gamma_{0n}|^2 P + |S_{21}|^2\,(4r_n - W)} \tag{3-82}$$

The output reflection coefficient of the device, when Γ_{0m} is presented to the input port, is given by

$$S'_{22} = \frac{S_{22} - \Delta\Gamma_{0m}}{1 - S_{11}\Gamma_{0m}} \tag{3-83}$$

Noise Correlation in Linear Two-Ports Using Correlation Matrices. In the introduction to two-port noise theory, it was indicated that noise correlation matrices form a general technique for calculating noise in n-port networks. Haus and Adler [6] have described the theory behind this technique. In 1976, Hillbrand and Russer [15] published equations and transformations that aid in supplying this method to two-port CAD.

This method is useful because it forms a base from which we can rigorously calculate the noise of linear two-ports combined in arbitrary ways. For many representations, the method of combining the noise parameters is as simple as that for combining the circuit element matrices. In addition, noise correlation matrices can be used to calculate the noise in linear frequency conversion circuits. The following is an introduction to this subject.

A linear, noisy two-port can be modeled as a noise-free two-port with two additional noise sources. These noise sources must be chosen so that they add directly to the resulting vector of the representation, as shown in Eqs. (3-84) and (3-85) and Figure 3-17:

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \tag{3-84}$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{3-85}$$

where the i and v vectors indicate noise sources for the y and z representations, respectively. This two-port example can be extended to n-ports in a straightforward, obvious way.

Since the noise vector for any representation is a random variable, it is much more convenient to work with the noise correlation matrix. The correlation matrix gives us deterministic numbers with which to calculate. The correlation matrix is formed by taking the mean value of the outer product of the noise vector. This is equivalent to multiplying the noise vector by its adjoint (complex conjugate transpose) and averaging the result:

$$\langle \bar{i}\ \bar{i}^{+} \rangle = \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} [i_1^* \quad i_2^*] = \begin{bmatrix} \langle i_1 i_1^* \rangle & \langle i_1 i_2^* \rangle \\ \langle i_1^* i_2 \rangle & \langle i_2 i_2^* \rangle \end{bmatrix} + [C_y] \tag{3-86}$$

where the angular brackets denote the average value.

Note that the diagonal terms are the "power" spectrum of each noise source and the off-diagonal terms are complex conjugates of each other and represent the "cross-power" spectra of the noise sources. "Power" is used because these magnitude-squared quantities are proportional to power.

To use these correlation matrices in circuit analysis, we must know how to combine them and how to convert them between various representations. An example using y matrices will illustrate the method for combining two-ports and their correlation matrices. Given two matrices y and y', when we parallel them we have the same port voltages, and the terminal currents add (Figure 3-26):

$$\begin{aligned} I_1 &= y_{11}V_1 + y_{12}V_2 + y'_{11}V_1 + y'_{12}V_2 + i_1 + i'_1 \\ I_2 &= y_{21}V_1 + y_{22}V_2 + y'_{21}V_1 + y'_{22}V_2 + i_2 + i'_2 \end{aligned} \tag{3-87}$$

or

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} + y'_{11} & y_{12} + y'_{12} \\ y_{21} + y'_{21} & y_{22} + y'_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 + i'_1 \\ i_2 + i'_2 \end{bmatrix} \tag{3-88}$$

Here we can see that the noise current vectors add just as the y parameters add. Converting the new noise vector to a correlation matrix yields

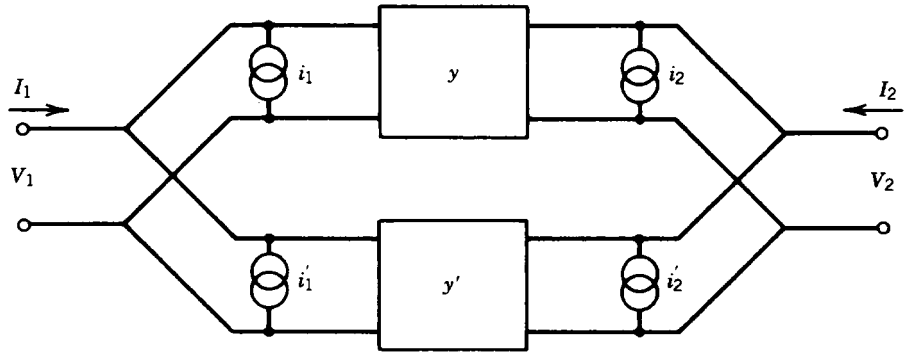


Figure 3-26 Parallel combination of two-ports using y parameters.

$$\langle \bar{i}_{\text{new}} \bar{i}^{+}_{\text{new}} \rangle = \left\langle \begin{bmatrix} i_1 + i_1' \\ i_2 + i_2' \end{bmatrix} [i_1^* + i_1'^* \quad i_2 i_2'^*] \right\rangle \tag{3-89}$$

$$= \begin{bmatrix} \langle i_1 i_1^* \rangle + \langle i_1' i_1'^* \rangle & \langle i_1 i_2^* \rangle + \langle i_1' i_2'^* \rangle \\ \langle i_2 i_1^* \rangle + \langle i_2' i_1'^* \rangle & \langle i_2 i_2^* \rangle + \langle i_2' i_2'^* \rangle \end{bmatrix} \tag{3-90}$$

The noise sources from different two-ports must be uncorrelated, so there are no cross products of different two-ports. By inspection, Eq. (3-90) is just the addition of the correlation matrices for the individual two-ports, so

$$[C_{y\text{new}}] = [C_y] + [C_y'] \tag{3-91}$$

The same holds true for g, h, and z parameters, but $ABCD$ parameters have the more complicated form shown below. If

$$[A_{\text{new}}] = [A][A'] \tag{3-92}$$

then

$$[C_{A\text{new}}] = \{C_A] + [A]\,[C_{A'}]\,[A]^{+} \tag{3-93}$$

The transformation of one representation to another is best illustrated by an example. Let us transform the correlation matrix for a Y representation to a Z representation. Starting with

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = [Y] \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \tag{3-94}$$

we can move the noise vector to the left side and invert y:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = [Y^{-1}] \begin{bmatrix} I_1 - i_1 \\ I_2 - i_2 \end{bmatrix} = [Y^{-1}] \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + [Y^{-1}] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-95}$$

Table 3-4 π Noise matrix $T_{\alpha\beta}$ transformation

		Original Form (α Form)					
		Y		Z		A	
Resulting Form (β Form)	Y	1	0	y_{11}	y_{12}	$-y_{11}$	1
		0	1	y_{21}	y_{22}	$-y_{21}$	0
	Z	z_{11}	z_{12}	1	0	1	$-z_{11}$
		z_{21}	z_{22}	0	1	0	$-z_{21}$
	A	0	A_{12}	1	$-A_{11}$	1	0
		1	A_{22}	0	$-A_{21}$	0	1

Since $(Y)^{-1} = (Z)$, we have

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = [Z] \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + [Z] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-96}$$

so

$$= [Z] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} = [T_{yz}] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-97}$$

where the signs of i_1 and i_2 are superfluous since they will cancel when the correlation matrix is formed. Here the transformation of the Y noise current vector to the Z noise voltage vector is done simply by multiplying by (Z). Other transformations are shown in Table 3-4.

To obtain the noise correlation matrix, we again form the mean of the outer product:

$$\langle vv^+ \rangle = \begin{bmatrix} \langle v_1 v_1^* \rangle & \langle v_1 v_2^* \rangle \\ \rangle v_1^* v_2 \rangle & \langle v_2 v_2^* \rangle \end{bmatrix} = [Z] \left\langle \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} [i_1^* \quad i_2^*] \right\rangle [Z]^+ \tag{3-98}$$

or

$$[C_z] = [Z]\,[C_y]\,[Z]^+ \tag{3-99}$$

where

$$v^+ = [i_1^* \quad i_2^*]\,[Z]^+ \tag{3-100}$$

This is called a congruence transformation. The key to all of these derivations is the construction of a correlation matrix from the noise vector, as shown in Eq. (3-90).

These correlation matrices may easily be derived from the circuit matrices of passive circuits with only thermal noise sources. For example,

$$[C_z] = 2kT\Delta f\,\mathrm{Re}([Z]) \tag{3-101}$$

and

$$[C_y] = 2kT\Delta f\,\mathrm{Re}([Y]) \tag{3-102}$$

The $2kT$ factor comes from the double-sided spectrum of thermal noise. The correlation matrix from the $ABCD$ matrix may be related to the noise figure, as shown by Hillbrand and Russer [15]. We have

$$F = 1 + \frac{\overline{Y}[C_a]\overline{Y}^+}{2kT\,\mathrm{Re}(Y_G)} \tag{3-103}$$

where

$$\overline{Y} = \begin{bmatrix} Y_G \\ 1 \end{bmatrix}$$

Expressing the noise factor in terms of the correlation matrix, we obtain a complete formula:

$$F = 1 + \frac{C_{22}^{A}(f) + 2\,\mathrm{Re}\{Y_g(f)\,C_{12}^{A}(f)\} + |Y_g(f)|^2\,C_{11}^{A}(f)}{2kT_0\,\mathrm{Re}\{Y_g(f)\}} \tag{3-104}$$

Once we transform this in the Y parameter form, we obtain the following equation:

$$F(f) = F_{\min}(f) + \frac{R_n(f)|Y_{\mathrm{opt}}(f) - Y_g(f)|^2}{\mathrm{Re}\{Y_g(f)\}} \tag{3-105}$$

It should be noted that all these values are frequency dependent, as expressed in this equation.

The *ABCD* correlation matrix can be written in terms of the noise figure parameters as (double-sided spectrum)

$$[C_a] = 2kT\begin{bmatrix} R_n & \dfrac{F_0 - 1}{2} - R_n Y_{0n}^{*} \\ \dfrac{F_0 - 1}{2} - R_n Y_{0n} & R_n|Y_{0n}|^2 \end{bmatrix} \tag{3-106}$$

The noise correlation matrix method forms an easy and rigorous technique for handling noise in networks. This technique allows us to calculate the total noise for complicated networks by combining the noise matrices of subcircuits. It should be remembered that although noise correlation matrices apply to n-port networks, noise figure calculations apply only to pairs of ports. The parameters of the C_a matrix can be used to give the noise parameters:

$$Y_{0n} = \sqrt{\frac{C_{ii*}}{C_{uu*}} - \left[\mathrm{Im}\left(\frac{C_{ui*}}{C_{uu*}}\right)\right]^2} + j\,\mathrm{Im}\left(\frac{C_{ui*}}{C_{uu*}}\right) \tag{3-107}$$

$$F_0 = 1 + \frac{C_{ui*} + C_{uu*}Y_{0n}^{*}}{kT} \tag{3-108}$$

$$R_u = C_{uu*} \tag{3-109}$$

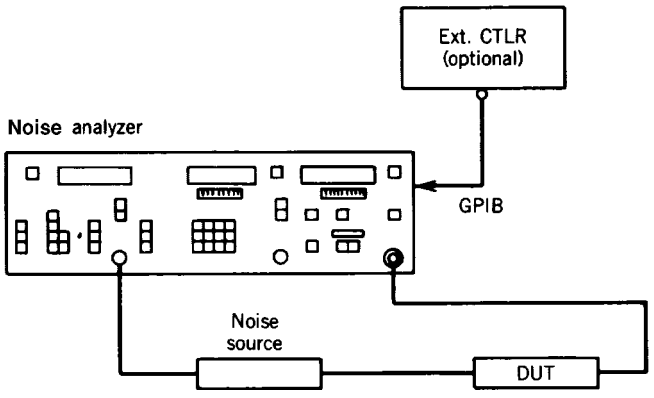


Figure 3-27 Noise figure measurement.

Noise Figure Test Equipment. Figure 3-27 shows the block diagram of a noise test setup. It includes the noise source and the other components. The metering unit has a special detector that is linear and over a certain dynamic range measures linear power. The tunable receiver covers a wide frequency range (e.g., 10–1800 MHz) and controls the noise source. The receiver is a double-conversion superheterodyne configuration with sufficient image rejection to avoid double-sideband noise measurements that would give the wrong results.

These receivers are microprocessor controlled and the measurement is a two-step procedure. The first is a calibration step that measures the noise figure of the receiver system and a reference power level. Then the device under test (DUT) is inserted and the system noise figure and total output power are measured. The noise factor is calculated by

$$F_1 = F_{\text{system}} - \frac{F_2 - 1}{G_1} \tag{3-110}$$

and the gain is given by the change in output power from the reference level [16]. The noise of the system is calculated by measuring the total noise power with the noise source on and off. With the excess noise ratio (ENR) known [16],

$$F_{\text{system}} = \frac{\text{ENR}}{Y - 1} \tag{3-111}$$

The noise bandwidth is usually set by the bandwidth of the receiver, which is assumed to be constant over the linear range. The ENR of the noise source is given by

$$\text{ENR} = \frac{T_{\text{hot}}}{T_{\text{cold}}} - 1 \tag{3-112}$$

where T_{cold} is usually room temperature (290 K). This ENR number is about 15 dB for noise sources with a 15-dB pad and 5 dB for noise sources with a 25-dB pad. Since both gain and noise were stored in the initial calibration, a noise/gain sweep can be performed.

For frequencies above 1800 MHz, we can extend the range with the help of the external signal generators, as shown in Figure 3-28. As shown, a filter ahead of the external mixer reduces the noise energy in the image band. If the DUT has a very broad frequency range

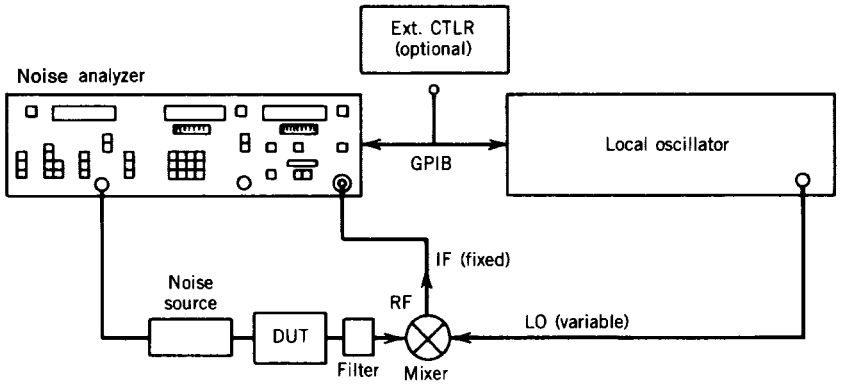


Figure 3-28 Single-sideband (SSB) noise figure measurements using an external mixer.

and has flat gain and noise over that range, a double-sideband (DSB) measurement is possible, with the image-rejection filter removed. However, a single-sideband (SSB) measurement is always more accurate [17].

How To Determine the Noise Parameters. The noise figure of a linear two-port network as a function of source admittance may be represented by

$$F = F_{\min} + \frac{R_n}{G_G}\left[(G_{0n} - G_G)^2 + (B_{0n} - B_G)^2\right] \tag{3-113}$$

where $G_G + jB_G$ = generator admittance presented to the input of the two-port
$G_{0n} + jB_{0n}$ = generator admittance at which optimum noise figure occurs
R_n = empirical constant relating sensitivity of the noise figure to generator admittance, with dimensions of resistance

It may be noted that for an arbitrary noise figure measurement with a known generator admittance, Eq. (3-113) has four unknowns—$F_{\min}$, R_n, G_{0n}, and B_{0n}. By choosing four known values of generator admittance, a set of four linear equations is formed and solutions of the four unknowns can be found [18, 19]. Eq. (3-113) may be transformed to

$$F = F_{\min} + \frac{R_n|Y_{0n}|^2}{G_G} - 2R_nG_{0n} + \frac{R_n|Y_G|^2}{G_G} - 2R_nB_{0n}\frac{B_G}{G_G} \tag{3-114}$$

or

$$F = F_{\min} + \frac{R_n}{G_G}|Y_G - Y_{0n}|^2 \tag{3-115}$$

Let

$$X_1 = F_{\min} - 2R_nG_{0n}$$

$$X_2 = R_n|Y_{0n}|^2$$

$$X_3 = R_n$$

$$X_4 = R_nB_{0n}$$

Then the generalized equation may be written

$$F_i = X_1 + \frac{1}{G_{si}}X_2 + \frac{|Y_{si}|^2}{G_{si}}X_3 - 2\frac{G_{si}}{B_{si}}X_4 \tag{3-116}$$

or, in matrix form,

$$[F] = [A]\,[X] \tag{3-117}$$

and the solution becomes

$$[X] = [A]^{-1}\,[F] \tag{3-118}$$

These parameters completely characterize the noise behavior of the linear two-port network. Direct measurement of these noise parameters by this method would be possible only if the receiver on the output of the two-port were noiseless and insensitive to its input admittance. In actual practice, the receiver itself behaves as a noisy two-port network and can be characterized in the same manner. What is actually being measured is the system noise figure of the two-port and the receiver. Thus, it becomes apparent that to do a complete two-port noise characterization, the gain of the two-port must be measured [20]. In addition, any losses in the input-matching networks must be carefully accounted for, because they add directly to the measured noise figure reading [21].

3-1-4 Linearity

Consistent with the elaborate discussion of linearity questions above, here is a quick refresher on what must be considered in amplifiers.

Dynamic Range, Compression, and IMD. An amplifier's dynamic range is related to the minimum discernible signal and the 1-dB compression point according to the equation

$$DR_n = \frac{(n-1)(IP_{n(\text{in})} - MDS)}{n} \tag{3-119}$$

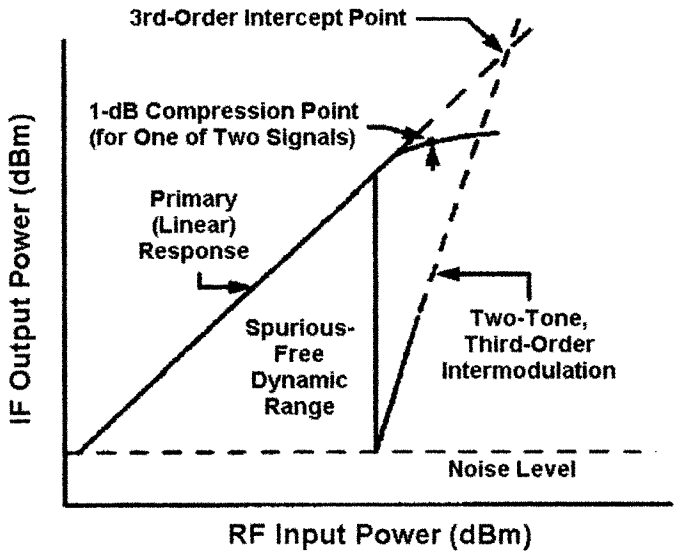


Figure 3-29 Amplifier linearity evaluation, including compression and two-tone IMD dynamic range. $P_{-1\text{dB}}$ for a single-tone cannot be read directly from this graph because the values shown are the result of two equal-power tones.

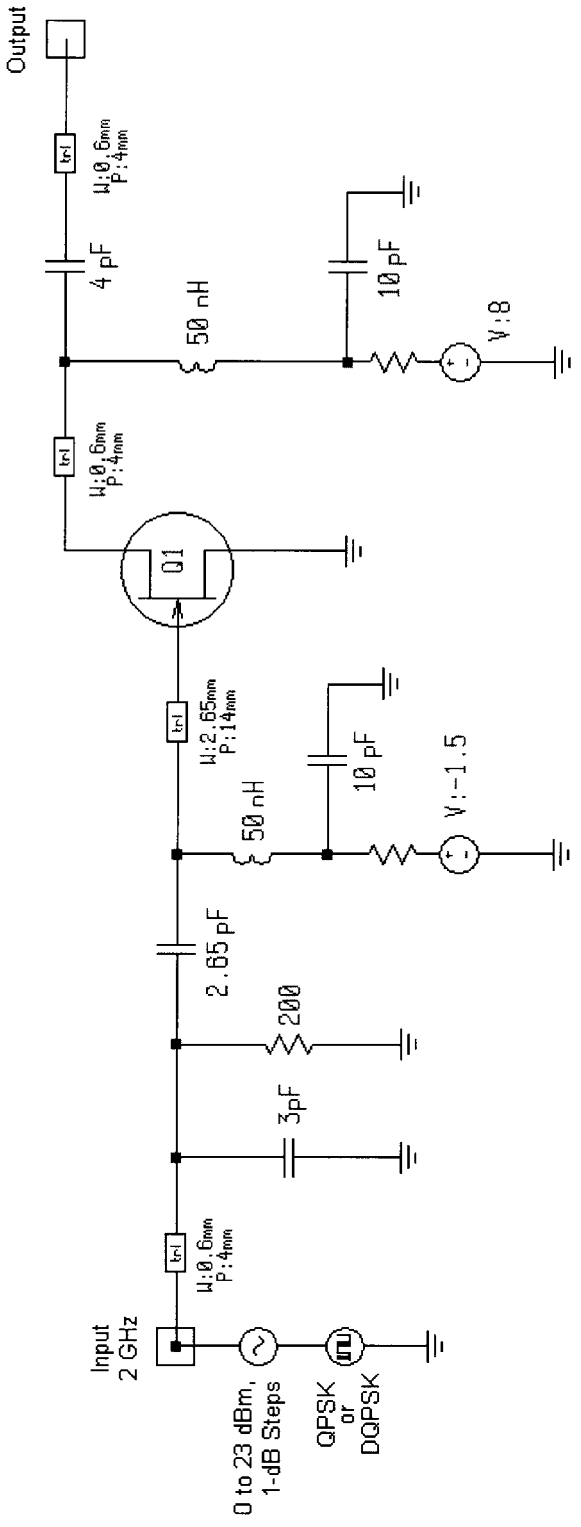


Figure 3-30 The power amplifier schematic.

where DR is the dynamic range in decibels, n is the order, $IP_{(in)}$ is the input intercept power in dBm, and MDS is the minimum detectable signal power in dBm. Once the 1-dB compression point is known, it is a fair assumption that the third-order intercept point, expressed in dBm, is 10 dB above this.

The intermodulation distortion gives the quality of an amplifier to withstand multiple signals without generating large intermodulation products or when using amplitude-modulated signals (cross-modulation). Figure 3-29 provides a quick reminder of IMD issues.

In some amplifiers, we would like to have control over the gain. While the better way to control gain while obtaining high linearity is a PIN-diode attenuator, special AGC circuits have become part of the amplifiers. Figure 3-48 shows a circuit of the SL610, made by Plessey, which has an AGC input provision. The noise figure of the PIN-diode attenuator is essentially equal to its attenuation. The AGC in multistage amplifiers is distributed over several stages and there is a lesser correlation between noise and AGC than in purely passive attenuators.

Special Case of Linearity Requirements for Digital Modulation. For cellular telephone systems we use digital modulation formats, such as QPSK and π/4-DQPSK, that combine phase and amplitude modulation. Amplifiers handling such signals must be carefully characterized and designed if adequate amplitude and phase linearity are to be maintained.

Examples of Power Amplifiers: Looking into the Effects of Distortion. This example presents modulation analysis using QPSK in a 2-GHz power amplifier (Figure 3-30). The simple amplifier was designed using the electronic Smith Tool of Serenade to determine the matching network for a narrowband design. The example demonstrates setup of modulation analysis and available results.

Siemens among others provides integrated amplifiers based on the concept depicted in Figure 3-30, the most popular ones being the CGY96 and CGY94. For the purpose of measurements, these amplifiers were actually put on a breadboard. Figure 3-31 shows the big brother, the CGY96. It is a GaAs MMIC intended as a power amplifier for GSM Class 4 phones with 3.2 W (35 dBm) at 3.5 V at an overall power added efficiency of 50%. Table 3-5 shows its electrical characteristics.

For the purpose of our experiments, we considered the CGY94 because it requires less external peripheral circuits. Table 3-6 shows its electrical characteristics. For those of us interested in seeing part of the internal circuit, Figure 3-32 shows the chip. This two-stage amplifier uses an interesting internal feedback scheme. Its schematic is shown in Figure 3-33.

For the purpose of looking at the waveforms, we will continue with the single-stage amplifier (Figure 3-30) and then compare the measured versus predicted performance of this CGY94-based amplifier.

A modulation source is connected to the RF source at the input port of the amplifier. The RF source specifies the carrier power (or voltage or current) and the modulation source specifies the modulation format and properties of the modulated signal.

A brief review of the modulation source used in this example indicates the following properties:

Property	Value	Description
NB	128	Number of bits
BR	1.2288E6	Bit rate
N	8	Number of samples per bit
BW1	590.E3	One-sided bandwidth of the main channel in the PIB or P1IB calculation
FS1	740.E3	Start baseband frequency of the fist adjacent channel
BW2	590.E3	One-sided bandwidth of the main channel in the P2IB calculation
FS2	1990.E3	Start baseband frequency of the second adjacent channel
BW3	590.E3	One-sided bandwidth of the main channel in the P3IB calculation
FS3	3240.E3	Start baseband frequency of the third adjacent channel
M	4	Order of the signal space; M = 4 is QPSK
DLY	0.0	Fractional bit delay
IASC	1.0	I channel amplitude scale; multiplier for the I waveform to model amplitude imbalance
QASC	1.0	Q channel amplitude scale; multiplier for the Q waveform to model amplitude imbalance
LPF	3	Baseband low-pass filter; 3 = root-raised cosine filter
LPFC	665.E3	–3 dB cutoff frequency of LPF
LPFN	3	Number of resonators if LPF = 1
LPFR	0.35	Roll-off factor if LPF = 3 or LPF = 4
FILE	N/A	File name of user-defined modulated signal file

The modulation source is a QPSK stream with a bit rate of 1.2288 Mbps. A total of 128 bits will be analyzed, and each bit is sampled eight times to construct an accurate analog waveform. The BW and FS parameters are defined here to specify the main (in-band) channel and adjacent channel for ACPR calculations. Imbalance of the I and Q amplitudes and phases can be described by the IASC, QASC, and DLY parameters, if desired. The modulation signal

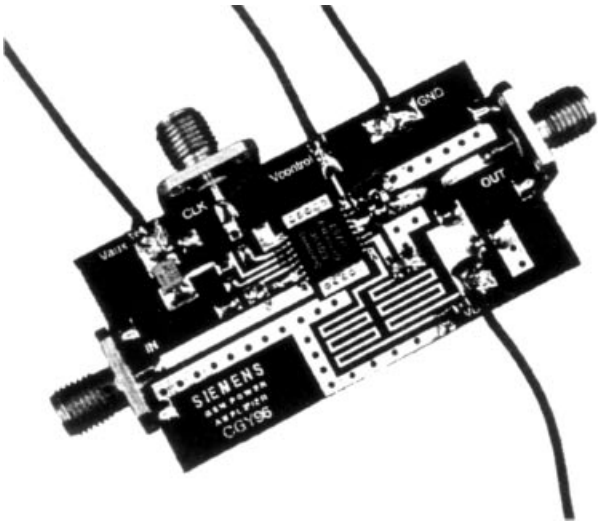


Figure 3-31 CGY96 GaAs MMIC power amplifier.

Table 3-5 CGY96 electrical characteristics

(T_A = 25 °C, V_{neg} = 5 V, $V_{control}$ = 2.2 V; duty cycle 12.5%, t_{on} = 577 µs)

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
Frequency range	f	880	—	915	MHz
Supply current P_{in} = 0 dBm	I_D	—	1.8	—	A
Supply current negative voltage generator V_{aux} = 3.5 V	I_{AUX}	—	10	—	mA
Gain (small signal)	G	—	40	—	dB
Power gain P_{in} = 0 dBm	G_P	—	35	—	dB
Output power P_{in} = 0 dBm, $V_{control}$ = 2.0–2.5 V	P_{OUT}	—	35	—	dBm
Overall power added efficiency P_{in} = 0 dBm	η	—	50	—	%
Dynamic range output power $V_{control}$ = 0.2–2.2 V		—	80	—	dB
Harmonics P_{in} = 0 dBm	$H(2f_0)$	—	–40	—	dBc
	$H(3f_0)$	—	–43	—	dBc
	$H(4f_0)$	—	–44	—	dBc
Noise power in RX (935–960 MHz) P_{in} = 0 dBm, P_{out} = 35 dBm, 100 kHz RBW	N_{RX}	—	–81	—	dBm
Stability *all spurious outputs* < –60 dBc, *VSWR load, all phase angles*		—	10:1	—	—
Input VSWR		—	1.7:1	—	—

Table 3-6 CGY94 electrical characteristics

T_A = 25 °C, f = 0.9 GHz, $Z_S = Z_L$ = 50 Ω, V_D = 3.6 V, V_G = 4 V, VTR pin connected to ground, unless otherwise specified; pulsed with a duty cycle of 10%, ton = 0.33 ms)

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
Supply current V_D = 3.0 V P_{in} = 10 dBm	I_{DD}	—	1.18	—	A
Negative supply current *(normal operation)*	I_G	—	2	—	mA
Shut-off current *VTR n.c.*	I_D	—	400	—	µA
Negative supply current (*shut off mode, VTR pin n.c.*)	I_G	—	10	—	µA
Gain P_{in} = *5* dBm	G	27.0	29.0	—	dB
Power gain V_D = 3.6 V; P_{in} = 10 dBm	G	22.8	23.6	—	dB
Output power V_D = 3.0 V*;* P_{in} = 10 dBm	P_O	31.5	32.3	—	dBm
Output power V_D = 3.6 V*;* P_{in} = 10 dBm	P_O	32.8	33.6	—	dBm
Output power V_D = 5 *V;* P_{in} = 10 dBm	P_O	34.5	35.5	—	dBm
Overall power added efficiency V_D = 3.0 V; P_{in} = 10 dBm	η	43	48	—	%
Overall power added efficiency V_D = 3.6 V; P_{in} = 10 dBm	η	42	47	—	%
Overall power added efficiency V_D = 5 V; P_{in} = 10 dBm	η	41	46	—	%
Harmonics (P_{in} = 10 dBm, *CW*) $2f_0$	—	—	–49	—	dBc
V_D = 3.6 V; (P_{out} = 33.1 dBm) $3f_0$	—	—	–45	—	dBc
Input VSWR V_D = 3.6 V	—	—	1.5:1	2.0:1	—

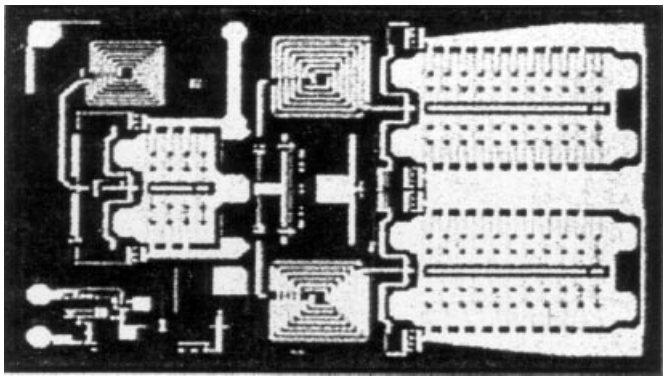


Figure 3-32 The CGY94 amplifier chip.

is often filtered, and several types of filters are available. Here a root-raised cosine filter is used with a cutoff frequency of 665 kHz and a roll-off factor of 0.35.

Analysis. Figure 3-34 shows the compression characteristics of the amplifier as the available RF source power (a sinusoid) is swept. Note that P_{-1dB} is 11.5 dBm referred to the input and 28.5 dBm referred to the output. We will look at modulation characteristics in the linear region of operation at a source power of 0 dBm, and the nonlinear region at a source power of 20 dBm.

Next, we will view the voltage across the transistor to examine its behavior under compression. Figure 3-35 shows the drain–source voltage of the FET as the source power is swept. It is clear from this graph that clipping due to pinchoff and forward-gate conduction

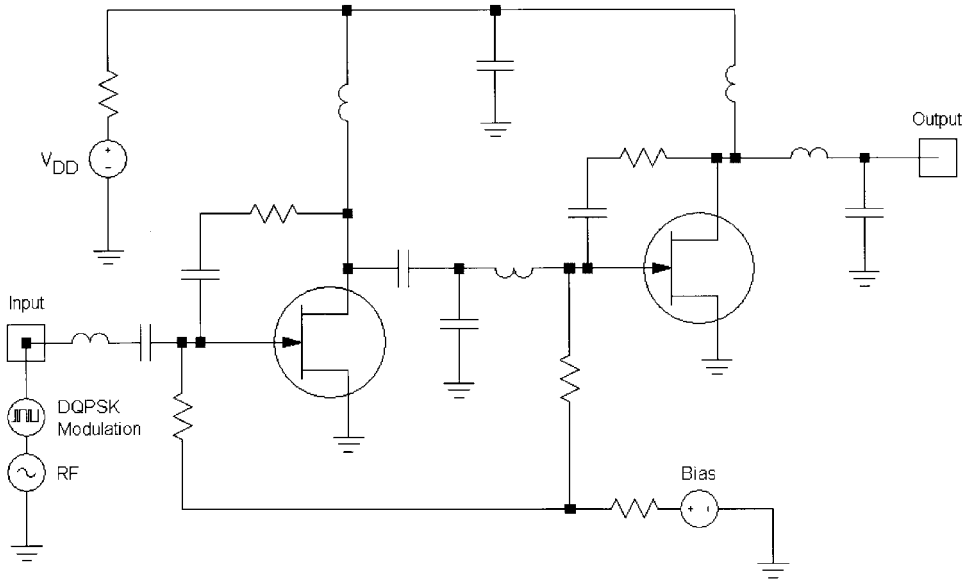


Figure 3-33 Schematic of the CGY94 GaAs MMIC power amplifier.

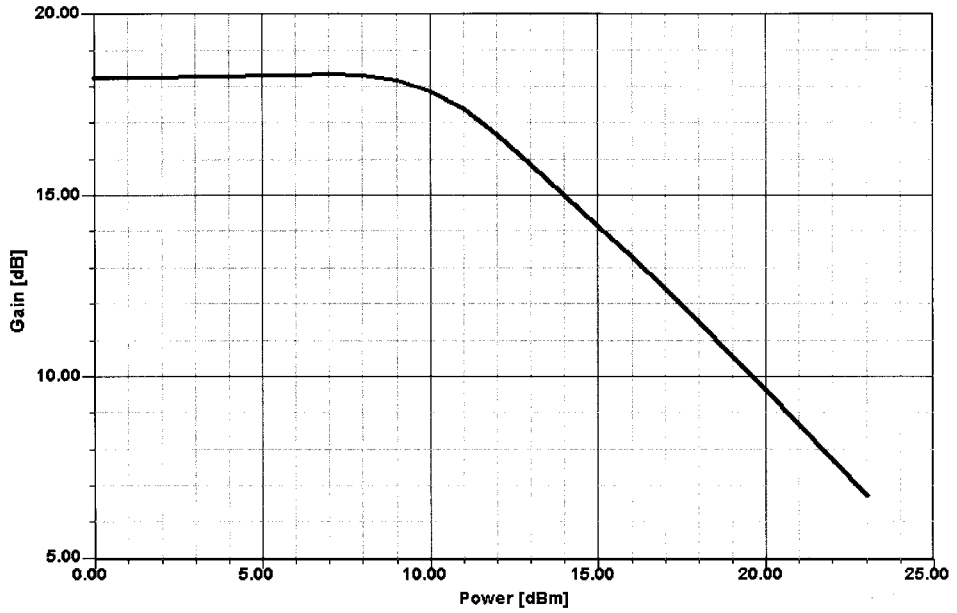


Figure 3-34 Single-tone RF power sweep analysis of the FET amplifier.

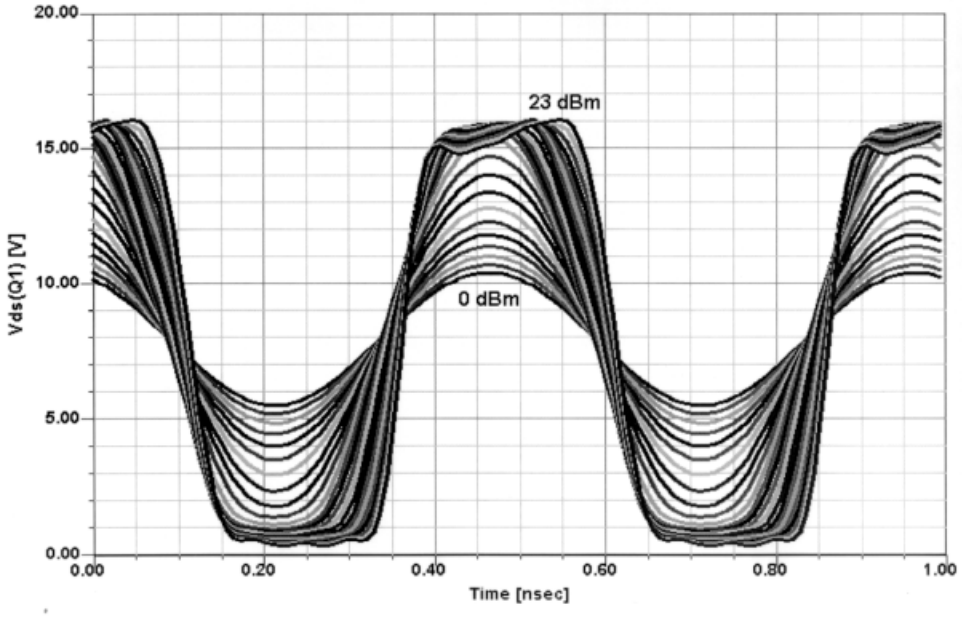


Figure 3-35 V_{DS} of the FET versus RF source power.

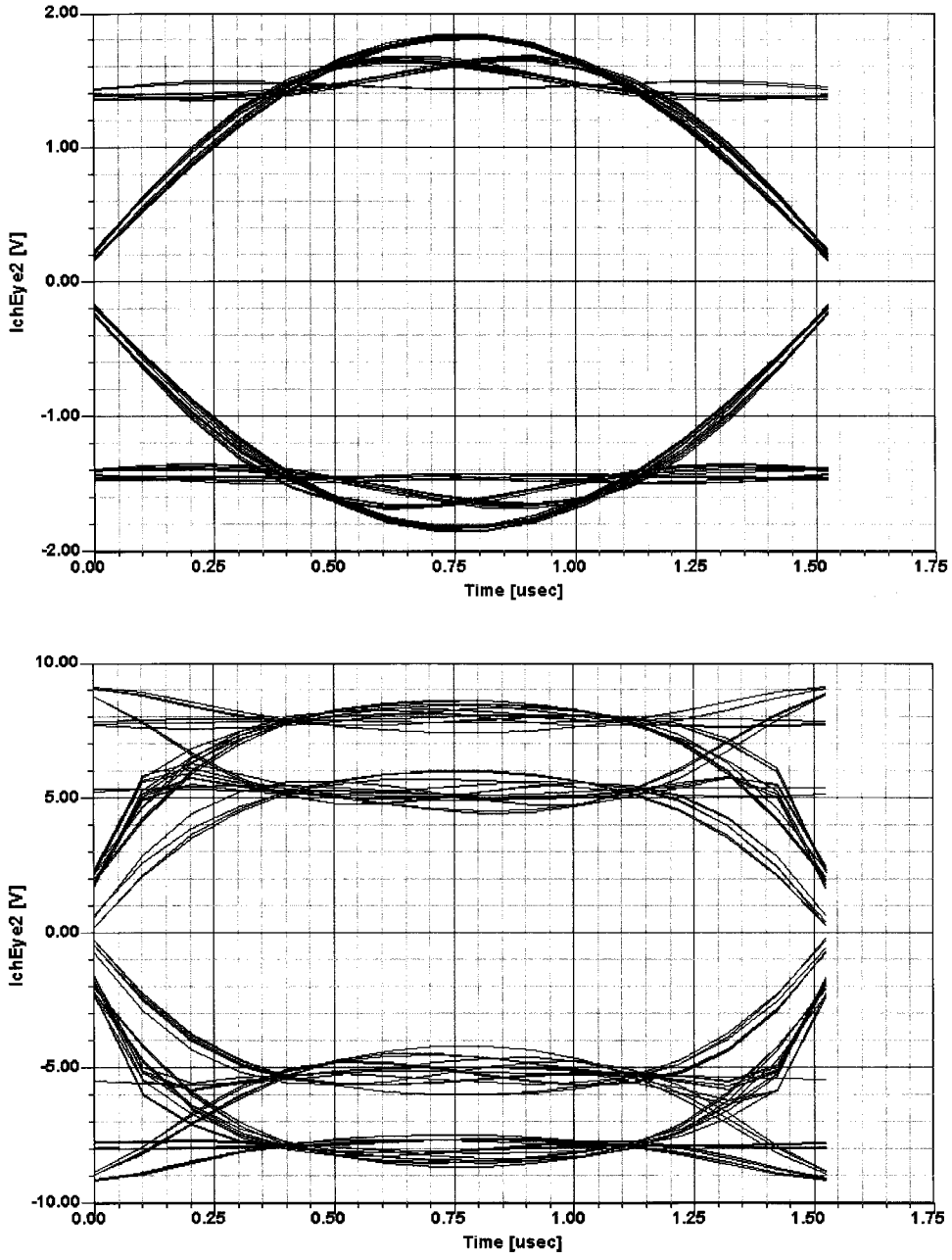


Figure 3-36 Eye diagrams of the I channel for source powers of (a) 0 dBm and (b) 20 dBm.

is limiting the performance of the device. A phase shift of the voltage waveform as power is increased is also apparent. The combination of the power compression and phase shift, or AM to PM conversion, will be used in the modulation analysis to determine the overall distortion of the modulated signal.

The output available from modulation analysis includes:

- I and Q channel waveforms
- Eye and constellation diagrams
- Spectral plots
- ACPR, in-band, and adjacent power

Figure 3-36 shows the eye diagrams for the 128-bit QPSK signal at two RF source powers, 0 and 20 dBm. Operating linearly at 0 dBm input, the amplifier doesn't distort the eye and it remains wide open. Note that in this example the filter bandwidth doesn't produce any intersymbol interference, as evidenced by the open eye. At an input power of 20 dBm, the amplifier compresses the signal, and distortion in the eye is evident.

We can also view the modulation spectrum at these two power levels to investigate the intermodulation distortion and spectral regrowth that takes place. Figure 3-37 shows the spectral plots. The lower trace corresponds to 0 dBm source power and shows almost no regrowth, while the upper trace at 20 dBm source power shows considerable regrowth.

The uncorrupted modulation source can also be shown. This corresponds to the source as a pure voltage without any circuit influence.

Now we evaluate the CGY94 board in a similar fashion. Figure 3-38 shows the waveform used for CAD purposes. Needless to say, such a signal cannot be produced by any signal

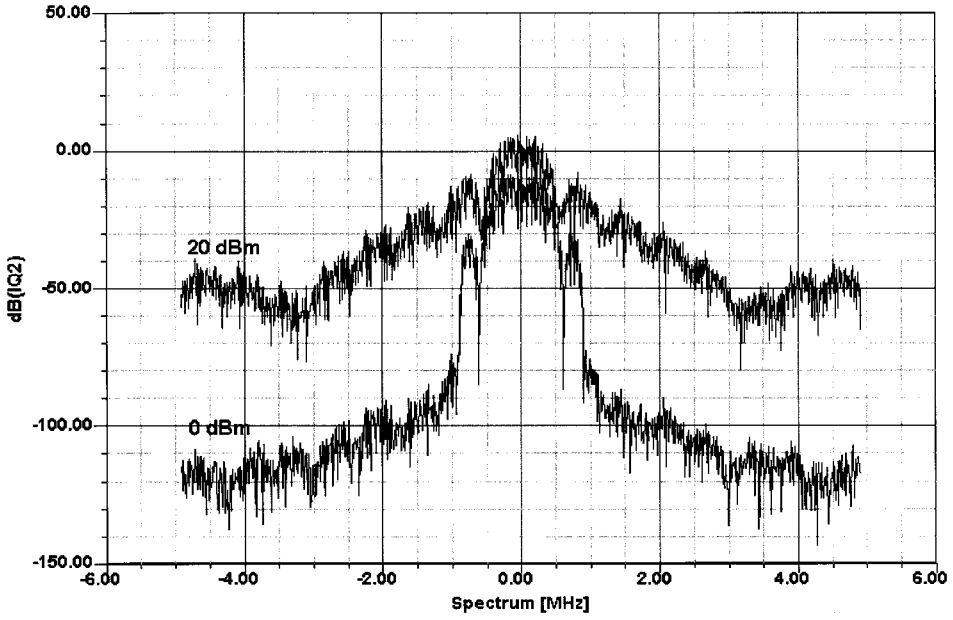


Figure 3-37 Spectrum of the modulation signal at the amplifier output. Source powers are 0 and 20 dBm.

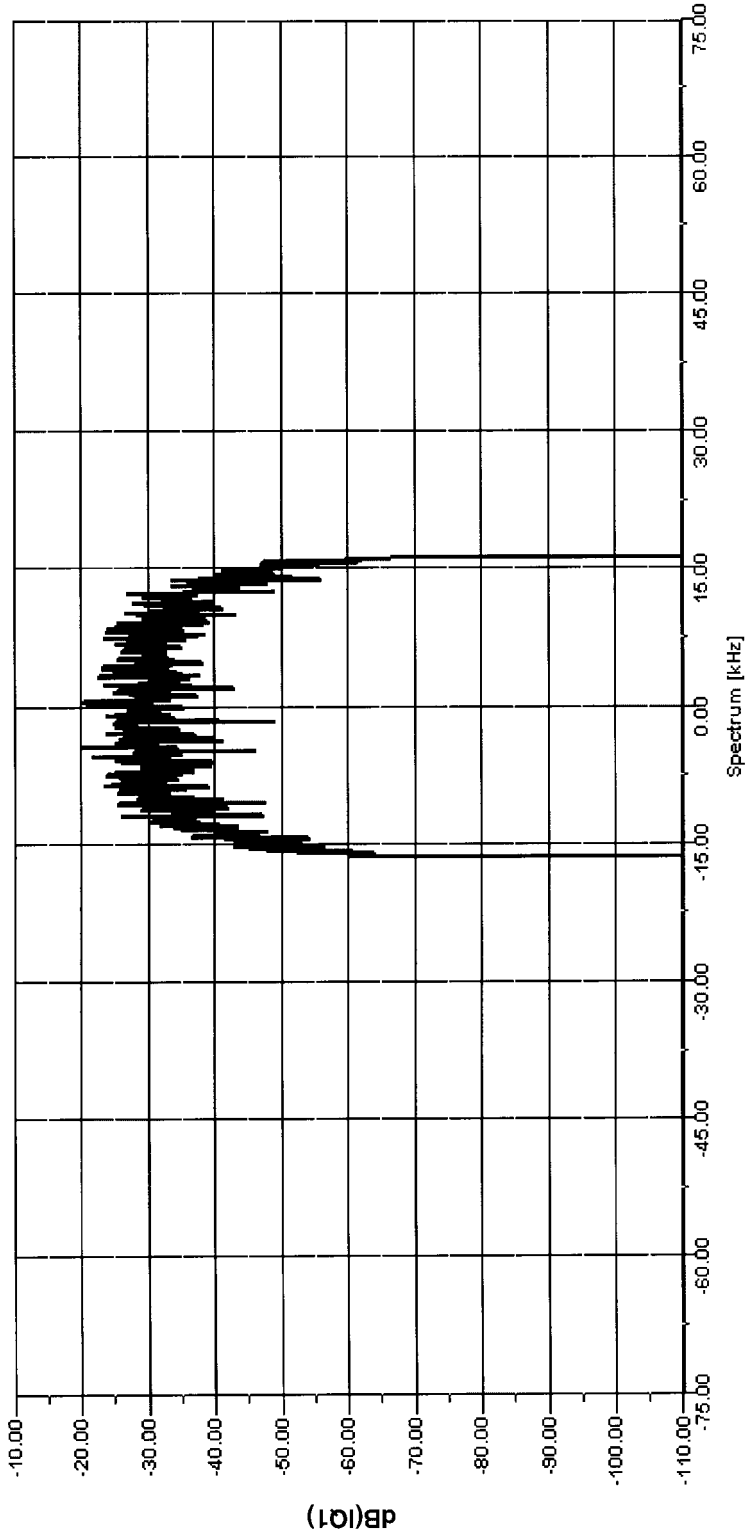


Figure 3-38 Simulated signal used for CAD analysis of the CGY94 amplifier.

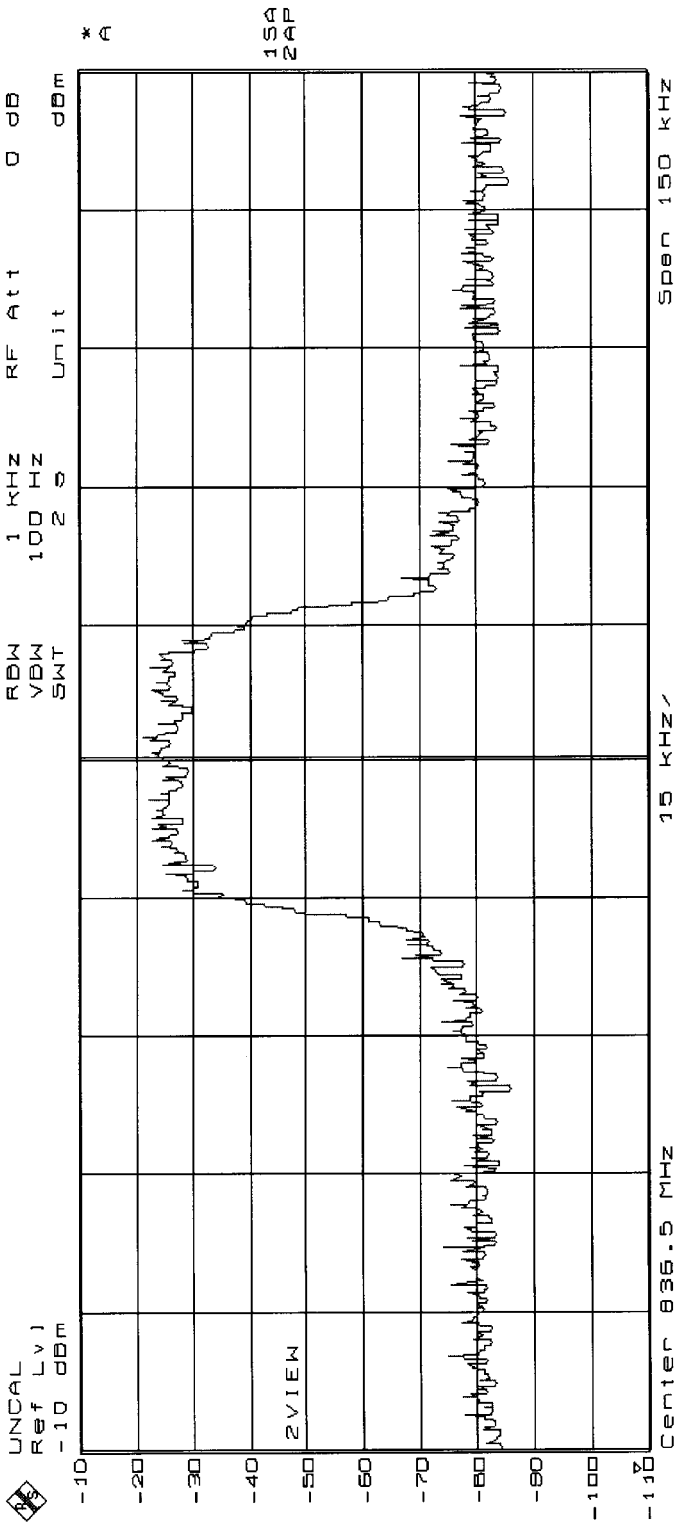


Figure 3-39 Real signal used for testing the actual amplifier.

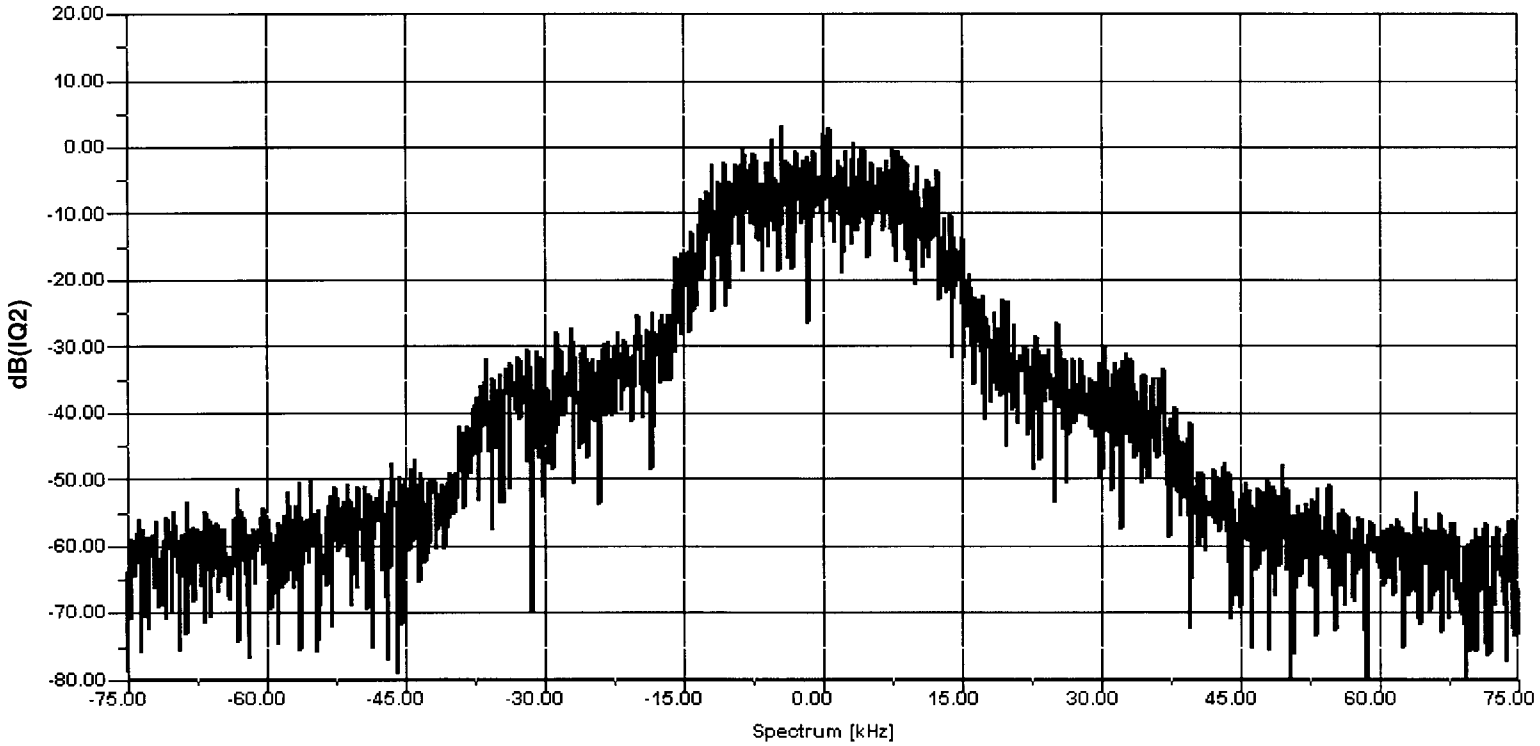


Figure 3-40 Spectral regrowth predicted by the simulation.

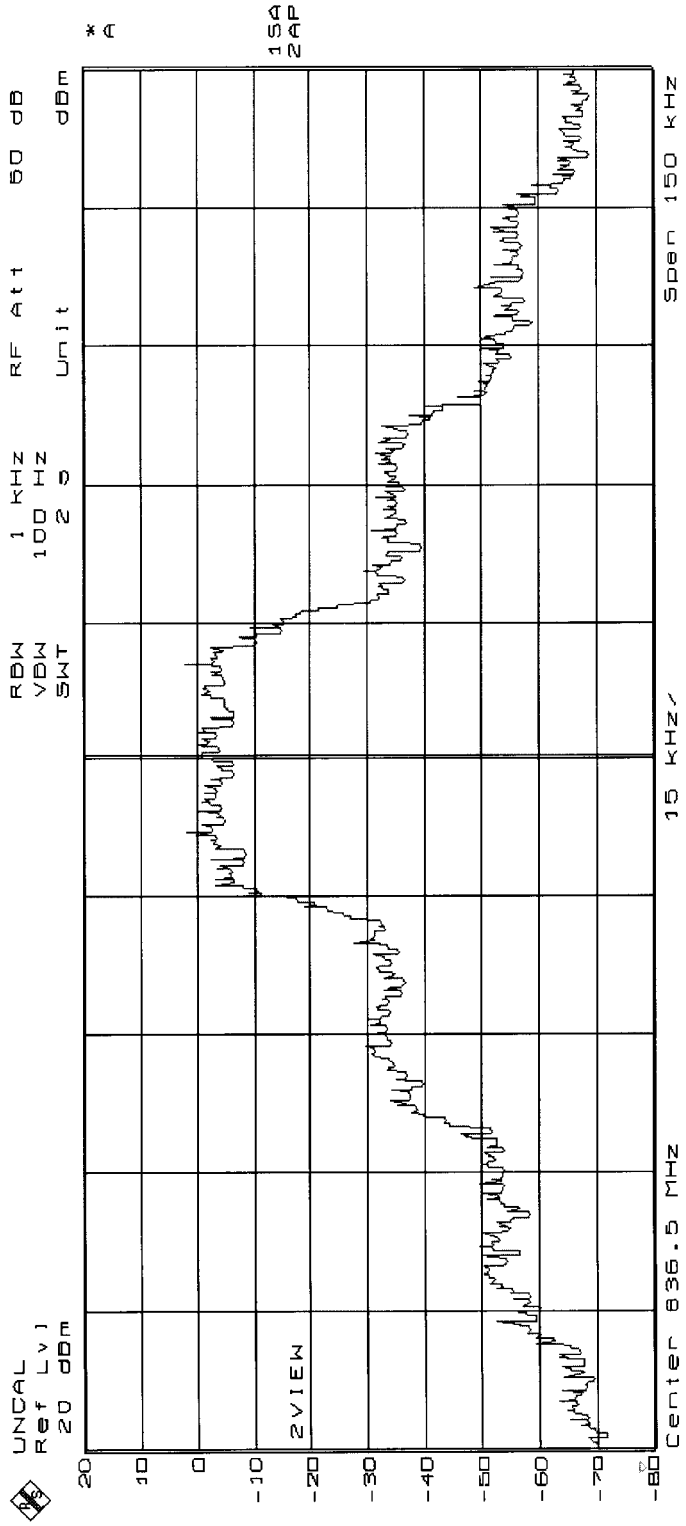


Figure 3-41 Measured spectral regrowth of the CGY94 amplifier.

generator. The one generated by an actual signal generator is shown in Figure 3-39. The wideband noise is around –80 dBm and the input level is about –25 dBm, consistent with the simulation signal. Looking at the output, Figure 3-40 shows the simulated signal with the expected increase in bandwidth. This is due to the nonlinearity of the amplifier and is consistent with Figure 3-37 for the single-stage amplifier. The actual measured output for the CGY94 shown in Figure 3-41 is quite close to its simulated output, indicating that the mathematical approach used for simulating is correct. This is only one example; in reality, many more waveforms can be analyzed and predicted using this method. Figure 3-41 shows essentially three distinct steps, which are sufficiently close to the response shown in the simulation. A dynamic range of 70 dB, while displayed in Figures 3-40 and 3-41, is not really necessary for good-quality transmission, and the discrepancy between the two pictures is in part due to the limited modeling quality of the active device as provided by the company that did the parameter extraction for this transistor.

Another useful view to aid in the understanding of the compression and distortion within the amplifier is to look at the magnitude of the complex waveform. Since the signal is composed of the in-phase and quadrature-phase components it is represented as

$$s(t) = i(t) + j \cdot q(t) \tag{3-120}$$

The time-domain magnitude is written

$$|s(t)| = \sqrt{i(t)^2 + q(t)^2} \tag{3-121}$$

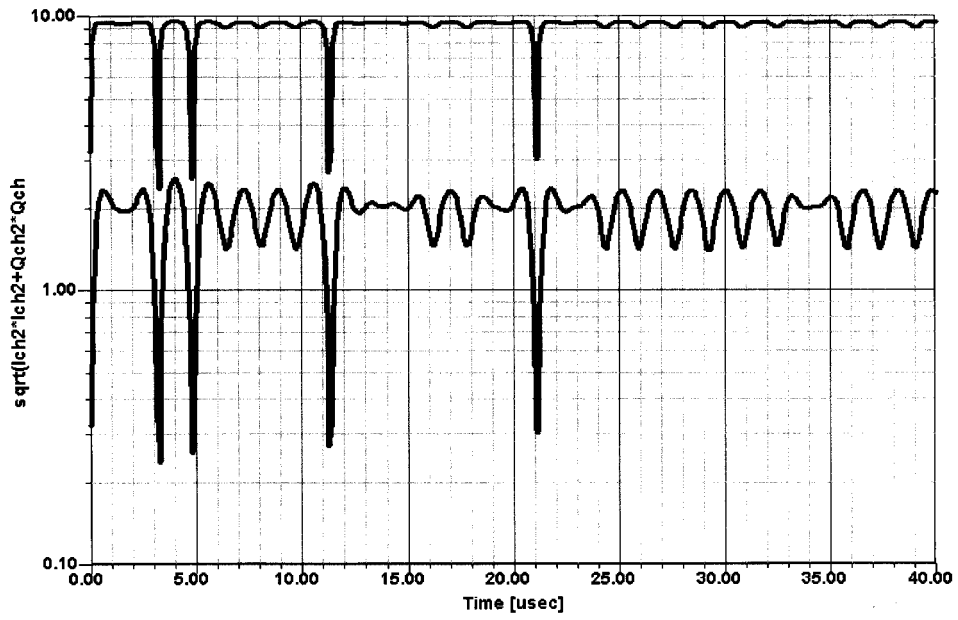


Figure 3-42 Time-domain magnitude of the complex modulation signal at source powers of 0 and 20 dBm.

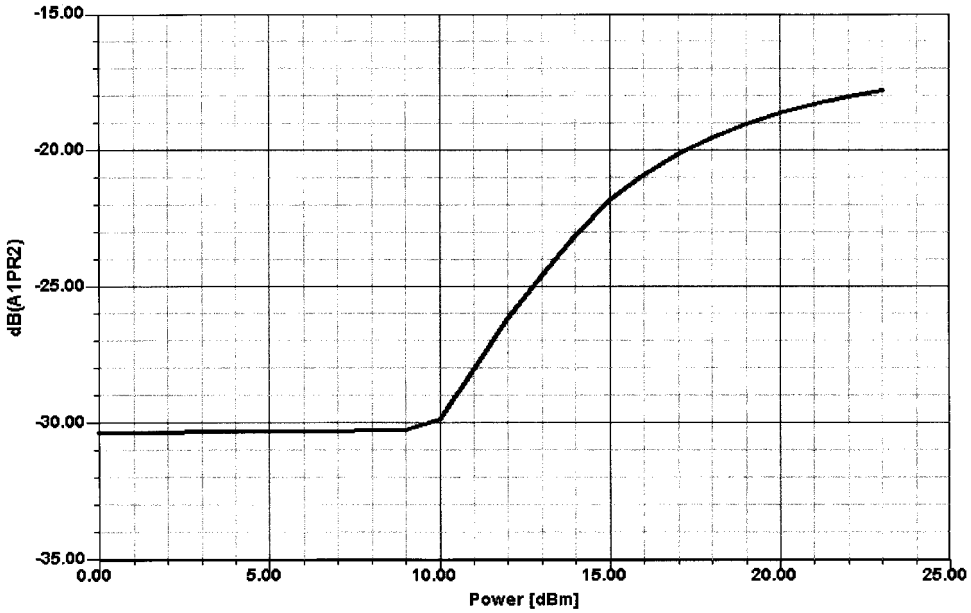


Figure 3-43 Adjacent-channel power ratio as function of RF source power.

Figure 3-42 shows the magnitude at source powers of 0 and 20 dBm. The *X* axis has been rescaled to 40 μs for improved viewing. Note the significant compression of the 20-dBm waveform at the higher signal levels. It is clear from this view that the signal is severely distorted.

We can also compute the adjacent-channel power ratio (ACPR) against the RF power sweep. ACPR is the ratio of the adjacent-channel power to the in-band channel power. The bandwidth (BW) and adjacent-channel start frequency (FS) are used for this calculation. For accurate ACPR calculation, a large number of bits are needed. This example uses 128 bits so the computation time is short, but 512 or more bits should be used for a more accurate computation.

Figure 3-43 shows the ACPR as a function of RF source power. The ACPR is nearly constant (and nonzero due to the gradual skirt of the baseband filter and inherent spillover to the adjacent channel) up to the P_{-1dB} compression point . As the intermodulation products spill into the adjacent channel and spectral regrowth occurs, the ACPR degrades.

π/4-DQPSK Circuit Analysis. Next, we will examine the same amplifier using a π/4-DQPSK modulation source. The task is identical to the previous one except for the source and the number of bits. The number of bits for this project has been increased from 128 to 256. Figure 3-44 shows the eye diagrams from this modulation format, at source powers of 0 and 20 dBm. The distortion and related intersymbol interference are clearly evident.

The constellation plot at 0 dBm is shown in Figure 3-45.

It may be interesting to take a look at the different waveforms currently in use. Figure 3-46 shows a spectrum analyzer picture of the most popular ones.

For reasons of time-division management, the European standard, GSM, which is now offered by Omnipoint at the 1.8-GHz band in the United States, uses gating to reduce the actual bandwidth of the spectrum. Figure 3-47 compares the GSM signal with and without gating.

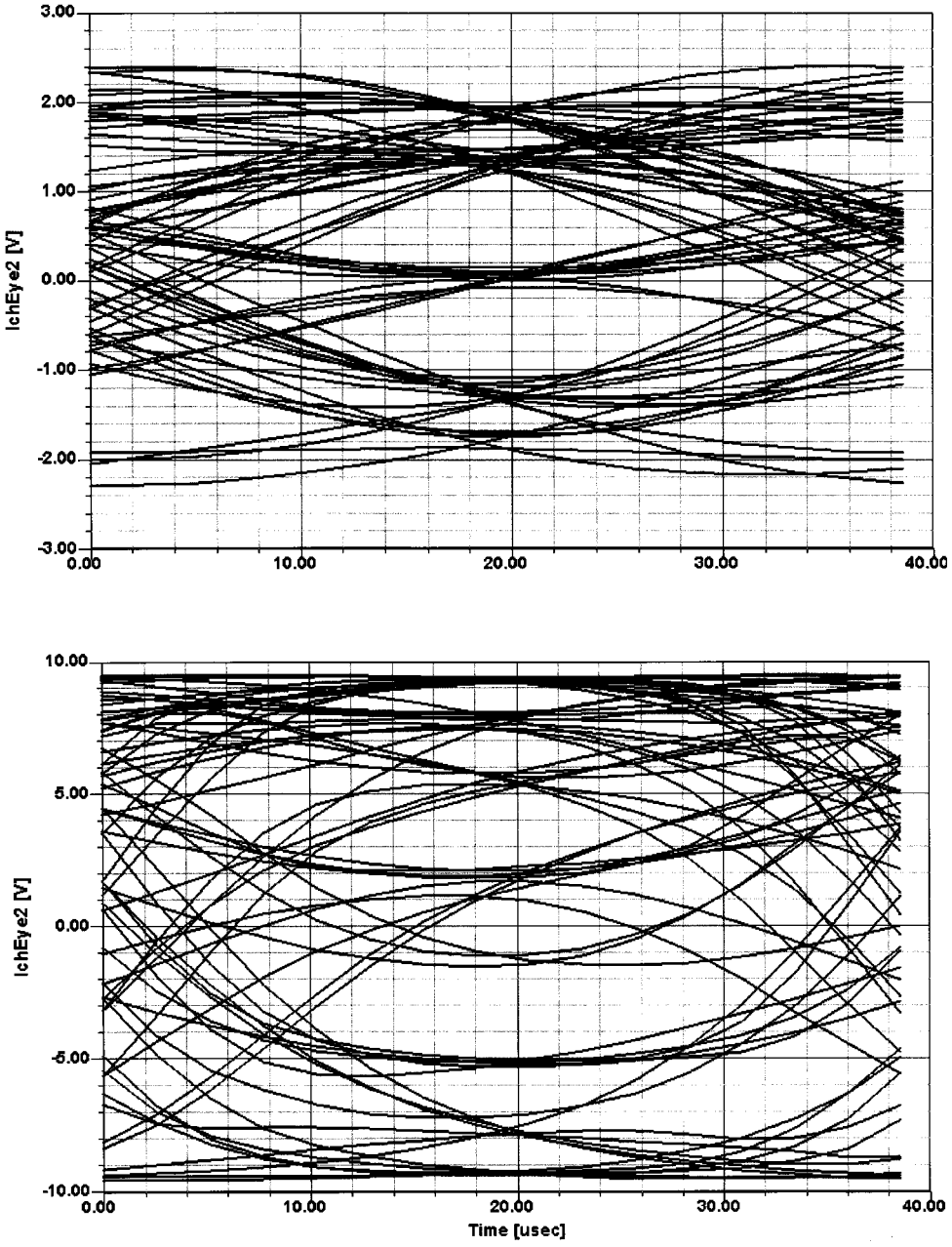


Figure 3-44 Eye diagrams for $\pi/4$-DQPSK at source powers of (a) 0 dBm and (b) 20 dBm. The distortion and related intersymbol interference are clearly visible.

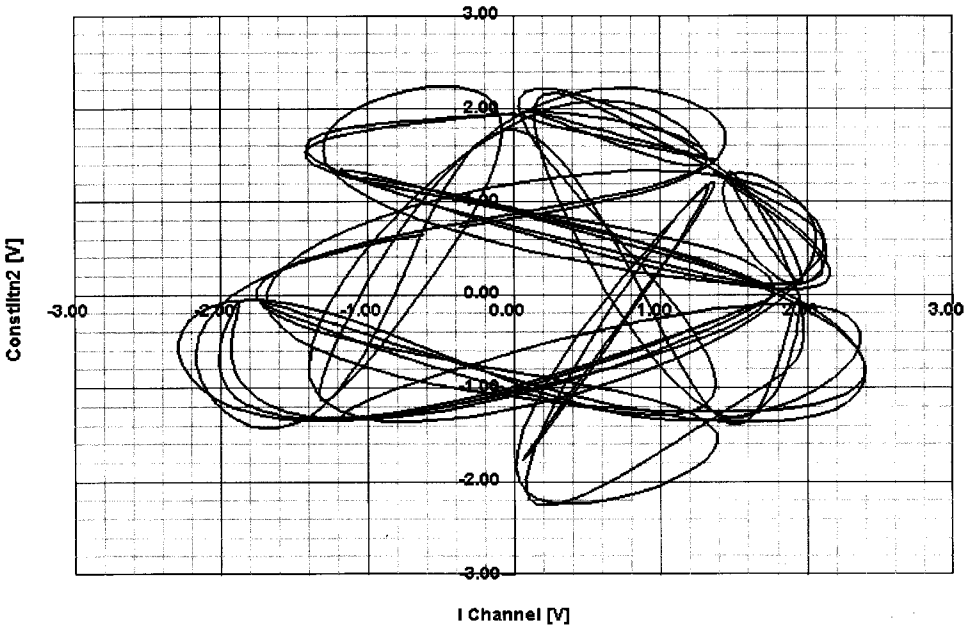


Figure 3-45 Constellation diagram for π/4-DQPSK at Port 2 and 0-dBm source power.

3-1-5 AGC

Many wireless systems consist of an up conversion or down conversion using an intermediate frequency, typically 10% or less of the receiving frequency (the FM broadcast IF has been set at 10.7 MHz for the same reason).

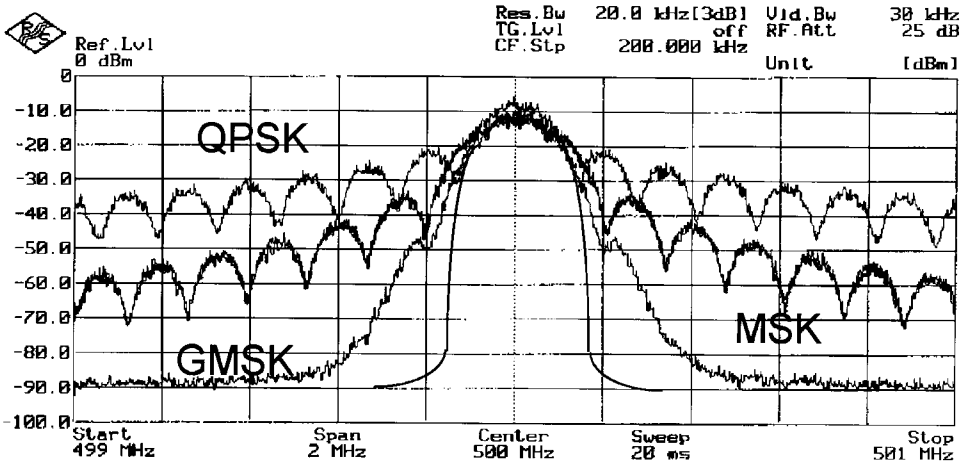


Figure 3-46 Spectrogram showing the characteristics of MSK, GMSK, and QPSK.

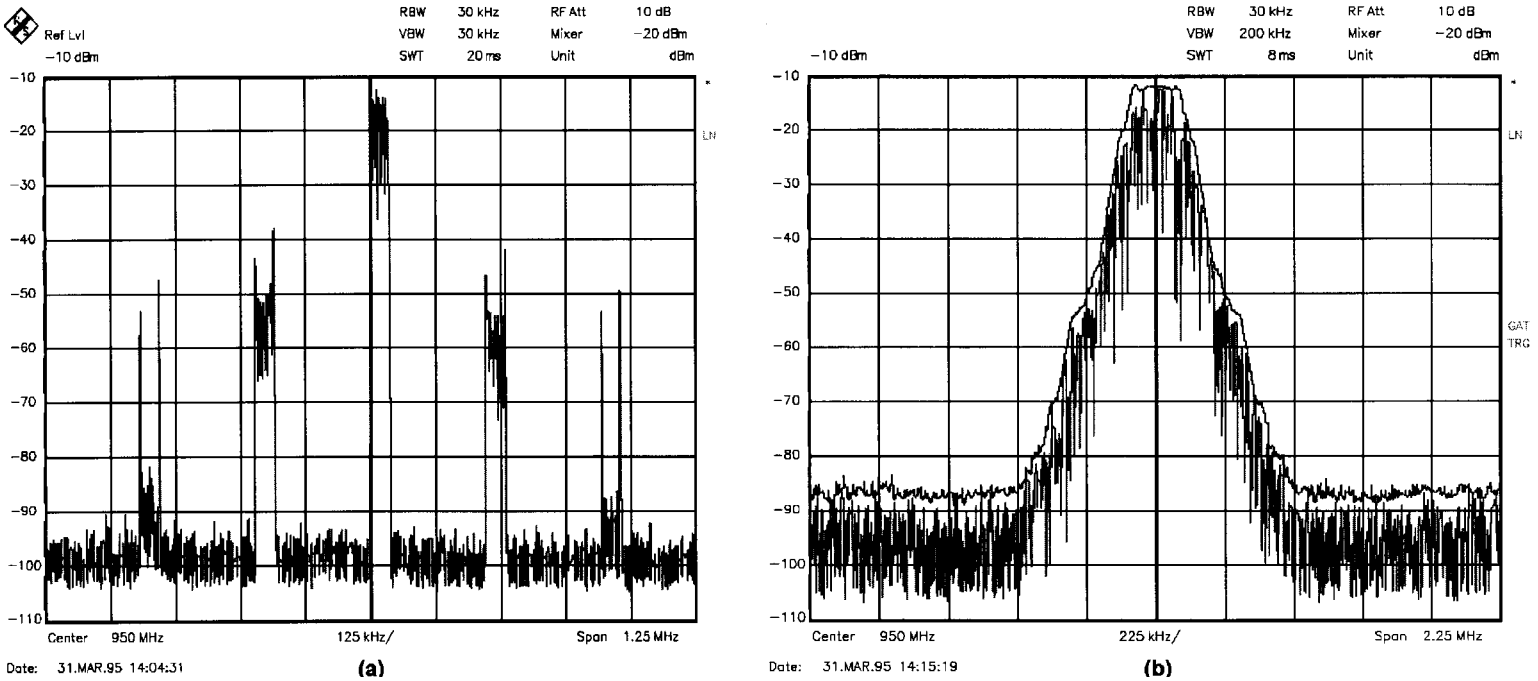


Figure 3-47 Pulsed GSM signal (a) without gating and (b) with gating.

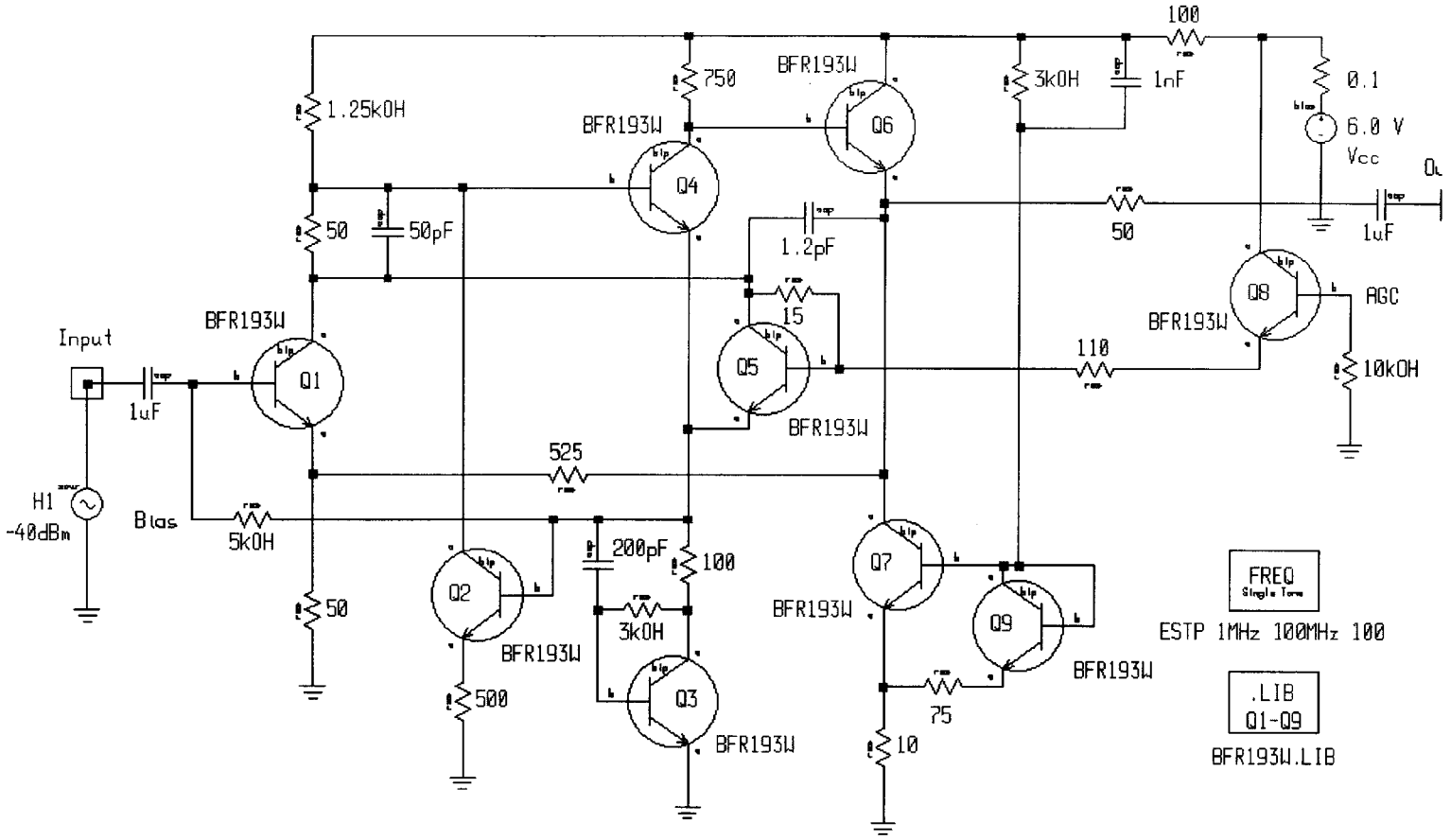


Figure 3-48 Schematic of the Plessey SL610 wideband amplifier with AGC. The 75-Ω resistor in the emitter of Q9 had to be added to obtain the appropriate dc current.

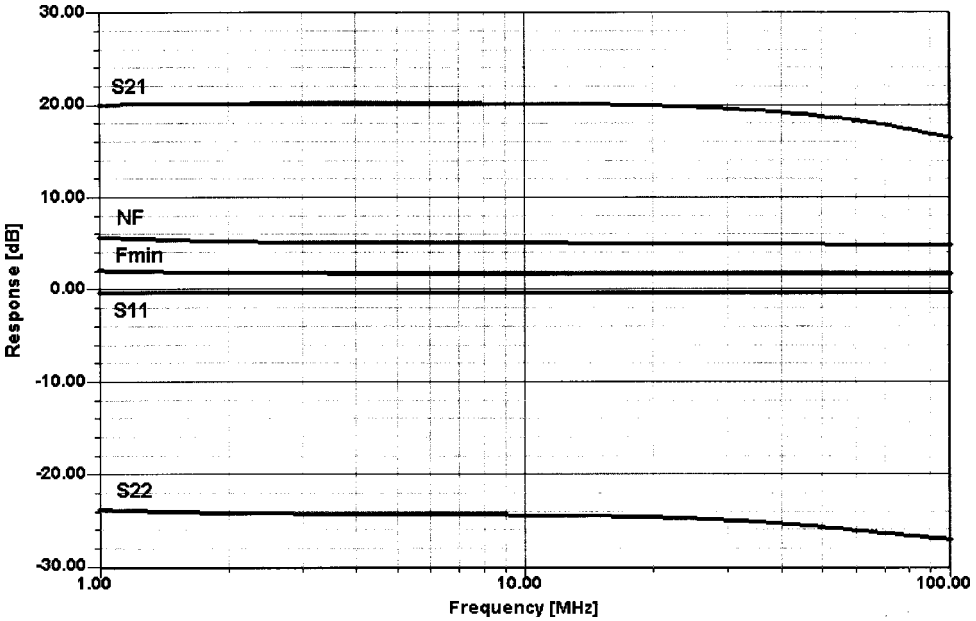


Figure 3-49 Simulated frequency response of the Plessey SL610 wideband amplifier IC with AGC = 0 V.

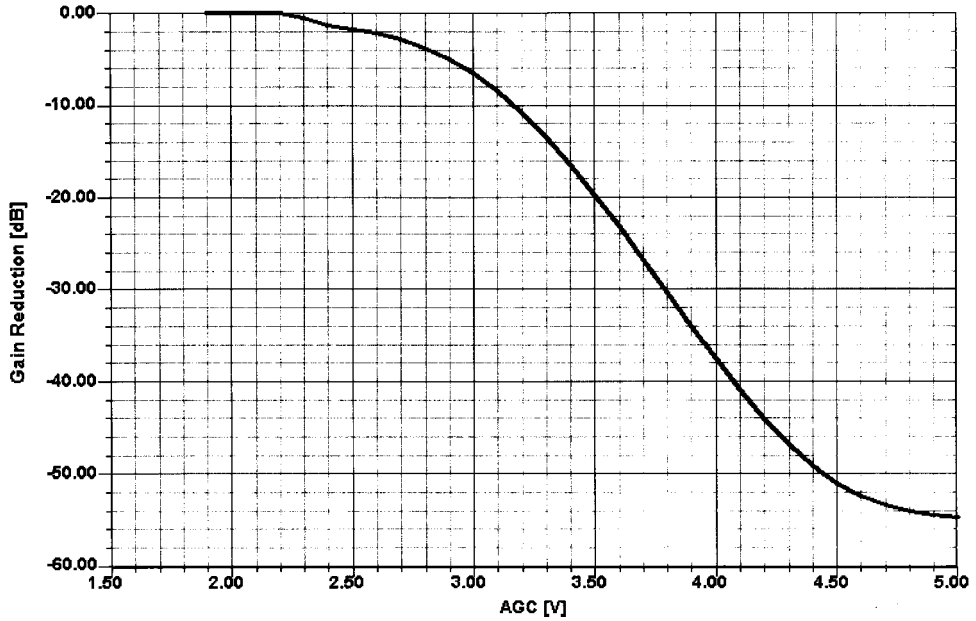


Figure 3-50 Simulated gain reduction versus AGC of the Plessey SL610 wideband amplifier IC.

Table 3-7 CGY121A features

- Variable gain amplifier (MMIC amplifier) for mobile communication
- Typical gain control range over 50 dB
- Positive control voltage
- 50-Ω input and output matched
- Low power consumption
- Operating voltage range: 2.7–6 V
- Frequency range 800 MHz to 2.5 GHz

We therefore will show two amplifiers with AGC and available data. The first one is a wideband amplifier type SL610, which has been manufactured by Plessey; its current fate is unknown. However, since we put the circuit into the simulator it provides significant insight into the operation of such AGC stages. Figure 3-48 shows the actual schematic of the SL610 entered in Ansoft's Serenade Design Environment for simulation.

The AGC action is derived from Q8 through Q5, which, together with Q4, forms a differential amplifier. Each consecutive stage (Q1, Q4, Q6) has increasing current, totaling about 20 mA; the last transistor takes about 13 mA. Based on the feedback surrounding the transistors, the actual transistor characteristics are less important. We have shown this by putting much higher f_T devices into the circuit than were available for the original design (20 years ago). The circuit is still interesting today because of its good large-signal-handling capability.

Table 3-8 CGY121A electrical characteristics at 900 MHz and 1.8 GHz

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
(T_A 25 °C, f = 900 MHz, V_g = –4 V, $RS = RL$ = 50 Ω unless otherwise specified)					
Power gain V_d = 3 V; I = 45 mA; V_{con} = 3 V	G	17	19	—	dB
Input return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{in}	—	11	—	dB
Output return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{out}	—	10	—	dB
Gain control range V_{con} = 3 V ... 0 V; V_d = 3 V; I = 45 mA	dG	48	53	—	dB
1-dB gain compression V_d = 3 V; I = 45 mA; V_{con} = 3 V	P_{1dB}	—	14	—	dBm
(T_A 25 °C, f = 1800 MHz, V_g = –4 V, $RS = RL$ = 50 Ω unless otherwise specified)					
Power gain V_d = 3 V; I = 45 mA; V_{con} = 3 V	G	15.5	17.5	—	dB
Input return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{in}	—	10	—	dB
Output return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{out}	—	8.5	—	dB
Gain control range V_{con} = 3 V ... 0 V; V_d = 3 V; I = 45 mA	dG	48	53	—	dB
1-dB gain compression V_d = 3 V; I = 45 mA; V_{con} = 3 V	P_{1dB}	—	14	—	dBm

Figure 3-49 shows the amplifier's frequency response which is extremely close to the measured data. The good simulation also can be seen from comparing the published AGC curve with the one we found in the simulation. The bias of the first transistor remains essentially constant, resulting in a noise figure more independent of the AGC than the PIN-diode attenuator, and also stabilizes the input impedance as a function of AGC. Figure 3-50 shows the SL610's predicted AGC response.

The same requirement is applicable for the wireless frequency range. Tables 3-7 and 3-8 and Figures 3-51 through 3-54 present specifications for the recently introduced CGY121A GaAs MMIC by Infineon Technologies AG, formerly the Semiconductor Group part of Siemens AG.

3-1-6 Bias and Power Voltage and Current (Power Consumption)

Biasing of transistors operating in Class A or B follows some standard rules. The only deviation from this tends to be when the operating voltage is very low, below 3 V, because then the luxury of using voltage drops for good dc stability as a function of temperature disappears. (We will discuss low-voltage design in detail in Section 3-2-6.) In our simple example (Figure 3-55), we are operating from a 12-V power supply, and the first rule of thumb is to make sure that the voltage between the emitter and ground is at least 0.7 V. This is necessary to compensate the base–emitter junction threshold voltage, which decreases by about 2 mV/°C as the temperature increases. A larger voltage drop in the emitter–ground connection greatly reduces the influence. In this amplifier the emitter–ground path consists of an emitter unbypassed resistor of 10 Ω in series with a bypassed 501-Ω resistor, resulting in an overall voltage drop of 2.6. We could reduce the supply voltage by about 1.9 V at the collector side and still maintain a safety margin of 0.7 V at the emitter–ground connection. The current 2.6 V, so to speak, is overkill.

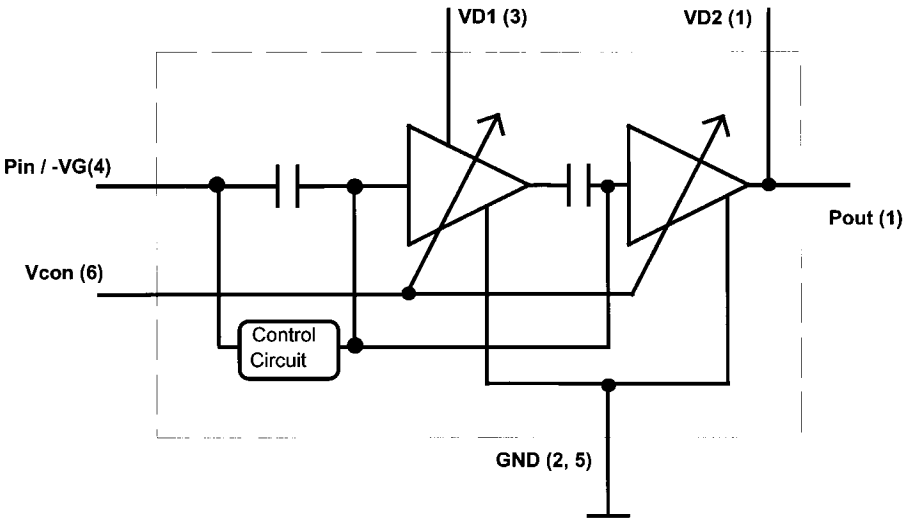
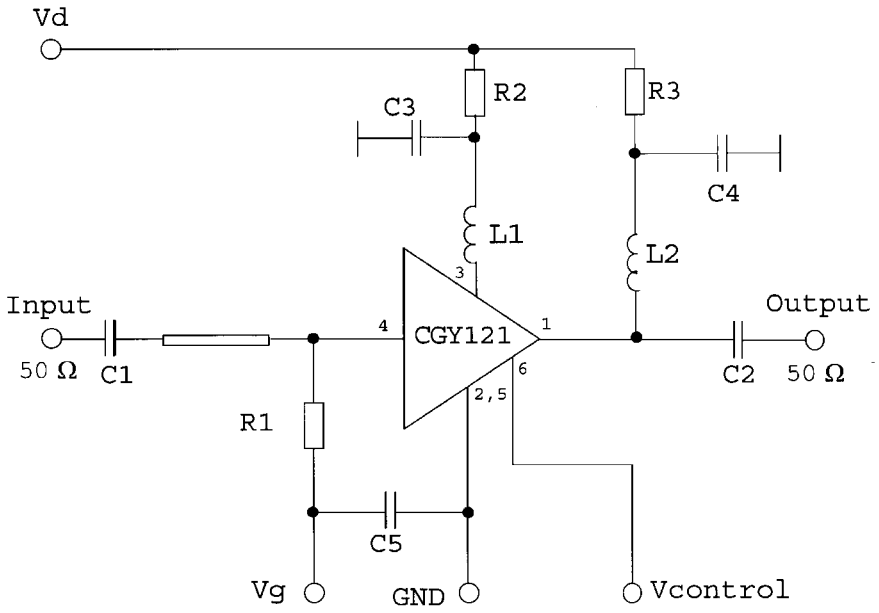


Figure 3-51 Functional block diagram of the gain-controllable CGY121A GaAs MMIC. Gain control is achieved by applying 0–3 V dc to the Vcon pin.



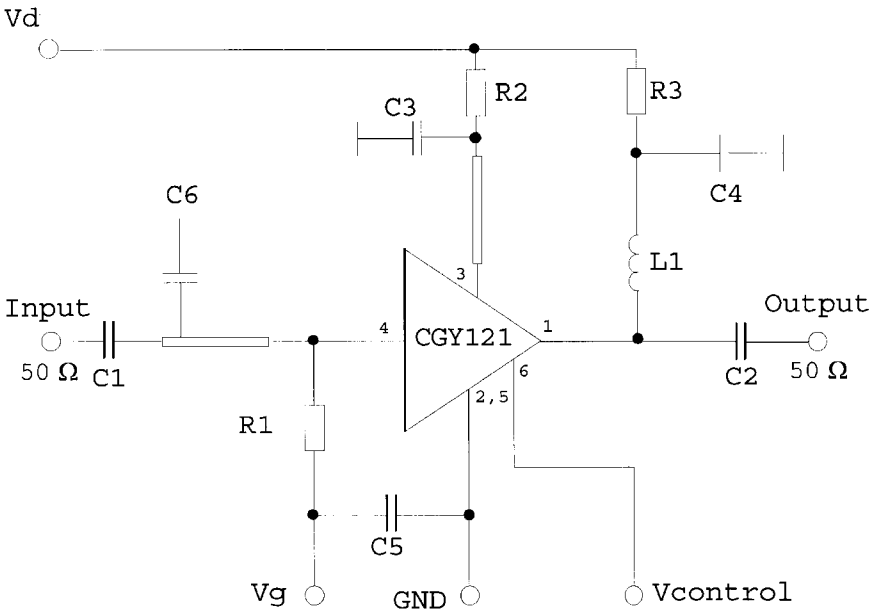
Parts List

Frequency	900 MHz
C1, C2 (Siemens Size 0603)	22 pF
C3, C4 (Siemens Size 0603)	100 nF
C5 (Siemens Size 0603)	47 nF
L1 (Coilcraft 0805CS-150XKBC)	15 nH
L2 (Coilcraft 0805CS-270XMBC)	27 nH
R1 (Siemens B 54102-A1271-J60)	270 Ω
R2 (Siemens B 54102-A1120-J60)	12 Ω
R3	6.8 Ω

Figure 3-52 CGY121A 900-MHz application circuit and parts list.

On the collector side, for an *RC* type of amplifier, it is not a bad idea to set the collector voltage at about half the supply, which allows the RF swing to be half the supply voltage. In this case, we have violated the rule because of our 2.6-V drop. Assuming we had chosen a 0.7-V drop, then the collector–emitter voltage would be 6.3 V available for the swing. The choice of 5 mA for the transistor has come from a combination of good linearity and low noise figure. While the 10-Ω unbypassed resistor will grade the noise figure somewhat, it reduces the emitter-current-dependent distortion. The base–emitter diffusion resistance equals 26 mV/5 mA, or roughly 5 Ω. The 10-Ω resistor is twofold larger and therefore adds to the linearity and to the noise.

To determine the base bias resistor it is a good assumption that the current going through the two resistors (in our case, 8.2 and 3.3 kΩ) is about 10% (or a slightly higher fraction) of



Parts List

Frequency	1.9 GHz
C1, C2 (Siemens size 0603)	12 pF
C3, C4 (Siemens size 0603)	100 nF
C5 (Siemens size 0603)	47 nF
C6 (Siemens size 0603)	1.2 pF
L1 (Coilcraft 0805CS-150XKBC)	15 nH
R1 (Siemens B 54102-A1271-J60)	270 Ω
R2 (Siemens B 54102-A1120-J60)	12 Ω
R3	2.7 Ω

Figure 3-53 CGY121A 1.9-GHz application circuit and parts list.

the collector current. Since the dc current gain of the transistor is typically between 50 and 100, it is sufficient if the current through the resistive network is between 5 and 10 times higher than the dc bias current taken off the network. This type of bias scheme is only valid for Class A to AB_1. Later, we will see that for monolithic circuits, additional transistors will be used to generate the required voltage drops. Lately, combinations of *npn* and *pnp* transistors have become popular to avoid using any resistance in the emitter for dc biasing purposes. Figure 3-56 shows such a circuit. This, of course, invites thermal runaway. To minimize the circuitry, there are bias ICs, such as the Siemens BCR400R, now available. Figures 3-57 and 3-58 show its application in biasing BJTs and FETs, respectively; it can also be used to control PIN-diode bias current for TR antenna switching, as shown in Figure 3-59.

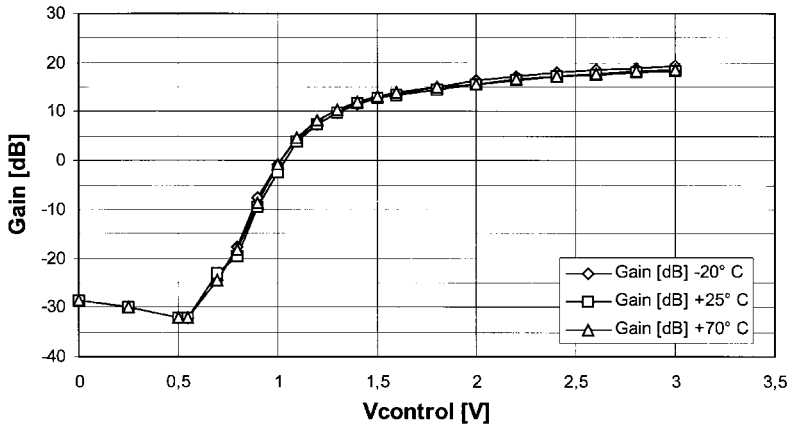


Figure 3-54 CGY121A gain versus $V_{control}$ at 1.9 GHz.

To validate this, we have used the Serenade 8.0 CAD tool, which has a built-in bias- and temperature-dependent noise model for BJTs and FETs. The SPICE-type dc analysis determines the actual currents, which were mentioned above. By maintaining the same bias and just bypassing both emitter resistors at the same time, the set of curves shown in Figure 3-60 will be obtained. As an exercise, we recommend the user to determine the optimum generator impedance for F_{min} with a 10-Ω resistor. Again, both the emitter current and collector voltage affect the noise figure, the intermodulation, and f_T. If a transistor is used

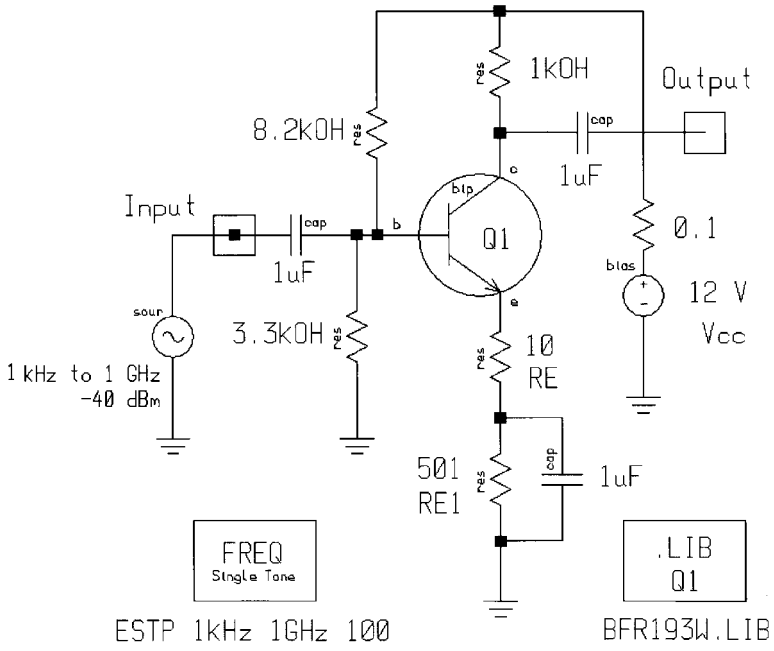


Figure 3-55 BJT amplifier example.

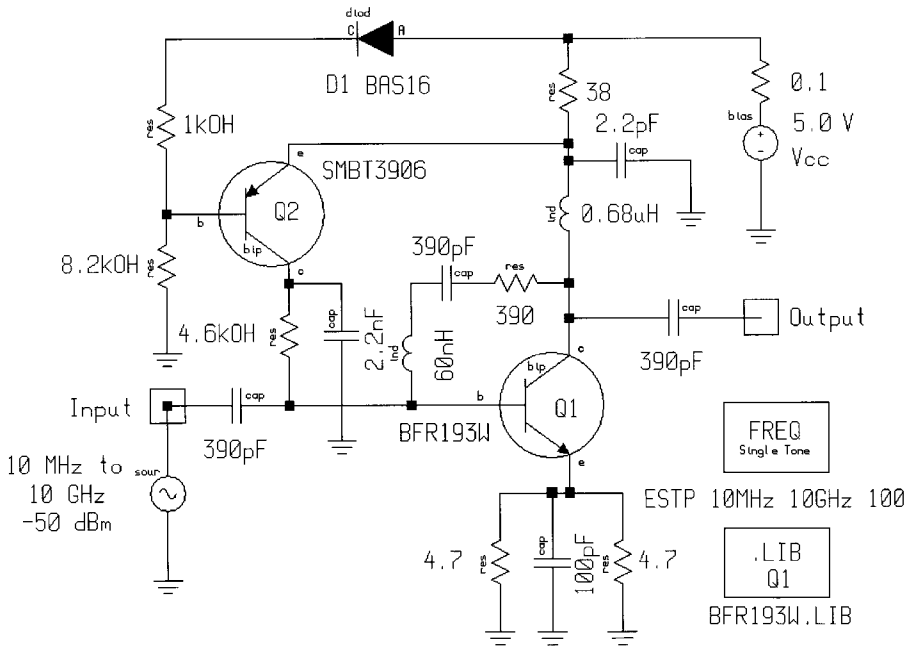


Figure 3-56 Wideband amplifier with active biasing. The 4.7-Ω resistors in the BFR193W emitter compensate the circuit's gain at low frequencies.

that is intended to be operated at significantly higher currents, then f_T will suffer while the other two parameters are less affected.

The dc input power translates into thermal dissipation, and Figure 3-60 is a plot of the noise figure as a function of temperature with and without emitter feedback. The heavy dc stabilization prevents a significant effect on the other parameters. The actual dc shift is from 5.0 to 5.2 mA for a temperature change of 70 K.

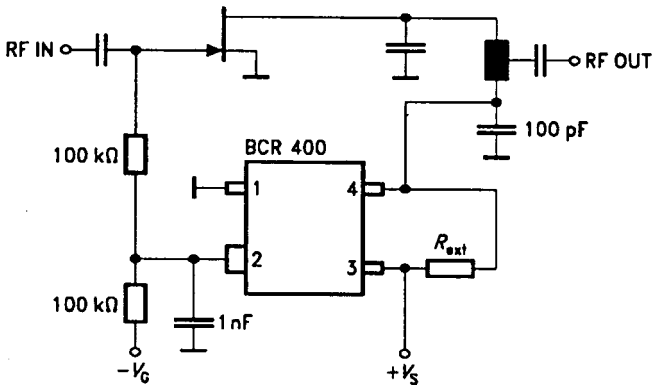


Figure 3-57 The Siemens BCR400 bias controller applied to a GaAsFET.

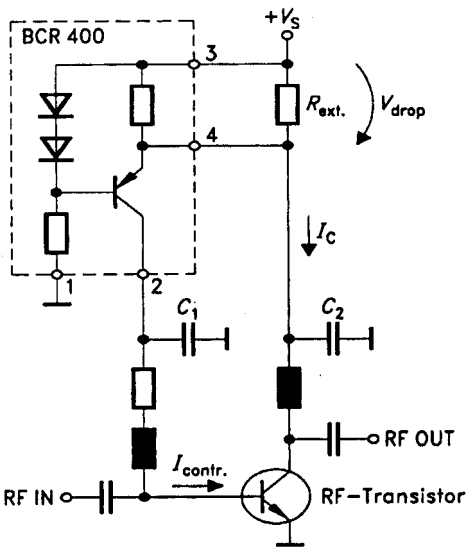


Figure 3-58 The Siemens BCR400 bias controller applied to a BJT.

As far as FETs are concerned, the enhancement types, typically PMOS, can use a similar type of bias, while the NMOS types complicate the designer's life as they require a negative gate voltage. For low-noise operation, these devices are operated somewhat close to the pinchoff voltage. Putting a source resistor in place causes mostly headaches simply because the gain is so high that the transistors are always ready to find a frequency at which they love to oscillate. The transistor will look for every possible parasitic to turn the amplifier into an oscillator.

3-2 AMPLIFIER GAIN, STABILITY, AND MATCHING

Most semiconductor manufacturers describe their devices these days by providing either their *S* parameters (small-signal) or SPICE-type parameters (large-signal). One of our headaches in writing this book has been that many of the published models don't translate

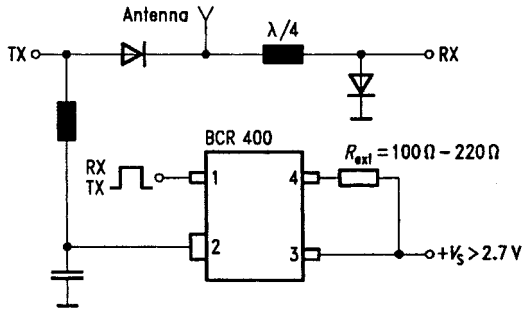


Figure 3-59 The Siemens BCR400 bias controller applied to TR switching.

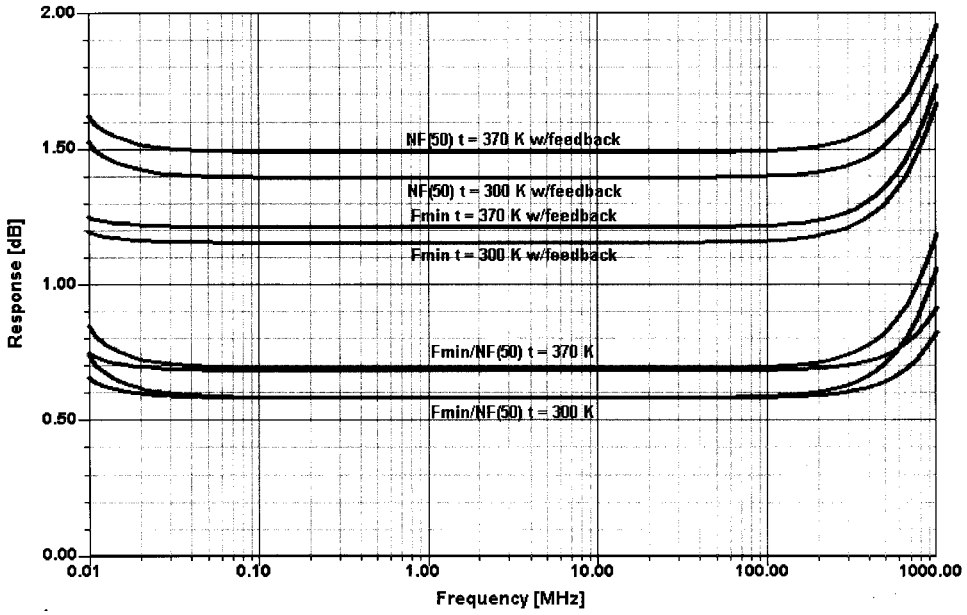


Figure 3-60 F_{min} and NF for the amplifier in Figure 3-55, with and without emitter feedback, at 300 and 370 K.

from large signal to small signal without giving unexpected or false results. Despite the availability of tools to do it right, many companies still can only provide data that are practically inadequate, and unfortunately the CAD manufacturers then get the blame for results for which they are not really responsible.

Tables 3-9 and 3-10 show typical data for a bipolar transistor (the BFP420 by Siemens). The next section presents a summary of the derivation and use of *S* parameters as published by Hewlett-Packard and others.

3-2-1 Scattering Parameter Relationships

Scattering (*S*) parameters characterize a network in terms of incident and reflected waves (Figures 3-61 and 3-62).

$$b_1 = s_{11}a_1 + s_{12}a_2 \tag{3-122}$$

$$b_2 = s_{21}a_1 + s_{22}a_2 \tag{3-123}$$

Notice that

$$s_{11} = \frac{b_1}{a_1} = \frac{V_1/I_1 - Z_0}{V_1/I_1 + Z_0} = \frac{Z_1 - Z_0}{Z_1 + Z_0} \tag{3-124}$$

and

Table 3-9 BFP420 common-emitter *S* parameters[a]

f	S_{11}		S_{21}		S_{12}		S_{22}	
(GHz)	(Magnitude)	(Angle)	(Magnitude)	(Angle)	(Magnitude)	(Angle)	(Magnitude)	(Angle)
				V_{CE} = 2 V, I_C = 20 mA				
0.01	0.452	–2.3	37.62	178.3	0.0011	94.4	0.956	–0.6
0.1	0.447	–52.1	36.30	164.7	0.0068	82.5	0.941	–12.4
0.5	0.386	–101.1	23.41	121.0	0.0262	61.7	0.632	–47.2
1.0	0.378	–146.2	13.99	96.0	0.0395	57.8	0.395	–63.9
2.0	0.405	173.5	7.18	70.8	0.0664	54.0	0.222	–87.3
3.0	0.446	149.4	4.77	52.6	0.0949	47.1	0.133	–111.3
4.0	0.501	130.0	3.52	36.8	0.1206	38.5	0.133	–158.5
6.0	0.599	104.8	2.27	8.2	0.1646	18.9	0.196	142.0
8.0	0.700	78.5	1.51	–20.8	0.1800	–2.4	0.289	99.3
9.0	0.758	67.6	1.25	–34.4	0.1820	–13.0	0.379	84.1
10.0	0.800	62.0	1.04	–43.5	0.1800	–19.3	0.465	76.6
				V_{CE} = 2 V, I_C = 5 mA				
0.01	0.790	–1.0	15.14	179.2	0.0012	83.4	0.988	–0.7
0.1	0.786	–11.6	14.98	171.8	0.0092	84.1	0.982	–6.5
0.5	0.702	–55.7	12.86	140.1	0.0398	62.8	0.857	–29.8
1.0	0.589	–99.1	9.63	112.6	0.0603	46.5	0.647	–48.6
2.0	0.507	–156.0	5.60	79.4	0.0798	34.6	0.401	–70.3
3.0	0.511	168.5	3.84	57.1	0.0957	29.8	0.267	–84.2
4.0	0.549	142.0	2.87	38.5	0.1121	25.1	0.207	–110.5
5.0	0.604	123.9	2.26	22.1	0.1285	19.4	0.150	–137.3
6.0	0.633	110.0	1.86	6.7	0.1442	13.1	0.173	–169.8

Table 3-10 BFP420 noise parameters

f	$F_{\min}$[a]	G_a[a]	Γ_{opt}		R_N	r_n	$F_{50\Omega}$[b]	$\lvert S_{21} \rvert^2$ [b]
(GHZ)	dB	(dB)	(Magnitude)	(Angle)	(Ω)		(dB)	(dB)
				V_{CE} = 2 V, I_C = 5 mA				
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	–175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	–157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	–142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	–123.0	15.0	0.30	3.30	5.3

[a]Input matched for minimum noise figure, output for maximum gain.
[b]$Z_S = Z_L = 50\ \Omega$.

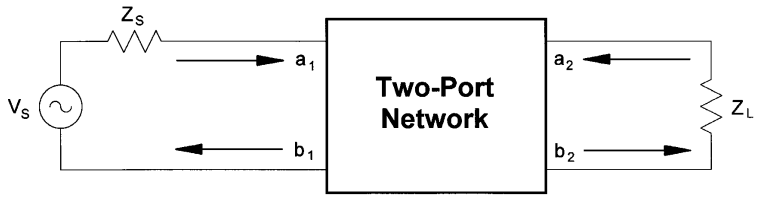


Figure 3-61 *S* parameters characterize a network (a two-port is shown) in terms of incident waves (a_1, a_2) and reflected waves (b_1, b_2). Figure 3-62 shows the flow graph for this network.

$$Z_1 = Z_0 \frac{(1 + s_{11})}{(1 - s_{11})} \tag{3-125}$$

where $Z_1 = V_1/I_1$ is the input impedance at Port 1.

This relationship between reflection coefficient and impedance is the basis of the Smith Chart transmission-line calculator. Consequently, the reflection coefficients S_{11} and S_{22} can be plotted on Smith charts, converted directly to impedance, and easily manipulated to determine matching networks for optimizing a circuit design.

The above equations show one of the important advantages of *S* parameters: They are simply gains and reflection coefficients, both familiar quantities to engineers. By comparison, some of the *Y* parameters described earlier in the article are not so familiar. For example, the *Y* parameter corresponding to insertion gain S_{21} is the "forward transadmittance," Y_{21}. Clearly, insertion gain gives by far the greater insight into the operation of the network.

S parameters are simply related to power gain and mismatch loss, quantities that are often of more interest than the corresponding voltage functions:

$$|s_{11}|^2 = \frac{\text{Power reflected from the network input}}{\text{Power incident on the network input}} \tag{3-126}$$

$$|s_{21}|^2 = \frac{\text{Power delivered to a } Z_0 \text{ load}}{\text{Power available from } Z_0 \text{ source}} \tag{3-127}$$

$$|s_{12}|^2 = \text{Reverse transducer power gain with } Z_0 \text{ load and source.} \tag{3-128}$$

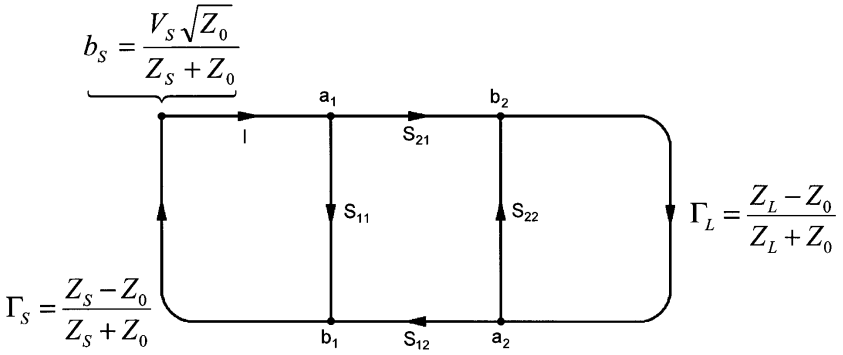


Figure 3-62 Flow graph for the network shown in Figure 3-61.

$$|s_{22}|^2 = \frac{\text{Power reflected from the network output}}{\text{Power incident on the network output}} \tag{3-129}$$

From this we obtain:
Input reflection coefficient with arbitrary Z_L:

$$s'_{11} = s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1 - s_{22}\Gamma_L} \tag{3-130}$$

Output reflection coefficient with arbitrary Z_S:

$$s'_{22} = s_{22} + \frac{s_{12}s_{21}\Gamma_S}{1 - s_{11}\Gamma_S} \tag{3-131}$$

Voltage gain with arbitrary Z_L and Z_S:

$$A_V = \frac{V_2}{V_1} = \frac{s_{21}(1 + \Gamma_L)}{(1 - s_{22}\Gamma_L)(1 + s'_{11})} \tag{3-132}$$

$$\text{Power gain} = \frac{\text{Power delivered to load}}{\text{Power input to network}}:$$

$$G = \frac{|s_{21}|^2 (1 - |\Gamma_L|^2)}{(1 - |s_{11}|^2) + |\Gamma_L|^2 (|s_{22}|^2 - |D|^2) - 2\,\text{Re}\,(\Gamma_L N)} \tag{3-133}$$

$$\text{Available power gain} = \frac{\text{Power available from network}}{\text{Power available from source}}:$$

$$G_A = \frac{|s_{21}|^2 (1 - |\Gamma_S|^2)}{(1 - |s_{22}|^2) + |\Gamma_S|^2 (|s_{11}|^2 - |D|^2) - 2\,\text{Re}\,(\Gamma_S M)} \tag{3-134}$$

$$\text{Transducer power gain} = \frac{\text{Power delivered to load}}{\text{Power available from source}}:$$

$$G_T = \frac{|s_{21}|^2 (1 - |\Gamma_S|^2)(1 - |\Gamma_L|^2)}{|(1 - s_{11}\Gamma_S)(1 - s_{22}\Gamma_L) - s_{12}s_{21}\Gamma_L\Gamma_S|^2} \tag{3-135}$$

Unilateral transducer power gain ($S_{12} = 0$):

$$G_{Tu} = \frac{|s_{21}|^2(1 - |\Gamma_S|^2)(1 - |\Gamma_L|^2)}{|1 - s_{11}\Gamma_S|^2|1 - s_{22}\Gamma_L|^2}$$

$$= G_0 G_1 G_2 \tag{3-136}$$

$$G_0 = |s_{21}|^2 \tag{3-137}$$

$$G_1 = \frac{1 - |\Gamma_S|^2}{|1 - s_{11}\Gamma_S|^2} \tag{3-138}$$

$$G_2 = \frac{1 - |\Gamma_L|^2}{|1 - s_{22}\Gamma_L|^2} \tag{3-139}$$

Maximum unilateral transducer power gain when $|S_{11}| < 1$ and $|S_{22}| < 1$:

$$G_u = \frac{|s_{21}|^2}{|(1 - |s_{11}|^2)(1 - |s_{22}|)^2|}$$

$$= G_0\, G_{1\max}\, G_{2\max} \tag{3-140}$$

$$G_{i\max} = \frac{1}{1 - |s_{ii}|^2} \quad i = 1, 2 \tag{3-141}$$

Constant-gain circles (unilateral case: $S_{12} = 0$):

- Center of constant-gain circle is on line between center of Smith Chart and point representing S_{ii}^*
- Distance of center of circle from center of Smith Chart:

$$r_i = \frac{g_i |s_{ii}|}{1 - |s_{ii}|^2(1 - g_i)} \tag{3-142}$$

- Radius of circle:

$$\rho_i = \frac{\sqrt{1 - g_i}(1 - |s_{ii}|^2)}{1 - |s_{ii}|^2(1 - g_i)} \tag{3-143}$$

where $i = 1, 2$ and

$$g_i = \frac{G_i}{G_{i\max}} = G_i(1 - |s_{ii}|^2) \tag{3-144}$$

Unilateral figure of merit:

$$u = \frac{|s_{11}s_{22}s_{12}s_{21}|}{|(1 - |s_{11}|^2)(1 - |s_{22}|^2)|} \tag{3-145}$$

Error limits of unilateral gain calculation:

$$\frac{1}{(1 + u^2)} < \frac{G_T}{G_{Tu}} < \frac{1}{(1 - u^2)} \tag{3-146}$$

Conditions for absolute stability. No passive source or load will cause a network to oscillate if a, b, and c are all satisfied:

$$\text{a.}\quad |s_{11}| < 1, |s_{22}| < 1 \tag{3-147}$$

$$\text{b.} \quad \left| \frac{|s_{12}s_{21}| - |M^*|}{|s_{11}|^2 - |D|^2} \right| > 1 \tag{3-148}$$

$$\text{c.} \quad \left| \frac{|s_{12}s_{21}| - |N^*|}{|s_{22}|^2 - |D|^2} \right| > 1 \tag{3-149}$$

where

$$D = s_{11}s_{22} - s_{12}s_{21} \tag{3-150}$$

$$M = s_{11} - Ds_{22}^* \tag{3-151}$$

$$N = s_{22} - Ds_{11}^* \tag{3-152}$$

Condition that a two-port network can be simultaneously matched with a positive real source and load:

$$K > 1 \tag{3-153}$$

$$K = \frac{1 + |D|^2 - |s_{11}|^2 - |s_{22}|^2}{2|s_{12}s_{21}|} \tag{3-154}$$

Source and load for simultaneous match

$$\Gamma_{mS} = M^* \left[\frac{B_1 \pm \sqrt{B_1^2 - 4|M|^2}}{2|M|^2} \right] \tag{3-155}$$

$$\Gamma_{mL} = N^* \left[\frac{B_2 \pm \sqrt{B_2^2 - 4|N|^2}}{2|N|^2} \right] \tag{3-156}$$

where

$$B_1 = 1 + |s_{11}|^2 - |s_{22}|^2 - |D|^2 \tag{3-157}$$

$$B_2 = 1 + |s_{22}|^2 - |s_{11}|^2 - |D|^2 \tag{3-158}$$

Maximum available power gain. If $K > 1$,

$$G_{A\max} = \left| \frac{s_{21}}{s_{12}} (K \pm \sqrt{K^2 - 1}) \right| \tag{3-159}$$

In conclusion, *a two-port network (active or passive) is unconditionally stable* when

$$K > 1$$

$$|S_{12}S_{21}| < 1 - |S_{11}|^2$$

and

$$|S_{12}S_{21}| < 1 - |S_{22}|^2$$

3-2-2 Low-Noise Amplifiers

A low-noise amplifier combines a low noise figure, reasonable gain, and stability without oscillation over its entire useful frequency range. These amplifiers are typically operated in Class A, which is characterized by a bias point more or less at the center of maximum current and voltage capability of the device used, and by RF current and voltages that are sufficiently small relative to the bias point that the bias point does not shift. This means that an amplifier operating at, say, 3 V and 1 mA for low-noise, high-gain operation will not be Class A if the RF current is larger than 0.1% of the current or the voltage swing is more than 0.1% of the current swing. These amplifiers can be designed using standard linear *S* parameters, which are (hopefully) available from the device manufacturer for the bias point of interest. If suitable *S* parameters are not available, we must resort to a CAD tool that allows us to determine a device's bias point as a function of its bias network and to generate small-signal *S* parameters for the device. With suitable *S* parameters in hand, and assuming the amplifier will be frequency selective, we then connect tuned circuits to the input and output to provide matching.

In designing amplifiers, the "power gain" of a circuit is mentioned frequently. This is actually deceptive because power gain is somewhat associated with dc power dissipation, and in considering it, we are on the verge of discussing large-signal performance. Therefore it would be better to define the power gain as a function of operating mode, such as "Class A power gain." We have also outlined that there is a difference between noise matching and gain matching. If we can be allowed to step back into the vacuum-tube era for a moment, the reader should be reminded that the difference between power gain and noise matching had to do with the grid-plate feedback capacitance which, via the Miller effect, showed up in parallel with the input. This actually detunes the input slightly off-center relative to the best gain. The reason for this is that as the feedback increases, Y_{12} or S_{12} starts playing a role, and the frequently assumed unilateral case no longer exists. The difference between noise matching and gain matching increases as a function of frequency. It can be brought closer together if appropriate reactive feedback is used. The two choices are: (1) increase the emitter or source inductance or (2) slightly increase the base-to-collector/gate-to-drain capacitance. Needless to say, the penalty for this is potential instability, and a sweep must be done from a few megahertz to several gigahertz to make sure the stage doesn't "take off."

The most efficiency this circuit provides is 50%; even this is a theoretical limit if one tries to keep distortion under control.

While the equation $P_{\text{OUT}} = (V_{\text{bat}} - V_{\text{sat}})^2/2R_L$ is valid, it really applies only to power amplifiers in the sense that for Class A, high-gain operation R_L can be made significantly higher as the absolute power output is not the important item. While the supply voltage V_{bat} is always known, the saturation voltage V_{sat} has to be obtained by the load line dc analysis. Note: The RF saturation voltage is always higher than the dc one. When looking at Figure 3-63, it becomes obvious that the conducting angle is 180°, which implies that the entire sine wave at the input is reproduced as an essentially undistorted sine wave at the output.

We can now calculate the components of the distorted current in separating them into the dc component, the fundamental frequency, and its first and second harmonics:

$$\text{Fundamental} \quad i_{c1} = i_{\text{peak}} \cdot f_1(\Theta) \tag{3-160}$$

$$\text{Second harmonic} \quad i_{c2} = i_{\text{peak}} \cdot f_2(\Theta) \tag{3-161}$$

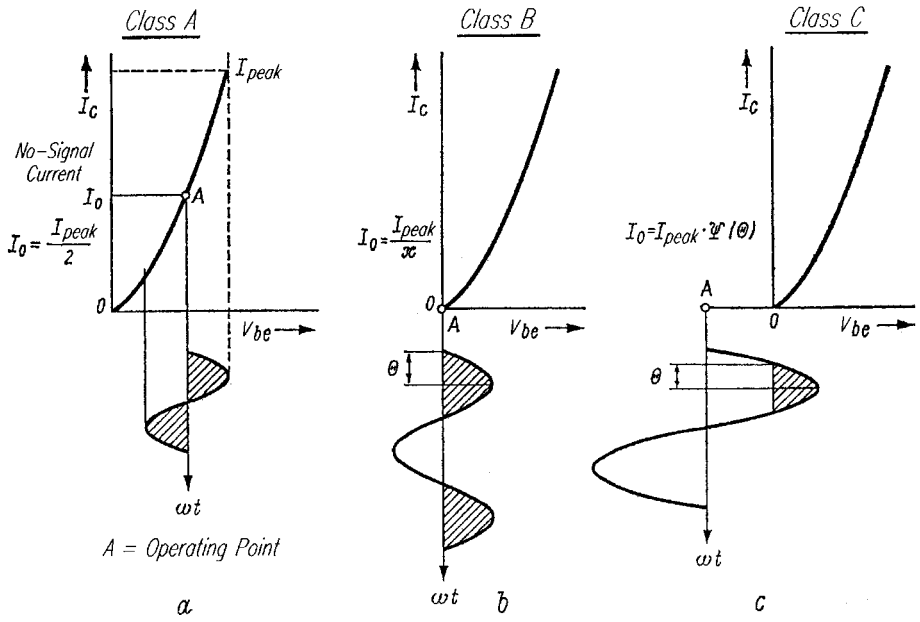


Figure 3-63 Definitions of Class A, B, and C operation, including conduction angle (Θ). Figure 3-64 shows the determination of Θ.

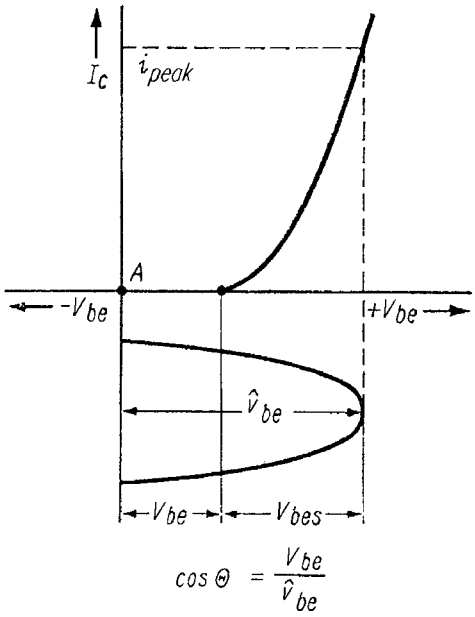


Figure 3-64 Determination of the conduction angle Θ from the base threshold voltage V_{be} and the base RF voltage $\hat{v}_{be}$.

$$\text{Third harmonic} \quad i_{c3} = i_{\text{peak}} \cdot f_3(\Theta) \tag{3-162}$$

For the collector dc current we obtain

$$I_c = i_{\text{peak}} \cdot \Psi(\Theta) \tag{3-163}$$

The frequency-dependent currents are a function of the conduction angle Θ. Figure 3-65 shows these transfer characteristics as a function of the conduction angle for a square-law dynamic transfer function; for higher-order transfer functions, the values don't change much. Tables for this can be found in various sources, including Clark and Hess [47]. Since the transistors exhibit a mixture of an exponential and a linear transfer characteristic (due to the unwanted parasitic emitter/source and collector/drain resistances when building the device), the plot in figure 3-65 is sufficiently valid for practical purposes. This function plot allows us to determine the amplitude of the current at various frequencies once the conducting angle has been set. For example, for Θ = 70°, the maximum collector current is $i_{\text{peak}} = I_c/0.253$, meaning that the peak current of the transistor is approximately four times larger than the dc value one would measure. Once the peak collector current is known, the fundamental and harmonic currents can be calculated:

$$i_{c1} = \frac{I_c}{0.253} f_1(\Theta)$$

$$i_{c1} = \frac{I_c}{0.253} 0.436$$

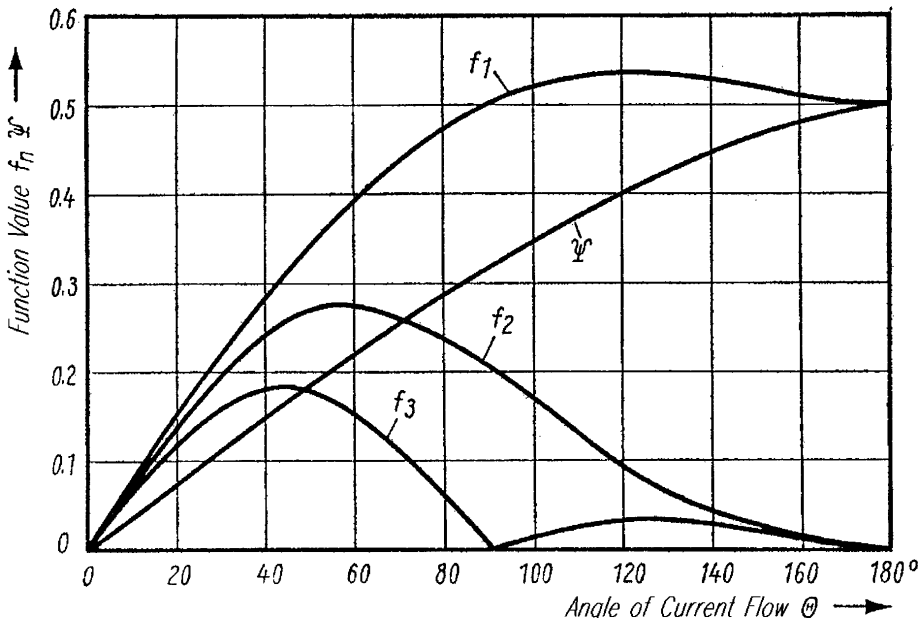


Figure 3-65 Function values for the fundamental, second harmonic, third harmonic, and dc components (f_1, f_2, f_3, and Ψ, respectively) versus device conduction angle.

In order to obtain i_{c2} and i_{c3} one needs to do the same calculations as above with the appropriate functions for f_2 and f_3.

It becomes obvious that the distorted collector current generates a lot of harmonics, which definitely requires sufficient low-pass filtering at the output, even for Class A operation. We already stated that for Class A operation the conduction angle is 180° while it is 90° for Class B operation. The conduction angle for Class C operation can be solved from the ratio of the knee voltage V_{be} and the peak value of the ac voltage at the base. We find

$$\cos \Theta = \frac{V_{be}}{\hat{V}_{be}} \tag{3-164}$$

where $\Theta = 90°$ and $V_{be} = 0$. Given the voltage V_{bes}, which determines the peak collector current (up to i_c peak), we can calculate the effective RF voltage at the base as

$$\hat{v}_{be} = \frac{v_{bes}}{1 - \cos \Theta} \tag{3-165}$$

Therefore, the peak value of v_b becomes v_{bes} for $\Theta = 90°$.

The necessary base voltage V_{be} for maximum current is obtained by the transistor model. Then one finds the threshold voltage at which the transistor becomes conductive and by using Eqs. (3-164) and (3-165) it can be checked whether this results in the appropriate conducting angle. An example for this will be shown in the high-power stage section.

Design Guidelines. The design of a one-stage amplifier consists of finding

- An input lossless matching network M_1
- An input lossless matching network M_2

so that the maximum or desired transistor gain is achieved over the operating bandwidth of the amplifier. Usually, the common-emitter or common-source configuration is chosen for highest gain per stage. If the stability factor K is greater than unity, these two networks can be found to give the maximum available gain $G_{\max}$. If the stability factor is less than or equal to unity, the amplifier could be terminated in a matching structure that causes oscillation; that is, $G_{\max}$ is infinite. This should be avoided by locating the regions of instability in the Γ_G and Γ_L planes. The input and output terminations (Γ_G and Γ_L) must be designed to avoid the instability regions. Usually, these unstable regions are near the conjugate match for S_{11} and S_{22}. Thus, a stable amplifier will require some input and/or output mismatch if K is less than or equal to unity.

There are at least two alternative approaches for potentially unstable amplifiers:

1. Add resistive matching elements to make $K \geq 1$ and $G_{\max} \approx G_{ms}$.
2. Add feedback to make $K \geq 1$ and $G_{\max \approx Gms}$.

For narrowband amplifiers, it is usually recommended that one accept a transistor with $K < 1$, design the amplifier for a gain approaching G_{ms}, and ensure that the Γ_G and Γ_L terminations provide stability at all frequencies, both inside and outside the amplifier passband.

The design of an amplifier would usually have the following specifications:

Gain and gain flatness
Bandwidth and center frequency ($f_2 - f_1, f_0$)
Noise figure
Linear output power
Input reflection coefficient ($VSWR_{in}$)
Output reflection coefficient ($VSWR_{out}$)
Bias voltage and current

For small-signal amplifiers, the small-signal S parameters are sufficient to complete the design. After selecting an appropriate transistor based on these specifications, a one-stage amplifier design should be considered if sufficient gain can be achieved; otherwise, a two-stage amplifier should be designed.

The circuit topology should be chosen to allow dc bias for the transistor. Usually, RF short-circuited stubs are placed near the transistor to allow dc biasing. If the topology does not allow dc bias, a broadband bias choke or high-resistance bias circuit that does not affect the amplifier performance must be used.

The following steps can be tabulated for the design of a one-stage amplifier (see Figure 3.66):

1. Select a transistor based on a datasheet description of the S parameters, noise figure, and linear output power.
2. Calculate K and G_{max} or G_{ms} versus frequency.
3. For $K > 1$, select the topologies that match the input and output (and allow dc biasing) at the upper band edge f_2. Ideally, this will give G_{max} and $S'_{11} = S'_{22} = 0$. Usually, $S_{12} = 0$ is assumed for the initial design. Next, the topology may be varied to flatten gain versus frequency at the expense of S'_{11} and S'_{22}.
4. For $K < 1$, plot the regions of instability on the Γ_G and Γ_L planes and select topologies that partially match the input and output at the upper band edge and avoid the unstable regions. The gain will approach G_{ms} as an upper limit. Next, the topology may be varied to flatten gain versus frequency.
5. After finding initial M_1 and M_2, plot the amplifier S parameters versus frequency; make adjustments in topology until the specifications for gain, input reflection coefficient, and output reflection are satisfied. Also plot Γ_G and Γ_L versus frequency to verify amplifier stability.
6. Design the dc bias circuit. Lay out the elements of the complete amplifier and check realizability.

As a beginning point in an amplifier design, the circuit topologies may be designed with the assumption that S_{12} is zero. Later, an exact design procedure will be described that includes the topologies with terminations in the stable region.

Design Examples. Throughout this book we always address the issue that we have to separate the battery-operated handheld applications and the more stationary designs like those for base stations. In the case of the handheld or low-voltage applications, the dynamic range requirements are significantly less than the front-end of a base station requires.

NE68133 Matched Amplifier. We will first consider the design of a low-voltage, low-current, low-noise amplifier based on the NEC BJT NE68133. This device was chosen

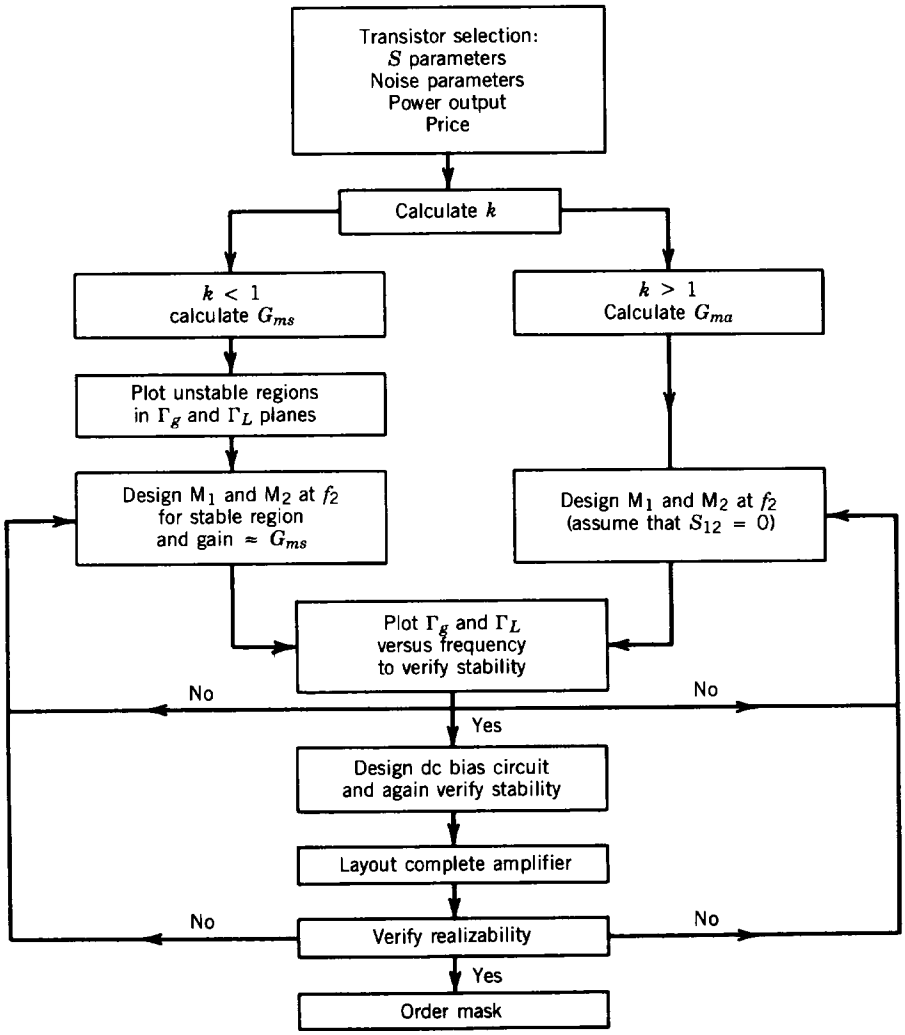


Figure 3-66 Simplified amplifier design procedure.

because it is unconditionally stable, has a minimum noise figure of about 1.2 dB at 7 mA, has 15.5 dB gain, and has an f_T of 8 GHz. While it is possible to start the design based on the published S and noise parameters, we will choose also to involve the nonlinear model, as this allows us more freedom in selecting the bias point including the collector voltage. The actual topology of the amplifier is shown in Figure 3-67.

Determining appropriate values for the input and output matching networks is essential to finalizing the design. The next several figures illustrate matching-network syntheses undertaken with the interactive Smith Tool in Ansoft's Serenade Design Environment. We evaluate the point of tangency as well as stability circles in the source and load planes, indicating that our impedance match will result in a stable amplifier.

In Figure 3-68, the center circle (with 1.1 marker) refers to the noise circle for which the noise figure of 1.1 dB can be obtained. The left circle (with the 14.3 marker) corresponds to

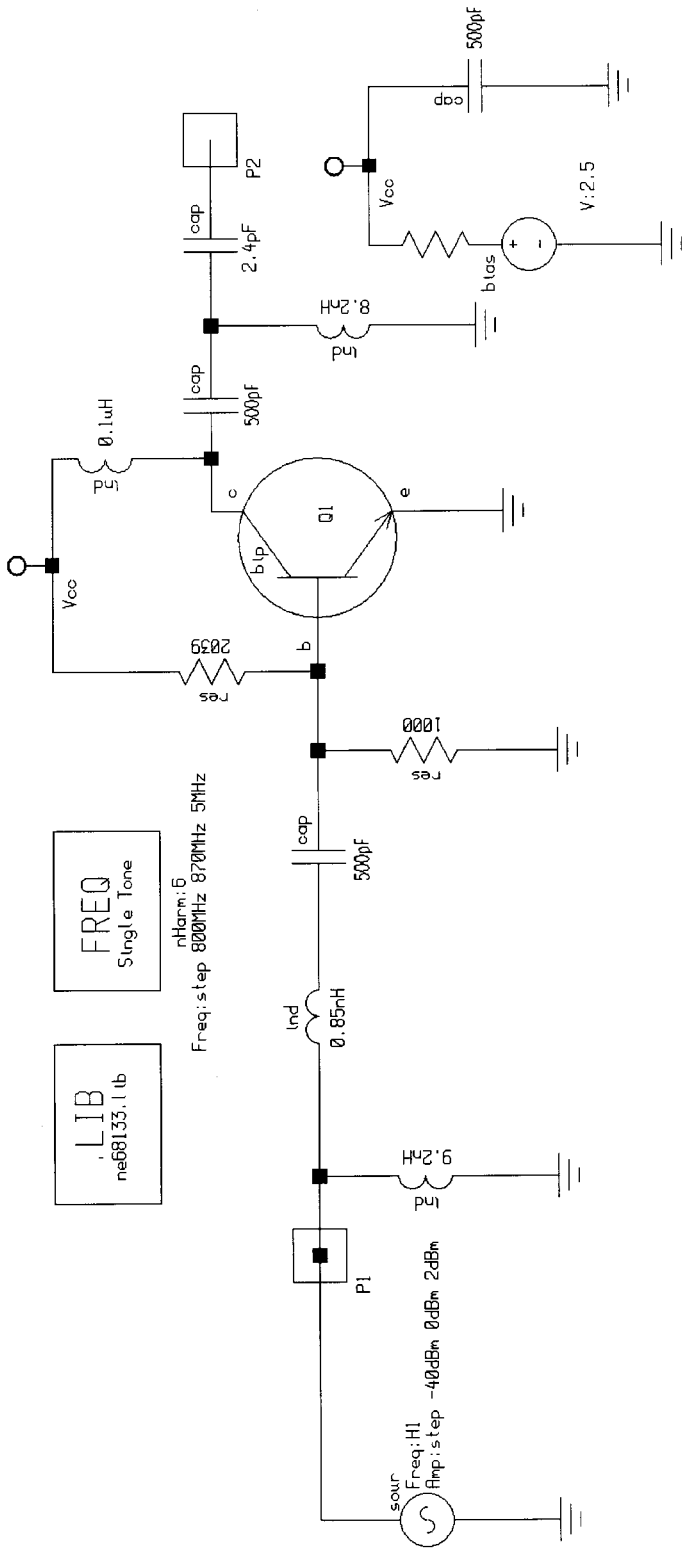


Figure 3-67 The NE68133 low-noise amplifier.

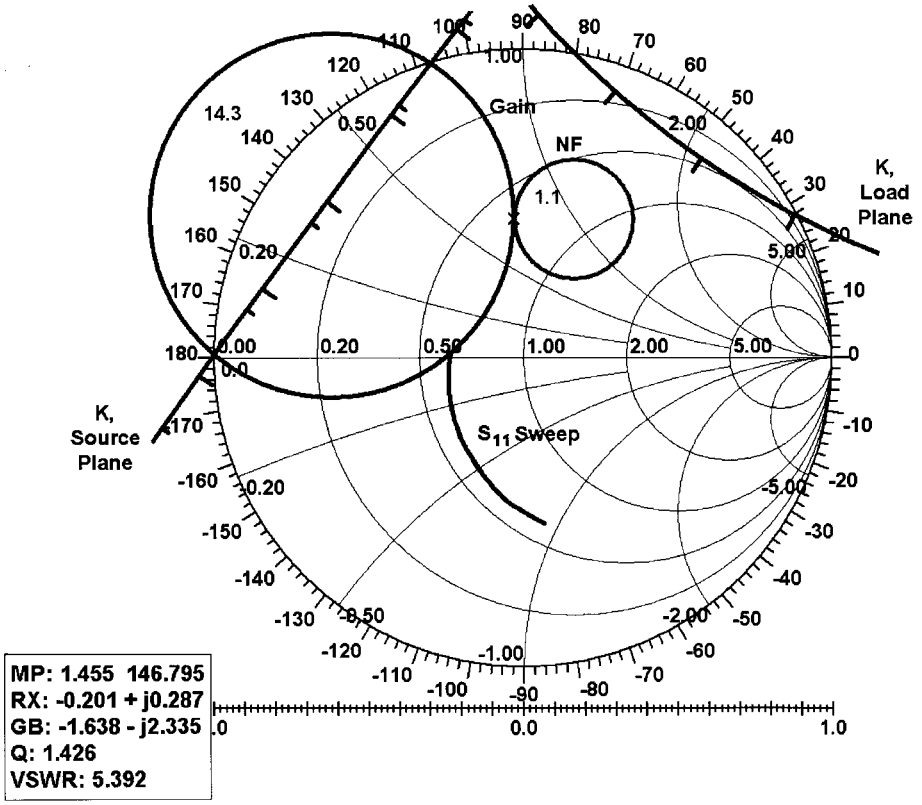


Figure 3-68 Display from the interactive Smith Tool, showing circles for gain, noise figure, and source- and load-plane stability, as well as a sweep of S_{11} versus frequency. In synthesizing an input matching circuit, we choose a single working frequency and work from the chart's origin to the point of tangency between the gain and noise figure circles. See Figure 3-69.

source-plane terminations that correspond to 14.3 dB of available gain. In order to get the required noise figure *and* gain, the input termination must be at the point of tangency of both circles. The arc touching the X axis close to 0.50 is the amplifier's S_{11} response from 800 to 875 MHz. The shallow arc that nearly bisects the 14.3-dB gain circle is actually a portion of the source-plane stability circle; the arc at the top right is a portion of the load-plane stability circle. Because we will be terminating the amplifier input with a value inside the source-plane stability circle, and terminating the amplifier output with a value outside the load-plane stability circle, the amplifier will be unconditionally stable at the frequency at which the match is obtained.

The next task is to obtain the values for the network circuit that matches the input for the 1.1-dB noise figure. A noise matching will not result in the best possible S_{11} matching (Figure 3-69). This translates into an inductor of 9.2 nH to ground and a series inductor of 0.85 nH based on the large-signal model of the NE68133 at the chosen bias point.

After we have determined our input matching network, it is time to find the output matching (Figure 3-70). First, find the termination point on the output plane that corresponds to 14.3-dB gain and 1.1-dB NF at the input plane. The output gain circle goes through the

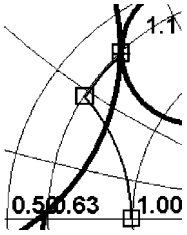


Figure 3-69 Input matching-network extraction, assuming lossless matching components and no parasitic effects.

1.1-dB NF circle at two points, and for reasons of easy matching we selected the lower point. The next step is to find the conjugate reflection coefficient $\Gamma_L = S_{22}^*$. This is the lower square marker in Figure 3-70. The remaining task is to find the matching network that brings this value to 50 Ω. By doing so, we first obtain a shunt inductance $L_{\text{shunt}} = 8.2$ nH and a series capacitor $C_{\text{series}} = 2.4$ pF. Again, these are the values applicable to the large-signal equivalent circuit.

The careful reader will have noticed that we have been avoiding the issue of small-signal exact matching using the manufacturer-supplied parameters. If this is done, the narrow band response shown in Figure 3-71 will be obtained. Therefore, we must optimize the circuit for a better frequency response to obtain the one shown in Figure 3-72.

After this optimization, we were well able to meet the specification of a center band gain of 14 dB and a corresponding NF of 1.1 dB. To show the flatness of the gain, Figure 3-73 shows a gain variation in the range of 0.25 dB.

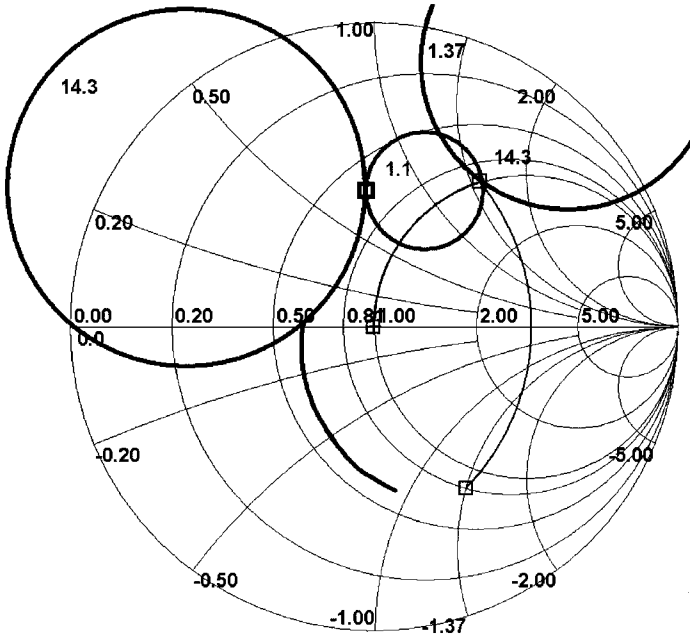


Figure 3-70 Output matching-network extraction.

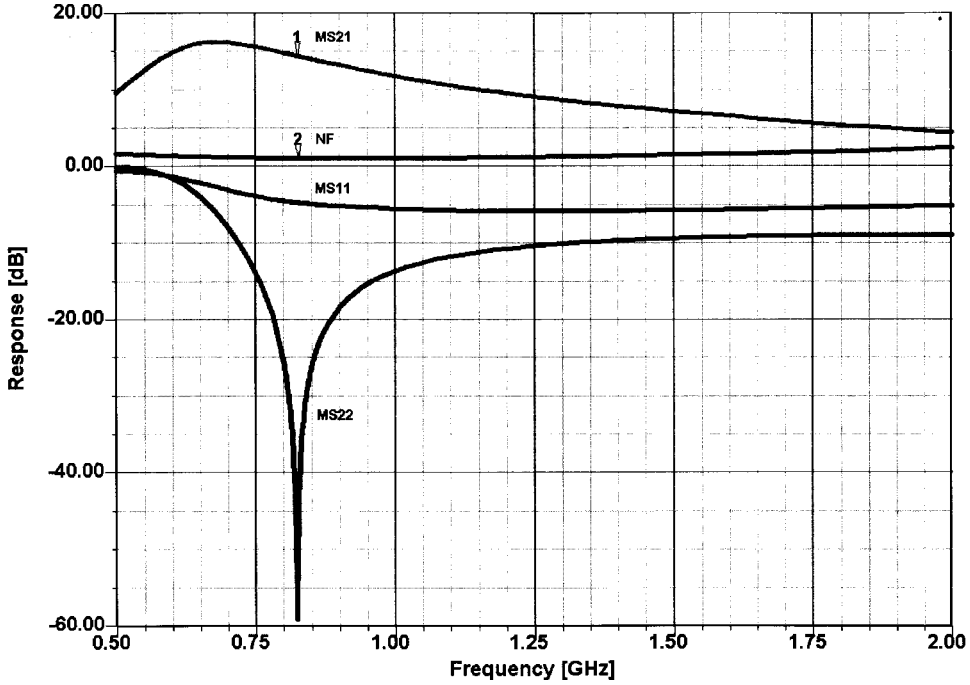


Figure 3-71 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-67.

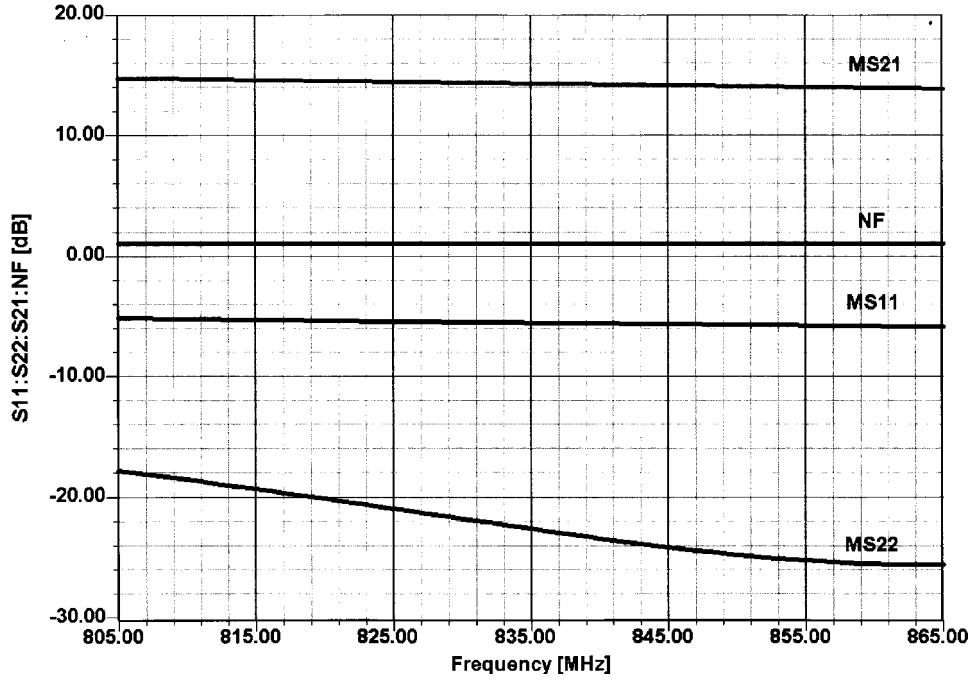


Figure 3-72 Optimized performance of input/output reflection, gain, and noise figure.

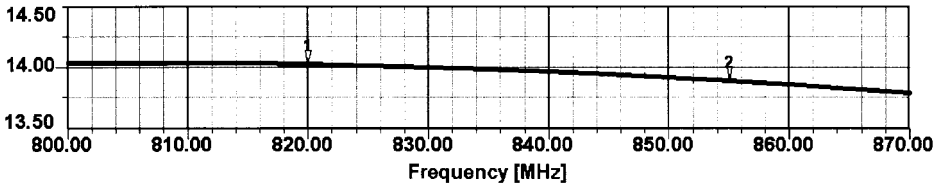


Figure 3-73 High-resolution display of gain as a function of frequency.

The next step is to look into the large-signal performance. The first order of business is to determine the third-order intercept point, which is based on a two-tone analysis. Given the fact that we run the device at low voltage and at a just-reasonable current, the actual intercept point of +13.5 dBm is quite a good number. Figure 3-74 shows the fundamental and third-order outputs. The crossover point between the two, as shown in Chapter 1, is defined as the third-order intercept point. If we had chosen the third harmonic output rather than the fundamental, we would have shown the fifth-order intercept point. The second-order IMD product ($f_1 \pm f_2$) typically can be minimized by input selectivity.

How an amplifier's linearity affects its ability to preserve the characteristics of digital signals is a major concern in wireless design. Figures 3-75 and 3-76 show constellation diagrams with the amplifier handling a PSK signal at drive levels of –40 dBm and +10 dBm, respectively. As expected, we see significant distortion at the higher drive levels due to circuit nonlinearities.

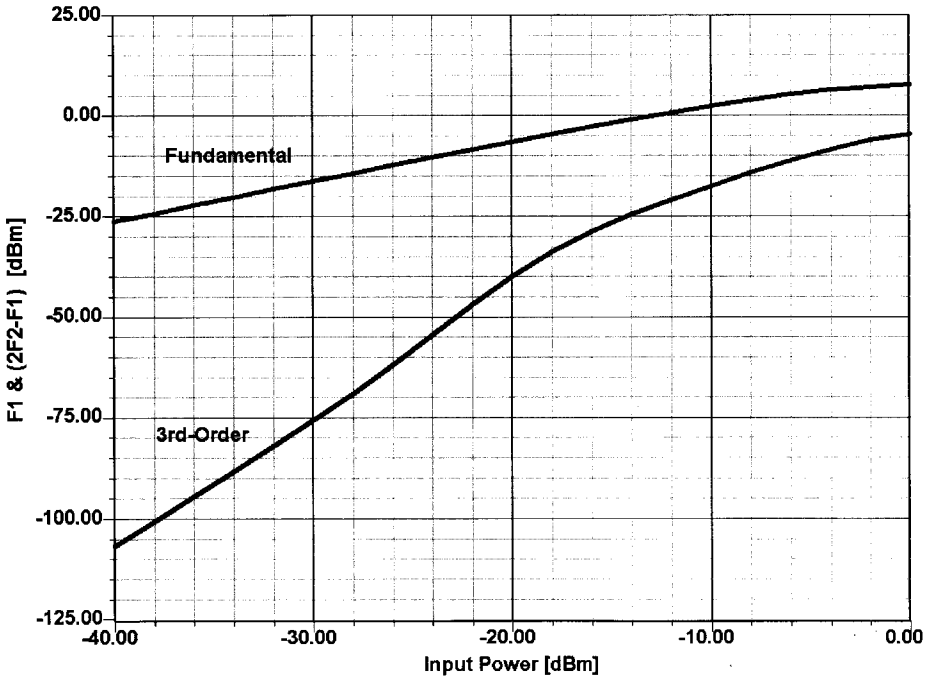


Figure 3-74 Simulated fundamental and third-order IMD outputs of the NE68133 amplifier.

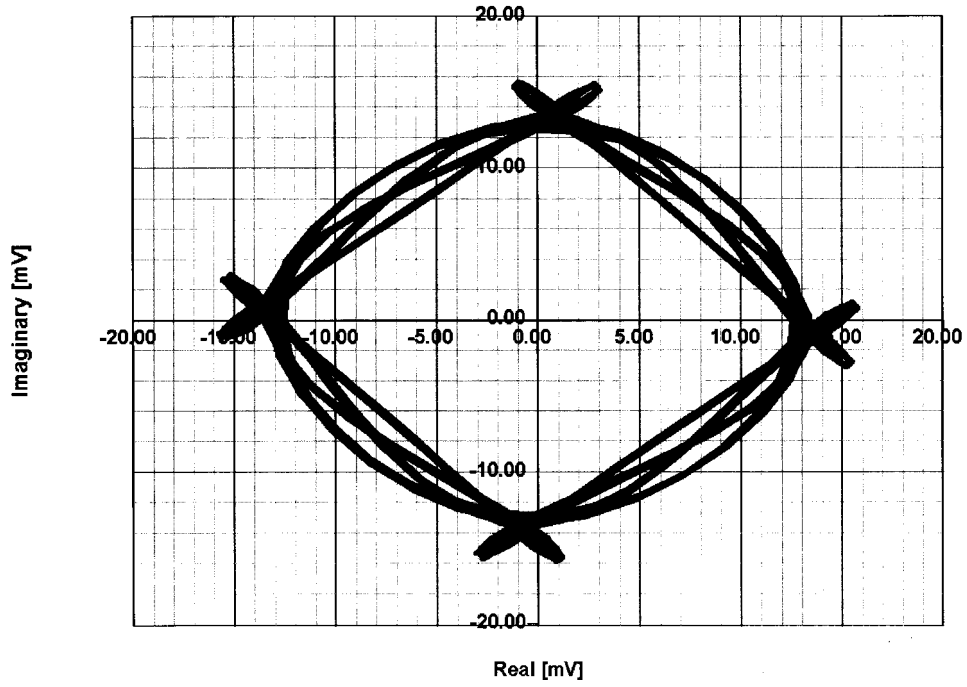


Figure 3-75 Output constellation at –40-dBm input.

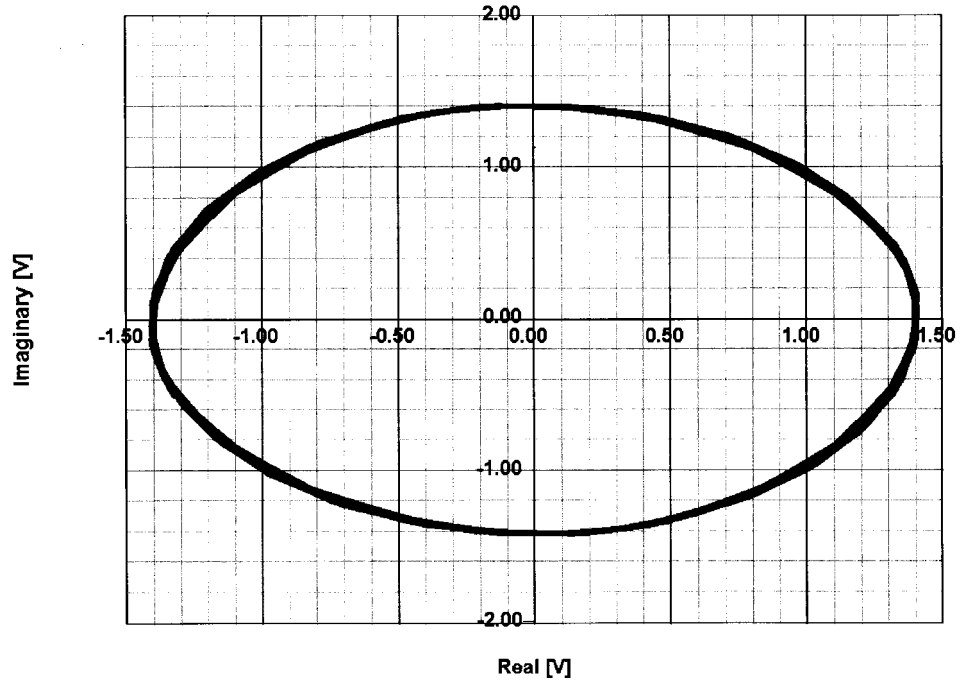


Figure 3-76 Output constellation at 10-dBm input.

BFP420 Matched Amplifier. The NE68133 has an attractive f_T of 8 GHz, but if more gain is needed, one may play with the concept of using a much "hotter" device, such as the BFP420. Recall that the NEC device was unconditionally stable, and our first worry is how the Siemens BFP420 will behave.

A study of Figure 3-77 shows that indeed the NEC device had to be stable all over and therefore was a good basis for a risk-free design. Things change dramatically with the BFP420, for which the K factor is essentially less than 1 up to a corner frequency of 2.75 GHz, where the transistor becomes unconditionally stable. Around 1.75 GHz, the K factor for the NEC transistor barely touches 1. We explained earlier that $G_{\max}$ deviating from the maximum stable gain is a more realistic number, since the conditions for maximum stable gain involve feedback and conjugate matching.

To improve the stability of the BFP420, we will now set out to use the same topology as the previous amplifier but with a resistive feedback between base and collector (Figure 3-78). A good value for this is always somewhere around 200–500 Ω. It must be remembered that this voltage feedback also stabilizes the input and output impedance and reduces the Miller effect because of the feedback being resistive and more dominant.

While the first example covered a fairly narrow frequency band, we are now attempting to cover a much wider frequency range; specifically, from 1.5 to 3 GHz reasonable gain and noise. This, by definition, requires a much higher gain-bandwidth product, and the simple matching network at the input and output provides a nonoptimal input and output matching (Figure 3-79). The feedback introduces some additional noise. This application shows the limits of single-stage UHF/SHF amplifiers with feedback.

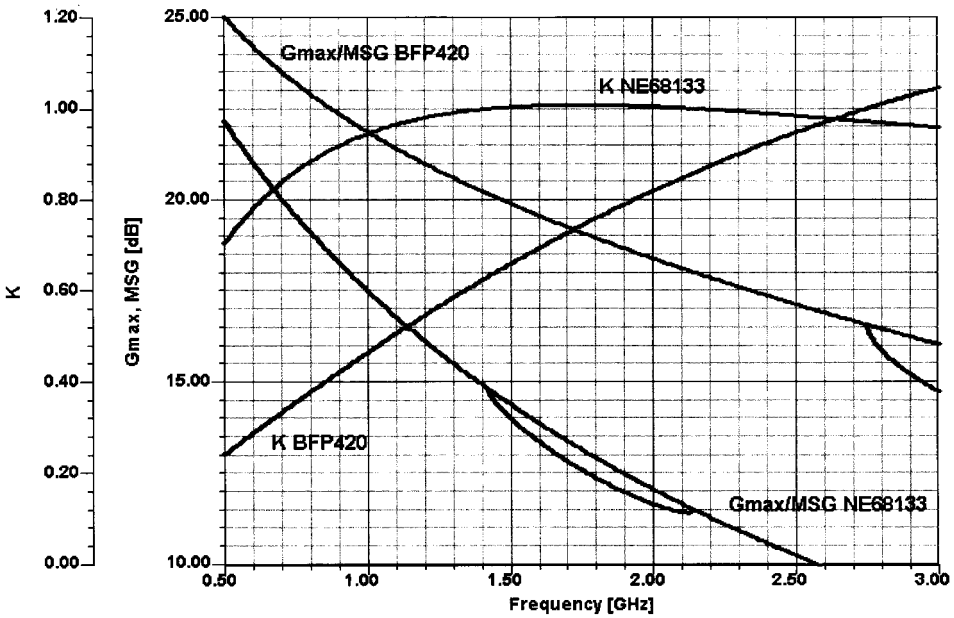


Figure 3-77 Comparison of K, $G_{\max}$, and maximum stable gain (MSG) for the NE68133 and BPF420 BJTs.

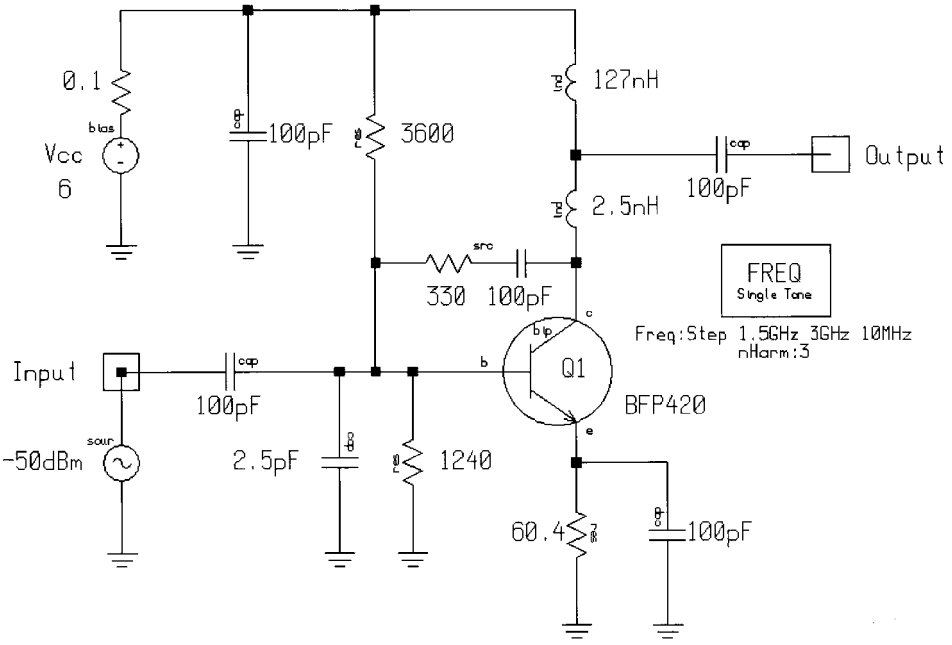


Figure 3-78 BFP420 amplifier with resistive voltage feedback from collector to base.

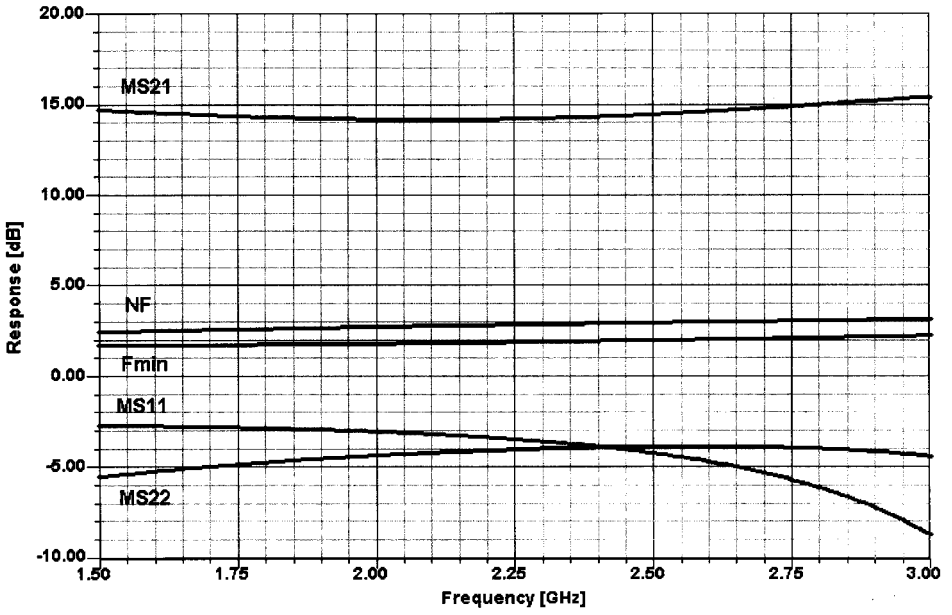


Figure 3-79 Frequency-dependent gain, matching, and noise performance of the BFP420 feedback amplifier.

Narrowband BFP420 Amplifier. Our next attempt will be to design a narrowband input stage that combines very high selectivity with good noise figure. This can only be achieved by using two tuned single-stage resonators with capacitive/magnetic coupling. The following amplifier is tailored to have a center frequency of 900 MHz. The transistor will operate at 12 V/10 mA to have a sufficient dynamic range. The input selectivity is achieved by the use of two top-coupled tuned circuits, followed by a BFP420 (Figure 3-80). The first order of business is to determine the input tuned circuit. The tuning capacitance for the two tuned circuits is arbitrarily set to a manageable 2 pF (small values have too much tolerances and larger values will require smaller inductances, resulting in lower Q). The reactance of the first tuned circuit is

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi \times 9\text{E}8 \times 2\text{E}{-}12} = 88.42\ \Omega$$

The necessary inductance for resonance will be

$$\omega L = 88.42$$

or

$$L = 88.42/\omega = 88.42/(2\pi \times 9\text{E}8) = 15.64\ \text{nH}$$

We will make both tuned circuits identical, but with different loaded Qs. By multiplying 88.42 times the loaded Q (20 for the input resonator and 50 for the output resonator), the resulting parallel-resonant resistance is $88.42 \times 20 = 1768\ \Omega$ at 900 MHz (for the input resonator) and $88.42 \times 50 = 4421\ \Omega$ (for the output resonator). These values are achieved by loading the input with the 50-Ω source and by loading the output with the input of the transistor. Therefore, the input transformation ratio $m_1 = 1768.4/50 = 35.37$. At the output, the equivalent number $m_2 = 4421/50 = 88.42$. For the purpose of simulation, and because it is extremely difficult to find the physical location of the appropriate tap on an air-wound coil for the necessary impedance, we do this "on paper" by incorporating an ideal transformer. We are also going to show later the actual transmission-line-based filter and its modeling. This filter will have a higher insertion loss based on the PC-board material used.

Using our nonlinear simulator, in small-signal ac mode, we obtain the frequency response and noise figure shown in Figure 3-81. In actually building this circuit, there is another problem besides these transformers—the top-coupling capacitor. We have used a grounded-T configuration to obtain a practically realizable value of 36 pF, whose lead inductance will not influence the circuit. (The actual coupling capacitance value for top coupling without the T configuration would have 0.1 pF. We would like to see this value repeatably realized.)

If the actual noise would be calculated relative to the output tuned circuit, the equivalent noise resistor R_n of 5 Ω would now be multiplied by m_2, resulting in 442 Ω. Fortunately, the Serenade simulator has a highly accurate BJT noise model that is both bias and temperature dependent, so the simulator will give us the answer directly.

We have not put any matching circuit at the output and leave this task to the reader. In practice, the slight tilt in the MS_{21} curve in Figure 3-81 would be tuned out. Also note that the minimum noise figure ($F_{\min}$) and the actual 50-Ω noise figure (NF) agree somewhat off center. This has to do with the Miller effect, which results in the well-known difference between noise matching and input-gain matching. By varying the coupling capacitor to move from the overcoupled response, which gives a low noise figure, to a critically coupled

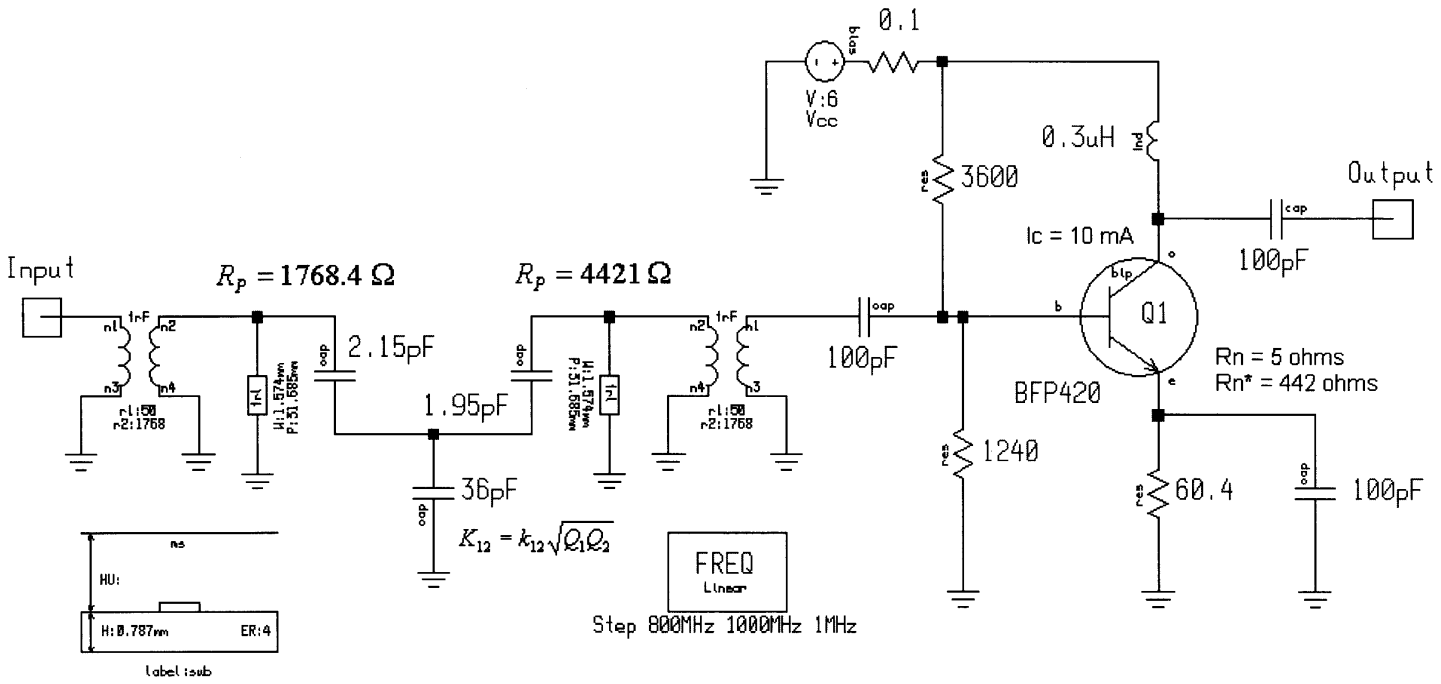


Figure 3-80 BFP420 amplifier with narrowband input filtering. Ideal transformers are used to match the filter's input and output to the input port and the BFP420 base, respectively.

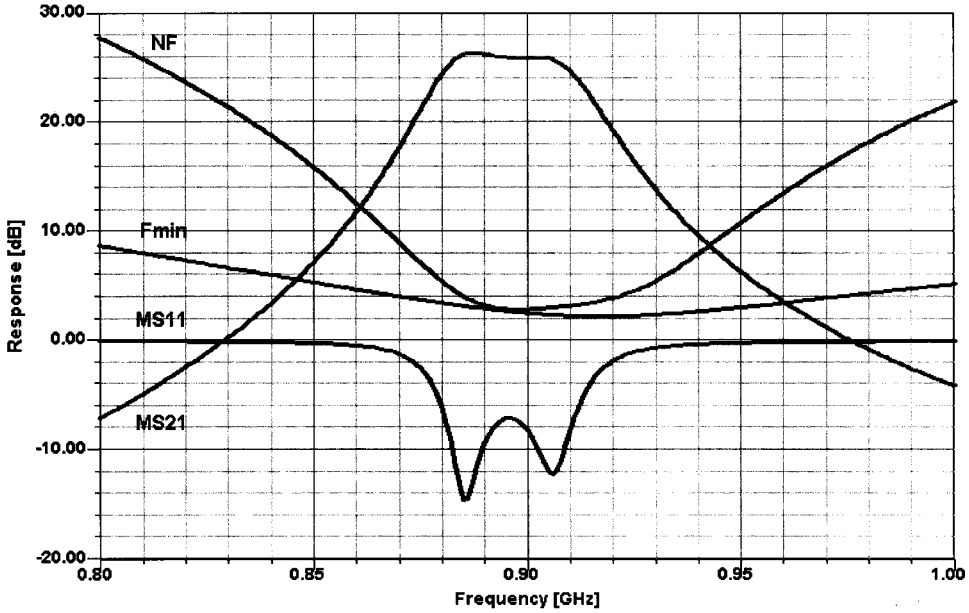


Figure 3-81 Frequency-dependent gain, matching, and noise performance of the narrowband BFP420 amplifier.

response ($K = 1$), we get significantly better selectivity and narrow bandwidth (at the obvious expense of the noise figure) as shown in Figure 3-82.

Since we somewhat compromised by not providing the reader with an actual input and output coupling that can be built, and furthermore, because the inductance in question (15.26 nH) is hard to realize with a high Q, we now show the printed circuit approach for the input filter. By using a Teflon-based material for the PC board, the losses would be less; however, the radiation would be more because ε would go down to 2.1 from the value we used (4). It is also convenient to introduce this filter to show that at higher frequencies, where lumped elements are really no longer available, one has to switch to distributed elements. A list of the commonly used distributed elements will be shown later; the transmission line and T junction are probably the most frequently used elements, followed by spiral and rectangular inductors, and bends.

The actual implementation of the input filter on a printed circuit board requires distributed elements (Figure 3-83). In practice, one has two parallel coupled lines with a tap somewhere toward the ground connection and, for modeling accuracy, we have to introduce a T junction with the appropriate impedance of 50 Ω for the input and a different impedance (if necessary) for the output. The coupling mechanism now is magnetic coupling, since no one can physically change the distance between the transmission lines by tuning some mechanical elements, one has to get it right the first time. This can only be achieved with a high-precision simulator. The reason why these simulators are so expensive is the investment of research time and personnel to produce valid models over a huge frequency and impedance range. Since we are now curious about the frequency response, we have plotted in Figure 3-84 the frequency response, which shows a higher (expected) insertion loss. The PC-board material is partially responsible for this.

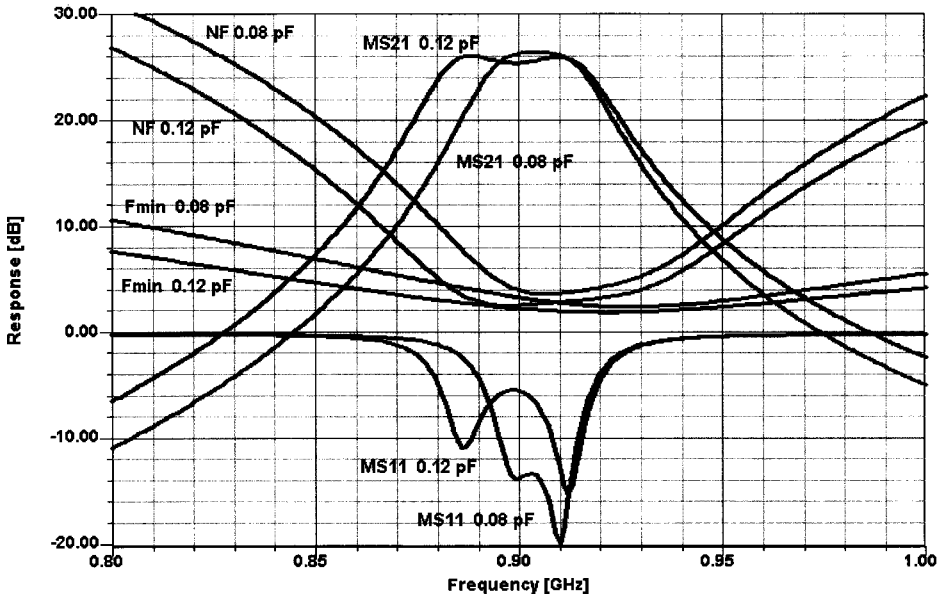


Figure 3-82 Comparison of filter responses with critical coupling (0.08-pF response) and overcoupling (0.12-pF response).

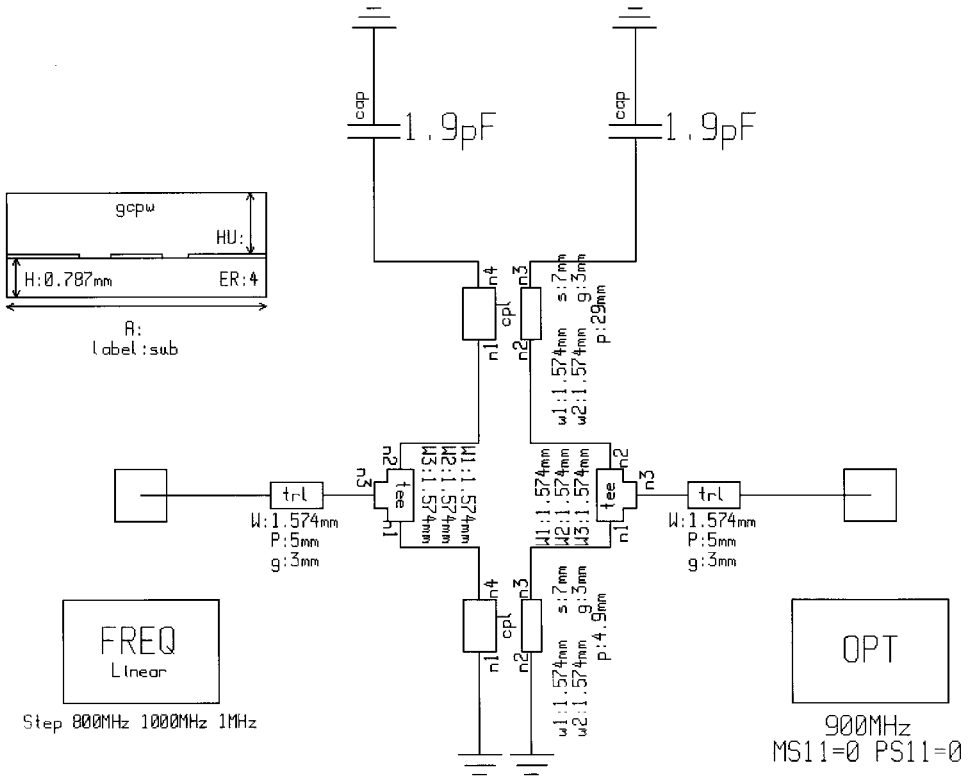


Figure 3-83 Input filter using distributed-element resonators and matching.

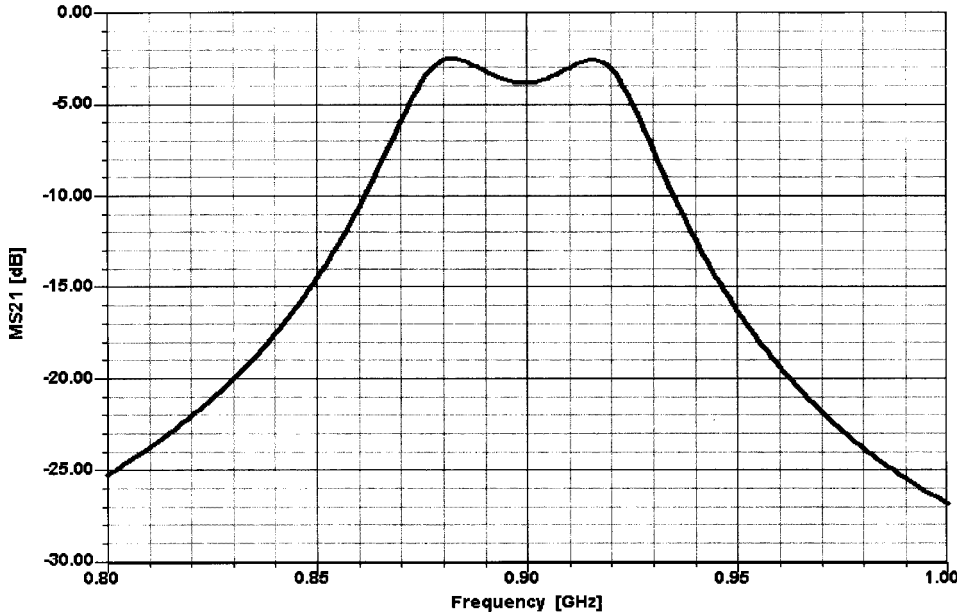


Figure 3-84 Gain versus frequency response for the distributed-element filter.

At the end of this chapter we will have a chance to discuss passive elements as needed for microwave frequencies, and one should consider frequencies above 1500 MHz as microwave frequencies.

GaAsFET Feedback Amplifier. The implementations we have shown so far have been based on bipolar transistors. Figure 3-85 shows a single-stage feedback amplifier using a GaAsFET. The feedback is based around an *LRC* network and some attempts at matching at the input and output are made. Figure 3-86 shows its frequency-dependent gain, matching, and noise performance.

3-2-3 High-Gain Amplifiers

In most cases, high-gain amplifiers receive an already amplified signal, which means we are less concerned about the signal-to-noise ratio but remain concerned about distortion. To reduce distortion, we have two options: (1) increase the dc current in the active device, which always improves the current-related distortion; and (2) add some voltage feedback, which reduces the voltage distortion. In simple terms again, only current feedback can reduce the current-related distortion of a square-law or exponential transfer characteristic and only voltage feedback can reduce the effect of overdrive as a function of the load, but only to some degree. Its main purpose remains to reduce the input impedance to useful values, specifically for FETs, since current feedback increases the input and output impedances. Some ultrasensitive circuits still require Class A operation, but now at higher current levels; these are typically CATV applications, and there are several transistors available that can be operated at 5–15 V but at 60 mA or higher. Because of the higher power consumption, this excludes handheld cordless or cellular telephones. The same circuit as used before (Figure 3-85) for

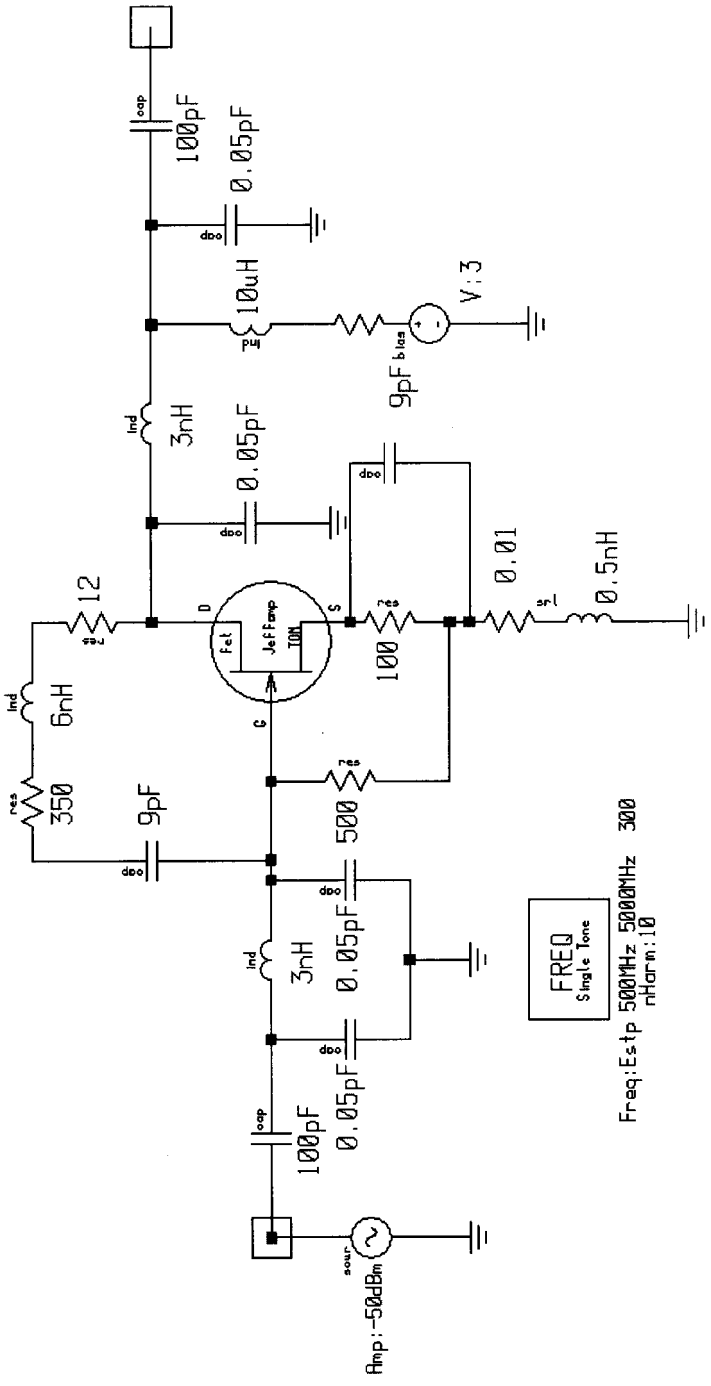


Figure 3-85 GaAsFET feedback amplifier.

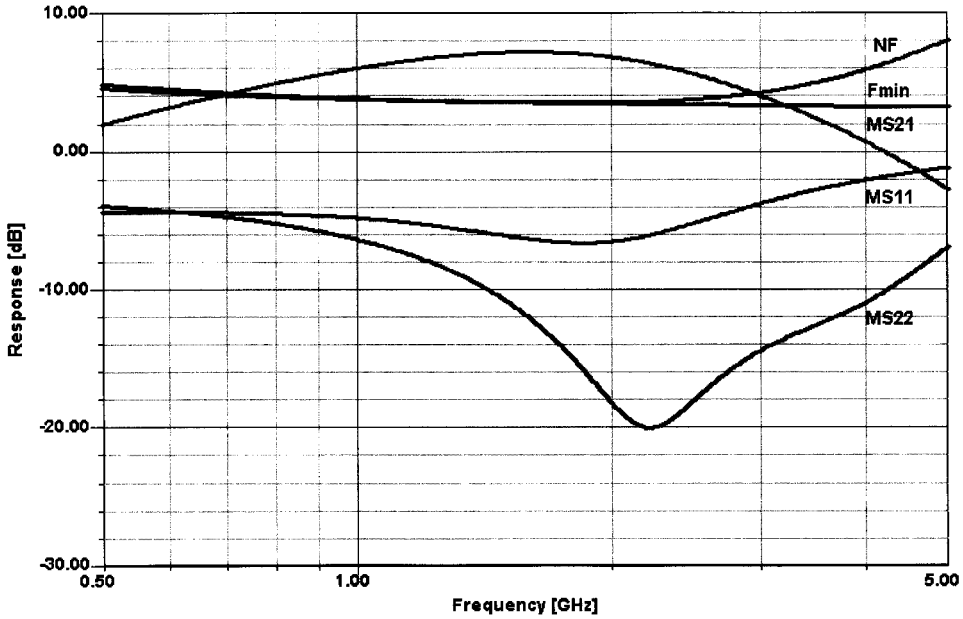


Figure 3-86 Frequency-dependent gain, matching, and noise responses for the GaAsFET amplifier.

Class A operation now can be operated at a higher current, which automatically increases the gain and the noise figure and reduces distortion. This is based on the fact that the transconductance (g_m) increases at higher currents. The second transistor in the TI circuit achieves this by being in reality a larger device, which could be modeled by several transistors, like the first stage, in parallel. This automatically increases the dc current by the scaling factor. The same is even more true for the output stage. Likewise, bipolar transistors can also be scaled. A visual inspection of some transistors in their packages will reveal that they are really not just one transistor, but several operating in parallel, each with its own ballast resistors (for bipolar transistors) or other means (for FETs).

Using tuned circuits rather than wideband matching, it is possible to go to Class AB or B operation. Class AB is really a hybrid between A and B, and Class B is probably better known. Figure 3-63 shows the Class B operating point. Most Class B stages are operated in push–pull because, together with the tuned circuit, one-half of the cycle is supplied by the other transistor. The definition of Class B operation is a conducting angle of 90° (see Figure 3-63). At this point it should be noted that some older texts define the conducting angle as being twice the value we have defined in this chapter.

Without going into further details, the highest efficiency in Class B operation is 78.5%. Based on the fact that modern wireless applications generally use a hybrid of amplitude and angle modulation, pure Class C operation finds little application today because of the severe nonlinear distortion it introduces. A Class B stage has less "power gain" than a Class A stage but provides much more output power. We have already mentioned the problem associated with characterizing a Class A stage in terms of power gain. That's the only time Kirchhoff's equations are valid, meaning that the sums of all currents and voltages have to be zero. Because Kirchhoff's equations do not consider nonlinearities and harmonic contributions,

the very moment we move into considering medium- and higher-power amplifiers, the nonlinearities will dominate and the performance of the stage is best analyzed with a high-quality circuit simulator with the appropriate dynamic range and good modeling capability. Without question, this applies to Class C operation also. As pointed out previously, the high-gain amplifier is optimized for gain rather than noise figure and typically requires more collector or drain current. This can easily be seen by rebiasing the FET amplifier as shown above to a significantly higher value of 33 mA. This is accomplished by changing the source resistor down to 3 Ω. As a result of this, we obtain the gain, matching, and noise figure responses shown in Figure 3-87.

It may be a surprise to the reader that the actual noise figure now is less than that shown in Figure 3-86. This device is now operating at 33 mA; the reason for this has to do with the fact that if the transconductance is too low, the noise figure also increases, and the NF curve for the FET has a similar-looking current-dependent minimum—a curve that is frequently not supplied by the manufacturer. This amplifier, now operating at 33 mA, should have quite good large-signal capabilities; we will examine this shortly.

Another interesting piece of information is the display of the dc I–V curves with the "load line." Since we have reactances in the circuit, we do not see a line but rather an ellipse, indicating the storage of energy. (The actual load line would be a line through the two foci of the ellipse.) This curve allows us to see immediately that at this operating point the transistor is neither going into current saturation nor becoming voltage limiting. The input power can be increased further until such a condition exists. Figure 3-88 shows the dc I–V curve for the amplifier shown in Figure 3-85.

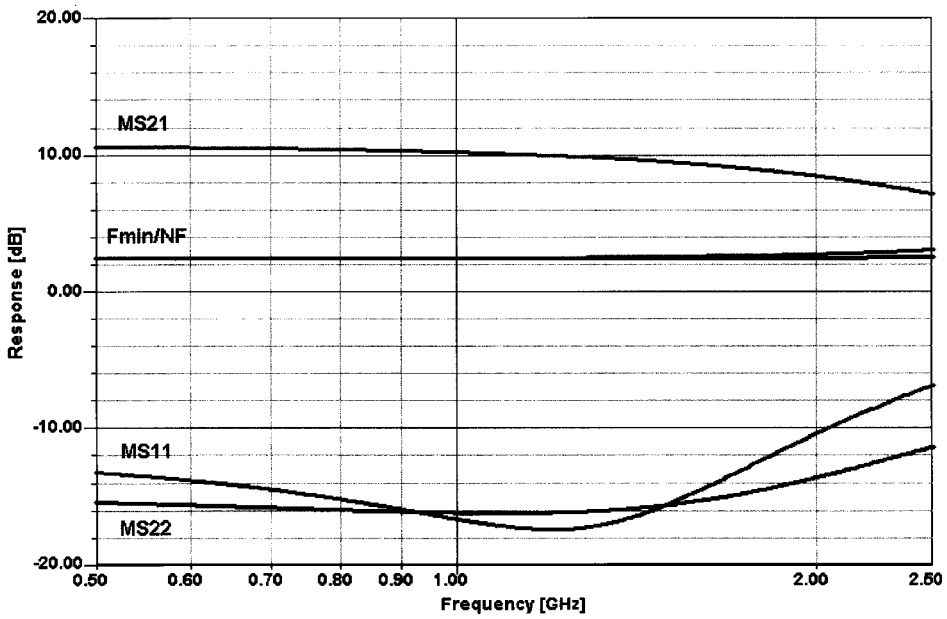


Figure 3-87 Frequency-dependent gain, matching, and noise performance for the GaAsFET amplifier of Figure 3-85 rebiased for a drain current of 33 mA.

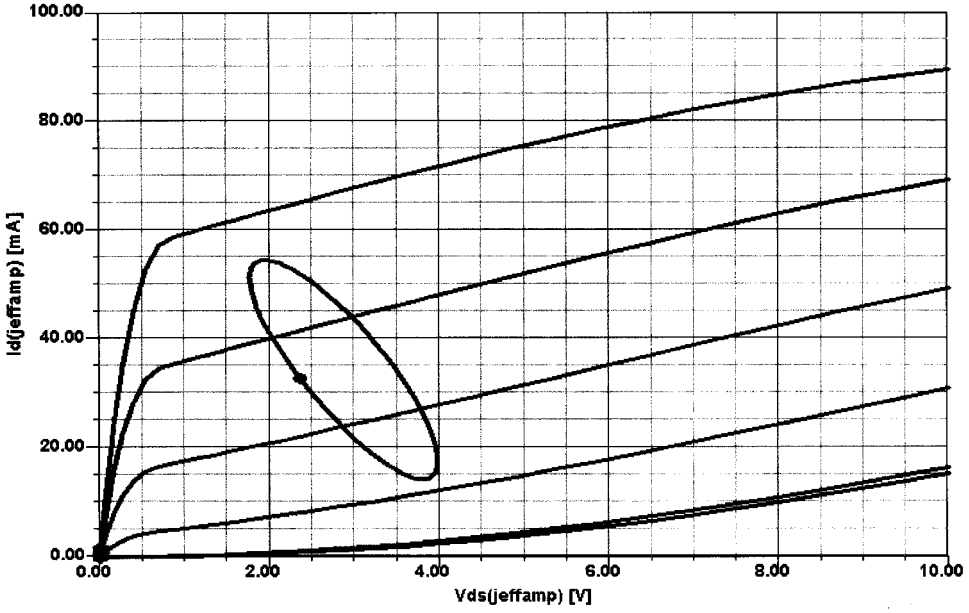


Figure 3-88 The dc *I–V* curves for the GaAsFET operating at I_D = 33 mA, including the ac load line. The load line is an ellipse rather than a straight line for the reason described in the text.

Besides looking for the adjacent-channel power ratio (ACPR), the other traditionally important points are third-order intercept point and the equivalent of a multitone simulation, as used for CATV application but relevant to the multisignal environment of wireless applications. Figure 3-89 shows the determination of the amplifier's third-order intercept point. It is most important for the reader to understand that very low values for input and output level have to be used to determine the crossover point. If this advice is not followed, totally erroneous numbers can occur.

Since an antenna supplies a large number of signals (this is true of handheld radios as well as base-station sites), a three-tone standard has been adopted as a good approximation of multisignal operation. This standard is in accordance with the German DIN 45004E,3.3 three-tone measurement. The requirement is that the intermodulation distortion products at the output are better than 60 dB suppressed relative to the largest of the three tones. One tone is used as the reference, and the other two tones are 6 dB less. This measurement has been derived from television, whereby the video transmitter, based on the vestigial sideband power, is 6 dB above the two sidebands produced by the FM sound transmitter using a modulation index of 5. The result of a three-tone test can be seen in Figure 3-90. This amplifier clearly meets the DIN specification for linearity at input levels up to –9/–15 dBm at the input.

Bipolar transistors are also quite popular for this purpose. Figure 3-91 shows a single-stage amplifier with an intrinsic transistor and all the other "hidden" intrinsics, normally not visible, added. The three diodes belong to those intrinsic elements, which are needed for accurate microwave modeling but which are not provided by the standard nonlinear models. The reader will see that we went into great detail to model all the necessary parts. This feedback amplifier, whose frequency response is shown in Figure 3-92, has been quite

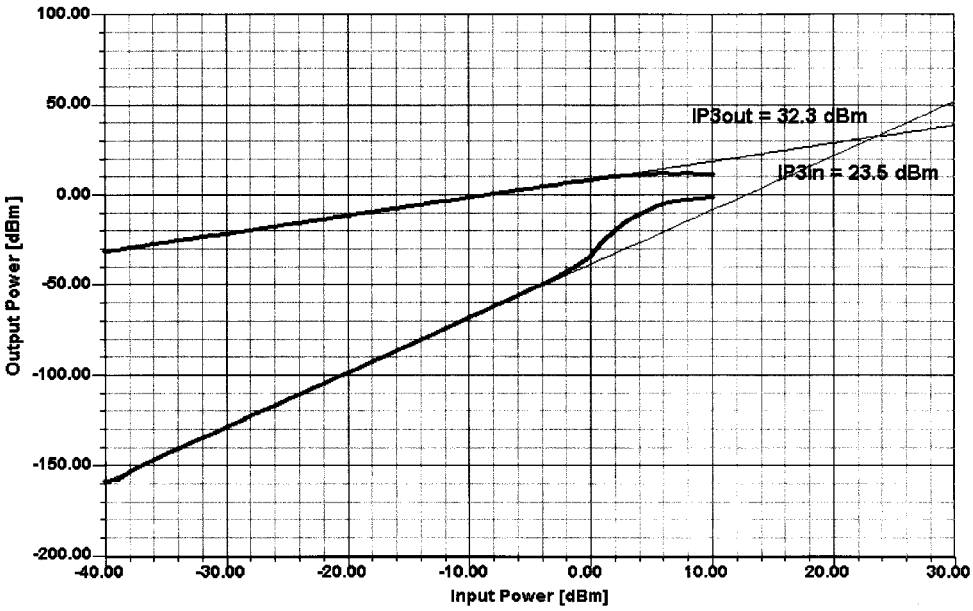


Figure 3-89 Determination of the third-order intercept point based on a two-tone measurement using tones spaced at 5% of the operating frequency (1.8 GHz).

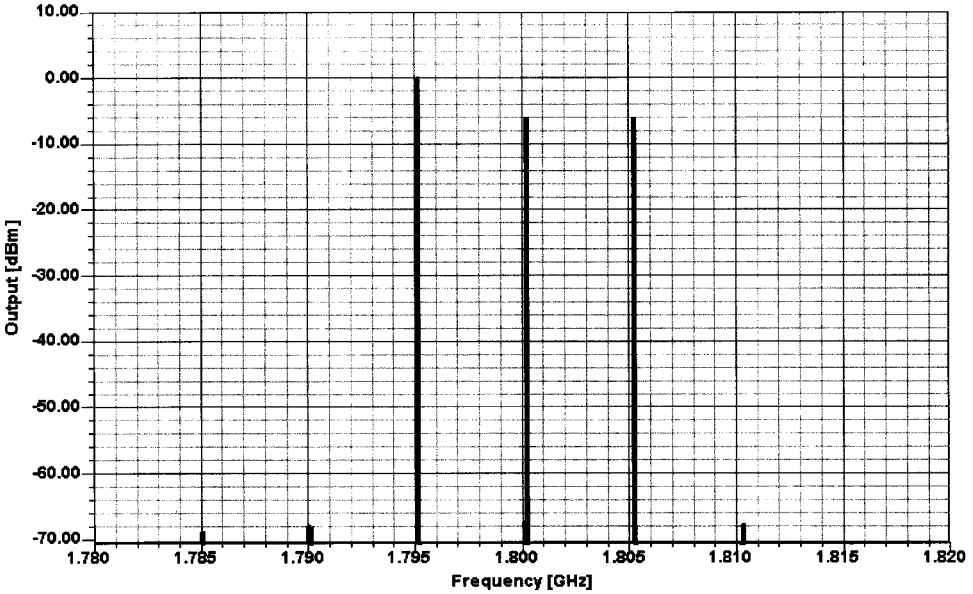


Figure 3-90 Simulated three-tone analysis of the GaAsFET amplifier with $I_D = 33$ mA.

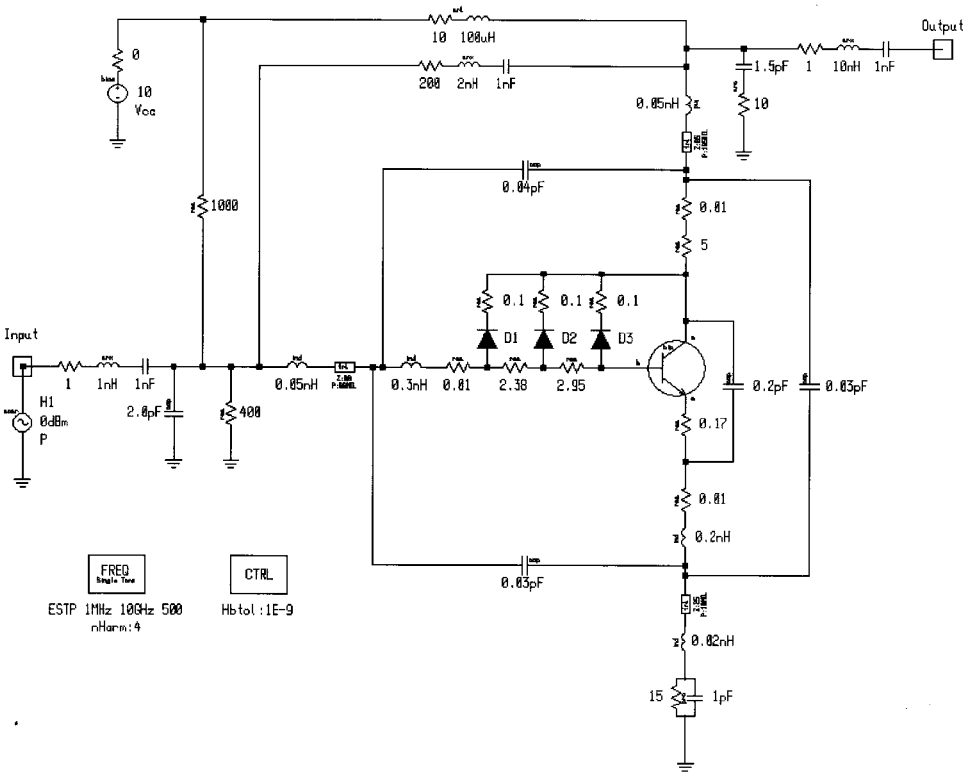


Figure 3-91 A high-linearity single-transistor amplifier modeled in great detail.

difficult to design, but there are monolithic circuits available where the higher gain–bandwidth product is being used to advantage.

Figure 3-92 shows the predicted frequency-dependent gain, matching, and noise figure at a bias point of 108 mA at 10 V. Using the same DIN three-tone test procedure, Figure 3-93 shows the expected good result.

Since this test requires two equal signals and one larger signal, we need to expect one IMD product larger than the rest. We have set our window for a dynamic range of 70 dB, and yes, in the lower left corner, we can see the product between the first two signals, while the other products are below the –50 dB level.

As a little task, Figure 3-94 shows a simple amplifier, based on a modern Siemens transistor, which is a feedback amplifier with fewer modeling details surrounding the transistor. Its associated frequency-dependent performance is shown in Figure 3-95. Interested readers are invited to improve its performance by using their skills and intuition.

Permitting higher currents and actually combining two transistors leads to our first MMIC. The following simulation is based on the MGA64135. Needless to say, there are some compensation networks inside the MMIC and distributed elements that affect the frequency response. However, with this arrangement, it is possible to obtain an appreciable gain at our wireless frequencies. The noise figure, based on the feedback, however, it not too impressive. This is a problem that occurs with most feedback amplifiers, which try to minimize the distortion and yet would like to keep the noise figure down. Figure 3-96 shows the schematic

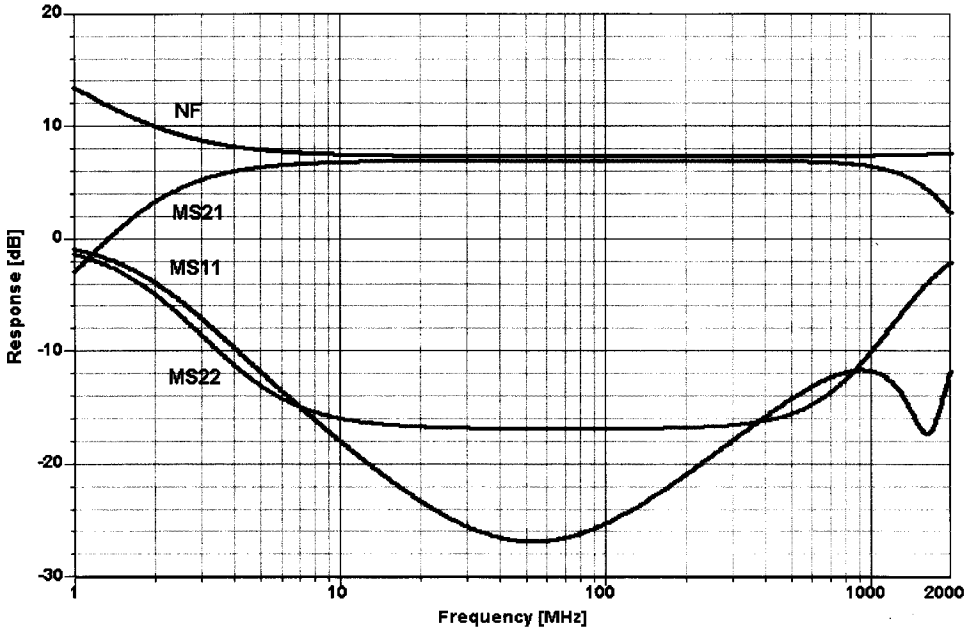


Figure 3-92 Simulated frequency-dependent gain, matching, and noise figure performance of the circuit in Figure 3-91.

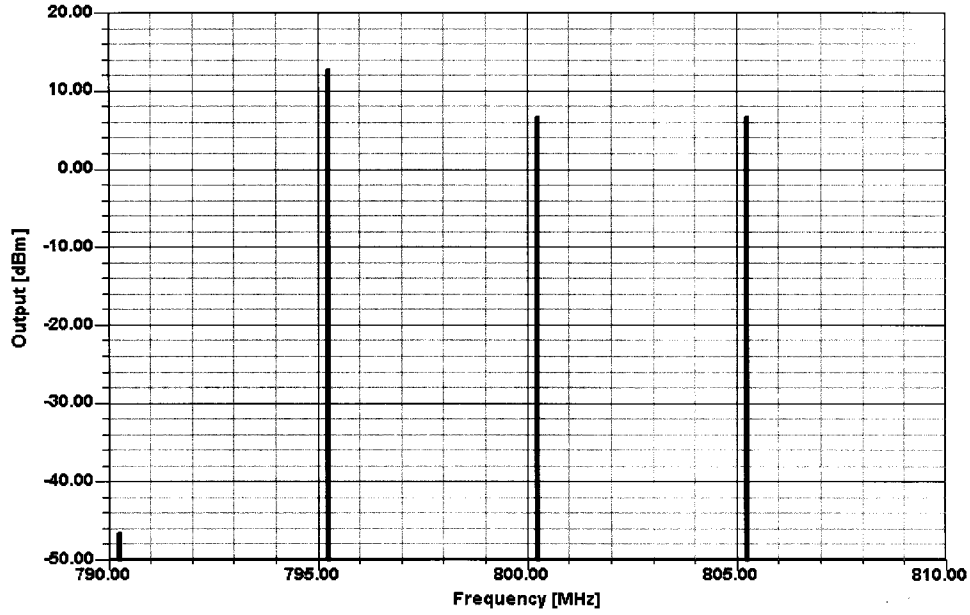


Figure 3-93 Simulated three-tone output spectrum of the amplifier in Figure 3-91.

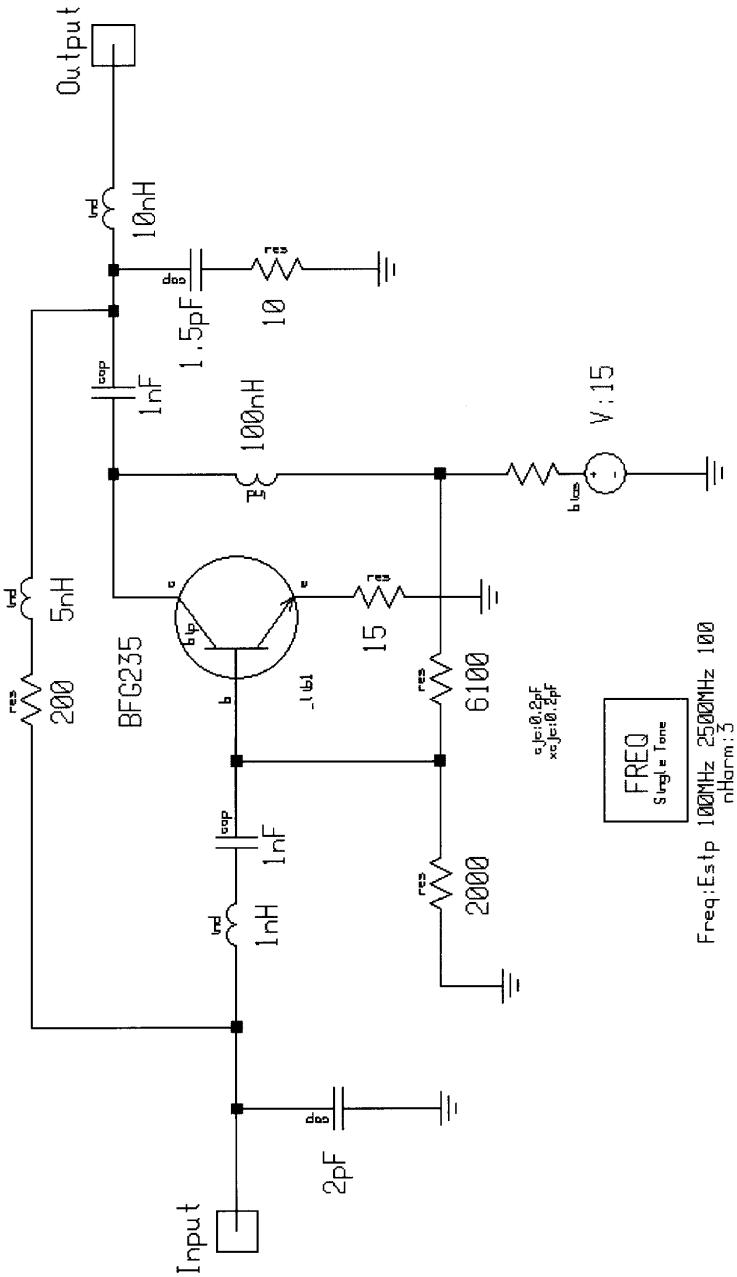


Figure 3-94 A high-linearity amplifier using the Siemens BFG235.

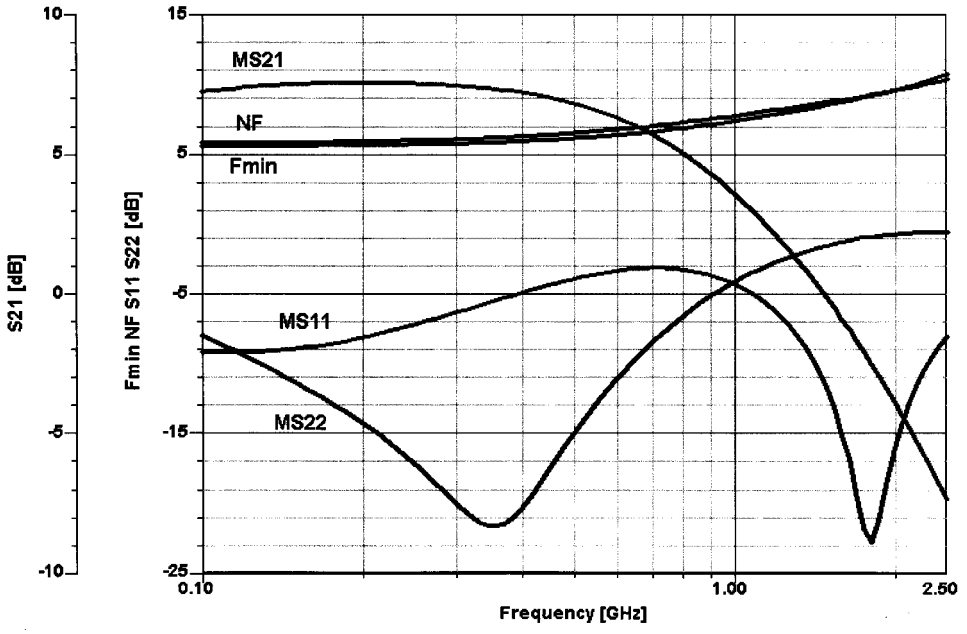


Figure 3-95 Frequency-dependent gain, matching, and noise performance of the amplifier in Figure 3-94.

of the GaAs MMIC; Figure 3-97 shows its frequency-dependent gain, matching, and noise performance; Figure 3-98 shows its two-tone response, which is used to obtain the intercept point ($IP_{3,in}$ = 8.5 dBm, $IP_{3,out}$ = 22 dBm); and, finally, Figure 3-99 shows the three-tone test results at an input RF level of –21/–27 dBm. While the gain–bandwidth product is considerably larger than the single-stage amplifier we examined before, the intercept point and three-tone performance do not reach the same values. As a general rule, the dynamic

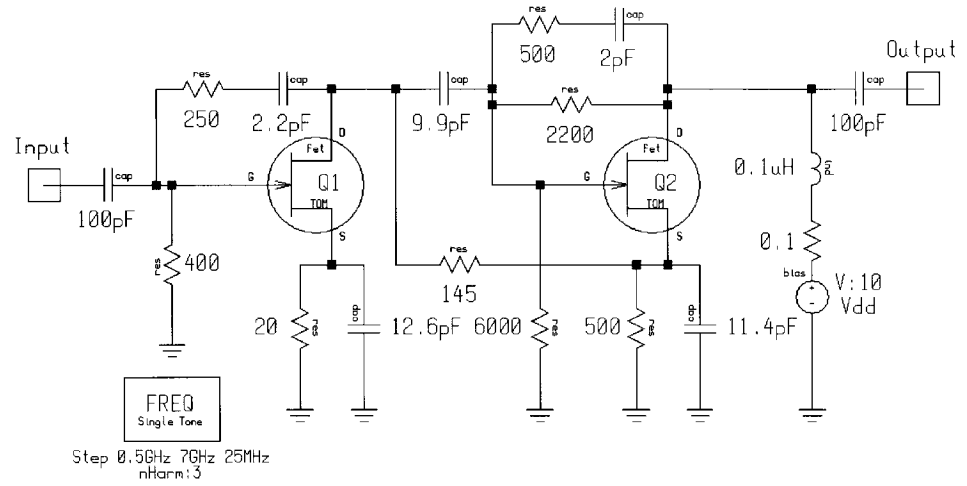


Figure 3-96 Simulation schematic for the MGA64135 MMIC amplifier. Q1 runs at 17.4 mA; Q2 at 32.2 mA.

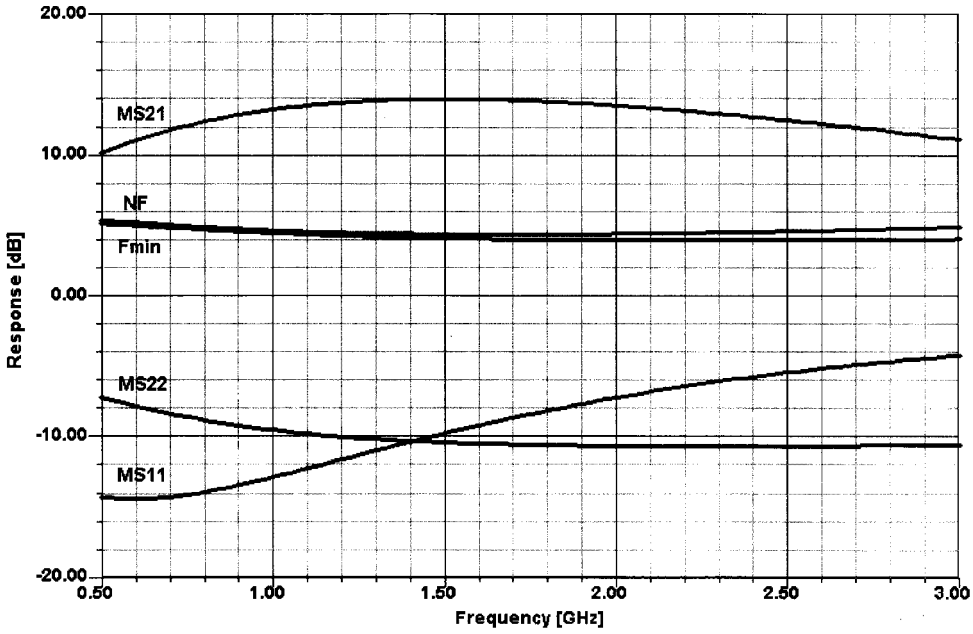


Figure 3-97 Frequency-dependent gain, matching, and noise figure performance for the MGA64135 MMIC amplifier.

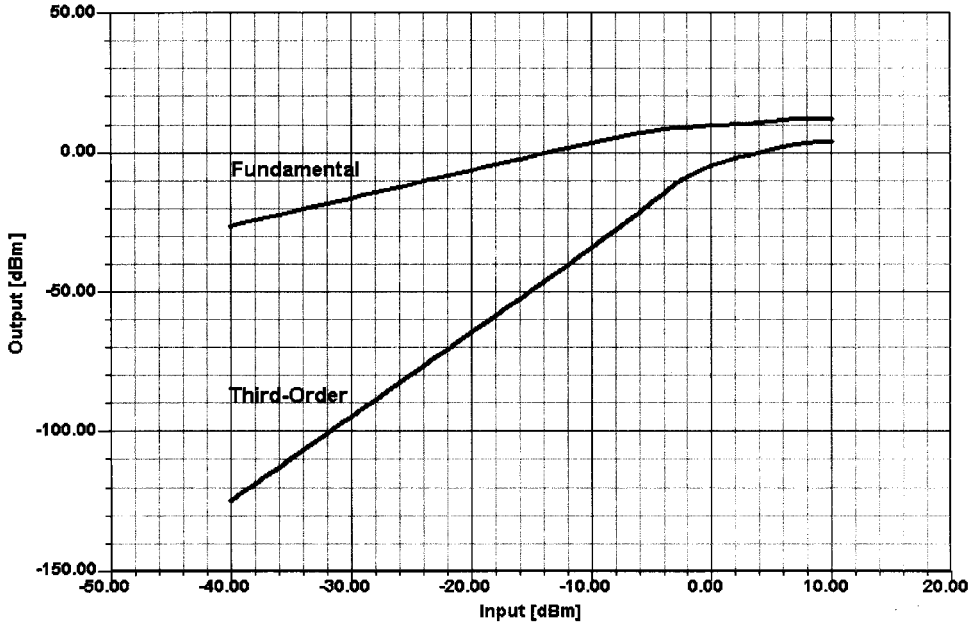


Figure 3-98 Simulated two-tone testing of the MGA64135 MMIC amplifier. $IP_{3,in}$ = 8.5 dBm and $IP_{3,out}$ = 22 dBm.

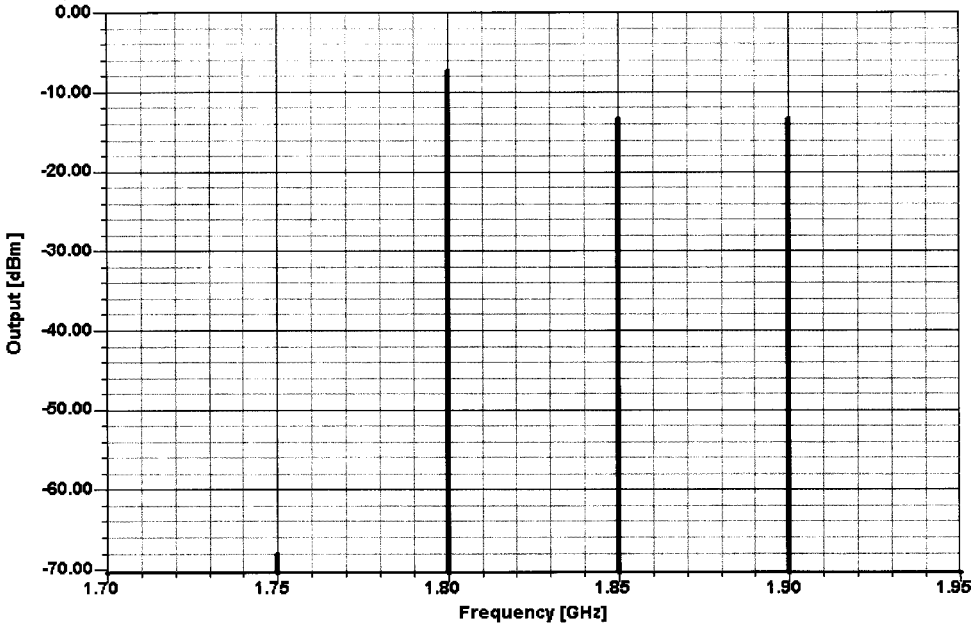


Figure 3-99 Simulated three-tone test of the MGA64135 MMIC. The test tones were 1.80 GHz (–21 dBm), 1.85 GHz (–27 dBm), and 1.90 GHz (–27 dBm). The 1.75-GHz IMD product is down 60.5 dB relative to the 1.80-GHz signal.

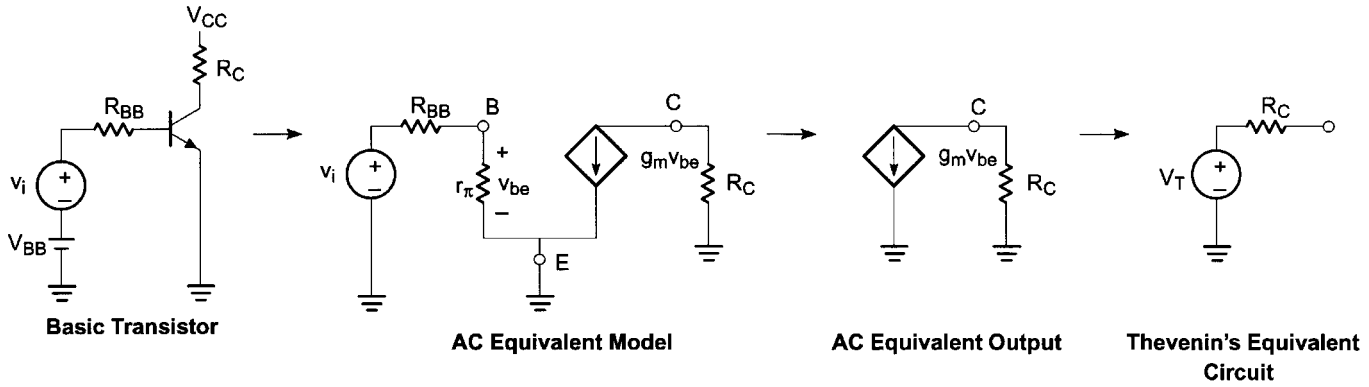
range is directly proportional to the power dissipated in the devices: specifically, the current. As explained earlier several times, the first-order nonlinearity comes from the voltage-to-current transfer characteristic, such as the transconductance itself, and can only be improved by a combination of resistive feedback in the emitter/source and by high currents. At these dc levels, the base–emitter (gate–source) junction gets linearized because the differential values are getting smaller compared to the parasitic loss resistances, including external dc/RF feedback. Once again, the voltage feedback reduces the voltage clipping but is initially more responsible for the input and output matching.

3-2-4 Low-Voltage Open-Collector Design [22–25]*

Why R_C Acts Like a Source Resistor. An open-collector output allows a designer the flexibility to choose the value of the R_C resistor. Choice of this resistor value not only sets the dc bias point of the device but also defines the source impedance value. Figure 3-100 shows the ac model of the transistor. Converting the output structure by applying a Norton and Thevenin transformation, one can conclude that R_C becomes the source resistance. Thus, by choosing R_C to be equal to the load, maximum power transfer will then occur.

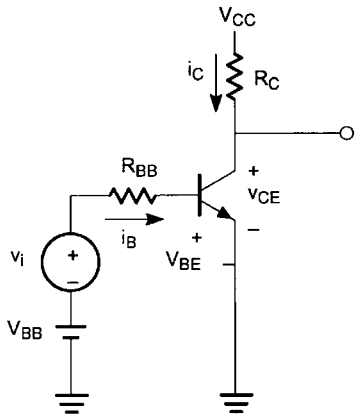
Figure 3-101 shows an active transistor with collector and base resistors. From basic transistor theory, Eq. A in Figure 3-101 is generated and has the same form as the general equation for a straight line ($y = mx + b$).

*Based on portions of the Philips Semiconductors/Signetics RF Communications Products Application Note AN1777, "Low-Voltage Front End Circuits: SA601, SA620," August 20, 1997. Used with permission.



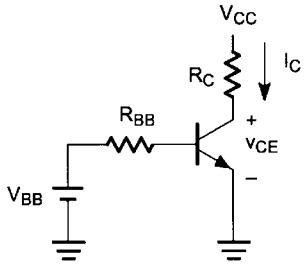
Symbol Convention:
DC: represented by uppercase symbols with uppercase subscripts, i.e., I_C
AC: represented by lowercase symbols with lowercase subscripts, i.e., i_C
Combined AC & DC: represented by lowercase symbol with uppercase subscript, i.e., i_C

Figure 3-100 *Rc* as source resistor.



DC Model

- Kill AC sources
- Open capacitors
- Short inductors

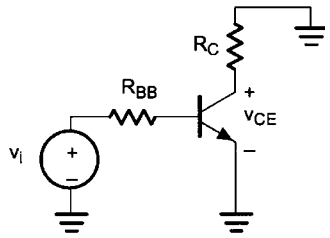


$$I_C = \left(\frac{-1}{R_C}\right) V_{CE} + \frac{V_{CC}}{R_C}$$

Eq. A

AC Model

- Kill DC sources
- Short capacitors
- Open inductors



$$i_C = \frac{-1}{R_C} (v_{CE})$$

Eq. B

Figure 3-101 Basic transistor analysis.

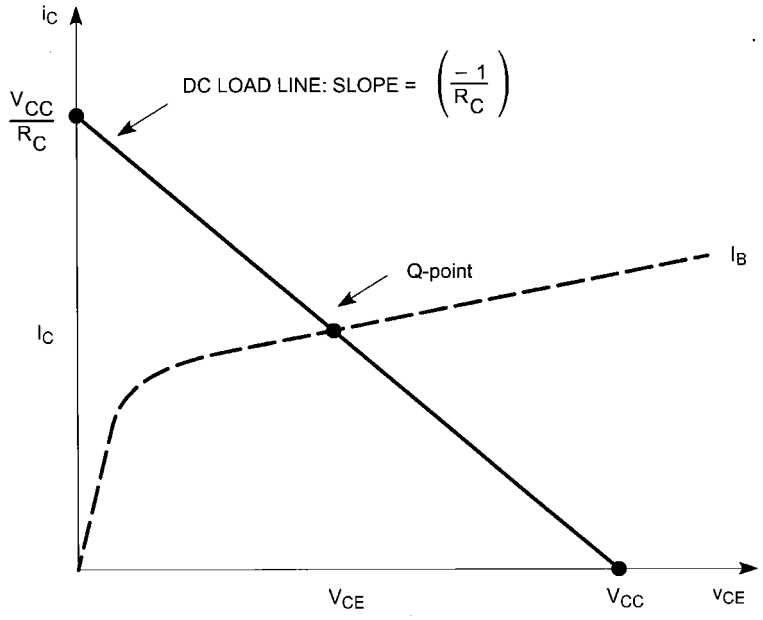


Figure 3-102 Load line and Q-point graph.

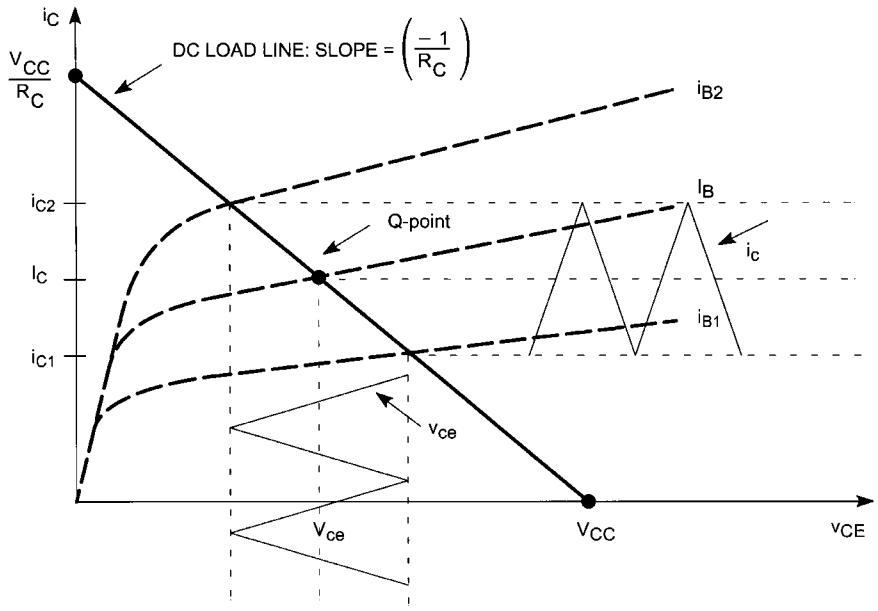


Figure 3-103 Graphical analysis for the circuitry in Figure 3-101.

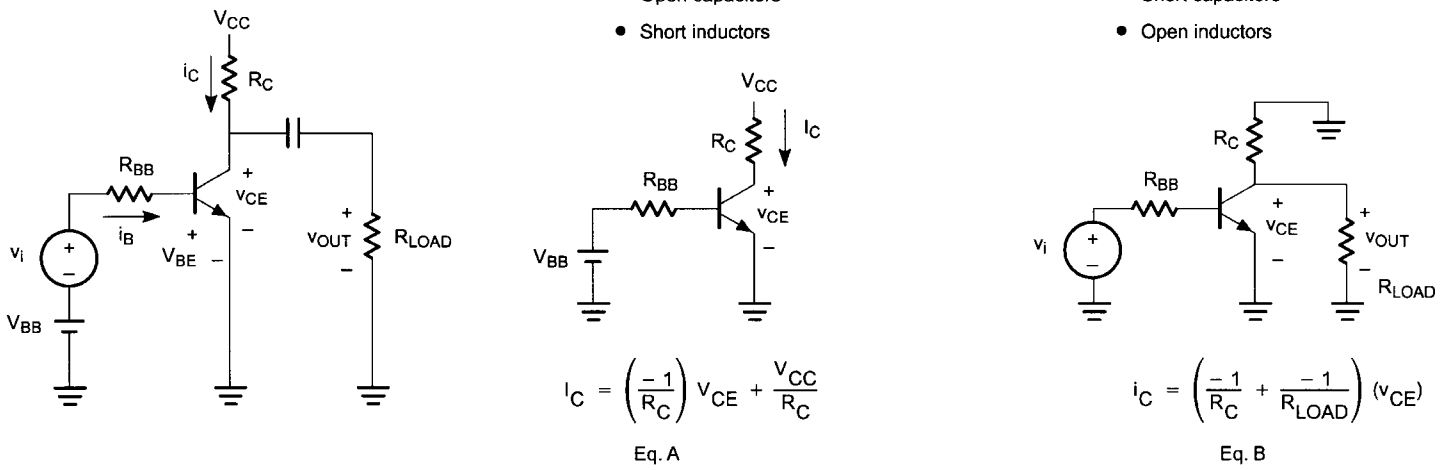


Figure 3-104 Basic transistor analysis with R_{LOAD}.

The slope of the dc load line is generated by the value of the collector resistor ($m = (1/R_C)$ and is shown in Figure 3-102. For a given small-signal base current, the collector current is shown by the dotted curve. The intersection of the dotted curve and the dc load line is called the *quiescent point* (Q-point) and dc bias determinant. The location of the Q-point is important because it determines whether the transistor is operated in the cutoff, active, or saturation region(s). In most cases, the Q-point should be in the active region because this is where the transistor acts like an amplifier.

Figure 3-103 shows the ac collector–emitter voltage (V_{CE}) output swing with respect to an ac collector current (I_C). Collector current is determined by the ac voltage presented to the input transistor's base (v_i) because it affects the base current (i_b), which then affects i_c. This is how v_i is amplified and seen at the output. Recall that this is with no external load (R_{LOAD}) present at the collector. Since no load is present, the slope of the ac load line is identical to that of the dc load line as seen in Eqs. A and B of Figure 3-101 ($m = -1/R_C$).

Open Collector with R_{LOAD}. A filter with some known input impedance is a typical load for the output of the transistor. For simplicity, we will assume a resistive load (R_{LOAD}) and neglect any reactance. Since a resistive load is used (see Figure 3-104), the ac output swing is measured at V_{OUT} or V_{CE}.

A dc blocking capacitor is used between the R_{LOAD} and the V_{CE} output to assure that the Q-point is not influenced by R_{LOAD}. It is also necessary to avoid passing dc to the load in applications where the load is a SAW filter. However, R_{LOAD} will affect the ac load line, which is seen in Eq. B in Figure 3-104. Note that the V_{CE} voltage swing is reduced; thus, the V_{OUT} signal is reduced (see Figure 3-105).

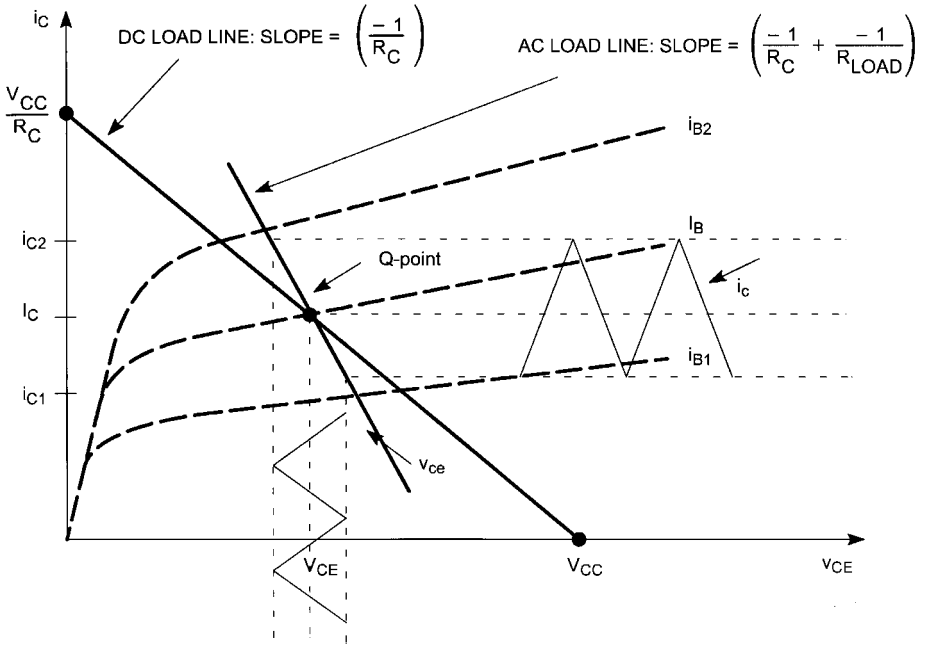


Figure 3-105 Graphical analysis for the circuitry in Figure 3-104. The ac load line intersects the Q point, and the slope is steeper than the dc load line due to R_{LOAD}, which makes V_{CE} smaller.

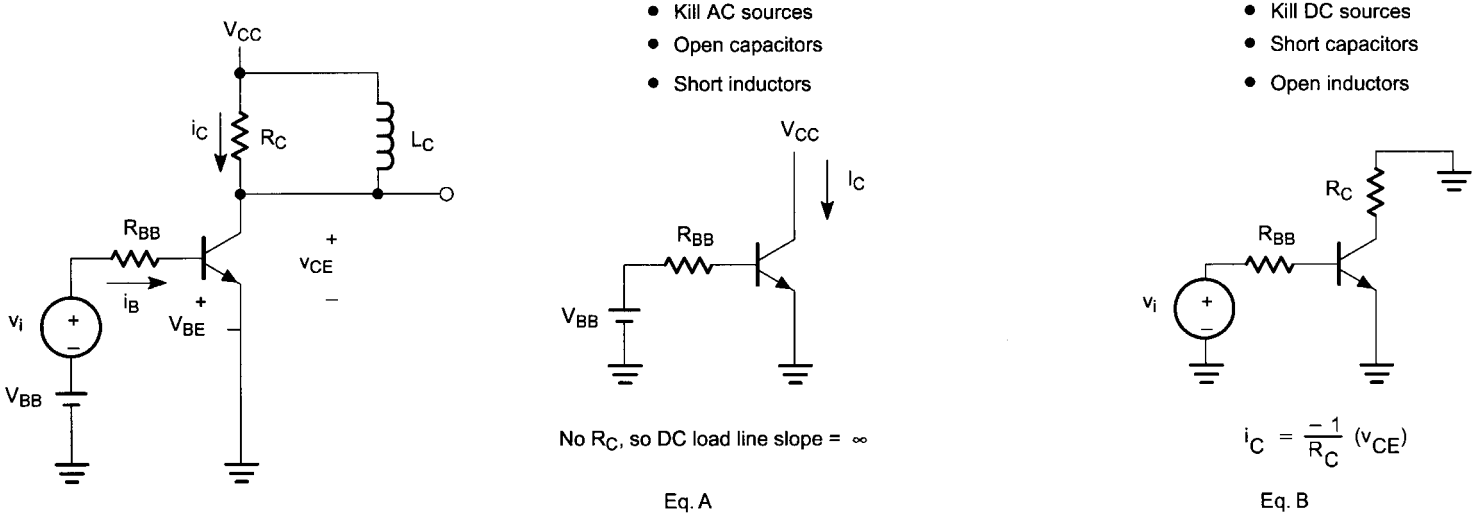


Figure 3-106 Basic transistor analysis with inductor added to collector.

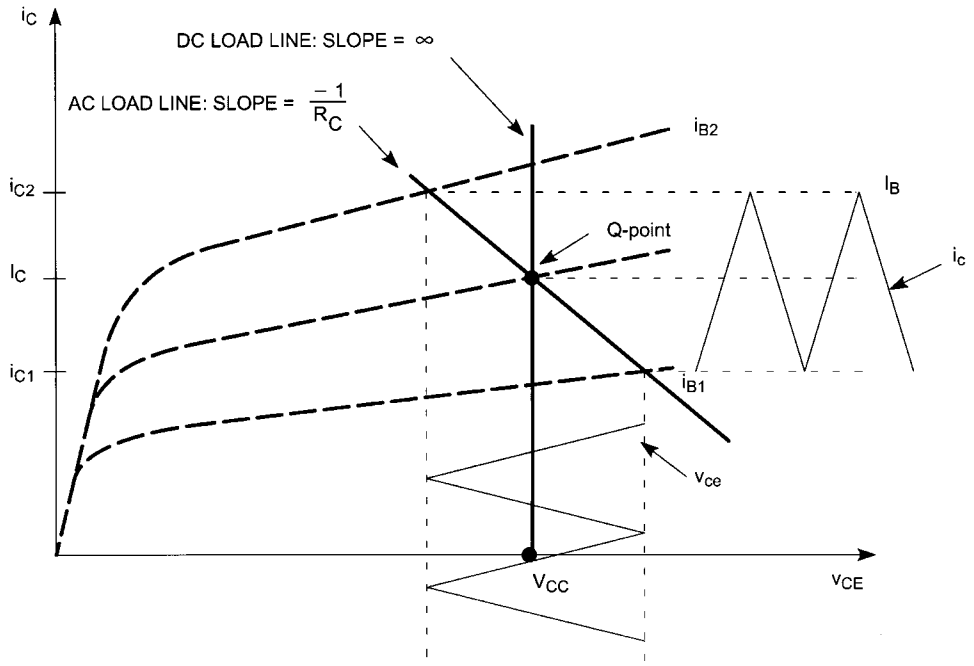


Figure 3-107 Graphical analysis for the circuitry in Figure 3-106.

Since the values of R_C and R_{LOAD} affect the ac load line slope, the value chosen is important. The higher the impedance of R_{LOAD} and R_C, the greater the ac output swing will be at the output, which means more conversion gain in a mixer. This is due to the slope getting flatter, thus allowing for more output swing.

Open Collector with Inductor (L_C). Adding an inductor in parallel with R_C can increase the ac output signal V_{CE}. Figure 3-106 shows the dc and ac analyses of this circuit configuration. In Eq. A in Figure 3-106, there is no R_C influence because the inductor acts like a short in the dc condition. This means the slope of the dc load is infinite and causes the Q-point to be centered around $V_{CC,}$ thus moving it to the right of the curve. The ac load line slope is set only by R_C because no load is present. Note that it has the same ac load line slope as the first condition in Eq. B of Figure 3-101.

Referring to Figure 3-107, one might notice that the base current (ac and dc) curves spread open as V_{CE} increases. This is caused by a noninfinite Early voltage (see Figure 3-108), which causes the collector current to be dependent on V_{CE}. Taking advantage of this nonideal condition, the peak-to-peak ac output swing V_{CE} can thereby be increased by moving the dc Q-point to the right due to the wider spreading between the curves corresponding to different base currents. Figure 3-109 combines Figures 3-103 and 3-107 to show the different ac output signals with different Q-points.

Looking at the ac output level, one might ask how the V_{CE} peak voltage can exceed the supply voltage V_{CC}. Recall that the inductor is an energy-storing device ($v = L\ di/dt$). Therefore, total instantaneous voltage is V_{CC} plus the voltage contribution of the inductor.

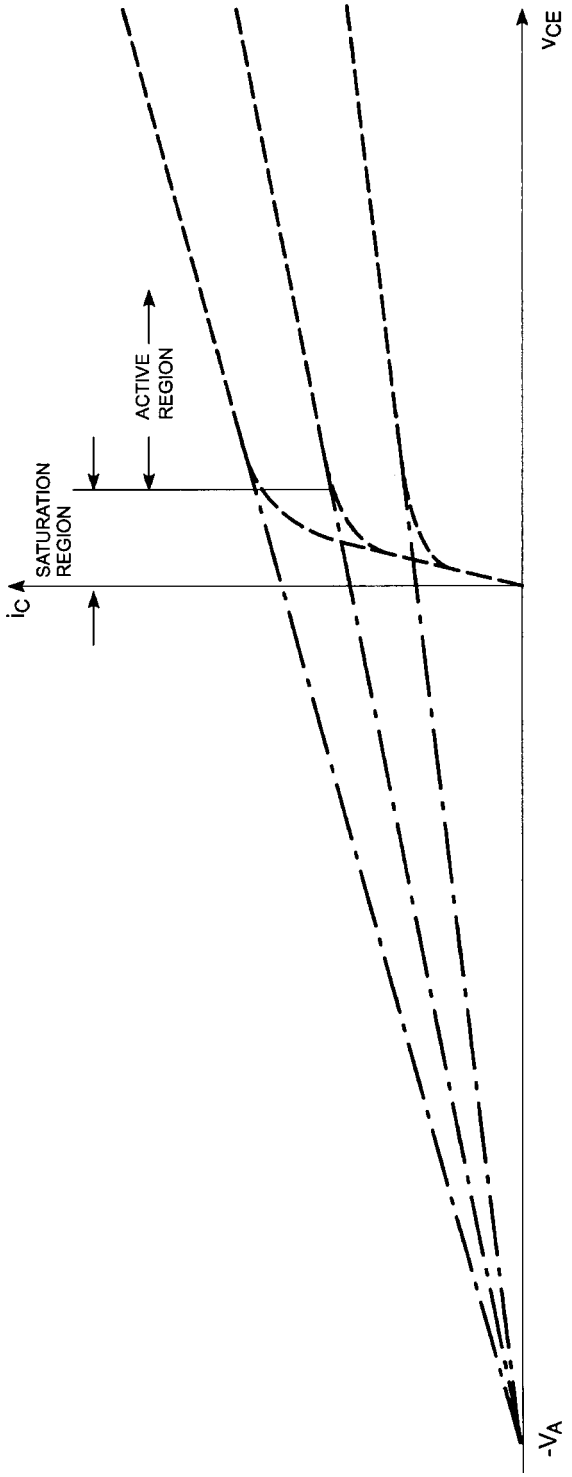


Figure 3-108 Graphical representation of Early voltage effect.

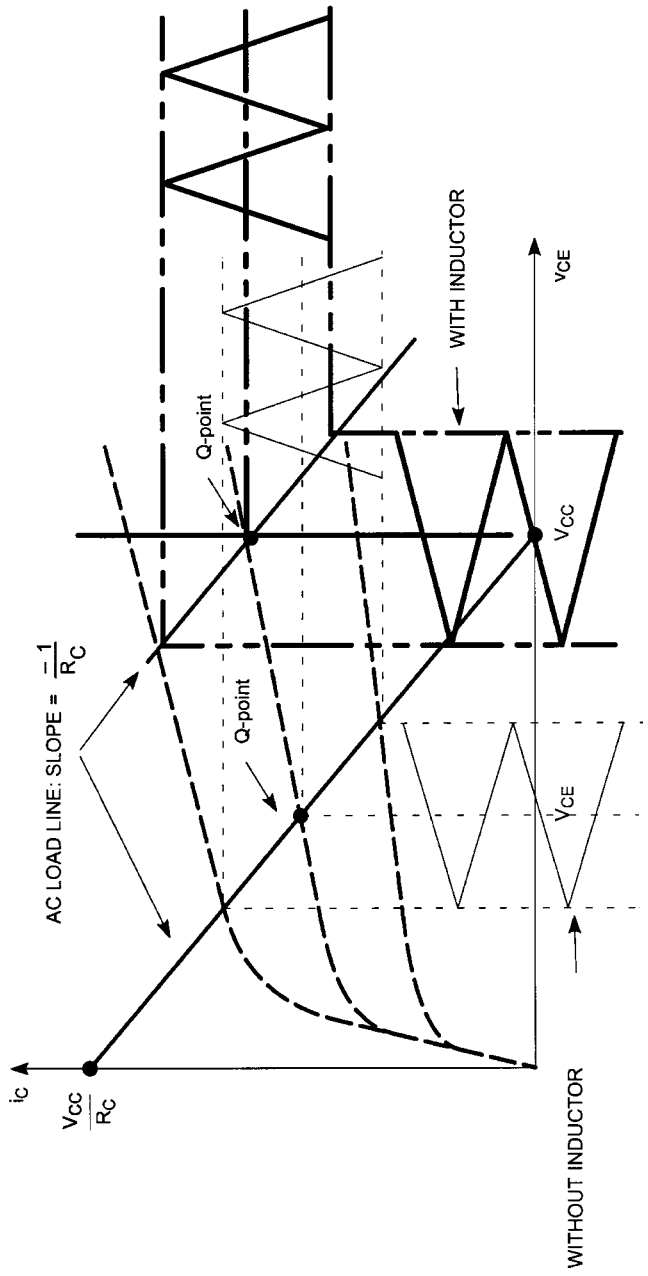


Figure 3-109 Comparison of open-collector circuit with inductor versus without inductor.

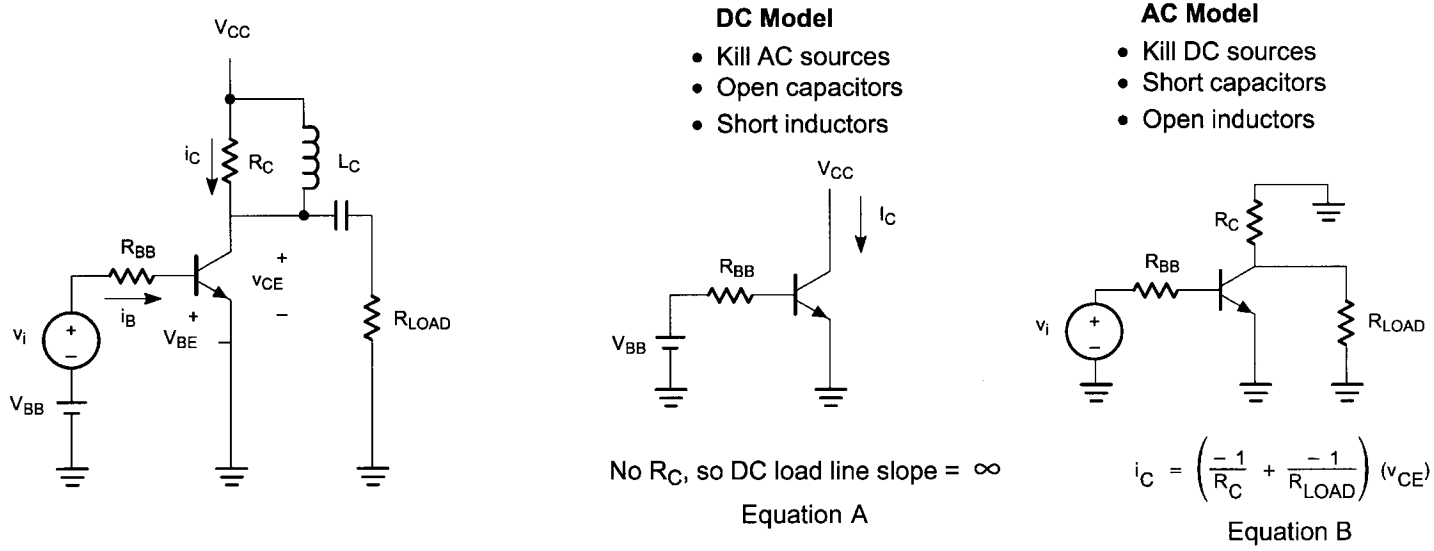


Figure 3-110 Basic transistor analysis with inductor and R_{LOAD}.

Open Collector with Inductor (L_C) and R_{LOAD}. Figure 3-110 shows the dc and ac analyses with the inductor and load resistor. Again, from the dc analysis, the inductor causes R_C to be nonexistent so the dc load line is vertical. In the ac analysis, the ac load line slope is influenced by both R_C and R_{LOAD} (see Eq. B in Figure 3-110). The ac load line slope is the same as the example of the open collector without the inductor. Figure 3-111 shows the response for the open collector with L_C and R_{LOAD}. Figure 3-112 combines Figures 3-105 and 3-111 showing the increase in ac output swing.

In conclusion, for the load line, R_C plays a role in setting up the bias-determined Q-point as well as the ac source impedance. However, when an inductor is placed in parallel with R_C, a different Q-point is set and the ac source impedance is altered. Moving the Q-point takes advantage of the transistor's nonideal i_c dependence on V_{CE} to get more signal output without having to change the base current. Since R_C is in parallel with R_{LOAD} in the ac condition, it influences the ac load line slope.

Flexible Matching Circuit [25]. A useful variation of the open-collector matching concepts previously outlined provided the capability of delivering equal power to two unequal resistive loads. This allows the power delivered to the load to be measured indirectly at another test point in the circuit where the impedance can be defined arbitrarily. If this impedance is defined to be 50 Ω, a spectrum analyzer can easily be placed directly into the circuit. This is an excellent troubleshooting technique and a valuable option to have available in high production environments.

Figure 3-113 shows the schematic for this flexible matching circuit. In this circuit, C_B functions only as a dc blocking capacitor and presents a negligible impedance at the frequency of interest. Recall from the previous open-collector matching discussions that,

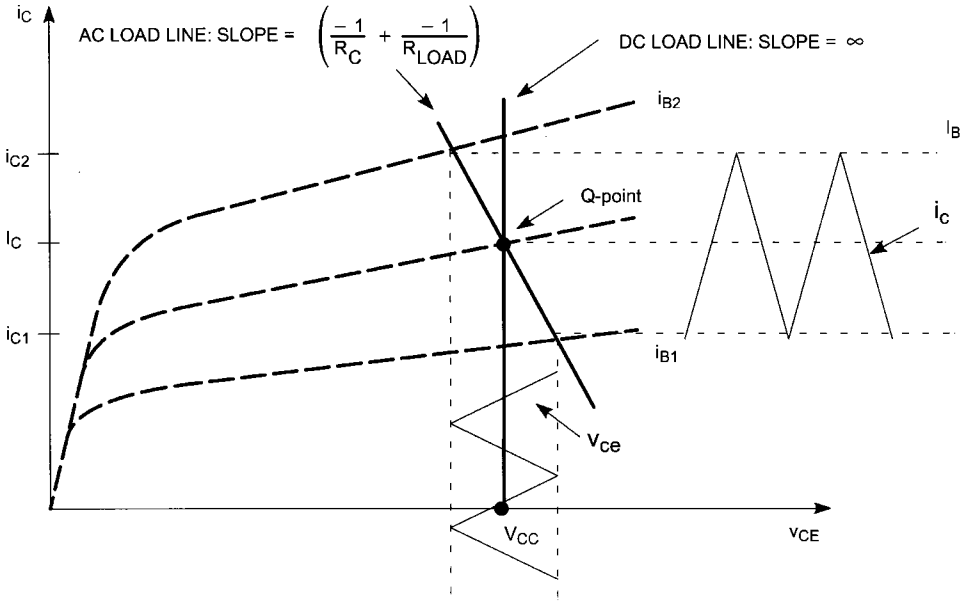


Figure 3-111 Graphical analysis for the circuitry in Figure 3-110.

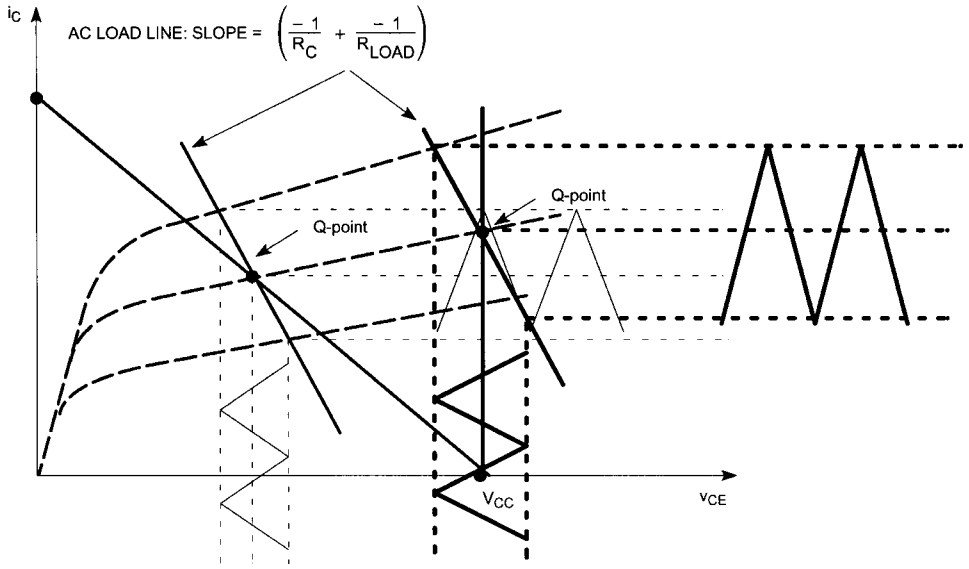


Figure 3-112 Comparison of open-collector R_{LOAD} circuit with inductor versus without inductor.

when R_C is placed in parallel with an inductor, it has no effect on the Q-point but does influence the slope of the ac load line as

$$\text{Slope} = \frac{-1}{R_C} + \frac{-1}{R_{\text{LOAD}}} \tag{3-166}$$

The capacitor C_S not only functions as a dc blocking capacitor but is also chosen such that the impedance presented by the combination of L, R_C, and C_S is equal to R_{LOAD} for optimum power transfer. The analysis is done in the following manner. First, note that inductor L is connected to V_{CC}, which is an affective ac ground. So L can be redrawn to ground. Next, R_C and C_S are converted to their parallel equivalent values as shown in Figure 3-114.

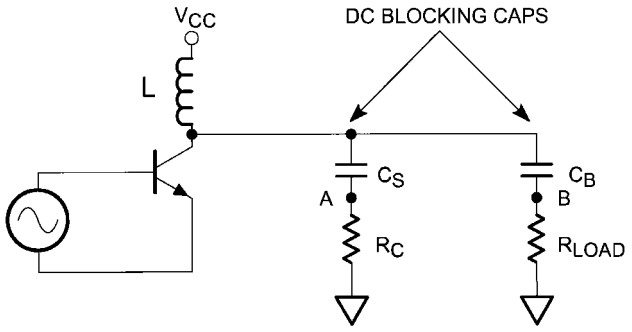


Figure 3-113 Flexible matching circuit.

Figure 3-114 Converting from series to parallel configuration.

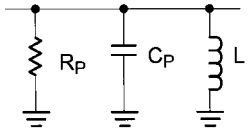


Figure 3-115 Converting from series to parallel configuration. Goal: Find L_C and C_p such that R_p looks like the R_{LOAD} value at the resonant frequency.

The resulting parallel *LCR* circuit is shown in Figure 3-115. At resonance, the parallel L–C_P combination will be an effective open circuit leaving only R_P. R_P is then simply chosen to be equal to R_{LOAD}.

Starting with the most simple transistor configuration and using *RC* types of dc-coupled amplifier circuits, we next present an overview of possible configurations for this application.

3-3 SINGLE-STAGE FEEDBACK AMPLIFIERS

This amplifier (Figure 3-116) is a standard minimal configuration amplifier with reduction of current distortion because of the emitter feedback. The voltage gain is

$$\text{VG} = \frac{R_L}{(26\text{mV})/I_C + R_e} \tag{3-167}$$

where R_L is the total load seen at the collector.

The feedback amplifier in Figure 3-117 uses voltage feedback to "correct" the input and output impedance to be close to 50 Ω but does not improve the current distortion. The voltage gain of this circuit is

$$\text{VG} = \frac{R_F}{R_g} \tag{3-168}$$

$$R_{\text{IN}} = R_{\text{OUT}}$$

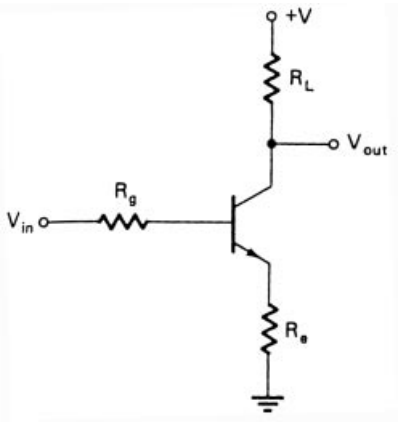


Figure 3-116 Single-stage amplifier with current feedback (provided by R_e).

$$= R_F R_D \quad \text{with} \quad R_D = \frac{26\text{mV}}{I_C} \tag{3-169}$$

Example. For $I_C = 10$ mA $\rightarrow R_D = 2.6\ \Omega$ and assuming $R_F = 200\ \Omega$, the input resistance $R_{\text{IN}} = 22.63\ \Omega$. The gain is calculated by getting the inverse of $R_D = g_m = 0.39$ A/V. Once this is multiplied with the output impedance of 22.63 Ω, the resulting gain becomes 10 log (8.6) or 9.3 dB. The gain can be increased by either using a larger collector current or a larger feedback resistor. It should be reminded that this circuit does not improve the distortion caused by the base–emitter junction.

To reduce the current distortion as well, we use the circuit shown in Figure 3-118, which contains both types of feedback. In reality, the collector–base resistor is frequently split into an unbypassed section responsible for the feedback and a larger resistor responsible for the appropriate biasing. The example circuit in Figure 3-119 shows this.

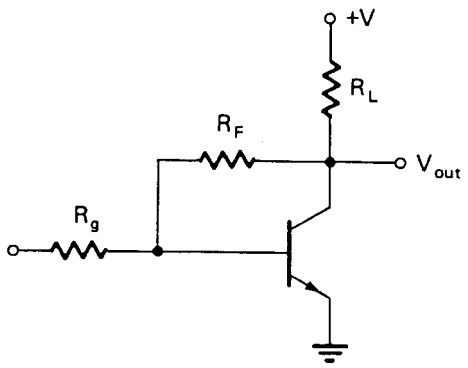


Figure 3-117 Single-stage amplifier with voltage feedback (provided by R_F).

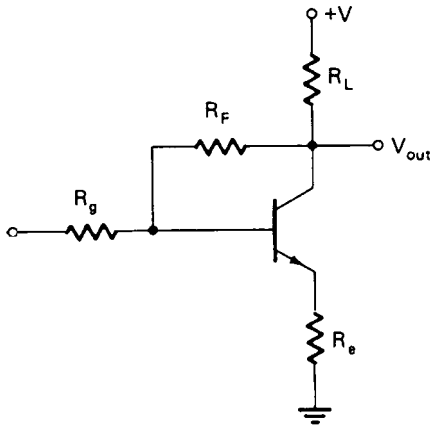


Figure 3-118 Single-stage amplifier with voltage and current feedback.

By using two of the fairly recent BFP420 transistors in parallel, and balancing the transistors by using the appropriate emitter resistors (which in total will result in 10 Ω) plus the bypassed value, we obtain the frequency-dependent gain, matching, and noise responses shown in Figure 3-120. To determine the appropriate feedback resistor R_F, we recommend the formula

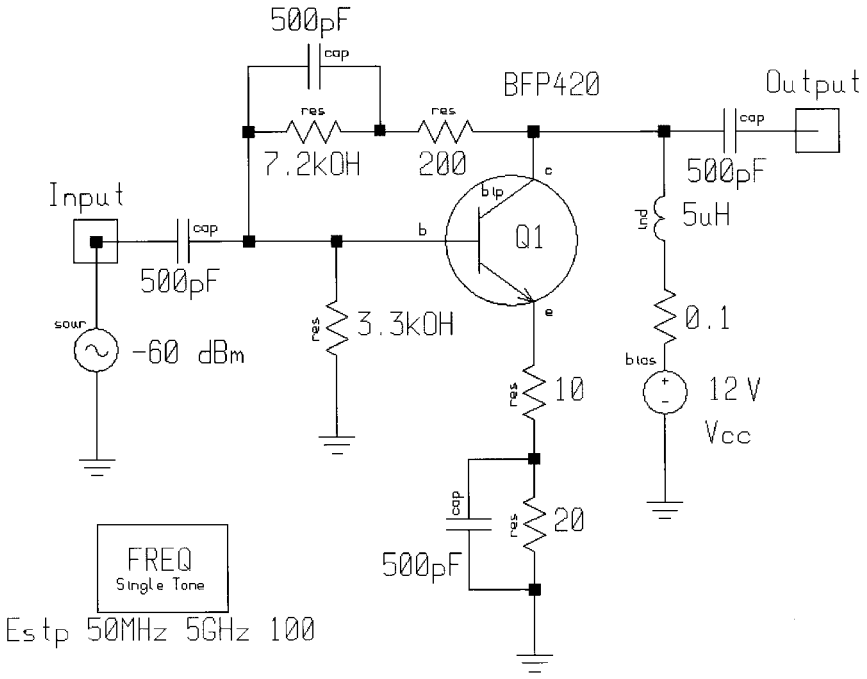


Figure 3-119 Practical amplifier circuit showing voltage and current feedback, and illustrating how reactances (in this case, 500-pF bypass capacitors) and multiple resistors can be used to provide different ac and dc resistances for the paths from collector to base and emitter to ground.

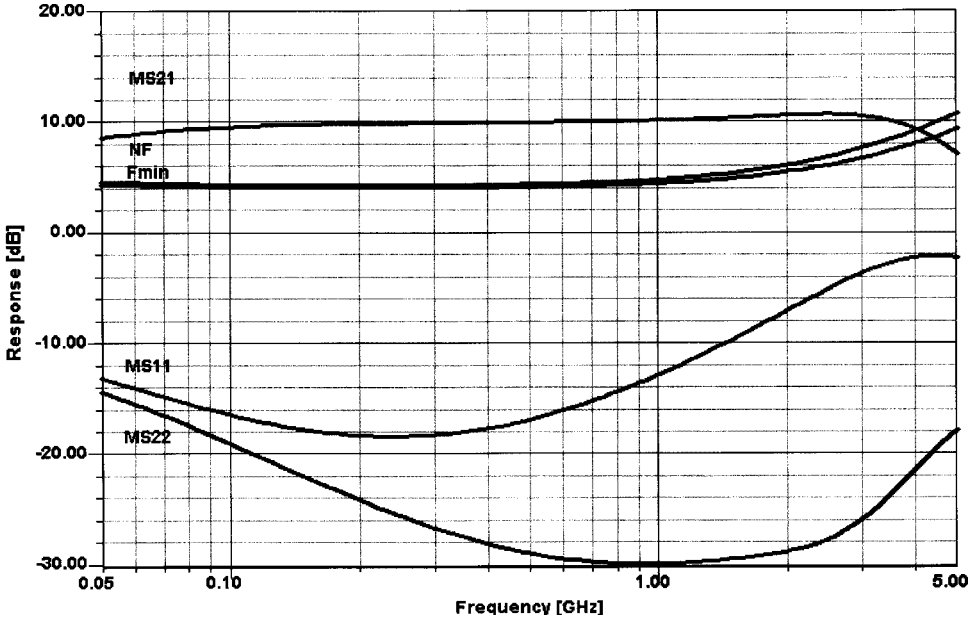


Figure 3-120 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-119.

$$R_F = \frac{(Z_{\text{in}}\, Z_{\text{out}})}{R_E} \tag{3-170}$$

This equation requires some additional comments. The combined shunt and series feedback only work as long as a sufficient transconductance is provided. In other words, there must be enough open-loop gain that will be reduced to obtain more bandwidth. A necessary condition for this approach, which really comes from the frequency range of less than 30 MHz, is that

$$Z_{\text{in}} = \frac{26\ \text{mV}}{I_C}\,\beta >> R_F \tag{3-171}$$

This requires a minimum transconductance of

$$g_m = \frac{1 + S_{21}}{Z_{\text{in}}} \tag{3-172}$$

where S_{21} must be a negative number that reflects the 180° phase shift. Of course, life becomes interesting when S_{21} goes through 1, but it (hopefully) never becomes negative and larger than 1 at the same time.

Example. The BFP420 has been used here frequently. A 10-dB gain would require an S_{21} of 3.16 ($\sqrt{10}$); therefore, we need to inspect the datasheet, which shows that the magnitude

of S_{21} is 3.16 around 4.5 GHz, yet the phase shift has changed from 180° to around 20°. In order to design an amplifier, we now set

$$R_F = Z_{\text{in}}(1 + |S_{21}|) \tag{3-173}$$

which is valid until S_{21} becomes too small. For a given 10 dB or 3.16 for S_{21}, the resulting RF is 4.16×50, or 208 Ω. The transconductance under these conditions becomes

$$g_m = \frac{1 + |S_{21}|}{Z_0} \tag{3-174}$$

where $Z_0 = 50$ Ω (the desired input impedance). g_m now becomes 83.2 mS. We now test the input and find

$$\frac{\beta}{g_m} = Z_{\text{in}} = \frac{50}{0.0832} = 600\ \Omega \tag{3-175}$$

Therefore, the emitter feedback resistor is not necessary to meet the previously mentioned condition whereby the input impedance needs to be higher than RF. Assuming the transconductance now is 0.00832, or 10% of the previous value, the input impedance would be 60 Ω, no longer meeting the prescribed requirement. Therefore, we need an unbypassed emitter resistor that fulfills the equation

$$R_E = \frac{Z_0^2}{R_F} - \frac{1}{g_{m_{\text{dc}}}} \tag{3-176}$$

where

$$g_{m_{\text{dc}}} = \frac{1 - S_{21_{\text{max}}}}{Z_0} \tag{3-177}$$

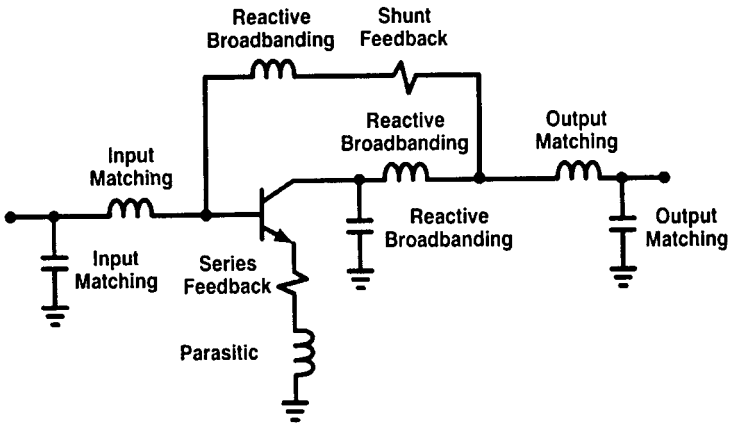


Figure 3-121 Elements of a feedback amplifier.

Again, the reader should be cautioned that in the common-emitter circuit, there is a 180° phase shift, and therefore $S_{21_{\max}}$ is a negative quantity, resulting in reality in

$$g_{m_{\text{dc}}} = \frac{1 + S_{21_{\max}}}{Z_0}$$

In our case,

$$g_{m_{\text{dc}}} = \frac{-S_{21}}{2Z_0} = \frac{38}{2 \times 50} = \frac{38}{100} = 0.38$$

so

$$R_E = \frac{50^2}{208} - \frac{1}{0.38} = 12 - 2.6 = 9.4\ \Omega \tag{3-178}$$

In practice, the designer would use a 10-Ω resistor.

Since the gain of the transistor continues to fall as a function of frequency, one can compensate to a degree for the cutoff frequency by putting an inductance in series that follows the equation

$$L_F = \frac{R_F}{2\pi f_{3\text{dB}}} \tag{3-179}$$

Assuming $f_{3\text{dB}}$ = 2 GHz, the resulting series inductance L_F = 16.6 nH. Basically, the same applies to circuits where the second transistor is used as an emitter follower to feed the signal back from the collector to the base of the first transistor.

In summary (see Figure 3-121):

- Resistive negative feedback at lower frequencies reduces open-loop gain and simultaneously matches input and output.
- Reactive elements set the gain and matching at high frequencies. Feedback is reduced at high frequencies.
- Emitter bypass capacitors are not used, as they would destroy the impedance match.

This is equally valid for FETs.

The next level of complexity can be reached by using two transistors with dc feedback.

3-3-1 Lossless or Noiseless Feedback

All the amplifiers shown above are based on a combination of resistive and reactive feedback. The very moment resistive feedback is used, the noise figure increases, as does the intercept point. As long as the intercept point increases overproportionally compared to the noise figure, this is an acceptable method. However, there are many cases in which an extremely low noise figure and a higher than normally achievable intercept point are required. To avoid this degradation of performance, Norton of Anzac developed a negative feedback structure based on the capabilities of the ferrite transformer. This patented technique (U.S. Patent Nos. 3,624,536 and 3,891,934, since expired), known as *lossless feedback*, provides lower noise figure and higher linear output than can be achieved using resistive feedback with the same transistor. Transformer feedback is used in which the transformer in the feedback network acts as a directional coupler. The coupling ratio between the input and output of the coupler

determines the magnitude of the feedback, and hence the gain. Across the frequency band over which the transformer operates as intended, the feedback is negative due to the in-phase coupling and the inversion of the active device.

This technique was used in amplifiers ranging from 5 to 500 MHz, 20 to 1000 MHz, and 300 to 1800 MHz. The difficulty associated with this type of amplifier is the fact that it is based on transformers that use twisted pairs of wires as transmission lines. In some implementations used at higher frequencies, actual directional couplers were used at the expense of bandwidth; in these circuits, the coupling was determined by the spacing of the coupler lines rather than transformer turns ratio. These microstrip couplers have much narrower bandwidths than transformers and are fairly large.

Another headache with the lossless feedback design is the fact that many of these transistor configurations are used in either grounded-base or grounded-emitter with transistors that have f_T values of 25 GHz and higher. Such devices are very hot, and achieving freedom from oscillation over a wide frequency range is a continuing problem, even for the circuits we have shown so far. A similar clever technique has been implemented in the termination-insensitive mixer, which we will examine in Chapter 4.

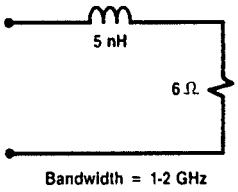
3-3-2 Broadband Matching

The circuitry above has enabled us to extend the frequency range of the amplifier and avoid the normal sharp cutoff. According to Fano [26], there is a limit to the bandwidth of matching. The concept is that impedance plus bandwidth determines the ability to match, since there is a Q of the circuit and there is a Q of the device. The following equations are valid:

$$|\rho_{\text{Bode}}| = e^{-(\pi Q_T / Q_{BP})} \tag{3-180}$$

$$Q_T = \frac{\sqrt{f_1 f_2}}{f_2 - f_1} \tag{3-181}$$

$$Q_{BP} = \frac{\text{Im}(Z)}{\text{Re}(Z)} = \frac{X_s}{R_s} \tag{3-182}$$



$$Q_{BP} = \frac{2\pi\sqrt{(2\text{ GHz})(1\text{ GHz})(5\text{ nH})}}{6\,\Omega} = 7.40$$

$$Q_T = \frac{\sqrt{(2\text{ GHz})(1\text{ GHz})}}{(2\text{ GHz} - 1\text{ GHz})} = 1.414$$

$$|\rho_{\text{Bode}}| = e^{-\left(\frac{\pi(1.414)}{7.40}\right)} = .549$$

$$\text{VSWR}_{\text{Bode}} = \frac{1 + .549}{1 - .549} = 3.43 : 1$$

Figure 3-122 Broadband matching example showing the Bode limit.

Figure 3-122 shows an example. This combination results in a VSWR of 3.43, which for practical purposes is already too high. The workaround is special circuitry (such as exponentially staggered impedance jumps in the matching network like 50 Ω > 20 Ω > 7 Ω > 3 Ω). This means we first transform the 50-Ω source of load impedance down to 3 Ω, and then look for a matching network that matches the transistor output/input, which is within a few ohms of 3 Ω. A detailed discussion of such matching can be found in the "Numerical Design" portion of Section 3-12-2 under the title "Broadband Matching Using Bandpass-Filter Networks—High-Q Case."

3-4 TWO-STAGE AMPLIFIERS

There are two popular configurations that have been used mostly for low-frequency applications. The first one (Figure 3-123, implemented in the circuit shown in Figure 3-124) shows a nice high-frequency performance assuming the emitter and collector currents are essentially the same. As for distortion products, this circuit behaves equivalently to the shunt impedance example from Figures 3-118 and 3-119, with the additional gain of the second transistor. Its frequency matching and noise performance are also shown, in Figure 3-125.

The next two-stage circuit (Figure 3-126, implemented in Figure 3-127) uses the collector current of the second stage as the feedback to the first. Proper biasing of this circuit is quite difficult and interactive. Figure 3-128 shows the frequency-dependent gain, matching, and noise figure. We believe that this circuit can be optimized further, but we leave such investigations to the reader.

A *cascode* amplifier is a special form of two-stage circuit. Although a cascaded pair acts like a single transistor, its merits are in its reduced feedback and resulting higher gain–bandwidth product. Figure 3-129 shows a simple cascode application that assumes the presence

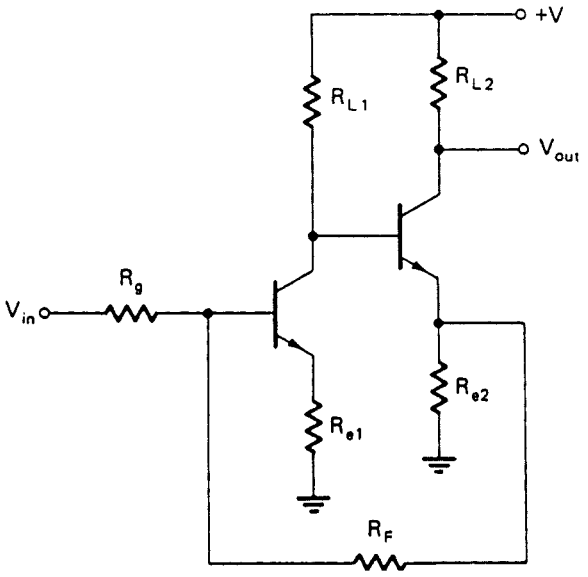


Figure 3-123 Two-stage amplifier with voltage feedback.

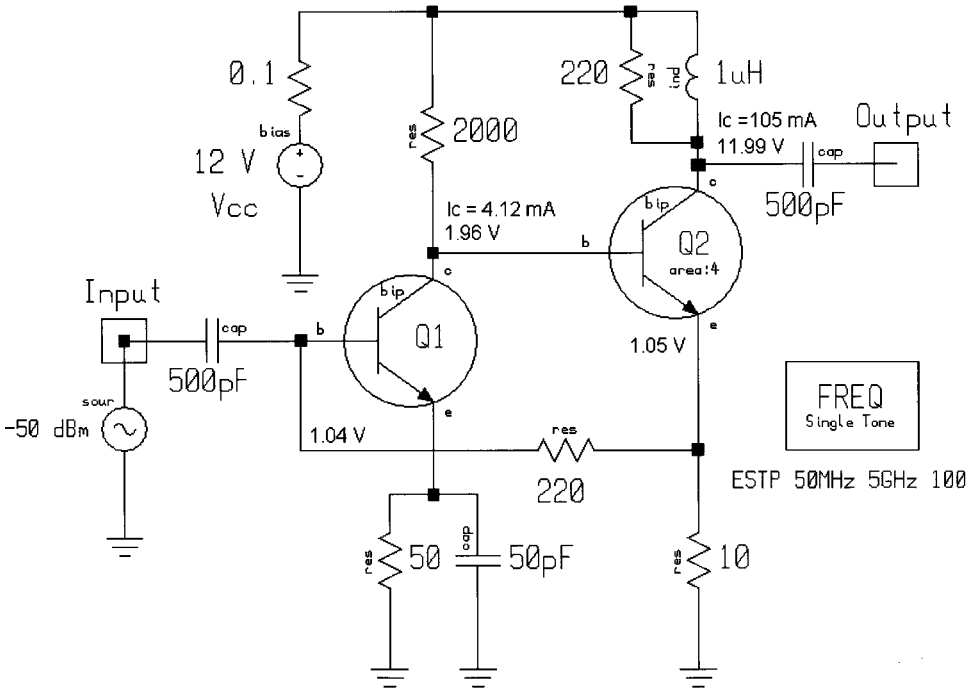


Figure 3-124 Implementation of the circuit in Figure 3-123.

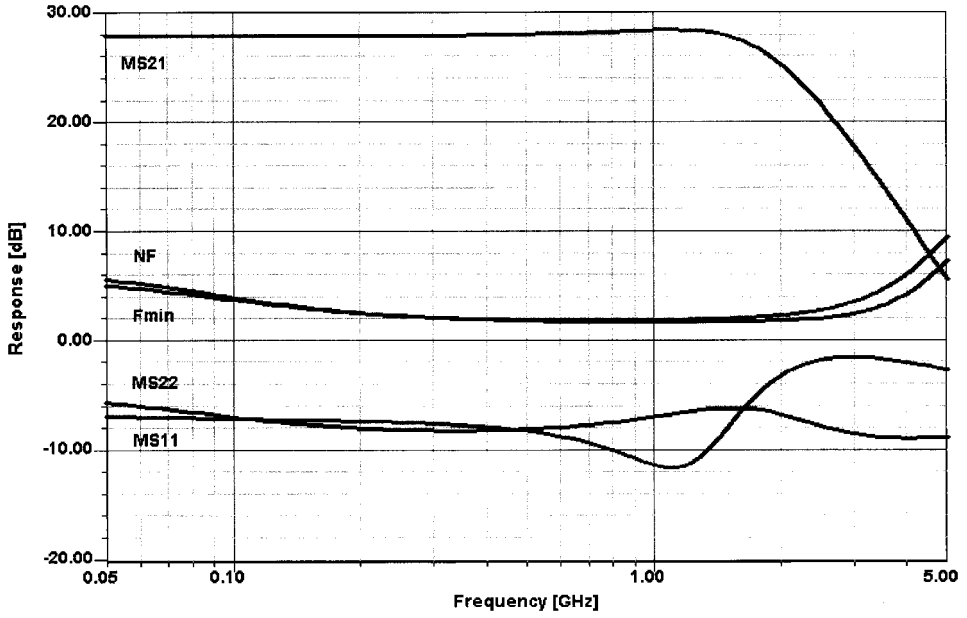


Figure 3-125 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-124.

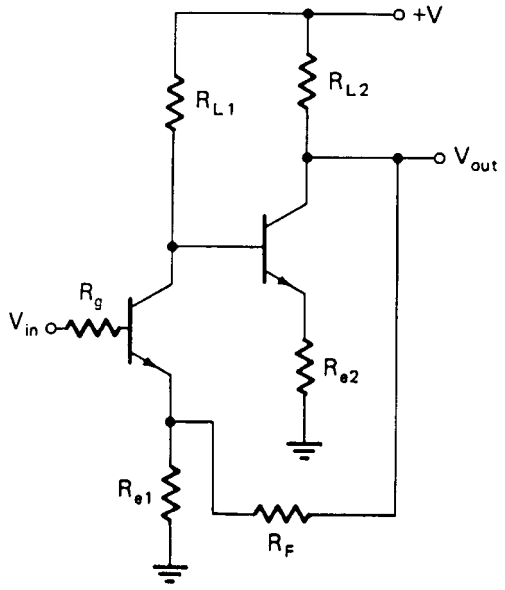


Figure 3-126 Two-stage amplifier with voltage feedback from Stage 2 to Stage 1, and current feedback in both stages.

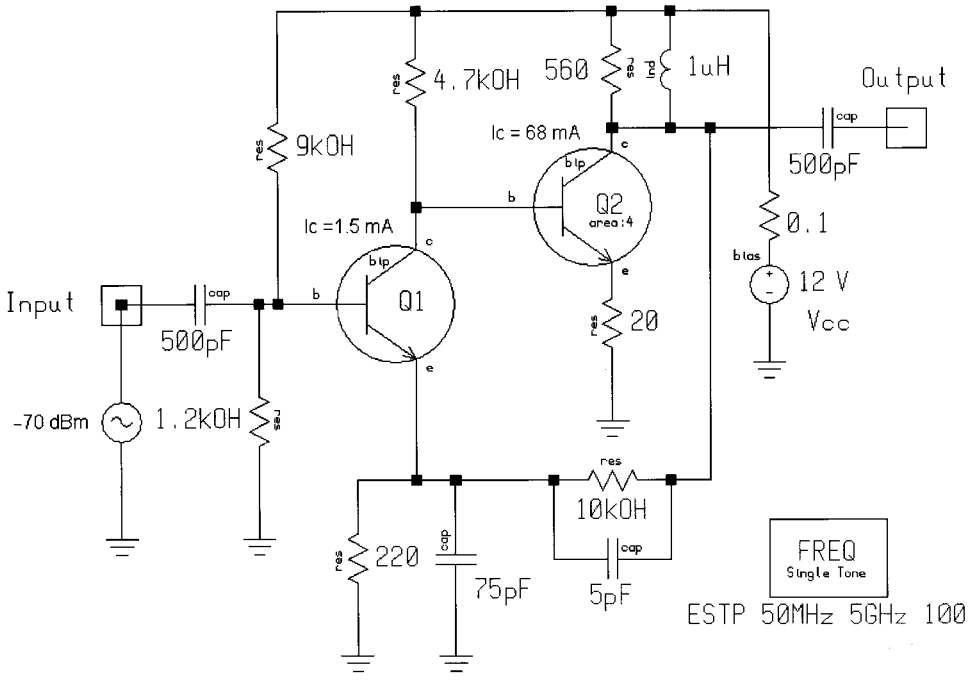


Figure 3-127 Simulated two-stage amplifier with feedback as shown in Figure 3-126.

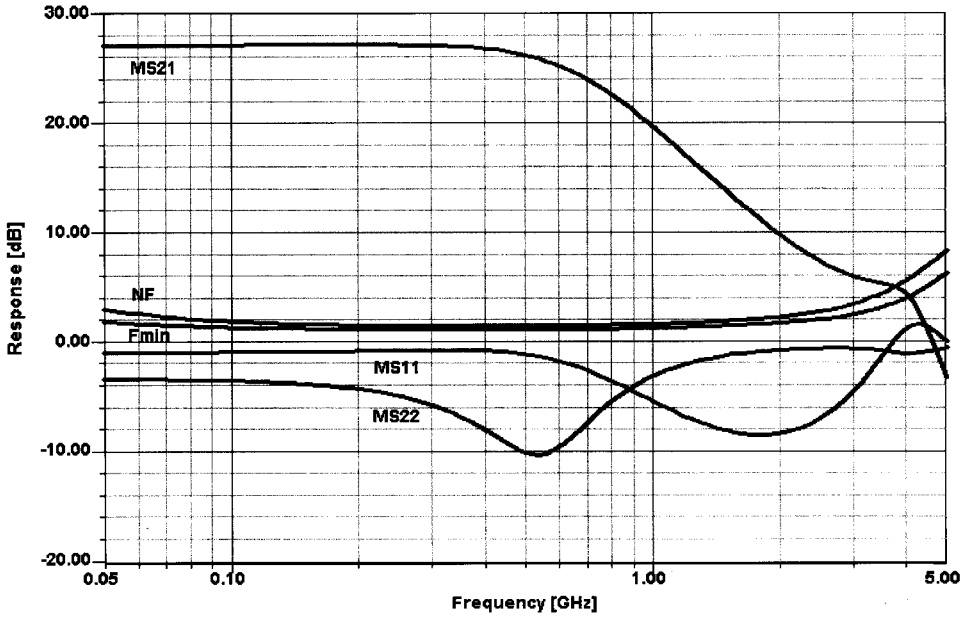


Figure 3-128 Gain, matching, and noise performance of the amplifier in Figure 3-127.

of an output transformer that absorbs the output capacitance. Figure 3-130 shows this circuit's frequency-dependent gain, matching, and noise performance.

It has been obvious to us that most published wireless application circuits are evaluated on a no-input-selectivity basis; that is, their reported performance does not take into account the filtering that is usually present in practical implementations. Evaluated in this way, a low-noise design may achieve a noise figure of less than 2 dB—1.2 dB or so is typical—but this is misleading because filters and connectors introduce insertion loss, resulting in a higher noise figure.

The cascode is a nice preamplifier combination. It combines (depending on the transistor) a low noise figure with essentially no feedback. Since the output of a grounded-base transistor provides a high-impedance source, this configuration is ideal to work into the SAW filters frequently used in wireless applications at the operating frequency. Special transistor pairs are available for use in cascode; as can be seen from this example, not every transistor combination is ideal. This one gives a higher noise figure than wanted and at frequencies above 2 GHz it tends to show possible instabilities. Both Siemens and Motorola are now offering transistors appropriate for these cascodes. Mostly, one would resort to transistors of less-exotic cutoff frequencies to guarantee stability. Typically, grounded-base/gate transistors tend to become oscillators, specifically because this configuration, like the emitter follower, allows a higher operating frequency while maintaining the same gain–bandwidth product; the actual gain for the last two circuit configurations is less. A grounded-base/gate stage has a high voltage gain and the emitter/source follower has power gain at low impedance levels. We invite our readers to experiment with the capacitor from the base of Q1 to ground in Figure 3-129 to get an interesting education about these transistors and their willingness to oscillate.

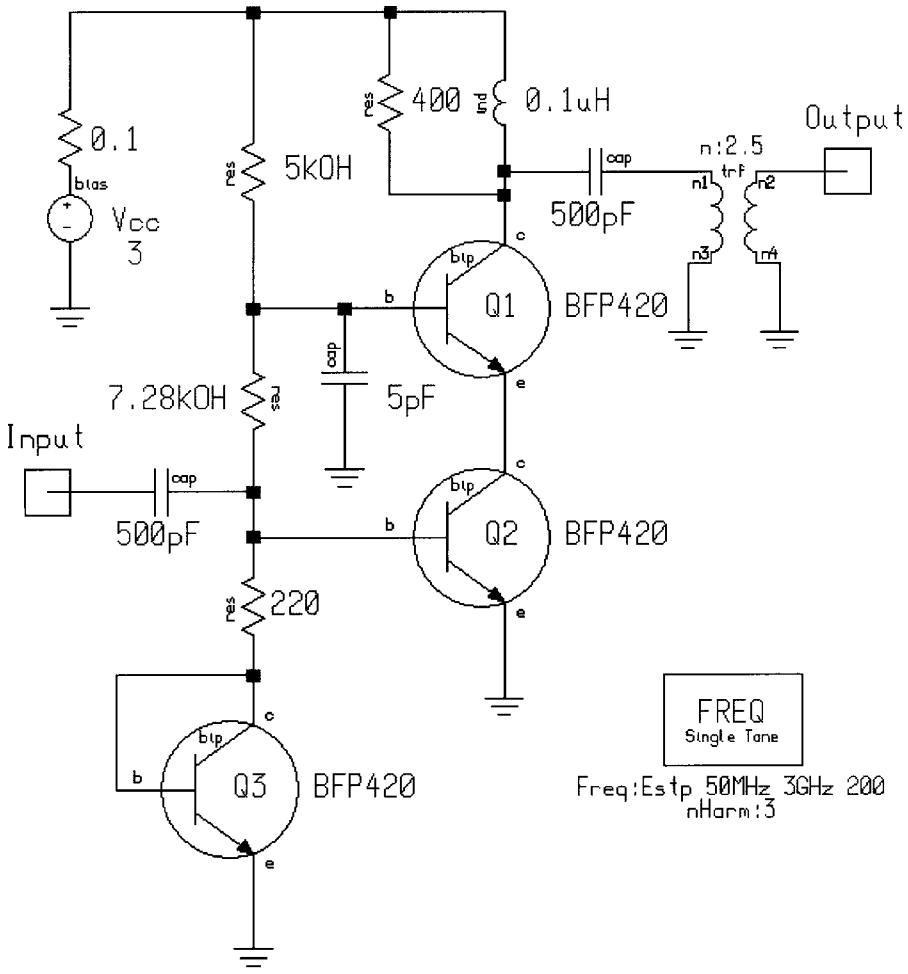


Figure 3-129 Schematic of the cascode amplifier.

The MC13144 is an example of a commercially available VHF to 2-GHz low-noise amplifier with programmable bias. As can be seen from Table 3-11, it is ideal for the 900-MHz range and its programmable gain/bias makes it very attractive. This design (Figure 3-131) is based on the cascode mentioned above and, because of high integration, uses many transistors for bias and level shifting.

The 900-MHz application circuit in Figure 3-132 shows the required external components for a successful design. Input and output are made tunable. The 5.6- and 39-nH inductances can be made from small air-wound coils. The output ceramic filter provides the necessary selectivity ahead of the mixer. With such a highly selective device in the input, the noise figure would deteriorate too much. The antenna is rarely in the vicinity of sufficiently strong transmitters to require such strong input filtering.

There are two more interesting applications for two-stage amplifiers: The one implemented in the Hewlett-Packard MSA-0735 MMIC has become a quite common one, and its usefulness up to higher frequencies depends mostly on the cutoff frequency of the transistors,

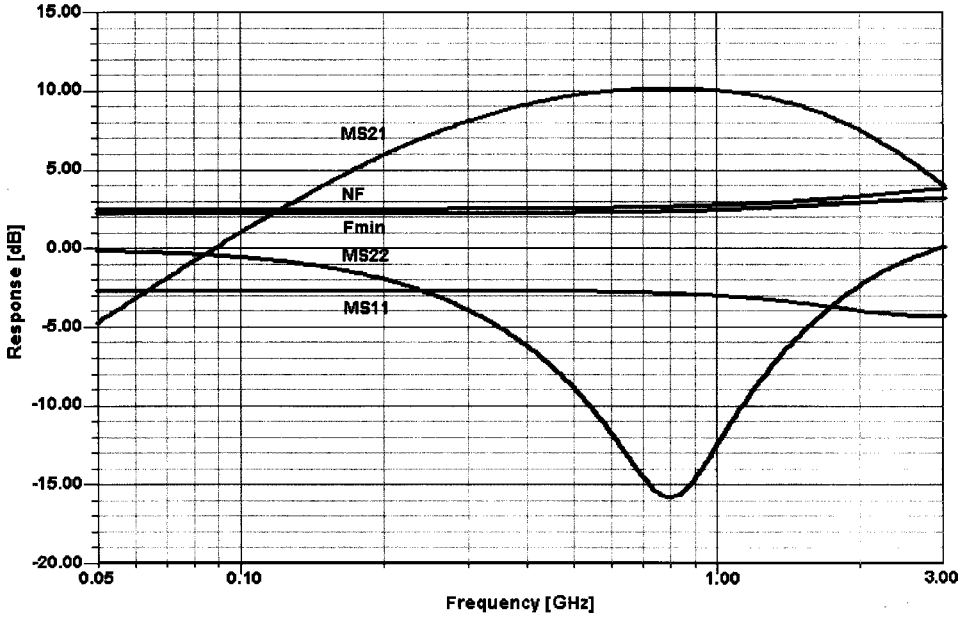


Figure 3-130 Frequency-dependent gain, matching, and noise performance of the cascode amplifier.

their parasitics, and the phase shift of the gain. The typical configuration is shown in Figure 3-133, and its frequency-dependent gain, match, and noise figure in Figure 3-134. There is an increase of S_{22} at higher frequencies; this indicates unwanted phase shift in the circuit. Some manufacturers hide the compensation circuits to overcome such peaking; most of the schematics shown by the manufacturers are really intended only to show operating principles, with proprietary details omitted.

Thanks to the properties of GaAs transistors, we can extend the frequency range significantly. The amplifier we are about to look at goes back several years and is described in great detail in Goyal [27]. Essentially what we have, seen in Figure 3-135, is an FET version of Figure 3-127. The main goal of this circuit was to have an all-monolithic circuit, avoiding all ac coupling and accomplishing all matching and interaction with dc coupling. The transistor Q1 is the main gain stage, which has the transistor Q2 as an active load. The output impedance of Q2 is $1/g_m$—typically in the area of 50–200 Ω, depending on the biasing,

Table 3-11 MC13144 features

Gain at 900 MHz, 17 dB
Noise figure at 900 MHz, 1.4 dB
P_{-1dB}, −7.0 dBm
$IP_{3,in}$, −5.0 dBm
Low operating supply voltage (1.8–6.0 Vdc)
Programmable bias with Enable 1 and Enable 2. [Enable 1 and Enable 2 programmed high for optimal noise figure and gain associated with NF. Enable can be overridden for external bias programming (up to 15 mA.)]

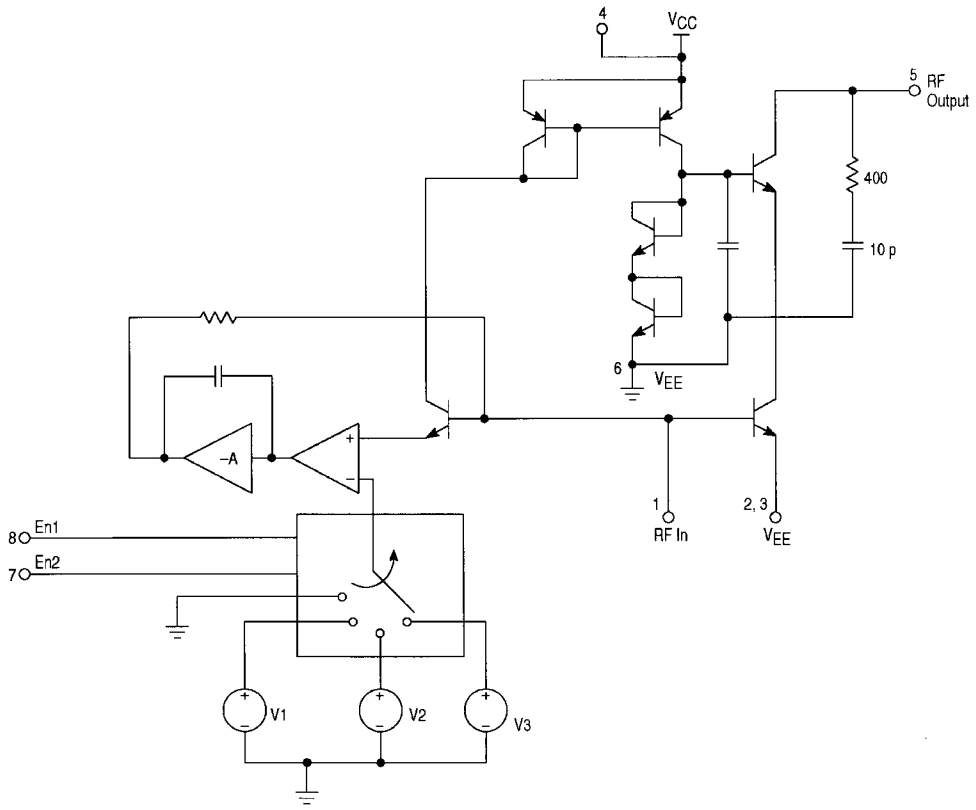


Figure 3-131 Schematic of the Motorola MC13144 cascode LNA.

which affects the transconductance. Transistor Q3 has the same function as an emitter/source follower; again, its dc bias determines its transconductance and therefore its output impedance. (Given this, a device that has a transconductance of 20 mS, or close to it, would establish a very good output match to 50 Ω.) The drains of Q1 and Q5 are at approximately 3.5 V. The diodes at the source of Q3 are level shifters that must shift the dc voltage at the gate of Q5 and at the drain of Q4 to approximately −1.5 V relative to ground—the necessary bias condition for Q1 and Q5. The standard shunt feedback circuit would show resistive feedback between the drain of Q1 (gate of Q2) and the gate of Q1; in the previous examples, we calculated the magic value of this part to be about 200 Ω. Adding in parallel to Q1 the transistor Q5 allows the designer to use a feedback scheme that isolates the feedback loop from the input. Dc-wise, the two transistors are tied together via a 10-kΩ resistor, which can be included in the actual packaged device. Insufficient information was published to allow us to duplicate the design, considering all the intricacies involved (we somewhat object to the fact that even in the *IEEE MTT-S Transactions*, so little information is given about the actual implementation of this design—using space constraints as an excuse—that its educational value is usually minimal); it shows a new principle without revealing the design steps necessary to obtain a practical circuit. Several variations of this type of circuit are available; the literature is full of discussions about them. The magnitude of the feedback is set by the ratio of the width of Q5 to that of Q1. Typical values for the Q5/Q1 ratio range

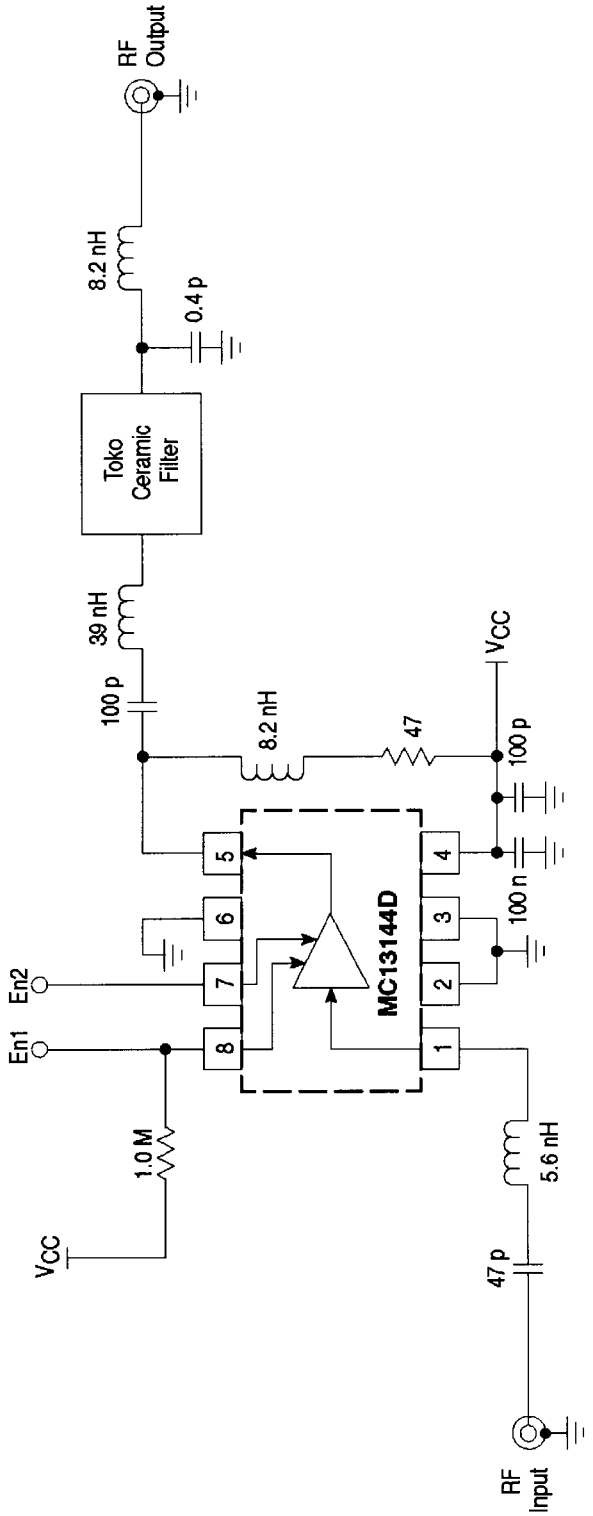


Figure 3-132 An MC13144 application circuit for 926.5 MHz.

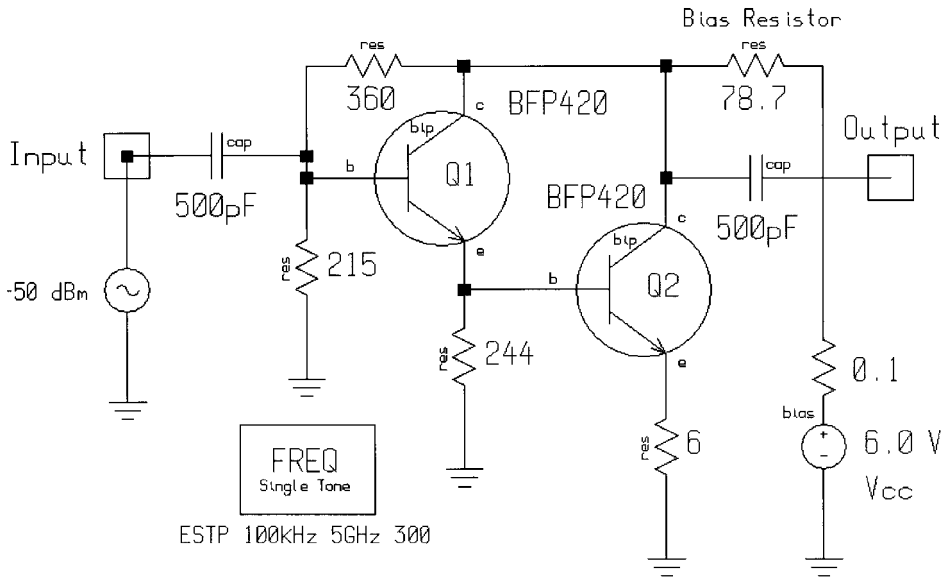


Figure 3-133 Schematic of the MSA-0375 MMIC amplifier.

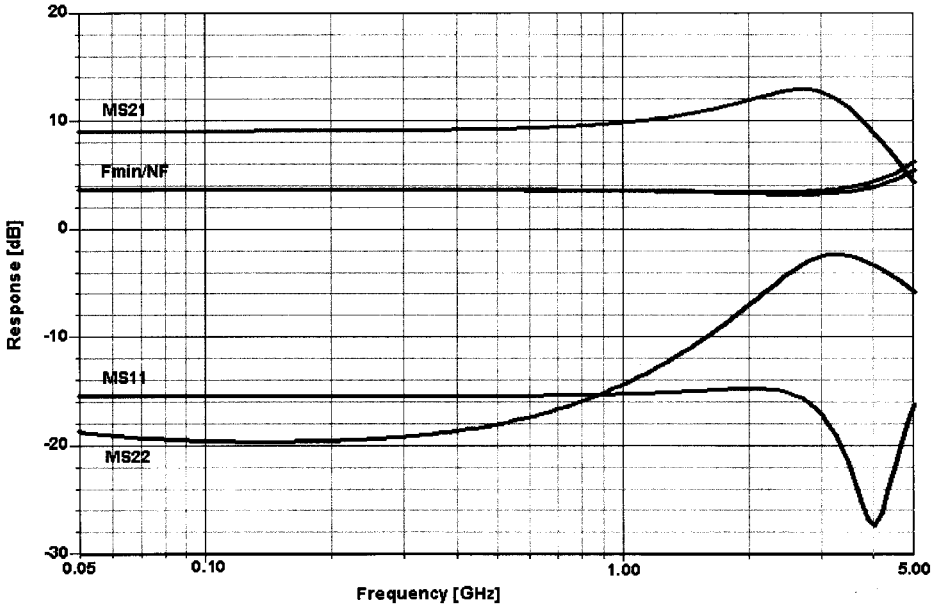


Figure 3-134 Frequency-dependent gain, matching, and noise performance of the MSA-0735 MMIC amplifier.

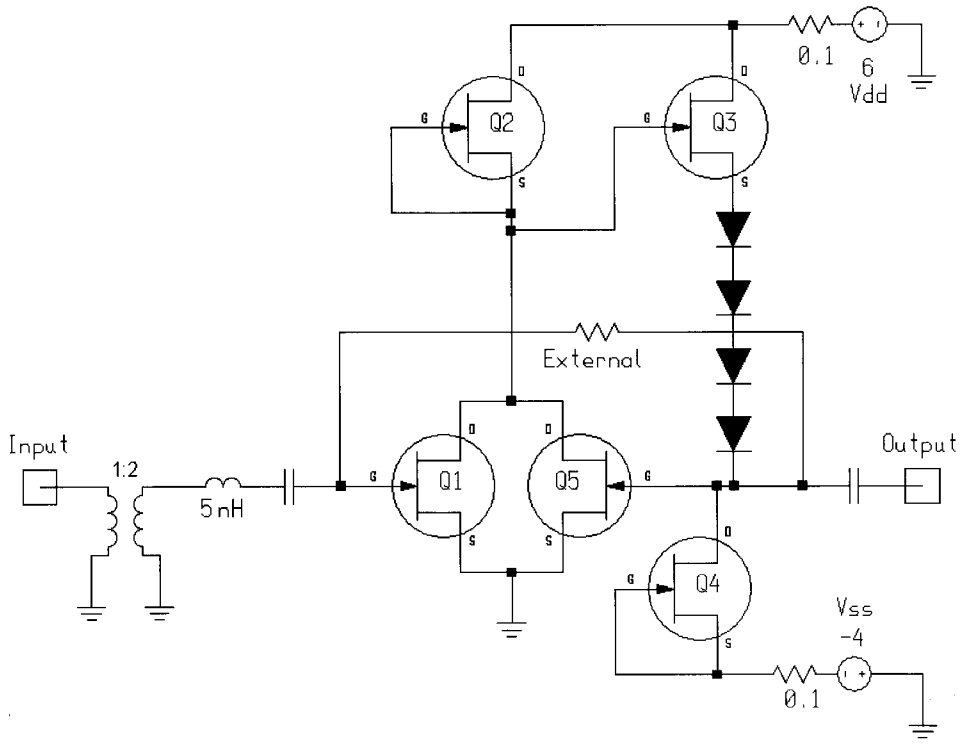


Figure 3-135 Schematic of the dc-coupled GaAsFET amplifier.

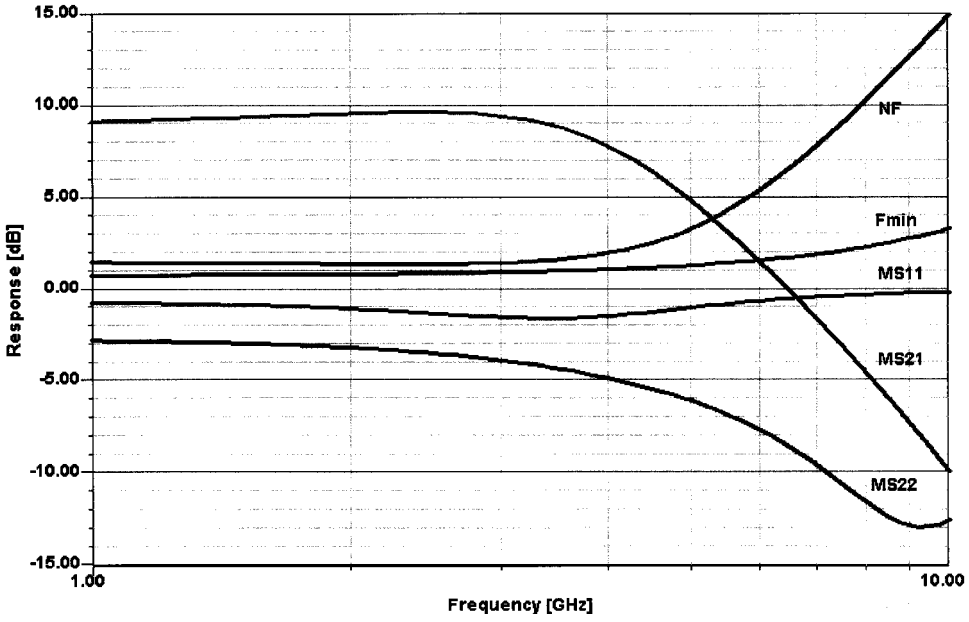


Figure 3-136 Frequency-dependent gain, matching, and noise performance of the dc-coupled GaAsFET amplifier.

from 0.15 to 0.30. In simulation, sizing the devices can best be accomplished by using the scaling factor for the FET model.

One of our questions was: "What would the perfect input termination for Q1 be?" By using an ideal transformer and playing with its turns ratio (needless to say, we did this with CAD), we determined experimentally that, relative to the gate, 200 Ω and a series inductor of 5–8 nH—the sort of details that designers often avoid telling their audiences—gave the ideal frequency response. Overall, we believe that the inventor of this circuit can be quite proud of the overall performance. The FETs we chose were taken from the Texas Instruments nonlinear foundry library; however, any similar device from another company would have worked equally well. The feedback loop used in this IC is an example of an approach frequently referred to as *active feedback*. The performance of this amplifier can be evaluated from Figure 3-136, which shows the frequency-dependent gain, matching, and noise figures (even including F_{min}).

3-5 AMPLIFIERS WITH THREE OR MORE STAGES

Most three-stage amplifiers are obtained by adding another buffer or power-gain stage to the previously described principal designs. As an example, we decided to evaluate the recent NEC UPC2749 IC, a silicon MMIC intended for application as a low-noise 1900-MHz amplifier operating at 3 V. Table 3-12 briefly summarizes its electrical specifications.

Table 3-12 NEC UPC2749 electrical characteristics

(T_A = 25 °C, $Z_L = Z_S$ = 50 Ω, V_{CC} = 3.0 V)			
	Part Number Package Outline		UPC2749T TO6
Symbols	Parameters and Conditions	Units	Typical
I_{CC}	Circuit current (no signal)	mA	6
G_S	Small-signal gain, f = 900 MHz dB	14.5	
	f = 1900 MHz	dB	16
f_{corner}	–3-dB Gain corner frequency	GHz	2.9
P_{1dB}	1 dB Compressed output power at 1900 MHz	dBm	–12.5
P_{SAT}	Saturated output power, f = 1900 MHz	dBm	–6
NF	Noise figure, f = 900 MHz	dB	3.2
	f = 1900 MHz	dB	4.0
RL_{IN}	Input return loss, f = 1900 MHz	dB	10
RL_{OUT}	Output return loss, f = 1900 MHz	dB	12.5
ISOL	Isolation, f = 1900 MHz	dB	30
OIP_3	SSB output third-order intercept, f_1 = 500 MHz, f_2 = 510 MHz	dBm	–3
	f_1 = 1000 MHz, f_2 = 1010 MHz	dBm	–3
	f_1 = 1900 MHz, f_2 = 1902 MHz	dBm	–3.5
	f_1 = 2000 MHz, f_2 = 2010 MHz	dBm	–4
$R_{TH(J-A)}$	Thermal resistance (junction to ambient)		
	Free air	°C/W	
	Mounted on a 50 × 50 × 1.6 mm epoxy glass PWB	°C/W	

In modeling the UPC2749's three transistors, we chose to use our favorite, the BFP420, which is not without its headaches at the higher frequencies. Figure 3-137 shows the circuit diagram, which closely resembles the one published by NEC; again, the only hard facts we had were the IC's basic circuit topology, gain, and dc current. Given those numbers, we staggered the current in the three devices by setting them at 0.64 mA for the first stage, 1.85 mA for the second stage, and 3 mA for the third stage. It should be noted that the two output transistors are not unlike the approach used for the Hewlett-Packard MSA-0735 (Figure 3-133). Taking (hopefully) all the right assumptions, we end up with a design that closely resembles the manufacturer's specifications for both power consumption, gain, and noise. This can be verified by looking at the results plotted in Figure 3-138.

The final example is a more complex derivative of the Philips NE/SA5204A amplifier, which was intended for standard 6-V operation and has a fairly "normal" circuit arrangement. This can be seen from Figure 3-139. A much more elaborate cousin of this is the low-voltage NEC UPC2710 amplifier. According to its specifications, it operates at 5 V but has significantly more gain than the NE/SA5204. Table 3-13 shows its electrical specifications and Figure 3-140 shows its measured gain as published by the manufacturer.

The actual interior of the circuit is shown in Figure 3-141. It consists of the standard preamplifier and then a Darlington configuration with the collectors tied together. The middle section of the circuit acts as a power supply responsible for the bias relative to the emitter of the first transistor influencing the rest of the circuit as well. Since NEC mentioned that they use their NESAT III process (f_T of 20 GHz at V_{ce} = 3) in producing the UPC2710T and we

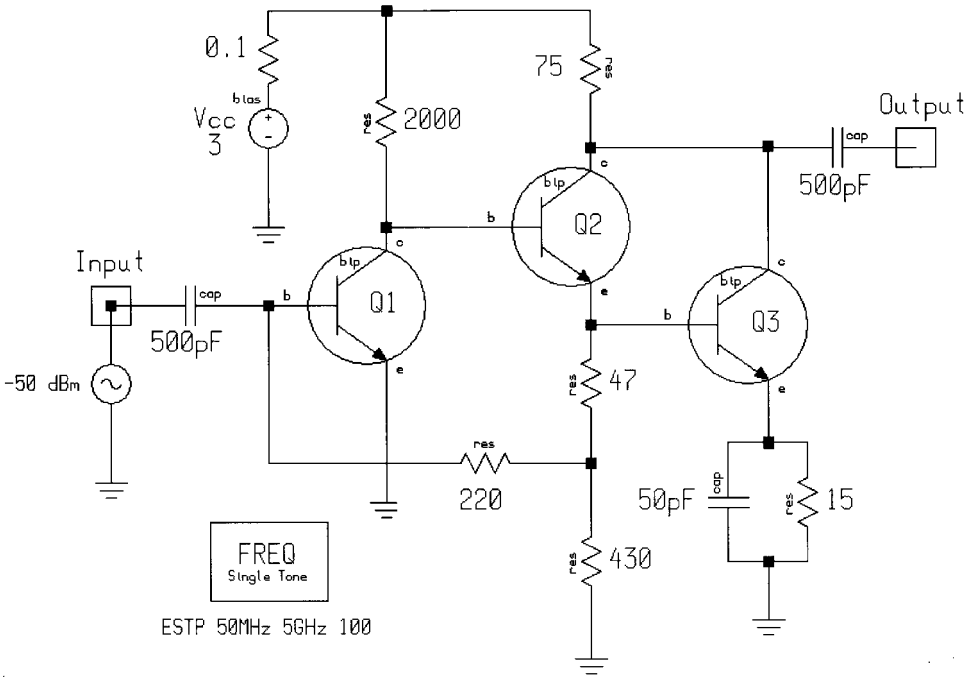


Figure 3-137 Circuit of the UPC2749 MMIC. The component values were not supplied by the manufacturer.

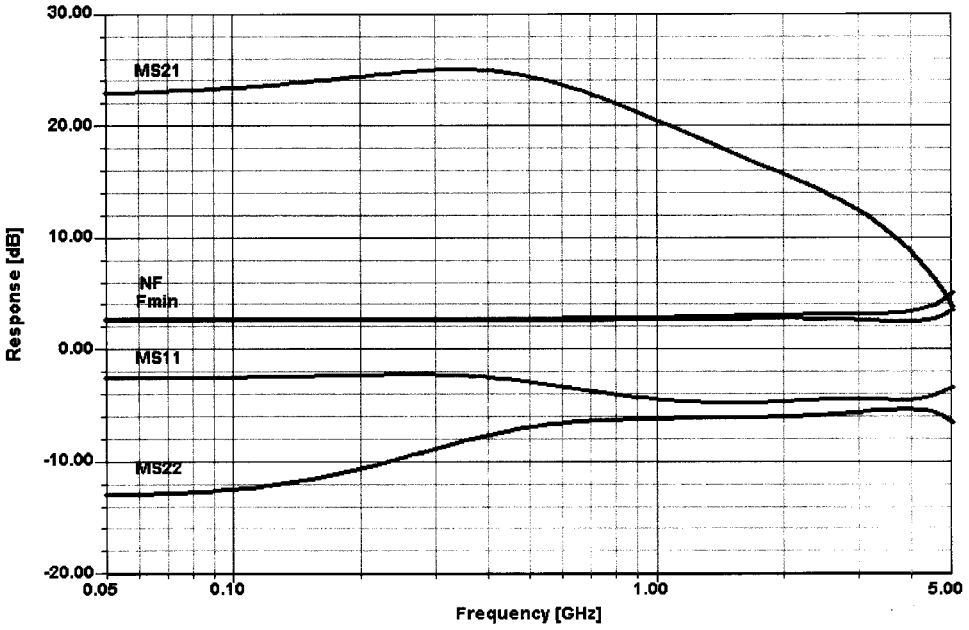


Figure 3-138 Simulated gain, matching, and noise performance of the NEC UP2749 MMIC.

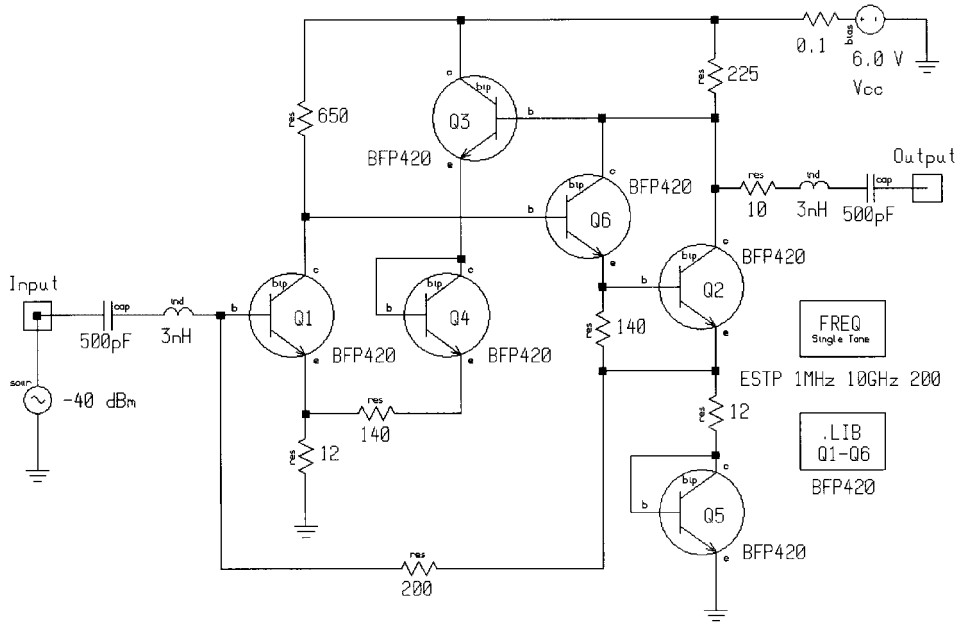


Figure 3-139 Schematic of the Philips NE/SA5204A amplifier IC entered for simulation.

Table 3-13 UPC2710 electrical specifications

	T_A = 25 °C, f = 0.5 GHz, V_{CC} = 5 V)		
	Part Number Package Outline		UPC2710T TO6
Symbols	Parameters and Conditions	Units	Typical
I_{CC}	Circuit current (no signal)	mA	22
G_S	Small-signal gain	dB	33
f_{corner}	–3-dB Gain corner frequency	GHz	1.0
ΔG_S	Gain flatness, f = 0.1–0.6 GHz f = 0.1–0.8 GHz	dB	±0.8
P_{SAT}	Saturated output power	dBm	13.5
P_{1dB}	Output power at 1dB compression point	dBm	7.5
NF	Noise figure	dB	3.5
RL_{IN}	Input return loss	dB	6
RL_{OUT}	Output return loss	dB	12
ISOL	Isolation	dB	39
ΔG_T	Gain–temperature coefficient	dB/°C	–0.006
R_{TH}	Thermal resistance (junction to ambient)	°C/W	

obtained some information regarding it, we decided to use our standard BFP420 (f_T of 25 GHz; see the datasheet in Chapter 2) for the purpose of simulation.

In our effort to simulate this circuit, we chose a combination of values that pretty much complies with the manufacturer's published specifications, but S_{22} is not close, as shown in Figure 3-142. We stopped at this point and invite our readers with CAD capabilities to provide the final touch in meeting all the published specifications for the circuit. We are particularly

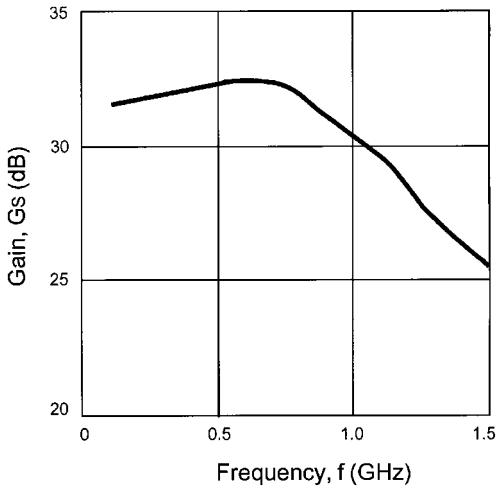


Figure 3-140 Manufacturer-supplied gain versus frequency for the NEC UPC2710T MMIC.

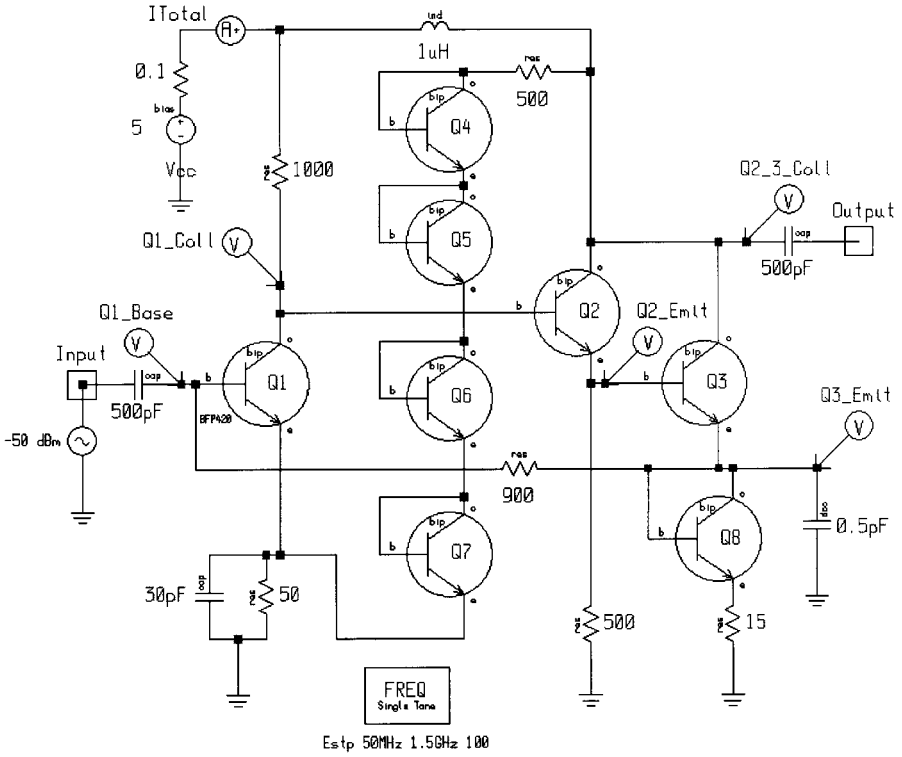


Figure 3-141 Schematic of the UPC2710T silicon MMIC entered for simulation.

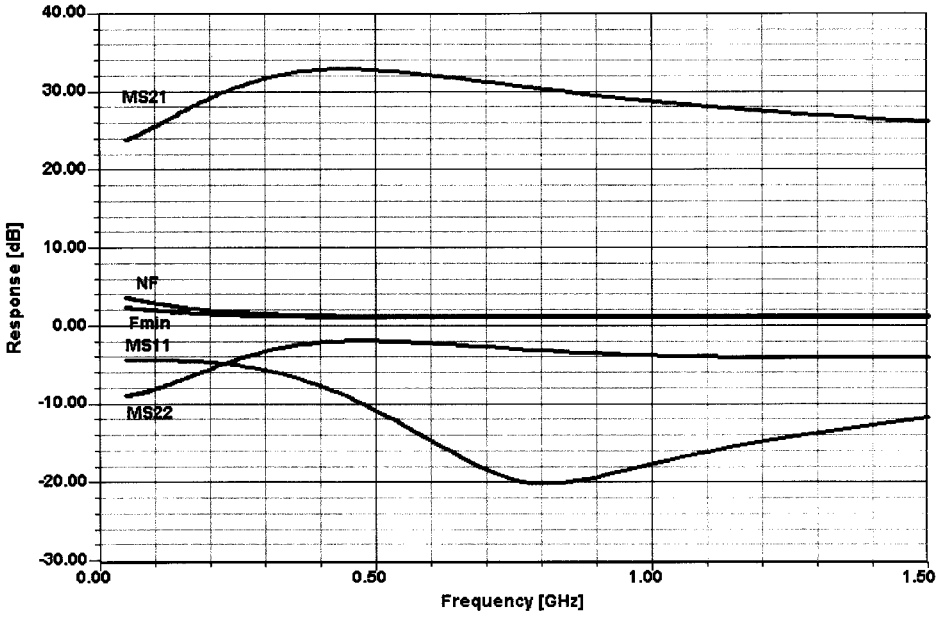


Figure 3-142 Simulated gain, match, and noise performance of the UPC2710T.

recommending such an effort because it shows very nicely the interaction of all the circuit elements in a fairly drastic way. Because of this, we admire the IC design even more, since the manufacturer has to sell a pretty consistent amplifier as a function of different production runs.

3-5-1 Stability of Multistage Amplifiers

Stability, needless to say, is a big issue, and the delay of the various stages adds to the phase shift, resulting in possible instabilities. The fact that S_{11} and/or S_{22} can become larger than 0 (in dB) means there is the potential of oscillating, depending on the reactances available. If a cable of inappropriate length is added, hell will break loose. Figure 3-143 shows the dangerous peaking of our simulation of the Philips NE5204A, in which the internal compensation was intentionally not assumed correctly. We will all agree that the high-end performance of this amplifier can generate high anxiety. As the production process of the aging device becomes more modern—meaning that transistors with much higher f_T will be the basis for the production—it will become quite difficult to maintain a previously achieved stability. This, by the way, applies also to discrete transistors, which for reasons of economics and multiple manufacturers will be improved over time. Such evolution results in generally smaller base-spreading resistances and capacitances, and other parameter changes, which, overall, cause headaches in manufacturing. We will leave the topic of three-stage amplifiers on this note.

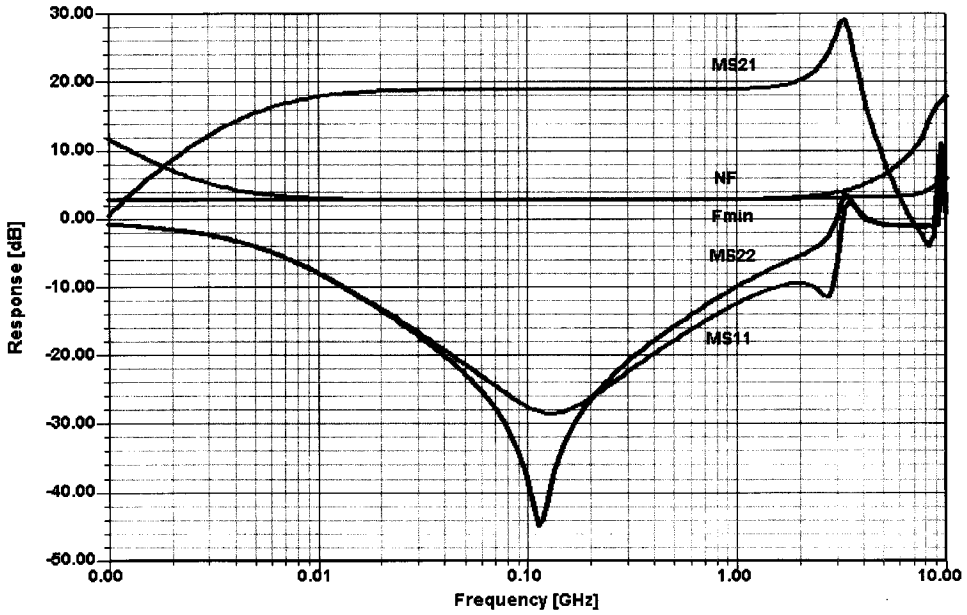


Figure 3-143 Simulated gain, match, and noise performance of the NE5204A IC.

3-6 A NOVEL APPROACH TO VOLTAGE-CONTROLLED TUNED FILTERS INCLUDING CAD VALIDATION [28]

Modern receivers control input stages as well as the oscillator band and frequency by electrical rather than mechanical means. Tuning is accomplished by voltage-sensitive capacitors (varactor diodes), and band switching by diodes with low forward conductance. Since the wireless band (essentially 400 MHz to 2.4 GHz) is so full of strong signals, the use of a tracking filter is desired as a solution to improve the performance and prevent second-order IMD products or other undesired overload effects. The dc control voltage needed for the filter can easily be derived from the VCO control voltage. There may be a small dc offset, depending on the IF used.

3-6-1 Diode Performance

The capacitance versus voltage curves of a varactor diode depend on the variation of the impurity density with the distance from the junction. When the distribution is constant, there is an "abrupt junction" and capacitance follows the law

$$C = \frac{K}{(V_d + V)^{1/2}} \tag{3-183}$$

where V_d is the contact potential of the diode and V is applied voltage.

Such a junction is well approximated by an alloyed junction diode. Other impurity distribution profiles give rise to other variations, and the above equation is usually modified to

$$C = \frac{K}{(V_d + V)^n} \tag{3-184}$$

where n depends on the diffusion profile and $C_0 = K/V_d^n$.

A so-called graded junction, having a linear decrease in impurity density with the distance from the junction, has a value of n. This is approximated in a diffused junction.

In all cases these are theoretical equations, and limitations on the control of the impurity pattern can result in a curve that does not have such a simple expression. In this case, the coefficient n is thought of as varying with voltage. If the impurity density increases away from the junction, a value of n higher than 0.5 can be obtained. Such junctions are called hyperabrupt. A typical n value for a hyperabrupt junction is about 0.75. Such capacitors are used primarily to achieve a large tuning range for a given voltage change. Figure 3-144 shows the capacitance–voltage variation for the abrupt and graded junctions as well as for a particular hyperabrupt-junction diode. Varactor diodes are available from a number of manufacturers, such as Motorola, Siemens, and Philips. Maximum values range from a few to several hundred picofarads, and useful capacitance ratios range from about 5 to 15.

Figure 3-145 shows three typical circuits that are used with varactor tuning diodes. In all cases, the voltage is applied through a large resistor R_e or, better, an RF choke in series with a small resistor. The resistance is shunted across the lower diode and may be converted to a shunt load resistor across the inductance to estimate Q. The diode also has losses that may result in lowering the circuit Q at high capacitance, when the frequency is sufficiently high. This must be considered in the circuit design.

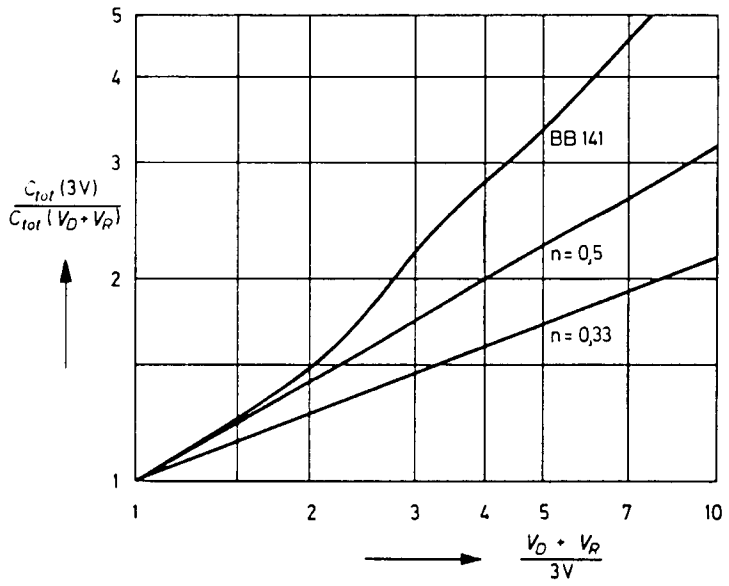


Figure 3-144 Voltage-dependent change of capacitance of different types of diodes. The BB141 is a hyperabrupt-junction diode with $n = 0.75$.

The frequency-dependent performance is not determined solely by applying the dc tuning voltage to Eq. (3-184). If the RF voltage is sufficient to drive the diode into conduction on peaks, an average current will flow in the circuits of Figure 3-145, which will increase the bias voltage. The current is impulsive, giving rise to various harmonics of the circuit. Even in the absence of conduction, Eq. (3-184) deals only with the small-signal capacitance. When the RF voltage varies over a relatively large range, the capacitance changes. In this case, Eq. (3-184) must be changed to

$$\frac{dQ}{dV} = \frac{K}{(V + V_d)^n} \tag{3-185}$$

Here Q is the charge on the capacitor. When this relation is substituted in the circuit differential equation, it produces a nonlinear differential equation, dependent on the parameter n. Thus, the varactor my generate direct current and harmonics of the fundamental frequency. Unless the diodes are driven into conduction at some point in the cycle, the direct current must remain zero.

The current of Figure 3-145c can be shown to eliminate the even harmonics, and it permits a substantially larger RF voltage without conduction than either circuit in Figure 3-145a or 3-145b. When $n = 0.5$, only the second harmonic is generated by the capacitor, and this can be eliminated by the back-to-back connection of the diode pair. Integrating Eq. (3-185), we find

$$Q + Q_A = \frac{K}{1 - n}(V + V_d)^{1-n}$$

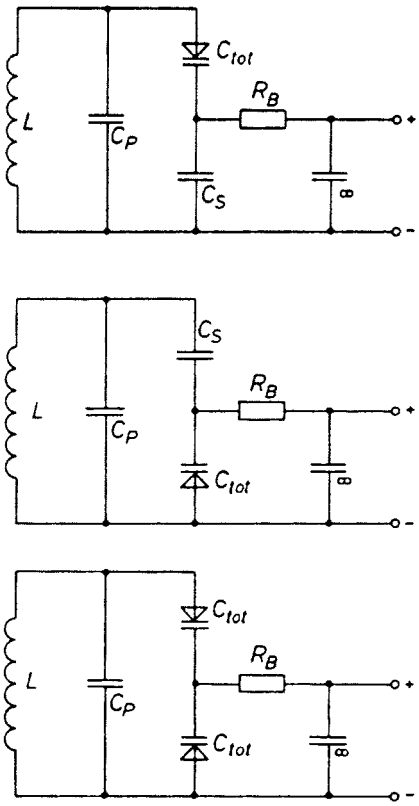


Figure 3-145 Various configurations to apply tuning diodes in a tuned circuit. The bottom version shows the lowest distortion.

$$= \frac{C_v}{1-n}(V + V_d) \tag{3-186}$$

C_v is the value of Eq. (3-184) for applied voltage V, and Q_A is a constant of integration. By letting $V = V_1 + v$ and $Q = Q_1 + q$, where the lowercase letters represent the varying RF and the uppercase letters indicate the values of bias when RF is absent, it follows that

$$q + Q_1 + Q_A = \frac{K}{1-n}[v + (V_1 + V_d)]^{1-n} \tag{3-187}$$

$$1 + \frac{v}{V'} = \left(1 + \frac{q}{Q'}\right)^{1/(1-n)} \tag{3-188}$$

where $V' = V_1 + V_d$ and $Q' = Q_1 + Q_A$. For the back-to-back connection of identical diodes, $K_{11} = K_{12} = K_1$, $V_1' = V_2' = V'$, $Q_1' = Q_2' = Q'$, $q = q_1 = -q_2$, and $v = v_1 - v_2$. Here the new subscripts 1 and 2 refer to the top and bottom diodes, respectively; v is the RF voltage across

the pair in series, and q is the charge transferred through the pair in series. This notation results in

$$\frac{v}{V'} \equiv \frac{v_1 - v_2}{V'} = \left(1 - \frac{q}{Q'}\right)^{1/(1-n)} - \left(1 - \frac{q}{Q'}\right)^{1/(1-n)} \tag{3-189}$$

For all n, this eliminates the even powers of q, and hence the even harmonics. This can be shown by expanding Eq. (3-189) in a series expansion and performing term by term combination of the equal powers of q. In the particular case $n = \frac{1}{2}$, $v/V' = 4q/Q'$, and the circuit becomes linear.

The equations hold as long as the absolute value of v_1/V' is less than unity, so that there is no conduction. At the point of conduction, the total value of v/V' may be calculated by noting that when $v_1/V' = 1$, $q/Q' = -1$, so $q_2/Q' = 1$, $v_2/V = 3$, and $v/V = -4$. The single-diode circuits conduct at $v/V' = -1$, so the peak RF voltage should not exceed this. The back-to-back configuration can provide a fourfold increase in RF voltage handling over the single diode. For all values of n, the back-to-back configuration allows an increase in the peak-to-peak voltage without conduction. For some hyperabrupt values of n, such that $1/(1 - n)$ is an integer, many of the higher-order odd harmonics are eliminated, although only $n = \frac{1}{2}$ provides elimination of the third harmonic. For example, $n = \frac{2}{3}$ results in $1/(1 - n) = 3$. The fifth harmonic and higher odd harmonics are eliminated, and the peak-to-peak RF without conduction is increased eightfold; for $n = \frac{3}{4}$ the harmonics 7 and above are eliminated, and the RF peak is increased 16 times. It must be noted in these cases that the RF peak at the fundamental may not increase so much, since the RF voltage includes the harmonic voltages.

Since the equations are only approximate, not all harmonics are eliminated, and the RF voltage at conduction, for the back-to-back circuit, may be different from that predicted. For example, abrupt-junction diodes tend to have n of about 0.46–0.48 rather than exactly 0.5. Hyperabrupt junctions tend to have substantial changes in n with voltage. The diode illustrated in Figure 3-144 shows a variation from about 0.6 at low bias to about 0.9 at higher voltages, with wiggles from 0.67 to 1.1 in the midrange. The value of V_d for varactor diodes tends to be in the vicinity of 0.7 V.

3-6-2 A VHF Example

The application of tuning diodes in double-tuned circuits in TV tuners has been common practice for many years. Figure 3-146 shows the circuit diagram. The input impedance of 50 Ω gets transformed up to 10 kΩ. The tuned circuits consist of the 0.3-μH inductor and two sets of antiparallel diodes. By dividing the RF current in the tuned circuit and using several diodes instead of just one pair, intermodulation distortion is reduced.

The coupling between the two tuned circuits is tuned via the 6-nH inductor that is common to both circuits. This type of inductance is usually printed on the circuit board. The diode parameters used for this application were equivalent to the Siemens BB515 diode. The frequency response of this circuit is shown in Figure 3-147.

The coupling is less than critical. This results in an insertion loss of about 2 dB and relatively steep passband sides. Once the circuit's large-signal performance (Figure 3-148) is seen, a third-order intercept point of about –2 dBm is not so unexpected. The reason for this poor performance is the high impedance (high L/C ratio), which provides a large RF voltage swing across the diodes. A better approach appears to be using even more diodes and at the same time changing the impedance ratio (L/C ratio).

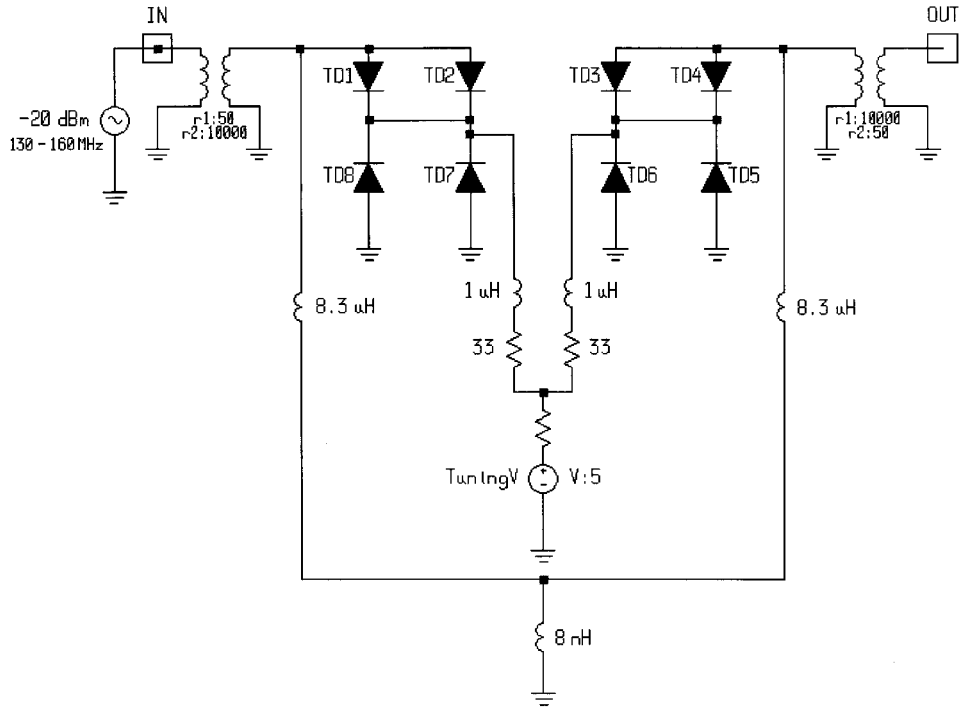


Figure 3-146 Double-tuned filter at 161 MHz using hyperabrupt-junction tuning diodes. By using several parallel diodes, the IMD performance improves.

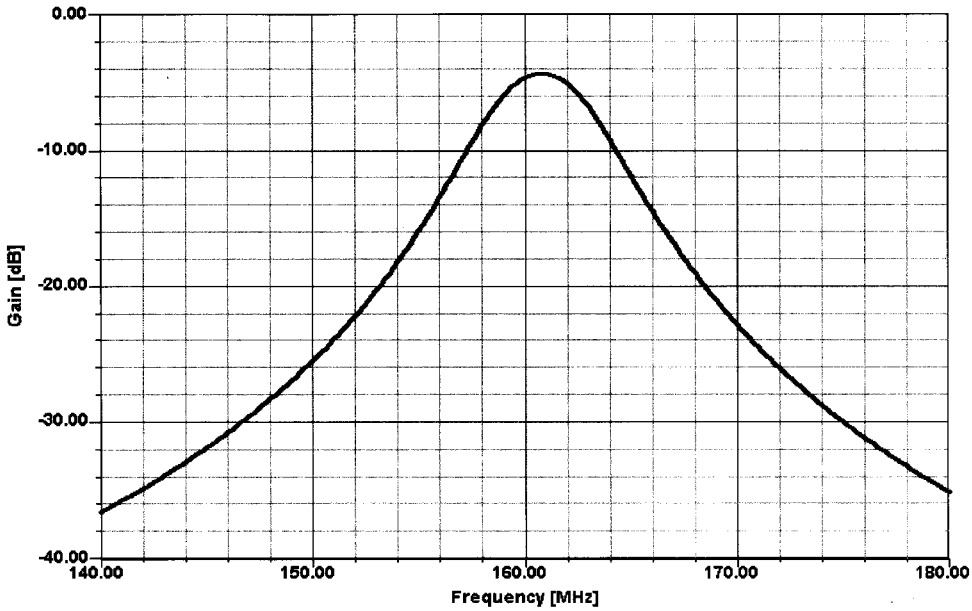


Figure 3-147 Frequency response of the tuned filter shown in Figure 3-146. The circuit is undercoupled (less than transitional coupling); $Q \times k < 1$.

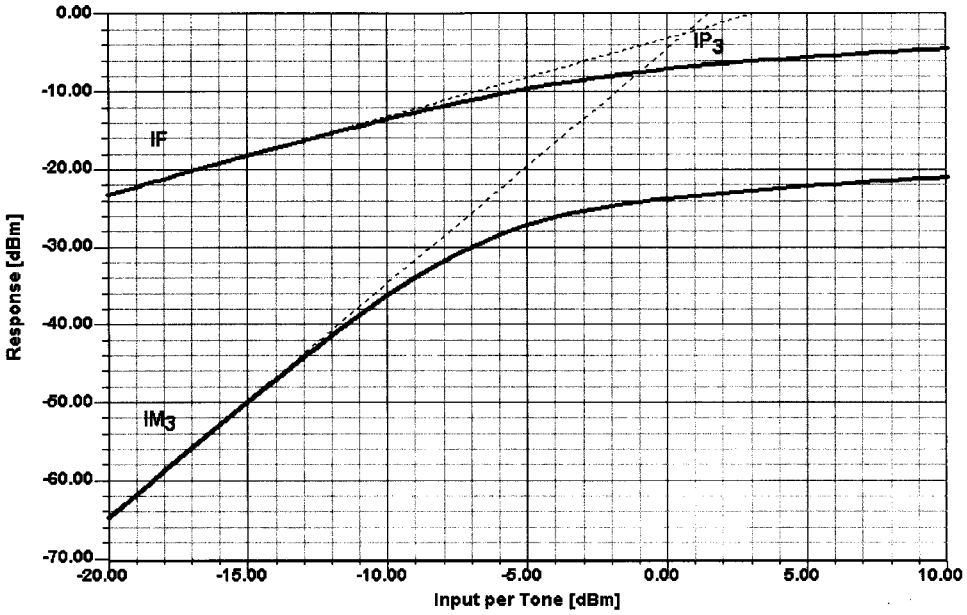


Figure 3-148 Prediction of intercept point of the double-tuned circuit shown in Figure 3-146. Note the compression of both the input signal and the IMD product.

3-6-3 An HF/VHF Voltage-Controlled Filter

The above example used a step-up procedure typically done by a tap at the input and output inductance. This method allows an impedance transformation; however, if one desires to change it into a series-tuned arrangement, it has to be done with a transformer. The large-signal conditions in the frequency range from 10 to 30 MHz on a medium- to large-sized antenna are equivalent to, if not worse than, the conditions for VHF operation. The only exception would be line of sight into the transmitter such as a tower in the middle of the city. Examples of such hostile conditions would be any large city, such as Munich, New York, Miami, Chicago, and San Francisco, where the authors had significant experience with intermodulation distortion problems.

By translating the circuit into a low-impedance-drive arrangement, building it symmetrically, and reducing the tuning range somewhat by capacitors and many series diode pairs, the filter's large-signal performance was significantly improved. The tuning diodes had about 125 pF at 1 Vdc per unit. Initial tests trying to use a high-capacitance diode, such as the BB112, resulted in much higher IMD products. The B112 has a 1-V capacitance of 470 pF down to about 20 pF at 8 V. Standard diodes show about 30 pF at 1 V, while the diodes used in the experiment had a 125-pF capacitance at 1 V. These experimental diodes were made available from a well-known manufacturer for this evaluation. Figure 3-149 shows the circuit diagram.

A step-down transformer drives the tuned circuit with a source impedance of 12.5 Ω. The circuit is symmetrical. There are two 0.8-μH inductors and 2×5 diodes in the loop. The output of the circuit is transformed back up to 50 Ω. The tuning voltage is supplied via a heavily filtered arrangement. This circuit has actually been implemented in Rohde & Schwarz field strength meter equipment. They guarantee an intercept point of about +20 dBm. The selectivity of these circuits is quite reasonable and mostly intended to reduce

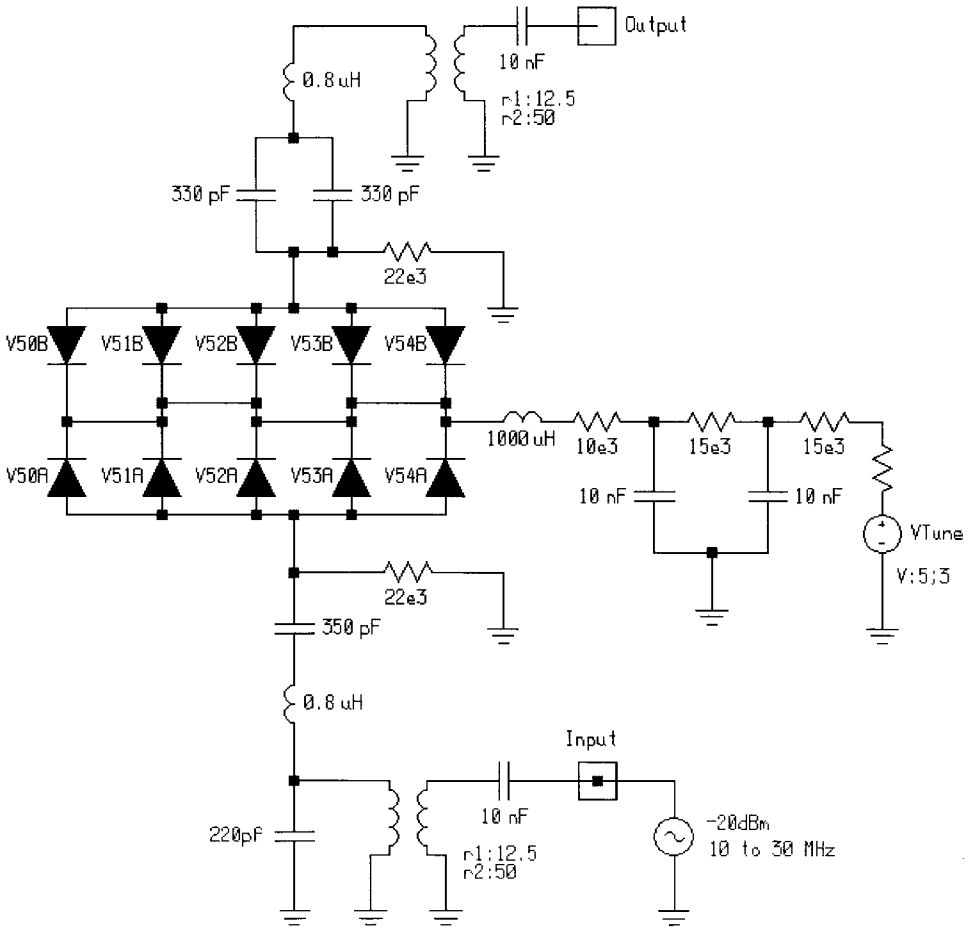


Figure 3-149 High dynamic range of a 10–20-MHz filter.

second-order intermodulation distortion products by about 10 dB, at the same time having a high third-order intercept point. Figure 3-150 shows the selectivity curves at two different tuning voltages.

The most interesting number, however, is the third-order intercept point (already determined as about 20 dBm), which still had to be simulated. This sometimes sounds like a contradiction: Once the measured values are available and are acceptable, why would one want—or, rather, need—to do a simulation besides the necessity to develop a high-input intercept filter? It was desirable to validate the nonlinear models and prove that the above-mentioned equations will hold true. This type of simulation is now more difficult because the number of nonlinear elements went from four to ten, and some numerical problems, such as convergence difficulties, can be expected, and the effect of harmonic frequency cancellation (compensation) can also be seen. By using diode combinations that result in $1(1 - n)$, n being an integer number, these IMD products can drastically be reduced. This implies that each of the five diode pairs has a selected value for n to meet this condition. Later, a final attempt will be made to improve the first VHF filter with the proper diode combinations.

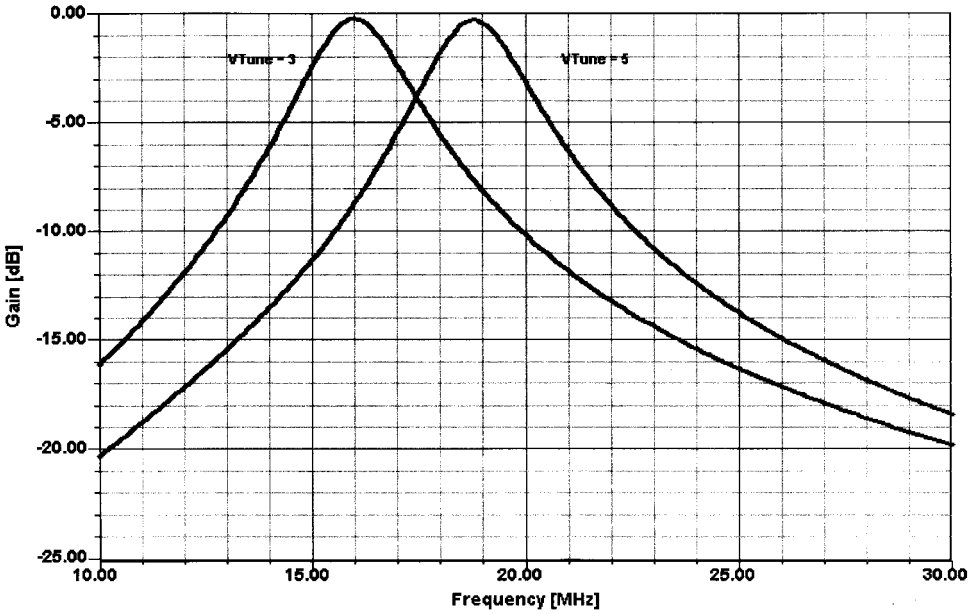


Figure 3-150 Frequency response of the filter shown in Figure 3-149. Filters of this type are intended to reduce second-order IMD by providing about 20-dB suppression at half the center frequency.

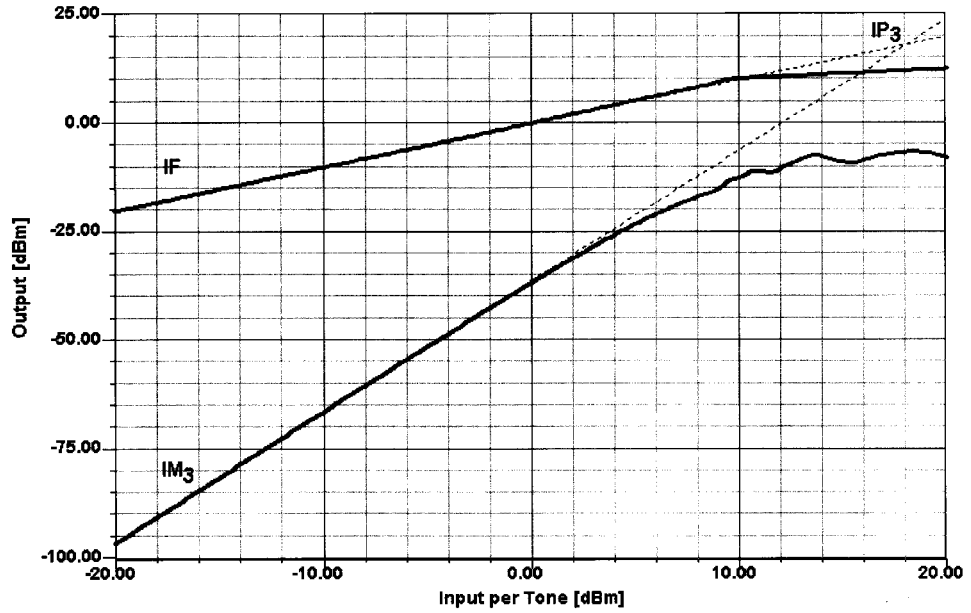


Figure 3-151 Prediction of third-order intercept point of the filter shown in Figure 3-149. The reason for the curves on the IMD plot lies in the interaction between the ten nonlinear devices.

Figure 3-151 shows the calculation of third-order intercept point for the arrangement shown in Figure 3-149. This number has now increased to +18 dBm. The difference between selection and measurement is approximately 2 dB. The reason for this more pessimistic value compared to the measured value probably has to do with slight variations of the exponent of the diodes' voltage-dependent capacitance. Therefore, the authors consider both the circuit performance and the simulation accuracy to be extremely good. This type of circuit, as mentioned earlier, has wide application in oscillator circuits. Danzeisen of Rohde & Schwarz was probably the first to have used this type of circuit by paralleling many diodes for improved performance.

3-6-4 Improving the VHF Filter

By selecting the appropriate diode combinations and circuit modifications as shown in Figure 3-152, a significant IMD improvement of the circuit shown in Figure 3-146 is obtained. A special shunt arrangement of several diodes with different exponents has been developed, which allows its value to be "adjusted." This circuit, for which a patent has been applied, showed an improvement from –2 dBm to 32 dBm. The same high operating Q was maintained. Figure 3-153 shows the frequency response after the modification. Note that a coupling slightly greater than critical ($Q \times k = 1.1$) has been selected.

3-6-5 Conclusion

After explaining some of the nonlinearities in mathematical terms, we have given some examples of voltage-tuned circuits and discussed their large-signal performance. This section has also shown that modern CAD tools can accurately predict the performance of such circuits.

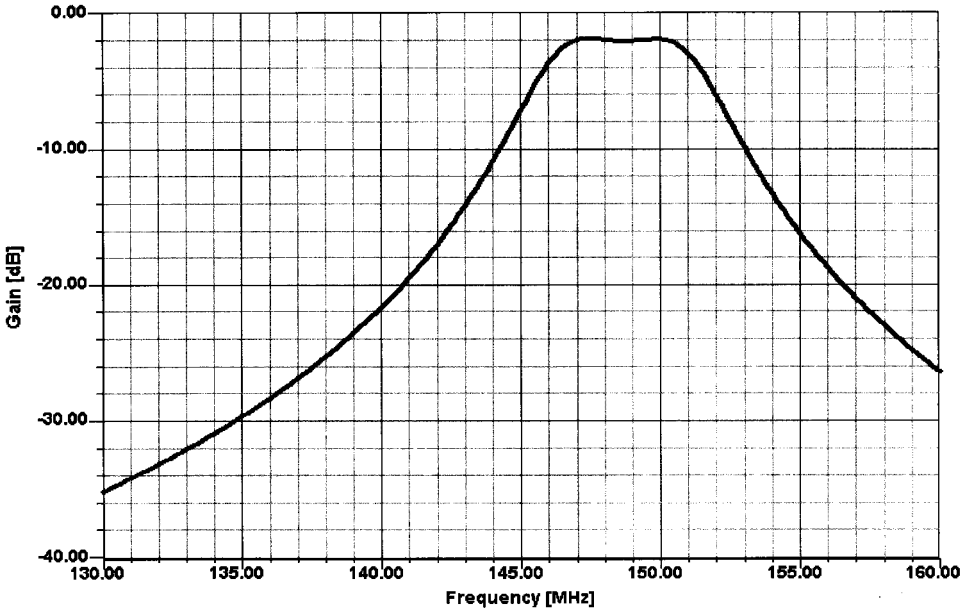


Figure 3-152 The frequency response of the improved circuit derived from Figure 3-146.

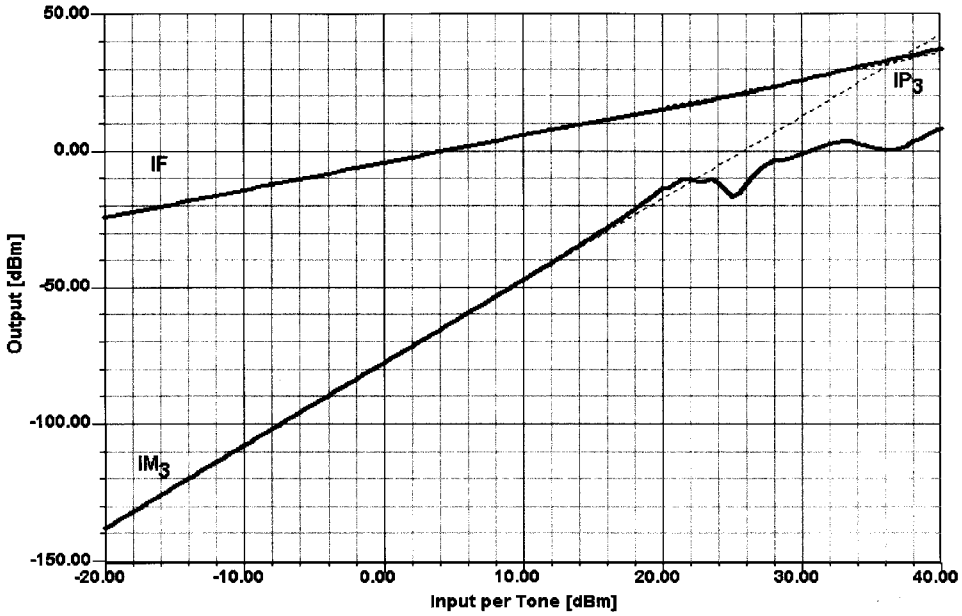


Figure 3-153 Predicted IMD performance of the improved version of the circuit shown in Figure 3-146. This type of improvement is significant for all applications. A patent is in the process of being obtained for this; therefore, the circuit cannot be disclosed at present.

3-7 DIFFERENTIAL AMPLIFIERS

The differential amplifier (Figure 3-154) goes back many years and also was the first step from TTL to ECL. The differential amplifier consists of an emitter follower (Q1) and a grounded-base stage (Q2). The constant-current generator (Q3) between ground and the emitters of Q1 and Q2 determines the current for these stages; the dc operating currents of Q1 and Q1 are the same. Most modern integrated circuits use this differential amplifier in one way or another.

If the circuit is operated about its point of symmetry, the dc component of the current through either Q1 or Q2 remains half of the emitter constant current through Q3 for all symmetrical driving signals. In addition, for input signals that are aperiodic, no even harmonics are generated. This assumes that the stage is driven symmetrically at both inputs. In our case, the differential amplifier has its second stage grounded, meaning that there is a base-to-ground capacitor, and the stage is asymmetrically fed. As the input voltage is increased, the differential stage becomes a limiter, which limits the current between plus and minus I_S, with I_S being the source current. The resulting output will become a square wave. This stage can also be used as a line receiver, translating a sine wave to an output logic suitable for driving emitter-coupled logic (ECL) stages.

One can assume that the differential amplifier has a transconductance of

$$g_m = \frac{\alpha g_{\text{in}}}{2} \tag{3-190}$$

where α is the ratio of collector to emitter current. Within the limits of the output being a sine wave, this stage has very low distortion. The output current now is determined by

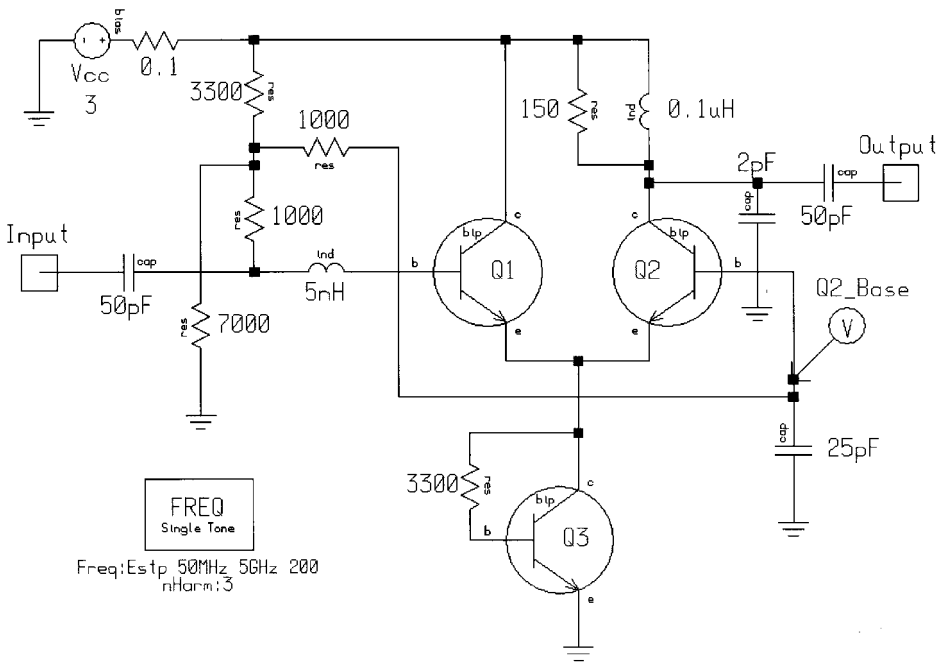


Figure 3-154 The differential amplifier schematic. The transistors are Siemens BFP420s. Figure 3-155 shows the circuit's frequency-dependent gain, matching, and noise performance.

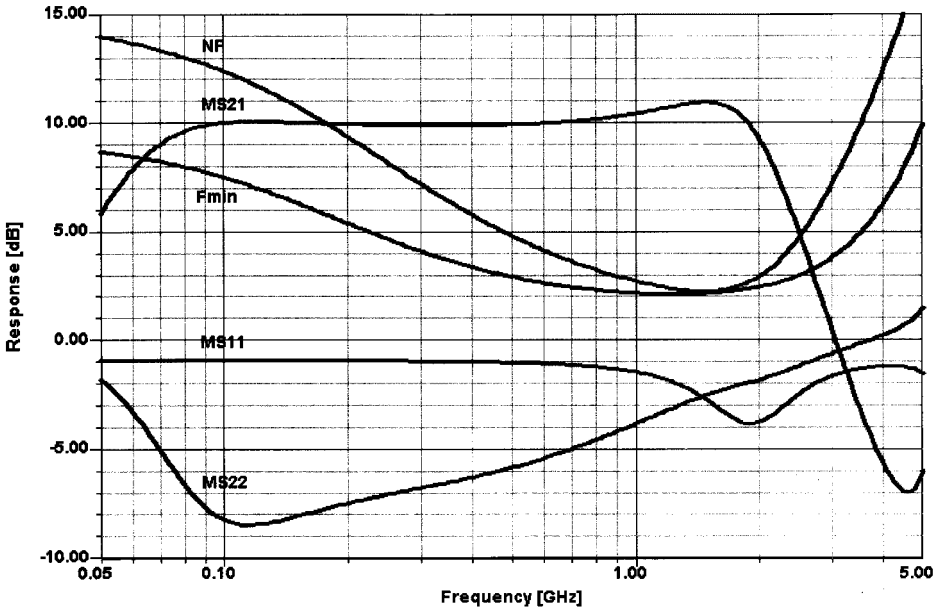


Figure 3-155 Frequency-dependent gain, matching, and noise performance of the differential amplifier.

$$i = \frac{I_{\text{in}}}{2} \tanh\left(\frac{x}{2} \cos \omega t\right) \tag{3-191}$$

where ωt is the input frequency and I_{in} is the current forced by Q3. The reason why the transconductance has to be divided by 2 comes from the fact that the current is split into two equal components for Q1 and Q2. Multiplying the new transconductance with the output load gives the voltage gain of this amplifier; its power gain can be obtained by multiplying this with $\sqrt{\text{output resistance} \div \text{load resistance}}$.

Having done the linear analysis and having learned that because of the hyperbolic tangent functions this amplifier can also be used as a line receiver, and converting sinusoidal waves to square waves, it may be useful to take a closer look at these issues. If we take a CAD oscilloscope, we see the resulting output waveform being a nonequal duty cycle output voltage, which results in high harmonic content. The amount of harmonic content can be controlled by the drive power and by the actual bias. Figure 3-156 shows the resulting output waveform.

If we now switch to a CAD spectrum analyzer, we can evaluate the harmonic content (Figure 3-157). Figure 3-158 shows the output transistor's dc I–V curves and ac load line under these conditions. Based on the previously defined conducting angle of the differential amplifier, one can optimize for harmonic content. Therefore, it is logical at this point to look at frequency multipliers, as they are sometimes needed. Finally, we are curious to see the so-called load line of the output transistor, which is, based on the interaction of all the stages, not looking the same way as the expected one, which we know from previous amplifier examples. Part of the reason for this is that for a certain voltage range, the collector voltage is in the negative region, and all kinds of saturation effects result in this surprising waveform. By reducing the input power, one would obtain a more "expected" load line. We invite readers with CAD capabilities to play with this example, as it gives a lot of insight into the operation of ECL stages.

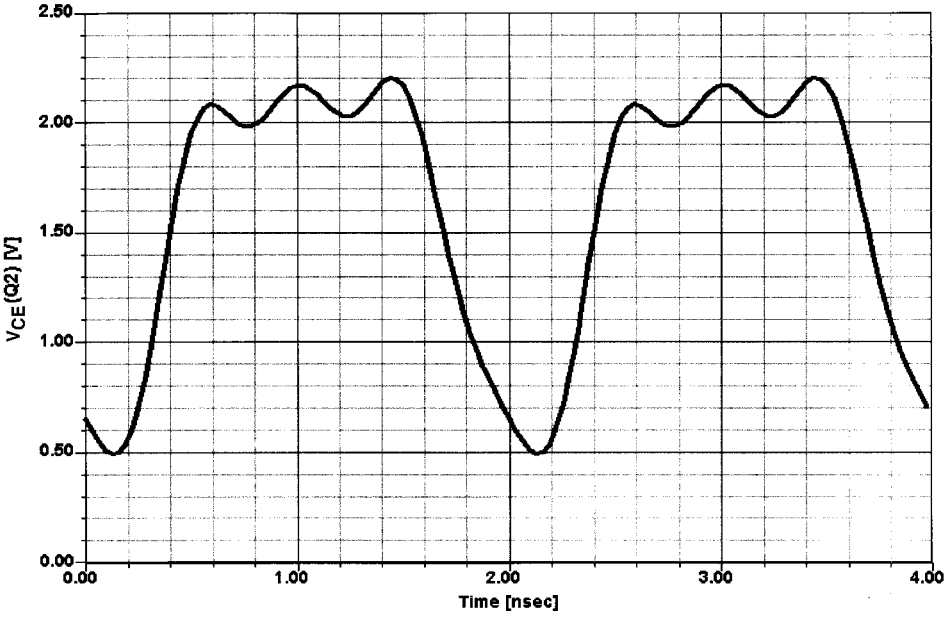


Figure 3-156 V_{CE} versus time for output transistor Q2 with a drive of 10 dBm.

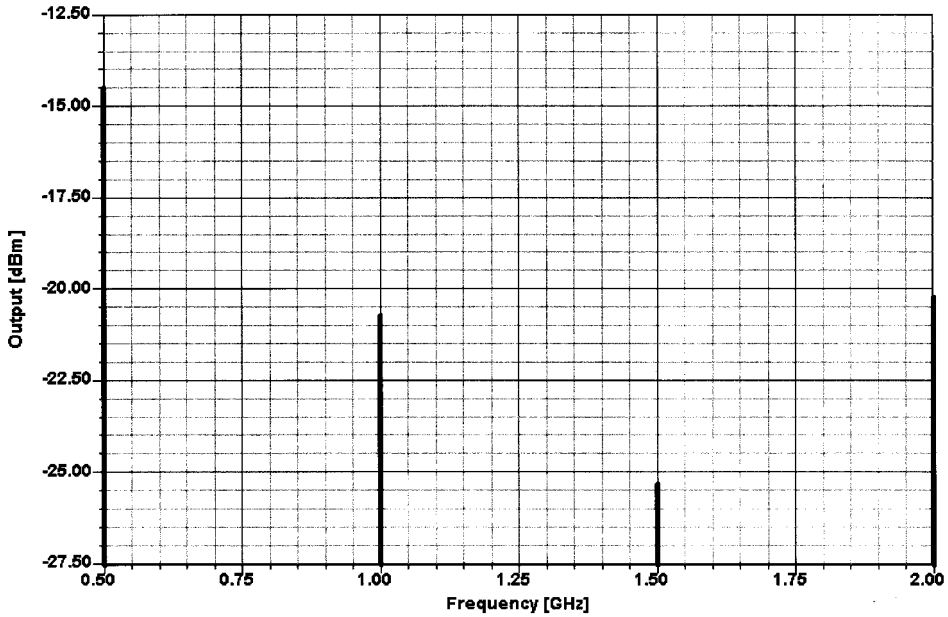


Figure 3-157 Output spectrum of the differential amplifier with a drive of 10 dBm. The biasing is not optimal for frequency multiplication.

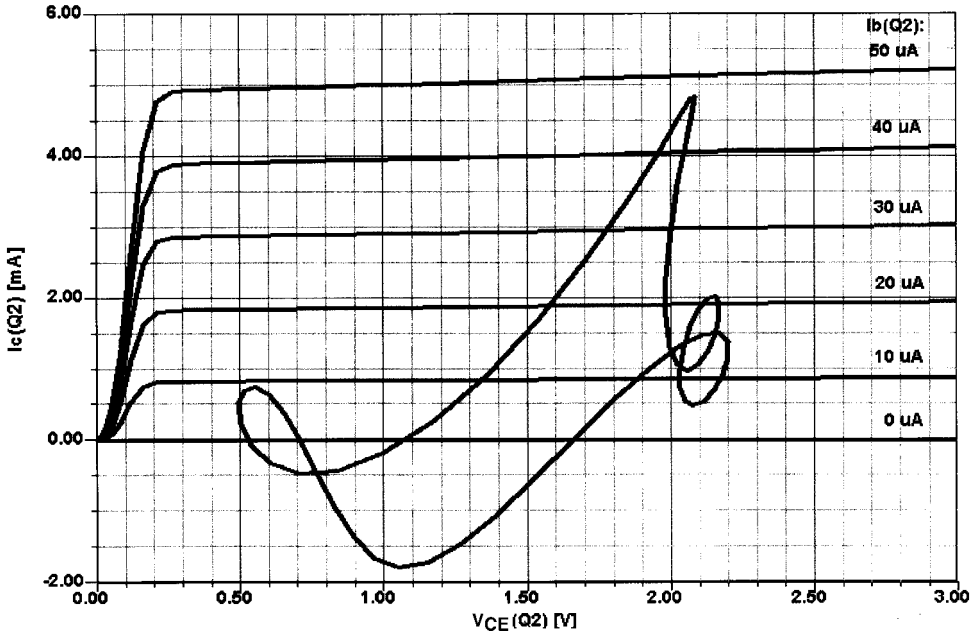


Figure 3-158 The dc *I–V* curves and ac load line for output transistor Q2 with a drive of 10 dBm.

3-8 FREQUENCY DOUBLERS

Having seen that the harmonic content invites us to bias a stage from Class A into Class B or C if our goal is a harmonically rich current, we will now evaluate a single stage that will be biased close to cutoff and driven by a relatively low frequency (500 MHz). To obtain the best results, we will use a parallel notch filter in the output, followed by a high-pass matching circuit to "catch" the desired harmonic. The input frequency is heavily suppressed by the notch filter. Based on its finite Q, there is a limited range over which this frequency can be varied with constant subharmonic attenuation. Figure 3-159 will be our test circuit.

This circuit consists of a bias transistor, Q2. This method has been discussed before and solves the temperature-dependent bias emitter thermal runaway problems through "temperature compensation." The purists among our readers may point out that the currents drawn by Q1 and Q2 are not quite identical (which would give 100% cancellation), but the approach shown is sufficient to do its job. Also, we have chosen a Class AB operation point, which is defined by Q1's 1-mA dc standing current. The conducting angle has not been optimized for the second harmonic—a task that the interested reader can do by using the conducting angle relationship shown at the beginning of this chapter. To give readers with access to nonlinear CAD some homework, we have not set this at the optimum point.

The second transistor, here called Q1 (because of its importance at RF), has a matching network at the input. This is recommended for optimum energy transfer. The matching condition can be validated by examining Figure 3-160. If we look at the higher frequency response, the input shows a possible trend of oscillation (negative loss); this could actually be eliminated by using a low-pass filter at the input instead of the high-pass filter used now. Findings of this type always make us nervous, and we hope that the nonlinear models are sufficiently accurate at higher frequencies to predict such behavior correctly. In the case of the FET doubler mentioned in Chapter 2, we found that the device modeling done by Compact (now Ansoft) gave the best results, the university answers were fairly close, and the device manufacturer data were worst. We have since become highly skeptical about device modeling unless we can validate it ourselves. Synergy Microwave has a complete set of test equipment in the laboratory, including network and spectrum analyzers with special software created by Rohde & Schwarz that can be used in addition to the Scout program to do in-house validation. This was achieved by the use of nonlinear optimization at a drive level of 0 dBm. One could say that this is similar to using a load-pull approach at the input, and then resorting to the large-signal S parameters obtained. Another note of caution: Capacitors at these low values tend to have a very large percentage error; for instance, a 1-pF capacitor, if not selected carefully, can vary between 0.7 and 1.3 pF—a 50% error. Needless to say, this is not acceptable for useful production.

The output consists of a high-pass filter and a notch filter with the purpose of reducing the subharmonic significantly. We have again done this on purpose, leaving room for the reader to come up with a more inventive output circuit like a Cauer-type (same as elliptical) low-pass filter with discrete notches at higher harmonics of the input frequency, such as 1.5 GHz, 2 GHz, and so on. In addition, we have shown three ways how to implement the inductance for the notch filter. This is equally applicable for other inductors in this value range. As frequency increases, this becomes more interesting but one needs to remember that many materials like silicon and gallium arsenide are extremely lossy. The same inductor approach has been used in Figure 5-61 of Chapter 5. For silicon, even the manufacturer becomes an issue. Figure 3-161 shows a top view of a "rectangular" inductor that actually approaches the spiral inductor; its implementation can be seen in Figure 3-162. Modeling this was a

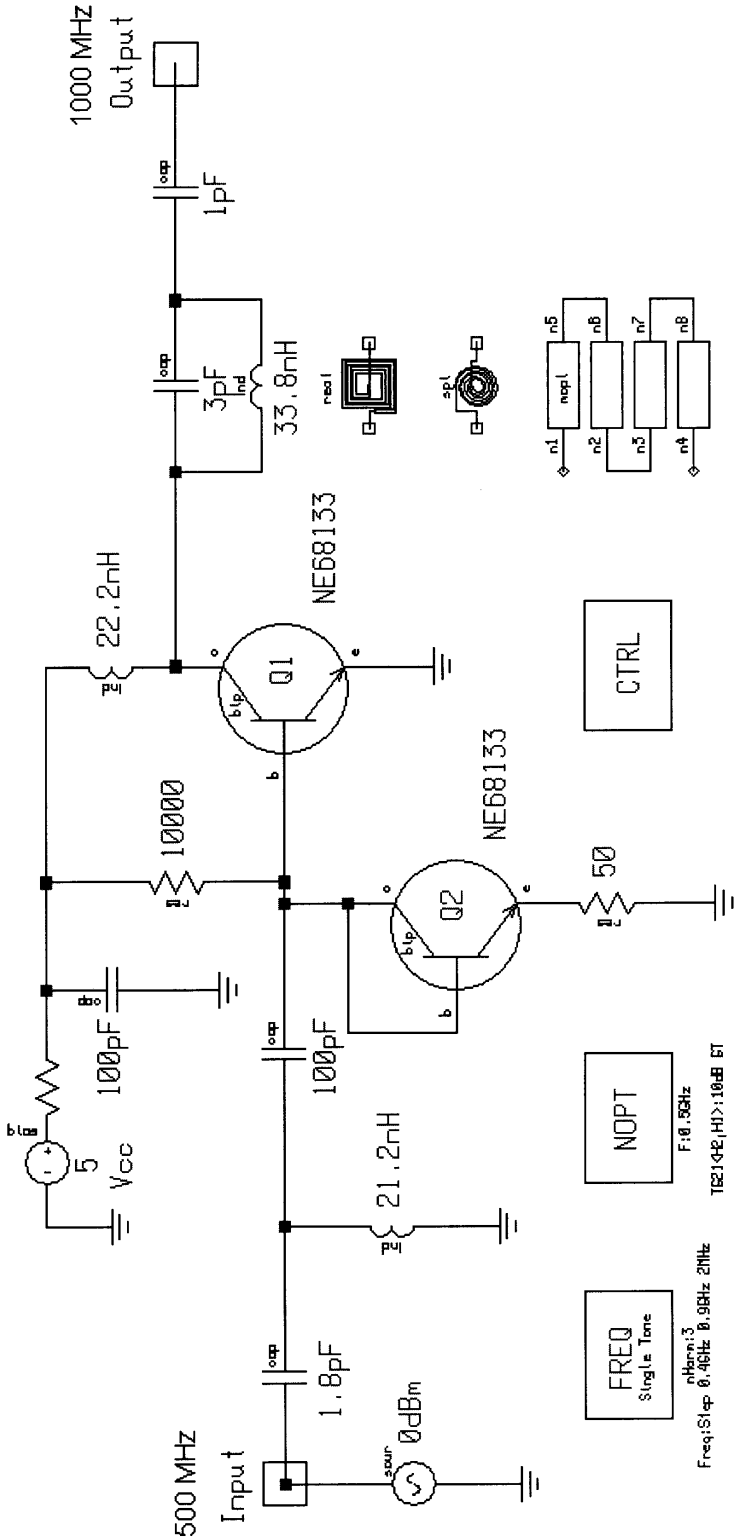


Figure 3-159 BJT frequency doubler schematic. A parallel-tuned trap in Q1's collector attenuates fundamental feedthrough.

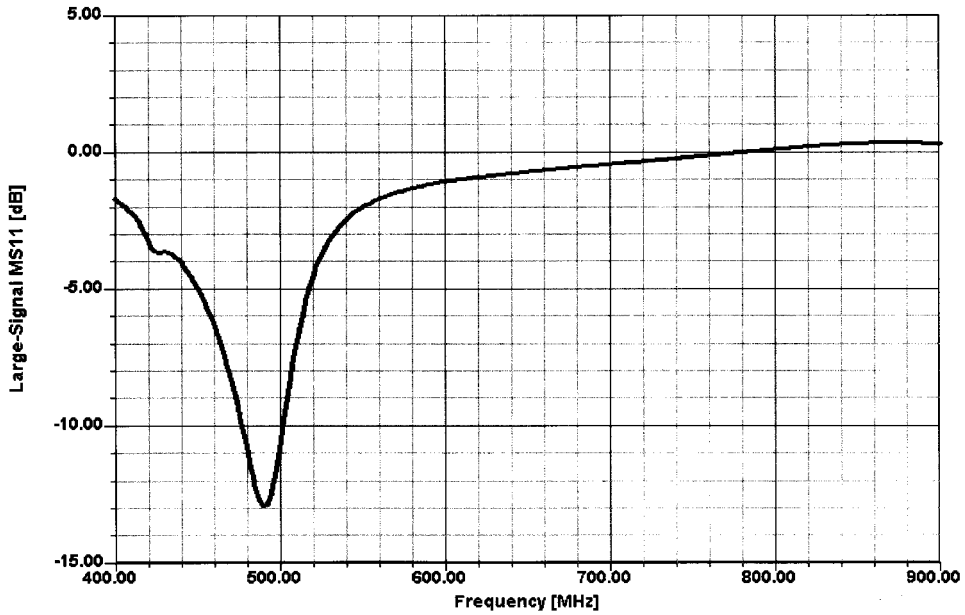


Figure 3-160 Large-signal MS_{11} for the BJT doubler. Note the positive region above 800 MHz.

nightmare, and during our Compact days, we actually had to modify the electromagnetic simulator to accommodate this and get results where measurements and simulation agreed. Figure 3-163 shows the measured and simulated Q as a function of frequency of a silicon-based inductor. In the case of other materials, such as PC board or GaAs, the models in the CAD tools are sufficient to predict the losses of the inductors. Following this initial exercise, Motorola then developed an improved silicon spiral inductor; its frequency-dependent Q is shown in Figure 3-164. Note that this Q was measured and not simulated.

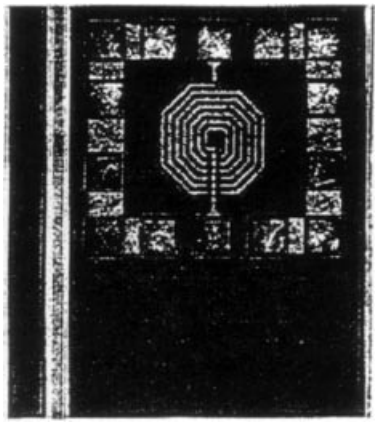


Figure 3-161 A 6.5-turn spiral inductor in silicon.

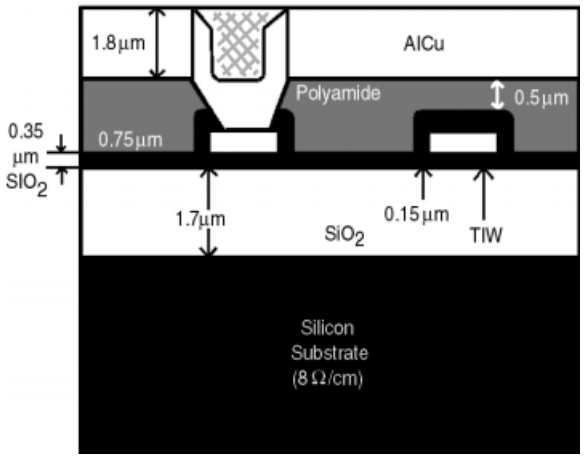


Figure 3-162 Cross section of the silicon inductor.

One of the reasons why we use a transistor multiplier is that we expect gain (because of the square-law characteristic, an FET would have been better tailored for a doubler); therefore, we have "measured" the gain of this circuit while driving it with 0 dBm. Another important factor is that, because of the way the CAD tool displays it, the gain is referenced to the 500-MHz input. Mentioning this is important; otherwise, the reader may look for the 1-GHz gain, which does not exist. Figure 3-165 shows the doubler gain.

By now we are curious to see the doubler in action, and by reviewing Figure 3-166 we can see the output spectrum. The output spectrum validates the statement that the output circuit, while doing a marvelous job of removing the 500-MHz input frequency, does not cure the problem of unwanted multiples of the input frequency. Again, a Cauer low-pass filter, while introducing some possible additional loss, should be considered for this circuit to have a bandpass response. On the other hand, many oscillators have very little harmonic suppression, and the same type of postfiltering needs to be considered. There are two-transistor solutions available also, such as push–push for frequency doubling (push–pull would

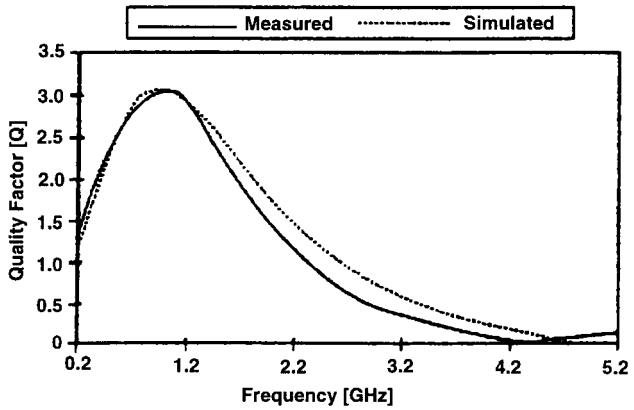


Figure 3-163 Measured and simulated *Q* versus frequency for the simulated inductor.

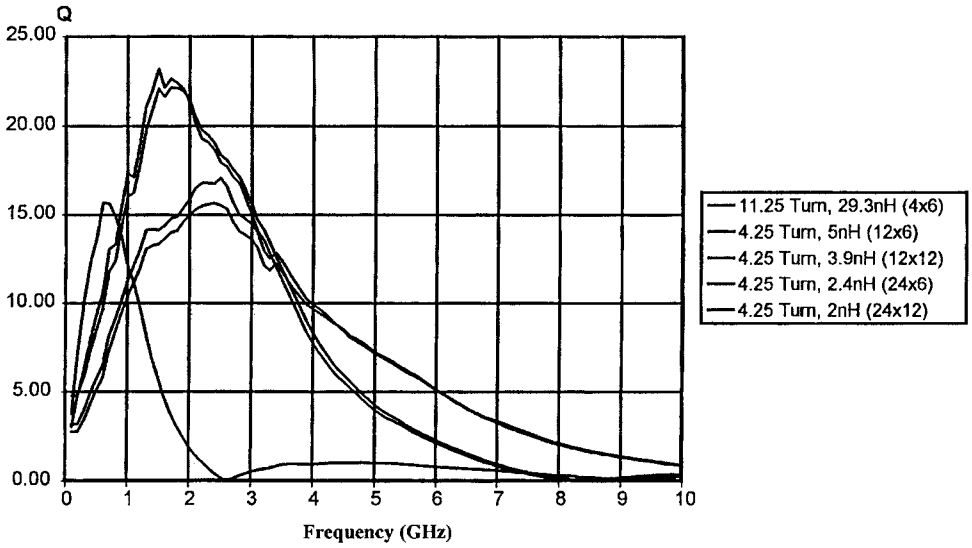


Figure 3-164 *Q* versus frequency for Motorola improved inductor.

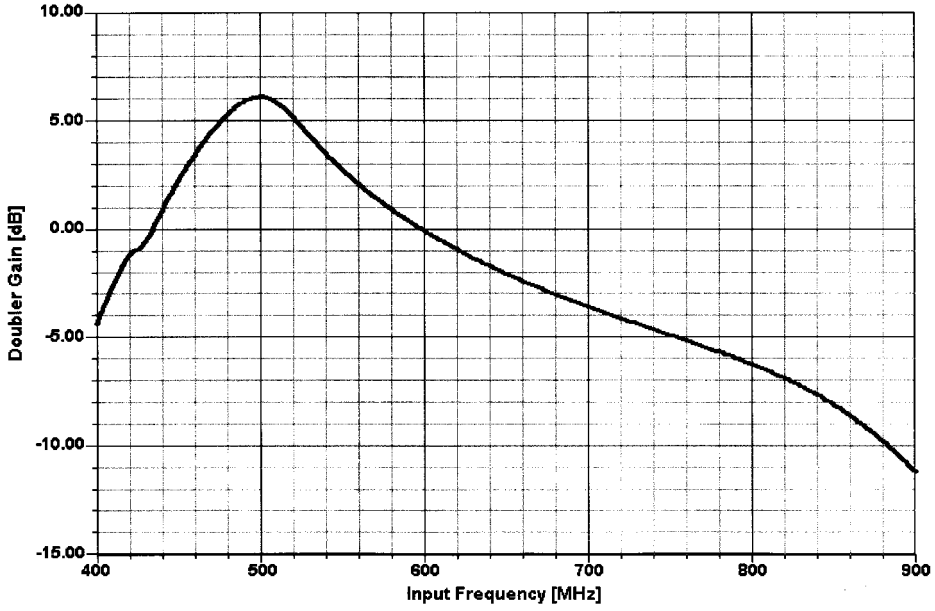


Figure 3-165 Frequency-dependent gain of the BJT doubler.

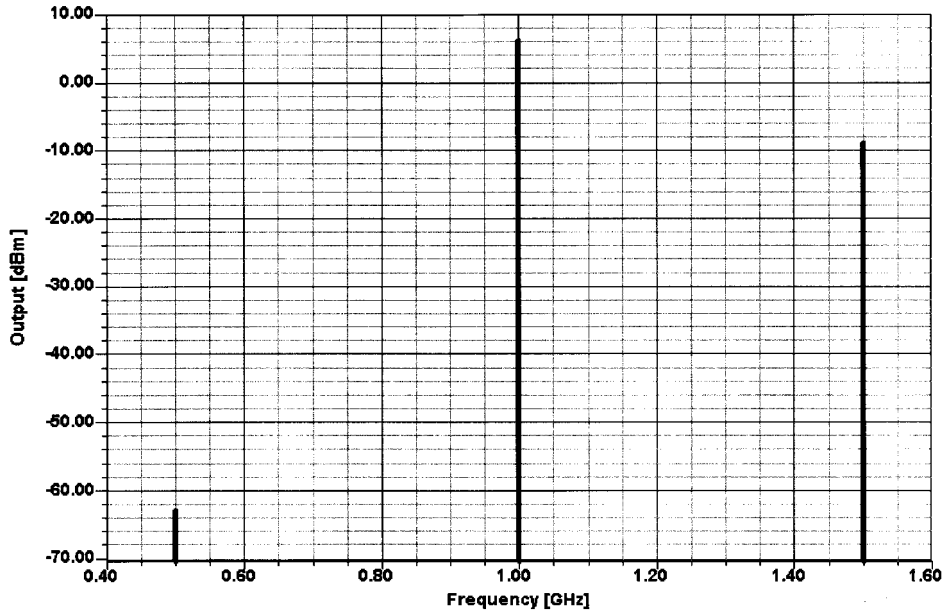


Figure 3-166 Output spectrum of the BJT doubler.

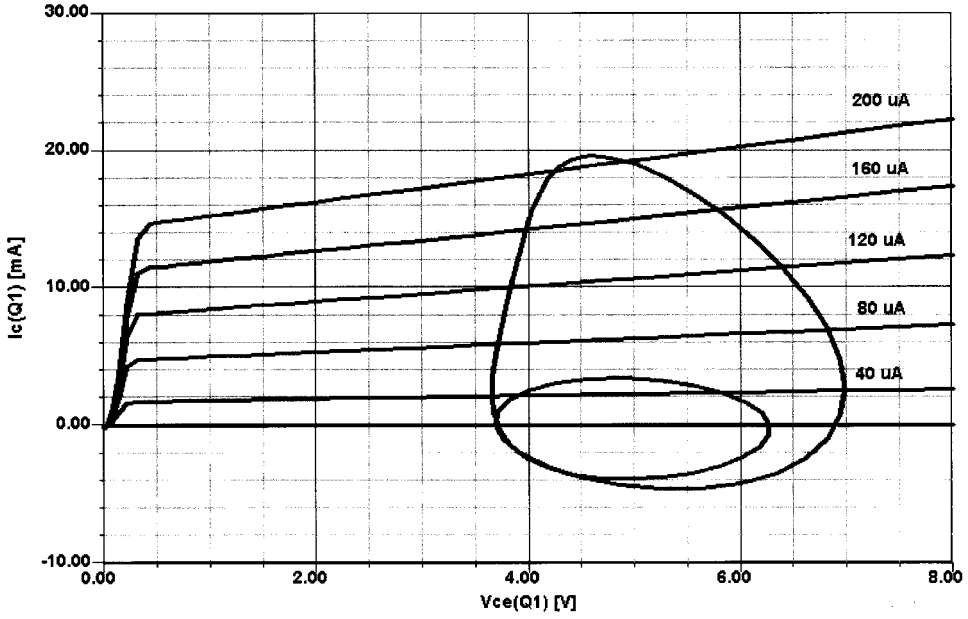


Figure 3-167 The dc *I–V* curves and ac load line for RF transistor Q1 in the BJT frequency doubler.

eliminate the desired second harmonic quite well). Actually, two diodes could also have been used for doubling, but since this requires transformers and a postamplifier, we have shown the preferred application. The only advantage of the diode application is that it is wideband, while this BJT circuit, as we have just learned, must be made narrowband for reasons of spectral purity.

Finally, we are curious to see the actual operating state of this amplifier, and this can be done by analyzing the load line (Figure 3-167). This is as complicated as the analysis for the differential amplifier. Because of the stored energy in the various tuned circuits, the load line looks quite impressive and would probably be difficult to predict without a nonlinear CAD tool.

3-9 MULTISTAGE AMPLIFIERS WITH AUTOMATIC GAIN CONTROL (AGC)

Having discussed the basic differential amplifier, we will now evaluate the performance of the MC1350/1490 video amplifier chip from Motorola. Most of the circuitry in the chip is self-explanatory. Figure 3-168 shows its inner circuit, and Figure 3-169 show a recommended application. By grounding one side of the input, we effectively get the same performance as the differential amplifier just discussed, with one difference being that the input is now fed to the constant-current generator. To minimize the distortion at the two lower transistors, they are decoupled by a 66-Ω resistor. Figure 3-170 shows the frequency-dependent gain, matching, and noise figure. Since this particular device has been specifically designed for AGC purposes, we also show in Figure 3-171 its gain reduction versus applied AGC voltage.

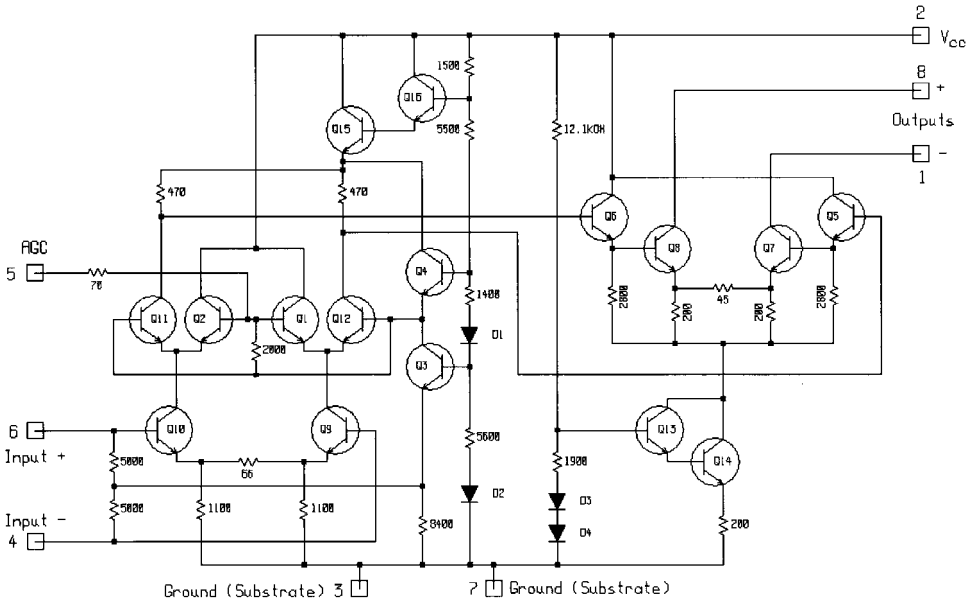


Figure 3-168 Schematic of the Motorola MC1350/1490 entered for simulation.

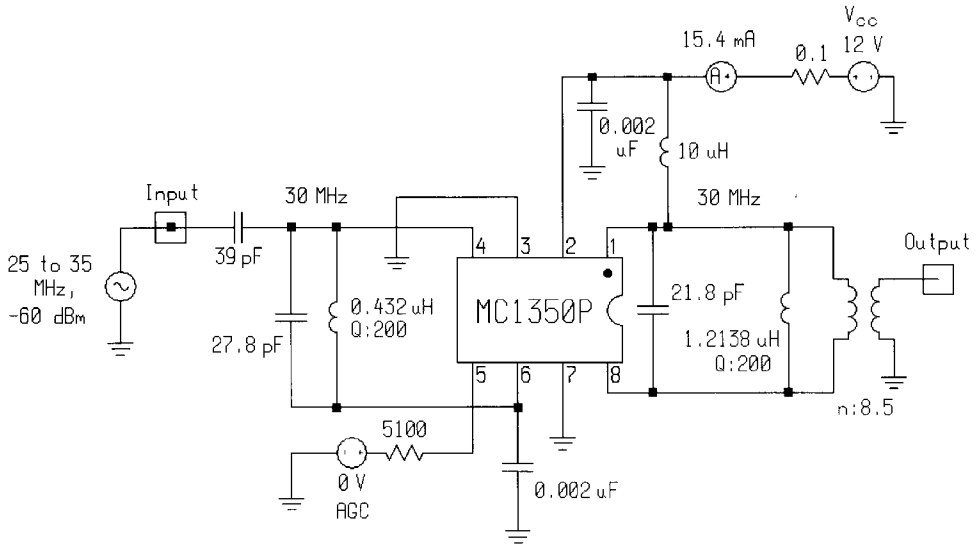


Figure 3-169 MC1350/1490 30-MHz application circuit.

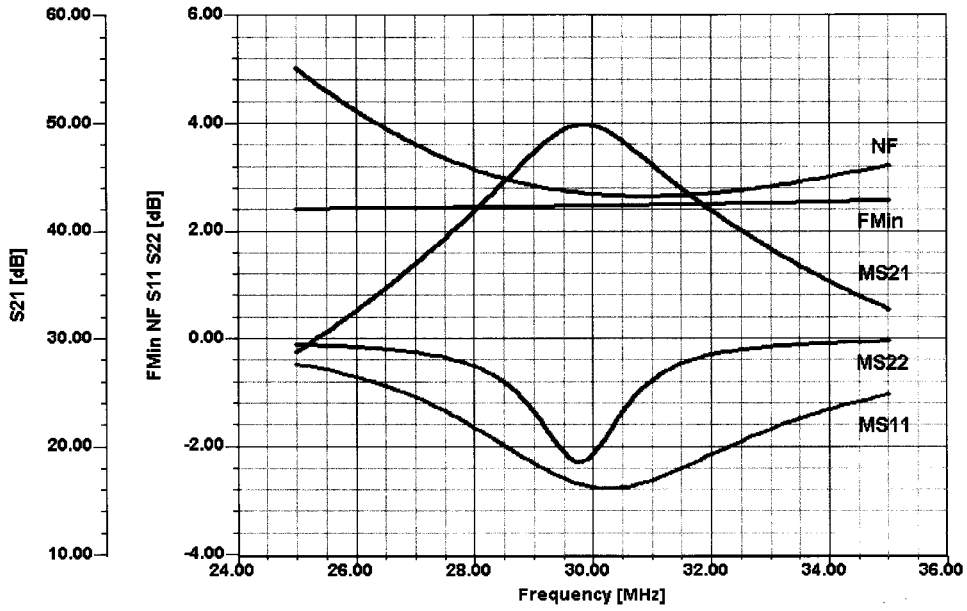


Figure 3-170 Frequency-dependent gain, matching, and noise performance of the MC1350/1490 in the circuit of Figure 3-169.

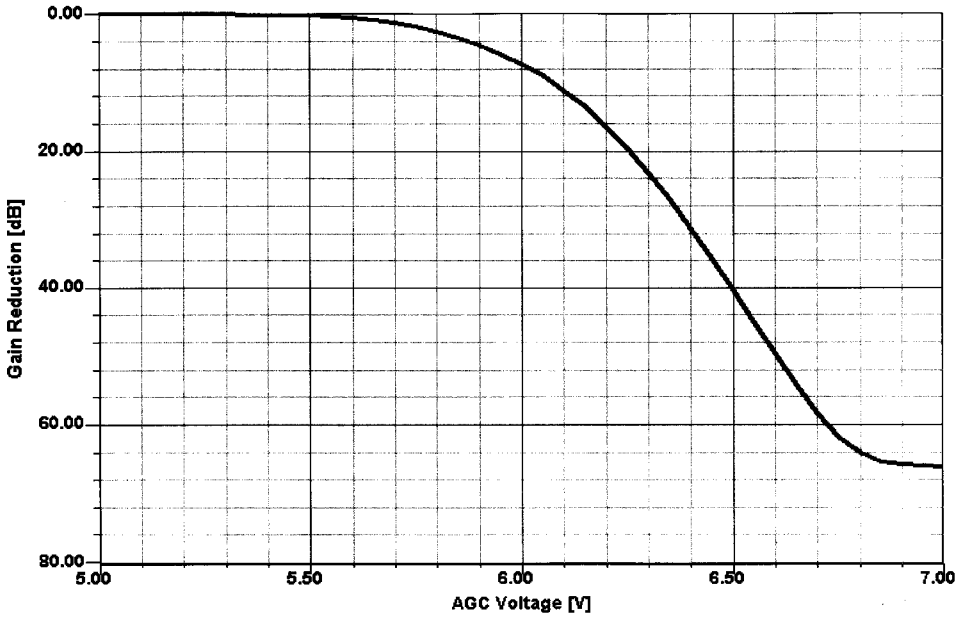


Figure 3-171 Gain reduction versus applied AGC voltage for the MC1350/1490 IC.

3-10 BIASING

A lot has been written about biasing, and since the purpose of this book is not to focus too much on dc, the following examples including RF consequences have been selected and considered sufficient. Early application reports by Motorola and others have provided a wealth of data. As a side effect of the introduction of the wireless era, we find that the engineers are so pressed for time that the number of complete application reports has approached zero. This is one of the reasons why we have tried to minimize the mathematical basis of this book (some is still needed), while still providing practical assistance. Figure 3-172 shows a textbook type of approach for RF.

In order to decouple the transistor from the biasing, we use RF chokes and dc decoupling (bypass) capacitors. This is the technique used in previous examples, so one might ask, "What's new?" As frequency increases, these inductors are either not manufacturable or have such a low Q that their use becomes questionable. This is the point where one may introduce the so-called distributed elements.

Figure 3-173 shows the very same circuit but resorting to distributed rather than lumped elements. The newly introduced elements (part of any good, up-to-date CAD tool; see element library for how to use their physical element descriptions) are described next.

- *Transmission Line.* Any printed connection between two points on a circuit board is a transmission line (Figure 3-174). Its characteristic electrical impedance is a function of the square root of the dielectric constant (ε_r), the width, metallization, thickness and height above substrate of the line, and the loss tangent of the substrate. Since lines frequently have to be laid out in the form of curved connections or have a bend in their direction, we have to add elements capable of describing the high-frequency conse-

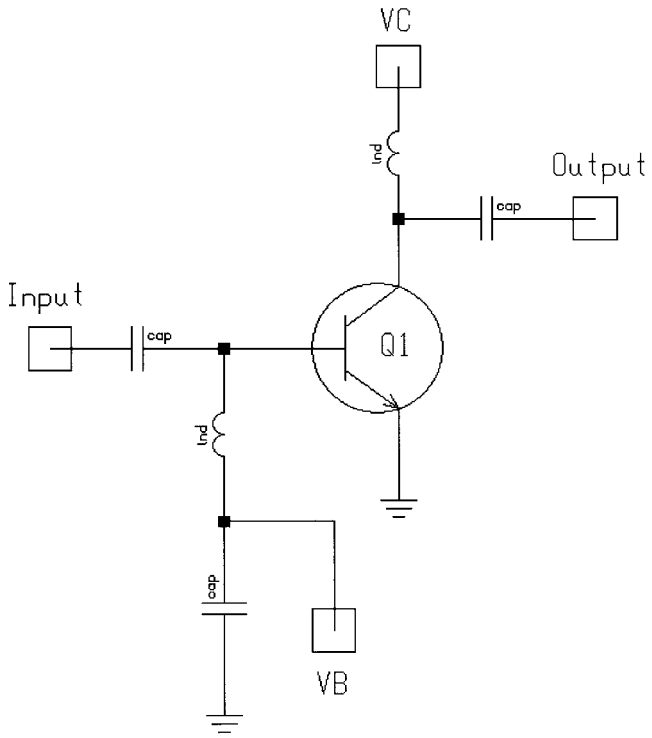


Figure 3-172 Simple BJT RF amplifier with lumped elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

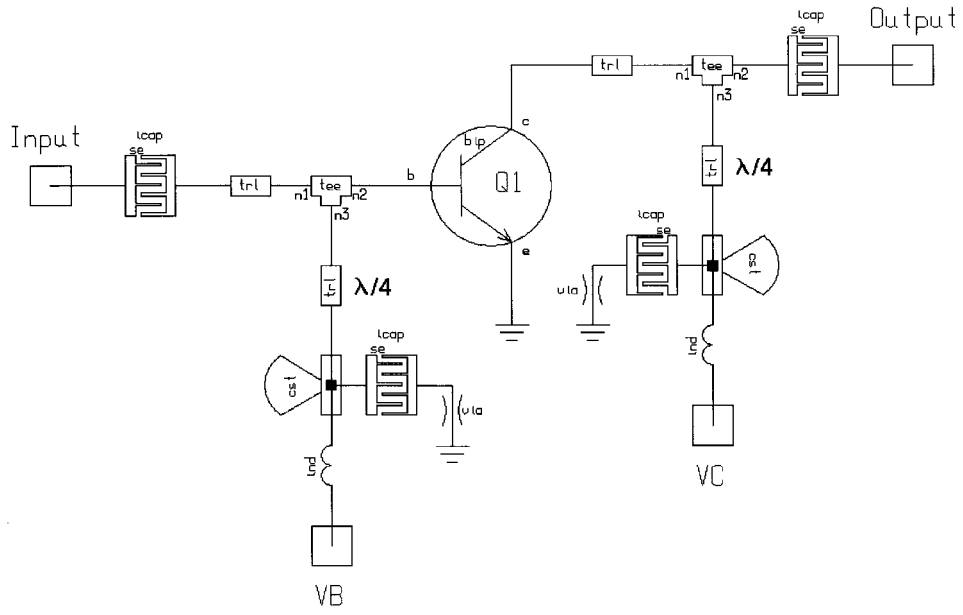


Figure 3-173 Simple BJT RF amplifier with distributed elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

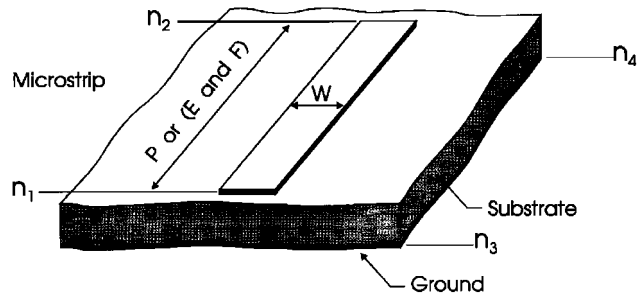


Figure 3-174 Transmission line in microstrip.

quences of such connections. Figures 3-175 and 3-176 show mitered and radial bend elements that perform this function.

- T *Junction, Cross, and* Y *Junction.* By the time a point like a collector or base, or its FET equivalent, spreads out into connections with other elements, we need additional modeling capability to describe ⊤ connections, crossings, and Y junctions. Figures 3-177, 3-178, and 3-179 show the way in which these connections must be modeled.

If the need exists, the standard inductances must be replaced with a transmission line whose length is λ/8 at the operating frequency. At higher frequencies, these transmission lines, however, then go into λ/4 resonant mode and later become capacitive. This type of design makes it fairly narrowband. A way around this is the use of printed inductors, as shown in Figures 3-180 and 3-181.

These inductors have a self-resonant frequency similar to the transmission line mentioned above, but the safety margin is significantly higher.

Talking about printed inductors, a logical extension of this is the printed transmission-line-based transformer, as shown in Figure 3-182. One can consider this as two interlaced rectangular inductors and, based on the substrate material, they are useful over a wide frequency range. Besides being used as a transformer, they can also be used to transit from

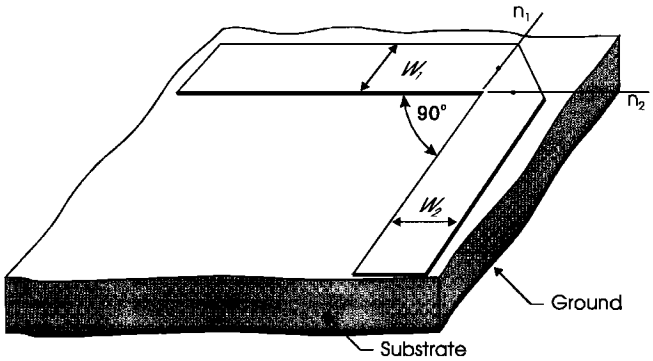


Figure 3-175 Mitered bend.

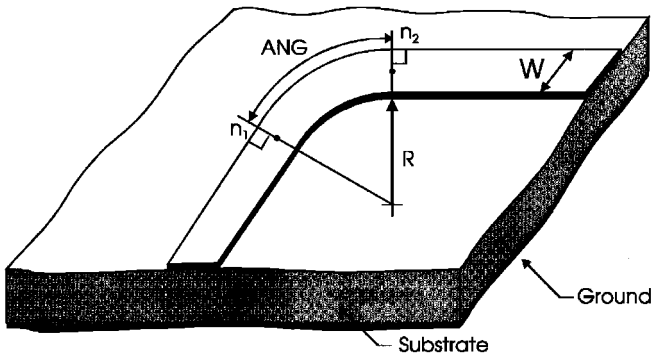


Figure 3-176 Radial bend.

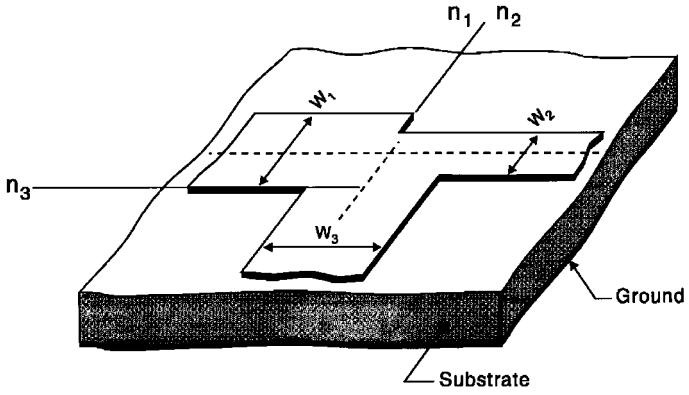


Figure 3-177 T junction.

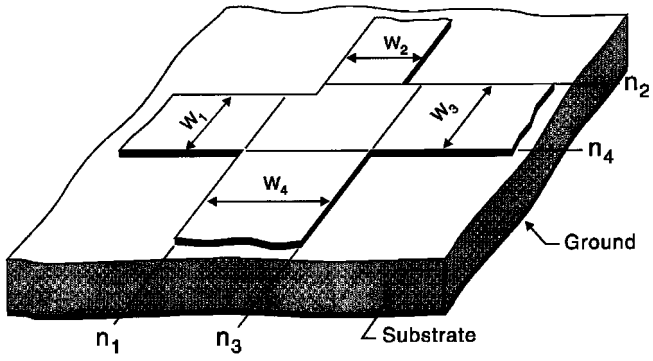


Figure 3-178 Cross.

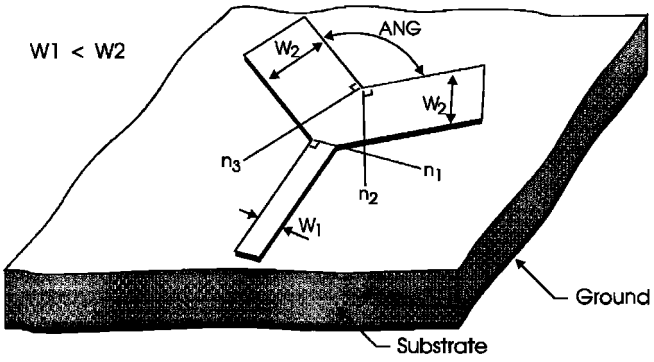


Figure 3-179 Y junction.

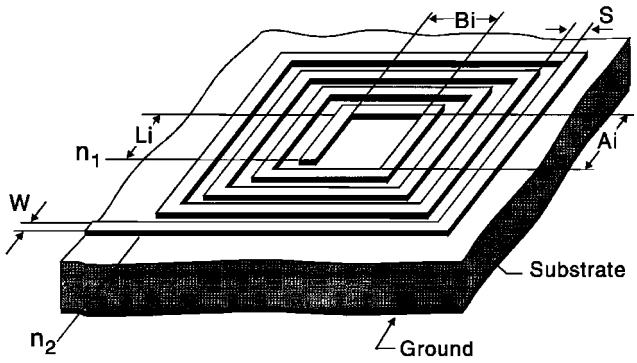


Figure 3-180 Rectangular inductor.

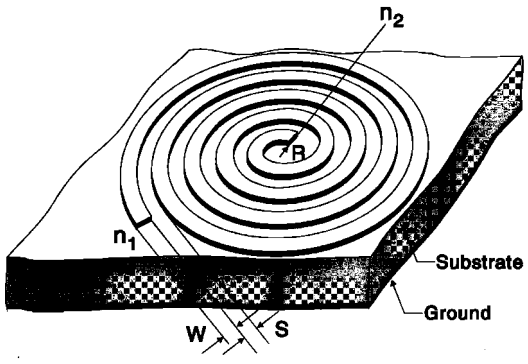


Figure 3-181 Spiral inductor.

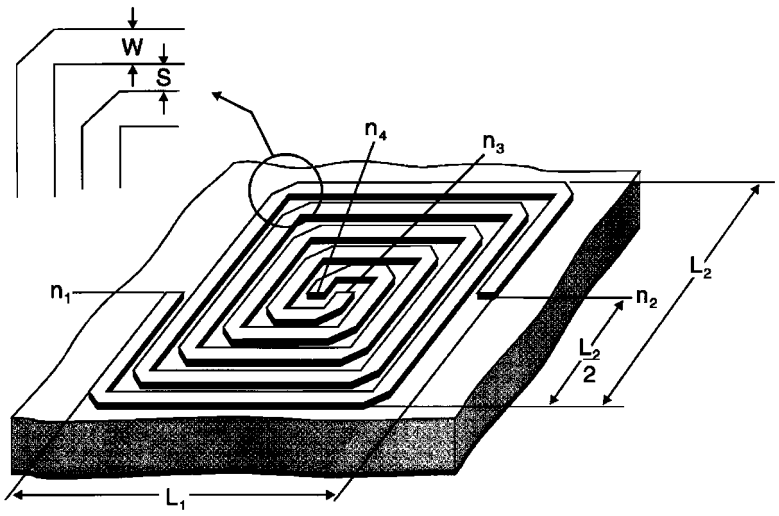


Figure 3-182 Transformer in microstrip.

unbalanced to balanced transmission provided that the difference in length from a connection point of view does not cause any problems (this is a layout issue).

A popular form of combining stages is the so-called Lange coupler (Figure 3-183) invented by the German Julius Lange. It is one of the major contributions in wideband applications.

Useful application of the Lange coupler probably starts at 4 GHz. It consists of parallel transmission lines with the appropriate connections as shown. Lange couplers are typically built with four, six, and eight fingers. The Ansoft Serenade product has a Lange coupler synthesis program that can be used to gain more insight into this coupler's application. We assume that other modern software has similar capabilities.

Where meander-type inductors are necessary, a neat way to implement and simulate them is to use the multiple coupled lines element (Figure 3-184) of the Serenade product, which both quickly and accurately calculates the behavior of the meander, including self-resonances and losses. We made use of this arrangement in our previous examples.

- *Interdigital Capacitors.* The issue of tolerances of small capacitors already has been brought up. The interdigital capacitor can be made on printed circuit board material as

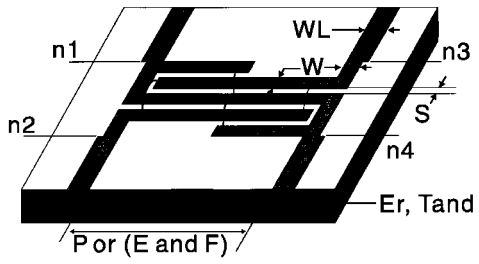


Figure 3-183 Four-strip Lange coupler.

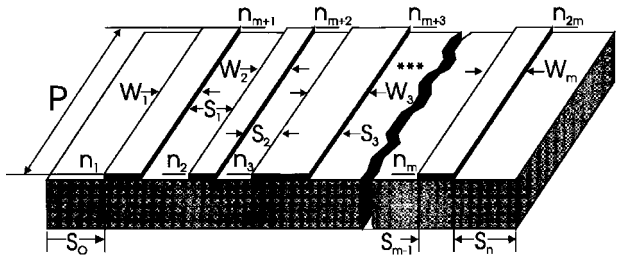


Figure 3-184 Multiple coupled lines element in microstrip.

well as gallium arsenide, and if its dimensions are continuous with the transmission-line width it does not cause any abrupt changes in impedance. This type of capacitor permits one to obtain very small values. An alternative to this is the use of transmission lines with a length of $\frac{3}{8}\lambda$. We learned earlier that a transmission line is inductive at frequencies below its fundamental resonance, is resistive at below resonance, and becomes capacitive above resonance. Again, bandwidth is also an issue. An interdigital capacitor consists of a number of parallel fingers as shown in Figure 3-185, and its capacitance can be varied by adjusting the number of fingers and their spacing. The advantage of the interdigital capacitor compared to discrete components is its low variation in value.

- *Radial Stubs.* The radial stub (Figure 3-186) is not much different from a $\lambda/4$ resonator, but its bandwidth is much greater than a simple transmission line. This is another way to ground the "cold" side of a transmission line or part of a circuit that needs to be grounded for RF. Of course, the interdigital capacitor comes in a version that is a combination of a capacitor and a via hole.
- *Via Holes.* The "cold" end of the transmission line, being considered either an inductor or a capacitor, needs to be connected to a sufficiently large copper backplane. One very efficient way to do this, especially if there is not enough copper left on the top of the board, is the use of via holes (Figure 3-187). One could theoretically generate

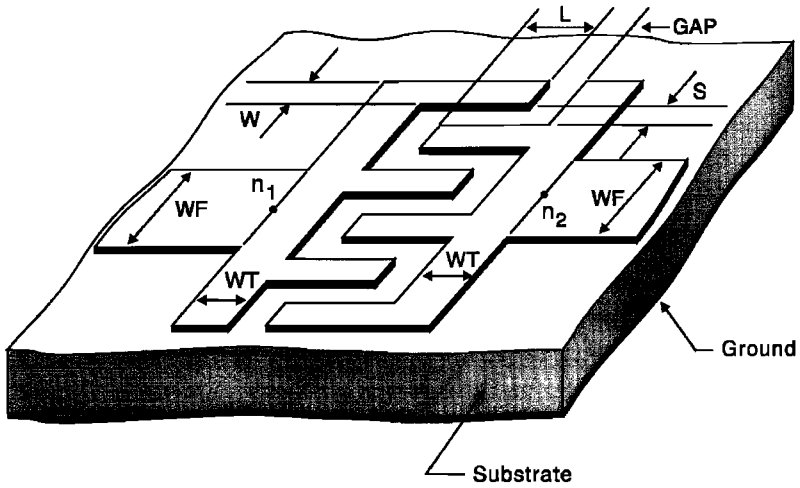


Figure 3-185 Interdigital capacitor.

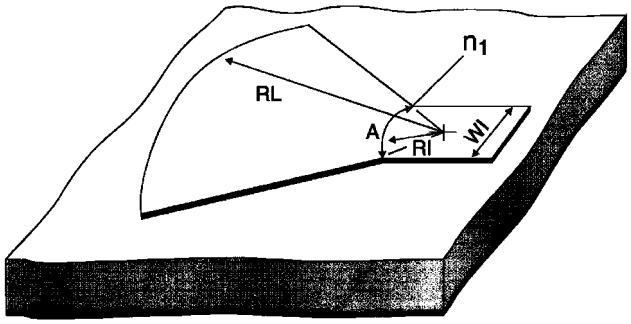


Figure 3-186 Radial stub.

a via hole with a rivet, but most manufacturing processes don't allow this; the normal solution is to use plated-through holes left open. In PC boards, via holes are typically cylindrical; on substrates like GaAs, they may be conical.

- *Correction Elements.* Although the behavior of an actual circuit proceeds regardless of our ability to measure and describe it, we do not enjoy this luxury in simulating circuit behavior in software. In a simulator, effects that are insufficiently described will be inaccurately simulated. For instance, segments of high-impedance transmission line (e.g., 120 Ω) are frequently used for dc feeds and RF chokes. In predicting the effect of a transition from 120-Ω line to 50-Ω line, a simulator must be alerted to the discontinuity so it can do the necessary mathematical corrections to account for the impedance jump. To do this, a specific circuit element, the STEP (Figure 3-188), must be inserted between the 120-Ω and 50-Ω line elements in the simulation circuit file. In addition to substrate data, we characterize a STEP by providing the widths of its input and output lines. The element itself has no physical length.

A similar correction is necessary if a transmission line is used as a resonator or just "left open" at one end. Such a transmission line tends to radiate and, because of its high-impedance properties, it reacts differently as far as its electrical length is concerned. A zero-length one-port element, the OPEN (Figure 3-189), must be added to such a line for mathematically correct calculation.

We end this excursion into distributed elements here. Interested readers may want to obtain a CD containing the element library of their favorite CAD tools that combine capabilities in the microwave area with a full set of distributed circuit elements. It is most important to keep

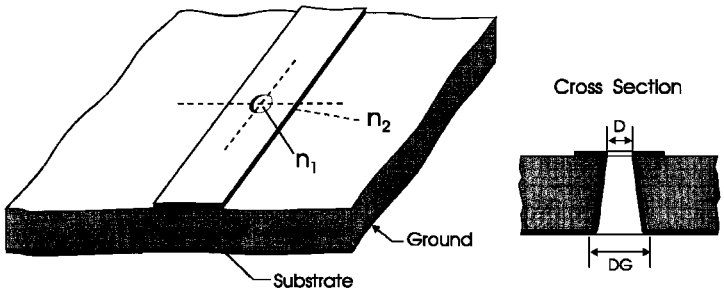


Figure 3-187 Via hole.

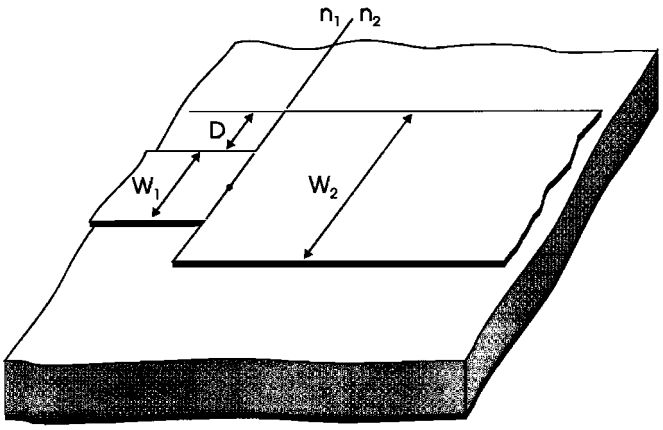


Figure 3-188 The STEP element tells the simulator to calculate the effects of joining transmission lines of differing characteristics.

in mind that, as frequency increases, we rapidly move into the area where we must consider all these distributed elements to achieve accurate simulations—even if doing so makes simulation a painful and time-consuming effort.

Finally, anyone who adventures into this area *must* obtain a foundry manual from the company that will build the integrated circuit or hybrid under design. There are basically two foundries, one applicable for MOS technology and one for GaAs. In the case of the bipolar transistor, the foundry service is not yet well-established. While we were able to interest one German company and one U.S. company in generating a custom bipolar IC within the activities of this book, the only "open" foundry we know of for bipolar (actually HBT) technology is the one operated by TRW. As of 1999, this seems to be the leading company for BJT-related products, and many currently available ICs have been developed with this foundry service.

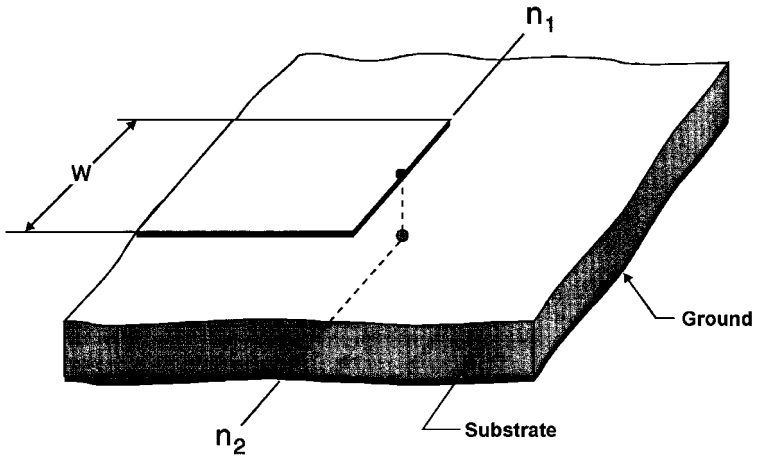


Figure 3-189 The OPEN element tells the simulator to calculate the effects of leaving the end of a transmission line unconnected.

We mention the issue of foundries here again because each foundry has its own proprietary approach to modeling discontinuities. The availability of a foundry service somewhat eases the requirement that a designer be fully up to speed on the nuances of discontinuities, because a foundry's designer service will help customers account for all relevant parasitics or discontinuities in their designs. In addition, there are tables of *S* parameters for standard cells of either capacitors, resistors, or inductors. The designer may then be forced to adjust the circuit so that it will work with a particular inductance value or value of another component within the resolution of the table that describes these elements. Information on the active parts, such as diodes and transistors, was given in Chapter 2.

3-10-1 RF Biasing

Applying the knowledge we have just acquired about distributed elements, Figure 3-190 shows a somewhat exaggerated case of using the various elements. While it is consistent with the simple RF case with which we started, modeling of this type is certainly necessary at higher frequencies. Many but not all of the elements we have just described are used, but we have not solved the question of the actual dc biasing.

3-10-2 dc Biasing

The following discussion is the recommended procedure for biasing a transistor. We will start with the collector current (in our example, Figure 3-191, 10 mA). To separate RF and dc, we have a combination of a load resistor (R_L) and an inductance (L_C). Information about this operation was given already in Section 3-2-6. The emitter resistor should have a voltage

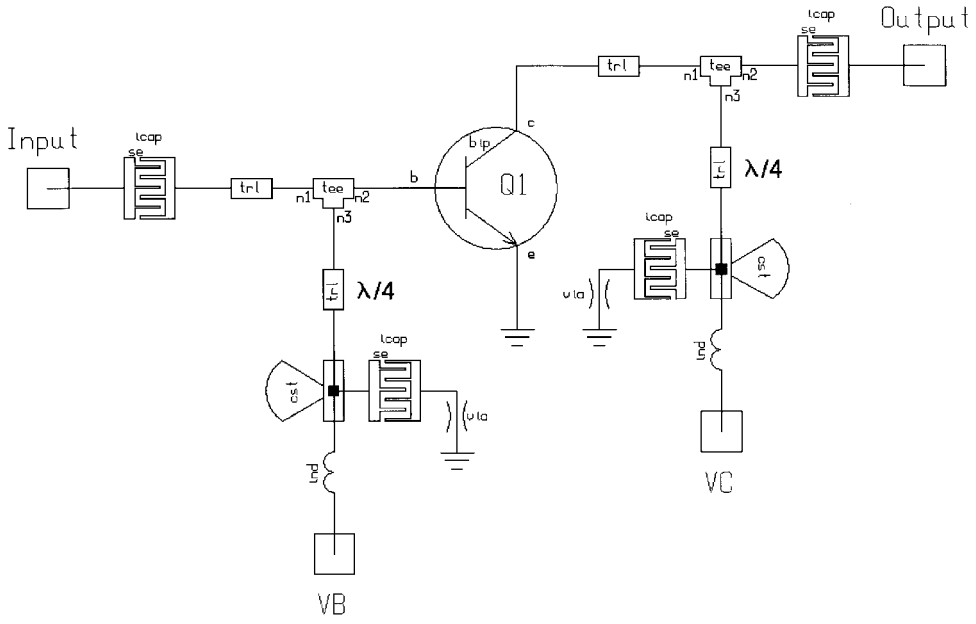


Figure 3-190 Simple BJT RF amplifier with distributed elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

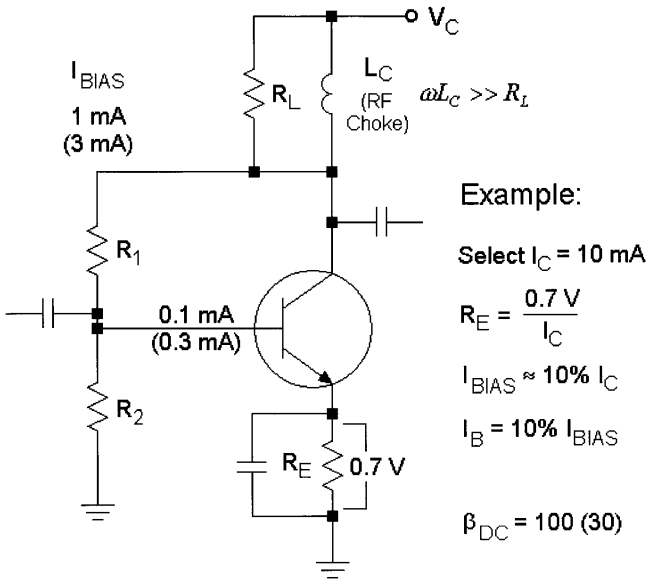


Figure 3-191 A dc biasing example.

drop of about 0.7 V, and therefore the emitter resistor is $0.7/I_C$. The next assumption is that the transistor has a dc current gain of 100, and therefore we decide to have the bias resistor chain (R_1 and R_2) draw 10% of I_C, or 1 mA. Assuming a collector voltage of 5 V, the value of $R_1 + R_2$ has to be 5 kΩ to result in the required 1 mA. Because the base current now is 10% of the divider current, the base will draw 0.1 mA. The voltage across R_2 has to be the voltage drop across R_E (0.7 V) + V_{BE}, which is also 0.7 V, or 1.4 V. Given the 1 mA, R_2 = 1.4/1E–3 = 1.4 kΩ. R_1 therefore is 5 kΩ – 1.4 kΩ = 3.6 kΩ. This concludes the calculation.

In the case of a dc beta of only 30, we need to raise I_{BIAS} from 1 mA to 3 mA since the base current now will be 0.3 mA. This changes the resistor values. $R_1 + R_2$ = 5 V/3 mA = (rounded) 1.6 kΩ. Following the same approach, we need 1.4/3E–3 = (rounded) 470 Ω. This makes R_1 = 1600 Ω – 470 Ω = (rounded) 1100 Ω. These roundings were necessary because resistors can only be bought in certain resistance steps.

However, for low battery voltages, the 0.7 V between the emitter and ground cannot be wasted. Therefore, a dc stabilization circuit, to which we have already referred, will be considered next. Figure 3-192 shows a combination of an RF transistor, Q1, and a dc stabilization transistor, Q2. The most important thing about this circuit is the fact that the difference between the supply voltage and the collector of Q1 is 0.3 V, which is approximately half of the voltage we had to "waste" across R_E in the above-mentioned circuit. Plus, the voltage drop is now in the collector, and given the fact that we feed Q1's collector through the inductor of a quasi-tuned circuit rather than through only a resistor, we get a much higher voltage swing. The base–collector diode temperature dependency of Q2 is compensated for with the silicon diode feeding the base. It is also fascinating to see that, because of the low currents, both the voltage drop of the diode and V_{BE} of Q2 are about the same, and yet Q1, operating at a reasonable current, shows the expected 0.7–0.8 V V_{BE}. We have supplied all the necessary dc voltages to give the reader an incentive to calculate the mechanism of this stabilization. The key equation to this is to determine the current in the voltage divider feeding

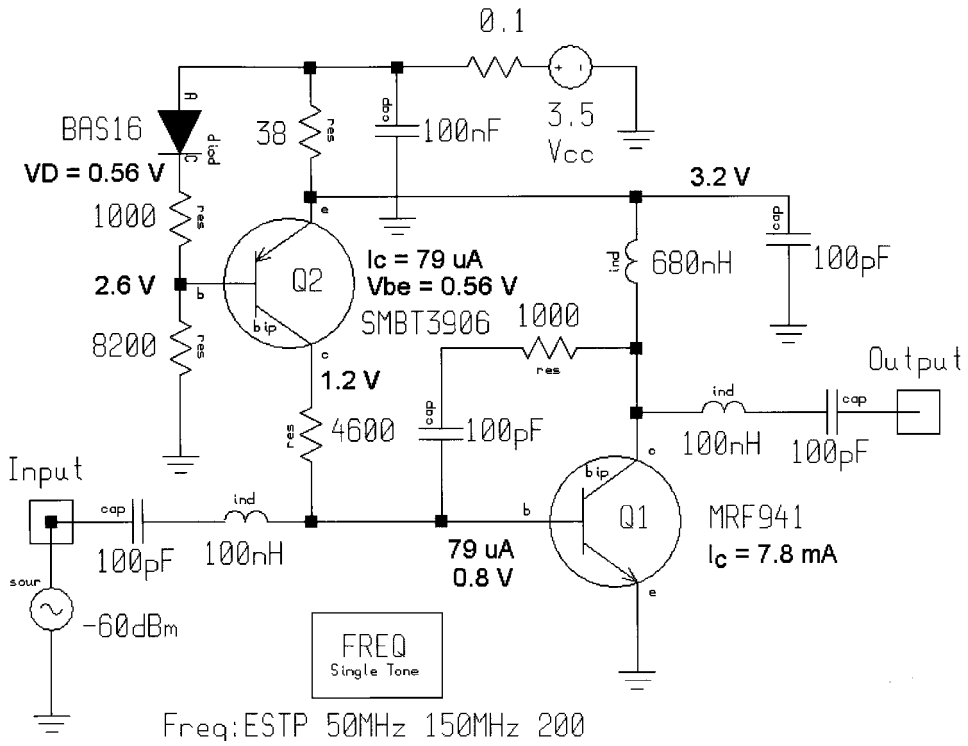


Figure 3-192 An RF amplifier with active biasing. (After the low-noise amplifier circuit in [29].)

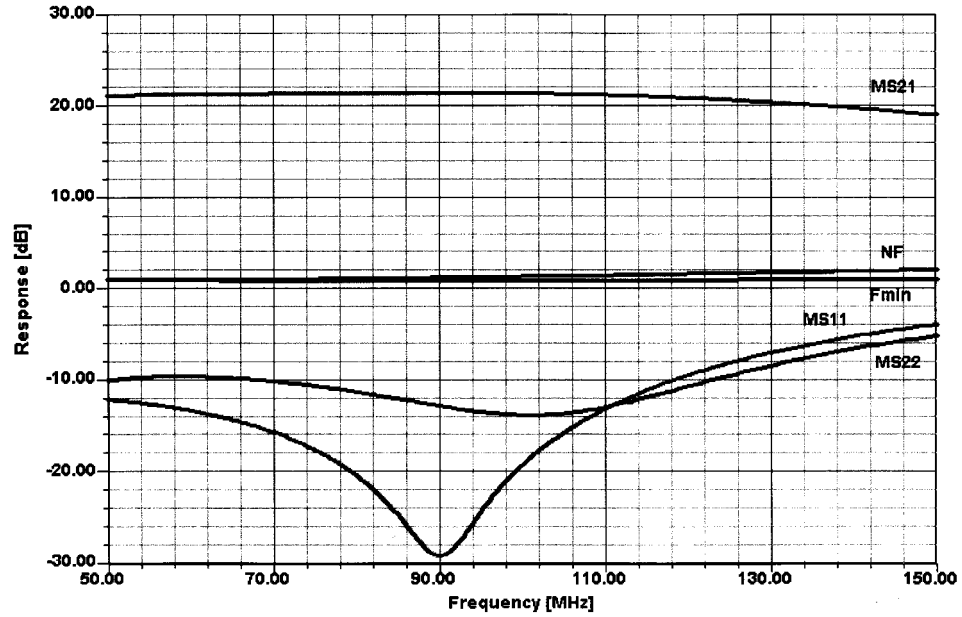


Figure 3-193 Frequency-dependent gain, matching, and noise responses for the amplifier in Figure 3-192.

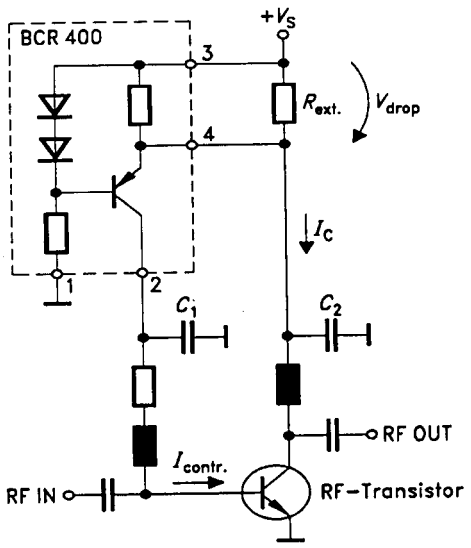


Figure 3-194 The Siemens BCR400 active bias controller can be applied to FETs as well as BJTs.

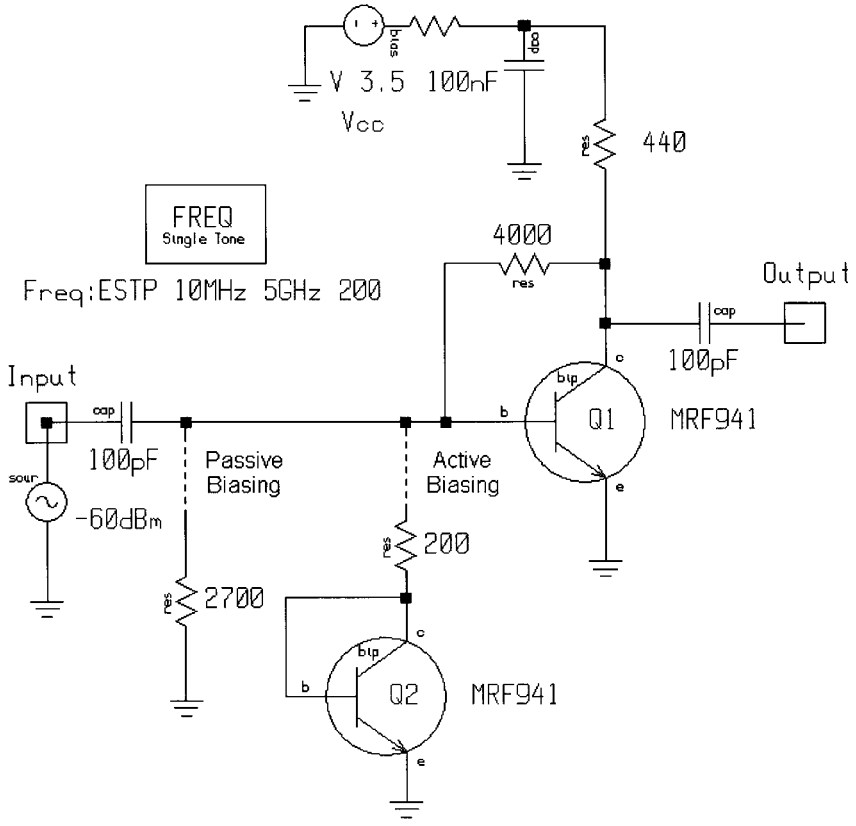


Figure 3-195 Amplifier stage (I_C = 3 mA) showing active and passive bias alternatives.

Q2, and proceeding from there. On the other hand, having a nice CAD tool that provides insight into the dc voltages and allows the addition of dc voltage probes makes life much easier. To prove that this circuit actually works, Figure 3-193 shows its frequency-dependent gain, matching, and noise figure performance. It is based on a Motorola application, which had excessive voltage drop for Q1 and which we have improved to be suitable for low-voltage operation [29].

Integrated active bias solutions are also available. Figure 3-194 shows the Siemens BCR400 active bias controller applied to a BJT; it can also be applied to FETs and TR switching diodes, as we saw in Figures 3-57 and 3-59, respectively.

3-10-3 dc Biasing of IC-Type Amplifiers

In integrated circuits, we do not have the unlimited flexibility described earlier. This results frequently in constant-current sources or transistors being used as diodes for biasing purposes. Our first suspicion is that this is going to cause a lot of noise in the IC; Figure 3-195 shows a test example.

As suspected, choosing the same bias point of 3 mA, the noise figure using active bias is quite a bit higher than the passive one, as shown in Figure 3-196.

3-11 PUSH–PULL/PARALLEL AMPLIFIERS

Since it is practically impossible to obtain $S_{11} = S_{22} = 0$, using a combiner at the input and output reduces the problem. Figure 3-197 shows the 3-dB hybrid approach. It has a practical

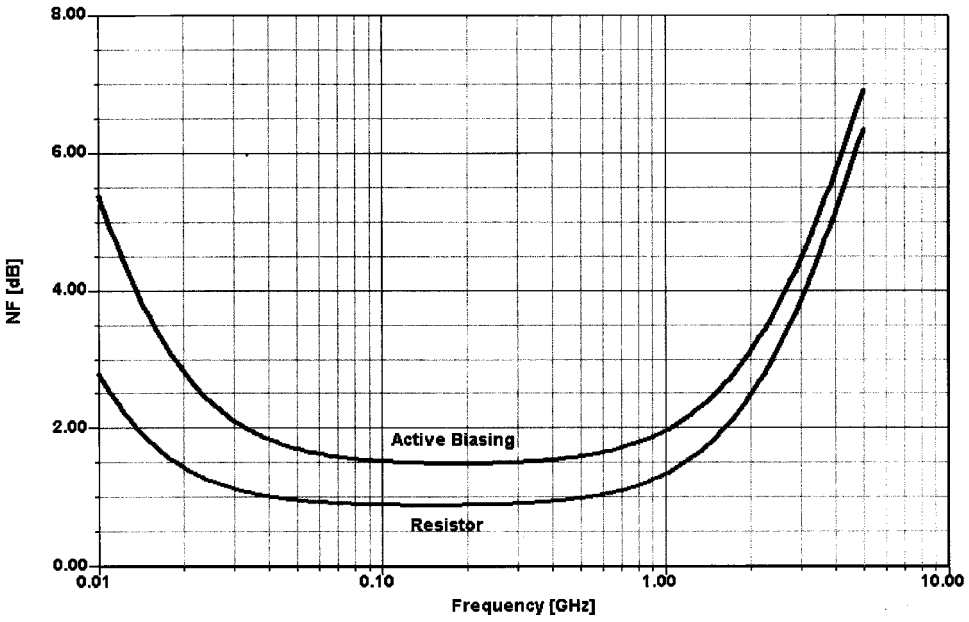


Figure 3-196 Frequency-dependent noise figure of the amplifier circuit shown in Figure 3-195. The active biasing curve is considerably noisier.

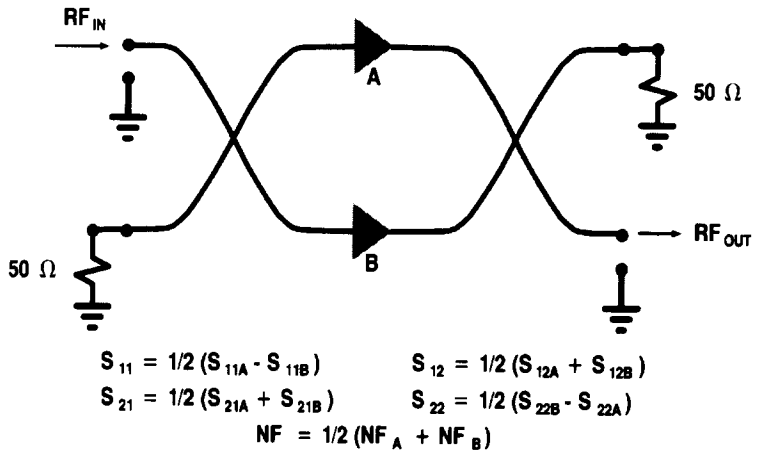


Figure 3-197 Combining power with 3-dB hybrid couplers.

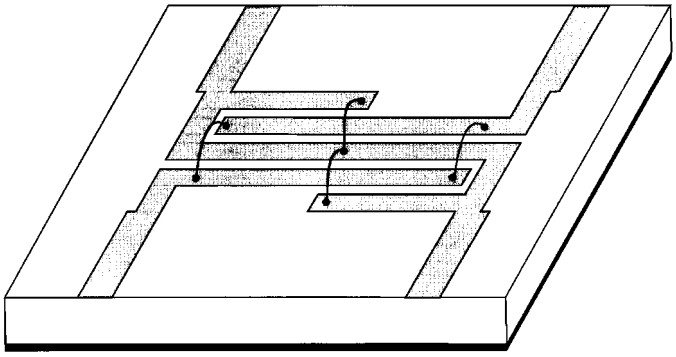


Figure 3-198 Four-strip version of the Lange coupler, a broadband 90° hybrid.

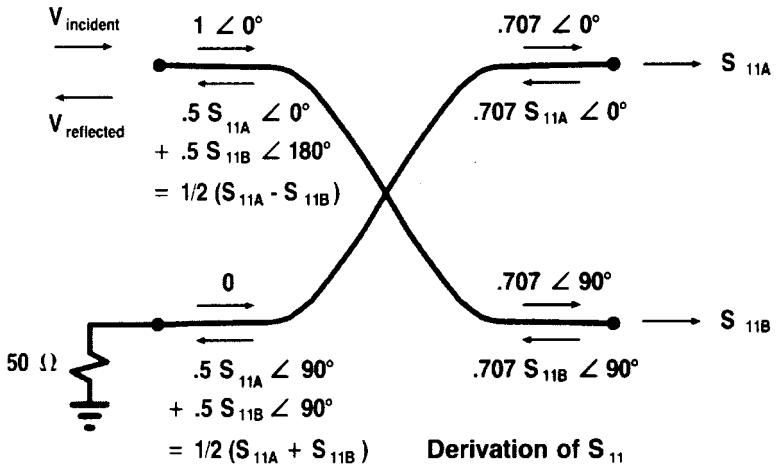


Figure 3-199 3-dB hybrids. This analysis method may be applied to any coupler.

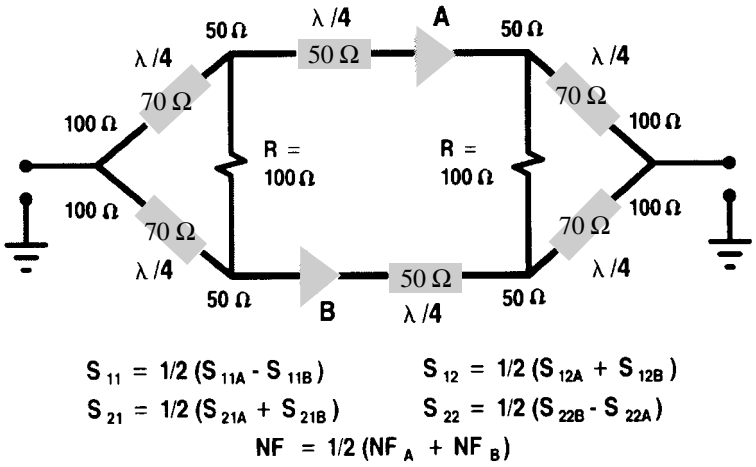


Figure 3-200 Wilkinson divider/combiners.

bandwidth limit of about 4:1; its advantage is that it is compact and ideal for cascading. If one stage opens, P_{OUT} and gain drop only 6 dB; however, it does not protect against load mismatch. A recommended distributed form is the Lange coupler (Figure 3-198). To analyze this approach, Figure 3-199 shows the resulting impedances and necessary phase shifts. If a Wilkinson coupler [30] is used instead of the 3-dB hybrid, the bandwidth is reduced. The Wilkinson coupler is larger in size than the 3-dB hybrid but more easily realized, leading to lower cost. Figure 3-200 shows the actual impedances necessary for matching, with A1 and A2 being the actual amplifiers combined.

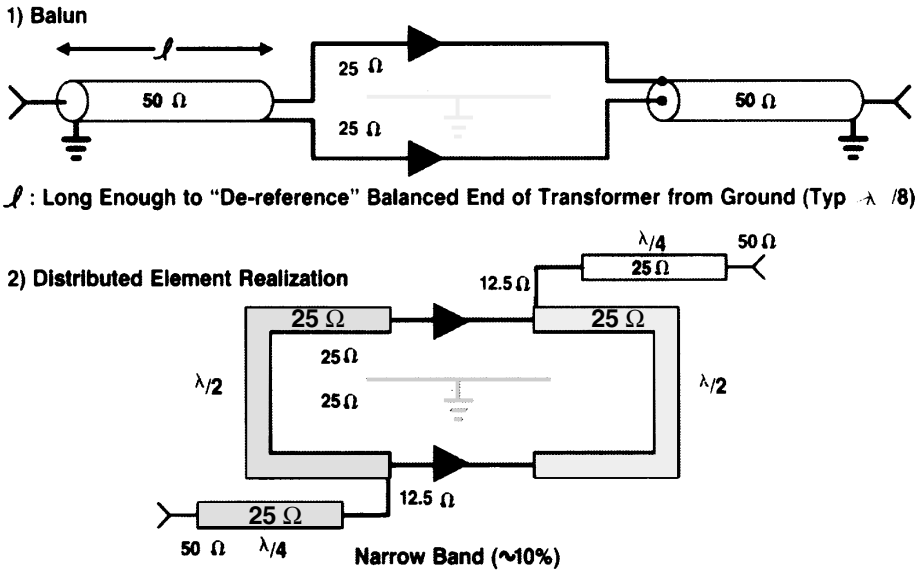


Figure 3-201 Push–pull amplifiers.

The solution for push–pull amplifiers resorts to a balun with the following results obtained: Bandwidths of greater than one decade are achievable; the balun structure gives 4:1 impedance advantage; there is no isolation; $S_{11} \neq 0$, $S_{22} \neq 0$; and cancellation of even-order products is provided.

Since most amplifiers have 50-Ω rather than 25-Ω impedances, Figure 3-201 shows a circuit that matches 25 Ω to 50 Ω on both sides. Since the standard amplifier operates at 50 Ω rather than 25 Ω, an additional network may be necessary to transform the impedance up to 50 Ω again. A similar problem occurs at the amplifier output.

3-12 POWER AMPLIFIERS

Power amplifiers are devices that are tasked to transform as much dc power into RF/microwave power as possible. Depending on the technology, devices are available that can produce at least 100 W up to 2 GHz. Since the transconductance of FETs at conducting angles of less than 180° becomes very small, the actual power gain sharply decreases compared to Class A operation while the power-added efficiency is optimized to values of 55% (Class B) or so. The dissipation of these devices is, of course, quite high, while in Class B operation this is even more of an issue. Power devices in Class C are dying to have big heat sinks to survive. One of the nice things about FETs is a much reduced thermal runaway effect compared to their bipolar brothers.

Power amplifier classes are defined in terms of the device conduction angle and/or the type of device operation:

Class A	"Linear" operation; 180° conduction angle
Class AB	"Linear" operation; conduction angle between 90° and 180°
Class B	"Linear" operation; 90° conduction angle
Class C	Fixed drive; less than 90° conduction angle
Class D	Switched operation, conduction range may vary with time from 0° to 180° or may be fixed.

There are also Classes E and F, which are currently under evaluation and showing some promising results.

Figure 3-202 shows Class C operation. The output of the amplifier is a current pulse train, which frequently leads to misunderstanding as to the calculation of its load. Given the fact that the amplifier is a current generator, it is silly to assume one has to do a conjugate matching looking backward into the amplifier. This is the condition for Class A operation. Unfortunately, this is still not understood by some engineers, but we hope we can clear this up today. The theoretical efficiency is 89.7% at a conducting angle of less than 180°; in practice, efficiencies more than 85% are not obtainable. (All of these efficiencies are those of the device itself; losses in matching networks are not taken into account in these efficiency values.) Considering optimal gain and optimal output power, the conducting angle actually settles somewhere between 100° and 140°. Again, the power gain in Class C operation is less than in Class A and Class B operation.

3-12-1 Example 1: 7-W Class C BJT Amplifier for 1.6 GHz

The following demonstrates how to design a 1.6-GHz, 7-W amplifier operating in Class C. We need to point out again that all modulation forms that contain the information in both

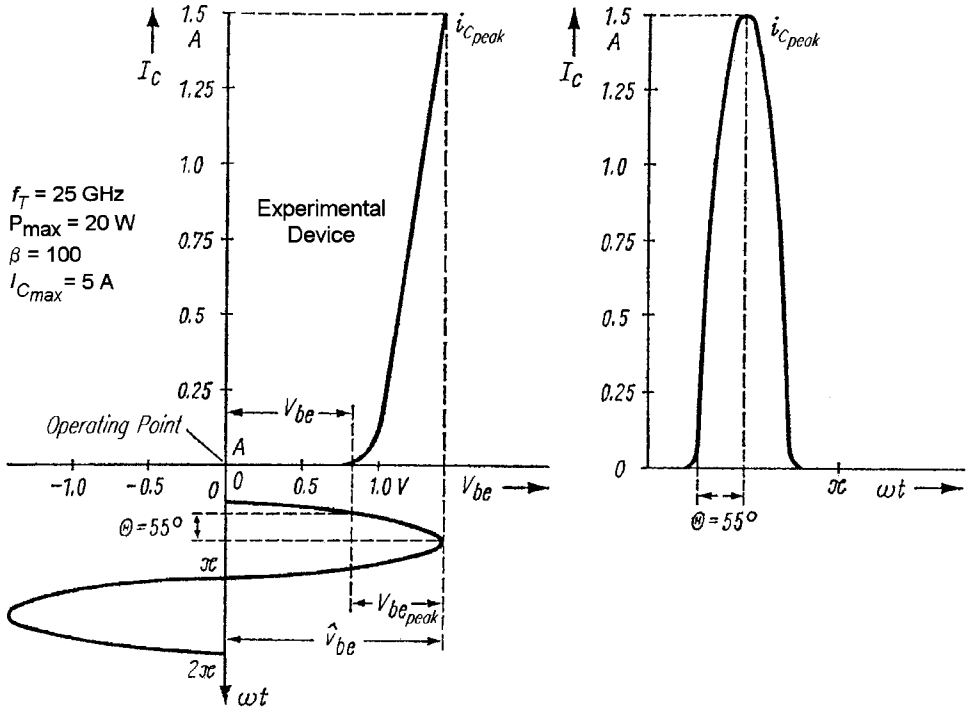


Figure 3-202 Output current as a function of drive voltage, taking into consideration the knee voltage of the base–emitter *pn* junction. The output is a series of current pulses with a duty cycle determined by the conduction angle.

amplitude and phase require more linearity than Class C affords. Needless to say, this affects the efficiency; therefore, at a later point we will also show a Class AB power amplifier. This is a tutorial, so we will present a UHF example and walk the reader through all the problems involved in the design. The selection of the transistor is done by first looking at the f_T, which in our case should be at least 4–5 GHz. The first order of business is to define the point of operation for the output amplifier. To do this, we need to determine the maximum voltage and current values supplied by the manufacturer. We have chosen a device that operates at 28 V and can sustain a peak collector current of 1.5 A. Figure 3-202 lets us determine the conducting angle Θ from the I_C/V_{BE} plane. This figure shows on the left side the I_C/V_{BE} plane and the necessary voltage to drive the amplifier.

Initially, we assume that the base–emitter junction is operated without any bias ($V_{BE} = 0$). The dc I–V curve shows that the collector current starts at about 650 mV. The amplifier itself, therefore, is already in Class C operation. For full output power, the drive voltage $V_{BE,\text{peak}}$ has to be 1.14 V. Therefore, the conducting angle can be determined by

$$\cos \Theta = \frac{V_{BE}}{\hat{\boldsymbol{v}}_{BE}} = \frac{0.65}{1.14} = 0.57$$

$$\cos 0.57 = 55^\circ$$

V_{BE} is the dc voltage at which the transistor starts conducting. From Figure 3-65 and Eqs. (3-160), (3-161), and (3-162), the fundamental current i_{c1} can be determined from

$$i_{c1} = i_{\text{peak}} \cdot f_1(\Theta)$$

or

$$i_{c1} = 1.5 \text{ A} \times 0.36 = 0.54 \text{ A}$$

The average collector dc current can be determined from

$$i_{c2} = i_{\text{peak}} \cdot f_2(\Theta)$$

or

$$i_{c2} = 1.5 \text{ A} \times 0.21 = 0.315 \text{ A}$$

The collector saturation voltage $V_{CE(\text{sat})}$, unfortunately, is around 2–3 V; for the purpose of the calculation, we assume that the saturation voltage is 3 V, and, therefore, we can assume that the maximum collector ac voltage $\hat{v}_{CE} \approx (V_B - 3) = 25$.

The appropriate collector load has to be

$$R_C = \frac{\hat{v}_{CE}}{i_{C1}} = \frac{25}{0.54} = 46.3 \ \Omega$$

The maximum output power is

$$P_{\text{out}} = \frac{(V_{CE} - V_{CE(\text{sat})})^2}{2R_C} = \frac{625}{92.6} = 6.75 \text{ W}$$

The only way to increase the power would be to have a smaller collector load at a higher current. In this family of Motorola transistors, there was no 10-W transistor; that's why we have chosen a 20-W experimental device, which for this purpose is overkill. The theoretical efficiency of the amplifier is defined as

$$\eta = \frac{P_{\text{out}}}{P_{\text{in}}} = \frac{6.75 \text{ W}}{0.315 \times 28} = 85\%$$

If the saturation voltage would be less, our efficiency would automatically increase, since for the given load, the actual output power would increase (6.75 W $\rightarrow$ 7.5 W). Assuming that this stage would operate at 1900 MHz with a gain of roughly 22 dB, the required drive would be significantly less than the output power, or roughly 40 mW. For the purpose of the following calculations, we will assume that the drive power is 0.7 W. The total dissipation of the device now adds up to be

$$P_T = P_{\text{in(dc)}} - P_{\text{out}} + P_{\text{drive}}$$

$$P_T = (28 \times 0.315) - 6.75 + 0.7 = 2.77$$

Power-added efficiency is derived from this. Knowing the power dissipation, one can calculate a thermal resistance R_{therm} of the heat sink and the transistor:

$$R_{\text{therm}} = \frac{T_J - T_A}{P_T}$$

Where T_J is the device junction temperature and T_A is the ambient temperature. For a maximum ambient temperature of 60 °C and a maximum junction temperature of 150 °C, the thermal resistance is

$$R_{\text{therm}} = \frac{120 - 60}{2.77} = 21.7\ ^\circ\text{C/W}$$

R_{therm} is the sum of thermal resistance of the transistor and its package, and the thermal resistance between the transistor package and the heat sink, and the thermal resistance between the heat sink and the ambient air. Therefore,

$$R_{\text{therm}} = R_{\text{th(case)}} + R_{\text{th(case/heat sink)}} + R_{\text{th(heat sink)}}$$

The required thermal resistance of the heat sink is calculated from

$$R_{\text{th(heat sink)}} = 21.7 - 10 - 1.4 = 15.6\ ^\circ\text{C/W}$$

Assuming an aluminum heat sink (a 2-mm-thick plate), the required heat-sink surface area A (cm^2) is

$$A = \frac{1}{\alpha R_{\text{th(heat sink)}}}; \quad \alpha \approx 1.5\ \frac{\text{mW}}{^\circ\text{Ccm}^2}$$

Therefore, the required area is

$$A = \frac{1}{1.5 \times 0.0156} = 42.7\ \text{cm}^2$$

The next task is to match both the input and output to the (most likely) 50-Ω source and termination. The following section will guide us through the procedure. The input impedance Z_{11} (after some lengthy search) was found to be $1 + j10\ \Omega$. This means that the input consists of the base-spreading resistance (approximately 1 Ω) and an input capacitance, which because of phase shift now turns out to be inductive. Since the popular method for input matching these days is a pi configuration, we are going to use a simple pi filter both for input and output matching. There are several equations in the literature that give the values for the two capacitors and inductors; most of them result in an incorrect resonance frequency. Since we have access to software tools, we decided to take three approximate values and use the optimizer to give us the appropriate matching. The only remaining variable in such an approach is the bandwidth. If we specify $S_{21} = 1$ and $S_{11} = 0$ (100% energy transfer and perfect return loss), we will realize that this is possible at only one frequency. We therefore need to identify a frequency band or bandwidth over which an acceptable match is needed. As the bandwidth increases, the match for such a simple circuit will get poorer; there are not enough elements available to do better. In this case, we took the center frequency ±100 MHz.

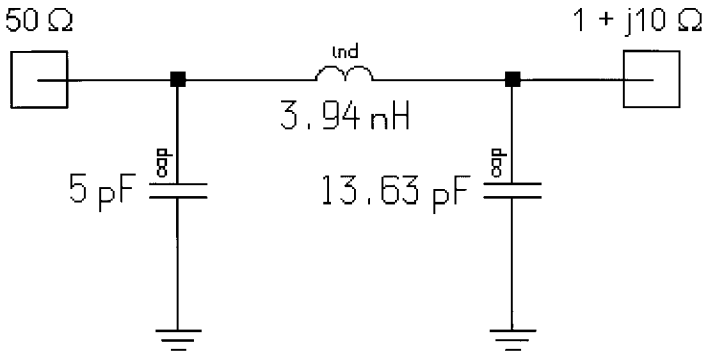


Figure 3-203 Input matching network for the amplifier used in this tutorial.

Figures 3-203 and 3-204 show the actual circuit with the element values and the resulting input match. Likewise, we did the same thing at the output, but the actual resulting dynamic output capacitance is not known *a priori*. For low frequencies, the integral equation gives a value of $2C_{CE}$; however, practical tests in frequencies above 500 MHz show that the actual increase is only 1.3 times the dc output capacitance. Therefore, we decided to design the filter with no output capacitance assumed, and the input capacitance of about 5 pF must be reduced for proper reactive matching. Mathematically, this could result in a negative number, specifically if the resulting output capacitance exceeds the 5-pF value. To compensate for this, one needs to use an RF choke that tunes out the resulting dynamic output capacitance and then one can revert to the original 5 pF as calculated.

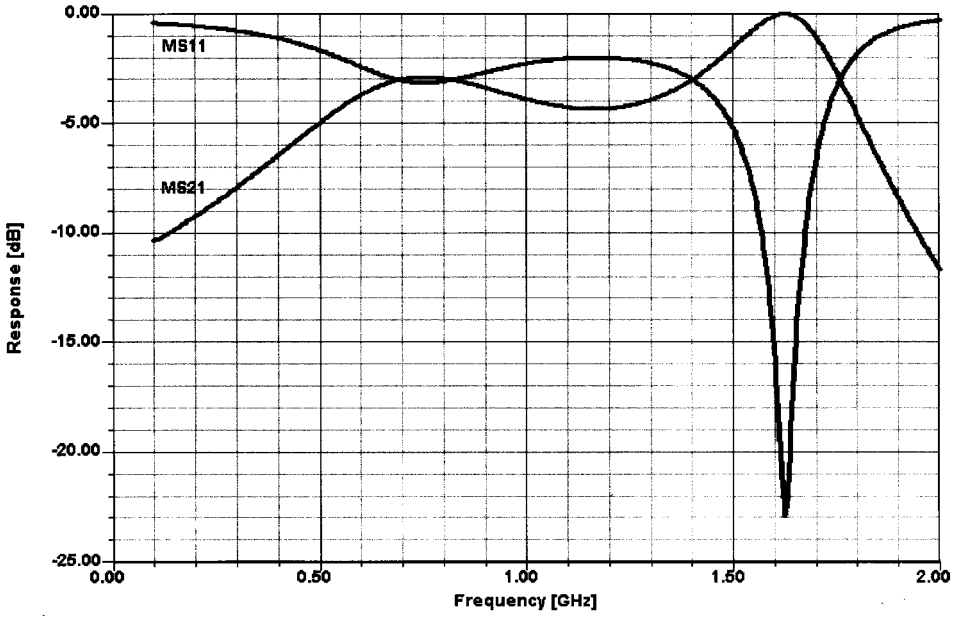


Figure 3-204 Frequency response of the input matching network.

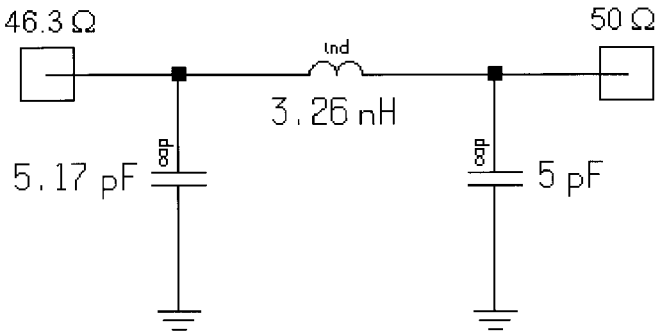


Figure 3-205 Output matching network for the amplifier.

The selection of the actual component is another issue, because these capacitors need to handle several amperes of RF current. It is not always easy to find capacitors with the appropriate value that can handle the current. The inductors can be printed. The paragraph following this section therefore will deal with the issue of how to translate the lumped values into distributed values, which will make the current handling much easier (we hope).

For this, we did the same optimizing approach by using a 50-Ω source and a 46.3-Ω load. Figures 3-205 and 3-206 show the circuit with the appropriate values and its simulated frequency-dependent responses.

The next step is to connect the input and output matching network to an "actual" transistor. This is a point where difficulties arise, since very few, if any, companies provide SPICE-type

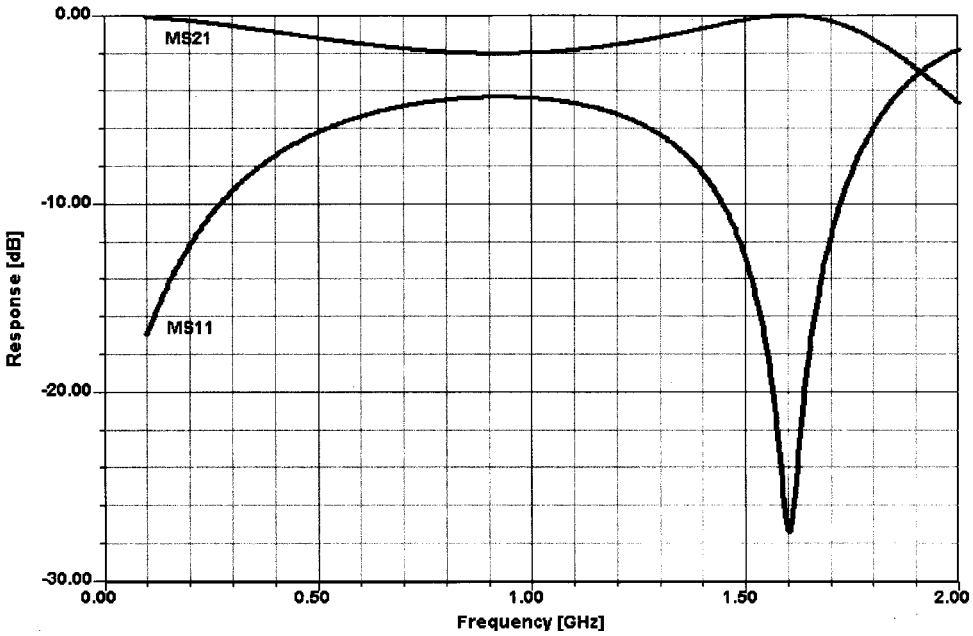


Figure 3-206 Frequency response of the output matching network.

parameters for transistors operating at currents of more than 200 mA. Figures 3-207 and 3-208 show the overall schematic and frequency response of the actual amplifier with all filters connected. Figure 3-209 details the amplifier's peak gain response. If more bandwidth is needed, the matching circuit must be increased in complexity, as will be shown in the following section.

Now is a good time to look at the final results of our simulation, which is based on the various assumptions as outlined above. Figure 3-210 shows the dc *I–V* curves with a straight load line and no reactances. This was accomplished by using a wideband transformer instead of an output matching circuit. Figure 3-211, however, shows the "load line" that results when the transformer is replaced with a real matching network. Based on the harmonics content and its improper termination, one can see how the output load line opens and shows ringing in the low-current area. If we expect to see the saturation voltage, we will be disappointed because the saturation voltage is frequency dependent (increases with frequency) and this figure gives an erroneous impression. Hence, we show dc *I–V* curves. To display the conduction angle, which we showed previously, we examine the output current as a function of time at a drive level of +16.5 dBm (Figure 3-212). During the negative half, the transistor is not conducting as predicted, so this validates our assumptions. Next, we are interested in determining the power gain. Therefore, we look at the output spectrum for a drive level of +16.5 dBm (Figure 3-213). It shows an output level of 38.6 dBm. If we subtract the drive level (16.5 dBm) from this, we see that the power gain is 22.1 dB at this drive level. Since we are used to thinking in watts, Figure 3-214 shows the absolute output power and the harmonics prior to filtering. Figure 3-215 shows the output spectrum after replacing the transformer with a simple pi output network.

During the conduction period, there will be a voltage across the transistor 180° out of phase, as shown in Figure 3-216. Finally, to highlight the nonlinearity of a Class C amplifier, Figure 3-217 shows the amplifier's output power versus drive, and Figure 3-218 shows the amplifier's gain versus drive. When the transistor turns on is clearly evident.

A particular difficulty arises in the case of FETs, where the charge energy makes for a difference between dc and pulsed measurements. Figure 3-219 shows this. Some of the more modern transistors, such as SiGe types, may need some pulsing also; as can be seen in Figure 3-219, even depending on the pulsing, one may get different dc *I–V* curves. This area is still under investigation, and even the "experts" do not agree on all aspects of this. In real life, it turns out that a lot of experimentation is the answer. Anyone who thinks that a pure CAD approach will give the right design answers immediately will encounter many surprises! This is not a fault of the CAD tools—assuming that they are properly engineered—but rather the availability of appropriately "tweaked" models for high-power applications, and the willingness of device suppliers to standardize on certain models and support them.

Because of the difficulties involved in cooling transistors under test and simulating RF pulse and CW conditions, depending on the application, manufacturers have so far stayed away from investing money in this area. The designer must therefore resort to application notes published by device manufacturers. Because even slight layer changes can have a large impact on the frequency response of a given device, application-note-based designs can run into difficulty if a device's fabrication has significantly evolved since the publication of its application notes.

Given the multitude of models outlined in Chapter 2, the CAD user is sometimes left hanging as to which is the best model to use. The latest indications are that the MEXTRAM (*m*ost *e*xquisite *t*ransistor *m*odel), developed by the University of Delft, Holland, with Philips of Holland, shows the most promising result, specifically in the area of IMD products. While

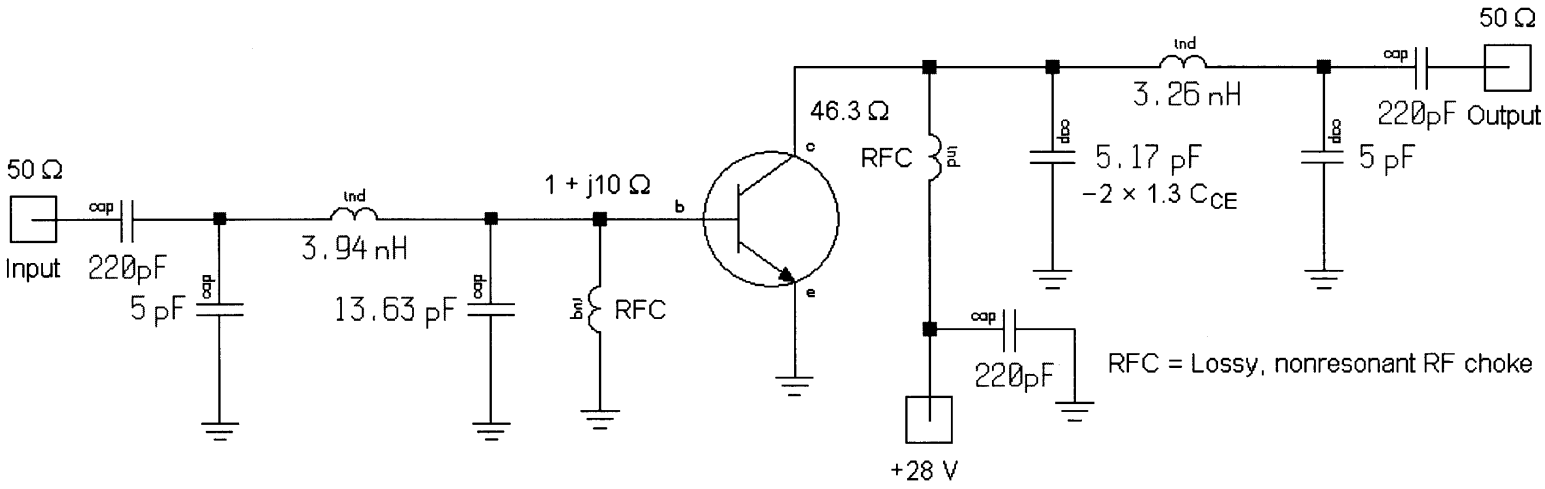


Figure 3-207 Actual schematic of our experimental 1.6-GHz amplifier. The RF chokes have enough RF losses not to be resonant and yet don't load the circuit. This is frequently accomplished with either ferrite beads or resistors in parallel with the chokes.

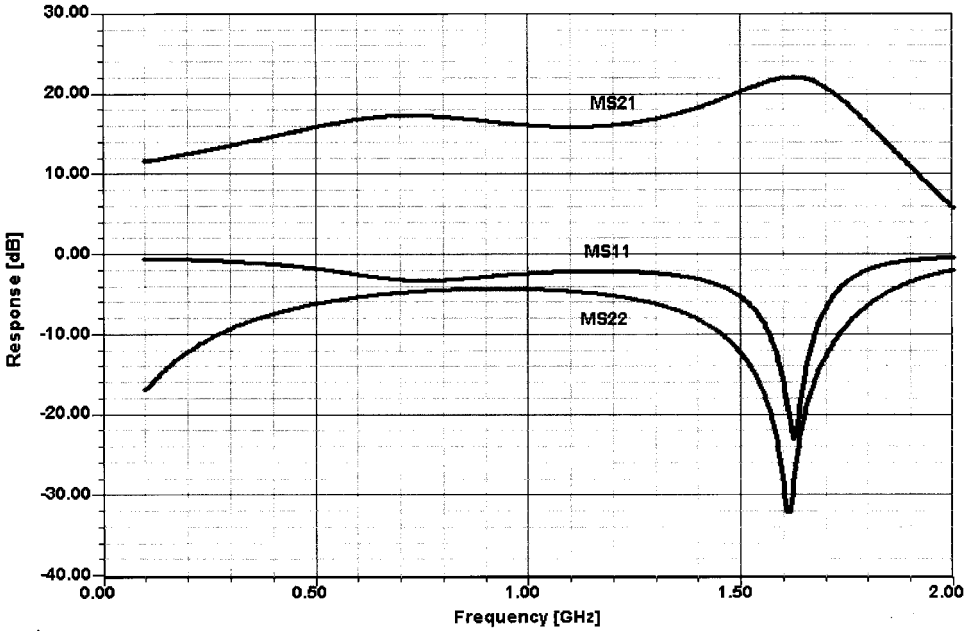


Figure 3-208 Overall frequency response of our experimental amplifier. Input and output matching, and gain, are shown as a function of frequency.

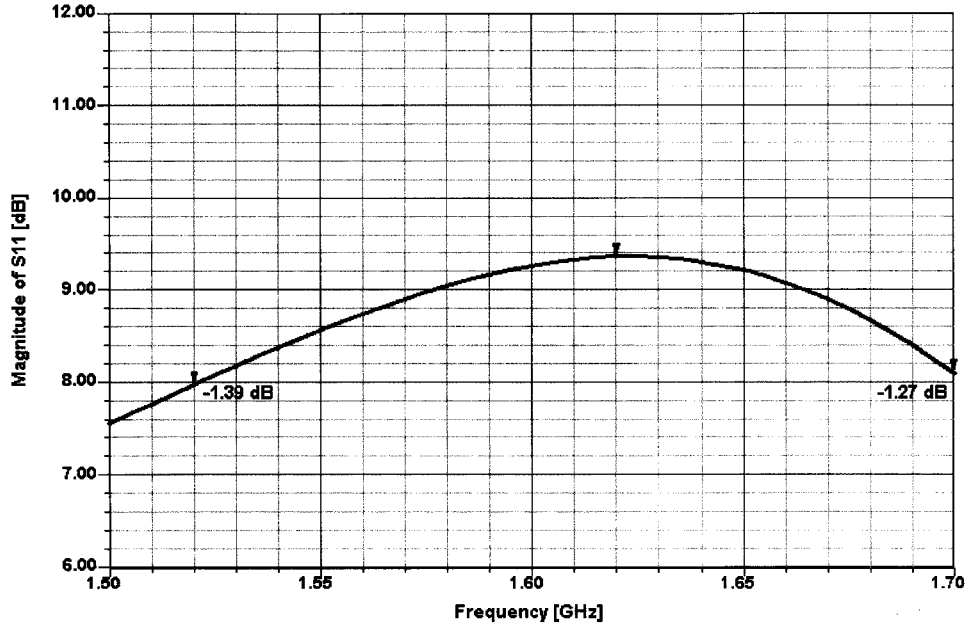


Figure 3-209 Close-in display of the frequency-dependent gain.

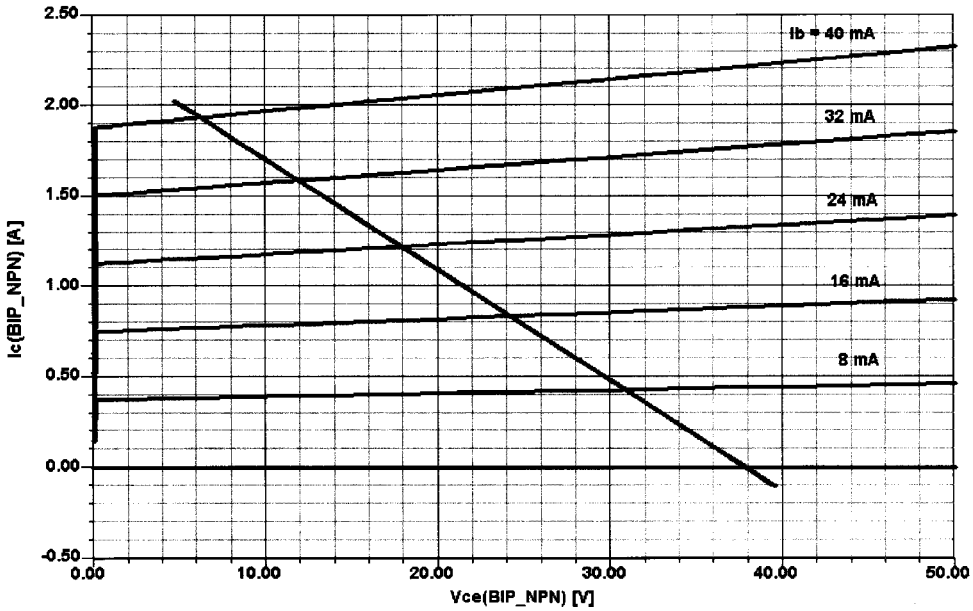


Figure 3-210 The dc *I–V* curves assuming a perfect resistive load. No RF saturation voltage or other RF saturation effects are evident.

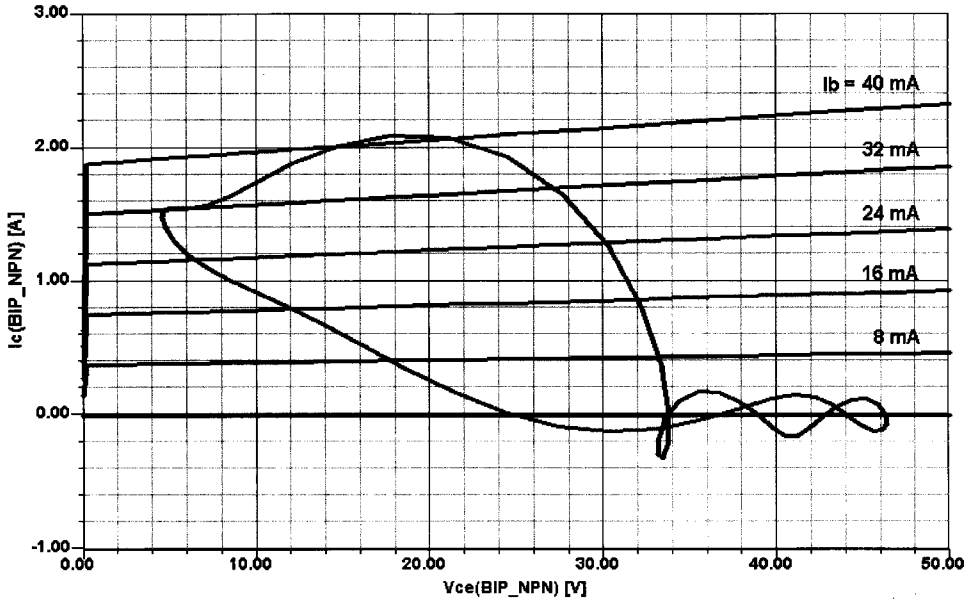


Figure 3-211 Load line with the output transformer replaced with a real matching circuit.

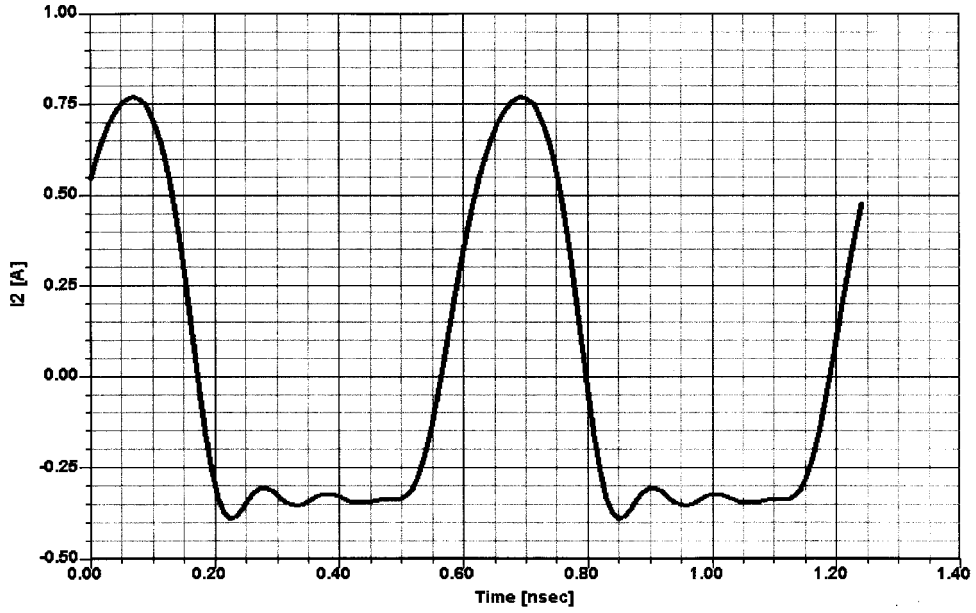


Figure 3-212 Output current as a function of time. The duty cycle is determined by the device conduction angle.

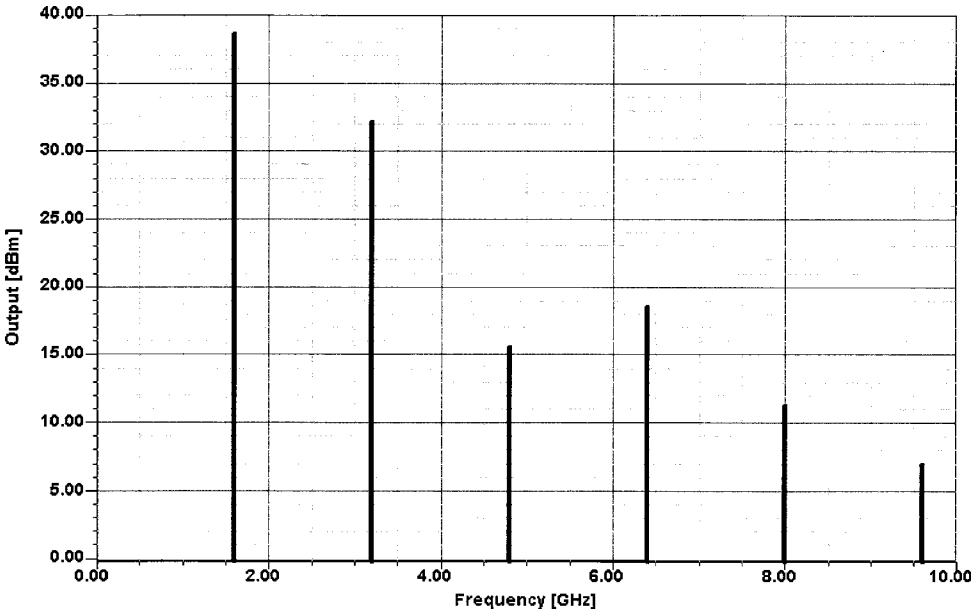


Figure 3-213 Power output at the collector prior to any filtering, assuming a real output termination.

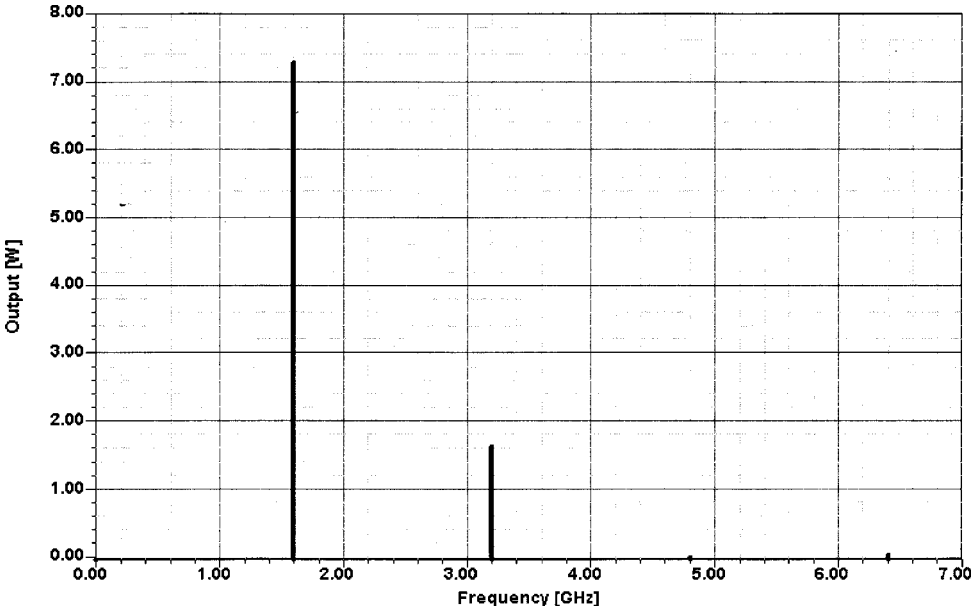


Figure 3-214 Available output power in watts, prior to any filtering, assuming a real output termination.

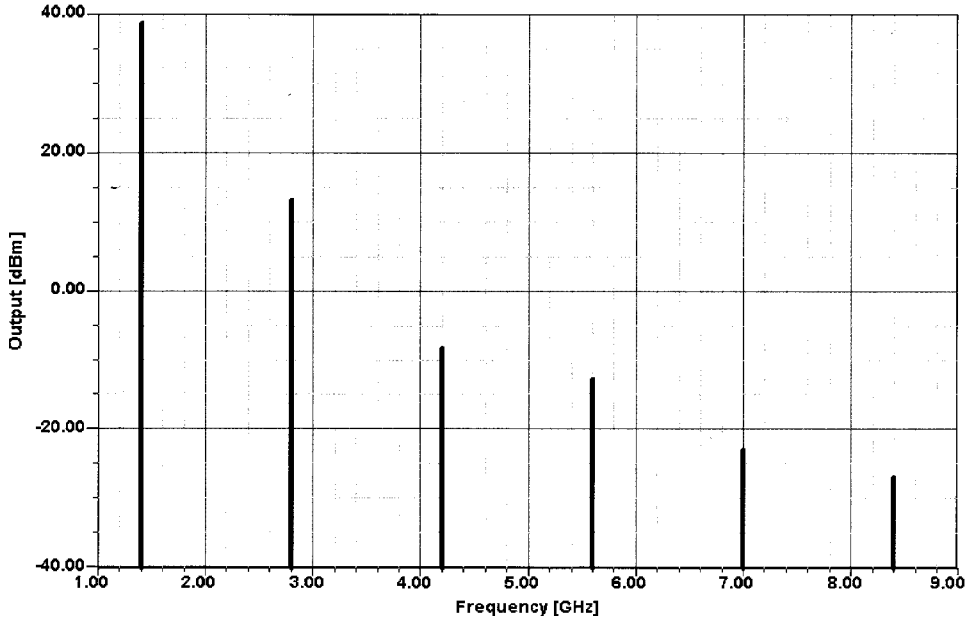


Figure 3-215 The output spectrum after replacing the transformer with a simple pi output network.

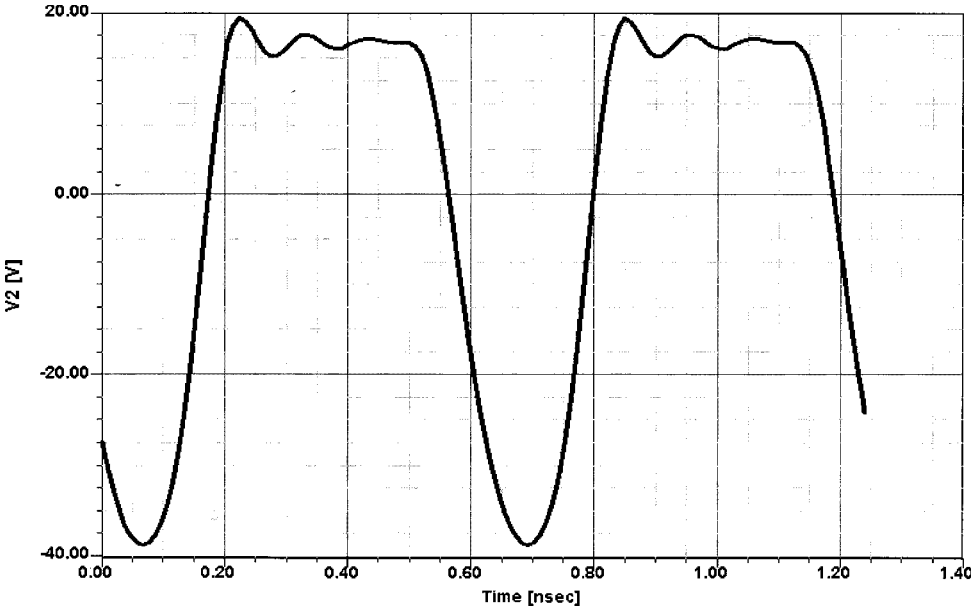


Figure 3-216 Output voltage at the collector. The voltage is asymmetrical compared to 0; this is important in the selection of a device capable of handling these voltages.

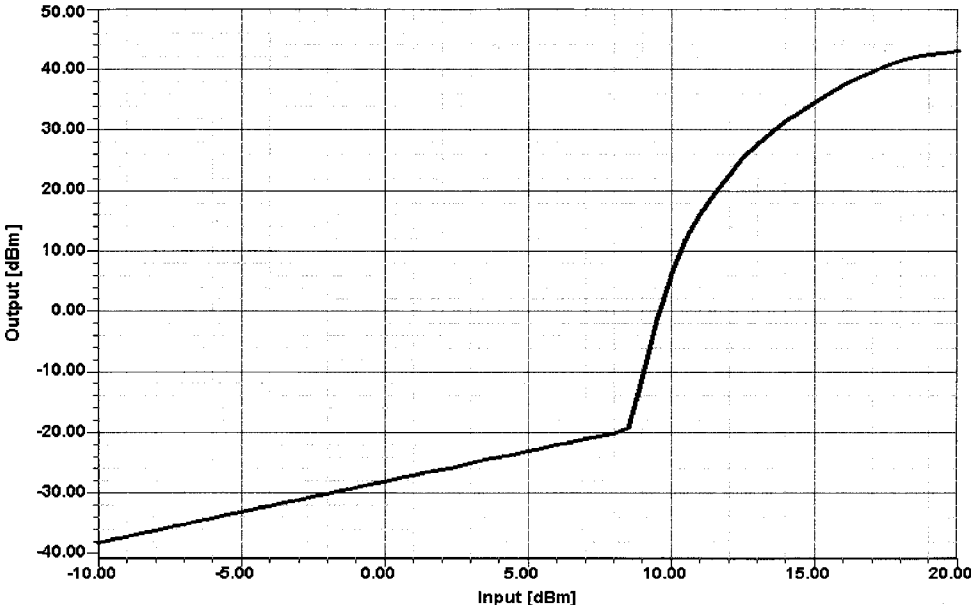


Figure 3-217 Output power versus input power for the Class C amplifier. Until the transistor turns on (at a drive level of about 8.5 dBm), only fedthrough power appears at the output.

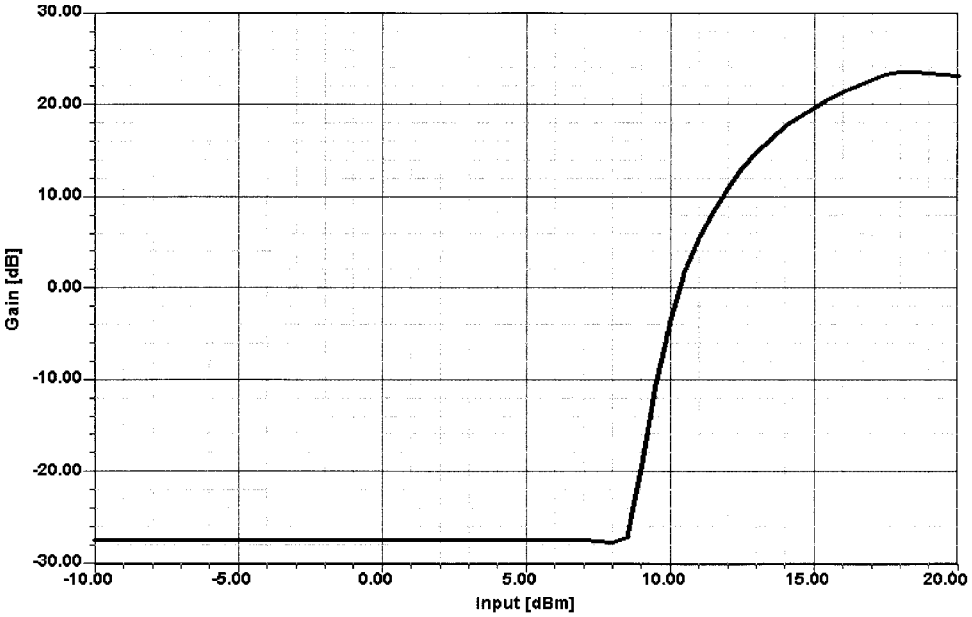


Figure 3-218 Amplifier gain as a function of drive. It is evident that one needs to cross the 0.65-V point for the transistor to become conductive and show gain. The initial amplification as a function of drive shows a very steep curve that then levels off and drops as the amplifier moves into saturation.

other later models have up to seven internal nodes, MEXTRAM fortunately uses only five. (The MEXTRAM model actually consists of three transistors, which makes modeling fun.) Power-added efficiency simulation is not available, since the MEXTRAM model was developed for small-signal applications, meaning circuits that operate at less than 100 mW of RF output power.

As if the absence of trustworthy high-power device models is not challenge enough, in developing this 1.6-GHz Class C amplifier example, we ran into another complication: Using a CAD optimizer to simultaneously match the circuit's input and output networks failed because of the device's input-impedance variation with drive. Every matching adjustment that improved the drive to the device resulted in an input-impedance change that reduced the drive to the device! In the end, the optimizer couldn't keep up. In actually building such a circuit, one would make the matching-network values adjustable on the initial breadboard and then, with the necessary input drive level applied, iteratively adjust the input and output matching for the desired output level. It appears that the CAD program could not handle such a degree of nonlinearity.

We then went to the approach of using large-signal *S* parameters at a drive level slightly less than 16 dBm, but this detuned the input filter, resulting in a poorer input match than the one initially predicted. Likewise, the output matching, which in a power stage must result in maximum output power rather than a conjugate match, was less impressive. The large-signal *S* parameters at the output for a power amplifier are somewhat dubious because Class C operation generates so many harmonics that the output termination will reflect a lot of energy and, depending on the output filter topology, different results can be expected. We firmly believe that there will be further mathematical improvement in the CAD tools, and we very much invite

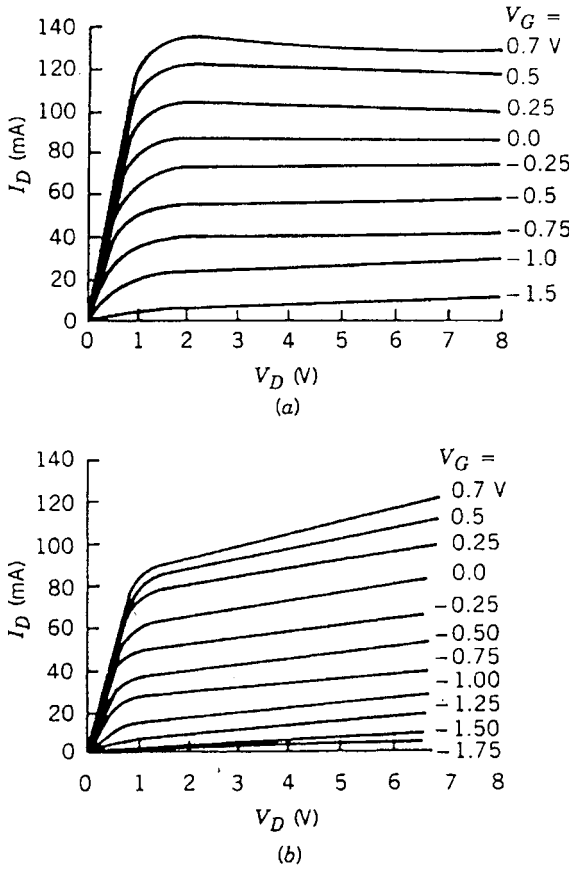


Figure 3-219 Comparison of FET drain current characteristics obtained from (a) pulsed *I–V* and (b) rectified sine-wave evaluation. The rectified sine-wave curves can be considered equivalent to curves obtained at dc.

our readers to use such a nonlinear example to test their tools, provided that a high-power model becomes available in this power range.

It has been rumored that the load-pull technique, rather than the application of large-signal *S* parameters, is a vehicle to solve such matching problems. However, the load-pull technique is really applied only to the fundamental frequency and does not deal with the issue of harmonics of the output current. More details about the load-pull technique and its capabilities are found in Appendix B.

Actual measurements of amplifiers of this type show that our simulated efficiency and drive level come quite close to reality, but at a frequency somewhat offset from that chosen for the design. As an example, the input and output filters tuned at 1.4 GHz instead of 1.6 GHz using the large-signal *S* parameters. Again, an iterative process, similar to manually trimming a prototype, will give a similar result. This is consistent with our earlier statement that most final designs are really hand-tweaked, and that CAD can only bring us close, but not all the way, to a good solution. For passive problems, the CAD does a significantly better job; the nonlinear cases are the ones that still cause headaches.

3-12-2 Impedance Matching Networks Applied to RF Power Transistors*

Introduction. Some graphical and numerical methods of impedance matching will be reviewed here. The examples given refer to high-frequency power amplifiers.

Although matching networks normally take the form of filters and therefore are also useful to provide frequency discrimination, this aspect will be considered only as a corollary of the matching circuit.

Matching is necessary for the best possible energy transfer from stage to stage. In RF power transistors, the input impedance is of low value, decreasing as the power increases or as the chip size becomes larger. This impedance must be matched either to a generator—of generally 50-Ω internal impedance—or to a preceding stage. Impedance matching has to be made between two complex impedances, which makes the design still more difficult, especially if matching must be accomplished over a wide frequency band.

Device Parameters

Input Impedance. The general shape of the input impedance of RF power transistors is as shown in Figure 3-220. It is a large-signal parameter, expressed here by the parallel combination of a resistance R_p and a reactance X_p [31].

The equivalent circuit shown in Figure 3-221 accounts for the behavior illustrated in Figure 3-220. With the presently used stripline or flange packaging, most of the power devices for VHF low band will have their R_p and X_p values below the series resonant point f_s. The input impedance will be essentially capacitive.

Most of the VHF high-band transistors will have the series-resonant frequency within their operating range; that is, they will be purely resistive at one single frequency f_s, which the parallel-resonant frequency f_p will be outside.

Parameters for 1- or 2-GHz transistors will be beyond f_s and approach f_p. They show a high value of R_p and X_p with inductive character.

A parameter that is very often used to judge on the broadband capabilities of a device is the input Q or Q_{IN}, defined simply as the ratio R_p/X_p. Practically, Q_{IN} ranges around 1 or less for VHF devices and around 5 or more for microwave transistors.

Q_{IN} is an important parameter to consider for broadband matching. Matching networks normally are low-pass or pseudo-low-pass filters. If Q_{IN} is high, it can be necessary to use bandpass filter-type matching networks and to allow insertion losses. But broadband matching is still possible. This will be discussed later.

Output Impedance. The output impedance of RF power transistors, as given by all manufacturers' datasheets, generally consists of only a capacitance C_{OUT}. The internal resistance of the transistor is supposed to be much higher than the load and is normally neglected. In the case of a relatively low internal resistance, the efficiency of the device would decrease by the factor

$$1 + R_L/R_T \tag{3-192}$$

*Based on Motorola application note AN-721 ("Impedance Matching Networks Applied to RF Power Transistors"). Copyright of Motorola; used by permission.

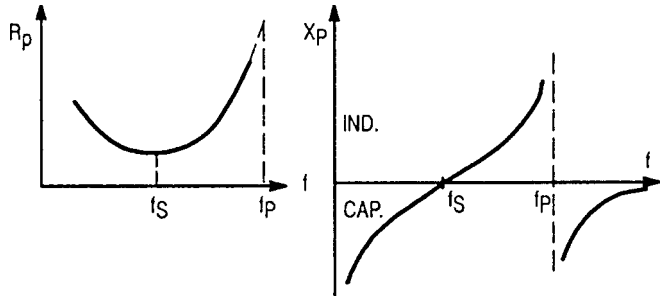


Figure 3-220 Input impedance of RF power transistors as a function of frequency.

where R_L is the load resistance, seen at the collector–emitter terminals, and R_T is the internal transistor resistance equal to

$$\frac{1}{\omega_T (C_{TC} + C_{DC})} \tag{3-193}$$

defined as a single parameter, where ω_T is the transit angular frequency and C_{TC} and C_{DC} are the transition and diffusion capacitances, respectively, at the collector junction.

The output capacitance, C_{OUT}, which is a large-signal parameter, is related to the small-signal parameter C_{CB}, the collector–base transition capacitance.

Since a junction capacitance varies with the applied voltage, C_{OUT} differs from C_{CB} in that it has to be averaged over the total voltage swing. For an abrupt junction and assuming certain simplifications, $C_{\text{OUT}} = 2C_{CB}$.

Figure 3-222 shows the variation of C_{OUT} with frequency. C_{OUT} decreases partly due to the presence of the collector lead inductance, but mainly because of the fact that the base–emitter diode does not shut off anymore when the operating frequency approaches the transit frequency, f_T.

Output Load. In the absence of a more precise indication, the output load R_L is taken as

$$R_L = \frac{[V_{CC} - V_{CE(\text{sat})}]^2}{2P_{\text{OUT}}} \tag{3-194}$$

with $V_{CE(\text{sat})}$ equal to 2 or 3 V, increasing with frequency.

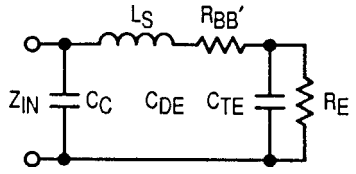


Figure 3-221 Equivalent circuit for the input impedance. R_E is the emitter diffusion resistance; C_{DE} and C_{TE} are the diffusion and transition capacitances, respectively, of the emitter junction; $R_{BB'}$ is the base-spreading resistance; C_C is the package capacitance; and L_S is the base lead inductance.

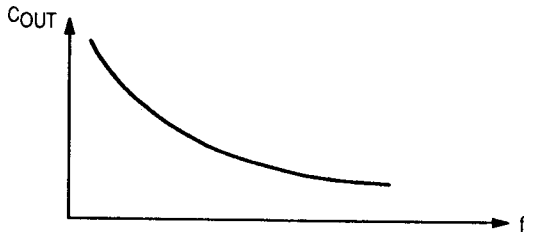


Figure 3-222 Output capacitance C_{OUT} as a function of frequency.

The above equation just expresses a well-known relation, but it also shows that the load, in first approximation, is not related to the device, except for $V_{CE(sat)}$. The load value is primarily dictated by the required output power and the peak voltage; it is not matched to the output impedance of the device.

At high frequencies this approximation becomes less exact, and for microwave devices the load that must be presented to the device is indicated on the datasheet.

Strictly speaking, impedance matching is accomplished only at the input. Interstage and load matching are more impedance transformations of the device input impedance and of the load into a value R_L (sometimes with additional reactive component) that depends essentially on the power demanded and the supply voltage.

Matching Networks. In the following, matching networks will be described by the order of complexity. These are ladder-type reactance networks. The different reactance values will be calculated and determined graphically. Increasing the number of reactances broadens the bandwidth. However, matching networks consisting of more than four reactances are rare. Above four reactances, the improvement is small.

Numerical Design

TWO-RESISTANCE NETWORKS. Resistance terminations will first be considered. Figure 3-223 shows the reactive L section and the terminations to be matched.

Matching or exact transformation from R_2 into R_1 occurs at a single frequency f_0. At f_0, X_1 and X_2 are equal to

$$X_1 = \pm R_1 \sqrt{\frac{R_2}{R_1 - R_2}} = R_1 \frac{1}{\sqrt{n-1}} \tag{3-195}$$

$$X_2 = \mp \sqrt{R_2(R_1 - R_{2)}} = R_1 \frac{\sqrt{n-1}}{1} \tag{3-196}$$

At f_0, $X_1X_2 = R_1R_2$. X_1 and X_2 must be of opposite sign. The shunt reactance is in parallel with the larger resistance.

The frequency response of the L section is shown in Figure 3-224, where the normalized current is plotted as a function of the normalized frequency.

If X_1 is capacitive and consequently X_2 is inductive, then

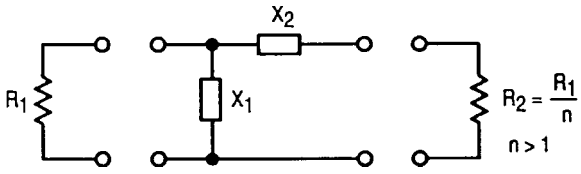


Figure 3-223 Two-reactance matching network.

$$X_1 = -\frac{f_0}{f} R_1 \sqrt{\frac{R_2}{R_1 - R_2}} = -\frac{f_0}{f} R_1 \frac{1}{\sqrt{n-1}} \tag{3-197}$$

and

$$X_2 = \frac{f}{f_0} \sqrt{R_2(R_1 - R_2)} = \frac{f}{f_0} R_1 \frac{\sqrt{n-1}}{1} \tag{3-198}$$

The normalized current absolute value is equal to

$$\left|\frac{I_2}{I_0}\right| = \frac{2\sqrt{n}}{\sqrt{(n-1)^2 \cdot (f/f_0)^4 - 2\,(f/f_0)^2 + (n+1)^2}} \tag{3-199}$$

where $I_0 = \sqrt{n}E/2R_1$ and is plotted in Figure 3-224 [32].

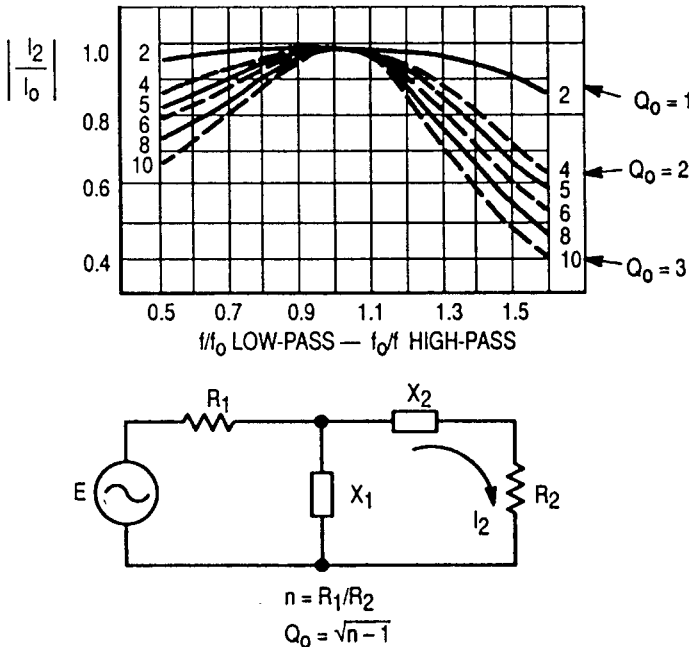


Figure 3-224 Normalized frequency response for the L section in low-pass or high-pass form.

If X_1 is inductive and consequently X_2 is capacitive, the only change required is a replacement of f by f_0 and vice versa. The L section has low-pass form in the first case and high-pass form in the second case.

The Q of the circuit at f_0 is equal to

$$Q_0 = \frac{X_2}{R_2} = \frac{R_1}{X_1} = \sqrt{n-1} \tag{3-200}$$

For a given transformation ratio n, there is only one possible value of Q. On the other hand, there are two symmetrical solutions for the network that can be either a low-pass filter or a high-pass filter.

The frequency f_0 does not need to be the center frequency, $(f_1 + f_2) / 2$, of the desired band limited by f_1 and f_2. In fact, as can be seen from the low-pass configuration of Figure 3-224, it may be interesting to shift f_0 toward the high-band edge frequency f_2 to obtain a larger bandwidth w, where

$$w = \frac{2(f_1 + f_2)}{f_2 - f_1} \tag{3-201}$$

This will, however, be at the expense of poor harmonic rejection.

Example. For a transformation ratio of $n = 4$, we can determine the following values from the above relations:

Bandwidth w	0.1	0.3
Maximum insertion losses	0.025	0.2
X_1/R_1	1.730	1.712

If the terminations R_1 and R_2 have a reactive component X, the latter may be taken as part of the external resistance as shown in Figure 3-225. This compensation is applicable as long as

$$Q_{int} = \frac{X_{\text{int}}}{R_2} \quad \text{or} \quad \frac{R_1}{X_{\text{int}}} < n - 1 \tag{3-202}$$

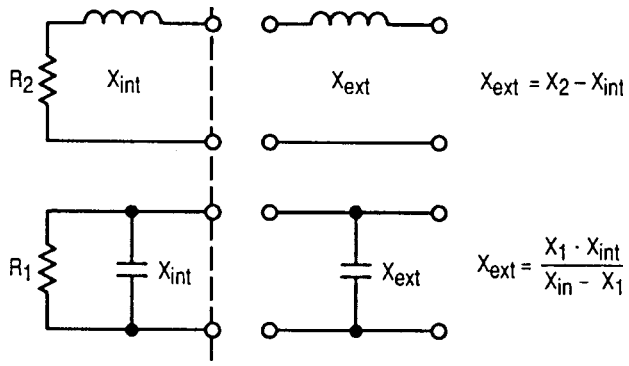


Figure 3-225 Termination reactance compensation.

Tables giving reactance values can be found in Matthaei et al. [33, 34].

USE OF TRANSMISSION LINES AND INDUCTORS. In the preceding section, the inductance was expected to be realized by a lumped element. A transmission line can be used instead (Figure 3-226).

As can be seen from the computed selectivity curves (Figure 3-227) for the two configurations, transmission lines result in a larger bandwidth. The gain is important for a transmission line having a length $L = \lambda/4$ ($\Theta = 90°$) and a characteristic impedance $Z_0 = \sqrt{R_1 R_2}$. It is not significant for lines short with respect to $\lambda/4$. One will notice that there is an infinity of solutions, one for each value of C, when using transmission lines.

THREE-REACTANCE MATCHING NETWORKS. The networks that will be investigated are shown in Figure 3-228. They are made of three reactances alternately connected in series and shunt.

A three-reactance configuration allows the designer to make the quality factor (Q) of the circuit and the transformation ratio $n = R_2/R_1$ independent of each other and consequently to choose the selectivity between certain limits.

For narrowband designs, one can use the following formulas (see [35], where tables are given):

Network (a)

$$X_{C1} = R_1/Q \quad Q \text{ must be selected first} \tag{3-203}$$

$$X_{C2} = R_2 \sqrt{\frac{R_1 R_2}{(Q^2 + 1) - R_1/R_2}} \tag{3-204}$$

$$X_L = \frac{QR_1 + (R_1 R_2/X_{C2})}{Q^2 + 1} \tag{3-205}$$

Network (b)

$$X_{L1} = R_1 Q \quad Q \text{ must be selected first} \tag{3-206}$$

$$X_{L2} = R_2 B \tag{3-207}$$

$$A = R_1(1 + Q^2) \tag{3-208}$$

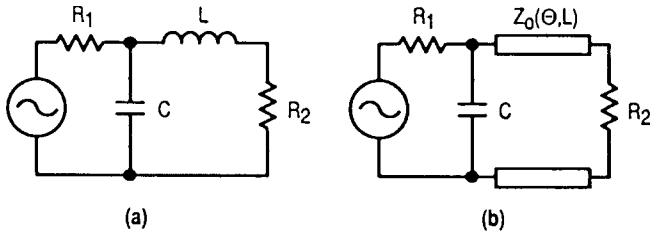


Figure 3-226 Use of a transmission line in the L section.

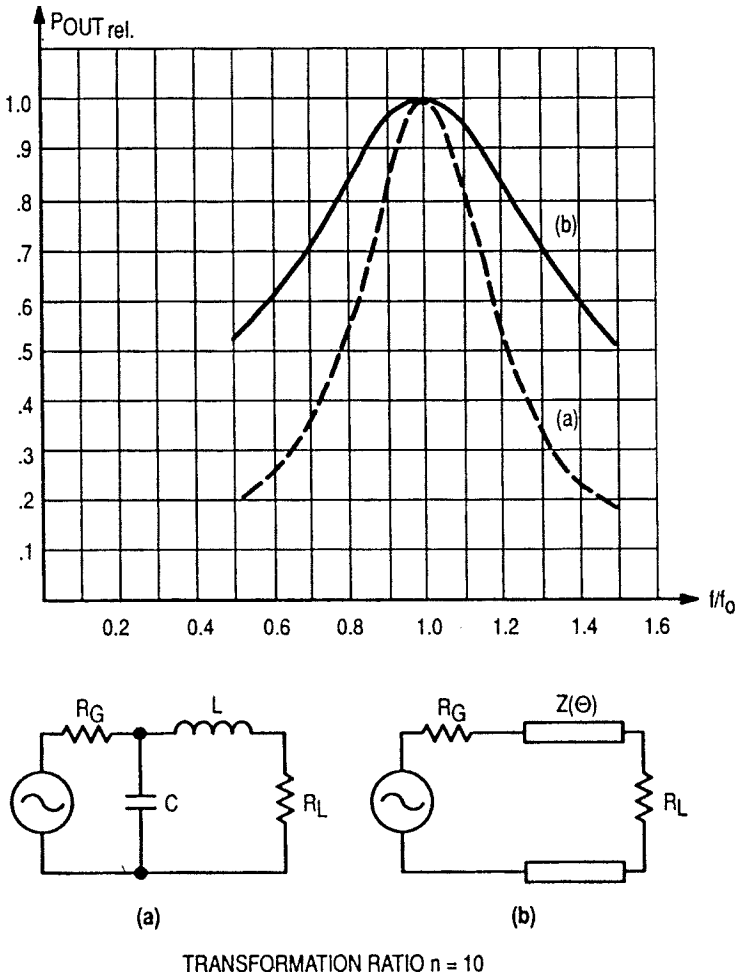


Figure 3-227 Bandwidth of the L section for $n = 10$ (a) with lumped constants and (b) with transmission line (λ/4).

$$X_{C1} = \frac{A}{Q + B} \tag{3-209}$$

$$B = \sqrt{\frac{A}{R_2} - 1} \tag{3-210}$$

Network (c)

$$X_{L1} = QR_1 \quad Q \text{ must be selected first} \tag{3-211}$$

$$X_{C2} = AR_2 \tag{3-212}$$

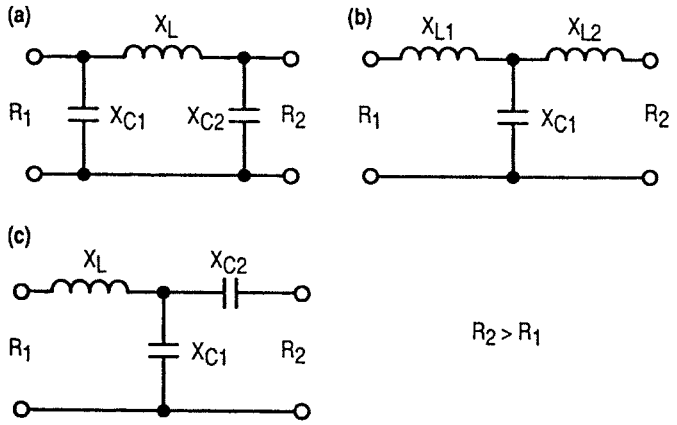


Figure 3-228 Three-reactance matching networks.

$$A = \sqrt{\frac{R_1(1+Q^2)}{R_2} - 1} \tag{3-213}$$

$$X_{C1} = \frac{B}{Q - A} \tag{3-214}$$

$$B = R_1(1+Q^2) \tag{3-215}$$

The network that yields the most-practical component values should be selected for a given application.

The three-reactance networks can be thought of as being formed of an L section (two reactances) and of a compensation reactance. The L section essentially performs the impedance transformation, while the additional reactance compensates for the reactive part of the transformed impedance over a certain frequency band.

Figure 3-229 shows a representation in the Z plane of the circuit of Figure 3-228a split into two parts: $R_1 - C_1 - L_1$ and $C_2 - R_2$. Exact transformation from R_1 into R_2 occurs at the points of intersection M and N. Impedances are then conjugate, or $Z' = R' + jX'$ and $Z'' = R'' + jX''$ with $R' = R''$ and $X' = -X''$.

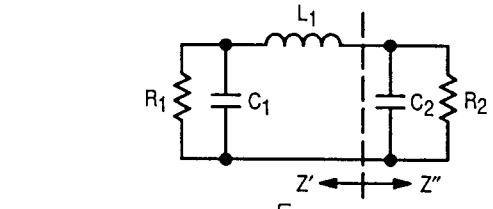
The only possible solution is obtained when X' and $-X''$ are tangential to each other. For the dashed curve, representing another value of L_1 or C_1, a wider frequency band could be expected at the expense of some ripple inside the band. However, this can only be reached with four reactances, as will be shown below.

With a three-reactance configuration, there are not enough degrees of freedom to permit $X' = -X''$ and simultaneously obtain the same variation of frequency on both curves from point M' to point N'. Exact transformation can, therefore, be obtained at only one frequency.

The values of the three reactances can be calculated by making

$$X' = -X'', \quad R' = R'', \quad \text{and} \frac{dX'}{dR'} = -\frac{dX''}{dR''} \tag{3-216}$$

The general solution of these equations leads to complicated calculations. Therefore, computed tables should be used.



$$Z' = \frac{R_1}{1 + \omega^2 R_1^2 C_1^2} + j\omega \left[L1 - \frac{R_1^2 C_1}{1 + \omega^2 R_1^2 C_1^2} \right] = R' + jX'$$

$$Z'' = \frac{R_2}{1 + \omega^2 C_2^2 R_2^2} - \frac{j\omega^2 C_2 R_2^2}{1 + \omega^2 C_2^2 R_2^2} = R'' + jX''$$

$$|x'| = \left(\frac{L_1}{C_1 R_1^2} - \frac{R'}{R_1} \right) \sqrt{\frac{R_1}{R'} - 1}$$

$$|x''| = \sqrt{R' R_2 - R'^2}$$

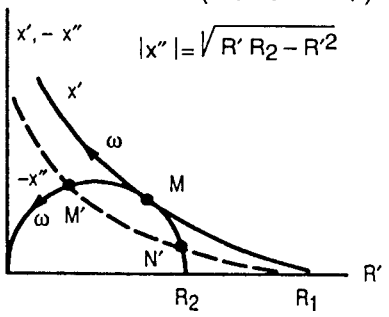


Figure 3-229 *Z*-plane representation of the circuit of Figure 3-228a.

One will note from Figure 3-229 that the compensation reactance contributes somewhat to the impedance transformation; that is, R' varies when going from M to R_2.

The circuit of Figure 3-228b is dual with respect to the first one and gives exactly the same results in a *Y*-plane representation. The circuit of Figure 3-228c is somewhat different since only one intersection M exists, as shown in Figure 3-230. Narrower frequency bands must be expected from this configuration. The widest band is obtained for $C_1 = \infty$.

Again, if one of the terminations has a reactive component, the latter can be taken as a part of the matching network, provided that it is not too large (see Figure 3-227).

FOUR-REACTANCE NETWORKS. Four-reactance networks are used essentially for broadband matching. The networks that will be considered in the following consist of two two-reactance sections in cascade. Some networks have pseudo-low-pass filter character; others, bandpass filter character. In principle, the former show narrower bandwidth since they extend the impedance transformation to very low frequencies unnecessarily, while the latter ensure good matching over a wide frequency band around the center frequency only (see Figure 3-233).

The two-reactance sections used in the networks in Figure 3-231 have either transformation properties or compensation properties. Impedance transformation is obtained with one series reactance and one shunt reactance. Compensation is made with both reactances in series or in shunt. If two cascaded transformation networks are used, transformation is accomplished partly by each one.

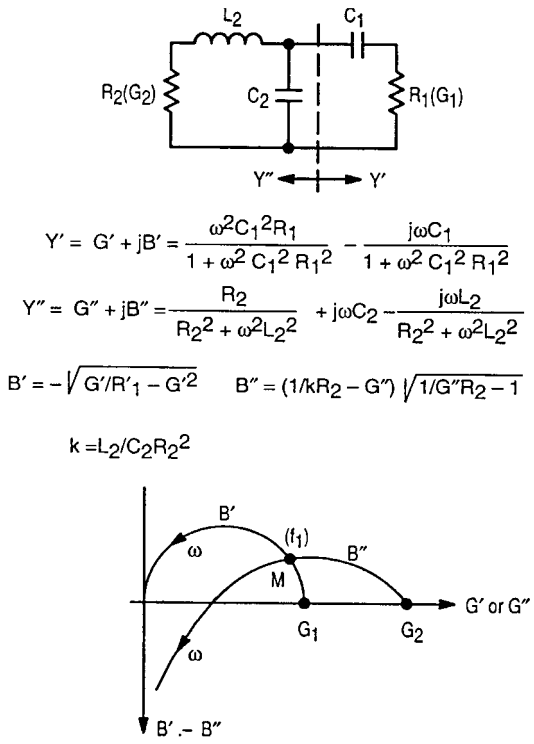


Figure 3-230 *Y*-plane representation of the circuit of Figure 3-228c.

With four-reactance networks there are two frequencies, f_1 and f_2, at which the transformation from R_1 into R_2 is exact. These frequencies may also coincide. For the network in Figure 3-231b, for instance, at point M, R_1 or R_2 is transformed into $\sqrt{R_1R_2}$ when both frequencies fall together. At all points (M), Z_1 and Z_2 are conjugate if the transformation is exact.

In the case of Figure 3-231b, the reactances are easily calculated for equal frequencies:

$$X_1 = \frac{R_1}{\sqrt{(n-1)^{1/2}}} \tag{3-217}$$

$$X_2 = R_1\sqrt{\frac{n^{1/2}-1}{n}} \tag{3-218}$$

$$X_1X_4 = R_1R_2 = X_2X_3 \tag{3-219}$$

$$X_3 = \frac{R_1}{\sqrt{n(n^{1/2}-1)}} \tag{3-220}$$

$$X_4 = \frac{R_1}{n}\sqrt{n^{1/2}-1} \tag{3-221}$$

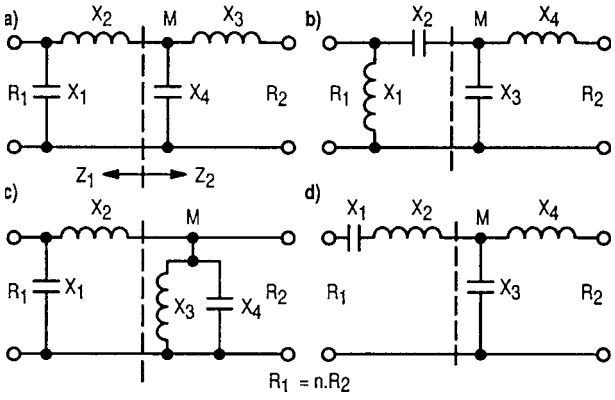


Figure 3-231 Four-reactance networks.

For the network in Figure 3-231a, normally at point M, Z_1 and Z_2 are complex. This pseudo-low-pass filter has been computed elsewhere [33]. Many tables can be found in the literature for networks of four and more reactances having Chebyshev character or maximally flat response [33, 34, 36].

Figure 3-232 shows the transformation path from R_1 to R_2 for networks (a) and (b) in Figure 3-231 on a Smith Chart.

Case (a) of Figure 3-232 has been calculated using tables in Matthaei [34]. Case (b) has been obtained from the relationship given above for $X_1, \ldots, X_4$. Both apply for a transformation ratio equal to 10 and for $R_1 = 1$. There is no simple relationship for $X_1', \ldots, X_4'$ of network (b) if f_1 is made different from f_2 for larger bandwidth. Figure 3-233 shows the respective bandwidths of networks (a) and (b) for the circuits shown in Figure 3-232.

If the terminations include a reactive component, the computed values for X_1 or X_4 may be adjusted to compensate for this.

For configuration (a), it can be seen from Figure 3-232 that, in the considered case, the Q values are equal to 1.6. For configuration (b), Q_1', which is equal to Q_2', is fixed for each transformation ratio

n	2	4	8	10	16
$Q'_1 = Q'_2$	0.65	1	1.35	1.46	1.73

$$Q' = \sqrt{n^{1/2} - 1} \qquad (3\text{-}222)$$

The maximum value of reactance that the terminations may have for use in this configuration can be determined from the above values of Q'.

If R_1 is the load resistance of a transistor, the internal transistor resistance may not be equal to R_1. In this case, the selectivity curve will be different from the curves given in Figure 3-233. Figure 3-234 shows the selectivity for networks (a) and (b) when the source resistance R_1 is infinite. From Figure 3-234 it can be seen that network (a) is more sensitive to changes in R_1 than network (b).

As mentioned earlier, the four-reactance network can also be thought of as two cascaded two-reactance sections: one used for transformation, the other for compensation. Figure 3-235 shows commonly used compensation networks, together with the associated L section.

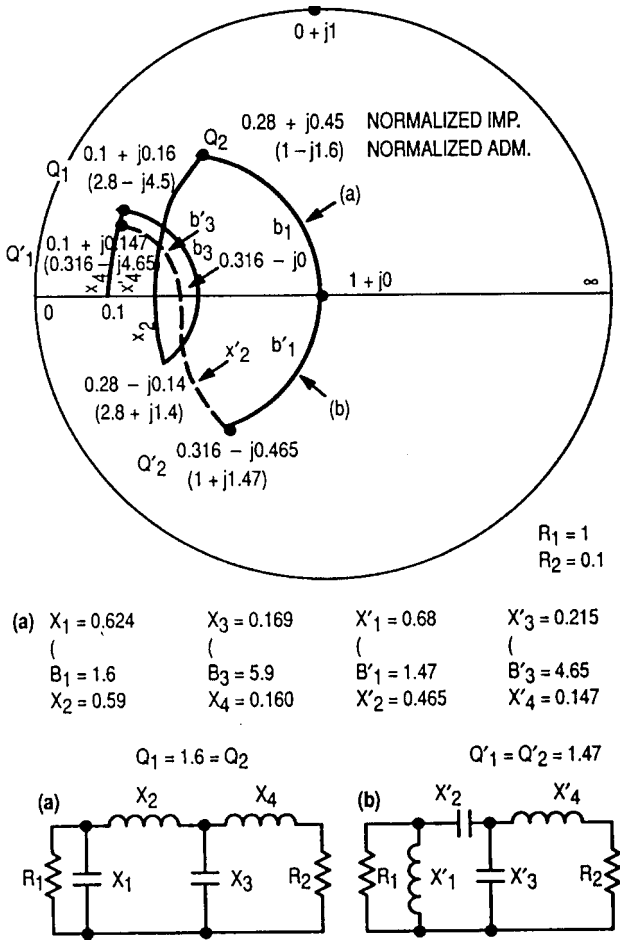


Figure 3-232 Transformation paths for networks (a) and (b).

The circuit of Figure 3-235a can be compared to the three-reactance network shown in Figure 3-228c. The difference is that capacitor C_2 of that circuit has been replaced by an *LC* circuit. The resulting improvement can be seen by comparing Figure 3-236 with Figure 3-230.

By adding one reactance, exact impedance transformation is achieved at two frequencies. It is now possible to choose component values such that the point of intersection M' occurs at the same frequency f_1 on both curves and simultaneously that N' occurs at the same frequency f_2 on both curves. Among the infinite number of possible intersections, only one allows the achievement of this.

When M' and N' coincide in M, the new $dX'/df = dX''/df$ condition can be added to the condition $X' = -X''$ (for three-reactance networks) and similarly $R' = R''$ and $dR'/df = dR''/df$.

Again, a general solution of the above equations leads to still more complicated calculations than in the case of three-reactance networks. Therefore, tables are preferable [33, 34, 36].

The circuit of Figure 3-235b is the dual of the circuit of Figure 3-233a and does not need to be treated separately. It gives exactly the same results in the *Z* plane. Figure 3-235c shows a higher-order compensation requiring six reactive elements.

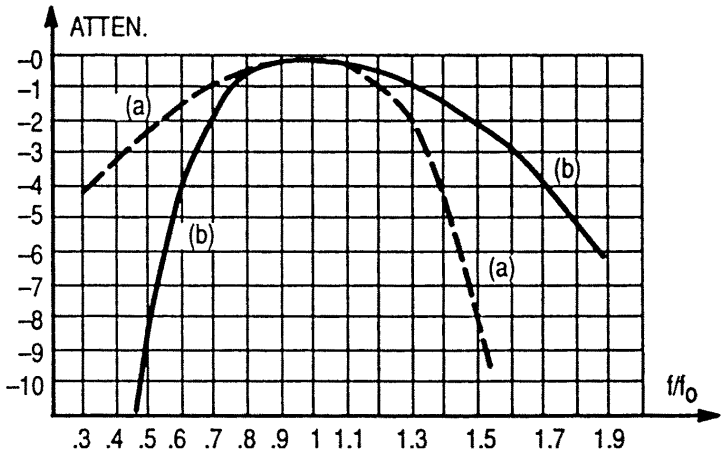


Figure 3-233 Selectivity curves for networks (a) and (b) of Figure 3-232.

The above-discussed matching networks employing compensation circuits result in narrower bandwidths than the former solutions (see p. 573) using two transformation sections. A matching arrangement with higher-order compensation, such as in Figure 3-235c, is not recommended. Better use can be made of the large number of reactive elements by using them all for transformation.

When the above configurations are realized using short portions of transmission lines, the equations or the usual tables no longer apply. The calculations must be carried out on a computer because of their complexity. However, a graphical method can be used (see the next section) that will consist essentially in tracing a transformation path on the Z–Y chart using the computed lumped element values and replacing it by the closest path obtained with distributed constants. The bandwidth change is not significant as long as short portions of lines are used.

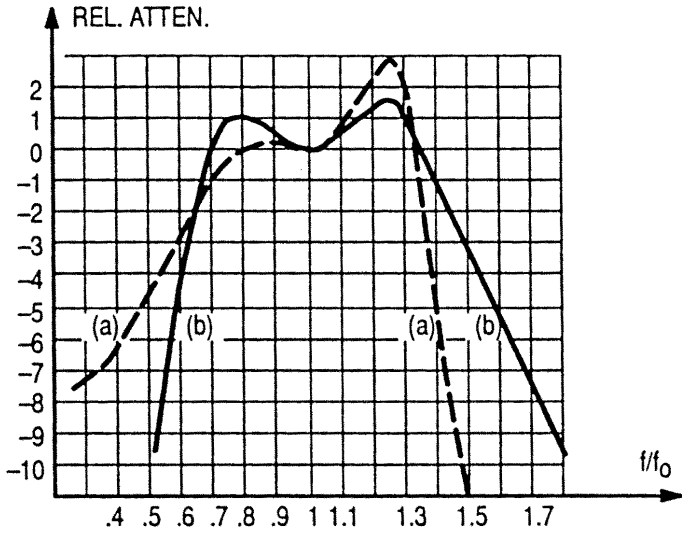


Figure 3-234 Selectivity curves for networks (a) and (b) of Figure 3-232 with infinite R_1.

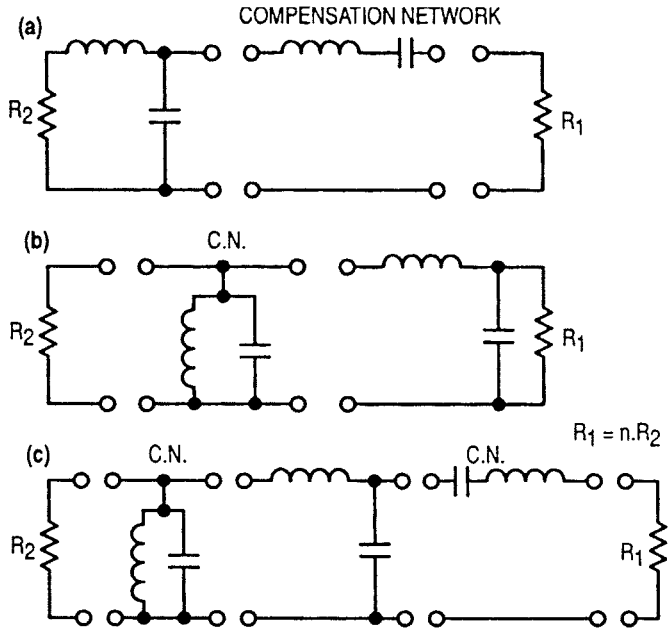


Figure 3-235 Compensation networks used with an L section.

MATCHING NETWORKS USING QUARTER-WAVE TRANSFORMERS. At sufficiently high frequencies, where λ/4-long lines of practical size can be realized, broadband transformations can easily be accomplished by the use of one or more λ/4 sections. Figure 3-237 summarizes the main relations for (a) one-section and (b) two-section transformations. A compensation network can be realized using a λ/2-long transmission line. Figures 3-238 and 3-239 show the selectivity curves for different transformation ratios and section numbers.

EXPONENTIAL LINES. Exponential lines have largely frequency-independent transformation properties. The characteristic impedance of such lines varies exponentially with their length l:

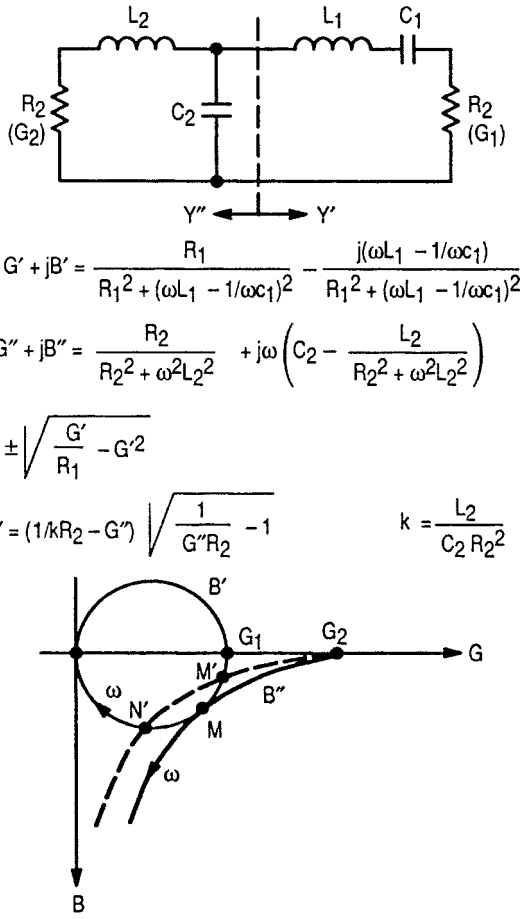
$$Z = Z_0 e^{kl} \tag{3-223}$$

where k is a constant, but these properties are preserved only if k is small.

BROADBAND MATCHING USING BANDPASS FILTER NETWORKS—HIGHQ CASE. The above circuits are applicable to devices having low input or output Q, if broadband matching is required. Generally, if the impedances to be matched can be represented, for instance, by a resistor R in series with an inductor L (sometimes a capacitor C) within the band of interest and if L is sufficiently low, the latter can be incorporated into the first inductor in the matching network. This is also valid if the representation consists of a shunt combination of a resistor and a reactance.

Practically, this is feasible for Q values around 1 or 2. For higher Q values or for input impedances consisting of a series or parallel resonant circuit (see Figure 3-221), as it appears to be for large bandwidths, a different treatment must be followed.

Let us first recall that, as shown by Bode [37] and Fano [38], limitations exist on the impedance matching of a complex load. In the example of Figure 3-240, the load to be matched consists of a capacitor C and a resistor R in shunt.



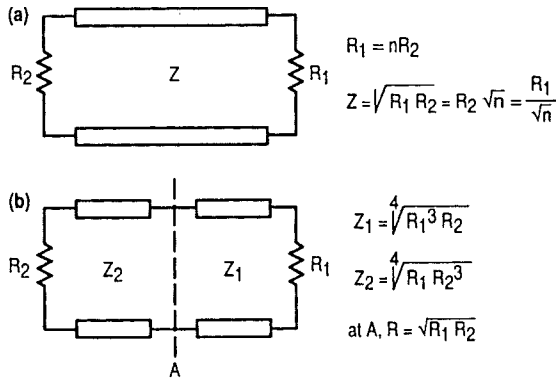
$$Y' = G' + jB' = \frac{R_1}{R_1^2 + (\omega L_1 - 1/\omega c_1)^2} - \frac{j(\omega L_1 - 1/\omega c_1)}{R_1^2 + (\omega L_1 - 1/\omega c_1)^2}$$

$$Y'' = G'' + jB'' = \frac{R_2}{R_2^2 + \omega^2 L_2^2} + j\omega\left(C_2 - \frac{L_2}{R_2^2 + \omega^2 L_2^2}\right)$$

$$B' = \pm\sqrt{\frac{G'}{R_1} - G'^2}$$

$$B'' = (1/kR_2 - G'')\sqrt{\frac{1}{G''R_2} - 1} \qquad k = \frac{L_2}{C_2 R_2^2}$$

Figure 3-236 *Y*-plane representation of the circuit of Figure 3-235a.



$$R_1 = nR_2$$

$$Z = \sqrt{R_1 R_2} = R_2\sqrt{n} = \frac{R_1}{\sqrt{n}}$$

$$Z_1 = \sqrt[4]{R_1^3 R_2}$$

$$Z_2 = \sqrt[4]{R_1 R_2^3}$$

$$\text{at A, } R = \sqrt{R_1 R_2}$$

Figure 3-237 Transformation networks using λ/4-long transmission lines.

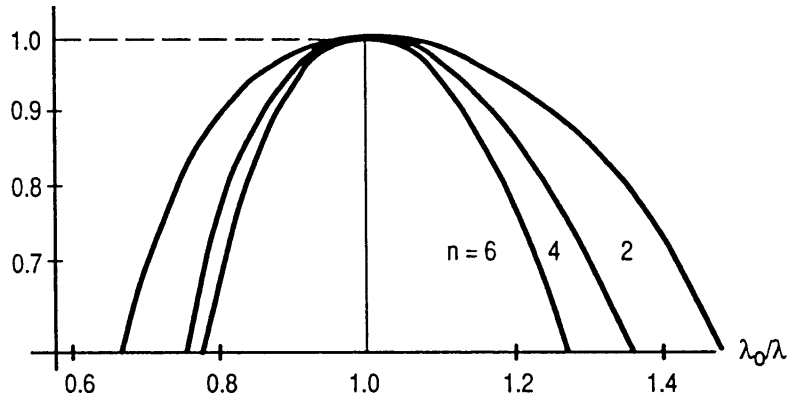


Figure 3-238 Selectivity curves for two λ/4-section networks at different transformation ratios.

The reflection coefficient between the transformed load and generator is equal to

$$\Gamma = \frac{Z_{\text{IN}} - R_G}{Z_{\text{IN}} + R_G} \tag{3-224}$$

When $\Gamma = 0$, there is perfect matching; when $\Gamma = 1$, there is total reflection. The ratio of reflected to incident power is

$$\frac{P_r}{P_i} = |\Gamma|^2 \tag{3-225}$$

The fundamental limitation on the matching takes the form

$$\int_{\omega=0}^{\infty} \ln\left(\frac{1}{|\Gamma|}\right) d\omega \le \frac{\pi}{RC} \qquad \text{Bode equation} \tag{3-226}$$

and is represented in Figure 3-241.

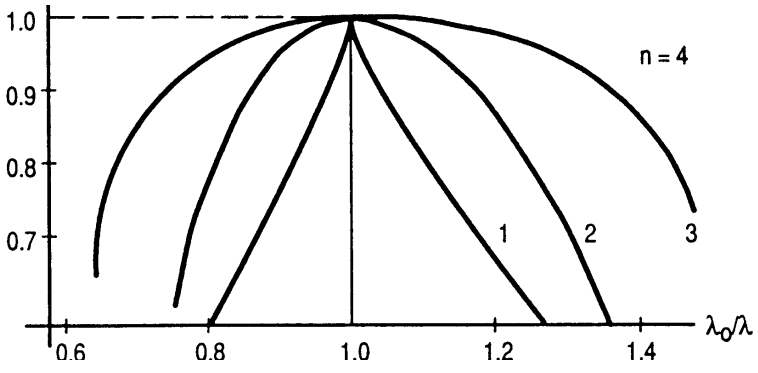


Figure 3-239 Selectivity curves for one, two, and three λ/4 sections.

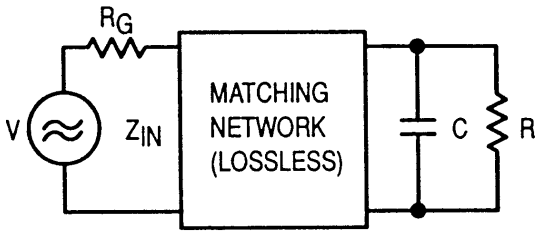


Figure 3-240 General matching conditions.

The meaning of the Bode equation is that the area S under the curve cannot be greater than π/RC and, therefore, if matching is required over a certain bandwidth, this can only be done at the expense of less power transfer within the band. Thus, power transfer and bandwidth appear as interchangeable quantities.

It is evident that the best utilization of the area S is obtained when $|\Gamma|$ is kept constant over the desired band ω_c and made equal to 1 over the rest of the spectrum. Then

$$|\Gamma| = e^{-\pi/\omega_c RC} \tag{3-227}$$

within the band and no power transfer happens outside. A network fulfilling this requirement cannot be obtained in practice because an infinite number of reactive elements would be necessary.

If the attenuation is plotted versus frequency for practical cases, one may expect to have curves like the ones shown in Figure 3-242 for a low-pass filter having Chebyshev character.

For a given complex load, an extension of the bandwidth from ω_1 to ω_2 is possible only with a simultaneous increase of the attenuation a. This is especially noticeable for Q values exceeding 1 or 2 (see Figure 3-243). Thus, devices having relatively high input Q values are usable for broadband operation, provided the consequent higher attenuation or reflection introduced is acceptable.

For average insertion losses or attenuation a (neglecting the ripple), the attenuation decreases if the number n of the network element increases. But above $n = 4$, the improvement is small. For a given attenuation a and bandwidth, the larger the n the smaller the ripple; for a given attenuation, the larger the n the larger the bandwidth.

Computations show that for $Q < 1$ and $n \leq 3$, the attenuation is below 0.1 dB (approximately). The impedance transformation ratio is not free here. The network is a

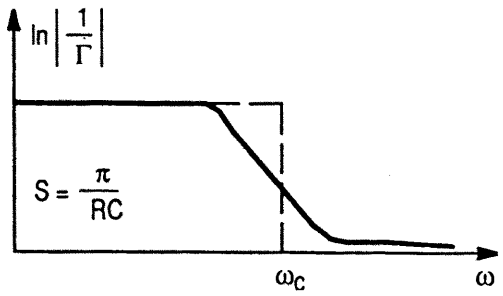


Figure 3-241 Representation of the Bode equation.

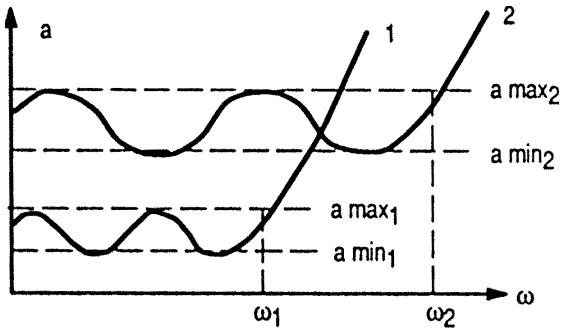


Figure 3-242 Attenuation versus angular frequency for different bandwidths with the same load.

true low-pass filter. For a given load, the optimum generator impedance will result from the computation.

Before impedance transformation is introduced, a conversion of the low-pass prototype into a bandpass filter network must be made. Figure 3-244 summarizes the main relations for this conversion. r is the conversion factor. For the bandpass filter, $Q_{\text{IN max}}$, or the maximum possible input Q of the device to be matched, has been increased by the factor r (from Figure 3-244, $Q'_{\text{IN max}} = rQ_{\text{IN max}}$).

Impedance inverters will be used for impedance transformation. These networks are suitable for insertion into a bandpass filter without affecting the transmission characteristics.

Figure 3-245 shows four impedance inverters. Note that one of the reactances is negative and must be combined in the bandpass network with a reactance of at least equal positive value. Insertion of the inverter can be made at any convenient place [32, 39].

When using the bandpass filter for matching the input impedance of a transistor, reactances $L'_1C'_1$ should be made to resonate at ω_0 by addition of a convenient series reactance.

As stated earlier, the series combination of R_0, L'_1 and C'_1 normally constitutes the equivalent input network of a transistor when considered over a large bandwidth. This is a good approximation up to about 500 MHz.

In practice, the normal procedure for using a bandpass filter matching network will be the following:

1. For a given bandwidth center frequency and input impedance of a device to be matched—for example, 50 Ω—first determine Q'_{IN} from the datasheet as

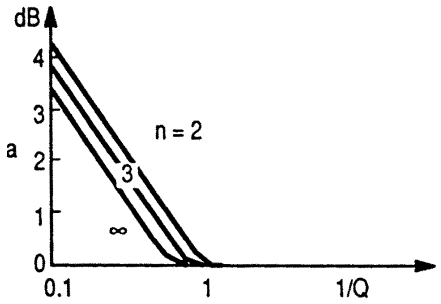


Figure 3-243 Insertion losses as a function of 1/*Q*.

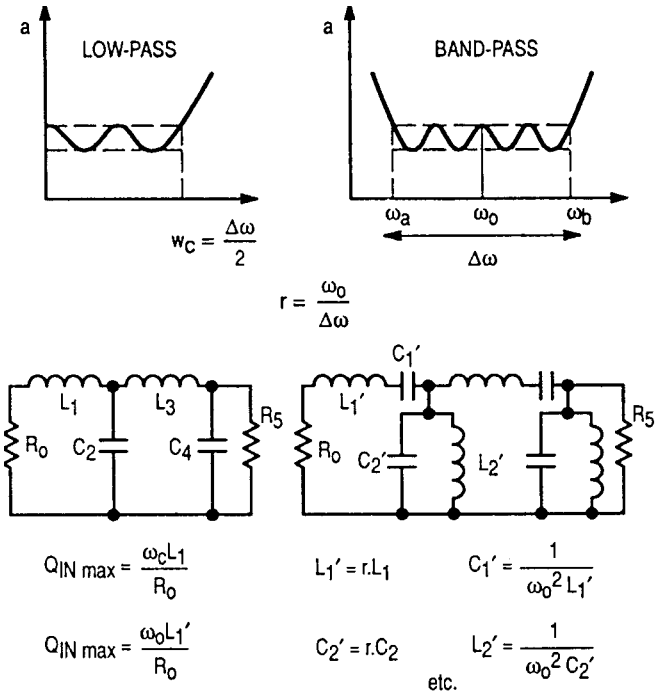


Figure 3-244 Conversion from low-pass filter into bandpass filter.

$$\frac{\omega_0 L_1'}{R_0} \tag{3-228}$$

after having eventually added a series reactor for centering.

2. Convert the equivalent circuit $R_0 L_1' C_1'$ into a low-pass prototype $R_0 L_1$ and calculate Q_{IN} using the formulas of Figure 3-244.
3. Determine the other reactance values from tables [33] for the desired bandwidth.
4. Convert the element values found by Step 3 into series- or parallel-resonant circuit parameters.
5. Insert the impedance inverter in any convenient place.

In the above discussions, the gain roll-off has not been taken into account. This is of normal use for moderate bandwidths (e.g., 30%). However, several methods can be employed to obtain a constant gain within the band despite the intrinsic gain decrease of a transistor with frequency. Tables have been computed elsewhere [40] for matching networks approximating 6 dB/octave attenuation versus frequency.

Another method consists of using the above-mentioned network and then adding a compensation circuit, as shown in Figure 3-246.

Resonance ω_b is placed at the high edge of the frequency band. Choosing Q correctly, roll-off can be made 6 dB/octave. The response of the circuit of Figure 3-246 is expressed by

$$\frac{1}{1 + Q^2(\omega/\omega_b - \omega_b/\omega)^2} \tag{3-229}$$

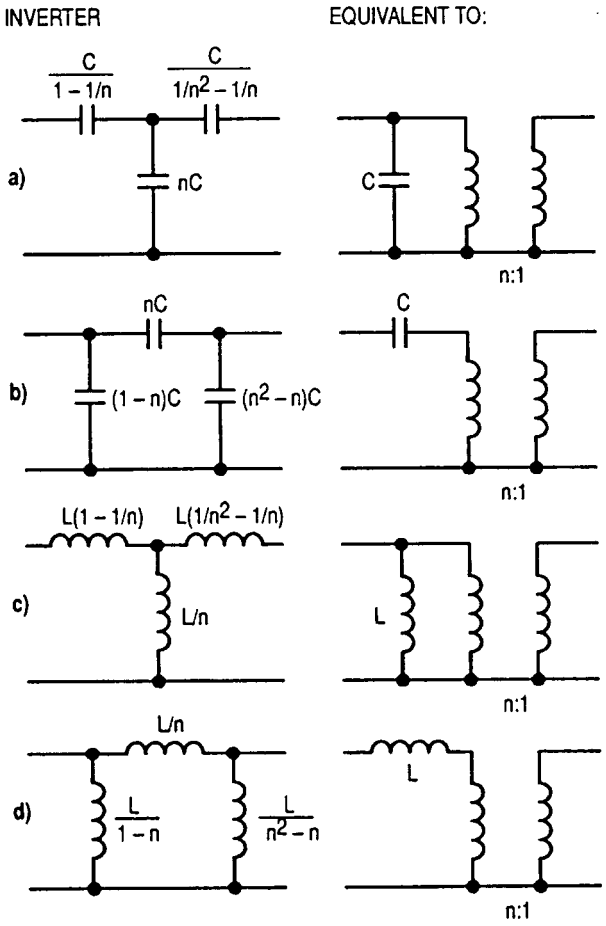


Figure 3-245 Impedance inverters.

where $\omega < \omega_b$. This must be equal to ω/ω_b for 6 dB/octave attenuation.

At the other band edge a, exact compensation can be obtained if

$$Q = \frac{(\omega_b/\omega_a)^2 - 1}{\omega_a/\omega_b - (\omega_b/\omega_a)^2} \tag{3-230}$$

There are several methods available for synthesis of matching networks. They should not be confused with filters because they can be either high-pass, low-pass, or bandpass, and their main purpose is to match one complex impedance to another. In many cases, one of the terminations happens to be 50 Ω and real, but this is just a "lucky" case of the general approach. An excellent paper on this [41] refers to the "real frequency" technique. Another equally important publication is the one by Cuthbert [42], since followed by his book [43]. Two more good references are Vendelin et al. [44] and Potter [45].

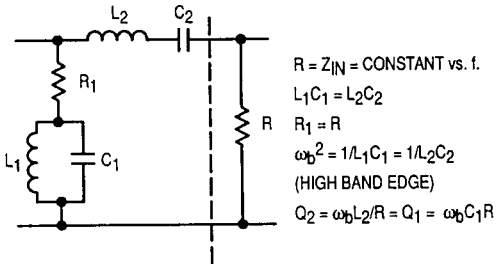


Figure 3-246 Roll-off compensation network.

3-12-3 Example 2: Low-Noise Amplifier Using Distributed Elements

In our previous examples, the amplifiers were built around lumped elements. Since a transmission line length of less than $\lambda/4$ is an inductor and $\frac{3}{4}\lambda$ is a capacitor, the whole amplifier can be built around microstrip technology.

If anyone wonders why we are going to show a low-noise amplifier in the middle of the power section, the reason is that we will select a low-power Siemens transistor, which, being operated in Class A, has an intercept point of +30 dBm and can easily achieve a noise figure of less than 2 dB. Such an amplifier is attractive when feeding an input stage from an antenna that sees many signals and curtails the bandwidth with an input filter of good selectivity that needs a good output termination, which in many cases works against the achievement of a good noise figure because when designing for best noise figure the input impedance of the transistor stage is far from 50 Ω.

First, we will start with the textbook approach and use the Smith diagram to obtain the best noise figure of which the device is capable. Figure 3-247 shows the actual approach to input and output matching.

We start our design by querying the Smith Tool about the minimum noise figure and the gain that comes with it. We mark the point of the lowest noise figure (labeled Γ_{opt}). The small

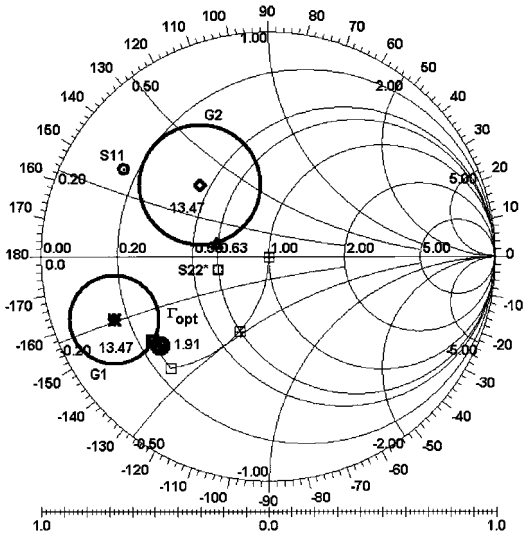


Figure 3-247 Input and output matches for 13.47 dB of gain and NF = 1.91 dB as displayed by the interactive Smith Tool.

noise circle yields the location for Γ_{opt} and tells us the minimum noise figure is going to be 1.91 dB. (By the way, this assumes no further losses, even in soldering the transistor to the PC board.) Then we construct the input gain circle (G1), which touches Γ_{opt} and turns out to represent 13.47 dB of available gain.

The next step is to find a matching network between Γ_{opt} and the 50-Ω source impedance. Immediately there is a conflict: If we look at S_{11}, which is the input reflection coefficient shown in the Smith diagram, it is on the other side of the complex plane and sufficiently far away, indicating that there is going to be quite a difference between noise matching and power-gain matching. We have two options for matching the input: Either we start from the origin and move toward the position of Γ_{opt}, or we start with Γ_{opt} and go to 50 Ω. Depending on the direction in which we do this, the result can be different topologies, such as a high-pass or low-pass filter. The high-pass filter is obtained by starting with Γ_{opt}; the low-pass topology is obtained by starting at the 50-Ω point. The Smith Tool then provides us with the corresponding values of capacitors and inductors after the transformation has been accomplished.

The final step is to do the same type of matching at the output. We start with S_{22}, and either select S_{22}^* (the complex conjugate for highest gain) or S_{22load}. As a rule, the second one is always a lower impedance (in terms of the real part) than the conjugate match. This graphic procedure is somewhat trial and error. As to the topology, a high-pass match should be preferred at the input and a low-pass match at the output. The reason for this is that the high-pass filter will tend to protect the stage from the enormous number of signals present below 1 GHz; using a low-pass filter at the output suppresses unwanted harmonics that the amplifier may generate. Figure 3-248 shows one of the approaches we like best simply because it gives a good input termination and we were trying for a simultaneous match for best noise figure and power gain.

The way we were able to achieve this was to take advantage of the existing feedback in the transistor and select a less-than-optimal output termination but gaining at the input. Taking advantage of the Miller effect, we rotate the phase angle of the feedback with the load to move the optimum noise and gain matching points closer. The danger in this is that the circuit is now highly sensitive to both input and output termination. The output *must* see 50 Ω to maintain this good input match, but since the second transistor stage can now be optimized for best input match, a two-stage combination can be configured to meet this

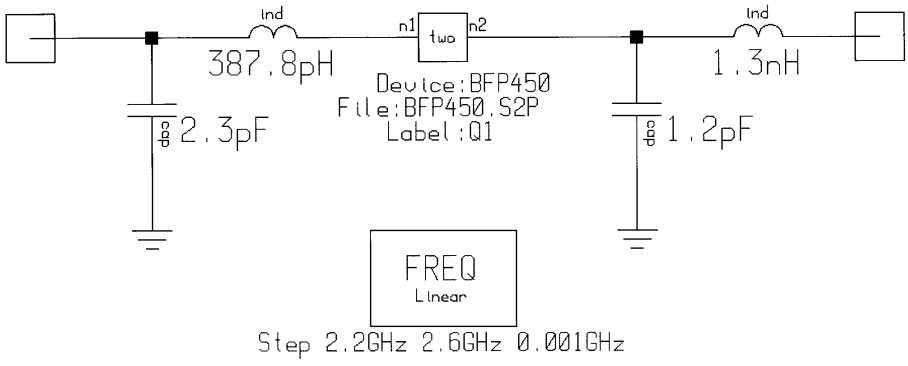


Figure 3-248 A 2.4-GHz amplifier using a BFP450 transistor and lumped-element input and output matching.

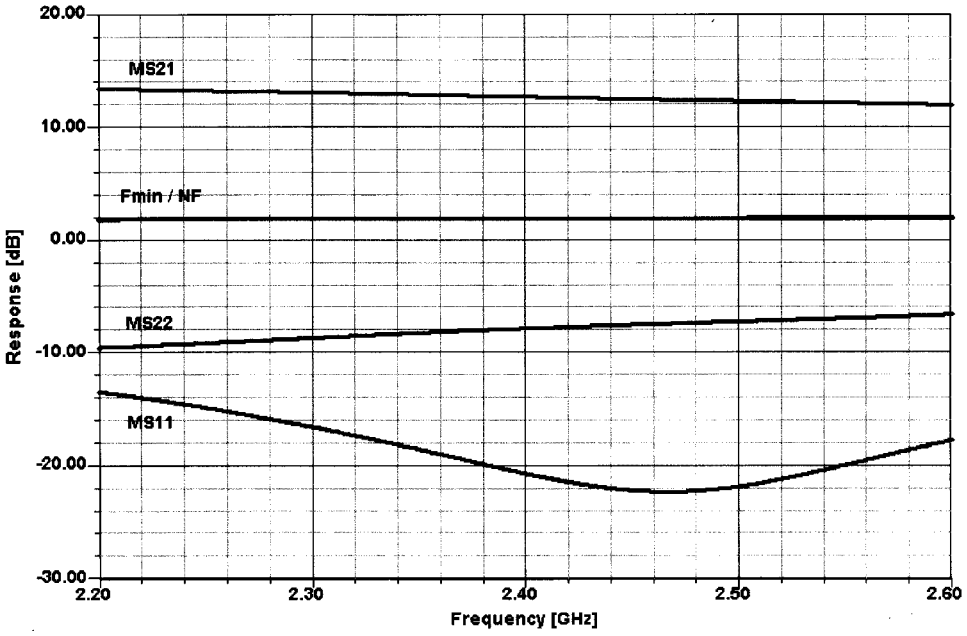


Figure 3-249 Frequency-dependent gain, matching, and noise characteristics of the matched high-dynamic-range BFP450 2.4-GHz amplifier.

requirement. Also, let us not forget that the device is running at 50 mA and as far as bipolar transistors are concerned, this is probably the best amplifier combination we have seen for getting the highest dynamic range and lowest noise figure. The resulting gain, matching, and noise figure characteristics are shown in Figure 3-249.

Now we turn to using distributed elements and at the same time use a shunt feedback at the output; to be specific, we will load it with 150 Ω. The reason for this is to be able to obtain a good compromise at the input that will depend less on the output and at the same

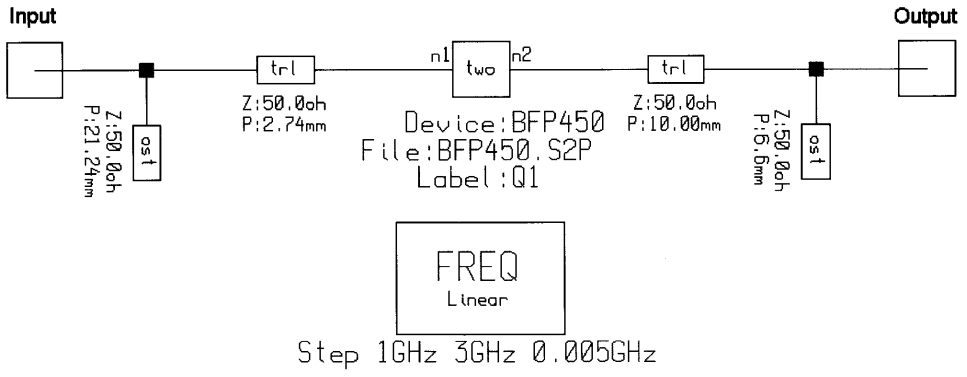


Figure 3-250 Distributed-element input and output matching for the 2.4-GHz BFP450 amplifier.

time improve the S_{22} matching at the output. The procedure for doing the matching with a transmission line offers only certain degrees of freedom, while doing the matching with lumped elements allows the choice of series or parallel inductors or capacitors. In the case of a transmission line, we have open stubs, shorted stubs, and transmission lines. This may result in somewhat more complex matching schemes. On the positive side, they can easily be trimmed, are reproducible, and certainly cost less than exotic capacitors, which need to have small tolerances and to be able to handle high RF currents.

The resulting input and output matching is shown in Figure 3-250, and its frequency response in Figure 3-251. We were able to maintain a good input and output match with an insignificant deterioration of the noise figure.

Since this is a high-dynamic-range amplifier, we also show its predicted two-tone response in Figure 3-252. The manufacturer specifies an $IP_{3,in}$ of 29 dB for a bias point of $V_{CE} = 3$ V and $I_C = 50$ mA with $Z_S = Z_{S,opt}$ and $Z_L = Z_{L,opt}$ at 1.8 GHz. Significantly, however, the manufacturer-supplied nonlinear BFP450 library we used exhibits considerably different characteristics at $V_{CE} = 4$ V and $I_C = 50$ mA than those reflected in the corresponding S parameters supplied by the same manufacturer. In our circuit, the S parameters resulted in about 13.5 dB of gain (based on a compromise between matching for optimum noise and maximum gain) versus about 9.5 dB of gain for the nonlinear library (match optimized for maximum gain). The difference between the 2.4-GHz F_{min} values returned for the two modeling approaches is similarly striking: 1.90 dB for the bare S parameter set and 7.76 dB for the bare nonlinear library biased to $V_{CE} = 4$ V and $I_C = 50$ mA. Clearly, the nonlinear modeling of the BFP450 needs work!

Having fully evaluated the amplifier at 2.4 GHz as a working example, we note that this frequency is used by radio amateurs for some applications. We will therefore supply

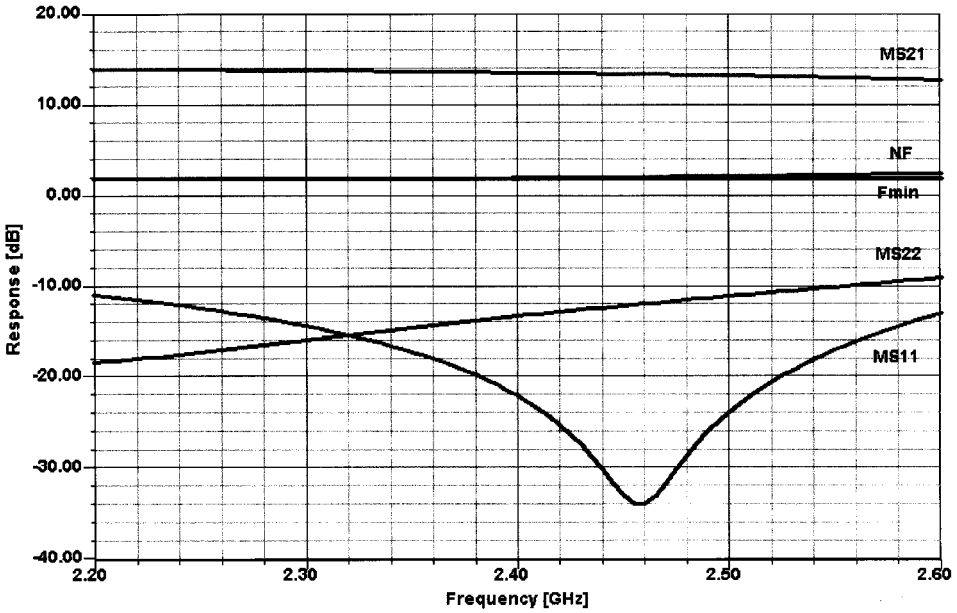


Figure 3-251 Predicted frequency-dependent gain, matching, and noise performance for the BFP450 amplifier with distributed-element matching.

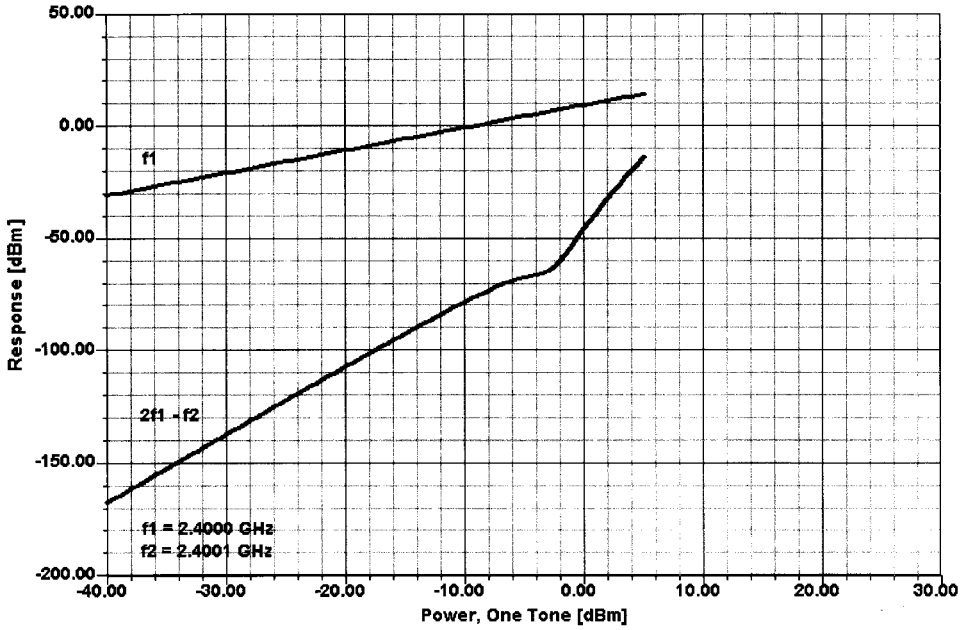


Figure 3-252 Predicted two-tone response of the BFP450 amplifier with the circuit adjusted for maximum gain (9.5 dB): $IP_{3,in} \approx 28.5$ dBm and $IP_{3,out} \approx 38$ dBm. That this version's gain is significantly lower than that of the *S*-parameter-based equivalent illustrates the need for better accuracy in parameter extraction.

examples for 432 and 1296 MHz, popular spot frequencies in the amateur 70- and 23-cm bands, respectively. Figures 3-253 and 3-254 show the 432-MHz version's schematic and frequency-dependent responses. Figures 3-255 and 3-256 show the schematic and frequency-dependent responses for the 1296-MHz version.

Using a bipolar transistor that operates at much less current, the noise figure can be improved at the expense of the dynamic range/intercept point. It is possible to reduce the noise figure by about 1 dB or so, but the reduction in intercept point will be much more dramatic. Another choice is the use of GaAsFETs for this application, but their matching in comparable circuits has been described as a nightmare simply because the devices have too much gain and very high input impedances (essentially 0.2 pF in series with 3 Ω) combined with reasonable output impedances in the vicinity of 240 Ω. Here is a wide field in which engineers can experiment and examine price/performance and performance/real estate issues. GaAsFETs also tend to require more than 30–40 mA to be alive, and even the low operating voltage of 4–5 V makes the device consume a lot of dc power.

3-12-4 Example 3: 1-W Amplifier Using the CLY15

The following is an application showing a fairly wideband amplifier with +30 dBm (1 W) output. We need to point out that in accordance with the various digital modulation modes used with this amplifier, this is not continuous operation, but a much less than 1:1 duty cycle operation. The device would probably not survive operating at full output under CW conditions. This is subject to a discussion with the manufacturer, Siemens, who has not

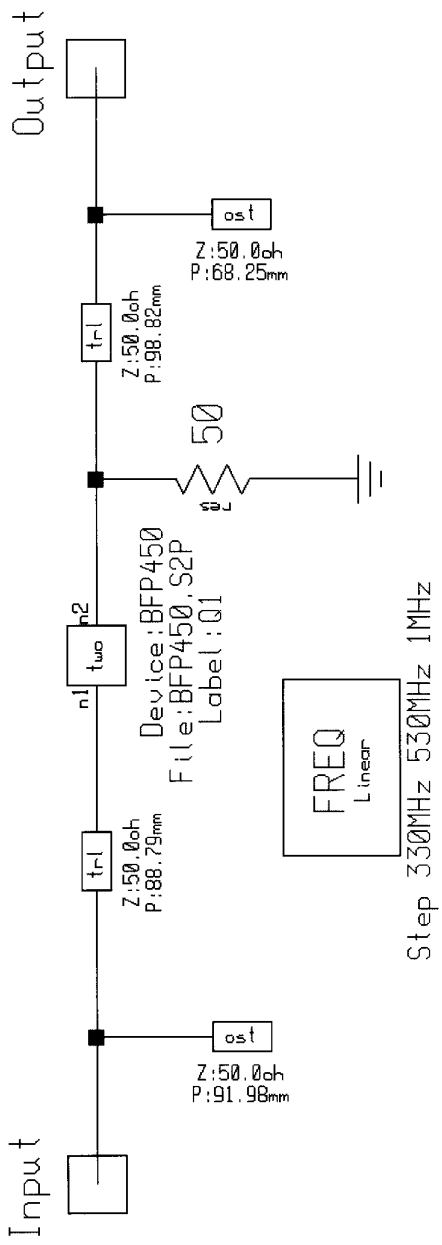


Figure 3-253 Schematic of BPF450 amplifier for 432 MHz. Because f_T = 25 GHz for the BFP450, at this lower frequency we must shunt the device output with a 50-Ω resistor to maintain stability.

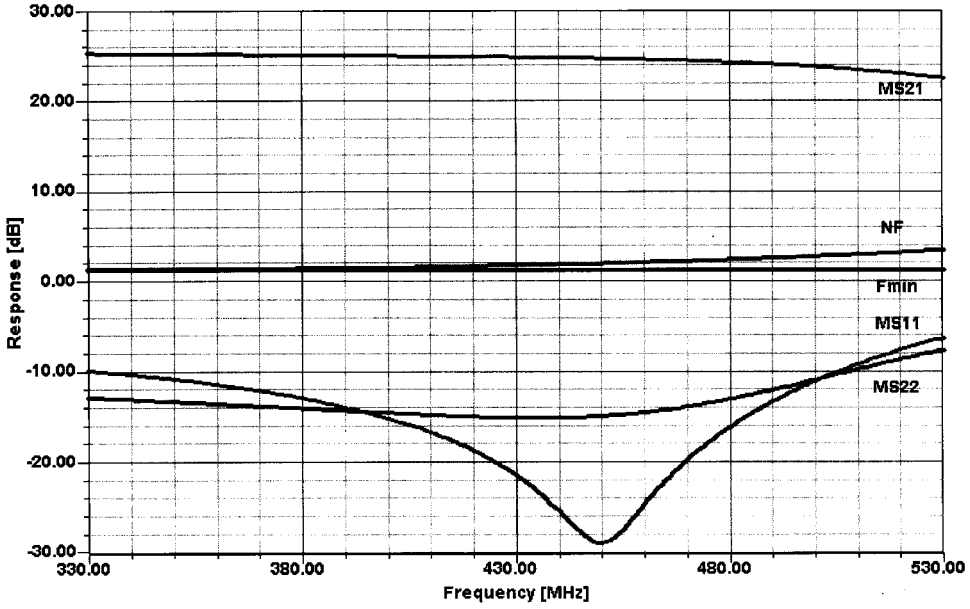


Figure 3-254 Frequency-dependent gain, matching, and noise performance for the BPF450 amplifier at 432 MHz. As with the circuit in Figure 3-253, the input and output network values have been chosen to achieve a low NF and good input matching at the expense of the output match.

released sufficient data for us to make a judgment as to the CLY15's continuous power capability. Since there is no useful large-signal model available at the time of this writing, we will design this Class A amplifier using large dc bias *S* parameters. In the actual design there will be some cut-and-try, but we find that in this type of application this is still the best way of handling the design.

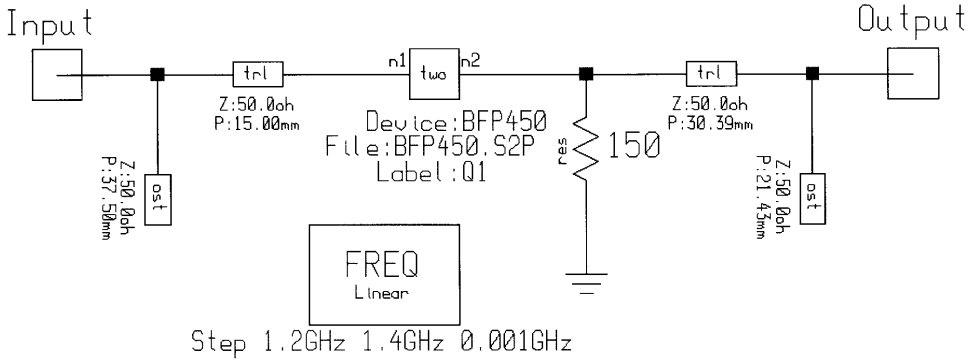


Figure 3-255 Schematic of the BPF450 amplifier for 1296 MHz. A loading resistor is still necessary to ensure stability, but because the transistor exhibits less gain at this frequency than at 432 MHz, we can increase the loading resistor to 150 Ω.

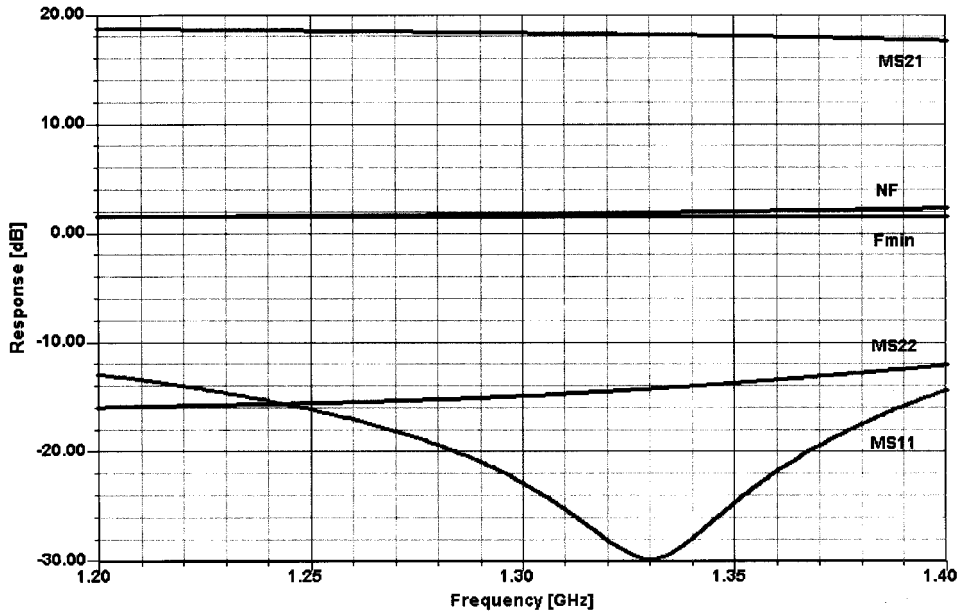


Figure 3-256 Frequency-dependent gain, matching, and noise responses of the BFP450 amplifier at 1296 MHz. As with the two previous versions, the input and output network values have been chosen to achieve a low NF and good input matching at the expense of the output match.

The CLY15 (gate width 16 mm) is the largest of a series of Siemens GaAsFETs; it was already mentioned in Chapter 2, where its datasheet is reproduced. Its device family also includes the CLY2 (gate width 2 mm), CLY5 (4 mm), and CLY10 (8 mm). It is possible that bigger devices will be offered in the future. One of the most critical parameters will be the saturation voltage; from the datasheet of the CLY15, we can see that at 1.5 A the saturation voltage is approximately 1 V. This drop of 1 V includes losses that result from other parasitic elements; the dc *I–V* curve (Figure 3-257a) indicates a saturation voltage of 0.75. Since we aim for 1-W output power, the resistive portion of the load will be $(5 - 1)^2 / (1 \times 2) = 8\ \Omega$. This assumes a supply voltage of 5 V.

The next step is to design the input and output matching network. Previewing the *S* parameters, it will be noticed that they are essentially inductive, which means that the input match, using lumped elements, will result in a combination of capacitors at both the input and the output. Figures 3-258 and 3-259 show the ways in which the input and output matches were done on the Smith diagram resulting in lumped matching elements. Inspecting the lumped input matching, we find the following. We start with S_{11}, which is in the lower part of the Smith diagram, and determine its conjugate match, S_{11}^*. We already pointed out that this device is inductive, which means that matching it from the conjugate point will require two capacitors. The first one is a series capacitor, as identified in the Smith diagram, and followed by a shunt capacitor to transform the impedance to 50 Ω. A close inspection of the Smith diagram shows that the resulting point is not quite on S_{11}; it is an iterative process to exactly find the crossover point, yet the input match of better than 20 dB is more than acceptable.

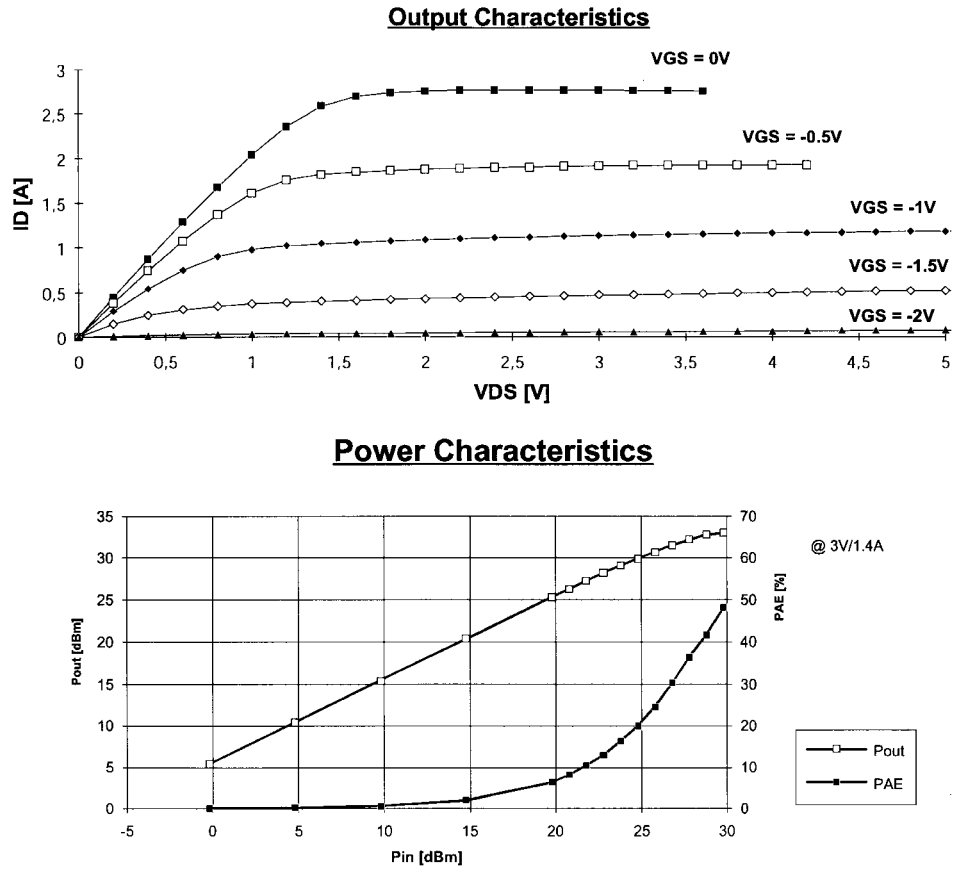


Figure 3-257 Output and power characteristics for the CLY15 medium-power GaAsFET.

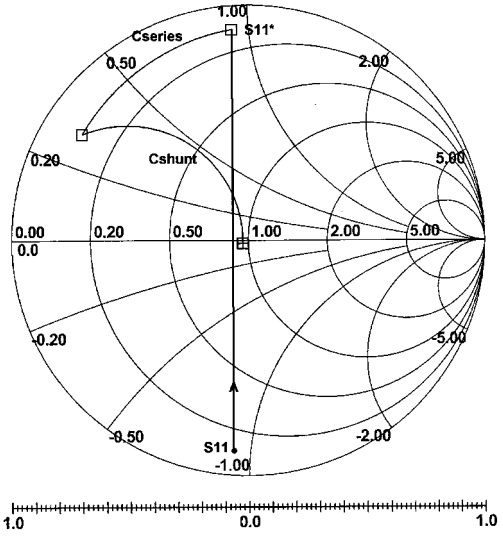


Figure 3-258 Input match for S_{11}^{*} using lumped elements.

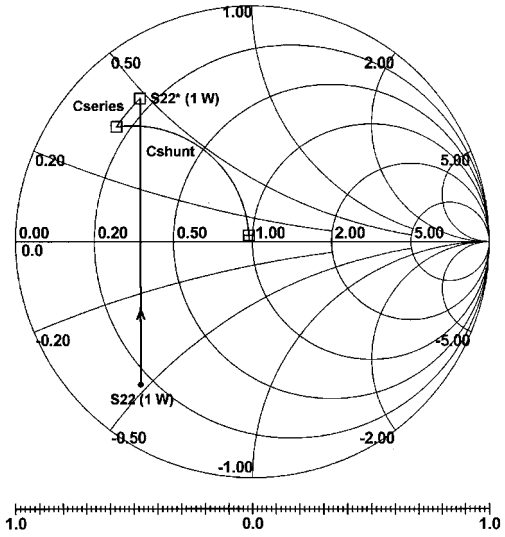


Figure 3-259 Output match for S_{22}^* (1 W) using ideal lumped elements.

As far as the output is concerned, we use the same technique by moving from S_{22} to S_{22}^*. S_{22} has been determined from the required output power; as we showed above, the real portion is 8 Ω. Using the lumped elements, the matching technique is identical; we first need to use a series capacitor and then a shunt capacitor for the transformation. The arc for the series capacitor this time is much longer, resulting in a series capacitance of 11.2 pF instead of the 2.1-pF capacitance used at the input. The actual schematic is shown Figure 3-260. The drawback of lumped elements is that it is not practical to buy components that can handle the heavy RF currents that flow in the input and output networks. Taking the values obtained

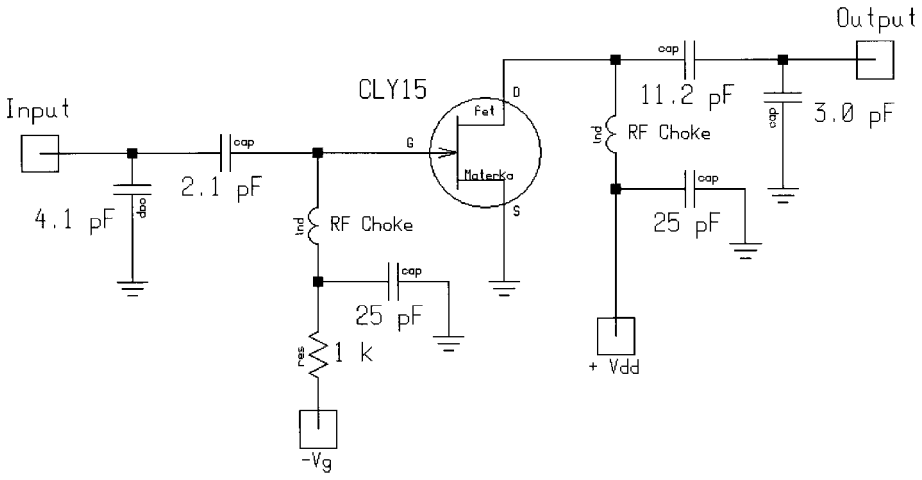


Figure 3-260 The complete amplifier using the CLY15 medium-power GaAsFET, with lumped-element input and output matching.

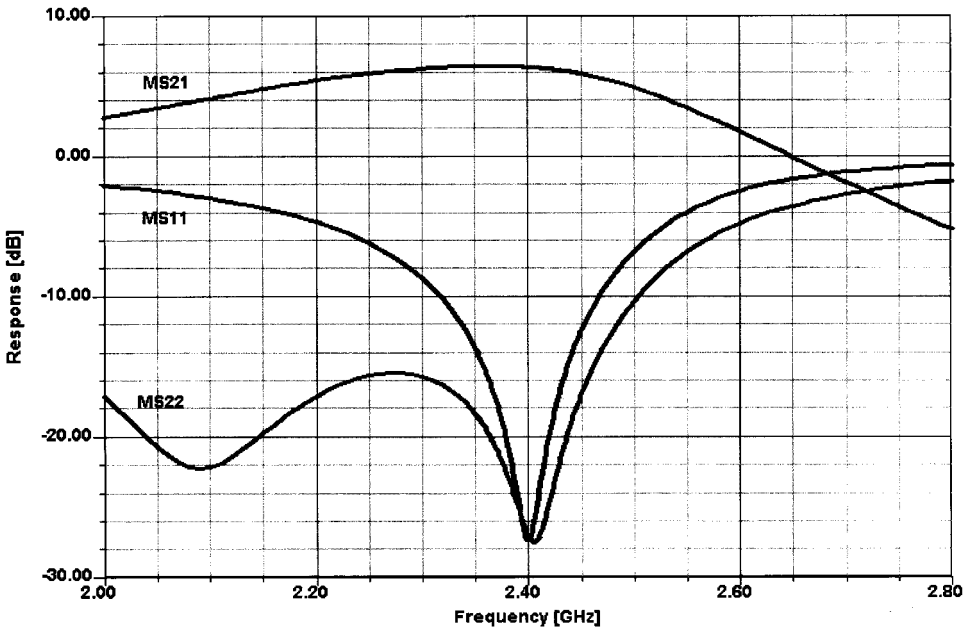


Figure 3-261 Frequency-dependent gain and matching responses of the CLY15 amplifier using ideal lumped elements.

from the lumped input and output matching approach, Figure 3-261 shows the resulting frequency response.

The next step is to provide the same matching condition using distributed elements, which of course have the advantage that they can be printed on the PC board. Instead of using ideal elements, we used a loss tangent of 2E–3, or a Q of 500. The manufacturer of the R4000 material (Rogers) provided this specification. Now we start with the same approach by moving from S_{11} to S_{11}^*. From there, we use an open stub with a transmission-line impedance of 50 Ω, followed by a transmission line to the point in the left lower corner, and, finally, a shorted transmission-line stub allows the 50-Ω matching. For this amplifier, we are not looking for noise matching but for input and output power matching. Figure 3-262 shows this match on the Smith Chart.

At the output, the procedure is again similar: We move from the S_{11} for 1 W to a conjugate value, then use an open stub that acts like a capacitor until we cross the real axis; finally, a λ/4 transmission line is used to match from approximately 75 Ω to 50 Ω. Figure 3-263 shows this match on the Smith Chart. This concludes the distributed matching design for the CLY15 amplifier.

Finally, the complete circuit using distributed elements is shown in Figure 3-264. From a dc point of view, we took advantage of the incoming shorted stub by using a transmission line of the appropriate length and a bypass capacitor to ground. This allows us to feed the gate with the required negative voltage without interfering with the frequency response of the amplifier. When evaluating the total frequency response (Figure 3-265), we noticed that this circuit seems to have a higher Q than the lumped equivalent. Again, to obtain a wider bandwidth more transformation stages will be needed. We leave this task to the interested reader.

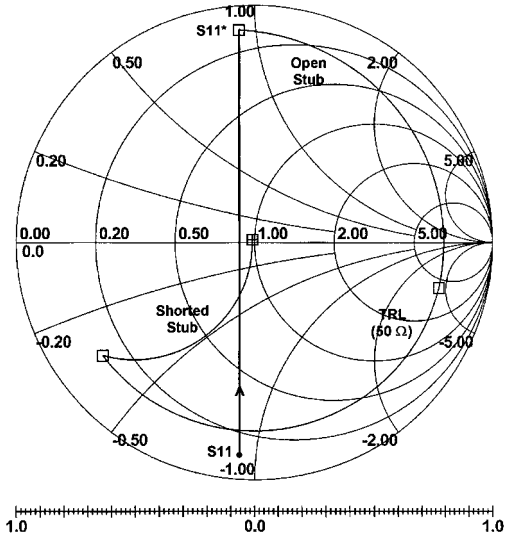


Figure 3-262 Input match for S_{11}^* (1 W) using distributed elements.

The 1-dB bandwidth of the amplifier using distributed, lossy elements is 120 MHz. If a wider bandwidth is required, matching networks with more elements are required. Because of the gain–bandwidth product, the peak gain for such a wider-bandwidth arrangement will be less than this narrowband example.

Having used a lumped to distributed transformation, we next present a power amplifier that uses coaxial lines, baluns, and other transformation steps capable of handling the increased power requirement.

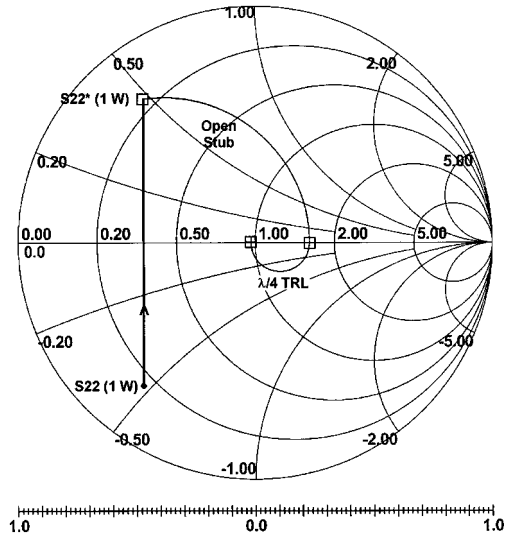


Figure 3-263 Output match for S_{22}^* (1 W) using distributed elements.

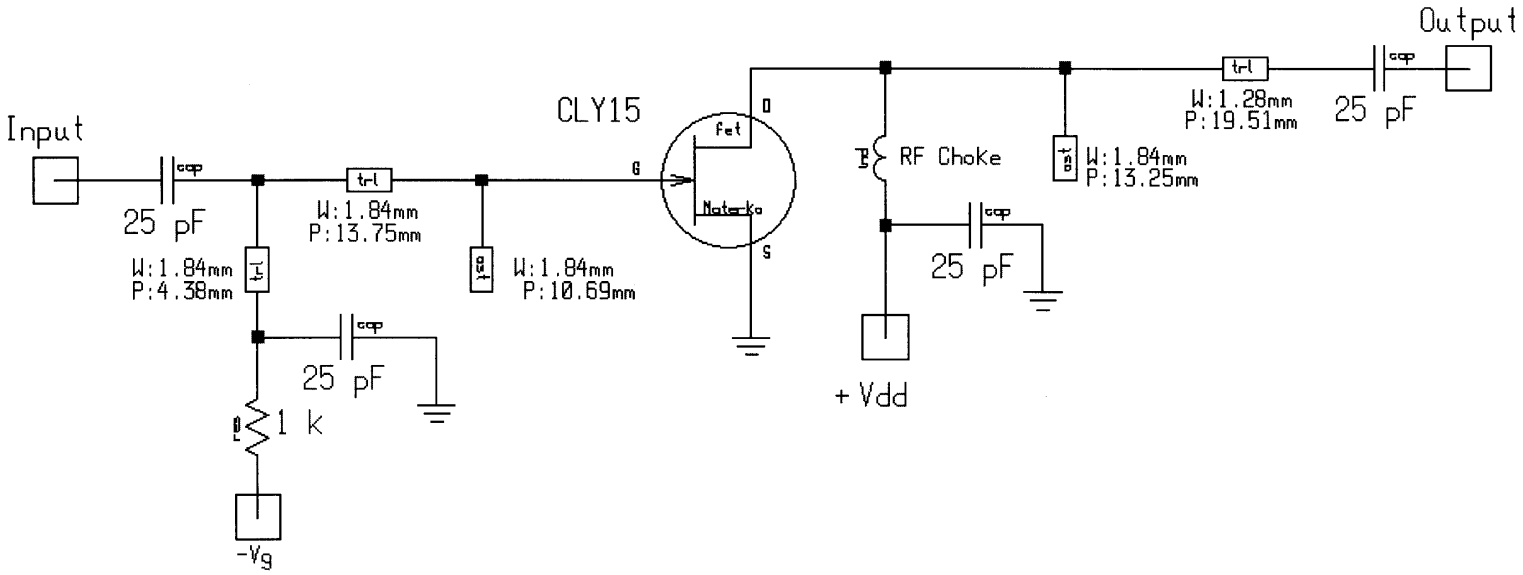


Figure 3-264 The complete amplifier using the CLY15 medium-power GaAsFET, with distributed-element input and output matching.

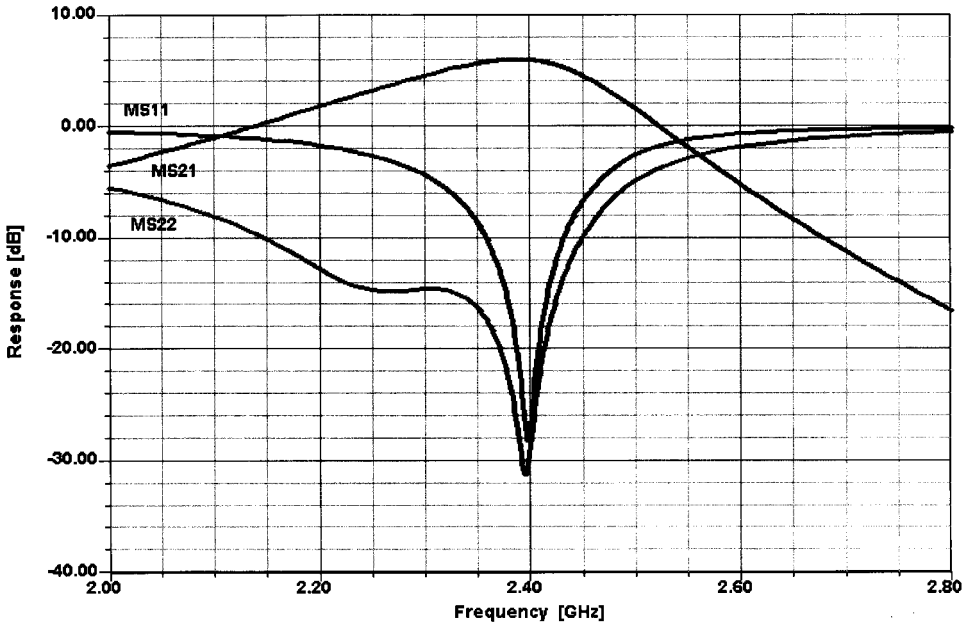


Figure 3-265 Frequency-dependent gain and matching using distributed, lossy elements.

3-12-5 Example 4: 90-W Push–Pull BJT Amplifier at 430 MHz

The last complete application we want to examine is a 90-W output, push–pull amplifier using two TRW transistors (that are no longer made—an all-too-frequent occurrence after a design is complete, resulting in a scramble for replacements). Figure 3-266 shows the circuit diagram of this UHF transistor amplifier, and Figure 3-267 shows its mechanical design. The amplifier requires 28 Vdc and will draw approximately 6 A, requiring 15 W of drive at 438 MHz. This amplifier operates in Class AB, which means it is not limited to amplifying FM and other angle-modulated signals, and can accommodate even single-sideband (SSB) operation.

The drive power coming from the source connected to the input N connector (15 W) is fed through a short piece of coaxial cable to the PC board sitting on top of the heat sink. To change from single ended to push–pull, a 50-Ω coaxial cable is coiled to provide sufficient electrical length to act as a balun that provides two symmetrical paths through the amplifier. The RC section of 10 Ω in parallel with a 220-pF capacitor stabilizes the amplifier at lower frequencies, while at higher frequencies the 10-Ω resistor has no contribution. There are two independent 4:1 transformers, one for each section; they are built using 25-Ω coaxial cable with a set of two binocular ferrite cores to force the transformation. The dc supply is also fed to the transistors at this point. The stabilization circuit consists of a temperature monitor using the *pn* junction of a transistor, side-mounted on the heat sink, and a combination of operational amplifiers and adjustments. This guarantees proper, temperature-independent biasing. The 4:1 input baluns result in 12.5 Ω impedance, or 3.125 Ω per transistor. Since the input impedance at this drive level is somewhere in the vicinity of $[(1 \ldots 2) + j(3 \ldots 5)]$ Ω; the final matching is accomplished by an *RC* section in which the inductance is actually printed

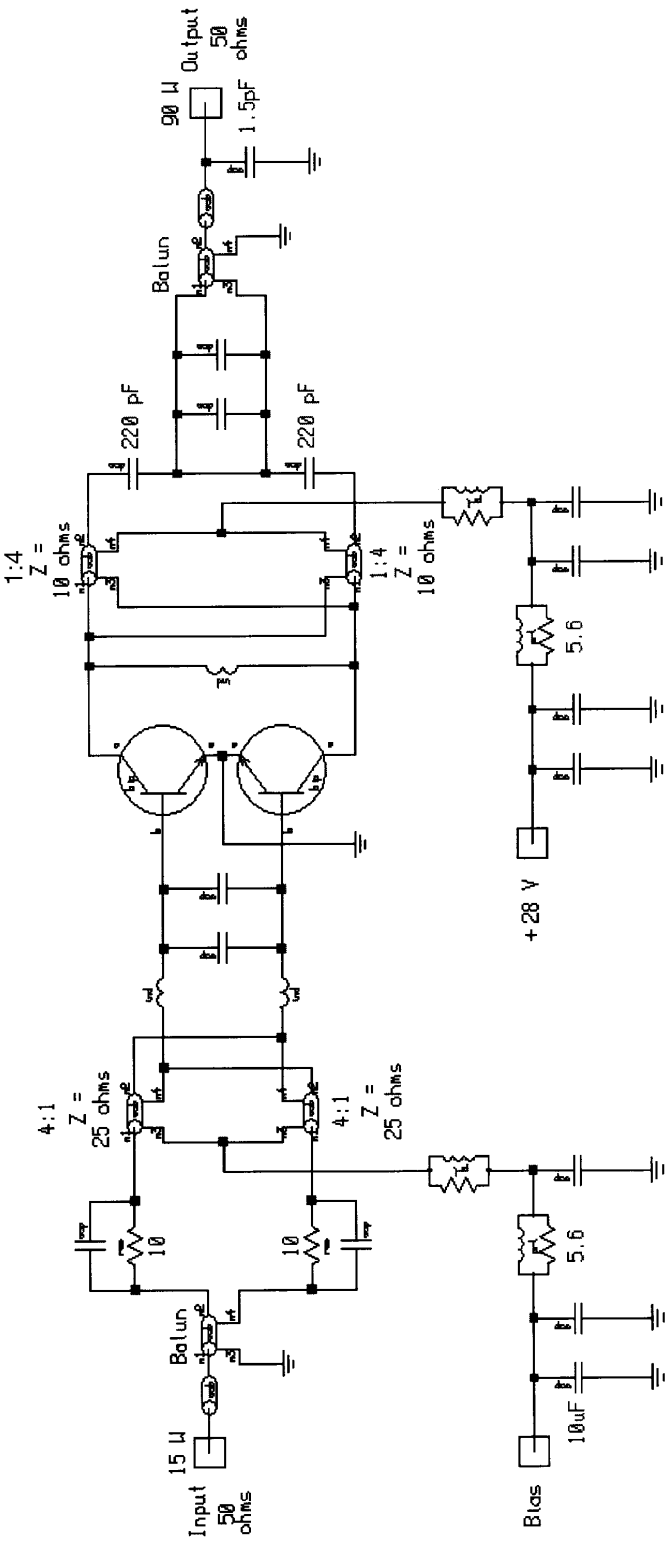


Figure 3-266 Circuit diagram of the 90-W, push–pull, 438-MHz linear amplifier.

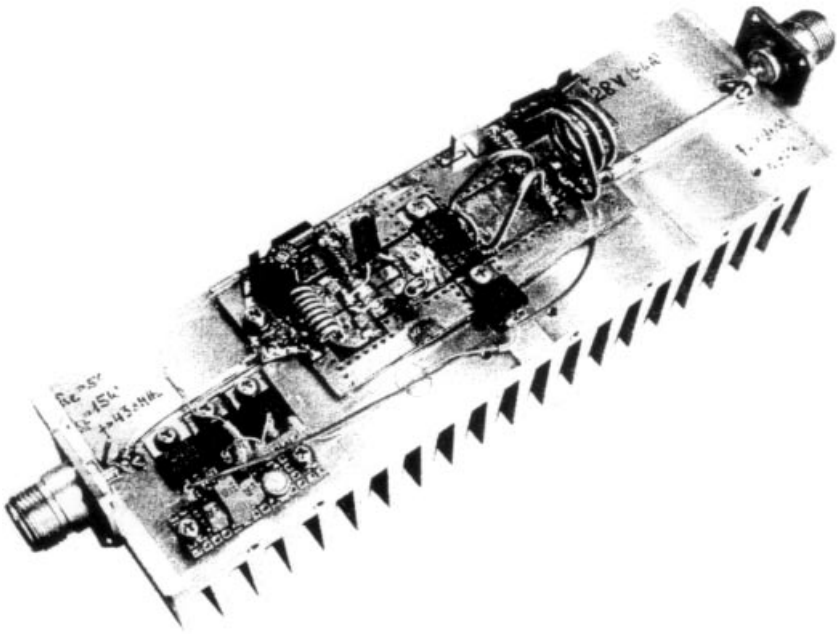


Figure 3-267 Mechanical assembly of the 90-W amplifier shown in Figure 3-266.

and the shunt capacitor across is soldered from base to base as close to the transistor bodies as possible. A ceramic-dielectric variable capacitor in parallel with this allows fine tuning.

The output consists of an inductor from collector to collector that tunes out the capacitive reactance and makes the transistor output impedance real. We then follow the same pattern by taking a 1:4 transformer, but this time built using 10-Ω rigid line instead of 25-Ω rigid line. Together with the output capacitance, prior to the output balun, this is a medium-bandwidth matching arrangement that transforms up to 50 Ω. The collector impedance per collector had to be $(28 - 2)^2/2 \times 90$, resulting in roughly 3.8 Ω—not that different from the input match at the balun point. This way, the matching networks at the input and output can be identical. At the output of the matching network, another coiled 50-Ω coaxial cable serves as balun; because it must operate at the 90-W level, larger-diameter line was used than in the input balun. Finally, another section of rigid line is used to connect the output balun on the PC board to the output N connector. The output capacitor again helps in the transformation.

Both the base and collector voltage supplies are heavily filtered using ferrite-bead-based RF chokes and heavy wire inductors wound on low-impedance resistors. A fairly large number of these amplifiers were actually built and are still in use.

3-12-6 Quasiparallel Transistors for Improved Linearity

Some applications require really low intermodulation distortion. One of the authors (Rohde), during his tenure at RCA Government Systems Division, and his team developed a parallel-device topology that is somewhat similar to the conventional parallel approach but shows vastly increased linearity. In creating higher-power devices, manufacturers parallel several transistors in one package, connecting them in a clever way so the overall drive, dissipation, and output power are equally divided between all the devices. Special nickel–chromium

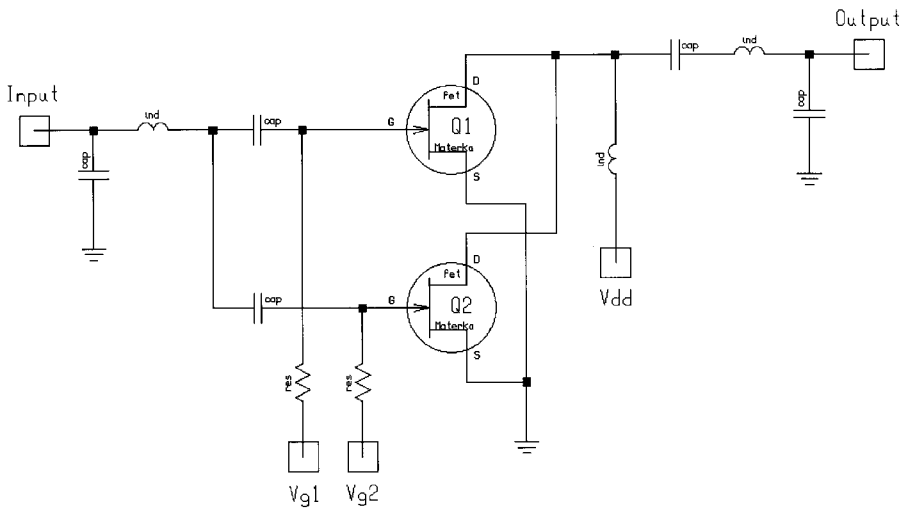


Figure 3-268 The quasiparallel amplifier topology looks like a standard parallel amplifier. Biasing the devices to different operating points nets an improvement in IMD performance.

ballasting resistors in the many fingers of the emitter connections of the multiple internal devices equalize any differences among the transistors while increasing the saturation voltage by only a small amount. In our high-linearity design, we used two transistors in parallel, one being biased somewhere between Class A and B, and the second one somewhere between Class B and C (Figure 3-268). The exact bias voltages and operating points had to be found through experiment. In practice, Q1, the "initially on" transistor, is driven to a point close to that at which its IMD products would increase. Q2 is biased such that at this same drive level it adds to the output power.

We frequently referred to this configuration as "turbocharged." The IMD products at full output power were down 40–45 dB, compared to 26–30 dB for a single transistor. While such an arrangement shows superior IMD performance, because of the somewhat abrupt kick-in of the second transistor, the transconductance response is semismooth, a characteristic that might make alert readers think that IMD will occur at this point. On the other hand, the transconductance for Q1, in effect, saturates or stays constant, so when the transconductance responses of both transistors are overlaid, the result is a smooth response overall.

This technique has been validated up to 100 W and up to at least 50 MHz. Its drawbacks include the need for two transistors with their associated costs, additional real estate, and higher input and output capacitances. Nonetheless, considering that MOS transistors capable of operation to 1 GHz were unavailable when this technique was developed, we believe that it should be revisited.

The Class D and higher operating modes require a large number of additional components and will make the efficiency higher (selective in frequency), but not necessarily reduce the IMD products. A further point of consideration is that it is incorrect to assume that reducing the input level to a given Class A will necessarily reduce its IMD products. As with tubes, the closer one gets to the knee, even in Class A operation, the worse the IMD. (This phenomenon was particularly troublesome in systems using 1-kW amplifiers up to 150 MHz;

underdrive with the intent of getting better IMD performance produced just the opposite result.) If a wide power range is considered, we recommend using a bias scheme that includes sufficient intelligence to find the best operating point.

Another hot amplifier issue is that of IMD as a function of load impedance. If an amplifier is used at output power considerably below its design value, its load impedance will be too low and, depending on the circuit configuration, the result may be additional IMD.

3-12-7 Distribution Amplifiers

It is not uncommon that a signal from a particular source like an antenna has to be distributed to several places. A condominium is a typical situation in which energy from a main signal source must be split for distribution to a larger number of users—there must be enough voltage or power to produce a noise-free picture. An amplifer capable of doing this is a *distribution* amplifier, not to be confused with a *distributed* (or traveling-wave) amplifier, such as that shown earlier in this chapter in Figure 3-6. An interesting means of achieving this is to use multiple Wilkinson couplers following an amplifier. Figure 3-269 shows an arrangement that makes heavy use of multiple power dividers to provide multiple outputs. Theoretically, the reverse is also possible: If we want to collect energy from a number of sources and combine them, as on a cable, we can use Wilkinson couplers at the *input* of a gain block. Note, however, that depending on the level and number of the combined signals, this may put enormous strain on the linearity of the amplifier. What we want to avoid is a lot of distortion, which, in the case of television, will show up as color confetti in the noise and signal.

3-12-8 Stability Analysis of a Power Amplifier

One of the hidden pitfalls in narrowband circuit design is that a circuit may be unstable outside the frequency range of interest. Circuits containing structures built from distributed elements, which exhibit multiple resonances, can be particularly susceptible to such instabilities. In this example, we analyze the stability of a simple FET power amplifier designed using small-signal concepts and the interactive Smith Tool utility in Ansoft's Serenade

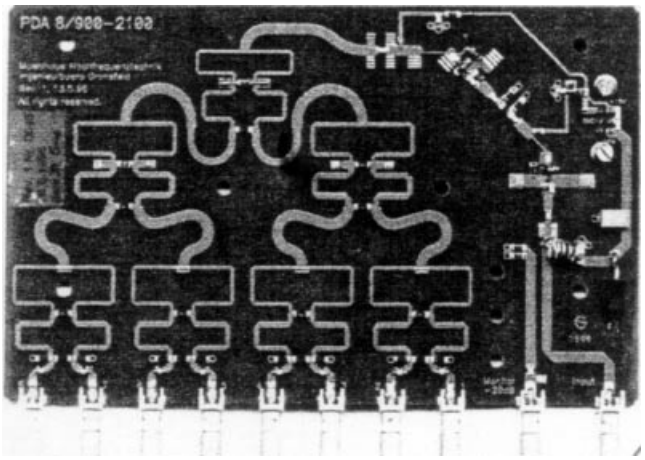


Figure 3-269 Distribution amplifier showing the use of multiple Wilkinson power dividers.

Design Environment. The evaluation shows how K-factor analysis alone is not sufficient to check for stability of a design.

Small-Signal ac Analysis. We begin by taking the Figure 3-270 amplifier, which was designed for about 16 dB of gain at 2.2 GHz. Due to the very low, capacitive input impedance of the FET and the proximity of the source-plane gain circles and stability circles to the edge of the Smith Chart, the matching network attempts to transform 50 Ω to an impedance of approximately $(2.5 + j15)$ Ω. The impedance lies close to the stability circles at 2 GHz but remains in a stable region. The bias point selected (–3.5 V) is chosen for Class AB operation. Higher linear gain can be achieved by reducing the gate voltage to –2.0 V; at this bias point, the amplifier gain will increase to about 19 dB.

First, we will do a small-signal ac analysis and examine the S parameters and K and B_1 factors [see Eqs. (3-154) and (3-157), respectively] from 1 GHz to 6 GHz—well beyond the bandwidth of the design. Figure 3-271 shows the circuit's frequency-dependent gain and matching characteristics.

If we examine B_1 and the stability factor K, we see that the amplifier is not unconditionally stable in the 2.2-GHz passband as shown in Figure 3-272. K dips slightly below 1, although B_1 remains above zero. In practice, a slight change in design would be required to bring K above 1 and remove this potential instability. However, for our example, this is of no consequence, as we shall see below.

Nyquist Stability Analysis. Now we will utilize a Nyquist stability analysis to check the stability of the amplifier. To fully realize the advantages of a Nyquist stability analysis, we must sweep frequency over a very wide range. We want to be able to pick up instabilities at *any* frequency, not just near the design frequency. In this example we sweep from 1 kHz to 20 GHz. A start frequency of 1 kHz is chosen so we can determine the starting point of the Nyquist plot and we ensure that any low-frequency resonances are picked up. The stop frequency is chosen past the f_{max} of the transistor (where we know it cannot oscillate).

The analysis results are shown in polar plot form in Figure 3-273 and in magnitude–angle form in Figure 3-274. The polar graph shows the real and imaginary parts of the system determinant as frequency is swept. The magnitude–angle graph shows the magnitude and cumulative angle (no cut at ±180°) of the system determinant as frequency is swept. The key to interpreting these graphs lies in the following statement drawn from the Nyquist criterion:

The circuit will be unstable if the Nyquist plot encircles the origin in a clockwise direction.

The term "unstable" refers to the existence of natural frequencies lying in the right-half of the complex frequency plane (RHP). These natural frequencies on the RHP will cause oscillations. To encircle the origin in a clockwise direction means that the Nyquist plot will travel in a clockwise path, cross the negative real axis, and continue to completely surround the origin. If the path appears to encircle the origin but then "unravels" itself in a counter-clockwise direction, the origin is not actually encircled. Figure 3-273 shows that the origin is not encircled and therefore this circuit is stable.

The magnitude–angle plot can assist in determining if a clockwise encirclement has actually been made with the help of the following:

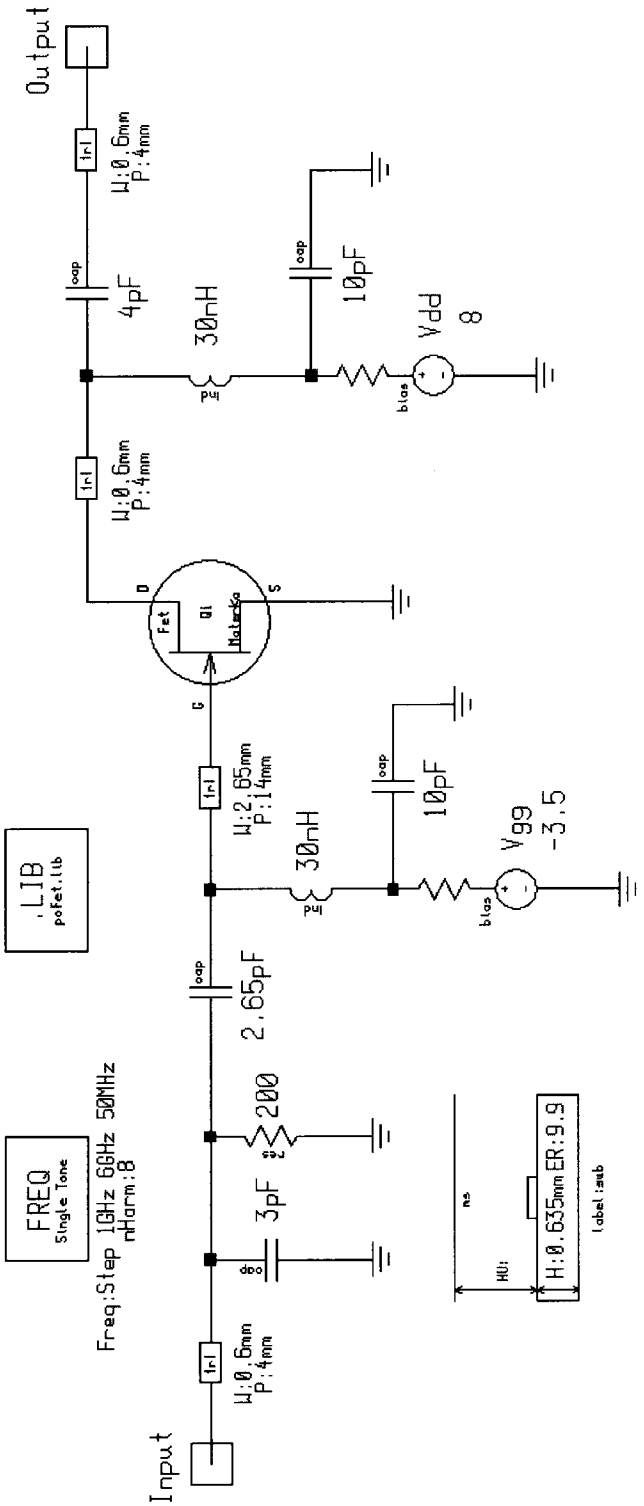


Figure 3-270 A 2.2-GHz amplifier design for stability analysis.

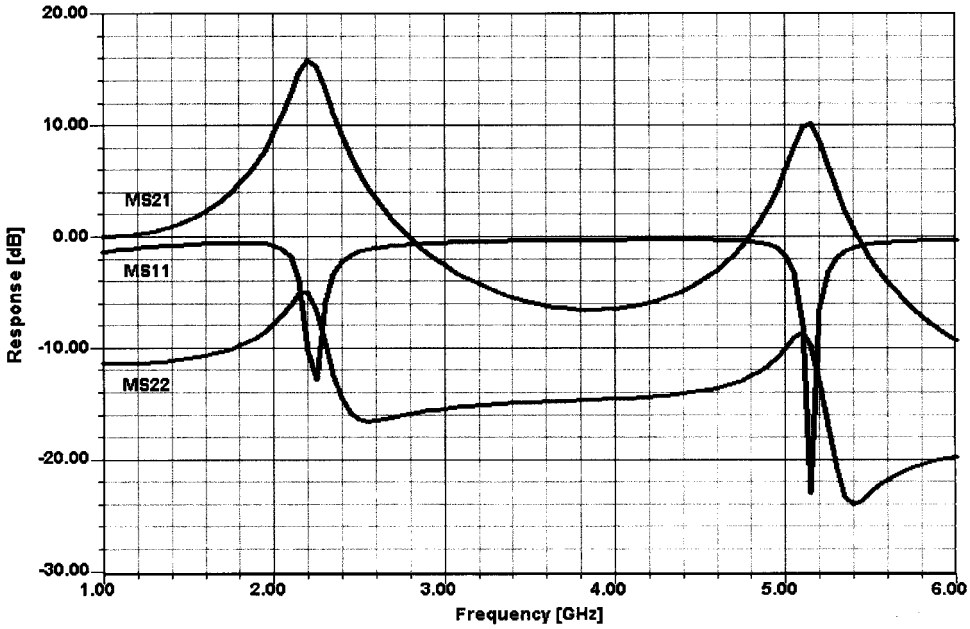


Figure 3-271 Frequency-dependent gain and matching characteristics of the amplifier. Although the circuit is designed for 2.2-GHz operation, the periodic responses of its distributed-element-based input and output matching networks result in a second gain peak near 5.1 GHz. As we shall see, however, instabilities may result at frequencies that cannot be intuited from this graph.

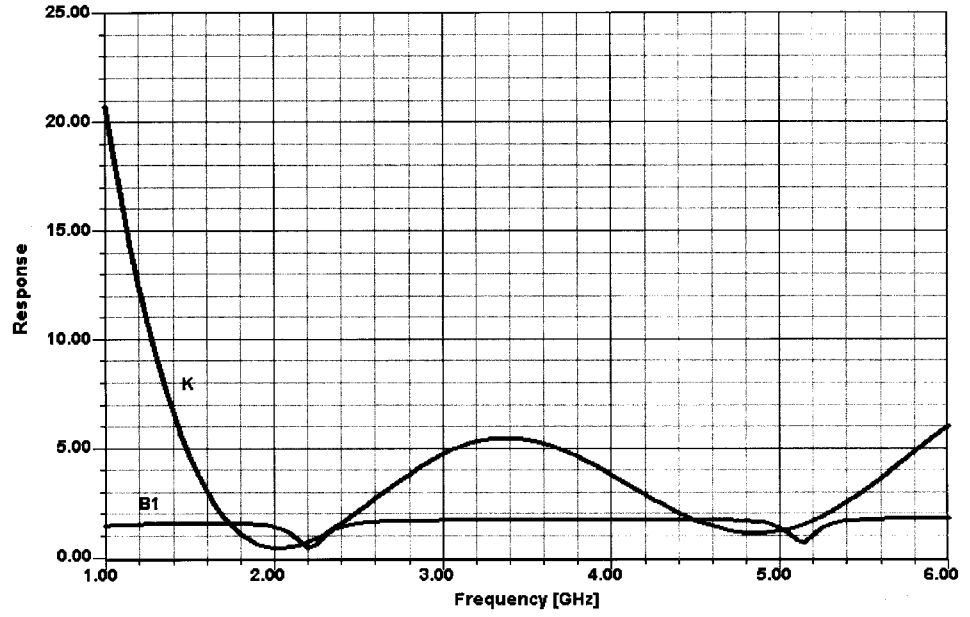


Figure 3-272 K and B_1 for the 2.2-GHz amplifier.

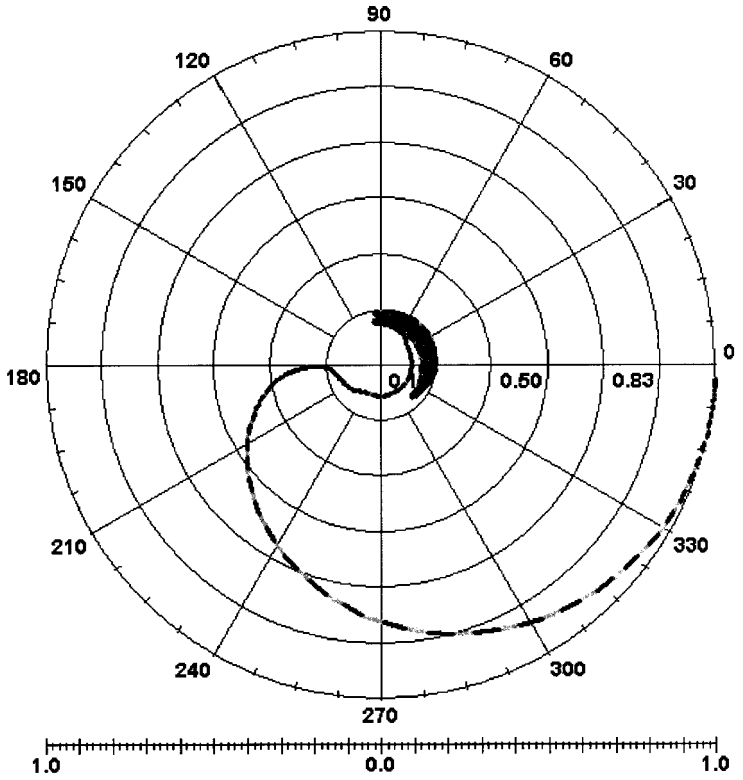


Figure 3-273 Polar Nyquist plot for the 2.2-GHz amplifier covering 1 kHz to 20 GHz.

- To travel in a clockwise direction on a polar chart means to decrease phase angle. On a magnitude–angle plot, the phase will become more negative as frequency is increased (although it may not be monotonic).
- If the path crosses the negative real axis on the polar chart, it will cross –180° on the magnitude–angle plot.
- Encirclement will occur if the phase angle continues to decrease to, say, –360°. If "unraveling" occurs, the phase angle will not stay below –360°. Figure 3-274 clearly shows that the phase angle decreases to about –180° but does not decrease any more. Rather, the angle returns to a positive value and oscillates between about –40° and +100°. Therefore an encirclement did not occur, and the circuit is stable.

An Unstable Case. Let us now change the gate bias to –2 V—a value that increases the amplifier gain enough to cause instability. Figures 3-275 and 3-276 show the results. From Figure 3-275 we can see that the origin is encircled in a clockwise direction. Figure 3-276 confirms this since the phase angle goes through –180° and continues without unraveling.

Approximate Frequency of Oscillation. The frequency where the Nyquist plot crosses –180° is often the approximate frequency of the unstable right-half-plane zeros and provides an estimate of the potential oscillation. For this example, there is a sharp crossing at 200 MHz that is characteristic of a circuit resonance. Note that the frequency is not related

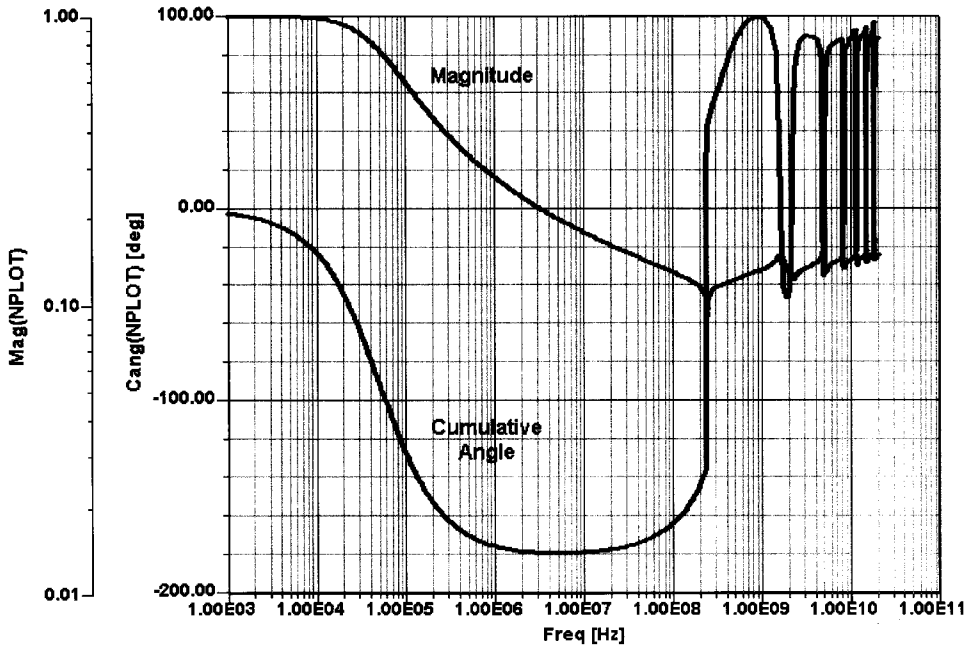


Figure 3-274 Magnitude and cumulative angle Nyquist plot for the 2.2-GHz amplifier covering 1 kHz to 20 GHz.

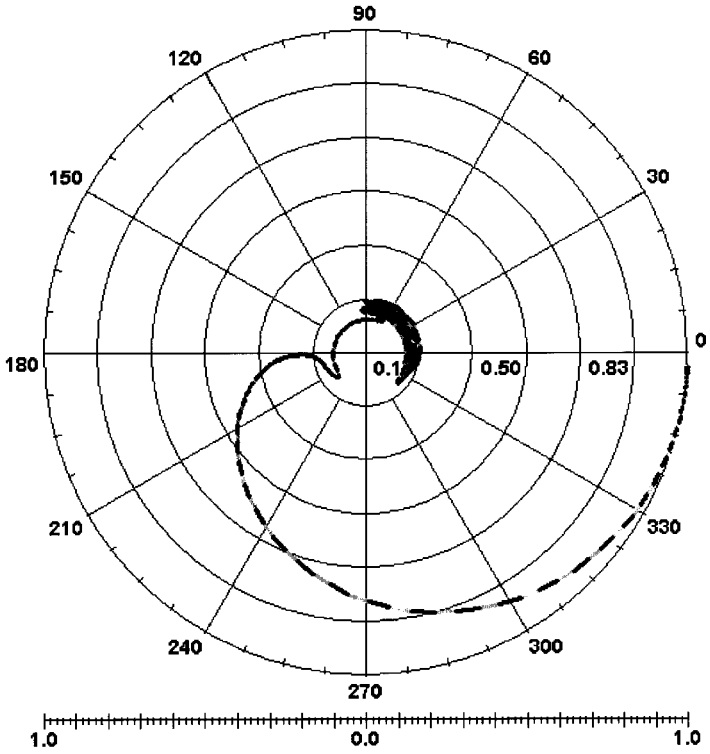


Figure 3-275 Polar Nyquist plot of the unstable amplifier.

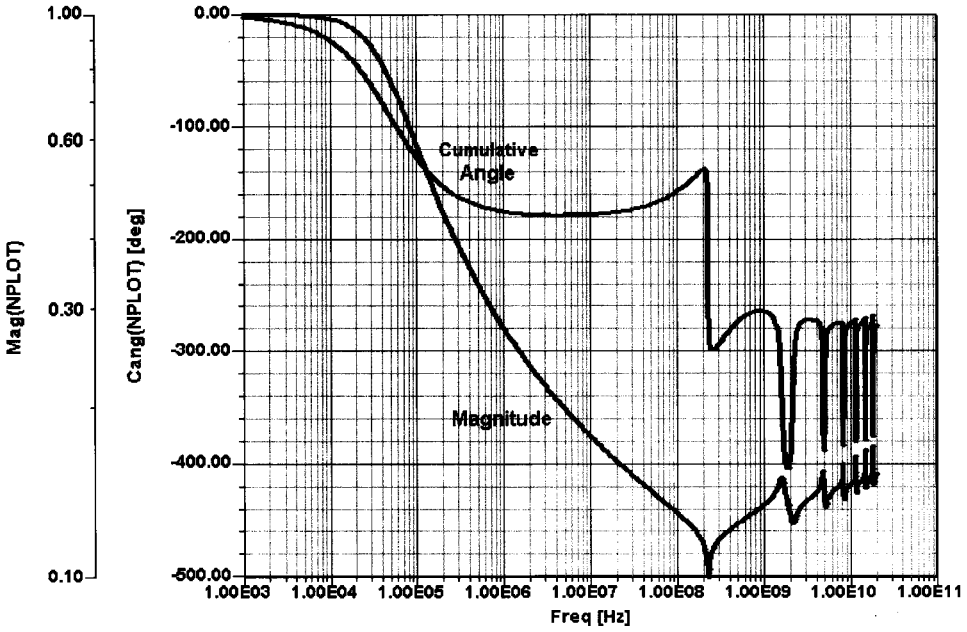


Figure 3-276 Magnitude–cumulative angle Nyquist plot of the unstable amplifier.

to a resonant frequency in a rigorous mathematical fashion, but it has been our experience that this frequency approximates the resonant frequency. To confirm the circuit's readiness for oscillation in the vicinity of this frequency, we evaluate it with Serenade's Oscillator Design Aid from 10 MHz to 1 GHz. The Design Aid finds an initial resonant frequency near 230.6 MHz—not too far from the –180° crossing frequency of 200 MHz of the Nyquist plot. Figure 3-277 shows this point.

Oscillator analysis finds a final oscillation frequency of 228.48 MHz. Figure 3-278 shows the resulting output spectrum at Port 2. At this point we decided to do a quick jump into Chapter 5 and take advantage of the CAD capability to predict the resulting phase noise, which is shown in Figure 3-279. As we compare this with other measured results, the approximately 228-MHz oscillator has a poor phase noise since the same phase noise is achievable at 1 GHz and higher. We remind the reader that this is an unwanted effect, but all of us who have built amplifiers will have experienced either low-frequency oscillations (motorboating) or very-high-frequency oscillation, sometimes caused by the components in the immediate vicinity of the transistor itself and found somewhere above 3 GHz. This case is really a demonstration case; a closer look reveals that the transconductance required for this condition is unrealistically high; but nonetheless it is a good example to demonstrate the stability issue.

Where Does Oscillation Begin? An interesting question to solve now is, "What gate bias voltage is needed for oscillation to begin?" Ansoft's Serenade Design Environment includes oscillator synchronous stability analysis, which uses a modification of the harmonic-balance technique to determine critical points, such as the voltage needed to "turn on" an oscillator. Mathematically, this is called a *Hopf bifurcation* because the behavior of

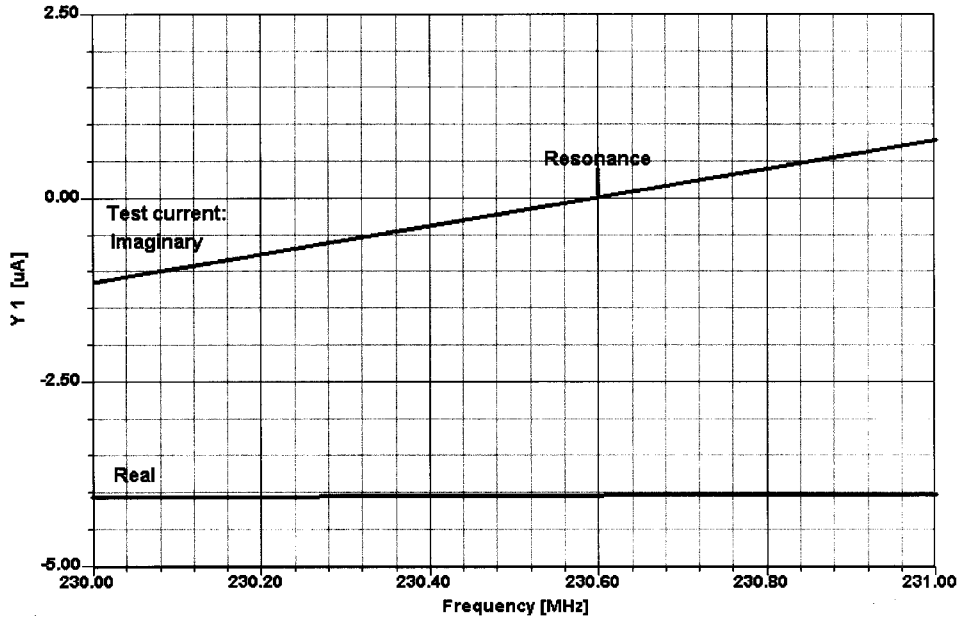


Figure 3-277 Graph from Serenade Oscillator Design Aid analysis of the amplifier. A resonant frequency is indicated where the imaginary part of the test current equals zero and the real part is negative.

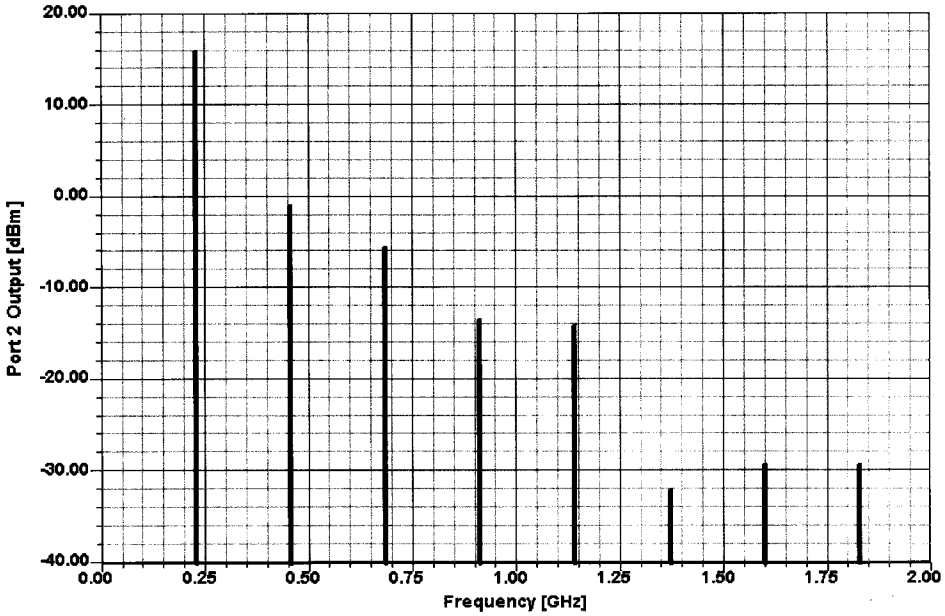


Figure 3-278 Predicted output spectrum of the unintended oscillator.

the circuit changes dramatically from a dc steady state to an oscillatory response. Figure 3-280 shows this effect by plotting output power versus gate bias. The turning point at –3.2 V (22 mW) is clearly visible, as is the bifurcation at –3.1 V (0 mW output).

Figure 3-280 can be interpreted as follows:

1. If the bias voltage is increased from, say, –4.0 V, the circuit will begin to oscillate at the bifurcation voltage and the output power will jump to 26 mW.
2. If the circuit is oscillating and the voltage is decreased from –2.0 V, the circuit will cease oscillating at the turning voltage of –3.2 V.
3. The branch between the turning point and the bifurcation point is unstable and cannot be realized in practice. This forms a hysteresis loop around points (–3.1 V, 0 mW), (–3.1 V, 26 mW), (–3.2 V, 22 mW), and (–3.2 V, 0 mW).

This amplifier stability analysis section is based on the stability analysis examples provided by Ansoft for the Serenade Design Environment.

An interesting example that combines both CAD and the feed-forward amplifier technique can be found in the Serenade Design Environment system simulator examples manual. This technique has been under evaluation for the last two years, but our feeling is that not enough measured data at frequencies above 500 MHz with reliable performance information have yet been made available in the literature. Along these lines, work done at Ansoft under a U.S. Air Force contract has provided very good insight into high-efficiency amplifiers at microwave frequencies [46].

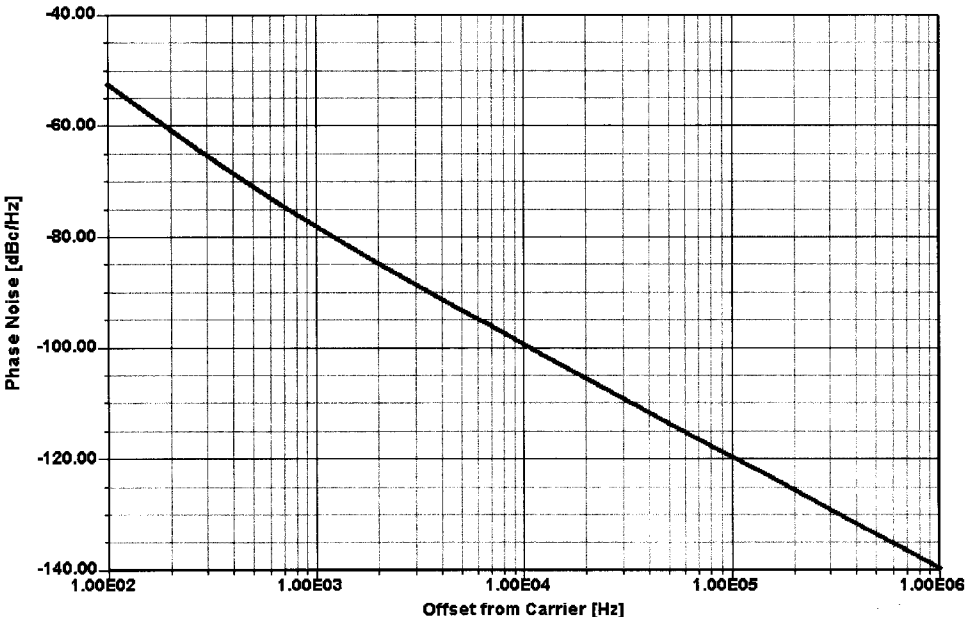


Figure 3-279 Predicted phase noise of the oscillating amplifier.

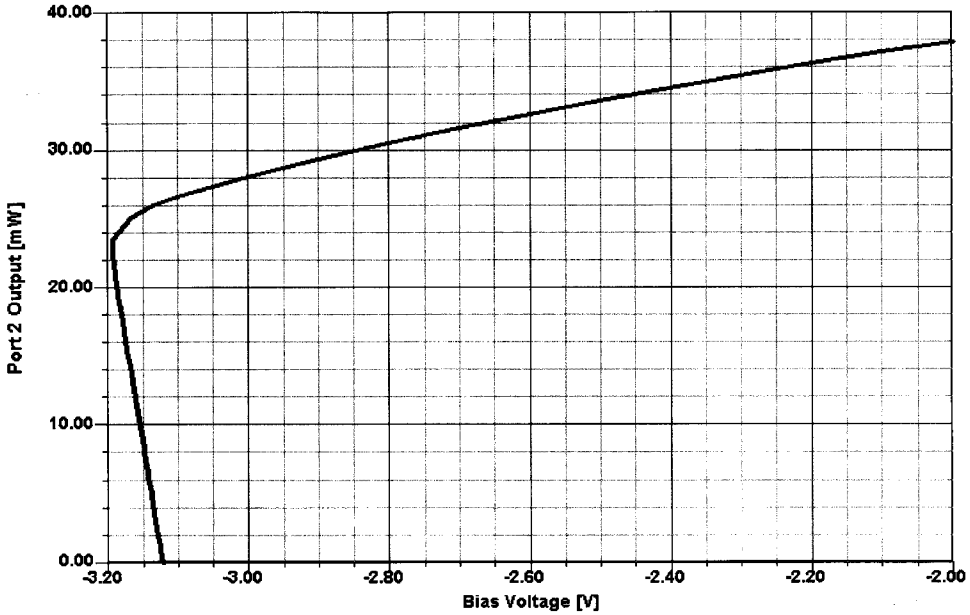


Figure 3-280 Output power versus gate bias for the unstable amplifier.

3-13 POWER AMPLIFIER DATASHEETS AND MANUFACTURER-RECOMMENDED APPLICATIONS

Having discussed the various aspects of BJTs and FETs, we have decided to use MOS examples from Ericsson and Motorola, and a bipolar example from Motorola, as good examples to show the technology required to build high-power transistor amplifiers in this frequency range. In the following pages, we reproduce, with permission, datasheets for Ericsson's PTF 10009 FET and Motorola's MRF186 FET, followed by Motorola's MRF899 BJT.

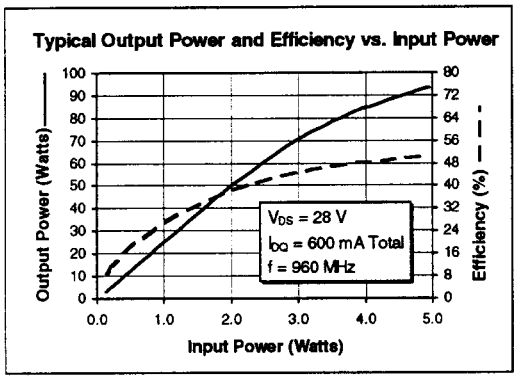
ERICSSON

PTF 10009
85 Watts, 1.0 GHz
LDMOS Field Effect Transistor

Description

The 10009 is a common source N-channel enhancement-mode lateral MOSFET intended for large signal amplifier applications to 1.0 GHz. It is rated at 85 watts minimum output power. Nitride surface passivation and gold metallization are used to ensure excellent device lifetime and reliability. 100% lot traceability is standard.

- Performance at 960 MHz, 28 Volts
  - Output Power = 85 Watts
  - Power Gain = 13.0 dB Typ
  - Efficiency = 50% Typ
- Gold Metallized
- Silicon Nitride Passivated
- Excellent Thermal Stability



Package 20230

Maximum Ratings

Parameter	Symbol	Value	Unit
Drain-Source Voltage	V_{DSS}	65	Vdc
Gate-Source Voltage	V_{GS}	±20	Vdc
Operating Junction Temperature	T_J	200	°C
Total Device Dissipation at T_{flange} = 25°C Above 25°C derate by	P_D	270 1.54	Watts W/°C
Storage Temperature Range	T_{STG}	-65 to 150	°C
Thermal Resistance (T_{flange} = 70°C)	$R_{\theta JC}$	0.65	°C/W

PTF 10009

ERICSSON

Electrical Characteristics (100% Tested)

Characteristic (per side)	Conditions	Symbol	Min	Typ	Max	Units
Drain-Source Breakdown Voltage	V_{GS} = 0 V, I_D = 25 mA	$V_{(BR)DSS}$	65	—	—	Volts
Drain-Source Leakage Current	V_{DS} = 28 V, V_{GS} = 0 V	I_{DSS}	—	—	1.0	mA
Gate Threshold Voltage	V_{DS} = 10 V, I_D = 75 mA	$V_{GS(th)}$	—	2.0	—	Volts
Forward Transconductance	V_{DS} = 10 V, I_D = 3 A	g_{fs}	—	2.8	—	Siemens

Dynamic Characteristics

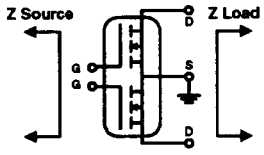
Characteristic (per side)	Symbol	Min	Typ	Max	Units
Input Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{iss}	—	90	—	pF
Output Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{oss}	—	36	—	pF
Reverse Transfer Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{rss}	—	1.9	—	pF

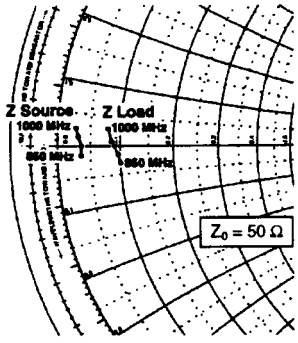
RF Specifications (100% Tested)

Characteristic	Symbol	Min	Typ	Max	Units
Gain (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz)	G_{ps}	12.0	13.0	—	dB
Drain Efficiency (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz)	η	47	50	—	%
Load Mismatch Tolerance (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz— all phase angles at frequency of test)	Ψ	—	—	5:1	—

Impedance Data (data shown for fixed-tuned broadband circuit)

(V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA)



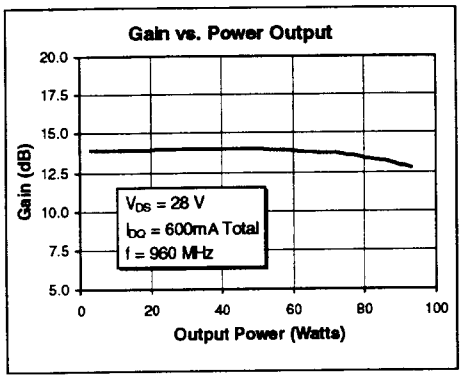


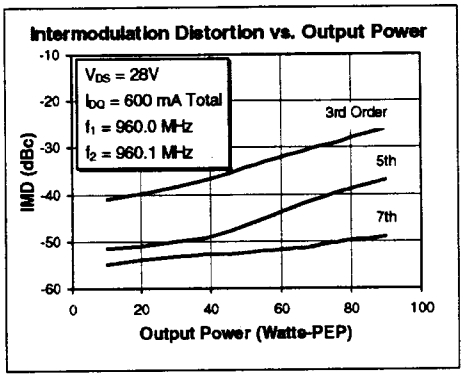
Frequency	Z Source		Z Load	
MHz	R	jX	R	jX
860	1.76	-0.78	5.00	-1.50
900	1.80	-0.05	4.80	-0.78
960	1.58	0.69	4.24	0.36
1000	1.39	1.35	3.95	1.41

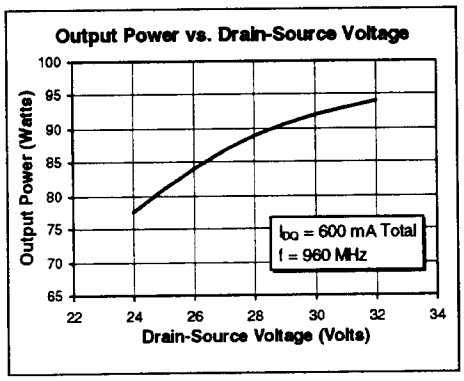
ERICSSON

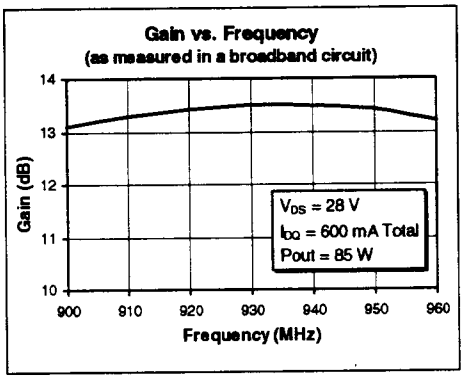
PTF 10009

Typical Performance





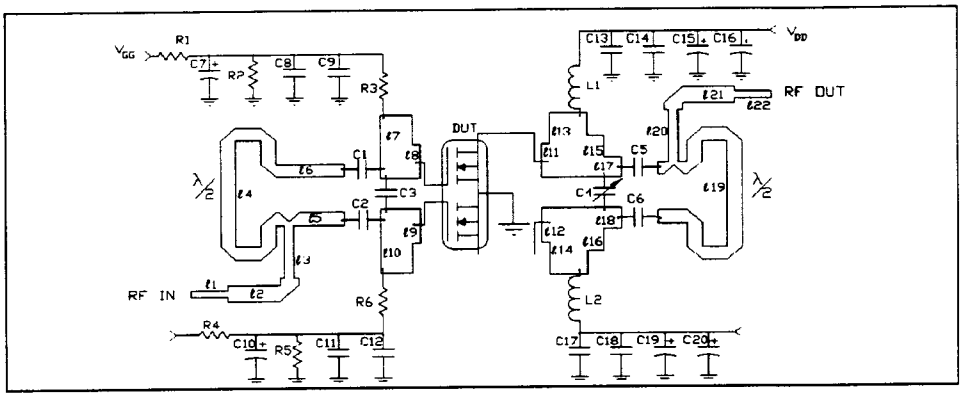




PTF 10009

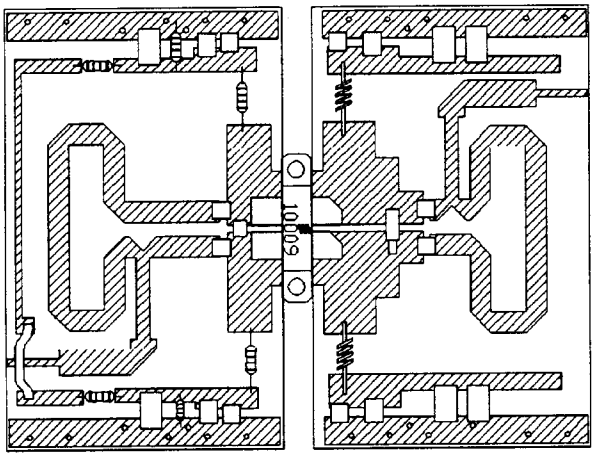
ERICSSON

Test Circuit



Schematic for f = 960 MHz

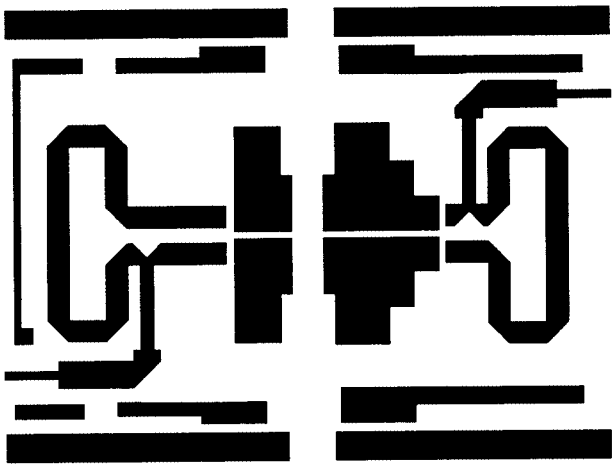
DUT	10009
C1-2, C5-6, C9, C12-13, C17	33 pF, Capacitor ATC 100 B
C3	11 pF, Capacitor ATC 100 B
C4	6.0 pF, Variable Capacitor, JMC 5701
C7, C10	10 µF, +10 V Electrolytic Capacitor
C8, C11, C14, C18	0.01 µF, Capacitor ATC 100 B
C15, C16, C19, C20	10 µF, +30 V Electrolytic Capacitor
L1. L2	4 Turn, #20 AWG, .120" I.D.
R1, R2, R4, R5	1.0 K, Ω Resistor
R3, R6	5.1 K, 1/4 Ω Resistor
ℓ1, ℓ22	50 Ω, .030 λ
ℓ2, ℓ21	20 Ω, .080 λ
ℓ3, ℓ20	32 Ω, .191 λ
ℓ4, ℓ19	25 Ω, .500 λ
ℓ5, ℓ6	25 Ω, .091 λ
ℓ7, ℓ10	7 Ω, .056 λ
ℓ8, ℓ9	13.0 Ω, .017 λ
ℓ11, ℓ12	13.0 Ω, .017 λ
ℓ13, ℓ14	7.0 Ω, .064 λ
ℓ15, ℓ16	10.0 Ω, .029 λ
ℓ17, ℓ18	19.0 Ω, .028 λ
Circuit Board	.031" Thick, ε_r = 4.0, AlliedSignal, G200



Parts Layout (not to scale)

ERICSSON

PTF 10009



Artwork (1 inch ——— *)*

Ericsson Components
RF Power Products
675 Jarvis Drive
Morgan Hill, CA 95037 USA
Telephone: 408-778-9434

1-877-GOLDMOS
(1-877-465-3667)
e-mail: rfpower@ericsson.com
www.ericsson.com/rfpower

Specifications subject to change
without notice.
L1
© 1997 Ericsson Inc.
EUS/KR 1301-PTF 10009 Uen Rev. B 09-29-98

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF186/D

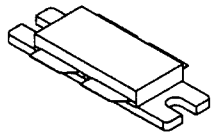
The RF MOSFET Line
RF Power Field-Effect Transistor
N–Channel Enhancement–Mode Lateral MOSFET

MRF186

120 W, 1.0 GHz
28 VOLTS
LATERAL N–CHANNEL
BROADBAND
RF POWER MOSFET

Designed for broadband commercial and industrial applications at frequencies from 800 MHz to 1.0 GHz. The high gain and broadband performance of this device makes it ideal for large–signal, common source amplifier applications in 28 volt base station equipment.

- Guaranteed Performance @ 960 MHz, 28 Volts
  Output Power — 120 Watts (PEP)
  Power Gain — 11 dB
  Efficiency — 30%
  Intermodulation Distortion — –28 dBc
- Excellent Thermal Stability
- 100% Tested for Load Mismatch Stress at all Phase Angles with 5:1 VSWR @ 28 Vdc, 960 MHz, 120 Watts CW



CASE 375B–02, STYLE 2

MAXIMUM RATINGS (2)

Rating	Symbol	Value	Unit
Drain–Source Voltage	V_{DSS}	65	Vdc
Drain–Gate Voltage (R_{GS} = 1 MΩ)	V_{DGR}	65	Vdc
Gate–Source Voltage	V_{GS}	±20	Vdc
Drain Current — Continuous	I_D	14	Adc
Total Device Dissipation @ T_C = 70°C Derate above 70°C	P_D	162.5 1.25	Watts W/°C
Storage Temperature Range	T_{stg}	– 65 to +150	°C
Operating Junction Temperature	T_J	200	°C

THERMAL CHARACTERISTICS (2)

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	0.8	°C/W

NOTE – **CAUTION** – MOS devices are susceptible to damage from electrostatic charge. Reasonable precautions in handling and packaging MOS devices should be observed.

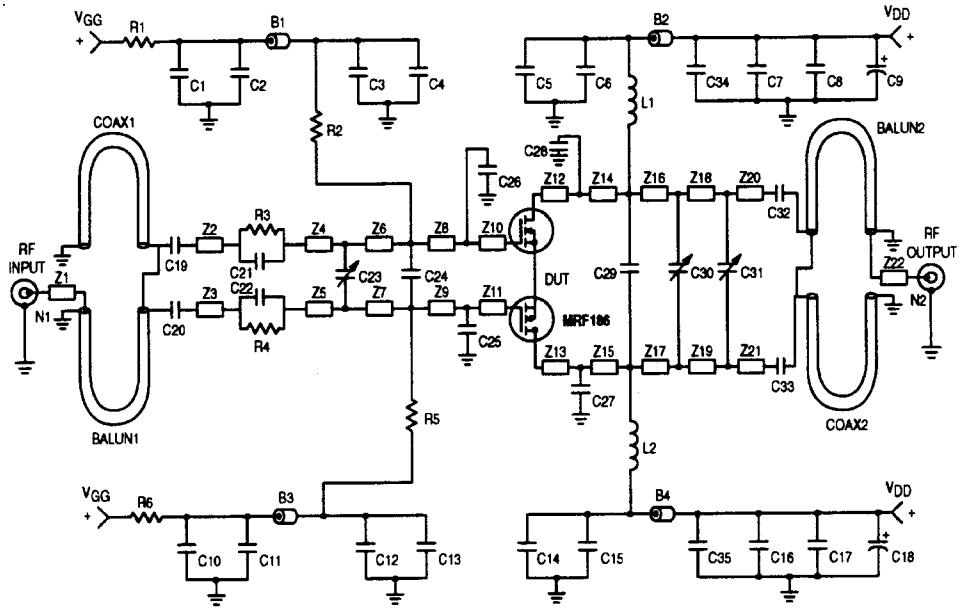
© Motorola, Inc. 1999

MOTOROLA

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted)

Characteristic	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS (1)					
Drain–Source Breakdown Voltage (V_{GS} = 0 Vdc, I_D = 50 μAdc)	$V_{(BR)DSS}$	65	—	—	Vdc
Zero Gate Voltage Drain Current (V_{DS} = 28 Vdc, V_{GS} = 0)	I_{DSS}	—	—	1	μAdc
Gate–Source Leakage Current (V_{GS} = 20 Vdc, V_{DS} = 0)	I_{GSS}	—	—	1	μAdc
ON CHARACTERISTICS (1)					
Gate Quiescent Voltage (V_{DS} = 26 Vdc, I_D = 300 μAdc Per Side)	$V_{GS(th)}$	2.5	3	4	Vdc
Gate Quiescent Voltage (V_{DS} = 26 Vdc, I_D = 300 mAdc Per Side)	$V_{GS(Q)}$	3.3	4.2	5	Vdc
Delta Gate Threshold Voltage (Side to Side) (V_{DS} = 28 V, I_D = 300 mA Per Side)	$\Delta V_{GS(Q)}$	—	—	0.3	Vdc
Drain–Source On–Voltage (V_{GS} = 10 Vdc, I_D = 3 Adc Per Side)	$V_{DS(on)}$	—	0.58	0.7	Vdc
Forward Transconductance (V_{DS} = 10 Vdc, I_D = 3 Adc Per Side)	g_{fs}	2.4	2.8	—	S
DYNAMIC CHARACTERISTICS (1)					
Input Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{iss}	—	177	—	pF
Output Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{oss}	—	45	—	pF
Reverse Transfer Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{rss}	—	3.4	—	pF
FUNCTIONAL CHARACTERISTICS (In Motorola Test Fixture) (2)					
Two–Tone Common Source Amplifier Power Gain (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	G_{ps}	11	12.2	—	dB
Two–Tone Drain Efficiency (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	η	30	35	—	%
3rd Order Intermodulation Distortion (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	IMD	—	–32	–28	dBc
Input Return Loss (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	IRL	9	16	—	dB
Two–Tone Common Source Amplifier Power Gain (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	G_{ps}	—	12	—	dB
Two–Tone Drain Efficiency (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	η	—	33	—	%
3rd Order Intermodulation Distortion (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	IMD	—	–32	—	dBc
Input Return Loss (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	IRL	—	16	—	dB
Output Mismatch Stress (V_{DD} = 28 Vdc, P_{out} = 120 W CW, I_{DQ} = 2 x 400 mA, f = 960 MHz, VSWR = 5:1, All Phase Angles at Frequency of Tests)	Ψ	No Degradation In Output Power Before and After Test			

(1) Each side of device measured separately.
(2) Device measured in push–pull configuration.



B1 – B4	Fair Rite Products Short Ferrit Bead, 2743021446	C31	0.8 – 8.0 pF, Variable Capacitor, Johanson Gigatrim
C1, C7, C8, C10, C16, C17	10 µF, 50 V, Tantalum	L1, L2	3 Turns, #20 AWG, IDIA 0.126", 24.7 nH
C2, C11, C34, C35	0.1 µF, Chip Capacitor	N1, N2	Type N Connectors
C3, C6, C12, C15	330 pF, Chip Capacitor	R1, R6	1 kΩ, 1/4 W, Carbon Resistor
C4, C5, C13, C14, C19, C20, C32, C33	47 pF, Chip Capacitor	R2, R5	1.2 kΩ, 0.1 W, Chip Resistor
C9, C18	250 µF, 50 V, Electrolytic Capacitor	R3, R4	75 Ω, 0.1 W, Chip Resistor
C21, C22	12 pF, Chip Capacitor	Z1 – Z22	Microstrip (See Component Placement)
C23, C30	0.6 – 4.5 pF, Variable Capacitor, Johanson Gigatrim	Balun1, Balun2, Coax1, Coax2	2.20" 50 Ω, 0.086" OD Semi–Rigid Coax
C24, C25, C26	5.1 pF, Chip Capacitor	Board	1/32" Glass Teflon®, ε_r = 2.55
C27, C28	3.9 pF, Chip Capacitor		

Figure 1. 930 – 960 MHz Test Circuit Schematic

TYPICAL CHARACTERISTICS

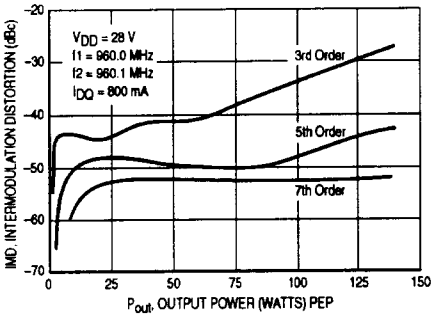


Figure 2. Intermodulation Distortion versus Output Power

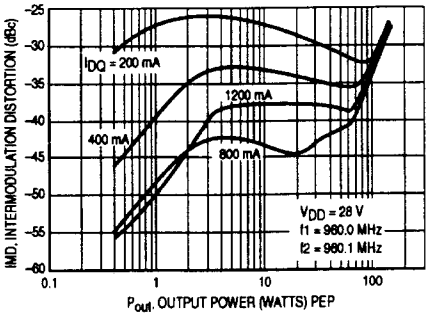


Figure 3. Intermodulation Distortion versus Output Power

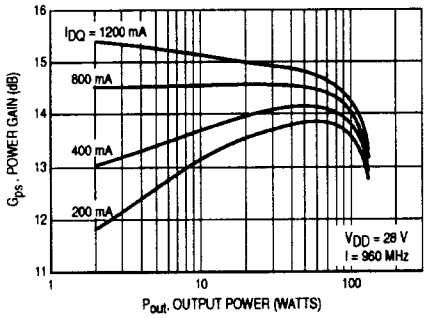


Figure 4. Power Gain versus Output Power

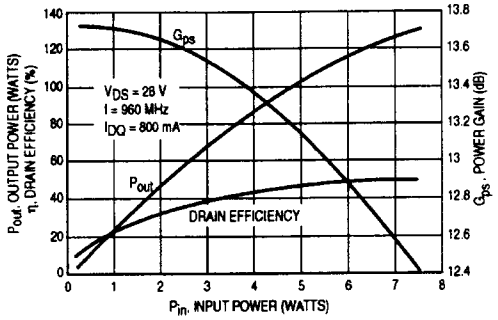


Figure 5. Output Power versus Input Power

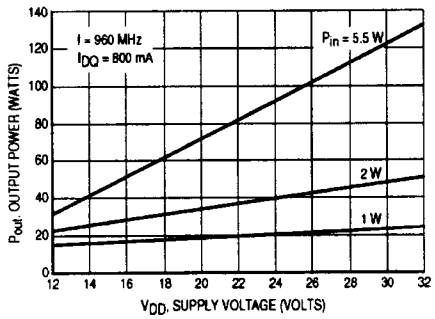


Figure 6. Output Power versus Supply Voltage

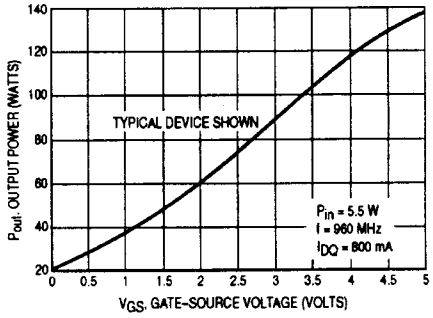


Figure 7. Output Power versus Gate Voltage

TYPICAL CHARACTERISTICS

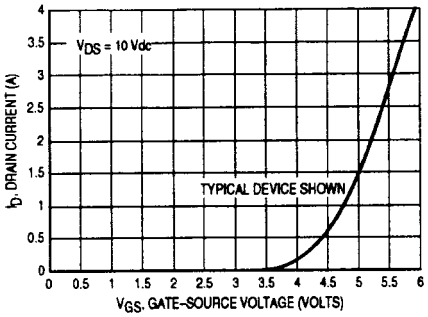


Figure 8. Drain Current versus Gate Voltage

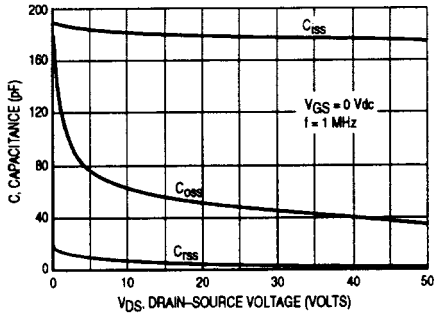


Figure 9. Capacitance versus Voltage

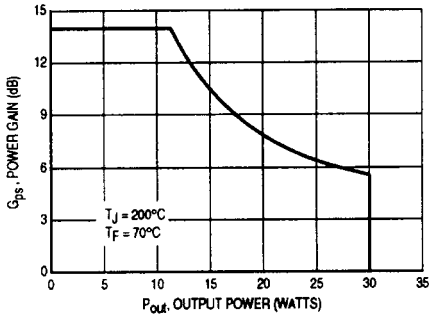


Figure 10. DC Safe Operating Area

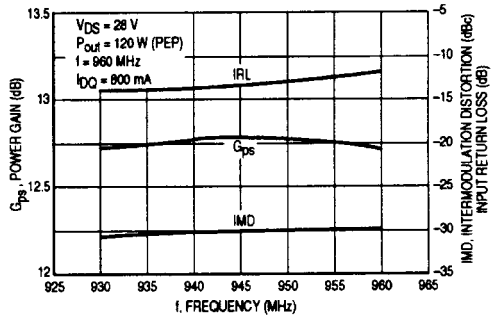
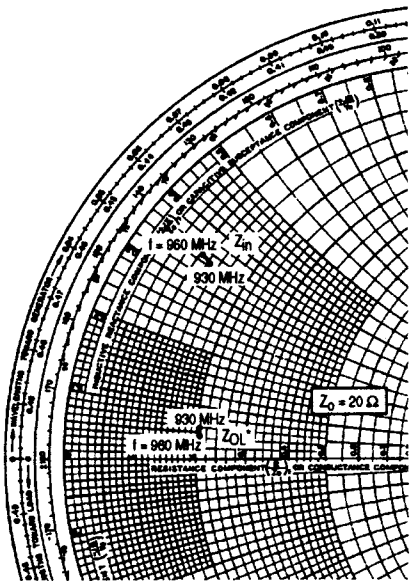


Figure 11. Broadband Circuit Performance



V_{CC} = 28 V, I_{DQ} = 2 × 400 mA, P_{out} = 120 Watts (PEP)

f MHz	Z_{in} Ω	Z_{OL}^* Ω
930	2.5 + j6.9	4.3 + j1.2
945	2.5 + j7.0	4.3 + j1.0
960	2.2 + j7.1	4.3 + j0.9

Z_{in} = Conjugate of impedance.

Z_{OL}^* = Conjugate of the optimum load impedance at a given output power, voltage, IMD, bias current, efficiency and frequency.

Note: Z_{OL}^* was chosen based on tradeoffs between gain, output power, drain efficiency and intermodulation performance. Impedances shown represent a single channel (1/2 of MRF186) impedance measurement.

Figure 12. Series Equivalent Input and Output Impedance

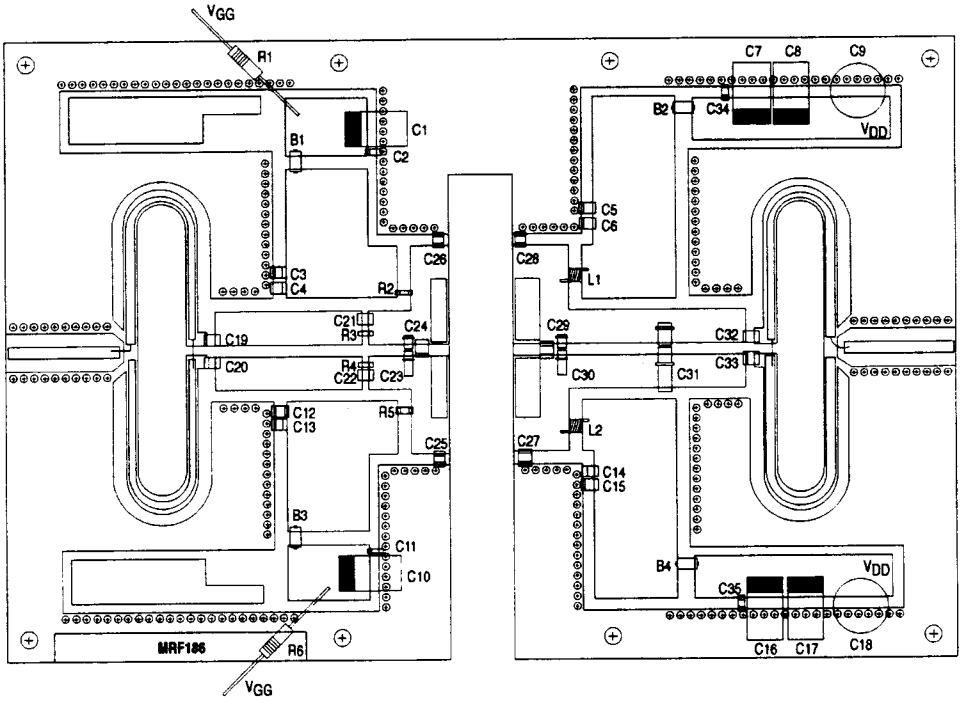
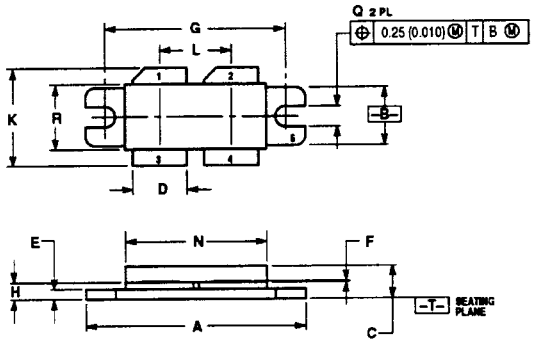


Figure 13. Component Placement Diagram of 930 – 960 MHz Broadband Test Fixture

PACKAGE DIMENSIONS



NOTES:
1. DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2. CONTROLLING DIMENSION: INCH.

DIM	INCHES MIN	INCHES MAX	MILLIMETERS MIN	MILLIMETERS MAX
A	1.330	1.350	33.78	34.29
B	0.375	0.395	9.52	10.03
C	0.180	0.205	4.57	5.21
D	0.320	0.340	8.13	8.64
E	0.060	0.070	1.52	1.77
F	0.004	0.006	0.11	0.15
G	1.100 BSC		27.94 BSC	
H	0.082	0.097	2.09	2.46
K	0.580	0.620	14.73	15.75
L	0.435 BSC		11.05 BSC	
N	0.845	0.875	21.46	22.23
Q	0.118	0.130	3.00	3.30
R	0.390	0.410	9.91	10.41

STYLE 1:
PIN 1. COLLECTOR
2. COLLECTOR
3. BASE
4. BASE
5. EMITTER

CASE 375A–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Motorola data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and Ⓜ are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

Mfax is a trademark of Motorola, Inc.

How to reach us:
USA/EUROPE/Locations Not Listed: Motorola Literature Distribution; P.O. Box 5405, Denver, Colorado 80217. 303–675–2140 or 1–800–441–2447

Mfax™: RMFAX0@email.sps.mot.com – TOUCHTONE 602–244–6609 – US & Canada ONLY 1–800–774–1848

INTERNET: http://motorola.com/sps

JAPAN: Nippon Motorola Ltd.: SPD, Strategic Planning Office, 4–32–1, Nishi–Gotanda, Shinagawa–ku, Tokyo 141, Japan. 81–3–5487–8488

ASIA/PACIFIC: Motorola Semiconductors H.K. Ltd.; 8B Tai Ping Industrial Park, 51 Ting Kok Road, Tai Po, N.T., Hong Kong. 852–26629298

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF899/D

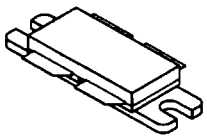
The RF Line
NPN Silicon RF Power Transistor

MRF899

150 W, 900 MHz
RF POWER
TRANSISTOR
NPN SILICON

Designed for 26 Volt UHF large–signal, common emitter, Class AB linear amplifier applications in industrial and commercial FM/AM equipment operating in the range 800–960 MHz.

- Specified 26 Volt, 900 MHz Characteristics
  Output Power = 150 Watts (PEP)
  Minimum Gain = 8.0 dB @ 900 MHz, Class AB
  Minimum Efficiency = 35% @ 900 MHz, 150 Watts (PEP)
  Maximum Intermodulation Distortion –28 dBc @ 150 Watts (PEP)
- Characterized with Series Equivalent Large–Signal Parameters from 800 to 960 MHz
- Silicon Nitride Passivated
- 100% Tested for Load Mismatch Stress at all Phase Angles with 5:1 VSWR @ 26 Vdc, and Rated Output Power
- Gold Metallized, Emitter Ballasted for Long Life and Resistance to Metal Migration
- Circuit board photomaster available upon request by contacting RF Tactical Marketing in Phoenix, AZ.



CASE 375A–01, STYLE 1

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Collector–Emitter Voltage	V_{CEO}	28	Vdc
Collector–Emitter Voltage	V_{CES}	60	Vdc
Emitter–Base Voltage	V_{EBO}	4.0	Vdc
Collector–Current — Continuous	I_C	25	Adc
Total Device Dissipation @ T_C = 25°C Derate above 25°C	P_D	230 1.33	Watts W/°C
Storage Temperature Range	T_{stg}	–65 to +150	°C

THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	0.75	°C/W

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS					
Collector–Emitter Breakdown Voltage (I_C = 100 mAdc, I_B = 0)	$V_{(BR)CEO}$	28	37	—	Vdc
Collector–Emitter Breakdown Voltage (I_C = 50 mAdc, V_{BE} = 0)	$V_{(BR)CES}$	60	85	—	Vdc
Emitter–Base Breakdown Voltage (I_E = 10 mAdc, I_C = 0)	$V_{(BR)EBO}$	4.0	4.9	—	Vdc
Collector Cutoff Current (V_{CE} = 30 Vdc, V_{BE} = 0)	I_{CES}	—	—	10	mAdc
ON CHARACTERISTICS					
DC Current Gain (I_{CE} = 1.0 Adc, V_{CE} = 5.0 Vdc)	h_{FE}	30	75	120	—
DYNAMIC CHARACTERISTICS					
Output Capacitance (V_{CB} = 26 Vdc, I_E = 0, f = 1.0 MHz) (1)	C_{ob}	—	75	—	pF

(1) For information only. This part is collector matched.

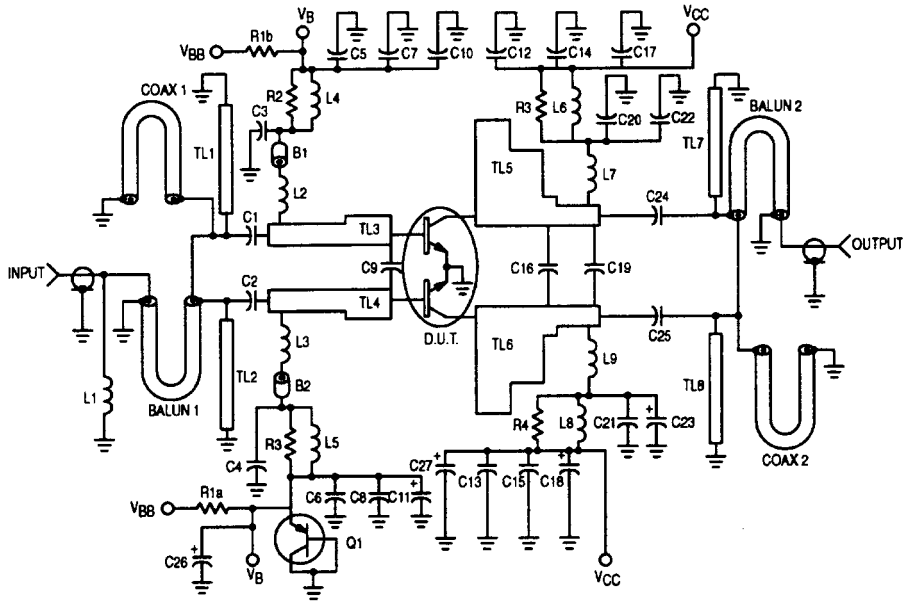
(continued)

© Motorola, Inc. 1997

MOTOROLA

ELECTRICAL CHARACTERISTICS — continued (T_C = 25°C unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
FUNCTIONAL CHARACTERISTICS					
Common–Emitter Amplifier Power Gain V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	G_{pe}	8.0	9.0	—	dB
Collector Efficiency V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	η	30	40	—	%
3rd Order Intermodulation Distortion V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	IMD	—	–32	–28	dBc
Output Mismatch Stress V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz, VSWR = 5:1 (all phase angles)	Ψ	No Degradation in Output Power Before and After Test			



B1, B2 — Ferrite Bead, Ferroxcube #56–590–65–3B
C1, C2, C24, C25 — 43 pF, B Case, ATC Chip Capacitor
C3, C4, C20, C21 — 100 pF, B Case, ATC Chip Capacitor
C5, C6, C12, C13 — 1000 pF, B Case, ATC Chip Capacitor
C7, C8, C14, C15 — 1800 pF, AVX Chip Capacitor
C9 — 9.1 pF, A Case, ATC Chip Capacitor
C10, C11, C17, C18, C22, C23 — 10 µF, Electrolytic Capacitor Panasonic
C16 — 3.9 pF, B Case, ATC Chip Capacitor
C19 — 0.8 pF, B Case, ATC Chip Capacitor
C26 — 200 µF, Electrolytic Capacitor Mallory Sprague
C27 — 500 µF Electrolytic Capacitor

L1 — 5 Turns 24 AWG IDIA 0.059″ Choke, 19.8 nH
L2, L3, L7, L9 — 4 Turns 20 AWG IDIA 0.163″ Choke
L4, L5, L6, L8 — 12 Turns 22 AWG IDIA 0.140″ Choke
N1, N2 — Type N Flange Mount, Omni Spectra
Q1 — Bias Transistor BD136 PNP
R2, R3, R4, R5 — 4.0 x 39 Ohm 1/8 W Chips in Parallel
R1a, R1b — 56 Ohm 1.0 W
TL1–TL8 — See Photomaster
Balun1, Balun2, Coax 1, Coax 2 — 2.20″ 50 Ohm 0.086″ o.d. Semi–rigid Coax, Micro Coax
Board — 1/32″ Glass Teflon, ε_r = 2.55″ Arlon (GX–0300–55–22)

Figure 1. 900 MHz Power Gain Test Circuit

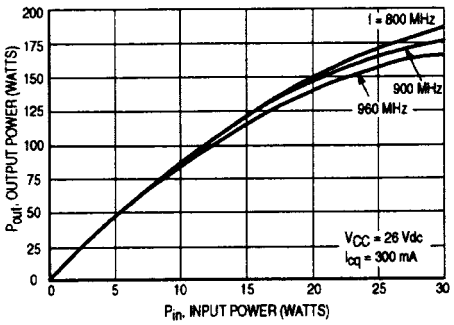


Figure 2. Output Power versus Input Power

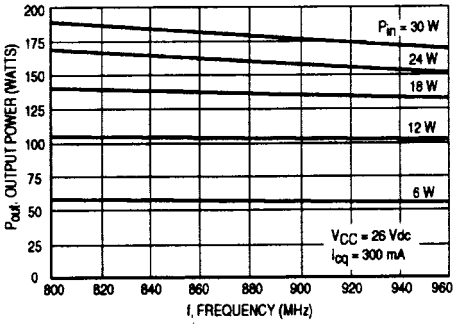


Figure 3. Output Power versus Frequency

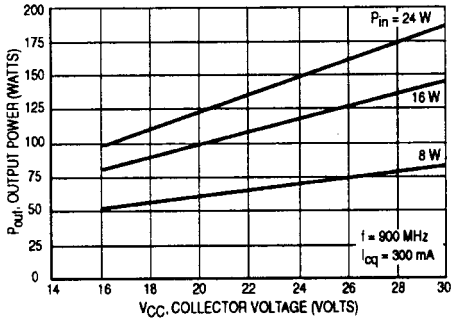


Figure 4. Output Power versus Supply Voltage

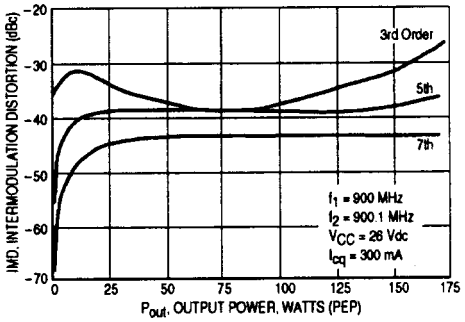


Figure 5. Intermodulation versus Output Power

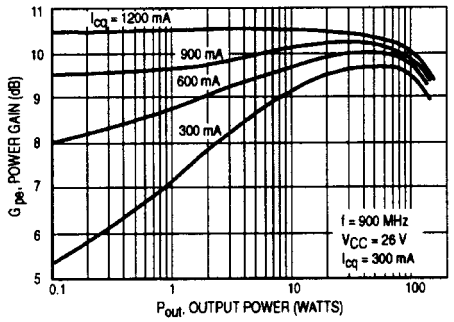


Figure 6. Power Gain versus Output Power

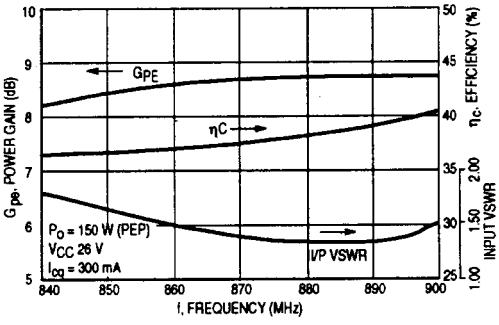
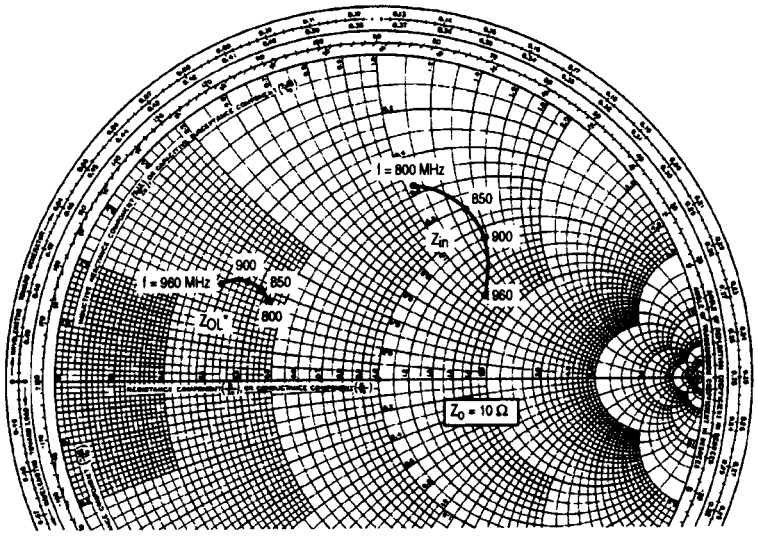


Figure 7. Broadband Test Fixture Performance



f MHz	Z_{in} Ohms	Z_{OL}^* Ohms
800	5.51 + j10.6	4.52 + j2.64
850	8.17 + j13.2	4.21 + j2.98
900	11.2 + j13.8	3.68 + j2.97
960	16.8 + j10.1	2.98 + j2.71

Z_{OL}^* = Conjugate of optimum load impedance into which the device operates at a given output power, voltage and frequency.

NOTE: Z_{in} & Z_{OL}^* are given from base–to–base and collector–to–collector respectively

Figure 8. Input and Output Impedances with Circuit Tuned for Maximum Gain @ P_O = 150 W (PEP), V_{CC} = 26 V

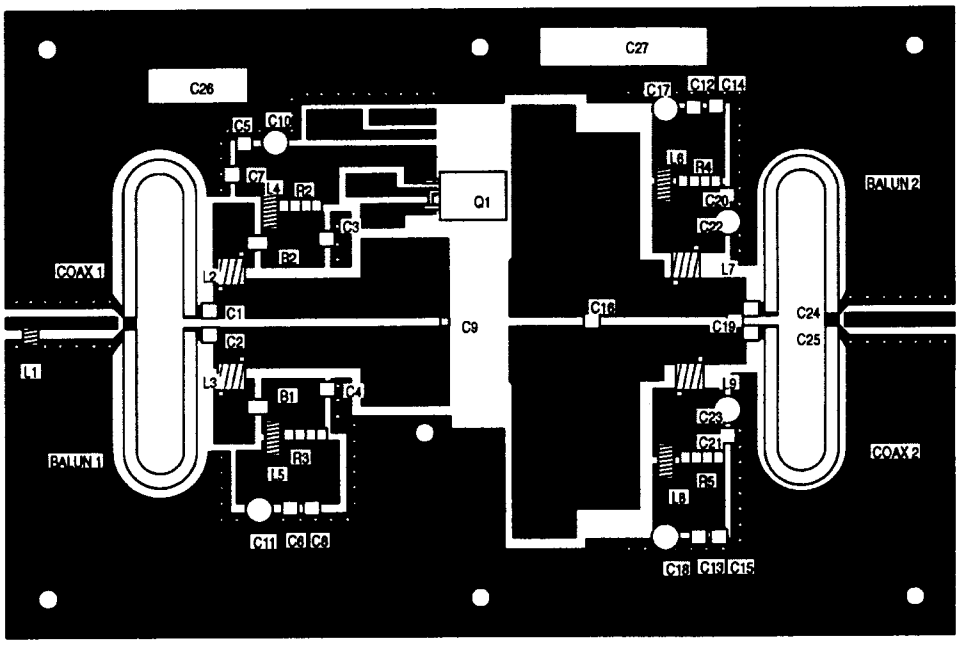
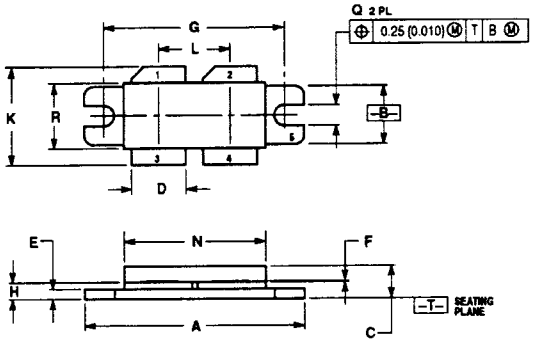


Figure 9. MRF899 Test Fixture Component Layout

PACKAGE DIMENSIONS



NOTES:
1. DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2. CONTROLLING DIMENSION: INCH.

DIM	INCHES MIN	INCHES MAX	MILLIMETERS MIN	MILLIMETERS MAX
A	1.330	1.350	33.79	34.29
B	0.375	0.395	9.52	10.03
C	0.160	0.205	4.57	5.21
D	0.320	0.340	8.13	8.64
E	0.060	0.070	1.52	1.77
F	0.004	0.006	0.11	0.15
G	1.100 BSC		27.94 BSC	
H	0.062	0.097	2.08	2.46
K	0.580	0.620	14.73	15.75
L	0.435 BSC		11.05 BSC	
N	0.845	0.875	21.46	22.23
Q	0.116	0.130	3.00	3.30
R	0.390	0.410	9.91	10.41

STYLE 1:
PIN 1. COLLECTOR
2. COLLECTOR
3. BASE
4. BASE
5. EMITTER

CASE 375A–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Motorola data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and M are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

Mfax is a trademark of Motorola, Inc.

How to reach us:
USA/EUROPE/Locations Not Listed: Motorola Literature Distribution;
P.O. Box 5405, Denver, Colorado 80217. 303–675–2140 or 1–800–441–2447

Mfax™: RMFAX0@email.sps.mot.com – TOUCHTONE 602–244–6609
– US & Canada ONLY 1–800–774–1848

INTERNET: http://motorola.com/sps

JAPAN: Nippon Motorola Ltd.: SPD, Strategic Planning Office, 4–32–1, Nishi–Gotanda, Shinagawa–ku, Tokyo 141, Japan. 81–3–5487–8488

ASIA/PACIFIC: Motorola Semiconductors H.K. Ltd.; 8B Tai Ping Industrial Park, 51 Ting Kok Road, Tai Po, N.T., Hong Kong. 852–26629298

REFERENCES

1. Ulrich L. Rohde, Anthony M. Pavio, and Robert A. Pucel, "Accurate Noise Simulation of Microwave Amplifiers Using CAD," *Microwave J.*, **31**(12), 130–141, December 1988.
2. H. T. Friis, "Noise Figures of Radio Receivers," *Proc. IRE*, **32**, 419–422, July 1944.
3. "IRE Standards on Methods of Measuring Noise in Linear Twoports," *Proc. IRE*, **48**, 60–68, January 1960.
4. H. Rothe and W. Dhalke, "Theory of Noisy Fourpoles," *Proc. IRE*, **44**, 811–818, June 1956.
5. "Representation of Noise in Linear Twoports," *Proc. IRE*, **48**, 69–74, January 1960.
6. H. A. Haus and R. B. Adler, *Circuit Theory of Linear Noisy Networks*, Cambridge, MA: MIT Press, 1959/New York: Wiley, 1959.
7. H. Fukui, "Noise Performance of Microwave Transistors," *IEEE Trans. Electron Devices*, **ED-13**, 329–341, March 1966.
8. S. Iversen, "The Effect of Feedback on Noise Figure," *Proc. IEEE*, **63**, 540–542, March 1975.
9. L. Besser, "Stability Considerations of Low-Noise Transistor Amplifiers with Simultaneous Noise and Power Match," *IEEE MTT-S International Microwave Symposium Digest*, 1975, pp. 327–329.
10. G. D. Vendelin, "Feedback Effects on the Noise Performance of GaAs Mesfets," *IEEE MTT-S International Microwave Symposium Digest*, 1975, pp. 324–326.
11. W. A. Suter, "Feedback and Parasitic Effects on Noise," *Microwave J.* **26**(2), 123–129, February 1983.
12. K. Hartmann and M. J. O. Strutt, "Changes in Four Noise Parameters Due to General Changes in Linear Two-Port Circuits," *IEEE Trans. Electron Devices*, **ED-20**(10), 874–877, October 1973.
13. H. Fukui, "Available Power Gain, Noise Figure, and Noise Measure of Two-Ports and Their Graphical Representation," *IEEE Trans. Circuit Theory*, **CT-13**, 137–142, June 1966.
14. C. R. Poole and D. K. Paul, "Optimum Noise Measure Terminations for Microwave Transistor Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**(11), 1254–1257, November 1985.
15. H. Hillbrand and P. H. Russer, "An Efficient Method for Computer Aided Noise Analysis of Linear Amplifier Networks," *IEEE Trans. Circuits Syst.*, **CAS-23**(4), 235–238, April 1976.
16. W. E. Pastori, *Topics on Noise*, Eaton Electronics Instrumentation Division (EID) Seminar Notes, Los Angeles, 1983.
17. W. E. Pastori, "Image and Second Stage Corrections Resolve Noise Figure Measurement Confusion," *Microwave Syst. News*, **13**(5), 67–86, May 1983.
18. R. Q. Lane, "The Determination of Device Noise Parameters," *Proc. IEEE*, **57**, 1461–1462, August 1969.
19. R. Q. Lane, "A Microwave Noise and Gain Parameter Test Set," *IEEE International Solid-State Circuits Conference*, 1978, pp. 172–173.
20. N. Kuhn, "Accurate and Automatic Noise Figure Measurements with Standard Equipment," *IEEE MTT-S International Microwave Symposium Digest*, 1980, pp. 425–427.
21. E. W. Strid, "Measurement of Losses in Noise-Matching Networks," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**, 247–252, March 1981.
22. Chris Bowick, *RF Circuit Design*. Indianapolis: Howard W. Sams & Co., 1982.
23. Adel S. Sedra and Kenneth C. Smith, *Microelectronic Circuits*. Chicago: Saunders College Publishing, 1991.
24. Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed. New York: Wiley, 1984.

25. S. H. Lee, A. K. Wong, and M. G. Wong, *A Current-Combiner Circuit for Better Mixer Conversion Gain*, Philips Semiconductors applications document.
26. F. M. Fano, "Theoretical Limitations in the Broadband Matching Arbitrary Impedances," *J. Franklin Inst.*, **249**, 57–85, 139–155, 1950.
27. Ravender Goyal, ed., *High-Frequency Analog Integrated-Circuit Design*. New York: Wiley, 1995, Chap. 6.
28. Stefan Georgi and Ulrich L. Rohde, "A Novel Approach to Voltage Controlled Tuned Filters Using CAD Validation," *Microwave Eng. Eur.*, 35–42, October 1997.
29. *Analog/Interface ICs Device Data*, Vol. II, Q4/96, DL128, Rev 6, Motorola, Inc., MC13158 Wideband FM IF Subsystem specifications, pp. 8-308 to 8-329.
30. E. J. Wilkinson, "An *N*-Way Power Divider," *IRE Trans. Microwave Theory Tech.*, **MTT-8**, 116, January 1960.
31. Motorola Application Note AN-282A, *Systemizing RF Power Amplifier Design*.
32. W. E. Everitt and G. E. Anner, *Communication Engineering*. New York: McGraw-Hill, 1956.
33. G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks and Coupling Structures*. New York: McGraw-Hill, 1980.
34. G. L. Matthaei, "Tables of Chebyshev Impedance-Transforming Networks of Low-Pass Filter Form," *Proc. IEEE*, August 1964.
35. Motorola Application Note AN-267, *Matching Network Designs with Computer Solutions*.
36. E. G. Cristal, "Tables of Maximally Flat Impedance-Transforming Networks of Low-Pass Filter Form," *IEEE Trans. Microwave Theory Tech.*, **MTT 3**(5), September 1965.
37. H. W. Bode, *Network Analysis and Feedback Amplifier Design*. New York: Van Nostrand, 1943.
38. R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," *J. Franklin Inst.*, 139–154, January-February 1950.
39. J. H. Horwitz, "Design Wideband UHF-Power Amplifiers," *Electron. Des.* **11**, x–x, May 24, 1969.
40. O. Pitzalis and R. A. Gilson, "Tables of Impedance Matching Networks Which Approximate Prescribed Attenuation Versus Frequency Slopes," *IEEE Trans. Microwave Theory Tech.*, **MTT-9**(4), April 1971.
41. Anthony N. Gerkis, "Broadband Impedance Matching Using the 'Real Frequency' Network Synthesis Technique," *Appl. Microwave & Wireless*, **10**(6), 26–36, July/August 1998.
42. T. Cuthbert, "Broadband Impedance Matching Methods," *RF Des.*, **17**(8), 64–71, August 1994.
43. T. Cuthbert, *Broadband Direct-Coupled and Matching RF Networks*. Greenwood, AR, T. R. Cuthbert, 1999.
44. George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990, pp. 279–283.
45. Andy Potter, "HP RF Compiler Automates Schematic Capture and Extends Capabilities of Circuit Synthesis." *Appl. Microwave & Wireless*, **11**(6), 106–118, June 1999.
46. Aron Kain, *Final Report for an Integrated CAD Tool for High-Power, High-Efficiency RF and Microwave Amplifiers*, STTR Phase II, Contract no. F33615-95-C-174, Dept. of the Air Force, Small Business Technology Transfer (STTR) Program.
47. Kenneth K. Clarke and Donald T. Hess, *Communication Circuits: Analysis and Design.* Reading, MA: Addison-Wesley, 1971.

FURTHER READING

P. L. D. Abrie, *The Design of Impedance-Matching Networks for Radio-Frequency and Microwave Amplifiers*. Norwood, MA: Artech House, 1985.

P. L. D. Abrie and P. Rademeyer, "A Method for Evaluating and the Evaluation of the Influence of the Reverse Transfer Gain on the Transducer Power Gain of Some Microwave Transistors," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**(8), 711–713, August 1985.

A. Ambrózy, *Electronic Noise*. New York: McGraw-Hill, 1982.

W. Baechtold and M. J. O. Strutt, "Noise in Microwave Transistors," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**(9), 578–585, September 1968.

A. Cappy, "Noise Modeling and Measurement Techniques," *IEEE Trans. Microwave Theory Tech.*, **MTT-36**, 1–10, January 1988.

R. S. Carson, *High Frequency Amplifiers*. New York: Wiley, 1975.

Cezar A. A. Carioca and Paulo H. P. de Carvalho, "Computer-Aided Design of Diode Frequency Multipliers," *Appl. Microwave & Wireless*, **11**(4), 38–46, April 1999.

H. J. Carlin, "A New Approach to Gain-Bandwidth Problems," *IEEE Trans. Circuits Syst.*, **CAS-23**, 170–175, April 1977.

H. J. Carlin and P. Amstutz, "On Optimum Broad-Band Matching," *IEEE Trans. Circuits Syst.*, **CAS-23**, 401–405, April 1977.

H. J. Carlin and P. Amstutz, "On Optimum Broadband Matching," *IEEE Trans. Circuits Syst.*, **CAS-28**(5), May 1981.

H. J. Carlin and J. J. Komiak, "A New Method of Broad-Band Equalization Applied to Microwave Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**, 93–99, February 1979.

H. J. Carlin and B. Yarman, "A Simplified Real Frequency Technique Applied to Broad-Band Multistage Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-30**(12), December 1982.

H. J. Carlin and B. S. Yarman, "The Double Matching Problem: Analytic and Real-Frequency Solutions," *IEEE Trans. Circuits Syst.*, **CAS-30**(1), 15–28, January 1983.

G. Caruso and M. Sannino, "Computer-Aided Determination of Microwave Two-Port Noise Parameters," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 639–642, September 1978.

Complex Math User's Guide, Compact Software, Paterson, NJ, 1988.

Steve C. Cripps, *RF Power Amplifiers for Wireless Communications*. Norwood, MA: Artech House, 1999.

T. Cuthbert, *Circuit Design Using Personal Computers*. New York: Wiley, 1983.

J. A. Eisenberg, "Designing Amplifiers for Optimum Noise Figure," *Microwaves*, **13**, 36–41, April 1974.

R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," *J. Franklin Inst.*, **249**, 57–83, January 1960/139–155, February 1960.

H. Fukui, *Low-Noise Microwave Transistors and Amplifiers*. New York: IEEE Press, 1981.

C. Gentili, *Microwave Amplifiers and Oscillators*. New York: McGraw-Hill, 1987.

Stefan Georgi and Ulrich L. Rohde, "A Novel Approach to Voltage-Controlled Tuned Filters Using CAD Validation," *Microwave Eng. Eur.*, 35–42, October 1997.

G. Gonzales, *Microwave Transistor Amplifier Analysis and Design*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

H. A. Haus and R. B. Adler, "Optimum Noise Performance of Linear Amplifiers," *Proc. IRE*, **46**, 1517–1533, August 1958.

Aron Kain, *Final Report for Bias Dependence Noise Modeling of Heterojunction Bipolar Transistors*, USAF SBIR Phase II (PIIN), F33615-95-C-1707, November 1997. Issued by USAF/AFMC/ASC, Wright Laboratory WL/AAKE BLD 7, 2530 C ST, Wright-Patterson AFB, OH 45433-7607.

M. J. Köppen, "A Single-Stage Power Amplifier for 17 watts Output in the 470-MHz Band Using the 266BLY, Operating from 12.5 volt Supply Voltage," C.A.B. Communications, Eindhoven, Holland, Electronic Application Laboratory Report, August 8, 1970.

William B. Kuhn and Naveen K. Yanduru, "Spiral Inductor Substrate Loss Modeling in SiliconRFICs," *Microwave J.*, **11**(3), 66–81, March 1999.

J. Lange, "Noise Characterization of Linear Two-Ports of Invariant Parameters," *IEEE J. Solid-State Circuits*, **SC-2**, 37–40, June 1967.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

David Lovelace, Natalino Camilleri, and George Kannell, "Silicon MMIC Inductor Modeling for High Volume, Low Cost Applications," *Microwave J.*, **39**(8), 60–71, August 1994.

Jarek Lucek and Robbin Damen, "Designing an LNA for a CDMA Front End," *RF Des.*, **22**(2), 20–30, February 1999.

Stephen A. Maas, "FET Models for Volterra-Series Analysis," *Microwave J.*, **42**(5), 260–270, May 1999.

Donald A. McClure, KB2Z, "Transmission-Line Transformers," *Commun. Q.*, 45–58, Summer 1997.

D. J. Mellor and J. G. Linvill, "Synthesis of Interstage Networks and Prescribed Gain Versus Frequency Slopes," *IEEE Trans. Microwave Theory Tech.*, **MTT-23**(12), 1013–1020, December 1975.

Sean Mercer, "Exact Simulation of LNAs Reduces Design Cycle Time," *Appl. Microwave & Wireless*, 74–82, January 1999.

R. P. Meys, "A Wave Approach to the Noise Properties of Linear Microwave Devices," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 34–37, January 1978.

C. K. S. Miller, W. C. Daywitt, and M. G. Arthur, "Noise Standards, Measurements and ReceiverNoise Definitions," *Proc. IEEE*, **55**, June 1967.

Brian J. Minnis, *Designing Microwave Circuits by Exact Synthesis* . Boston: Artech House, 1996.

R. Minton, *Design of Large Signal VHF Transistor Power Amplifiers* , RCA Application Note SMA-36, July 1964.

R. Minton and H. C. Lee, *Frequency Multiplication Using Overlay Transistors* , RCA Application Note SMA-40, September 1965.

M. Mitama and H. Katoh, "An Improved Computational Method for Noise Parameter Measurement," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**, 612–615, June 1979.

Motorola tutorial, "The Cascode Amplifier: A Useful RF Building Block," *Microwave Prod. Digest*, 39, 59, March 1999.

T. Mukaihata, "Applications and Analysis of Noise Generation in *N*-Cascaded Mismatched Two-Port Networks," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 699–708, September 1968.

W. W. Mumford and E. H. Scheibe, *Noise Performance Factors in Communications Systems* . Dedham, MA: Horizon House, 1968.

D. E. Norton, "High Dynamic Range Transistor Amplifier Using Lossless Feedback," *Microwave J.*, **19**(5), 53–57, May 1976.

D. E. Norton, "Strong Signal Survivability Design for Lossless Feedback Amplifiers," *Proceedings of the Eleventh Annual European Microwave Conference* , 1981.

H. S. Okean and P. O. Lombardo, "Noise Performance of M/W and MM-Wave Receivers," *Microwave J.*, **16**(1), January 1973.

T. Y. Otoshi, "The Effect of Mismatched Components on Microwave Noise-Temperature Calibrations," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 675–686, September 1968.

P. Penfeld, "Wave Representations of Amplifier Noise," *IRE Trans. Circuit Theory*, 84–86, March 1962.

Ulrich L. Rohde, Jerry C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*, 2nd ed. New York: McGraw-Hill, 1997, Chap. 3.

C. Rorabaugh, *Circuit Design and Analysis—Featuring C Routines* . New York: McGraw-Hill, 1992.

M. Sannino, "On the Determination of Device Noise and Gain Parameters," *Proc. IEEE*, **67**, 1364–1366, September 1979.

SYNMAT, matching synthesis program by Radom, Inc., developed by Douglas Mellor.

A. van der Ziel, *Noise in Solid State Devices and Circuits* . New York: Wiley, 1986.

Dave Vondran, "Noise Figure Measurement: Corrections Related to Match and Gain," *Microwave J.*, **42**(3), 22–38, March 1999.

D. F. Wait, "Thermal Noise from a Passive Linear Multiport," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 687–691, September 1968.

B. S. Yarman and H. J. Carlin, "A Simplified 'Real Frequency' Technique Applied to Broad-Band Multistage Microwave Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-30**(12), 2216–2222, December 1982.

G. P. Young and S. O. Scanlan, "Matching Network Design Studies for Microwave Transistor Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**(10), 1027–1035, October 1981.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

3

AMPLIFIER DESIGN WITH BJTs AND FETs

3-1 PROPERTIES OF AMPLIFIERS

3-1-1 Introduction

The goal of this chapter is to move from the details on the semiconductor devices themselves to our first practical application. One of the more interesting features of amplifiers is that they can easily become what will be covered in Chapter 5: *oscillators*. This is due to the high gain of the devices, unaccounted for parasitic elements, and other design flaws. The amplifiers we need will fall into three categories:

- Low-noise amplifiers
- High-gain amplifiers
- Medium- to high-power amplifiers

See Figure 3-1. The low-noise amplifier always operates in Class A, typically at 15–20% of its maximum useful current. The high-gain amplifier can operate in Class A, as well as Class B (mostly push–pull). The higher dc current for the same device results in a higher noise figure and more gain, and ultimately more output power. Class C operating mode is really reserved for either FM transmissions or constant-carrier modes like CW. Some of the modern digital modulation types are sensitive to phase distortion rather than amplitude changes and because of the resulting output spectrum designers have stayed away from Class C operation.

Some of these amplifiers, such as bipolar versions, can be dc-coupled with very few difficulties; others, like those in the FET families, will cause more headaches. An interesting example, although much too high in frequency for the purpose of this book, is shown in Figures 3-2, 3-3, 3-4, and 3-5. It shows a three-stage amplifier that has noise feedback for the first two stages and resistive feedback for the output stage. These are the three types of amplifiers, regardless of technology, we will be evaluating. The achievement of accurate noise analysis of this circuit, reported in Rohde et al. [1], marked the first time a complete

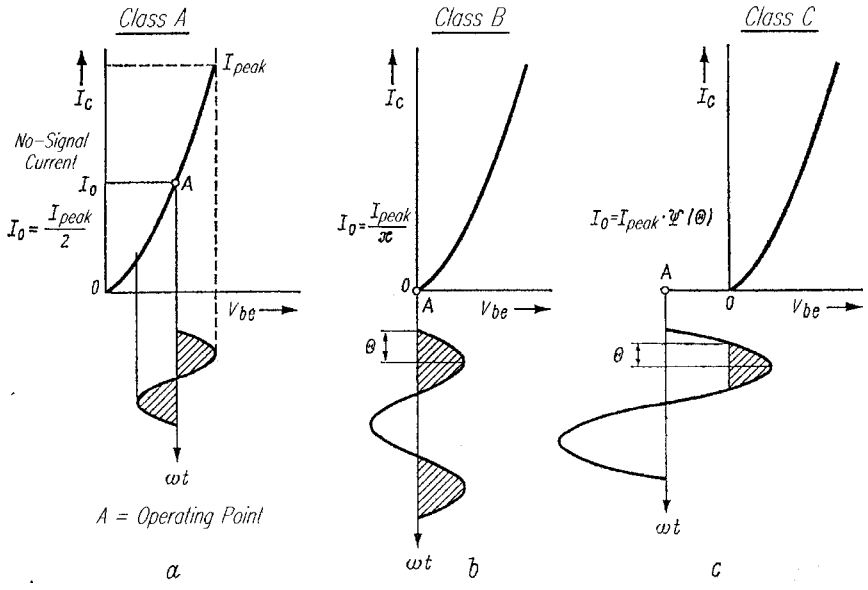


Figure 3-1 Definition of Class A, B, and C operation and resulting bias, including conduction angle (Θ, to be discussed in Section 3-2-2). The transfer characteristic can be either quadratic or exponential. This results in different distortion but does not change the basic operating mode.

linear noise model with essentially no frequency limitation had been developed. Its accuracy depends solely on the accuracy of the element values of the linear equivalent circuit, which can be obtained from measured *S*-parameters by an optimization process. Needless to say, this technique is also applicable at lower frequencies, such as 1.5–3 GHz, and its deviation from measured values has never been worse than 0.2 dB or, relative to the noise figure, 10% expressed in dB.

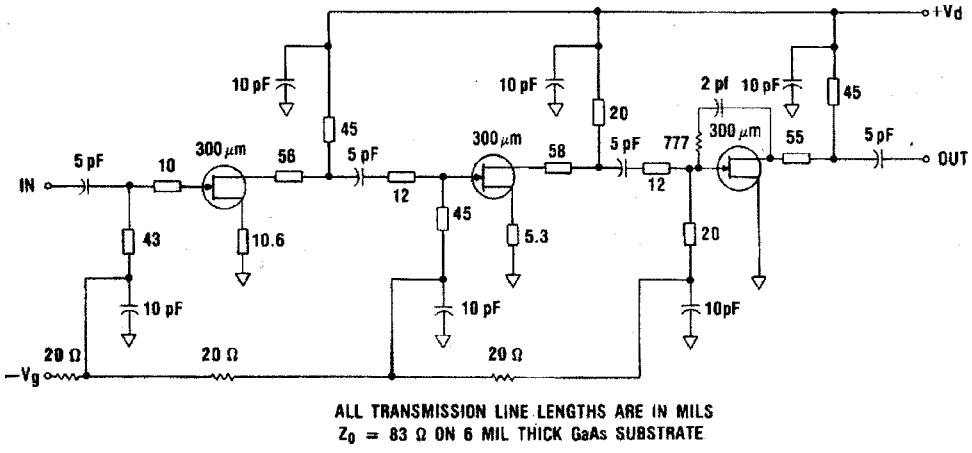


Figure 3-2 Schematic of the X-band GaAs monolithic low-noise amplifier (Texas Instruments EG8021). For reasons of linearity, all three stages operate in Class A.

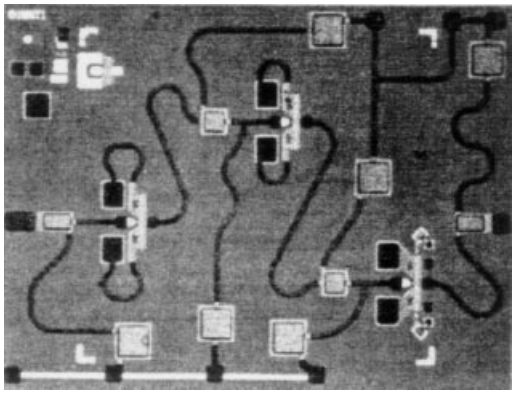


Figure 3-3 Photograph of the EG8021 monolithic amplifier chip. The area pictured is 0.09 inch by 0.12 inch in size.

The results of Figure 3-4 were obtained by using the linear FET model as supplied by Texas Instruments. It becomes somewhat obvious that a match for S_{11} of –60 dB is not likely. We then replaced the linear transistor with the SPICE-type nonlinear model, resulting in the responses shown in Figure 3-5. S_{11} now looks much more realistic in the center of the range, but at the same time exhibits a tendency toward instability at frequencies around 8 GHz. The actual circuit, when measured, did not exhibit this potential instability; on the other hand,

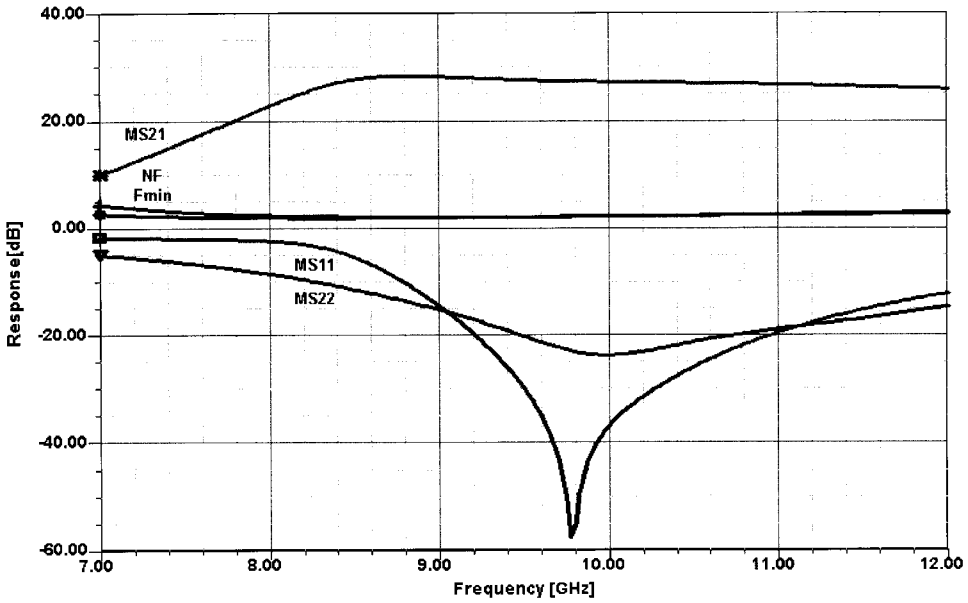


Figure 3-4 Simulated F_{min}, noise figure, S_{11}, S_{21}, and S_{22} responses for the three-stage GaAsFET amplifier using the TI linear FET model. The values at 10 GHz are: F_{min}, 2.20 dB; NF, 2.21 dB; S_{11}, –36.8 dB; S_{21}, 27.2 dB; and S_{22}, –23.7 dB.

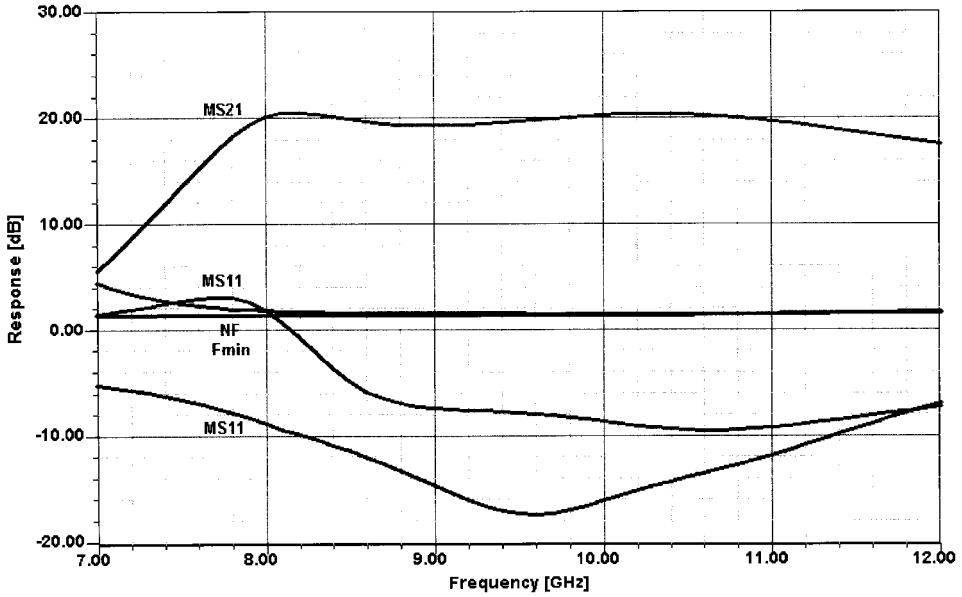


Figure 3-5 Simulated F_{min}, noise figure, S_{11}, S_{21}, and S_{22} responses for the three-stage GaAsFET amplifier using the nonlinear FET model. The values at 10 GHz are: F_{min}, 1.53 dB; NF, 1.65 dB; S_{11}, –8.5 dB; S_{21}, 20.3 dB; and S_{22}, –15.9 dB.

the noise figure and gain agreed quite well. The discrepancies between the results with the linear and nonlinear models emphasize the importance of accurate device modeling.

Figure 3-6 shows a wideband amplifier, specifically a distributed amplifier, that covers 1–20 GHz. Figure 3-7 shows its simulated frequency-dependent gain, matching, and noise performance. The example given here as the upper frequency for wireless systems is rapidly expanding and the technology to accommodate this is already in place. Modern CAD tools can handle such high frequencies accurately, which means one can trust them at the significantly lower wireless frequencies of 500 MHz to 3 GHz. Above 1500 MHz, one really should model using distributed elements. We will come to this later in this chapter.

Input/output and interstage matching will become another challenging task, specifically if one requires sufficient bandwidth. As circuits become more complex and as the operating frequency increases, passive components play an increasingly important role; the use of

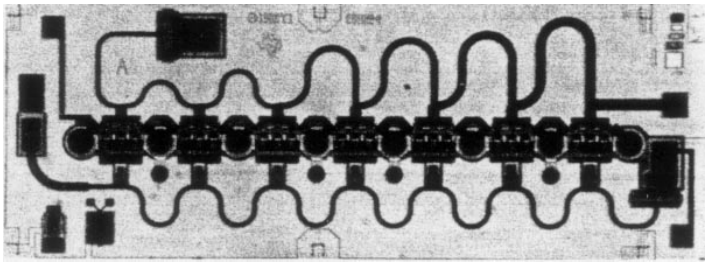


Figure 3-6 Photo of the 1–20-GHz distributed amplifier.

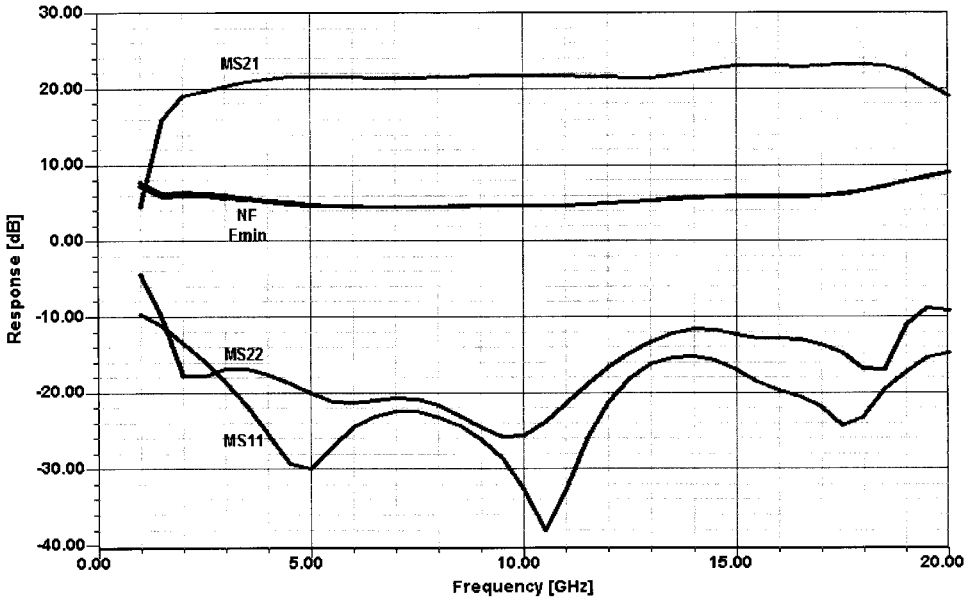


Figure 3-7 Simulated F_{min}, NF, S_{11}, S_{21}, and S_{22} responses. At the center of the range, the following numbers were found: F_{min}, 4.80 dB; NF, 4.80 dB; S_{11}, –32.8 dB; S_{21}, 21.8 dB; and S_{11}, –25.2 dB.

distributed elements is practically a necessity above 1500 MHz. For this reason, we begin this chapter by covering the application of transistors as amplifiers, following this with a discussion of the surrounding external circuitry, such as transmission lines, inductors, capacitors, and resistors. Monolithic ICs on silicon are a totally different ball game, and neither of the authors has sufficient expertise in this area. For this reason, we refer readers who will concentrate on MMIC design to the literature mentioned at the end of this chapter.

The goal of the applications presented here is to address, where possible, high-performance wireless stages that are less subject to constraints of space, power consumption, and cost, than portable/mobile applications. This is consistent with our preface, in which we have clearly stated our goal. The designers of handsets typically cannot consider the full variety of circuits we will present here. These types of stages (low noise, high gain and power) frequently will be used with filtering, and we will give some examples of input filters, including a tracking high-performance design.

A key topic besides the frequently mentioned gain is the whole noise issue. We will give the reader a complete introduction to this important area, followed by a similar introduction to the gain and matching of transistors. *Caution*: The majority of applications published go through the exercise of using *S*-parameters manually to design low-noise and small-signal amplifiers. While we have been tempted to repeat the information found in several textbooks, we feel that today's engineers will have access to CAD tools one way or another. It is highly unlikely that readers will undertake the inefficient process of going through all the design steps manually (a procedure that can take days); hopefully, engineers will resort to a linear/nonlinear CAD tool to do so. Only for the purpose of summarizing the necessary steps will we provide the relevant information regarding gain, stability, and related topics, demonstrating their usefulness or possible misuse with a modern CAD tool—in our case,

the Serenade Design Environment product from Ansoft. At this time, several additional companies beside Ansoft and Hewlett-Packard have emerged, but it is dangerous to list all of them as they come and go.

3-1-2 Gain

In practical terms, when we talk about gain we mean the output power relative to the input power. The industry standard now uses 50 Ω; however, at the point at which we have to match high-impedance devices such as monolithic crystal filters where the impedance jumps, things are not so obvious. The transformation equation that allows us to transform from one impedance to another is

$$m = \sqrt{\frac{R_2}{R_1}} \tag{3-1}$$

This correction factor, m, is required to do the transformation from R_1 (50 Ω) to R_2 (e.g., 1.2 kΩ). It needs to be understood that while the power gain remains the same, the voltage gain requires this correction factor. Since the power gain is expressed as 10 log A, A being the loop gain, 20 log A is the voltage gain, both expressed in dB. The reason for 10 versus 20 is that the power gain is always the voltage squared, which accounts for the 2. As a reminder of these relationships we present the following:

Distinction: Power Ratios–Voltage Ratios

$$n \text{ dB} = 10 \log_{10}(P_2/P_1) \quad \xrightarrow{PP = V^2/Z} \quad n = 10 \log_{10}\left(\frac{V_2^2/Z_2}{V_1^2/Z_1}\right)\text{dB}$$

$$Z_2 = Z_1 \rightarrow n \text{ dB} = 10 \log_{10}(V_2^2/V_1^2) = 10 \log_{10}(V_2/V_1)^2$$

$$\Longrightarrow \boldsymbol{n \text{ dB} = 20 \log_{10}(V_2/V_1) \text{ dB}}$$

The gain is typically defined under the condition of real resistive source and load terminations, which means that all the available power is being supplied to the amplifier, and likewise through proper matching at the output all the available power is fed to the termination. This gain figure is typically measured at a single spot frequency, but since most amplifiers are not totally flat, one needs to specify the gain tolerances or gain flatness. A good number is x dB ± 0.5 dB gain variation. In the case of a narrowband amplifier, the gain typically is shaped like a tuned circuit or like a critically coupled bandpass filter frequency response, while in the case of a real wideband application, the desired gain has some kind of ripple, which is due to the various compensation components in the circuit. An extreme case of a wideband amplifier is a 6–18-GHz single-stage amplifier that uses some very clever matching. This is shown in the following example. Figure 3-8 shows the circuit diagram, and Figure 3-9 shows its frequency-dependent performance.

While this is a cute approach for generating a single-stage amplifier over such a wide bandwidth, in real life it is not acceptable. The reason for this is that neither S_{11} nor S_{22} meets any useful specifications over the entire range. Typically, S_{11}/S_{22} values are only a few dB, while in reality more than 10 dB is required. We have chosen this example to show that gain

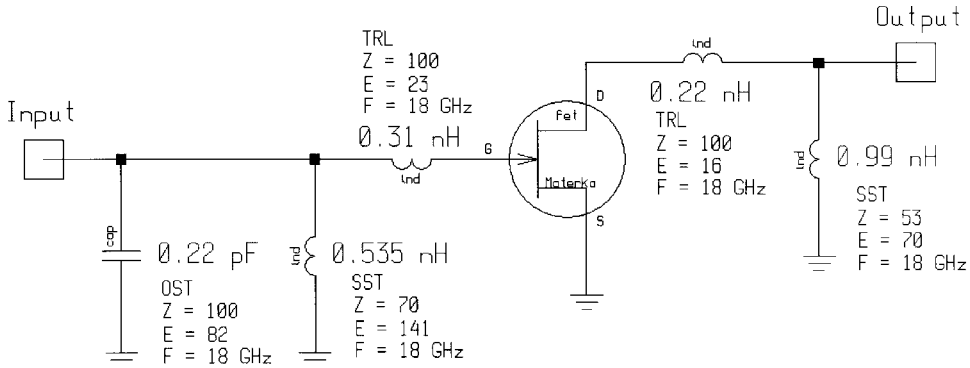


Figure 3-8 A 6–18-GHz FET amplifier.

and matching, depending on the circuit topology, are connected. We get similar results for the noise figure, whereby the best achievable noise figure ($F_{\min}$) and the actual 50-Ω noise figure are only a few dB apart; however, its frequency dependence in a wideband stage is certainly not acceptable.

Another parameter associated with gain is the stability factor K of the circuit. A value of $K > 1$ is mandatory to guarantee stability together with the S_{11} and S_{22} magnitudes being less than 1. Figure 3-10 plots the K stability factor, which for absolute stability must be >1, and maximum stable gain (MSG) as well as $G_{\max}$. It turns out that MSG is a highly artificial

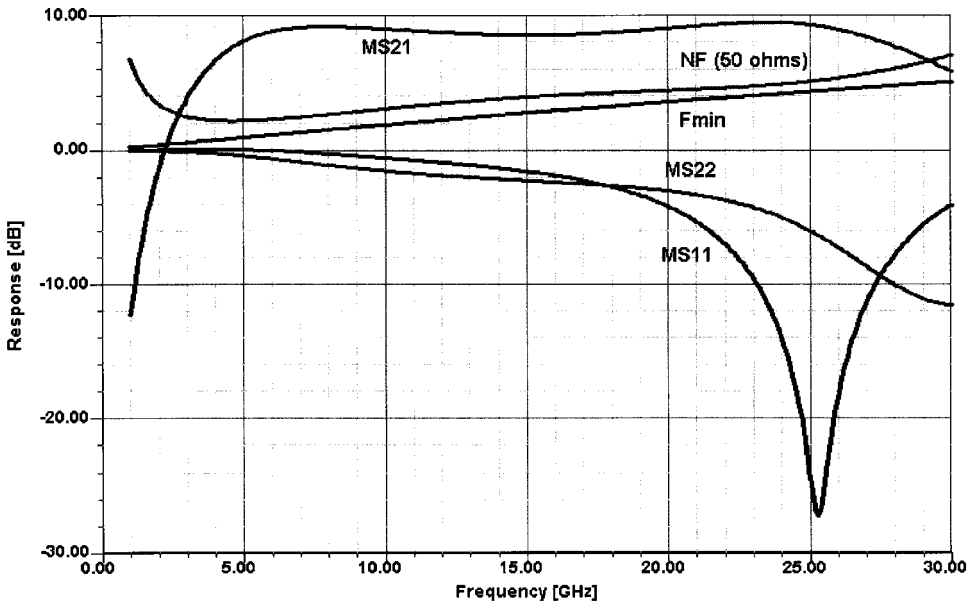


Figure 3-9 Frequency-dependent gain, matching, and noise performance of the amplifier shown in Figure 3-8.

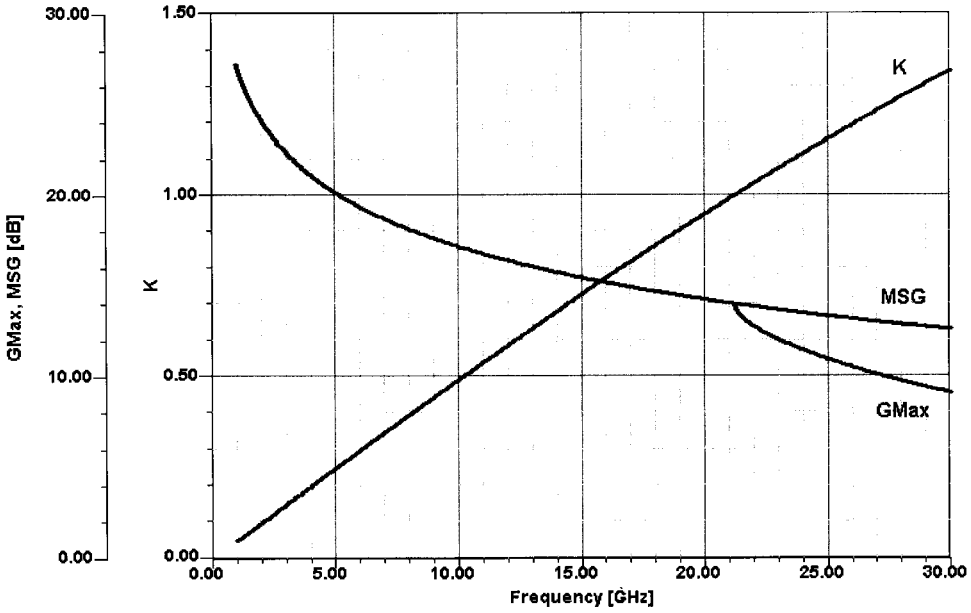


Figure 3-10 G_{max}, maximum stable gain (MSG), and K performance of the amplifier in Figure 3-8.

definition of stability that implies that input and output are terminated with pure resistances and then conjugately matched. There are four ways of achieving this resistive loading, either in the emitter as a shunt resistor between gate/drain or base/collector, or in parallel to the input and output to ground. Needless to say, this is not a real operating mode because it deteriorates both noise and gain. In our case, the entity G_{max} is a function of S_{12} of the circuit, not just the transistor, and depends highly on S_{12}. If we bring the system close to oscillation, then G_{max} will increase at the expense of bandwidth. Unfortunately, some of these microwave definitions of gain (see Table 3-1) have become widely used in industry without always being useful. Evaluating these equations for S_{12} is not legitimate because other quantities like K become infinite, and the entire system of equations falls apart. The power gain with input and output conjugately matched, or the transducer power gain for arbitrary source and load matching, still give more insight into the system.

In the above example, we specifically stated that the input and output VSWR was totally unacceptable, and this was due to a one-stage design. Figure 3-11 shows an amplifier chip built around two transistors with standard feedback techniques applied; Figures 3-12 and 3-13 show its frequency-dependent gain, matching, stability, and noise performance. This amplifier was one of the early examples of GaAsFET wideband amplifiers, but it is interesting because while it has wild (and wide) gain swings, it exhibits a fairly constant input and output impedance or reflection coefficient. A further evaluation of the circuit shows that not only is the circuit unconditionally stable, but that the maximum stable gain has an extremely high value based on essentially minimal S_{12}, the K factor is always significantly higher than 1, and G_{max} follows, to some degree, the actual gain response of the amplifier. (Ultimately, designers gave up on simple *RCL* feedback amplifiers for such a wide bandwidth, replacing them with distributed amplifiers previously called traveling wave amplifiers.) The problem of gain and stability starts at audiofrequency amplifiers, where the

Table 3-1 Nine power gains

Transducer power gain in 50-Ω system	$G_T = \lvert S_{21}\rvert^2$
Transducer power gain for arbitrary Γ_G and Γ_L	$G_T = \dfrac{(1-\lvert\Gamma_G\rvert^2)\,\lvert S_{21}\rvert^2\,(1-\lvert\Gamma_L\rvert^2)}{\lvert(1-S_{11}\Gamma_G)\,(1-S_{22}\Gamma_L)-S_{12}S_{21}\Gamma_G\Gamma_L\rvert^2}$
Unilateral transducer power gain	$G_{\mathrm{TU}} = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_G\rvert^2)\,(1-\lvert\Gamma_L\rvert^2)}{\lvert 1-S_{11}\Gamma_G\rvert^2\lvert 1-S_{22}\Gamma_L\rvert^2}$
Power gain with input conjugate matched	$G = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_L\rvert^2)}{\lvert 1-S_{22}\Gamma_L\rvert^2(1-\lvert S'_{11}\rvert^2)} = \dfrac{\lvert S_{21}\rvert^2}{1-\lvert S_{11}\rvert^2}$ (for $\Gamma_L = 0$)
Available power gain with output conjugate matched	$G_A = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_G\rvert^2)}{\lvert 1-S_{11}\Gamma_G\rvert^2(1-\lvert S'_{22}\rvert^2)} = \dfrac{\lvert S_{21}\rvert^2}{1-\lvert S_{22}\rvert^2}$ (for $\Gamma_G = 0$)
Maximum available power gain[a]	$G_{ma} = \left\lvert\dfrac{S_{21}}{S_{12}}\right\rvert\,(k-\sqrt{k^2-1})$
Maximum unilateral transducer power gain	$G_{\mathrm{TUmax}} = \dfrac{\lvert S_{21}\rvert^2}{(1-\lvert S_{11}\rvert^2)(1-\lvert S_{22}\rvert^2)}$
Maximum stable power gain	$G_{ms} = \dfrac{\lvert S_{21}\rvert}{\lvert S_{12}\rvert}$
Unilateral power gain	$U = \dfrac{1/2\lvert S_{21}/S_{12}-1\rvert^2}{k\lvert S_{21}/S_{12}\rvert-\mathrm{Re}(S_{21}/S_{12})}$

[a]Definition for $G_{\max}$: If $k > 1$, it is called G_{ma}, and if $k < 1$, it is called G_{ms}.

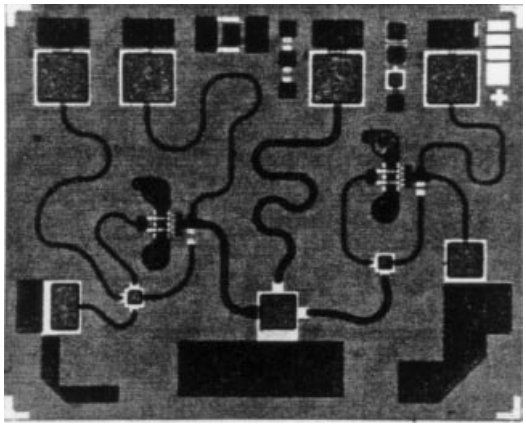


Figure 3-11 A two-stage, wideband feedback amplifier.

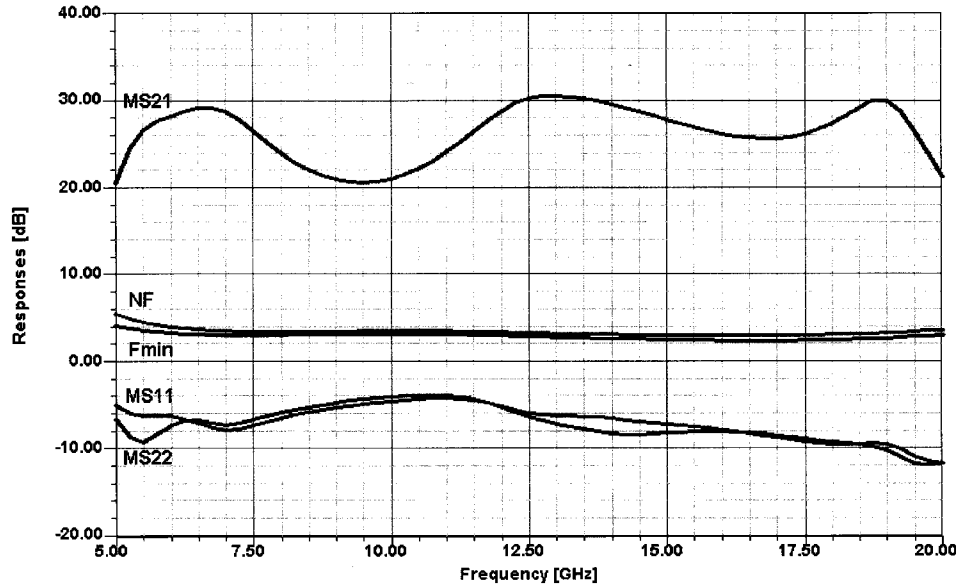


Figure 3-12 Frequency-dependent gain, matching, and noise performance for the amplifier in Figure 3-11.

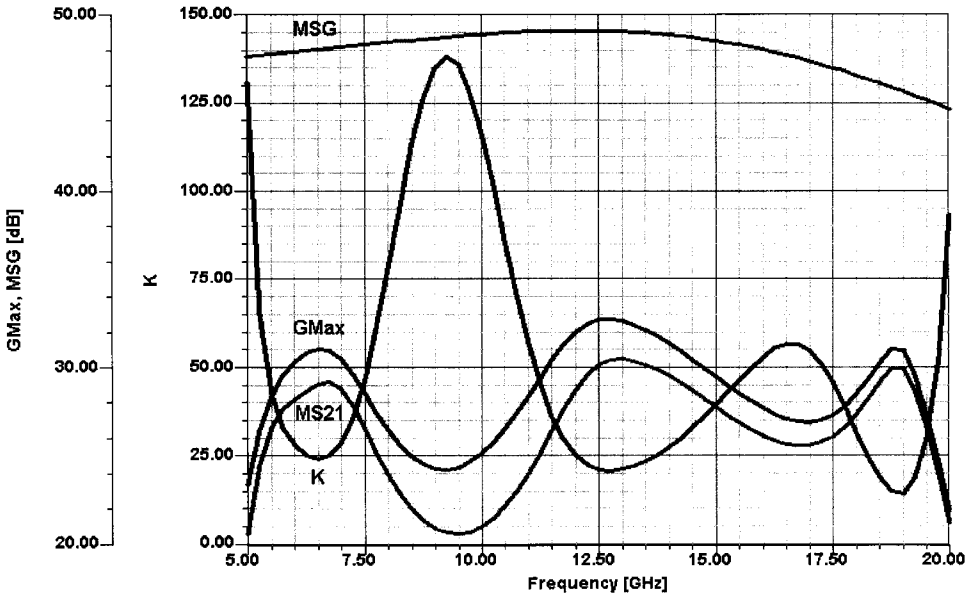


Figure 3-13 Frequency-dependent G_{max}, K, maximum stable gain, and gain performance for the amplifier in Figure 3-11.

motorboating phenomenon has been known for years, and the problem moves up in frequency with more modern technology and is largely dependent not only on the device and the circuit, but also on the layout. The motorboating effect has to do with ground loops and the fact that if the printed circuit traces from the power supply connection to the output power stage present too high a resistance, then at high currents the voltage will drop to a point below useful operation and then, as the volume is lowered, become stable again. In this case the amplitude of the signal through the amplifier controls the stability of the output stage. Needless to say, this most unpleasant effect applies to high-power stages for all output transistors, where we use Class B or C operation, meaning that the dc operating current as a function of output does not remain constant.

In terms of gain, there is also an important gain characteristic called *differential gain* and *differential phase*. These expressions were first used in television circuits, where linearity in both areas is an absolute must. The amplitude of the signal determines the black and white contrast, and the phase determines the color. As an example, transmitting an image of one of our presidents playing golf in California to the New York area is done by many microwave hops. Any multipath reception, such as that caused by reflection from airplanes, may have little effect on the signal amplitude, but the resulting phase shift can transform a green lawn to a red lawn if a total shift of 180° occurs. Given a particular operating point, if the 1-dB compression point of multiple amplifiers in the chain is approached or exceeded, their differential gain, which is defined by the IEEE as "the difference between (a) the ratio of the output of a small high-frequency sine-wave signal at two different levels of a low-frequency signal on which it is superimposed, and (b) unity," may increase. This mode of saturation charges all the capacitances of the amplifier transistors, causing huge phase shifts that result in the known problem of changing colors. The same applies, of course, in test equipment amplifiers, in which such changes cannot be tolerated. A somewhat similar, but not quite the same, problem is spectral regrowth [evaluated in terms of adjacent-channel power ratio (ACPR)], which is also due to nonlinearities. The question of linearity has already been covered in detail in Chapter 1. Needless to say, any overdrive condition also changes an amplifier's input and output SWR.

3-1-3 Noise Figure (NF)*

Introduction. In this section we look at the best way to describe noise in active devices (black boxes). Even when a two-port is linear, the output waveform will differ from the input, because of the failure to transmit all spectral components with equal gain (or attenuation) and delay. By careful design of the two-port, or by limitation of the bandwidth of the input waveform, such distortions can largely be avoided. However, noise generated within the two-port can still change the waveform of the output signal. In a linear passive two-port, noise arises only from the losses in the two-port; thermodynamic considerations indicate that such losses result in the random changes that we call noise. When the two-port contains active devices, such as transistors, there are other noise mechanisms present. A very important consideration in a system is the amount of noise that it adds to the transmitted signal. This is often judged by the ratio of the output signal power to the output noise power

*Section 3-1-3 is reprinted from Chapter 2 of George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*, copyright 1990 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

(S/N). The ratio of signal plus noise power to noise power [$(S + N)/N$] is generally easier to measure and approaches S/N when the signal is large.

In the evaluation of a two-port it is important to know the amount of noise added to a signal passing through it. An important parameter for expressing this characteristic is the noise factor. The signal energy coming from a generator or antenna is amplified or attenuated in passing from the input to the output of a two-port, as is the noise that accompanies the input signal energy. A system generally includes a cascade of two-port networks that constitute one overall two-port that amplifies the signal to a high-enough power level for its intended use. The *noise factor* of a system is defined as the ratio of signal-to-noise ratios available at input and output:

$$F = \frac{(S/N)_{\text{input}}}{(S/N)_{\text{output}}} \geq 1 \tag{3-2}$$

The noise figure (or factor) of a receiver is an easily measured quantity that describes the signal-to-noise ratio reduction of that receiver.

When this ratio of powers is converted to decibels, it is generally referred to as the *noise figure* rather than noise factor. Various conventions are used to distinguish the symbols used for noise factor and noise figure. Here we use F to represent the noise factor and NF to represent the noise figure, although the terms are usually used interchangeably.

For an amplifier with the power gain G, the noise factor can be rearranged as

$$F = \frac{S_i/N_i}{GS_i/G(N_i + N_a)} \tag{3-3}$$

when N_a is the additional noise power added by the amplifier referred to the input. This can be computed to be

$$F = 1 + N_a/N_i \tag{3-4}$$

The noise factor is often replaced by the noise figure (NF), which is defined in decibels as

$$\text{NF} = 10 \log_{10} F \tag{3-5}$$

In applications like satellite receivers, the noise factor becomes such a small number that it is inconvenient to work with. Many people have adopted the use of an equivalent noise temperature for a circuit to remedy this situation. Since the thermal noise power available from a resistor at temperature T_e is

$$N = kT_eB \tag{3-6}$$

where k is Boltzmann's constant (1.38×10^{-23} J/K), T_e is the effective temperature in kelvins, and B is the bandwidth in hertz. The equation above may be used to associate an effective noise temperature with circuits containing more than just thermal noise sources. This allows Eq. (3-4) to be written as

$$F = 1 + \frac{kT_eB}{kT_0B} = 1 + \frac{T_e}{T_0} \tag{3-7}$$

where T_e is the effective noise temperature of the circuit and T_0 is the temperature of the generator resistor in kelvins. The noise temperature T_e now characterizes our circuit noise contribution and can directly be related to the noise factor.

Assuming a reference noise temperature of 290 K (–273 + 290 = 17 °C), let us determine the noise temperature of the system with a noise factor of 2.6 (4.15 dB):

$$T_e = (2.6 - 1)\ (290) = 464 \text{ K}$$

This temperature T_e should not be confused with the environmental operating temperature T_0. It is quite common to operate low-noise amplifiers with T_e below 100 K at an ambient temperature of 290 K.

Signal-to-Noise Ratio. Let us consider the signal-to-noise ratio of power delivered from a generator to a load as shown in Figure 3-14. The signal power delivered to the input is given by

$$S_{\text{in}} = P_{\text{in}} = \frac{E_g^2 \text{Re}(Z_{\text{in}})}{|Z_g + Z_{\text{in}}|^2} \tag{3-8}$$

where E_g is the rms voltage of the input signal supplied to the system, and the noise power supplied to the input is expressed by

$$N_{\text{in}} = \frac{\overline{v_n^2} \text{Re}(Z_{\text{in}})}{|Z_g + Z_{\text{in}}|^2} \tag{3-9}$$

where the noise power at the input is provided by the noise energy of the real part of Z_g. The input impedance Z of the system in the form $Z = R_{\text{in}} + jX_{\text{in}}$ is assumed to be complex.

The Johnson noise of a resistor [here $\text{Re}(Z_g)$], given by the mean-square voltage

$$\overline{v_n^2} = 4kTRB \tag{3-10}$$

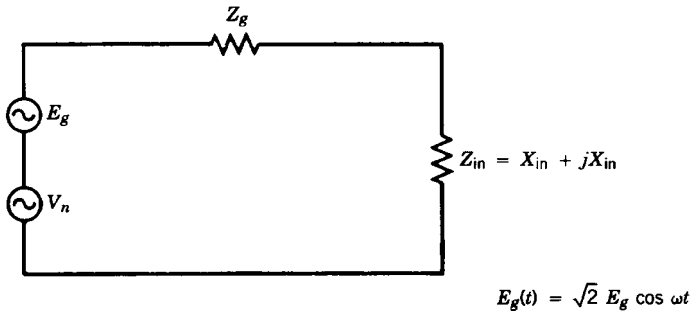


Figure 3-14 Combination of signal and noise voltages supplied to a complex termination.

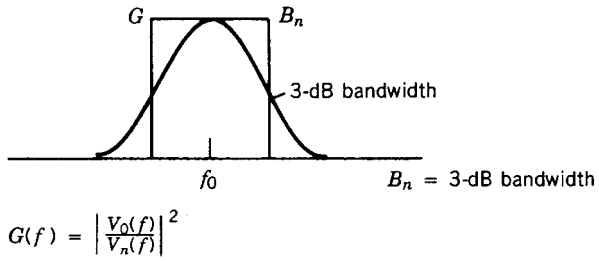
where k (Boltzmann's constant) = 1.38×10^{-23} J/K, T is the absolute temperature of the resistor, and B is the bandwidth, is sufficiently small that the resistive component of impedance does not change. The available signal power from the generator has a lower limit, even if the signal is attenuated by the highest possible attenuation. The generator resistor acts as a Johnson noise generator, its power being

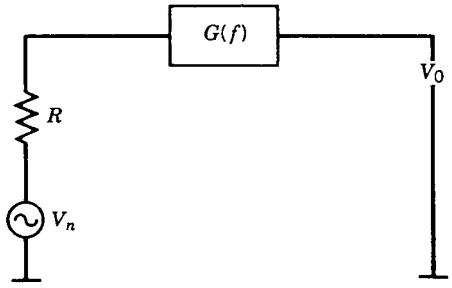
$$P_A = \frac{4kTRB}{4R} = kTB \tag{3-11}$$

where k is Boltzmann's constant, T is the absolute temperature, and B is the bandwidth. This power is the maximum available output power.

For an ambient temperature of 290 K, $kT = 4 \times 10^{-21}$ W/Hz. This expression is also given as kT = –204 dBW/Hz = –174 dBm/Hz = –114 dBm/MHz. We can combine Eqs. (3-8) to (3-10) to obtain

$$\left(\frac{S}{N}\right)_{\text{in}} = \frac{E_g^2}{4kT\text{Re}(Z_g)B} \tag{3-12}$$





$$V_0^2 = \int_0^\infty 4kT_0RG(f)\,dF$$
$$= 4kT_0R \int_0^\infty G(f)\,dF$$
$$B_n = \frac{1}{G} \int_0^\infty G(f)\,dF$$
B_n = noise bandwidth

Figure 3-15 Graphical and mathematical explanations of the noise bandwidth from a comparison of the Gaussian-shaped bandwidth to the rectangular filter response.

This is the value of *S/N* contributed by the generator, which does not include the noise generated by the load, in this case $\mathrm{Re}(Z_{\mathrm{in}})$, which would need to be included in the measurement of the total *S/N* across the input impedance.

A critical parameter is the noise bandwidth, B_n, which is defined as the equivalent bandwidth, as shown in Figure 3-15. For reasons of group delay correction, most practical filters have round rather than sharp corners. The noise figure measurements shown later can be used to determine the "integrated" bandwidth, which is B_n.

An active system such as a combination of amplifiers and mixers will add noise to the input signals, and the noise factor that describes this is defined as the *S/N* ratio at the input to the *S/N* ratio at the output, which is always greater than unity [2]. In practice, a certain minimum signal-to-noise ratio is required for operation. For example, in a communication system such a minimum is required for intelligible transmission, either voice or data. For high-performance TV reception, to provide a noise-free picture to the eye, a typical requirement is for a 60-dB *S/N*. In the case of a TV system, a large dynamic range is required, as well as a very large bandwidth to reproduce all colors truthfully and all shades from high-intensity white to black. Good systems will have a bandwidth of 8 MHz or more.

Noise Figure Measurements. Some of the noise equations are based on mathematical models and physics. To understand some of these expressions, it is useful to look at a practical system, with amplifiers, that has to be evaluated.

Let us look at Figure 3-16, which consists of a signal generator, the system or device under test (DUT), and a selective receiver with a built-in root-mean-square (rms) voltmeter to determine the signal and noise voltages. It is necessary that the system have enough gain so that the noise voltage supplied by the generator will be indicated [3].

If we assume that our selective receiver is a video noise meter calibrated in rms voltage levels, we can perform two measurements. With an input termination connected to the TV system (typically, 75 Ω for cable TV, 50 Ω for satellite TV), the noise receiver/meter will read a value for proper termination that can easily be calculated. Since one-half of the mean-square noise voltage appears across the input,

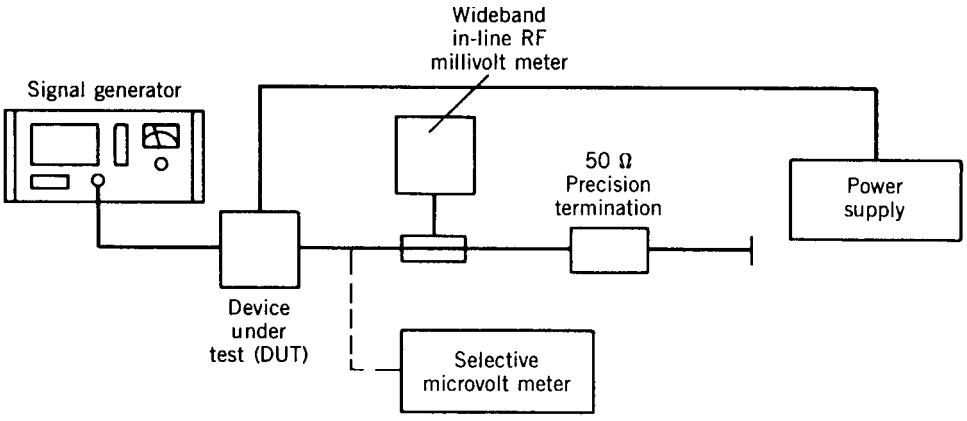


Figure 3-16 Test setup to measure signal-to-noise ratio.

$$v_{\text{in}} = \frac{v_n}{2} = \frac{\sqrt{4kTRB}}{2} \tag{3-13}$$

With $B = 10$ MHz, $T = 290$ K (T is always expressed in absolute temperature; $T_0 = -273$ °C), and $k = 1.38 \times 10^{-23}$ J/K, then for $R = 75\ \Omega$,

$$v_{\text{in}} = \frac{v_n}{2} = 1.73\ \mu\text{V} \tag{3-14}$$

where the rms noise voltage has been referred to the input port. We can verify this with our first measurement.

Now we increase the input voltage of the signal generator to a value that indicates a 60-dB *S/N* ratio at the output port. This should be about

$$E_g = \frac{v_n}{2}\sqrt{F} \times 1000 = 1.73\sqrt{F} \quad \text{mV}$$

where F is the noise factor of the receiver. For a receiver noise factor of 10, we would obtain $E_g = 5.48$ mV (rms value). If the noise energy equivalent to a noise factor of F is assumed, we need $\sqrt{F}$ times more voltage. For a 60-dB ratio, this means that $E_g = 1000 \times (v_n/2) \times \sqrt{F}$.

As they are done here, over a power range of 60 dB, the measurements can be performed over such a wide range only if special equipment is available. In cases where the internal detector of a piece of communication equipment is used, the signal-to-noise ratio measurements are performed over much smaller power ranges.

Let us assume that for the above-mentioned case ($F = 10$) we find a *S/N* ratio of 10 dB at the output for an input signal of 5.47 μV. By rewriting Eq. (3-7) as

$$E_g = \frac{v_n}{2}\sqrt{F} = \sqrt{KTRBF} \tag{3-15}$$

with F being the noise factor, we can solve for F with

$$F = \frac{P_s}{P_n} = \frac{E_g^2/R}{kTB} \tag{3-16}$$

While the input power from the thermal energy of the input termination resistor was $kTB = 4 \times 10^{-4}$ W, the input power required for the 10-dB *S/N* ratio was

$$P_s = \frac{(5.47 \times 10^{-6})^2}{75} = 3.98 \times 10^{-13}\ \text{W} \tag{3-17}$$

The noise factor is defined as the ratio P_s/P_n:

$$F = \frac{3.98 \times 10^{-13}}{4 \times 10^{-14}} = 10 \tag{3-18}$$

which is the proof.

This method is used more frequently at the 3-dB point, or double the input power if the dynamic range of the detector is small or only a linear indicator is available. Because of hum and other pickup, this is not an easy measurement. Using a signal generator is very expensive because in a laboratory or production environment a wide frequency range requires several generators.

Another method is the use of a wideband noise generator. Modern gas discharge diodes or avalanche diodes are available that provide essentially white noise energy over a large frequency range. These microwave diodes typically have an output of 30 dB above kT when switched on and kT when switched off. To provide good matching at microwave frequencies, a 15-dB attenuator is cascaded. This means that the noise power of the source in the ON condition is about 15 dB above kT.

In the early 1960s, low-cost noise figure test equipment was built around vacuum diodes whose operating range was limited to 1200 MHz due to the resonance effects of the structure. Today, automatic noise gain analyzers use calibrated solid-state noise sources up to 26.5 GHz. It appears that the upper frequency limit has to do with matching, and the lower-frequency limit with $1/f$ noise.

Noisy Two-Port Description. Based on the convention by Rothe and Dalke [4], any linear two-port can be in the form shown in Figure 3-17. This general case of a noisy two-port can be redrawn showing noise sources at the input and at the output. Figure 3-17b shows this in admittance form and Figure 3-17c in impedance form. The internal noise sources are assumed to produce very small currents and voltages, and we assume that linear two-port equations are valid. The internal noise contributions have been expressed by using external noise sources:

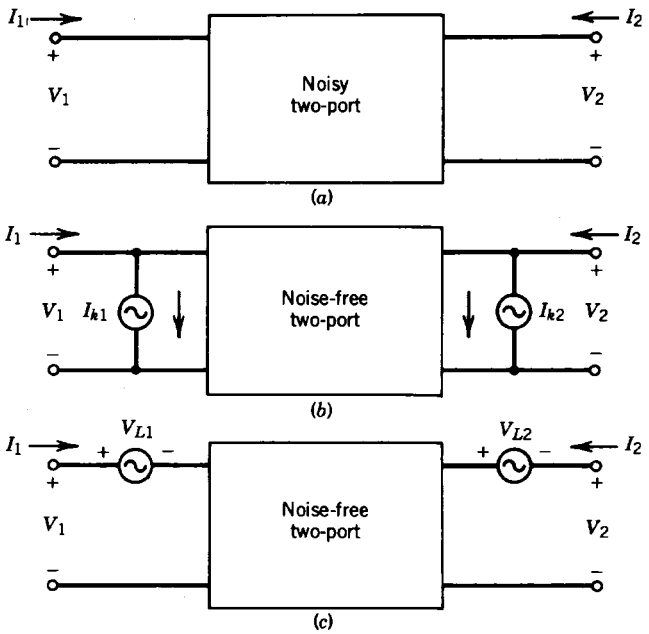


Figure 3-17 Noisy linear two-ports: (a) general form, (b) admittance form, and (c) impedance form.

$$I_1 = y_{11}V_1 + y_{12}V_2 + I_{K1}$$
$$I_2 = y_{21}V_1 + y_{22}V_2 + I_{K2} \tag{3-19}$$

$$V_1 = z_{11}I_1 + z_{12}I_2 + V_{L1}$$
$$V_2 = z_{21}I_1 + z_{22}I_2 + V_{L2} \tag{3-20}$$

where the external noise sources are I_{K1}, I_{K2}, V_{L1}, and V_{L2}.

Since we want to describe our noisy circuit in terms of the noise figure, the *ABCD*-matrix description will be more convenient since it refers both noise sources to the input of the two-port [5]. This representation is given below (note the change in direction of I_2):

$$V_1 = AV_2 + BI_2 + V_A$$
$$I_1 = CV_2 + DI_2 + I_A \tag{3-21}$$

where V_A and I_A are the external noise sources.

It is important to remember that all of these matrix representations are interrelated. For example, the noise sources for the *ABCD*-matrix description can be obtained from the *z*-matrix representation shown in Eq. (3-20). This transformation is

$$V_A = -\frac{I_{K2}}{y_{21}} = V_{L1} - \frac{V_{L2}z_{11}}{z_{21}} \tag{3-22}$$

$$I_A = I_{K1} - \frac{I_{K2}y_{11}}{y_{21}} = -\frac{V_{L2}}{z_{21}} \tag{3-23}$$

The *ABCD* representation is particularly useful because it allows us to define a noise temperature for the two-port referenced to the input. The two-port itself (shown in Figure 3-18) is assumed to be noise-free.

In the past, *z* and *y* parameters have been used, but in microwave applications it has become common to use *S*-parameter definitions. This is shown in Figure 3-19. The previous equations can be rewritten in their new form using *S*-parameters:

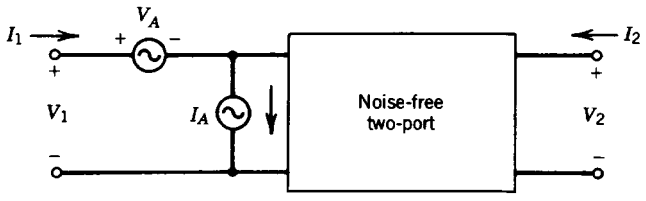


Figure 3-18 Chain-matrix form of linear noisy two-ports.

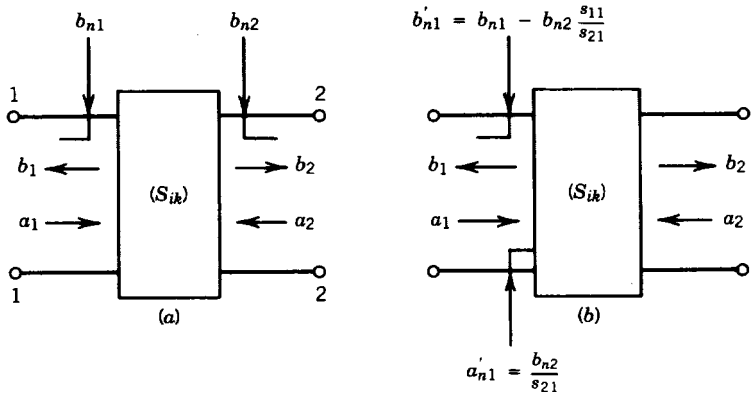


Figure 3-19 *S*-parameter form of linear noisy two-ports.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_{n1} \\ b_{n2} \end{bmatrix} \tag{3-24}$$

There are different physical origins for the various sources of noise. Typically, thermal noise is generated by resistances and loss in the circuit or transistor, whereas shot noise is generated by current flowing through semiconductor junctions and vacuum tubes. Since these many sources of noise are represented by only two noise sources at the device input, the two equivalent noise sources are often a complicated combination of the circuit internal noise sources. Often, some fraction of V_A and I_A is related to the same noise source. This means that V_A and I_A are not independent in general. Before we can use V_A and I_A to calculate the noise figure of the two-port, we must calculate the correlation between the V_A and I_A shown in Figure 3-18.

The noise source V_A represents all the device noise referred to the input when the generator impedance is zero; that is, when the input is short-circuited. The noise source I_A represents all the device noise referred to the input when the generator admittance is zero; that is, the input is open-circuited.

The correlation of these two noise sources considerably complicates analysis. By defining a correlation admittance, we can simplify the mathematics and get some physical intuition for the relationship between noise figure and generator admittance. Since some fraction of I_A will be correlated with V_A, we split I_A into correlated and uncorrelated parts as follows:

$$I_A = I_n + I_u \tag{3-25}$$

I_u is the part of I_A uncorrelated with V_A. Since I_n is correlated with V_A, we can say that I_n is proportional to V_A and the constant of proportionality is the correlation admittance:

$$I_n = Y_{\text{cor}} V_A \tag{3-26}$$

This leads us to

$$I_A = Y_{\text{cor}} V_A + I_u \tag{3-27}$$

The following derivation of noise figure will use the correlation admittance. Y_{cor} is not a physical component located somewhere in the circuit. Y_{cor} is a complex number derived by correlating the random variables I_A and V_A. To calculate Y_{cor}, we multiply each side of Eq. (3-27) by V_A^* and average the result. This gives

$$\overline{V_A^* I_A} = Y_{\text{cor}} \overline{V_A^2} \tag{3-28}$$

where the I_u term averaged to zero since it was uncorrelated with V_A. The correlation admittance is thus given by

$$Y_{\text{cor}} = \frac{\overline{V_A^* I_A}}{V_A^2} \tag{3-29}$$

Often, people use the term "correlation coefficient." This normalized quantity is defined as

$$c = \frac{\overline{V_A^* I_A}}{\sqrt{\overline{V_A^2} I_A^2}} = Y_{\text{cor}} \sqrt{\frac{\overline{V_A^2}}{I_A^2}} \tag{3-30}$$

Note that the dual of this admittance description is the impedance description. Thus, the impedance representation has the same equations as above with Y replaced by Z, I replaced by V, and V replaced by I.

V_A and I_A represent internal noise sources in the form of a voltage source acting in series with the input voltage and a source of current flowing in parallel with the input current. This representation conveniently leads to the four noise parameters needed to describe the noise performance of the two-port. Again using the Nyquist formula, the open-circuit voltage of a resistor at the temperature T is

$$\overline{V_A^2} = 4kTRB \tag{3-31}$$

This voltage is a mean-square fluctuation (or spectral density). It is the method used to calculate the noise identity. We could also define a noise equivalent resistance for a noise voltage as

$$R_n = \frac{\overline{V_A^2}}{4kTB} \tag{3-32}$$

The resistor R_n is not a physical resistor but can be used to simulate different portions of the noise equivalent circuit.

In a similar manner a mean-square current fluctuation can be represented in terms of an equivalent noise conductance G_n, which is defined by

$$G_n = \frac{\overline{I_A^2}}{4kTB} \tag{3-33}$$

and

$$G_u = \frac{\overline{I_u^2}}{4kTB} \tag{3-34}$$

for the case of the uncorrelated noise component. The input generator to the two-port has a similar contribution:

$$G_G = \frac{I_G^2}{4kTB} \tag{3-35}$$

with Y_G being the generator admittance and G_G being the real part. With the definition of F above, we can write

$$F = 1 + \left| \frac{I_A + Y_G V_A}{I_G} \right|^2 \tag{3-36}$$

The use of the voltage V_A and the current I_A has allowed us to combine all the effects of the internal noise sources.

We can use the previously defined [Eq. (3-29)] correlation admittance, $Y_{\text{cor}} = G_{\text{cor}} + jB_{\text{cor}}$, to simplify Eq. (3-36). First, we determine the total noise current:

$$\overline{I_A^2} = 4kT(|Y_{\text{cor}}|^2 R_n + G_u)B \tag{3-37}$$

where R_n and G_u are defined in Eqs. (3-32) and (3-34). The noise factor can now be determined:

$$F = 1 + \frac{G_u}{G_g} + \frac{R_n}{G_g}[(G_G + G_{\text{cor}})^2 + (B_G + B_{\text{cor}})^2] \tag{3-38}$$

$$F = 1 + \frac{R_u}{R_g} + \frac{G_n}{R_g}[(R_G + R_{\text{cor}})^2 + (X_G + X_{\text{cor}})^2] \tag{3-39}$$

The noise factor is a function of various elements, and the optimum impedance for the best noise figure can be determined by minimizing F with respect to generator reactance and resistance. This gives

$$R_{0n} = \sqrt{\frac{R_u}{G_n} + R_{\text{cor}}^2} \tag{3-40}$$

$$X_{0n} = -X_{\text{cor}} \tag{3-41}$$

and

$$F_{\min} = 1 + 2G_n R_{\text{cor}} + 2\sqrt{R_u G_n + (G_n R_{\text{cor}})^2} \tag{3-42}$$

(To distinguish between optimum noise and optimum power, we have introduced the convention 0*n* instead of the more familiar abbreviation *opt*.) At this point we see that the optimum condition for minimum noise figure is not a conjugate power match at the input port. We can explain this by recognizing that the noise source V_A and I_A represent all the

two-port noise, not just the thermal noise of the input port. We should observe that the optimum generator susceptance, $-X_{corr}$, will minimize the noise contribution of the two noise generators.

In rearranging for conversion to S parameters, we write

$$F = F_{\min} + \frac{g_n}{R_G}|Z_G - Z_{0n}|^2 \tag{3-43}$$

$$F = F_{\min} + \frac{R_n}{G_G}|Y_G - Y_{0n}|^2 \tag{3-44}$$

From the definition of the reflection coefficient,

$$\Gamma_G = \frac{Y_0 - Y_G}{Y_0 + Y_G} \tag{3-45}$$

and with

$$g_G = \frac{G_G}{Y_0} \tag{3-46}$$

$$r_n = \frac{R_n}{Z_0} \tag{3-47}$$

the normalized equivalent noise resistance,

$$F = F_{\min} + \frac{4r_n|\Gamma_G - \Gamma_{0n}|^2}{g_G(1 - |\Gamma|^2)|1 + \Gamma_{0n}|^2} \tag{3-48}$$

$$r_n = (F_{50} - F_{\min})\frac{|1 + \Gamma_{0n}|^2}{4|\Gamma_{0n}|^2} \tag{3-49}$$

$$\Gamma_{0n} = \frac{Z_{0n} - Z_0}{Z_{0n} + Z_0} \tag{3-50}$$

The noise performance of any linear two-port can now be determined if the values of the four noise parameters, $F_{\min}$, r_n (= $R_n/50$), and Γ_{0n} (magnitude and phase), are known.

Figure 3-20 shows the noise factor of a high-frequency transistor as a function of B_g for G_g = constant and as a function of G_g for $B_g = B_{opt}$.

Noise Figure of Cascaded Networks. In a system with many circuits connected in cascade (Figure 3-21), we must consider the contributions of the various circuits. In considering the equivalent noise resistor R_n in series with the input circuit,

$$F = \frac{R_G + R_n}{R_G} \tag{3-51}$$

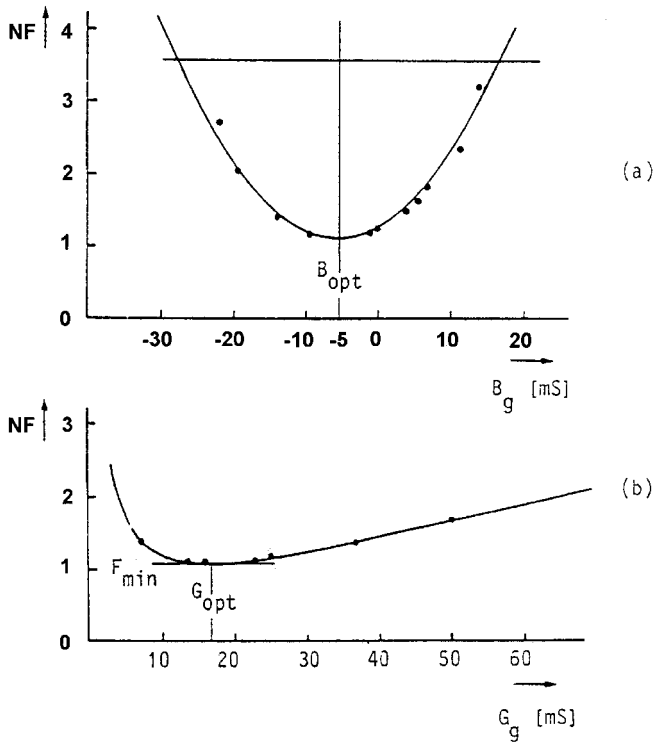


Figure 3-20 Noise factor in high-frequency BJTs for $f = 600$ MHz: (a) as a function of B_g for G_g = constant and (b) as a function of G_g for $B_g = B_{opt}$.

$$F = 1 + \frac{R_n}{R_G} \tag{3-52}$$

The excess noise added by the circuit is R_n/R_G.

In considering two cascaded circuits a and b, by definition the available noise at the output of b is

$$N_{ab} = F_{ab} G_{ab} kTB \tag{3-53}$$

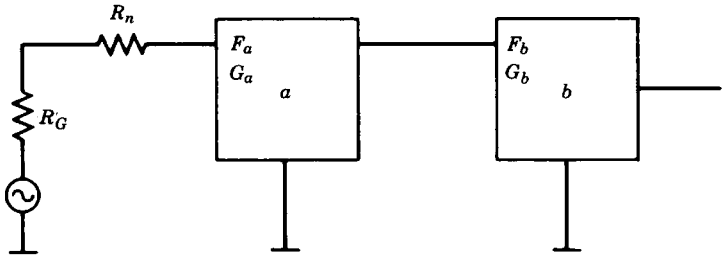


Figure 3-21 Cascaded noisy two-ports with the noise factors F_a and F_b and the gain figures G_a and G_b.

with B the equivalent noise bandwidth in which the noise is measured. The total available gain G is the product of the individual available gains, so

$$N_{ab} = F_{ab} G_a G_b kTB \tag{3-54}$$

The available noise from network a at the output of network b is

$$N_{a/b} = N_a G_b = F_a G_a G_b kTB \tag{3-55}$$

The available noise added by network b (its excess noise) is

$$N_{b/b} = (F_b - 1) G_b kTB \tag{3-56}$$

The total available noise N_{ab} is the sum of the available noise contributed by the two networks:

$$\begin{aligned} N_{ab} &= N_{a/b} + N_{b/b} = F_a G_a G_b kTB + (F_b - 1) G_b kTB \\ &= \left(F_a + \frac{F_b - 1}{G_a} \right) G_a G_b kTB \end{aligned} \tag{3-57}$$

$$F_{ab} = F_a + \frac{F_b - 1}{G_a} \tag{3-58}$$

For any number of circuits, this can be extended to

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \cdots \tag{3-59}$$

When considering a long chain of cascaded amplifiers, there will be a minimum noise figure achievable for this chain. This is a figure of merit and was proposed by Haus and Adler [6]. It is calculated by rearranging Eq. (3-59):

$$(F_{\text{tot}})_{\min} = (F_{\min} - 1) + \frac{F_{\min} - 1}{G_A} + \frac{F_{\min} - 1}{G_A^2} + \cdots + 1 \tag{3-60}$$

where $F_{\min}$ is the minimum noise figure for each stage and G_A is the available power gain of the identical stages. Using

$$\frac{1}{1 - X} = 1 + X + X^2 + \cdots \tag{3-61}$$

we find a quantity $(F_{\text{tot}} - 1)$, which is defined as the noise measure M. The minimum noise measure

$$(F_{\text{tot}})_{\min} - 1 = \frac{F_{\min} - 1}{1 - 1/G_A} = M_{\min} \tag{3-62}$$

refers to the noise of an infinite chain of optimally tuned, low-noise stages, so it represents a lower limit on the noise of an amplifier.

The minimum noise measure $M_{\min}$ is an invariant parameter and is not affected by feedback. It is somewhat similar to a gain–bandwidth product, in its use as a system invariant. The minimum noise measure is achieved when the amplifier is tuned for the available power gain and $\Gamma_G = \Gamma_{0n}$, given by Eq. (3-50).

Influence of External Parasitic Elements. Mounting an active two-port such as a transistor usually adds stray capacitance and lead inductance to the device, as shown in Figure 3-22. These external components consisting of transmission lines and parasitic reactances modify the noise parameters and the gain. Some researchers have published the results of these parasitic effects and have made manual computations or used some limited computer programs.

In a paper by Fukui [7], an attempt was made to determine the necessary equations, but the formulas are too involved even for pocket calculators. A more generic study by Iversen [8] is also very involved because of the various matrix manipulations and is more suitable for a computer. Besser's paper [9] in the *IEEE MTT-S* in 1975 and Vendelin's paper [10] in the same issue have shown for the first time some practical results using computers and even optimization methods, using an early version of COMPACT. The intention of these investigations was to find feedback that modifies the device noise and scattering parameters such that a noise match could also provide a low-input VSWR. It can be seen from these discussions that some feedback, besides resulting in some gain reduction, may improve the noise matching at the input for a limited frequency range.

A more recent paper by Suter [11] based on a report by Hartmann and Strutt [12] has given a simple transformation starting from the *S*-parameters and the noise parameters from common-source (or common-emitter) measurements. The noise parameters for the "packaged" device are calculated. This means that the parameters for the "new" device, including

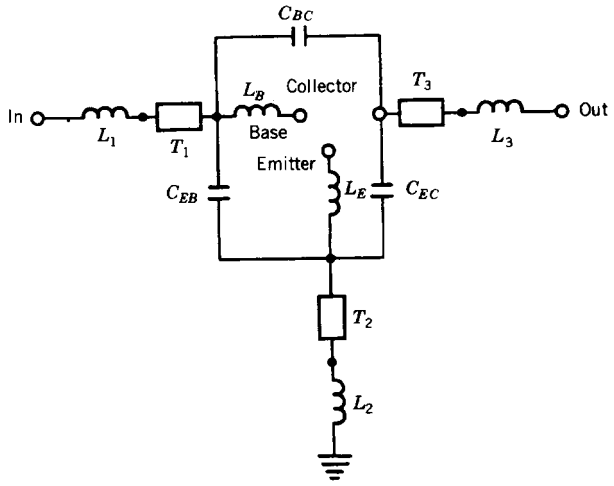
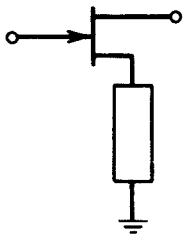


Figure 3-22 Equivalent circuit of the transistor package.

Table 3-2 Transformation matrix parameters

Series Feedback



$$Z_S = R_S + jX_S = Z_S$$

$$[n] = \begin{bmatrix} n_{11} = 1 & n_{12} = Z_0 \dfrac{S_{21}M - S'_{21}N}{S_{21}C'_1 + S'_{21}C_1}\ \Omega \\ n_{21} = 0 & n_{22} = \dfrac{S_{21}C'_1}{S_{21}C'_1 + S'_{21}C_1} \end{bmatrix}$$

where

$$S'_{11} = S'_{22} = \frac{-1}{1 + 2Z_S}$$

$$S'_{12} = S'_{21} = \frac{2Z_S}{1 + 2Z_S}$$

$$M = (1 + S'_{11})(1 - S'_{22}) + S'_{12}S'_{21}$$

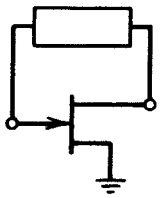
$$N = (1 + S_{11})(1 - S_{22}) + S_{12}S_{21}$$

$$C_1 = (1 - S_{11})(1 - S_{22}) - S_{12}S_{21}$$

$$C'_1 = (1 + S'_{11})(1 - S'_{22}) - S'_{12}S'_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

Shunt Feedback



$$Z_P = R_P + jX_P = Z_P$$

$$[n] = \begin{bmatrix} n_{11} = \dfrac{S_{21}C'_2}{S_{21}C'_2 + S'_{21}C_2} & n_{12} = 0\ \Omega \\ n_{21} = \dfrac{1}{Z_0}\dfrac{S_{21}P - S'_{21}Q}{S_{21}C'_2 + S'_{21}C_2}\,S & n_{22} = 1 \end{bmatrix}$$

where

$$S'_{11} = S'_{22} = \frac{Z_P}{2 + Z_P}$$

(*continued*)

Table 3-2 (*Continued*)

$$S'_{12} = S'_{21} = \frac{2}{2 + Z_P}$$

$$P = (1 - S'_{11})(1 + S'_{22}) + S'_{12}S'_{21}$$

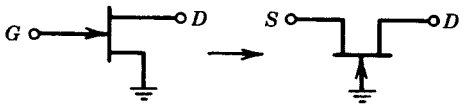
$$Q = (1 - S_{11})(1 + S_{22}) + S_{12}S_{21}$$

$$C_2 = (1 + S_{11})(1 + S_{22}) - S_{12}S_{21}$$

$$C'_2 = (1 + S'_{11})(1 + S'_{22}) - S'_{12}S'_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

Common Gate



$$[n] = \begin{bmatrix} n_{11} = \dfrac{2S_{21}}{-2S_{21} + C_4} & n_{12} = 0\ \Omega \\ n_{21} = \dfrac{1}{Z_0}\dfrac{C_3C_4 - 4S_{12}S_{21}}{V(-2S_{21} + C_4)}S & n_{22} = -1 \end{bmatrix}$$

where

$$V = (1 + S_{11})(1 + S_{22}) - S_{12}S_{21}$$

$$C_3 = (1 - S_{11})(1 + S_{22}) + S_{12}S_{21}$$

$$C_4 = (1 + S_{11})(1 - S_{22}) + S_{12}S_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \text{common-source } S \text{ parameters}$$

the common-gate (or common-base) case, are calculated. The equations are device dependent. They are valid for any active two-port.

A transformation matrix, n, may be used to combine the noise sources of the various circuit configurations. The transformation matrix parameters are given in Table 3-2 for series feedback, shunt feedback, and the common-gate (base) case, which will be important for oscillator analysis. The transformation matrix gives the new noise parameters as follows:

$$R'_n = R_n|n_{11} + n_{12}Y_{\text{cor}}|^2 + G_n|n_{12}|^2 \tag{3-63}$$

$$G'_n = \frac{G_nR_n}{R'_n}|n_{11}n_{22} - n_{12}n_{21}|^2 \tag{3-64}$$

$$Y'_{\text{cor}} = \frac{R_n}{R'_n}(n_{21} + n_{22}Y_{\text{cor}})(n^*_{11} + n^*_{12}Y^*_{\text{cor}}) + \frac{G_n}{R'_n}n_{22}n^*_{12} \tag{3-65}$$

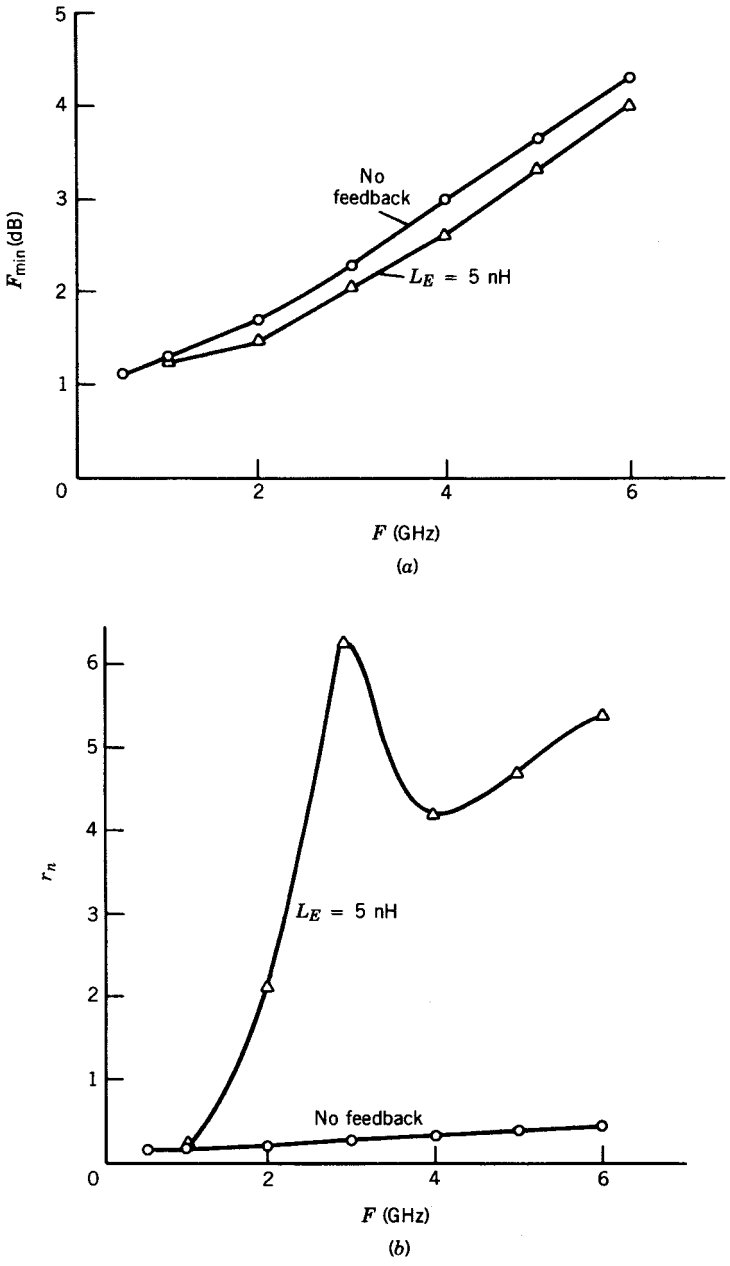


Figure 3-23 Noise parameters versus feedback for the AT-41435 silicon bipolar transistor: (a) F_{min} versus frequency and feedback; (b) r_n versus frequency and feedback; (c) Γ_{0n} versus frequency and feedback.

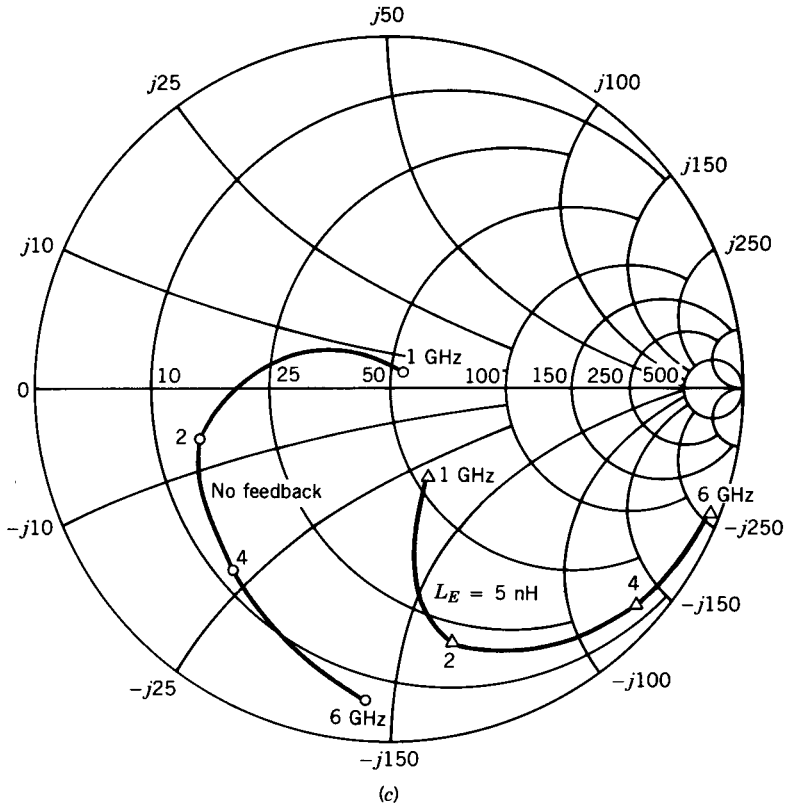


Figure 3-23 (*Continued*)

A final transformation to the more common noise-parameter format given by Eq. (3-44) is still needed [10]:

$$F_{\min} = 1 + 2R_n'(G_{\text{cor}}' + G_{0n}') \tag{3-66}$$

$$R_n = R_n' \tag{3-67}$$

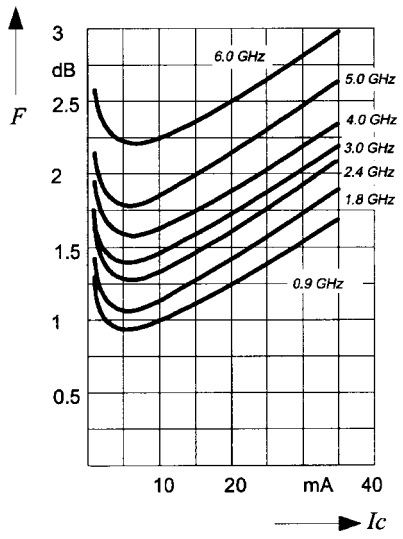
$$G_{0n} = \sqrt{\frac{G_n'}{R_n'} + G_{\text{cor}}'^2} \tag{3-68}$$

$$B_{0n} = -B_{\text{cor}}' \tag{3-69}$$

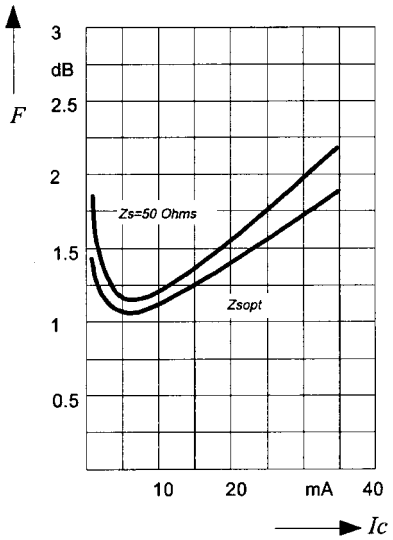
Figure 3-23 shows the noise figure as a function of external feedback for a low-noise microwave bipolar transistor, the AT-41435.

Bias-Dependent Noise Parameters. Since we have elaborated on the effect of parasitics, it is now useful to point out that noise is also a strong function of bias point and frequency. Figure 3-24 shows the noise figure expressed in dB as a function of different dc currents and frequencies for a BFP420 BJT by Siemens. A change in the collector voltage will also affect the noise parameters, not quite as strongly as shown here, but still significantly. The reason

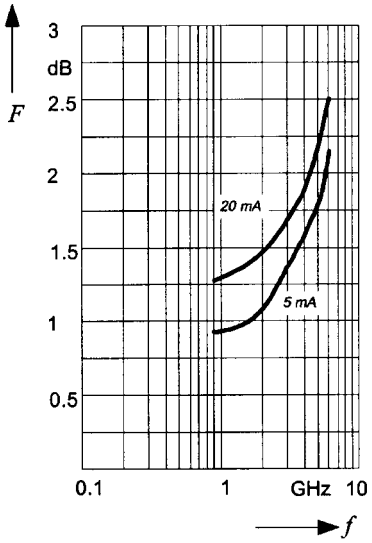
Noise Figure
versus Collector Current
V_{CE} = 2 V, $Z_S = Z_{Sopt}$



Noise Figure
versus Collector Current
VCE = 2 V, f = 1.8 GHz



Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA,
$Z_S = Z_{Sopt}$



Source Impedance for min.
Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA

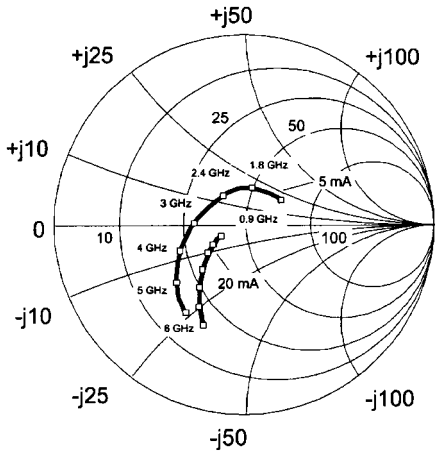


Figure 3-24 Noise figure versus collector current, noise figure versus frequency, and source impedance for minimum noise figure versus frequency for the Siemens BFP420 transistor.

Table 3-3 BFP420 common-emitter noise parameters

f	$F_{\min}{}^a$	$G_a{}^a$	Γ_{opt}		R_N	r_n	$F_{50\Omega}{}^b$	$\lvert S_{21} \rvert^{2b}$
(GHz)	dB	(dB)	(Magnitude)	(Angle)	(Ω)		(dB)	(dB)
V_{CE} = 2 V, I_C = 5 mA								
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	–175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	–157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	–142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	–123.0	15.0	0.30	3.30	5.3

[a]Input matched for minimum noise figure, output for maximum gain.
[b]$Z_S = Z_L = 50\ \Omega$.

for the turning point of the noise figure has to do with the fact that for lower currents the f_T cutoff frequency has not reached its peak yet, and therefore the noise increases. The increase of the noise at higher currents is due to the Schottky noise calculated from the equation

$$P_n = 2Iq \tag{3-70}$$

where P_n is the noise power, I is the saturation, collector, or drain current, and q is the charge of an electron (1.60×10^{-19} coulomb).

Table 3-3 shows the common-emitter noise parameters for the BFP420.

Noise Circles. From the section entitled "Influence of External Parasitic Elements," we see that the noise factor is a strong function of the generator admittance (or impedance) presented to the input terminals of the noisy two-port. Noise tuning is used to change the values of the input admittance to obtain the best noise performance. There is a range of values of input reflection coefficient over which the noise figure is constant. In plotting these points of constant noise figure, we obtain the so-called noise circles, which can be drawn on the Smith chart Γ_G plane [13]. Starting with the noise equation [see Eq. (3-44)] for a 50-Ω generator impedance, we find that

$$F_{50} = F_{\min} + 4r_n \frac{|\Gamma_{0n}|^2}{|1 + \Gamma_{0n}|^2} \tag{3-71}$$

We want to find the position of the reflection coefficient on the Smith chart, as in the case of the gain circles, for which F is constant. First we rearrange Eq. (3-67) to read

$$r_n = (F_{50} - F_{\min}) \frac{|1 + \Gamma_{0n}|^2}{4|\Gamma_{0n}|^2} \tag{3-72}$$

By introducing

$$N_i = \frac{F_i - F_{\min}}{4r_n} |1 + \Gamma_{0n}|^2 \tag{3-73}$$

we can find an expression for a circle of constant noise figure as introduced by Rothe and Dahlke [4] (see also Fukui [13]). The center for the noise circle is

$$C_i = \frac{|\Gamma_{0n}|}{1 + N_i} \tag{3-74}$$

and the radius is

$$r_i = \frac{\sqrt{N_i^2 + N_i(1 - |\Gamma_{0n}|^2)}}{1 + N_i} \tag{3-75}$$

with the definition of N used previously. However, if we consider only the minimum noise figure for a given device, we will not obtain the minimum noise figure for the multistage amplifier system. This was explained when the noise measure was introduced. (See Figure 3-25.) Therefore, a better way to design the amplifier would be to use circles at constant noise measure instead of circles of constant noise figure. This was recently done by Poole

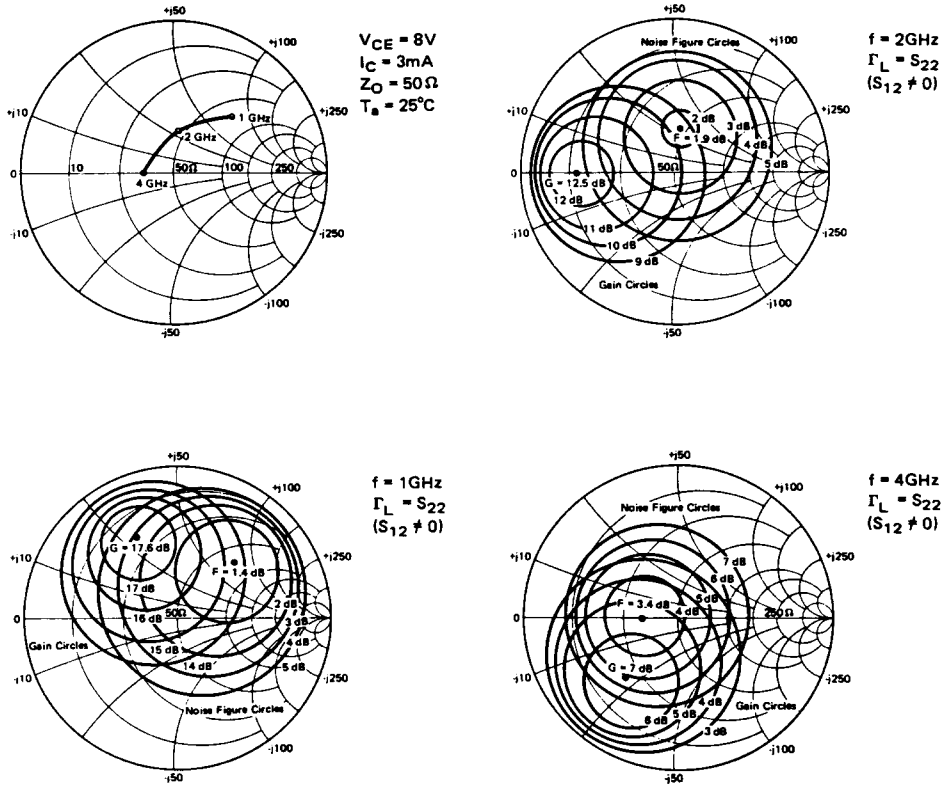


Figure 3-25 Typical noise figure circles and gain circles.

and Paul [14]. They derived the expressions for the noise measure circles as a function of S parameters, noise parameters, and Γ_G, using

$$G_A = \frac{|S_{21}|^2(1 - |\Gamma_G|^2)}{(1 - |S_{22}|^2) + |\Gamma_G|^2 (|S_{11}|^2 - |\Delta|^2) - 2\,\mathrm{Re}(\Gamma_G C_1)} \tag{3-76}$$

where

$$C_1 = S_{11} - S_{22}^* \Delta$$

$$\Delta = S_{11}S_{22} - S_{12}S_{11}$$

In terms of the generator reflection coefficient, the noise measure can be expressed as

$$M = [(F_{\min} - 1)(1 - |\Gamma_G|^2)|S_{21}|^2|1 + \Gamma_{0n}|^2 + 4r_n|S_{21}|^2|\Gamma_G - \Gamma_{0n}|^2]$$

$$\times \left\{|1 + \Gamma_{0n}|^2[|S_{21}|^2(1 - |\Gamma_G|^2) - (1 - |S_{22}|^2) - |\Gamma_G|^2(|S_{11}|^2 - |\Delta|^2) + 2\,\mathrm{Re}(\Gamma_G C_1)]\right\}^{-1} \tag{3-77}$$

Equation (3-77) can be shown to represent circles in the source reflection coefficient plane described by the following:

$$|\Gamma_G|^2 + |\Gamma_m|^2 - \Gamma_G\Gamma_m^* - \Gamma_G^*\Gamma_m = r_m^2 \tag{3-78}$$

The centers and radii of the constant-noise-measure circles are given by

$$C_m = \frac{M|1 + \Gamma_{0n}|^2 C_1^* + 4r_n|S_{21}|^2\Gamma_{0n}}{M|1 + \Gamma_{0n}|^2 P + |S_{21}|^2(4r_n - W)} \tag{3-79}$$

$$r_m = \frac{\sqrt{M^2M_a + MM_b + M_c}}{M|1 + \Gamma_{0n}|^2 P + |S_{21}|^2(4r_n - W)} \tag{3-80}$$

where

$$P = |S_{21}|^2 + |S_{11}|^2 - |\Delta|^2$$

$$Q = |S_{21}|^2 + |S_{22}|^2 - 1$$

$$W = |1 + \Gamma_{0n}|^2(F_{\min} - 1)$$

$$M_a = |1 + \Gamma_{0n}|^4(PQ + |C_1|^2)$$

$$M_b = |1 + \Gamma_{0n}|^2|S_{21}|^2[8r_n\,\mathrm{Re}(\Gamma_{0n}C_1)$$

$$- (4r_n|\Gamma_{0n}|^2 + W)P - (W - 4r_n)Q]$$

$$M_c = |S_{21}|^4[W - 4r_n(1 - |\Gamma_{0n}|^2)]$$

The value of the minimum noise measure can be found by considering the noise measure circle of zero radius; that is, set r_m equal to zero in Eq. (3-80). This results in

$$M_{\min} = \frac{-M_b \pm \sqrt{M_b^2 - 4M_a M_c}}{2M_a} \tag{3-81}$$

Equation (3-81) yields the same value of $M_{\min}$ as would have been obtained by using the immittance parameter equation given by Fukui [13] and can, therefore, be considered as the reflection coefficient plane analog of Fukui's expression. The elimination of the need for the parameter R_{eg}, however, results in a considerable simplification as compared with the earlier approach [13].

The value of the minimum noise measure is taken as the smallest nonnegative value of $M_{\min}$ given by Eq. (3-81). The source reflection coefficient that results in the minimum noise measure can now be obtained by employing Eq. (3-79):

$$\Gamma_{0m} = \frac{M_{\min}|1 + \Gamma_{0n}|^2 C_1^* + 4r_n|S_{21}|^2\Gamma_{0n}}{M_{\min}|1 + \Gamma_{0n}|^2 P + |S_{21}|^2\,(4r_n - W)} \tag{3-82}$$

The output reflection coefficient of the device, when Γ_{0m} is presented to the input port, is given by

$$S'_{22} = \frac{S_{22} - \Delta\Gamma_{0m}}{1 - S_{11}\Gamma_{0m}} \tag{3-83}$$

Noise Correlation in Linear Two-Ports Using Correlation Matrices. In the introduction to two-port noise theory, it was indicated that noise correlation matrices form a general technique for calculating noise in n-port networks. Haus and Adler [6] have described the theory behind this technique. In 1976, Hillbrand and Russer [15] published equations and transformations that aid in supplying this method to two-port CAD.

This method is useful because it forms a base from which we can rigorously calculate the noise of linear two-ports combined in arbitrary ways. For many representations, the method of combining the noise parameters is as simple as that for combining the circuit element matrices. In addition, noise correlation matrices can be used to calculate the noise in linear frequency conversion circuits. The following is an introduction to this subject.

A linear, noisy two-port can be modeled as a noise-free two-port with two additional noise sources. These noise sources must be chosen so that they add directly to the resulting vector of the representation, as shown in Eqs. (3-84) and (3-85) and Figure 3-17:

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \tag{3-84}$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{3-85}$$

where the i and v vectors indicate noise sources for the y and z representations, respectively. This two-port example can be extended to n-ports in a straightforward, obvious way.

Since the noise vector for any representation is a random variable, it is much more convenient to work with the noise correlation matrix. The correlation matrix gives us deterministic numbers with which to calculate. The correlation matrix is formed by taking the mean value of the outer product of the noise vector. This is equivalent to multiplying the noise vector by its adjoint (complex conjugate transpose) and averaging the result:

$$\langle \bar{i}\ \bar{i}^{+} \rangle = \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} [i_1^* \quad i_2^*] = \begin{bmatrix} \langle i_1 i_1^* \rangle & \langle i_1 i_2^* \rangle \\ \langle i_1^* i_2 \rangle & \langle i_2 i_2^* \rangle \end{bmatrix} + [C_y] \tag{3-86}$$

where the angular brackets denote the average value.

Note that the diagonal terms are the "power" spectrum of each noise source and the off-diagonal terms are complex conjugates of each other and represent the "cross-power" spectra of the noise sources. "Power" is used because these magnitude-squared quantities are proportional to power.

To use these correlation matrices in circuit analysis, we must know how to combine them and how to convert them between various representations. An example using y matrices will illustrate the method for combining two-ports and their correlation matrices. Given two matrices y and y', when we parallel them we have the same port voltages, and the terminal currents add (Figure 3-26):

$$\begin{aligned} I_1 &= y_{11}V_1 + y_{12}V_2 + y'_{11}V_1 + y'_{12}V_2 + i_1 + i'_1 \\ I_2 &= y_{21}V_1 + y_{22}V_2 + y'_{21}V_1 + y'_{22}V_2 + i_2 + i'_2 \end{aligned} \tag{3-87}$$

or

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} + y'_{11} & y_{12} + y'_{12} \\ y_{21} + y'_{21} & y_{22} + y'_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 + i'_1 \\ i_2 + i'_2 \end{bmatrix} \tag{3-88}$$

Here we can see that the noise current vectors add just as the y parameters add. Converting the new noise vector to a correlation matrix yields

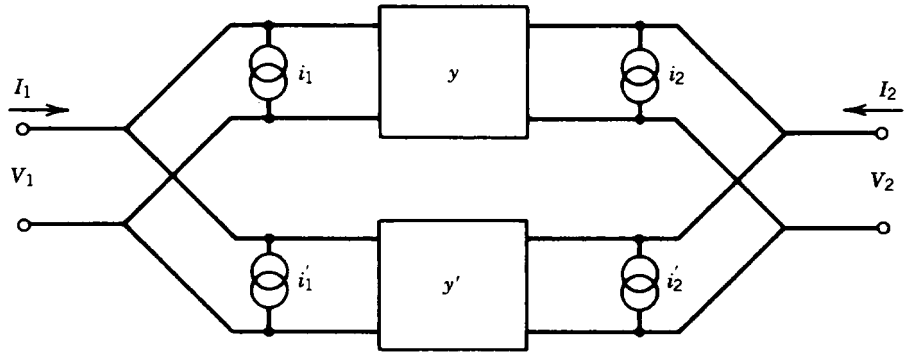


Figure 3-26 Parallel combination of two-ports using y parameters.

$$\langle \bar{i}_{\text{new}} \bar{i}^{+}_{\text{new}} \rangle = \left\langle \begin{bmatrix} i_1 + i_1' \\ i_2 + i_2' \end{bmatrix} [i_1^* + i_1'^* \quad i_2 i_2'^*] \right\rangle \tag{3-89}$$

$$= \begin{bmatrix} \langle i_1 i_1^* \rangle + \langle i_1' i_1'^* \rangle & \langle i_1 i_2^* \rangle + \langle i_1' i_2'^* \rangle \\ \langle i_2 i_1^* \rangle + \langle i_2' i_1'^* \rangle & \langle i_2 i_2^* \rangle + \langle i_2' i_2'^* \rangle \end{bmatrix} \tag{3-90}$$

The noise sources from different two-ports must be uncorrelated, so there are no cross products of different two-ports. By inspection, Eq. (3-90) is just the addition of the correlation matrices for the individual two-ports, so

$$[C_{y\text{new}}] = [C_y] + [C_y'] \tag{3-91}$$

The same holds true for g, h, and z parameters, but $ABCD$ parameters have the more complicated form shown below. If

$$[A_{\text{new}}] = [A][A'] \tag{3-92}$$

then

$$[C_{A\text{new}}] = \{C_A] + [A]\,[C_{A'}]\,[A]^{+} \tag{3-93}$$

The transformation of one representation to another is best illustrated by an example. Let us transform the correlation matrix for a Y representation to a Z representation. Starting with

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = [Y] \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \tag{3-94}$$

we can move the noise vector to the left side and invert y:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = [Y^{-1}] \begin{bmatrix} I_1 - i_1 \\ I_2 - i_2 \end{bmatrix} = [Y^{-1}] \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + [Y^{-1}] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-95}$$

Table 3-4 π Noise matrix $T_{\alpha\beta}$ transformation

		Original Form (α Form)					
		Y		Z		A	
Resulting Form (β Form)	Y	1	0	y_{11}	y_{12}	$-y_{11}$	1
		0	1	y_{21}	y_{22}	$-y_{21}$	0
	Z	z_{11}	z_{12}	1	0	1	$-z_{11}$
		z_{21}	z_{22}	0	1	0	$-z_{21}$
	A	0	A_{12}	1	$-A_{11}$	1	0
		1	A_{22}	0	$-A_{21}$	0	1

Since $(Y)^{-1} = (Z)$, we have

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = [Z] \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + [Z] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-96}$$

so

$$= [Z] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} = [T_{yz}] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-97}$$

where the signs of i_1 and i_2 are superfluous since they will cancel when the correlation matrix is formed. Here the transformation of the Y noise current vector to the Z noise voltage vector is done simply by multiplying by (Z). Other transformations are shown in Table 3-4.

To obtain the noise correlation matrix, we again form the mean of the outer product:

$$\langle vv^+ \rangle = \begin{bmatrix} \langle v_1 v_1^* \rangle & \langle v_1 v_2^* \rangle \\ \rangle v_1^* v_2 \rangle & \langle v_2 v_2^* \rangle \end{bmatrix} = [Z] \left\langle \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} [i_1^* \quad i_2^*] \right\rangle [Z]^+ \tag{3-98}$$

or

$$[C_z] = [Z]\,[C_y]\,[Z]^+ \tag{3-99}$$

where

$$v^+ = [i_1^* \quad i_2^*]\,[Z]^+ \tag{3-100}$$

This is called a congruence transformation. The key to all of these derivations is the construction of a correlation matrix from the noise vector, as shown in Eq. (3-90).

These correlation matrices may easily be derived from the circuit matrices of passive circuits with only thermal noise sources. For example,

$$[C_z] = 2kT\Delta f\,\mathrm{Re}([Z]) \tag{3-101}$$

and

$$[C_y] = 2kT\Delta f\,\mathrm{Re}([Y]) \tag{3-102}$$

The $2kT$ factor comes from the double-sided spectrum of thermal noise. The correlation matrix from the $ABCD$ matrix may be related to the noise figure, as shown by Hillbrand and Russer [15]. We have

$$F = 1 + \frac{\overline{Y}[C_a]\overline{Y}^+}{2kT\,\mathrm{Re}(Y_G)} \tag{3-103}$$

where

$$\overline{Y} = \begin{bmatrix} Y_G \\ 1 \end{bmatrix}$$

Expressing the noise factor in terms of the correlation matrix, we obtain a complete formula:

$$F = 1 + \frac{C_{22}^{A}(f) + 2\,\mathrm{Re}\{Y_g(f)\,C_{12}^{A}(f)\} + |Y_g(f)|^2\,C_{11}^{A}(f)}{2kT_0\,\mathrm{Re}\{Y_g(f)\}} \tag{3-104}$$

Once we transform this in the Y parameter form, we obtain the following equation:

$$F(f) = F_{\min}(f) + \frac{R_n(f)|Y_{\mathrm{opt}}(f) - Y_g(f)|^2}{\mathrm{Re}\{Y_g(f)\}} \tag{3-105}$$

It should be noted that all these values are frequency dependent, as expressed in this equation.

The *ABCD* correlation matrix can be written in terms of the noise figure parameters as (double-sided spectrum)

$$[C_a] = 2kT\begin{bmatrix} R_n & \dfrac{F_0 - 1}{2} - R_n Y_{0n}^{*} \\ \dfrac{F_0 - 1}{2} - R_n Y_{0n} & R_n|Y_{0n}|^2 \end{bmatrix} \tag{3-106}$$

The noise correlation matrix method forms an easy and rigorous technique for handling noise in networks. This technique allows us to calculate the total noise for complicated networks by combining the noise matrices of subcircuits. It should be remembered that although noise correlation matrices apply to n-port networks, noise figure calculations apply only to pairs of ports. The parameters of the C_a matrix can be used to give the noise parameters:

$$Y_{0n} = \sqrt{\frac{C_{ii*}}{C_{uu*}} - \left[\mathrm{Im}\left(\frac{C_{ui*}}{C_{uu*}}\right)\right]^2} + j\,\mathrm{Im}\left(\frac{C_{ui*}}{C_{uu*}}\right) \tag{3-107}$$

$$F_0 = 1 + \frac{C_{ui*} + C_{uu*}Y_{0n}^{*}}{kT} \tag{3-108}$$

$$R_u = C_{uu*} \tag{3-109}$$

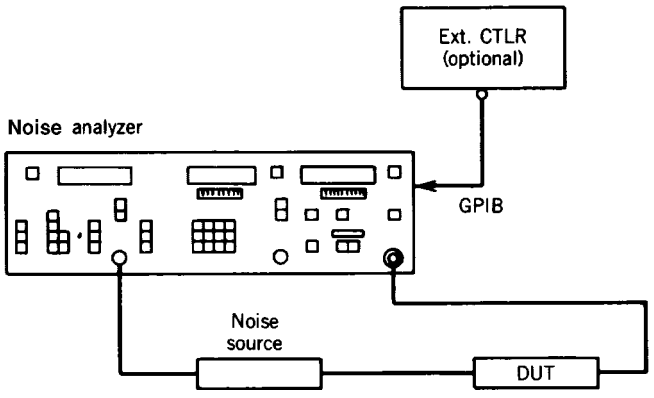


Figure 3-27 Noise figure measurement.

Noise Figure Test Equipment. Figure 3-27 shows the block diagram of a noise test setup. It includes the noise source and the other components. The metering unit has a special detector that is linear and over a certain dynamic range measures linear power. The tunable receiver covers a wide frequency range (e.g., 10–1800 MHz) and controls the noise source. The receiver is a double-conversion superheterodyne configuration with sufficient image rejection to avoid double-sideband noise measurements that would give the wrong results.

These receivers are microprocessor controlled and the measurement is a two-step procedure. The first is a calibration step that measures the noise figure of the receiver system and a reference power level. Then the device under test (DUT) is inserted and the system noise figure and total output power are measured. The noise factor is calculated by

$$F_1 = F_{\text{system}} - \frac{F_2 - 1}{G_1} \tag{3-110}$$

and the gain is given by the change in output power from the reference level [16]. The noise of the system is calculated by measuring the total noise power with the noise source on and off. With the excess noise ratio (ENR) known [16],

$$F_{\text{system}} = \frac{\text{ENR}}{Y - 1} \tag{3-111}$$

The noise bandwidth is usually set by the bandwidth of the receiver, which is assumed to be constant over the linear range. The ENR of the noise source is given by

$$\text{ENR} = \frac{T_{\text{hot}}}{T_{\text{cold}}} - 1 \tag{3-112}$$

where T_{cold} is usually room temperature (290 K). This ENR number is about 15 dB for noise sources with a 15-dB pad and 5 dB for noise sources with a 25-dB pad. Since both gain and noise were stored in the initial calibration, a noise/gain sweep can be performed.

For frequencies above 1800 MHz, we can extend the range with the help of the external signal generators, as shown in Figure 3-28. As shown, a filter ahead of the external mixer reduces the noise energy in the image band. If the DUT has a very broad frequency range

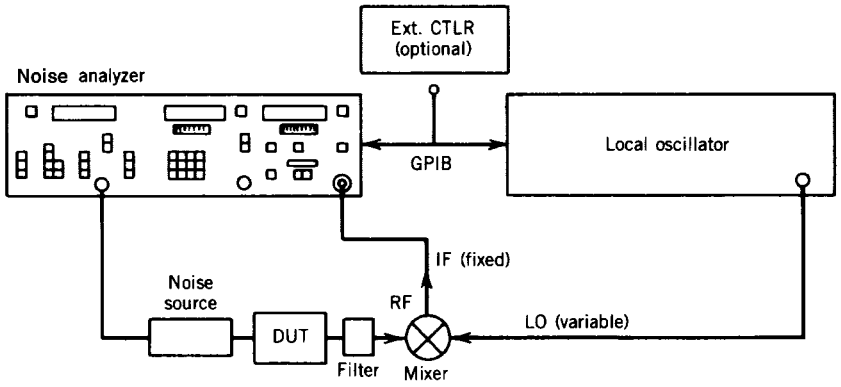


Figure 3-28 Single-sideband (SSB) noise figure measurements using an external mixer.

and has flat gain and noise over that range, a double-sideband (DSB) measurement is possible, with the image-rejection filter removed. However, a single-sideband (SSB) measurement is always more accurate [17].

How To Determine the Noise Parameters. The noise figure of a linear two-port network as a function of source admittance may be represented by

$$F = F_{\min} + \frac{R_n}{G_G}\left[(G_{0n} - G_G)^2 + (B_{0n} - B_G)^2\right] \tag{3-113}$$

where $G_G + jB_G$ = generator admittance presented to the input of the two-port
$G_{0n} + jB_{0n}$ = generator admittance at which optimum noise figure occurs
R_n = empirical constant relating sensitivity of the noise figure to generator admittance, with dimensions of resistance

It may be noted that for an arbitrary noise figure measurement with a known generator admittance, Eq. (3-113) has four unknowns—$F_{\min}$, R_n, G_{0n}, and B_{0n}. By choosing four known values of generator admittance, a set of four linear equations is formed and solutions of the four unknowns can be found [18, 19]. Eq. (3-113) may be transformed to

$$F = F_{\min} + \frac{R_n|Y_{0n}|^2}{G_G} - 2R_nG_{0n} + \frac{R_n|Y_G|^2}{G_G} - 2R_nB_{0n}\frac{B_G}{G_G} \tag{3-114}$$

or

$$F = F_{\min} + \frac{R_n}{G_G}|Y_G - Y_{0n}|^2 \tag{3-115}$$

Let

$$X_1 = F_{\min} - 2R_nG_{0n}$$

$$X_2 = R_n|Y_{0n}|^2$$

$$X_3 = R_n$$

$$X_4 = R_nB_{0n}$$

Then the generalized equation may be written

$$F_i = X_1 + \frac{1}{G_{si}}X_2 + \frac{|Y_{si}|^2}{G_{si}}X_3 - 2\frac{G_{si}}{B_{si}}X_4 \tag{3-116}$$

or, in matrix form,

$$[F] = [A]\,[X] \tag{3-117}$$

and the solution becomes

$$[X] = [A]^{-1}\,[F] \tag{3-118}$$

These parameters completely characterize the noise behavior of the linear two-port network. Direct measurement of these noise parameters by this method would be possible only if the receiver on the output of the two-port were noiseless and insensitive to its input admittance. In actual practice, the receiver itself behaves as a noisy two-port network and can be characterized in the same manner. What is actually being measured is the system noise figure of the two-port and the receiver. Thus, it becomes apparent that to do a complete two-port noise characterization, the gain of the two-port must be measured [20]. In addition, any losses in the input-matching networks must be carefully accounted for, because they add directly to the measured noise figure reading [21].

3-1-4 Linearity

Consistent with the elaborate discussion of linearity questions above, here is a quick refresher on what must be considered in amplifiers.

Dynamic Range, Compression, and IMD. An amplifier's dynamic range is related to the minimum discernible signal and the 1-dB compression point according to the equation

$$DR_n = \frac{(n-1)(IP_{n(\text{in})} - MDS)}{n} \tag{3-119}$$

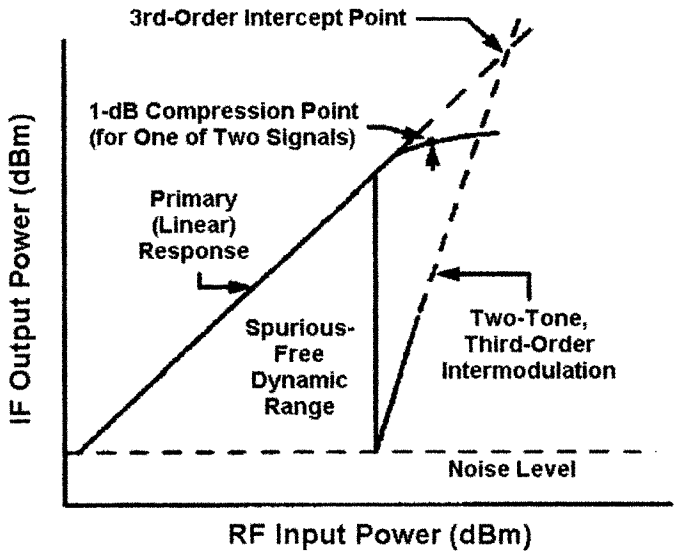


Figure 3-29 Amplifier linearity evaluation, including compression and two-tone IMD dynamic range. $P_{-1\text{dB}}$ for a single-tone cannot be read directly from this graph because the values shown are the result of two equal-power tones.

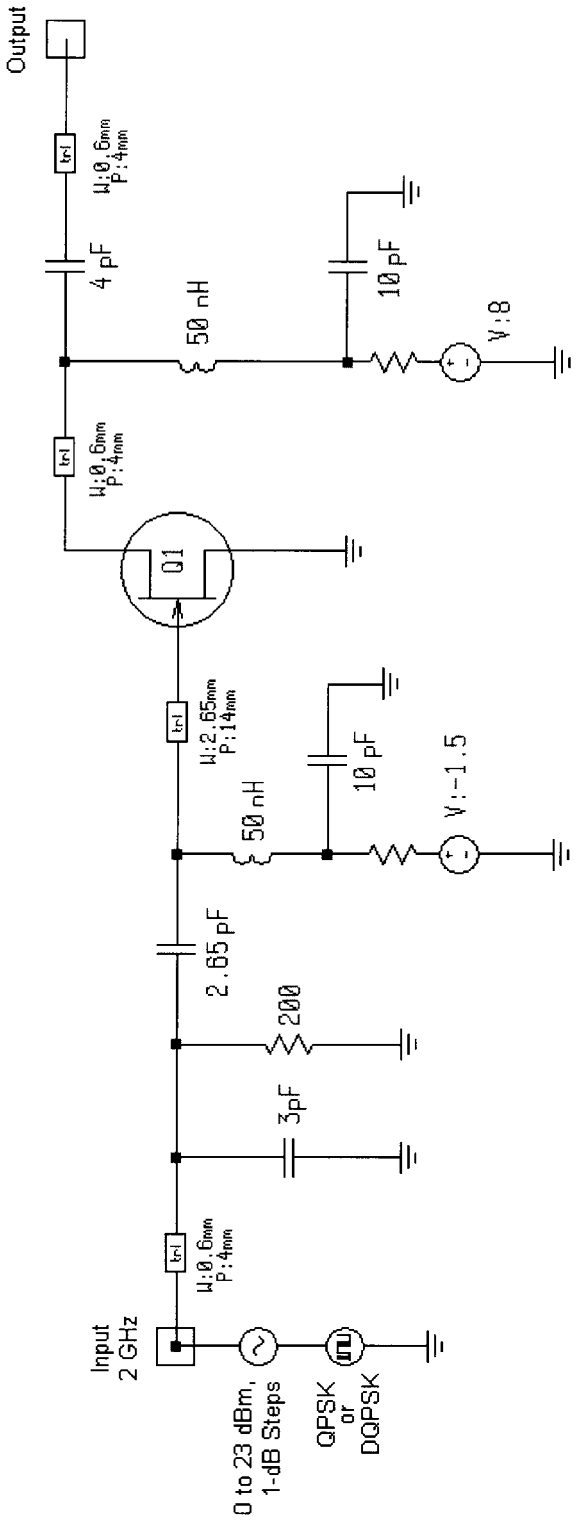


Figure 3-30 The power amplifier schematic.

where DR is the dynamic range in decibels, n is the order, $IP_{(in)}$ is the input intercept power in dBm, and MDS is the minimum detectable signal power in dBm. Once the 1-dB compression point is known, it is a fair assumption that the third-order intercept point, expressed in dBm, is 10 dB above this.

The intermodulation distortion gives the quality of an amplifier to withstand multiple signals without generating large intermodulation products or when using amplitude-modulated signals (cross-modulation). Figure 3-29 provides a quick reminder of IMD issues.

In some amplifiers, we would like to have control over the gain. While the better way to control gain while obtaining high linearity is a PIN-diode attenuator, special AGC circuits have become part of the amplifiers. Figure 3-48 shows a circuit of the SL610, made by Plessey, which has an AGC input provision. The noise figure of the PIN-diode attenuator is essentially equal to its attenuation. The AGC in multistage amplifiers is distributed over several stages and there is a lesser correlation between noise and AGC than in purely passive attenuators.

Special Case of Linearity Requirements for Digital Modulation. For cellular telephone systems we use digital modulation formats, such as QPSK and π/4-DQPSK, that combine phase and amplitude modulation. Amplifiers handling such signals must be carefully characterized and designed if adequate amplitude and phase linearity are to be maintained.

Examples of Power Amplifiers: Looking into the Effects of Distortion. This example presents modulation analysis using QPSK in a 2-GHz power amplifier (Figure 3-30). The simple amplifier was designed using the electronic Smith Tool of Serenade to determine the matching network for a narrowband design. The example demonstrates setup of modulation analysis and available results.

Siemens among others provides integrated amplifiers based on the concept depicted in Figure 3-30, the most popular ones being the CGY96 and CGY94. For the purpose of measurements, these amplifiers were actually put on a breadboard. Figure 3-31 shows the big brother, the CGY96. It is a GaAs MMIC intended as a power amplifier for GSM Class 4 phones with 3.2 W (35 dBm) at 3.5 V at an overall power added efficiency of 50%. Table 3-5 shows its electrical characteristics.

For the purpose of our experiments, we considered the CGY94 because it requires less external peripheral circuits. Table 3-6 shows its electrical characteristics. For those of us interested in seeing part of the internal circuit, Figure 3-32 shows the chip. This two-stage amplifier uses an interesting internal feedback scheme. Its schematic is shown in Figure 3-33.

For the purpose of looking at the waveforms, we will continue with the single-stage amplifier (Figure 3-30) and then compare the measured versus predicted performance of this CGY94-based amplifier.

A modulation source is connected to the RF source at the input port of the amplifier. The RF source specifies the carrier power (or voltage or current) and the modulation source specifies the modulation format and properties of the modulated signal.

A brief review of the modulation source used in this example indicates the following properties:

Property	Value	Description
NB	128	Number of bits
BR	1.2288E6	Bit rate
N	8	Number of samples per bit
BW1	590.E3	One-sided bandwidth of the main channel in the PIB or P1IB calculation
FS1	740.E3	Start baseband frequency of the fist adjacent channel
BW2	590.E3	One-sided bandwidth of the main channel in the P2IB calculation
FS2	1990.E3	Start baseband frequency of the second adjacent channel
BW3	590.E3	One-sided bandwidth of the main channel in the P3IB calculation
FS3	3240.E3	Start baseband frequency of the third adjacent channel
M	4	Order of the signal space; M = 4 is QPSK
DLY	0.0	Fractional bit delay
IASC	1.0	I channel amplitude scale; multiplier for the I waveform to model amplitude imbalance
QASC	1.0	Q channel amplitude scale; multiplier for the Q waveform to model amplitude imbalance
LPF	3	Baseband low-pass filter; 3 = root-raised cosine filter
LPFC	665.E3	–3 dB cutoff frequency of LPF
LPFN	3	Number of resonators if LPF = 1
LPFR	0.35	Roll-off factor if LPF = 3 or LPF = 4
FILE	N/A	File name of user-defined modulated signal file

The modulation source is a QPSK stream with a bit rate of 1.2288 Mbps. A total of 128 bits will be analyzed, and each bit is sampled eight times to construct an accurate analog waveform. The BW and FS parameters are defined here to specify the main (in-band) channel and adjacent channel for ACPR calculations. Imbalance of the I and Q amplitudes and phases can be described by the IASC, QASC, and DLY parameters, if desired. The modulation signal

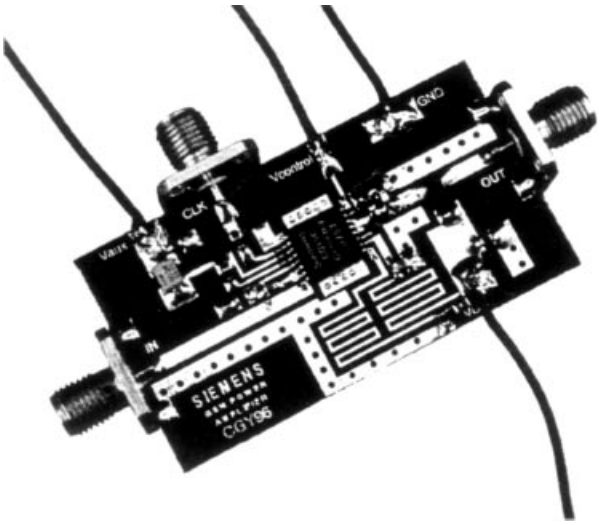


Figure 3-31 CGY96 GaAs MMIC power amplifier.

Table 3-5 CGY96 electrical characteristics

(T_A = 25 °C, V_{neg} = 5 V, $V_{control}$ = 2.2 V; duty cycle 12.5%, t_{on} = 577 µs)

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
Frequency range	f	880	—	915	MHz
Supply current P_{in} = 0 dBm	I_D	—	1.8	—	A
Supply current negative voltage generator V_{aux} = 3.5 V	I_{AUX}	—	10	—	mA
Gain (small signal)	G	—	40	—	dB
Power gain P_{in} = 0 dBm	G_P	—	35	—	dB
Output power P_{in} = 0 dBm, $V_{control}$ = 2.0–2.5 V	P_{OUT}	—	35	—	dBm
Overall power added efficiency P_{in} = 0 dBm	η	—	50	—	%
Dynamic range output power $V_{control}$ = 0.2–2.2 V		—	80	—	dB
Harmonics P_{in} = 0 dBm	$H(2f_0)$	—	–40	—	dBc
	$H(3f_0)$	—	–43	—	dBc
	$H(4f_0)$	—	–44	—	dBc
Noise power in RX (935–960 MHz) P_{in} = 0 dBm, P_{out} = 35 dBm, 100 kHz RBW	N_{RX}	—	–81	—	dBm
Stability *all spurious outputs* < –60 dBc, *VSWR load, all phase angles*		—	10:1	—	—
Input VSWR		—	1.7:1	—	—

Table 3-6 CGY94 electrical characteristics

T_A = 25 °C, f = 0.9 GHz, $Z_S = Z_L$ = 50 Ω, V_D = 3.6 V, V_G = 4 V, VTR pin connected to ground, unless otherwise specified; pulsed with a duty cycle of 10%, ton = 0.33 ms)

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
Supply current V_D = 3.0 V P_{in} = 10 dBm	I_{DD}	—	1.18	—	A
Negative supply current *(normal operation)*	I_G	—	2	—	mA
Shut-off current *VTR n.c.*	I_D	—	400	—	µA
Negative supply current (*shut off mode, VTR pin n.c.*)	I_G	—	10	—	µA
Gain P_{in} = *5* dBm	G	27.0	29.0	—	dB
Power gain V_D = 3.6 V; P_{in} = 10 dBm	G	22.8	23.6	—	dB
Output power V_D = 3.0 V*;* P_{in} = 10 dBm	P_O	31.5	32.3	—	dBm
Output power V_D = 3.6 V*;* P_{in} = 10 dBm	P_O	32.8	33.6	—	dBm
Output power V_D = 5 *V;* P_{in} = 10 dBm	P_O	34.5	35.5	—	dBm
Overall power added efficiency V_D = 3.0 V; P_{in} = 10 dBm	η	43	48	—	%
Overall power added efficiency V_D = 3.6 V; P_{in} = 10 dBm	η	42	47	—	%
Overall power added efficiency V_D = 5 V; P_{in} = 10 dBm	η	41	46	—	%
Harmonics (P_{in} = 10 dBm, *CW*) $2f_0$	—	—	–49	—	dBc
V_D = 3.6 V; (P_{out} = 33.1 dBm) $3f_0$	—	—	–45	—	dBc
Input VSWR V_D = 3.6 V	—	—	1.5:1	2.0:1	—

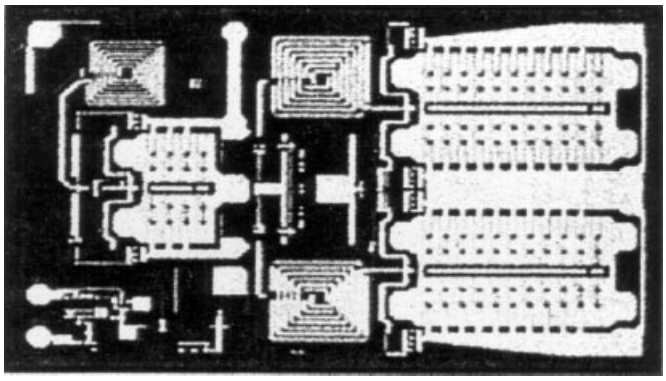


Figure 3-32 The CGY94 amplifier chip.

is often filtered, and several types of filters are available. Here a root-raised cosine filter is used with a cutoff frequency of 665 kHz and a roll-off factor of 0.35.

Analysis. Figure 3-34 shows the compression characteristics of the amplifier as the available RF source power (a sinusoid) is swept. Note that P_{-1dB} is 11.5 dBm referred to the input and 28.5 dBm referred to the output. We will look at modulation characteristics in the linear region of operation at a source power of 0 dBm, and the nonlinear region at a source power of 20 dBm.

Next, we will view the voltage across the transistor to examine its behavior under compression. Figure 3-35 shows the drain–source voltage of the FET as the source power is swept. It is clear from this graph that clipping due to pinchoff and forward-gate conduction

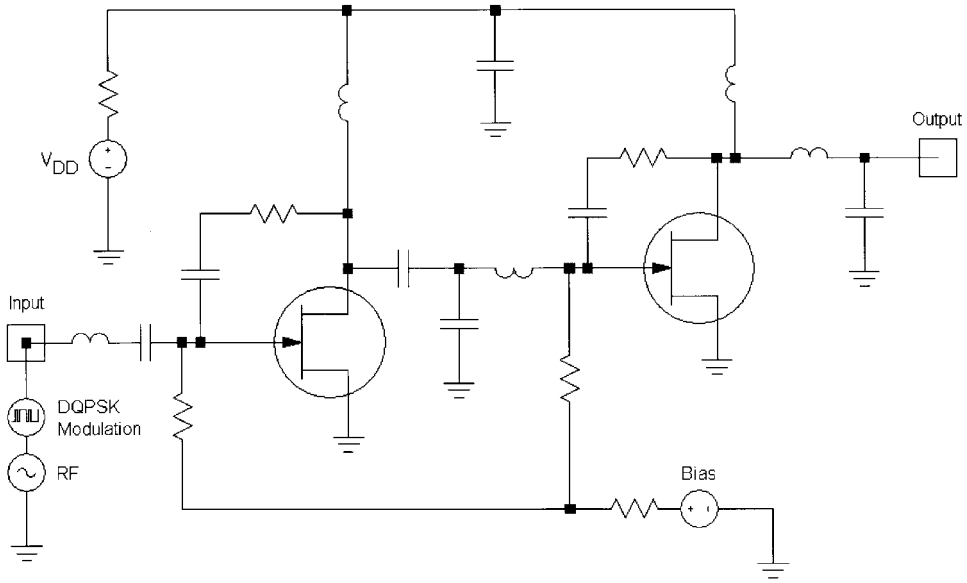


Figure 3-33 Schematic of the CGY94 GaAs MMIC power amplifier.

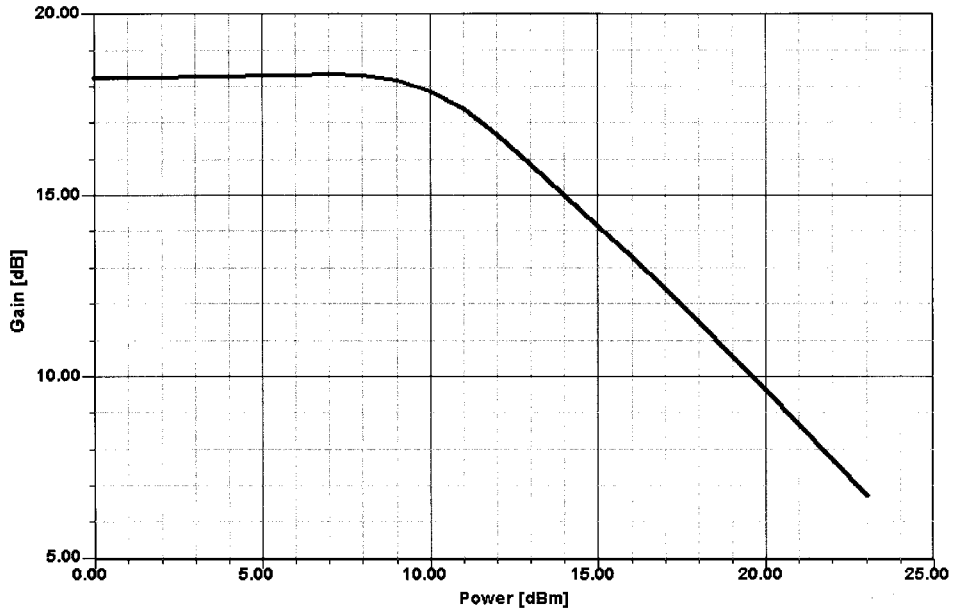


Figure 3-34 Single-tone RF power sweep analysis of the FET amplifier.

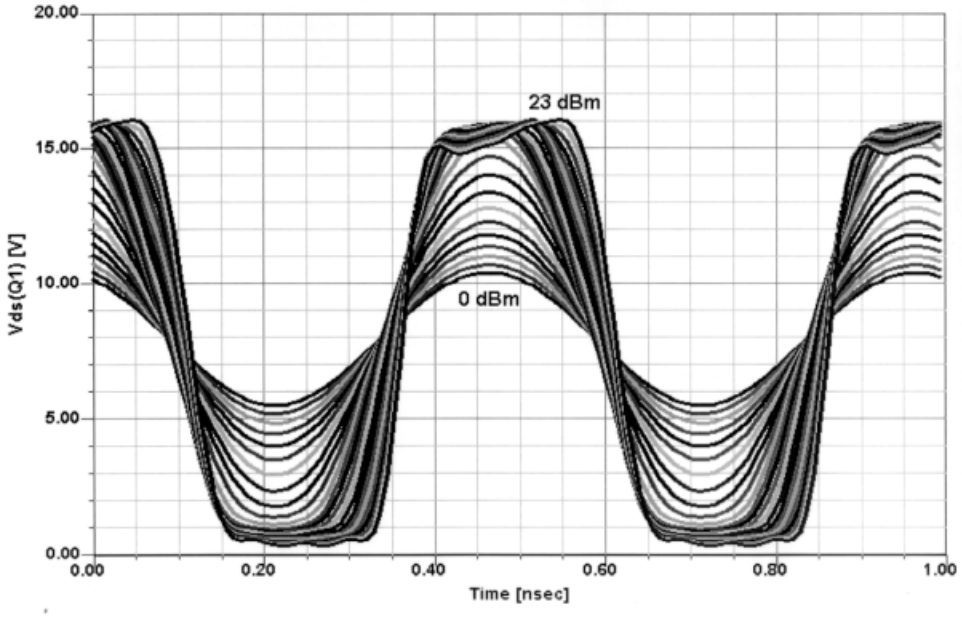


Figure 3-35 V_{DS} of the FET versus RF source power.

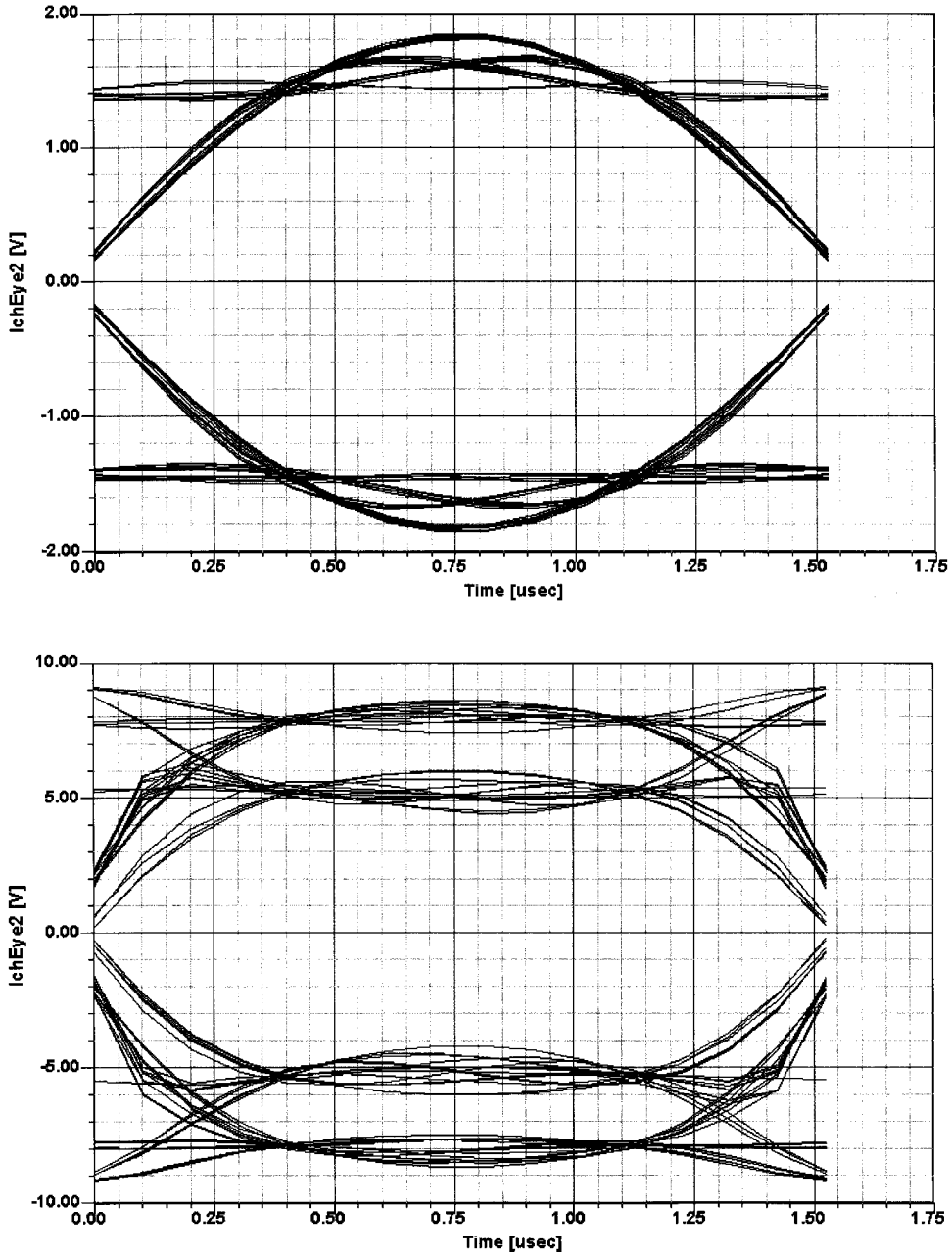


Figure 3-36 Eye diagrams of the I channel for source powers of (a) 0 dBm and (b) 20 dBm.

is limiting the performance of the device. A phase shift of the voltage waveform as power is increased is also apparent. The combination of the power compression and phase shift, or AM to PM conversion, will be used in the modulation analysis to determine the overall distortion of the modulated signal.

The output available from modulation analysis includes:

- I and Q channel waveforms
- Eye and constellation diagrams
- Spectral plots
- ACPR, in-band, and adjacent power

Figure 3-36 shows the eye diagrams for the 128-bit QPSK signal at two RF source powers, 0 and 20 dBm. Operating linearly at 0 dBm input, the amplifier doesn't distort the eye and it remains wide open. Note that in this example the filter bandwidth doesn't produce any intersymbol interference, as evidenced by the open eye. At an input power of 20 dBm, the amplifier compresses the signal, and distortion in the eye is evident.

We can also view the modulation spectrum at these two power levels to investigate the intermodulation distortion and spectral regrowth that takes place. Figure 3-37 shows the spectral plots. The lower trace corresponds to 0 dBm source power and shows almost no regrowth, while the upper trace at 20 dBm source power shows considerable regrowth.

The uncorrupted modulation source can also be shown. This corresponds to the source as a pure voltage without any circuit influence.

Now we evaluate the CGY94 board in a similar fashion. Figure 3-38 shows the waveform used for CAD purposes. Needless to say, such a signal cannot be produced by any signal

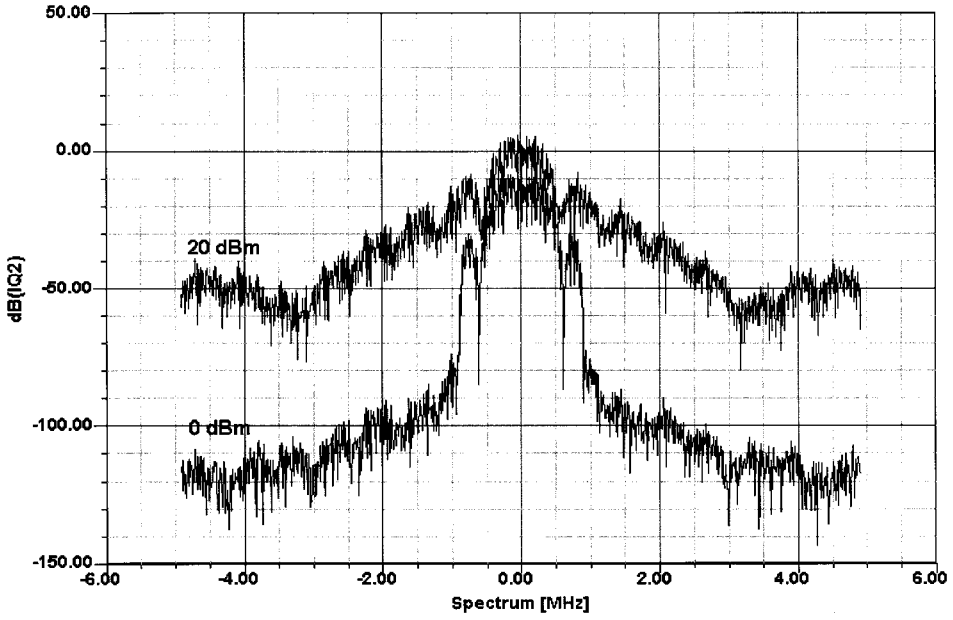


Figure 3-37 Spectrum of the modulation signal at the amplifier output. Source powers are 0 and 20 dBm.

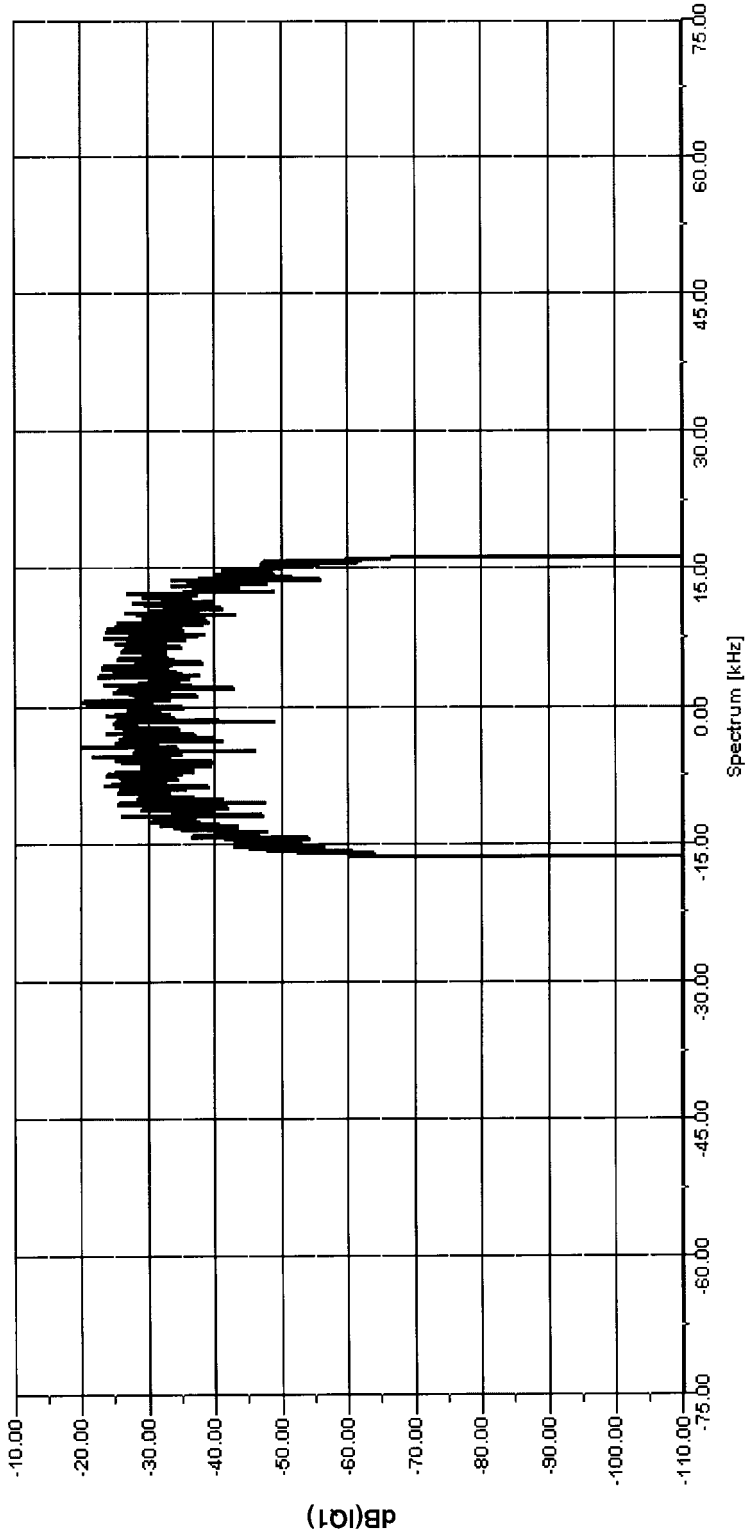


Figure 3-38 Simulated signal used for CAD analysis of the CGY94 amplifier.

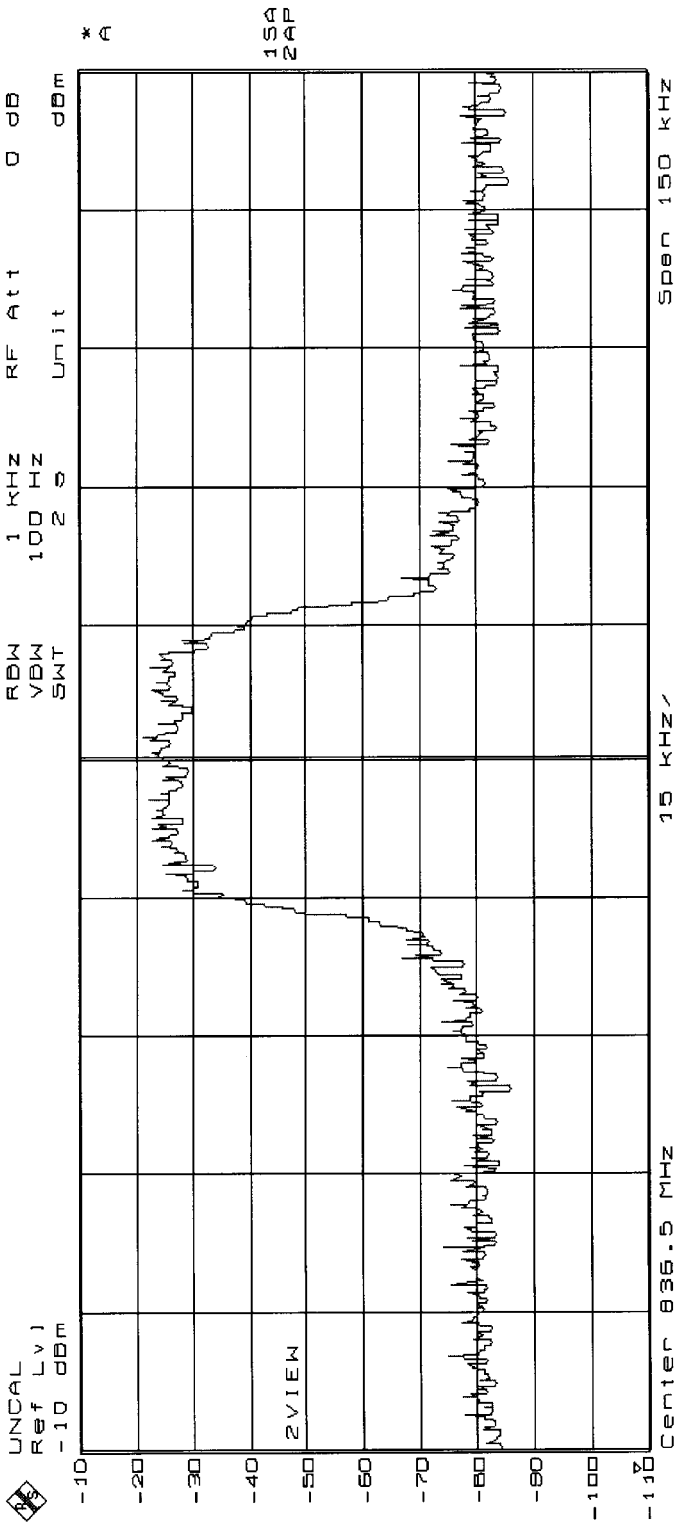


Figure 3-39 Real signal used for testing the actual amplifier.

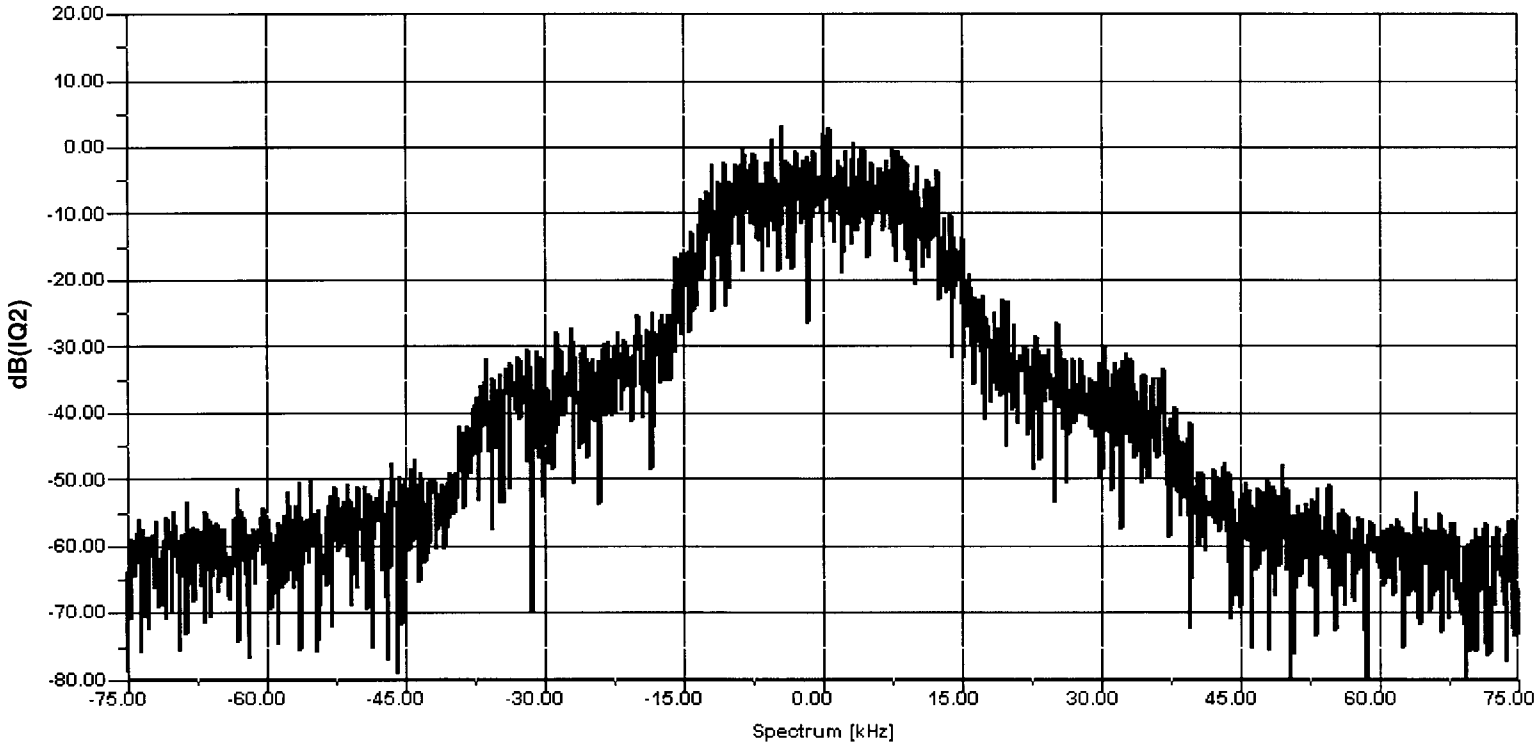


Figure 3-40 Spectral regrowth predicted by the simulation.

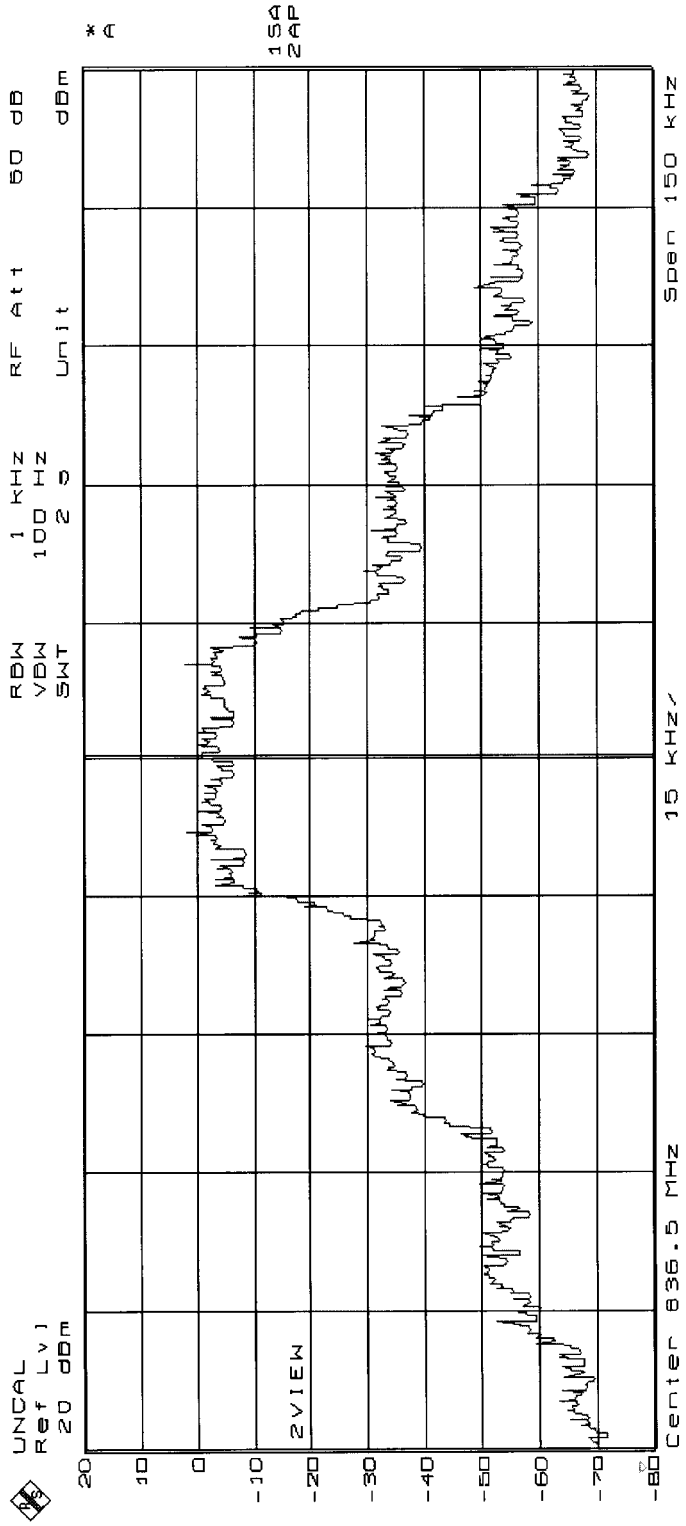


Figure 3-41 Measured spectral regrowth of the CGY94 amplifier.

generator. The one generated by an actual signal generator is shown in Figure 3-39. The wideband noise is around –80 dBm and the input level is about –25 dBm, consistent with the simulation signal. Looking at the output, Figure 3-40 shows the simulated signal with the expected increase in bandwidth. This is due to the nonlinearity of the amplifier and is consistent with Figure 3-37 for the single-stage amplifier. The actual measured output for the CGY94 shown in Figure 3-41 is quite close to its simulated output, indicating that the mathematical approach used for simulating is correct. This is only one example; in reality, many more waveforms can be analyzed and predicted using this method. Figure 3-41 shows essentially three distinct steps, which are sufficiently close to the response shown in the simulation. A dynamic range of 70 dB, while displayed in Figures 3-40 and 3-41, is not really necessary for good-quality transmission, and the discrepancy between the two pictures is in part due to the limited modeling quality of the active device as provided by the company that did the parameter extraction for this transistor.

Another useful view to aid in the understanding of the compression and distortion within the amplifier is to look at the magnitude of the complex waveform. Since the signal is composed of the in-phase and quadrature-phase components it is represented as

$$s(t) = i(t) + j \cdot q(t) \tag{3-120}$$

The time-domain magnitude is written

$$|s(t)| = \sqrt{i(t)^2 + q(t)^2} \tag{3-121}$$

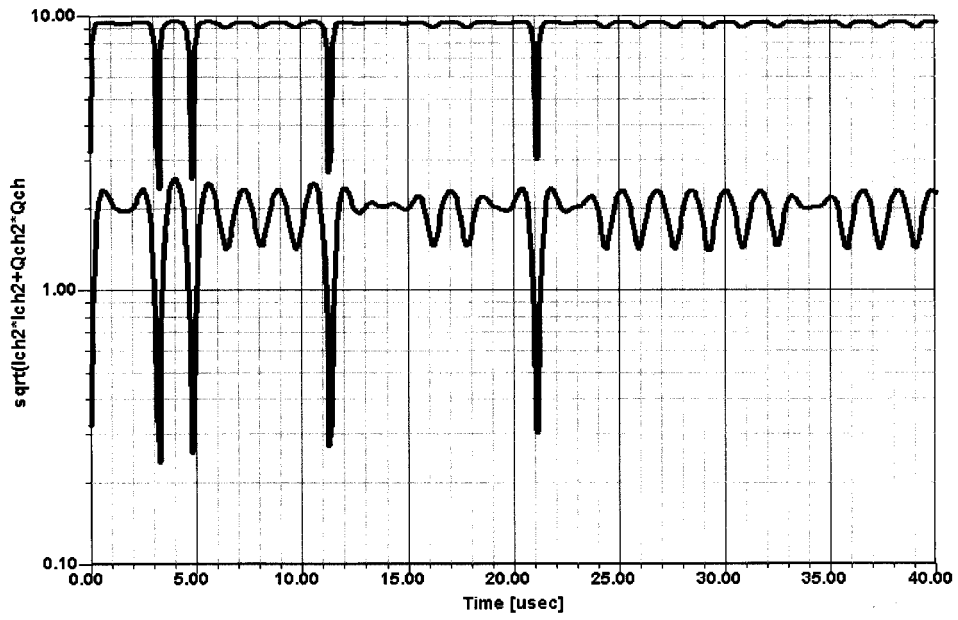


Figure 3-42 Time-domain magnitude of the complex modulation signal at source powers of 0 and 20 dBm.

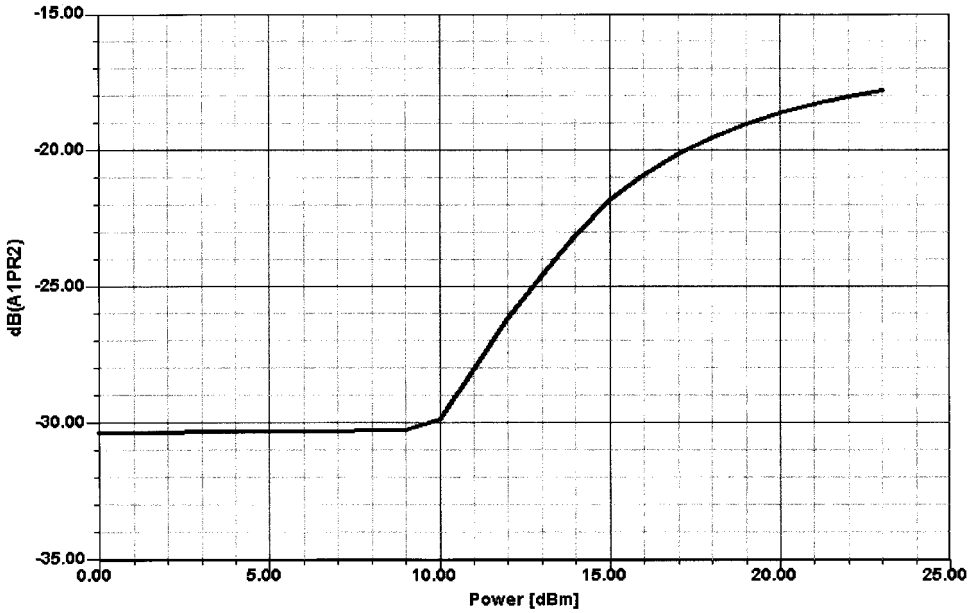


Figure 3-43 Adjacent-channel power ratio as function of RF source power.

Figure 3-42 shows the magnitude at source powers of 0 and 20 dBm. The *X* axis has been rescaled to 40 μs for improved viewing. Note the significant compression of the 20-dBm waveform at the higher signal levels. It is clear from this view that the signal is severely distorted.

We can also compute the adjacent-channel power ratio (ACPR) against the RF power sweep. ACPR is the ratio of the adjacent-channel power to the in-band channel power. The bandwidth (BW) and adjacent-channel start frequency (FS) are used for this calculation. For accurate ACPR calculation, a large number of bits are needed. This example uses 128 bits so the computation time is short, but 512 or more bits should be used for a more accurate computation.

Figure 3-43 shows the ACPR as a function of RF source power. The ACPR is nearly constant (and nonzero due to the gradual skirt of the baseband filter and inherent spillover to the adjacent channel) up to the P_{-1dB} compression point . As the intermodulation products spill into the adjacent channel and spectral regrowth occurs, the ACPR degrades.

π/4-DQPSK Circuit Analysis. Next, we will examine the same amplifier using a π/4-DQPSK modulation source. The task is identical to the previous one except for the source and the number of bits. The number of bits for this project has been increased from 128 to 256. Figure 3-44 shows the eye diagrams from this modulation format, at source powers of 0 and 20 dBm. The distortion and related intersymbol interference are clearly evident.

The constellation plot at 0 dBm is shown in Figure 3-45.

It may be interesting to take a look at the different waveforms currently in use. Figure 3-46 shows a spectrum analyzer picture of the most popular ones.

For reasons of time-division management, the European standard, GSM, which is now offered by Omnipoint at the 1.8-GHz band in the United States, uses gating to reduce the actual bandwidth of the spectrum. Figure 3-47 compares the GSM signal with and without gating.

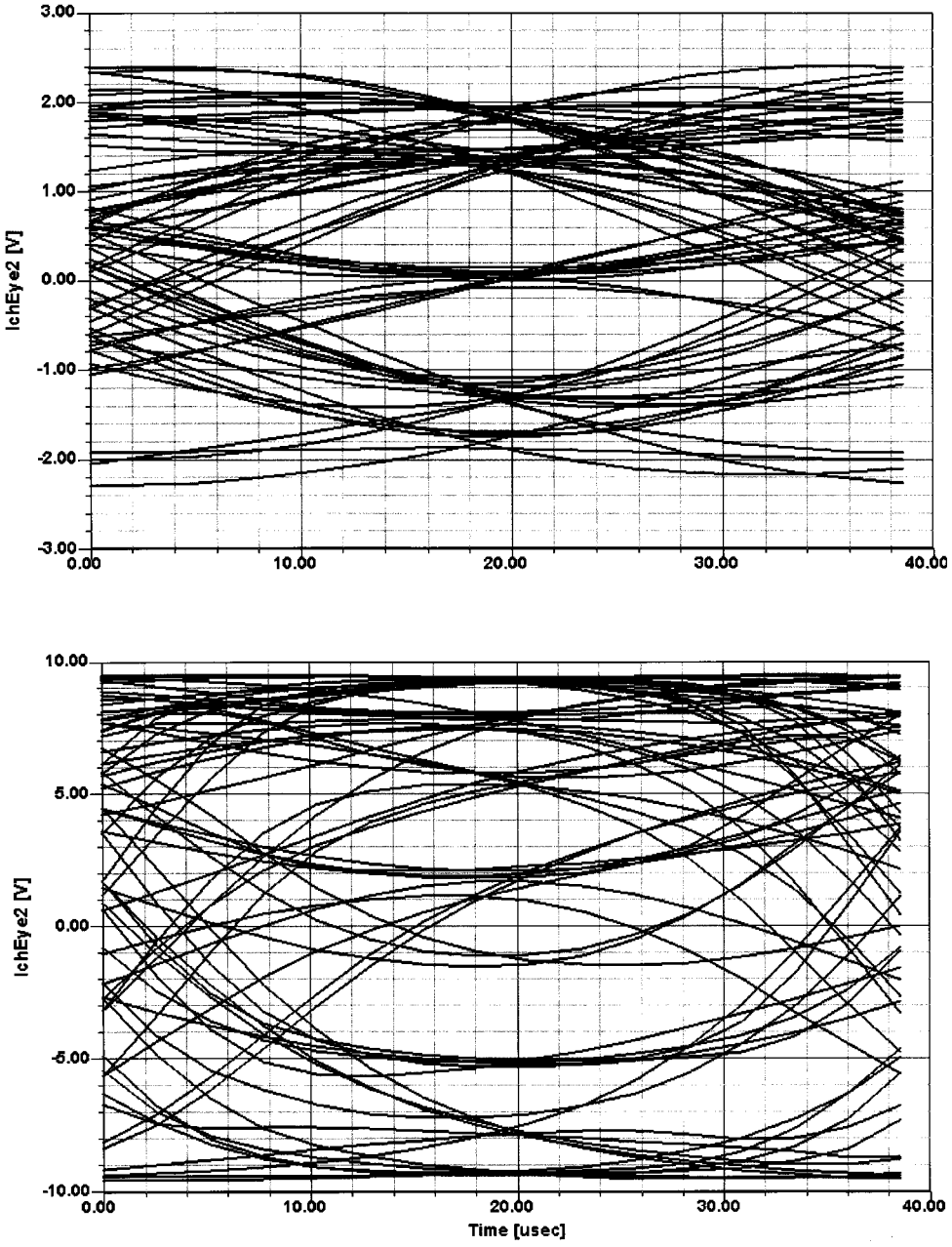


Figure 3-44 Eye diagrams for $\pi/4$-DQPSK at source powers of (a) 0 dBm and (b) 20 dBm. The distortion and related intersymbol interference are clearly visible.

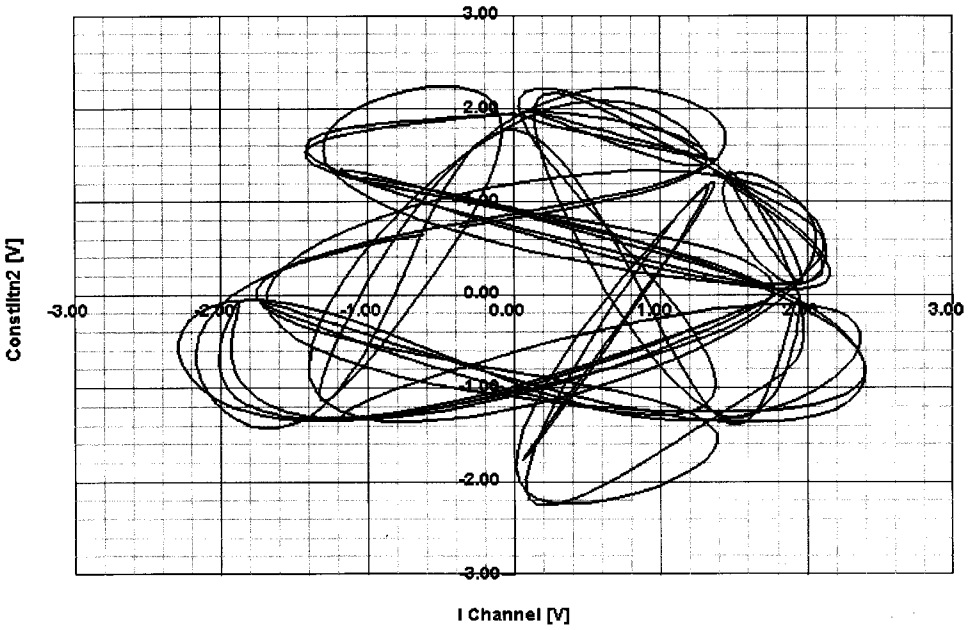


Figure 3-45 Constellation diagram for π/4-DQPSK at Port 2 and 0-dBm source power.

3-1-5 AGC

Many wireless systems consist of an up conversion or down conversion using an intermediate frequency, typically 10% or less of the receiving frequency (the FM broadcast IF has been set at 10.7 MHz for the same reason).

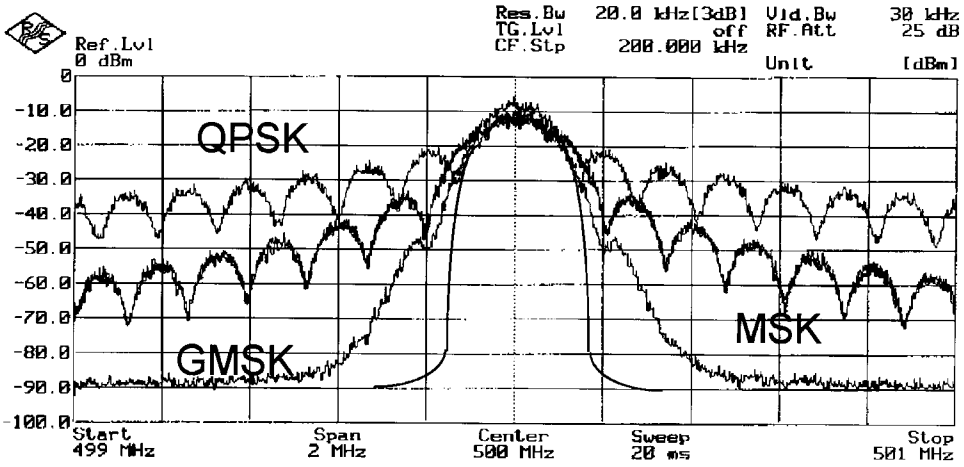


Figure 3-46 Spectrogram showing the characteristics of MSK, GMSK, and QPSK.

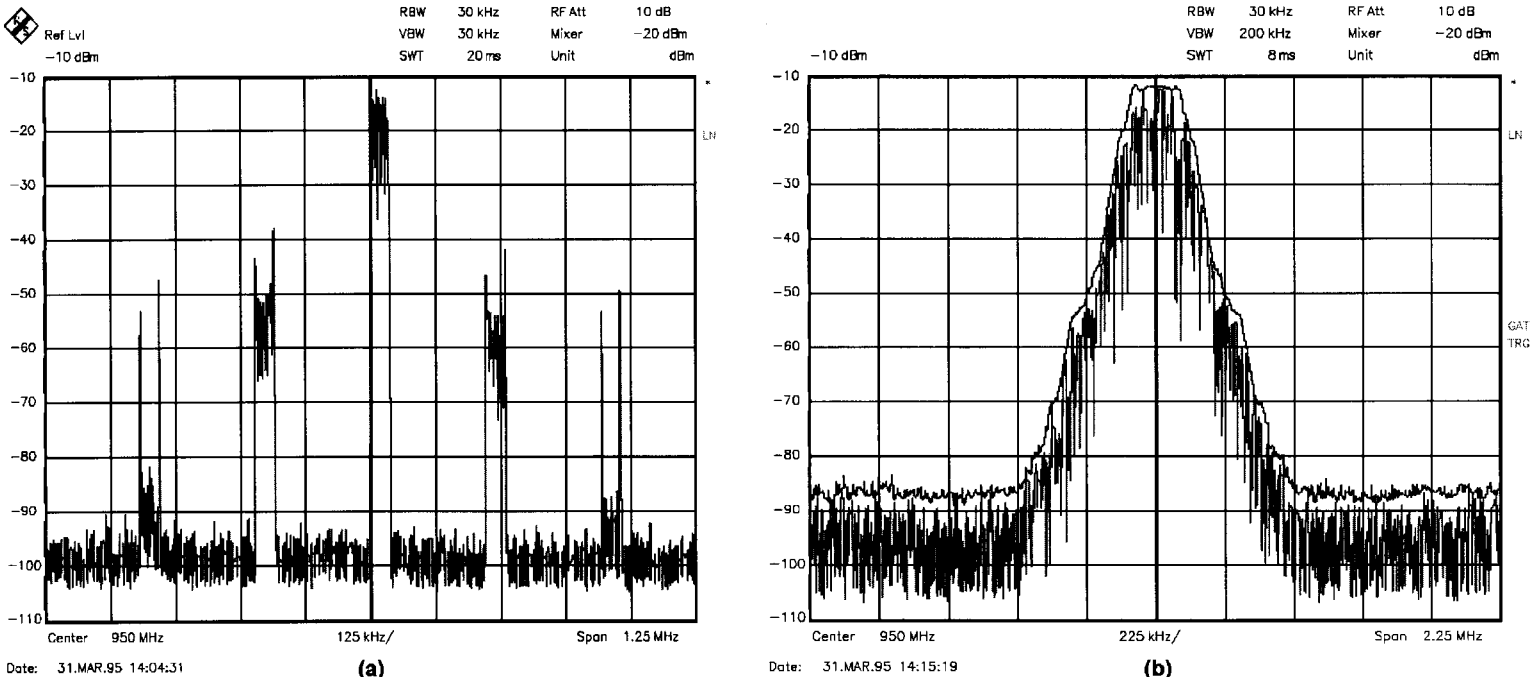


Figure 3-47 Pulsed GSM signal (a) without gating and (b) with gating.

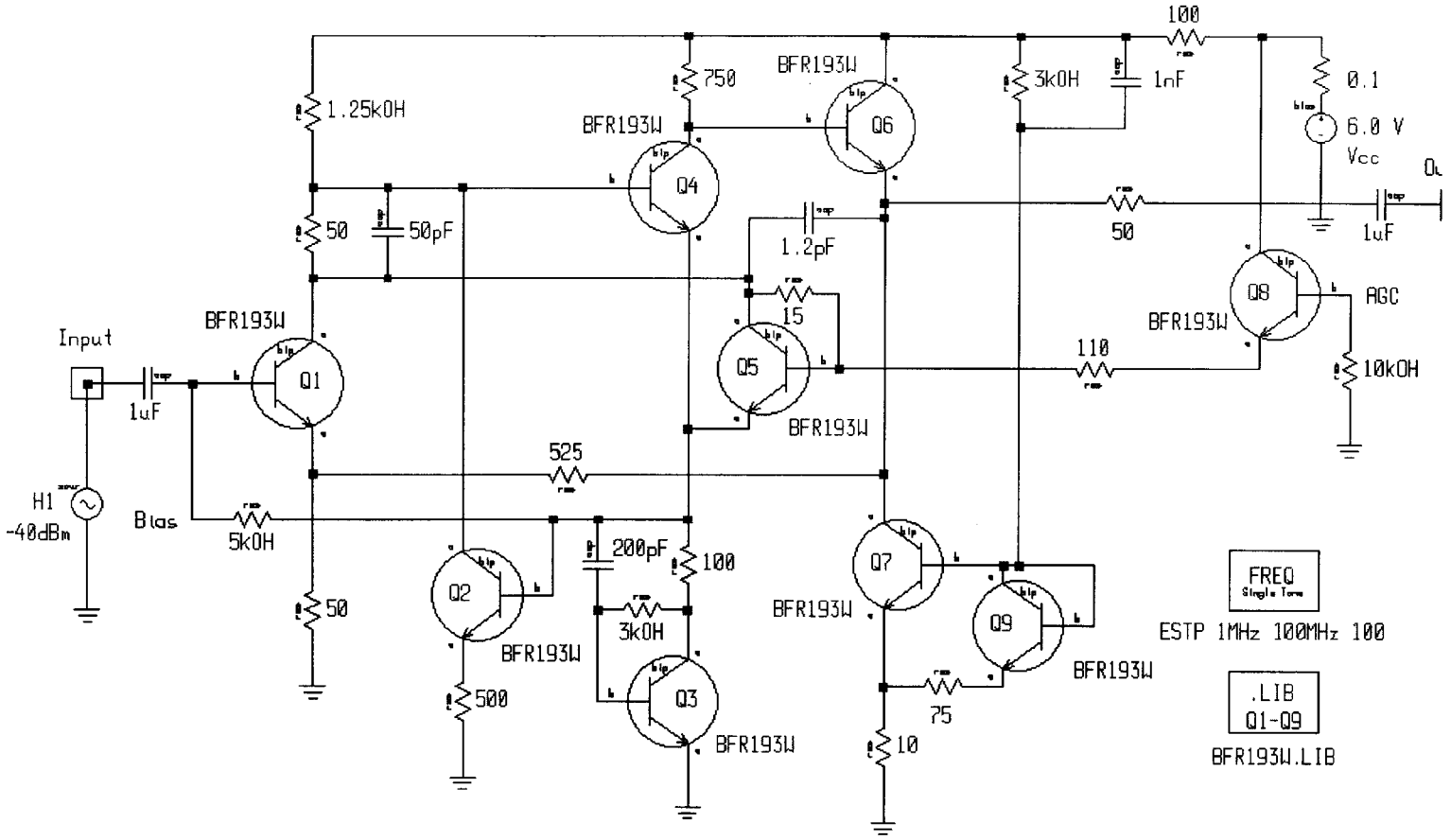


Figure 3-48 Schematic of the Plessey SL610 wideband amplifier with AGC. The 75-Ω resistor in the emitter of Q9 had to be added to obtain the appropriate dc current.

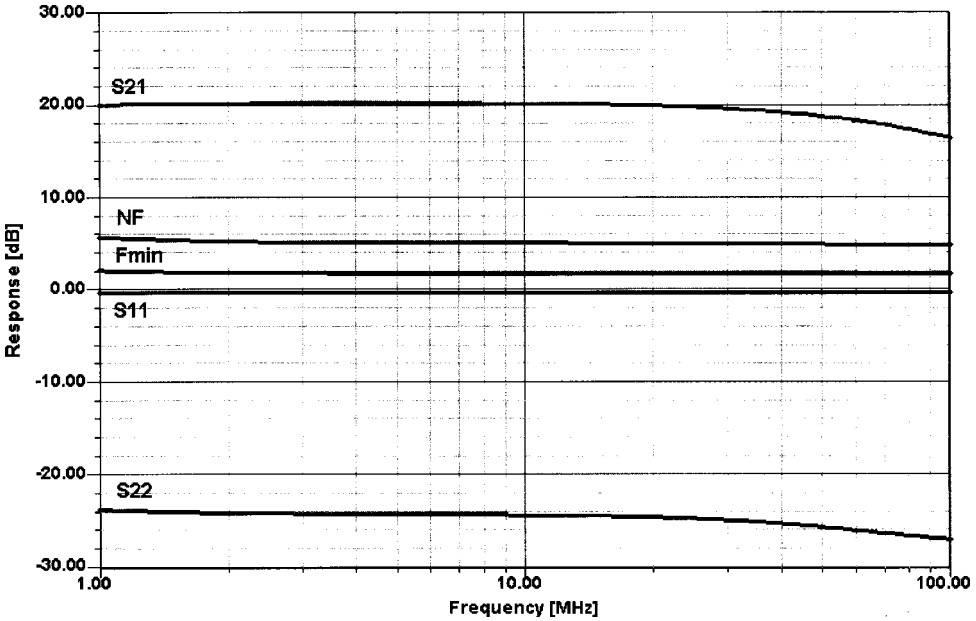


Figure 3-49 Simulated frequency response of the Plessey SL610 wideband amplifier IC with AGC = 0 V.

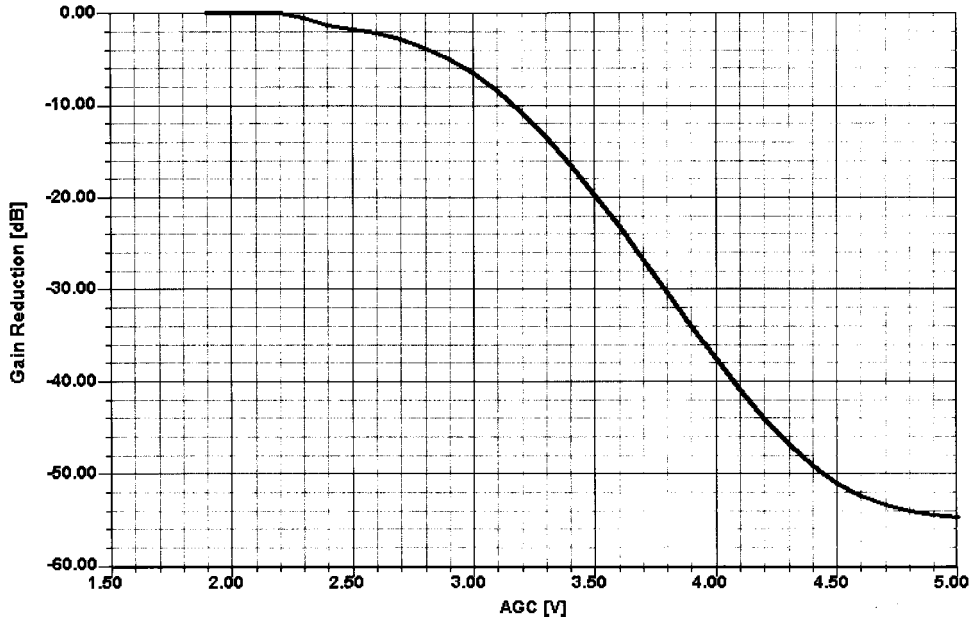


Figure 3-50 Simulated gain reduction versus AGC of the Plessey SL610 wideband amplifier IC.

Table 3-7 CGY121A features

- Variable gain amplifier (MMIC amplifier) for mobile communication
- Typical gain control range over 50 dB
- Positive control voltage
- 50-Ω input and output matched
- Low power consumption
- Operating voltage range: 2.7–6 V
- Frequency range 800 MHz to 2.5 GHz

We therefore will show two amplifiers with AGC and available data. The first one is a wideband amplifier type SL610, which has been manufactured by Plessey; its current fate is unknown. However, since we put the circuit into the simulator it provides significant insight into the operation of such AGC stages. Figure 3-48 shows the actual schematic of the SL610 entered in Ansoft's Serenade Design Environment for simulation.

The AGC action is derived from Q8 through Q5, which, together with Q4, forms a differential amplifier. Each consecutive stage (Q1, Q4, Q6) has increasing current, totaling about 20 mA; the last transistor takes about 13 mA. Based on the feedback surrounding the transistors, the actual transistor characteristics are less important. We have shown this by putting much higher f_T devices into the circuit than were available for the original design (20 years ago). The circuit is still interesting today because of its good large-signal-handling capability.

Table 3-8 CGY121A electrical characteristics at 900 MHz and 1.8 GHz

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
(T_A 25 °C, f = 900 MHz, V_g = –4 V, $RS = RL$ = 50 Ω unless otherwise specified)					
Power gain V_d = 3 V; I = 45 mA; V_{con} = 3 V	G	17	19	—	dB
Input return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{in}	—	11	—	dB
Output return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{out}	—	10	—	dB
Gain control range V_{con} = 3 V ... 0 V; V_d = 3 V; I = 45 mA	dG	48	53	—	dB
1-dB gain compression V_d = 3 V; I = 45 mA; V_{con} = 3 V	P_{1dB}	—	14	—	dBm
(T_A 25 °C, f = 1800 MHz, V_g = –4 V, $RS = RL$ = 50 Ω unless otherwise specified)					
Power gain V_d = 3 V; I = 45 mA; V_{con} = 3 V	G	15.5	17.5	—	dB
Input return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{in}	—	10	—	dB
Output return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{out}	—	8.5	—	dB
Gain control range V_{con} = 3 V ... 0 V; V_d = 3 V; I = 45 mA	dG	48	53	—	dB
1-dB gain compression V_d = 3 V; I = 45 mA; V_{con} = 3 V	P_{1dB}	—	14	—	dBm

Figure 3-49 shows the amplifier's frequency response which is extremely close to the measured data. The good simulation also can be seen from comparing the published AGC curve with the one we found in the simulation. The bias of the first transistor remains essentially constant, resulting in a noise figure more independent of the AGC than the PIN-diode attenuator, and also stabilizes the input impedance as a function of AGC. Figure 3-50 shows the SL610's predicted AGC response.

The same requirement is applicable for the wireless frequency range. Tables 3-7 and 3-8 and Figures 3-51 through 3-54 present specifications for the recently introduced CGY121A GaAs MMIC by Infineon Technologies AG, formerly the Semiconductor Group part of Siemens AG.

3-1-6 Bias and Power Voltage and Current (Power Consumption)

Biasing of transistors operating in Class A or B follows some standard rules. The only deviation from this tends to be when the operating voltage is very low, below 3 V, because then the luxury of using voltage drops for good dc stability as a function of temperature disappears. (We will discuss low-voltage design in detail in Section 3-2-6.) In our simple example (Figure 3-55), we are operating from a 12-V power supply, and the first rule of thumb is to make sure that the voltage between the emitter and ground is at least 0.7 V. This is necessary to compensate the base–emitter junction threshold voltage, which decreases by about 2 mV/°C as the temperature increases. A larger voltage drop in the emitter–ground connection greatly reduces the influence. In this amplifier the emitter–ground path consists of an emitter unbypassed resistor of 10 Ω in series with a bypassed 501-Ω resistor, resulting in an overall voltage drop of 2.6. We could reduce the supply voltage by about 1.9 V at the collector side and still maintain a safety margin of 0.7 V at the emitter–ground connection. The current 2.6 V, so to speak, is overkill.

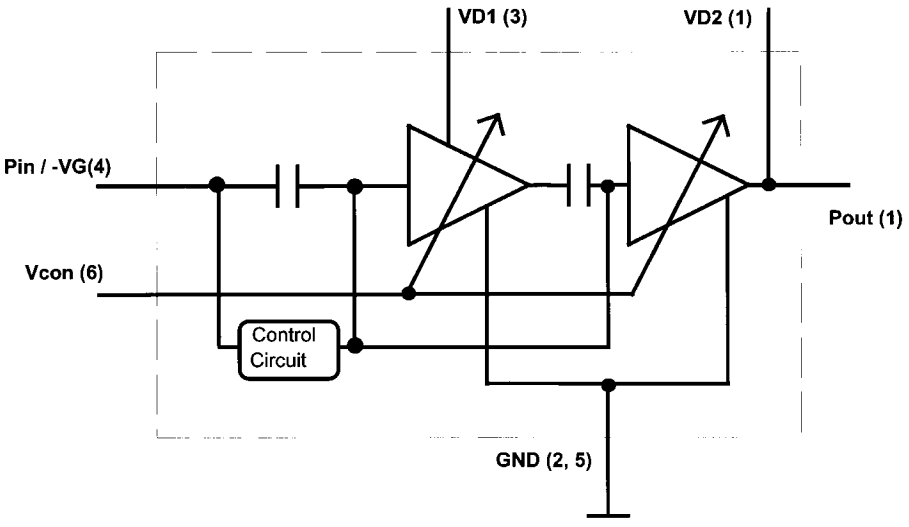
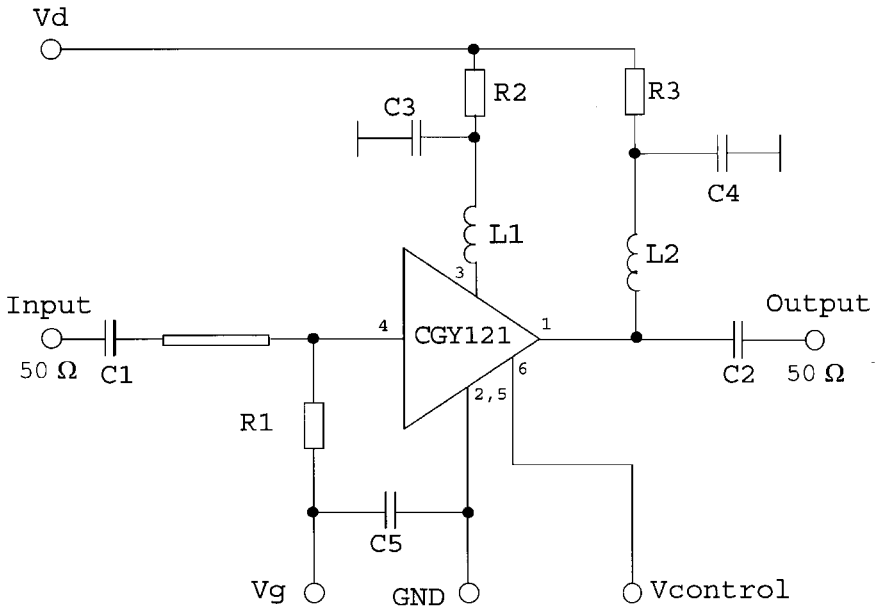


Figure 3-51 Functional block diagram of the gain-controllable CGY121A GaAs MMIC. Gain control is achieved by applying 0–3 V dc to the Vcon pin.



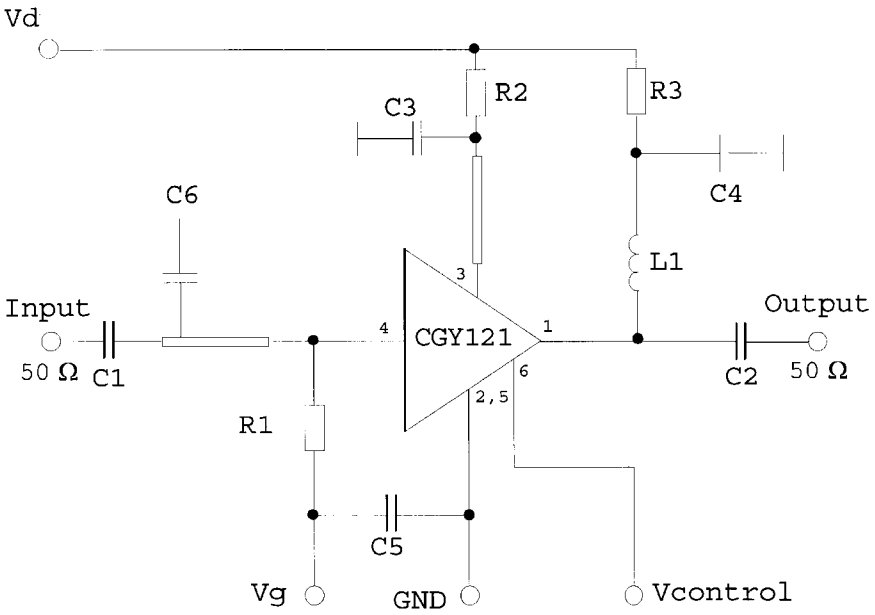
Parts List

Frequency	900 MHz
C1, C2 (Siemens Size 0603)	22 pF
C3, C4 (Siemens Size 0603)	100 nF
C5 (Siemens Size 0603)	47 nF
L1 (Coilcraft 0805CS-150XKBC)	15 nH
L2 (Coilcraft 0805CS-270XMBC)	27 nH
R1 (Siemens B 54102-A1271-J60)	270 Ω
R2 (Siemens B 54102-A1120-J60)	12 Ω
R3	6.8 Ω

Figure 3-52 CGY121A 900-MHz application circuit and parts list.

On the collector side, for an *RC* type of amplifier, it is not a bad idea to set the collector voltage at about half the supply, which allows the RF swing to be half the supply voltage. In this case, we have violated the rule because of our 2.6-V drop. Assuming we had chosen a 0.7-V drop, then the collector–emitter voltage would be 6.3 V available for the swing. The choice of 5 mA for the transistor has come from a combination of good linearity and low noise figure. While the 10-Ω unbypassed resistor will grade the noise figure somewhat, it reduces the emitter-current-dependent distortion. The base–emitter diffusion resistance equals 26 mV/5 mA, or roughly 5 Ω. The 10-Ω resistor is twofold larger and therefore adds to the linearity and to the noise.

To determine the base bias resistor it is a good assumption that the current going through the two resistors (in our case, 8.2 and 3.3 kΩ) is about 10% (or a slightly higher fraction) of



Parts List

Frequency	1.9 GHz
C1, C2 (Siemens size 0603)	12 pF
C3, C4 (Siemens size 0603)	100 nF
C5 (Siemens size 0603)	47 nF
C6 (Siemens size 0603)	1.2 pF
L1 (Coilcraft 0805CS-150XKBC)	15 nH
R1 (Siemens B 54102-A1271-J60)	270 Ω
R2 (Siemens B 54102-A1120-J60)	12 Ω
R3	2.7 Ω

Figure 3-53 CGY121A 1.9-GHz application circuit and parts list.

the collector current. Since the dc current gain of the transistor is typically between 50 and 100, it is sufficient if the current through the resistive network is between 5 and 10 times higher than the dc bias current taken off the network. This type of bias scheme is only valid for Class A to AB_1. Later, we will see that for monolithic circuits, additional transistors will be used to generate the required voltage drops. Lately, combinations of *npn* and *pnp* transistors have become popular to avoid using any resistance in the emitter for dc biasing purposes. Figure 3-56 shows such a circuit. This, of course, invites thermal runaway. To minimize the circuitry, there are bias ICs, such as the Siemens BCR400R, now available. Figures 3-57 and 3-58 show its application in biasing BJTs and FETs, respectively; it can also be used to control PIN-diode bias current for TR antenna switching, as shown in Figure 3-59.

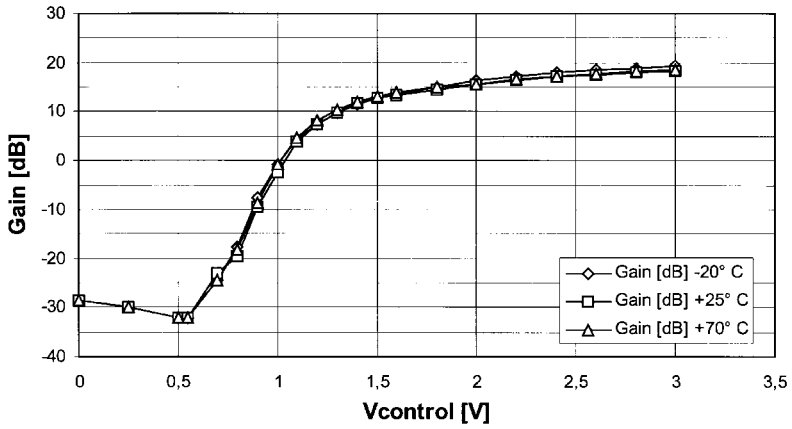


Figure 3-54 CGY121A gain versus $V_{control}$ at 1.9 GHz.

To validate this, we have used the Serenade 8.0 CAD tool, which has a built-in bias- and temperature-dependent noise model for BJTs and FETs. The SPICE-type dc analysis determines the actual currents, which were mentioned above. By maintaining the same bias and just bypassing both emitter resistors at the same time, the set of curves shown in Figure 3-60 will be obtained. As an exercise, we recommend the user to determine the optimum generator impedance for F_{min} with a 10-Ω resistor. Again, both the emitter current and collector voltage affect the noise figure, the intermodulation, and f_T. If a transistor is used

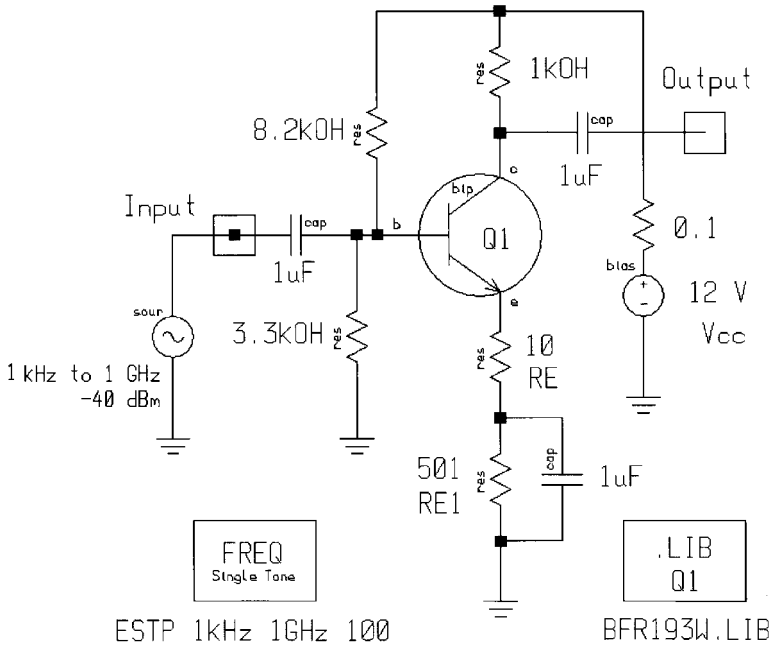


Figure 3-55 BJT amplifier example.

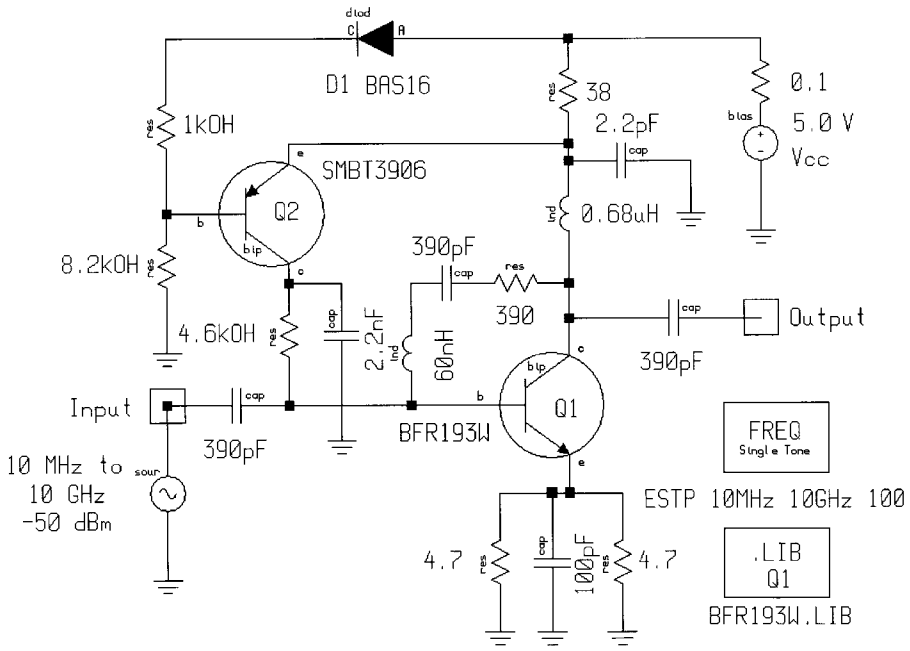


Figure 3-56 Wideband amplifier with active biasing. The 4.7-Ω resistors in the BFR193W emitter compensate the circuit's gain at low frequencies.

that is intended to be operated at significantly higher currents, then f_T will suffer while the other two parameters are less affected.

The dc input power translates into thermal dissipation, and Figure 3-60 is a plot of the noise figure as a function of temperature with and without emitter feedback. The heavy dc stabilization prevents a significant effect on the other parameters. The actual dc shift is from 5.0 to 5.2 mA for a temperature change of 70 K.

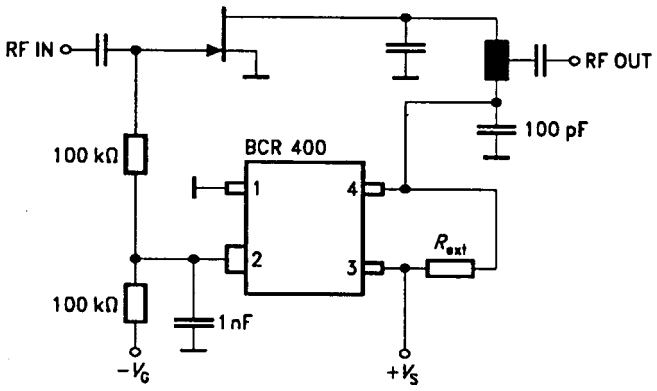


Figure 3-57 The Siemens BCR400 bias controller applied to a GaAsFET.

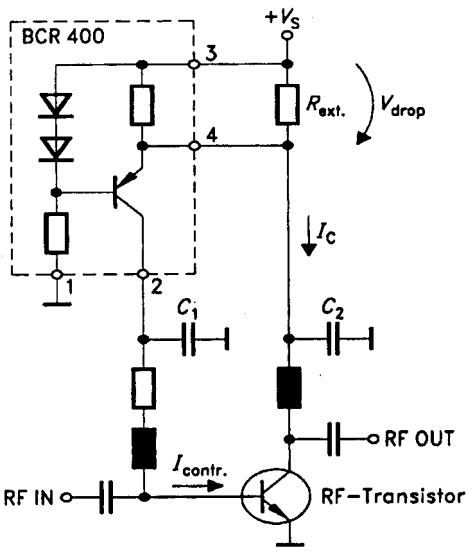


Figure 3-58 The Siemens BCR400 bias controller applied to a BJT.

As far as FETs are concerned, the enhancement types, typically PMOS, can use a similar type of bias, while the NMOS types complicate the designer's life as they require a negative gate voltage. For low-noise operation, these devices are operated somewhat close to the pinchoff voltage. Putting a source resistor in place causes mostly headaches simply because the gain is so high that the transistors are always ready to find a frequency at which they love to oscillate. The transistor will look for every possible parasitic to turn the amplifier into an oscillator.

3-2 AMPLIFIER GAIN, STABILITY, AND MATCHING

Most semiconductor manufacturers describe their devices these days by providing either their *S* parameters (small-signal) or SPICE-type parameters (large-signal). One of our headaches in writing this book has been that many of the published models don't translate

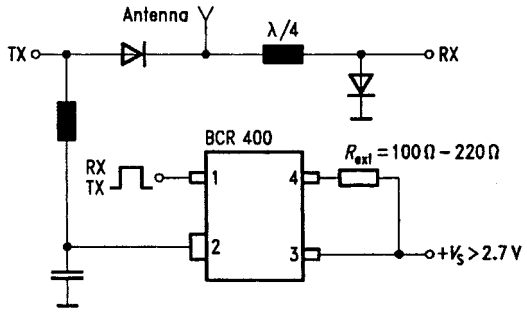


Figure 3-59 The Siemens BCR400 bias controller applied to TR switching.

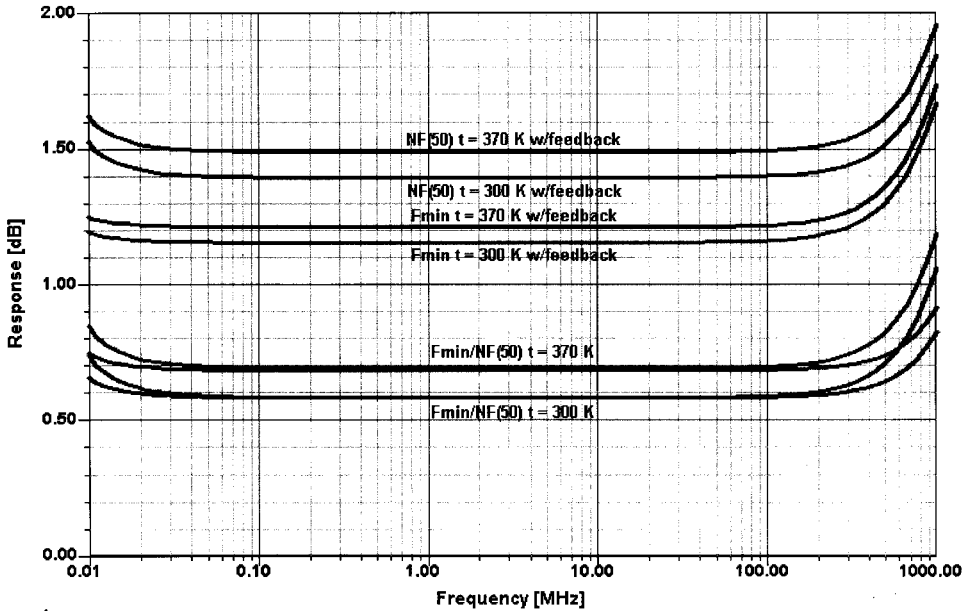


Figure 3-60 F_{min} and NF for the amplifier in Figure 3-55, with and without emitter feedback, at 300 and 370 K.

from large signal to small signal without giving unexpected or false results. Despite the availability of tools to do it right, many companies still can only provide data that are practically inadequate, and unfortunately the CAD manufacturers then get the blame for results for which they are not really responsible.

Tables 3-9 and 3-10 show typical data for a bipolar transistor (the BFP420 by Siemens). The next section presents a summary of the derivation and use of *S* parameters as published by Hewlett-Packard and others.

3-2-1 Scattering Parameter Relationships

Scattering (*S*) parameters characterize a network in terms of incident and reflected waves (Figures 3-61 and 3-62).

$$b_1 = s_{11}a_1 + s_{12}a_2 \tag{3-122}$$

$$b_2 = s_{21}a_1 + s_{22}a_2 \tag{3-123}$$

Notice that

$$s_{11} = \frac{b_1}{a_1} = \frac{V_1/I_1 - Z_0}{V_1/I_1 + Z_0} = \frac{Z_1 - Z_0}{Z_1 + Z_0} \tag{3-124}$$

and

Table 3-9 BFP420 common-emitter *S* parameters[a]

f	S_{11}		S_{21}		S_{12}		S_{22}	
(GHz)	(Magnitude)	(Angle)	(Magnitude)	(Angle)	(Magnitude)	(Angle)	(Magnitude)	(Angle)
				V_{CE} = 2 V, I_C = 20 mA				
0.01	0.452	–2.3	37.62	178.3	0.0011	94.4	0.956	–0.6
0.1	0.447	–52.1	36.30	164.7	0.0068	82.5	0.941	–12.4
0.5	0.386	–101.1	23.41	121.0	0.0262	61.7	0.632	–47.2
1.0	0.378	–146.2	13.99	96.0	0.0395	57.8	0.395	–63.9
2.0	0.405	173.5	7.18	70.8	0.0664	54.0	0.222	–87.3
3.0	0.446	149.4	4.77	52.6	0.0949	47.1	0.133	–111.3
4.0	0.501	130.0	3.52	36.8	0.1206	38.5	0.133	–158.5
6.0	0.599	104.8	2.27	8.2	0.1646	18.9	0.196	142.0
8.0	0.700	78.5	1.51	–20.8	0.1800	–2.4	0.289	99.3
9.0	0.758	67.6	1.25	–34.4	0.1820	–13.0	0.379	84.1
10.0	0.800	62.0	1.04	–43.5	0.1800	–19.3	0.465	76.6
				V_{CE} = 2 V, I_C = 5 mA				
0.01	0.790	–1.0	15.14	179.2	0.0012	83.4	0.988	–0.7
0.1	0.786	–11.6	14.98	171.8	0.0092	84.1	0.982	–6.5
0.5	0.702	–55.7	12.86	140.1	0.0398	62.8	0.857	–29.8
1.0	0.589	–99.1	9.63	112.6	0.0603	46.5	0.647	–48.6
2.0	0.507	–156.0	5.60	79.4	0.0798	34.6	0.401	–70.3
3.0	0.511	168.5	3.84	57.1	0.0957	29.8	0.267	–84.2
4.0	0.549	142.0	2.87	38.5	0.1121	25.1	0.207	–110.5
5.0	0.604	123.9	2.26	22.1	0.1285	19.4	0.150	–137.3
6.0	0.633	110.0	1.86	6.7	0.1442	13.1	0.173	–169.8

Table 3-10 BFP420 noise parameters

f	$F_{\min}$[a]	G_a[a]	Γ_{opt}		R_N	r_n	$F_{50\Omega}$[b]	$\lvert S_{21} \rvert^2$ [b]
(GHZ)	dB	(dB)	(Magnitude)	(Angle)	(Ω)		(dB)	(dB)
				V_{CE} = 2 V, I_C = 5 mA				
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	–175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	–157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	–142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	–123.0	15.0	0.30	3.30	5.3

[a]Input matched for minimum noise figure, output for maximum gain.
[b]$Z_S = Z_L = 50\ \Omega$.

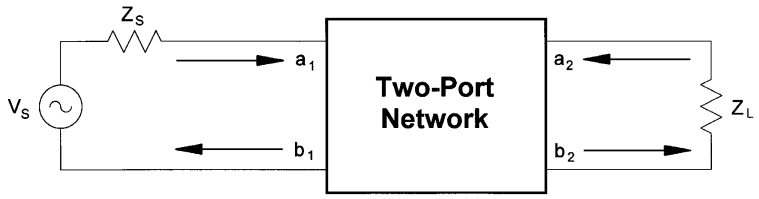


Figure 3-61 *S* parameters characterize a network (a two-port is shown) in terms of incident waves (a_1, a_2) and reflected waves (b_1, b_2). Figure 3-62 shows the flow graph for this network.

$$Z_1 = Z_0 \frac{(1 + s_{11})}{(1 - s_{11})} \tag{3-125}$$

where $Z_1 = V_1/I_1$ is the input impedance at Port 1.

This relationship between reflection coefficient and impedance is the basis of the Smith Chart transmission-line calculator. Consequently, the reflection coefficients S_{11} and S_{22} can be plotted on Smith charts, converted directly to impedance, and easily manipulated to determine matching networks for optimizing a circuit design.

The above equations show one of the important advantages of *S* parameters: They are simply gains and reflection coefficients, both familiar quantities to engineers. By comparison, some of the *Y* parameters described earlier in the article are not so familiar. For example, the *Y* parameter corresponding to insertion gain S_{21} is the "forward transadmittance," Y_{21}. Clearly, insertion gain gives by far the greater insight into the operation of the network.

S parameters are simply related to power gain and mismatch loss, quantities that are often of more interest than the corresponding voltage functions:

$$|s_{11}|^2 = \frac{\text{Power reflected from the network input}}{\text{Power incident on the network input}} \tag{3-126}$$

$$|s_{21}|^2 = \frac{\text{Power delivered to a } Z_0 \text{ load}}{\text{Power available from } Z_0 \text{ source}} \tag{3-127}$$

$$|s_{12}|^2 = \text{Reverse transducer power gain with } Z_0 \text{ load and source.} \tag{3-128}$$

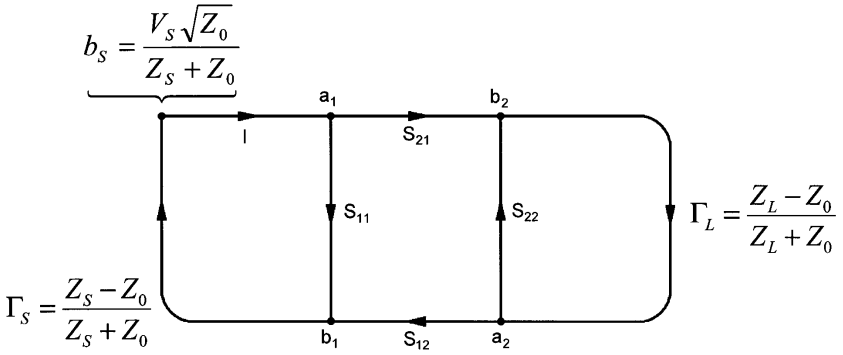


Figure 3-62 Flow graph for the network shown in Figure 3-61.

$$|s_{22}|^2 = \frac{\text{Power reflected from the network output}}{\text{Power incident on the network output}} \tag{3-129}$$

From this we obtain:
Input reflection coefficient with arbitrary Z_L:

$$s'_{11} = s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1 - s_{22}\Gamma_L} \tag{3-130}$$

Output reflection coefficient with arbitrary Z_S:

$$s'_{22} = s_{22} + \frac{s_{12}s_{21}\Gamma_S}{1 - s_{11}\Gamma_S} \tag{3-131}$$

Voltage gain with arbitrary Z_L and Z_S:

$$A_V = \frac{V_2}{V_1} = \frac{s_{21}(1 + \Gamma_L)}{(1 - s_{22}\Gamma_L)(1 + s'_{11})} \tag{3-132}$$

$$\text{Power gain} = \frac{\text{Power delivered to load}}{\text{Power input to network}}:$$

$$G = \frac{|s_{21}|^2 (1 - |\Gamma_L|^2)}{(1 - |s_{11}|^2) + |\Gamma_L|^2 (|s_{22}|^2 - |D|^2) - 2\,\text{Re}\,(\Gamma_L N)} \tag{3-133}$$

$$\text{Available power gain} = \frac{\text{Power available from network}}{\text{Power available from source}}:$$

$$G_A = \frac{|s_{21}|^2 (1 - |\Gamma_S|^2)}{(1 - |s_{22}|^2) + |\Gamma_S|^2 (|s_{11}|^2 - |D|^2) - 2\,\text{Re}\,(\Gamma_S M)} \tag{3-134}$$

$$\text{Transducer power gain} = \frac{\text{Power delivered to load}}{\text{Power available from source}}:$$

$$G_T = \frac{|s_{21}|^2 (1 - |\Gamma_S|^2)(1 - |\Gamma_L|^2)}{|(1 - s_{11}\Gamma_S)(1 - s_{22}\Gamma_L) - s_{12}s_{21}\Gamma_L\Gamma_S|^2} \tag{3-135}$$

Unilateral transducer power gain ($S_{12} = 0$):

$$G_{Tu} = \frac{|s_{21}|^2(1 - |\Gamma_S|^2)(1 - |\Gamma_L|^2)}{|1 - s_{11}\Gamma_S|^2|1 - s_{22}\Gamma_L|^2}$$

$$= G_0 G_1 G_2 \tag{3-136}$$

$$G_0 = |s_{21}|^2 \tag{3-137}$$

$$G_1 = \frac{1 - |\Gamma_S|^2}{|1 - s_{11}\Gamma_S|^2} \tag{3-138}$$

$$G_2 = \frac{1 - |\Gamma_L|^2}{|1 - s_{22}\Gamma_L|^2} \tag{3-139}$$

Maximum unilateral transducer power gain when $|S_{11}| < 1$ and $|S_{22}| < 1$:

$$G_u = \frac{|s_{21}|^2}{|(1 - |s_{11}|^2)(1 - |s_{22}|)^2|}$$

$$= G_0\, G_{1\max}\, G_{2\max} \tag{3-140}$$

$$G_{i\max} = \frac{1}{1 - |s_{ii}|^2} \quad i = 1, 2 \tag{3-141}$$

Constant-gain circles (unilateral case: $S_{12} = 0$):

- Center of constant-gain circle is on line between center of Smith Chart and point representing S_{ii}^*
- Distance of center of circle from center of Smith Chart:

$$r_i = \frac{g_i |s_{ii}|}{1 - |s_{ii}|^2(1 - g_i)} \tag{3-142}$$

- Radius of circle:

$$\rho_i = \frac{\sqrt{1 - g_i}(1 - |s_{ii}|^2)}{1 - |s_{ii}|^2(1 - g_i)} \tag{3-143}$$

where $i = 1, 2$ and

$$g_i = \frac{G_i}{G_{i\max}} = G_i(1 - |s_{ii}|^2) \tag{3-144}$$

Unilateral figure of merit:

$$u = \frac{|s_{11}s_{22}s_{12}s_{21}|}{|(1 - |s_{11}|^2)(1 - |s_{22}|^2)|} \tag{3-145}$$

Error limits of unilateral gain calculation:

$$\frac{1}{(1 + u^2)} < \frac{G_T}{G_{Tu}} < \frac{1}{(1 - u^2)} \tag{3-146}$$

Conditions for absolute stability. No passive source or load will cause a network to oscillate if a, b, and c are all satisfied:

$$\text{a.}\quad |s_{11}| < 1, |s_{22}| < 1 \tag{3-147}$$

$$\text{b.} \quad \left| \frac{|s_{12}s_{21}| - |M^*|}{|s_{11}|^2 - |D|^2} \right| > 1 \tag{3-148}$$

$$\text{c.} \quad \left| \frac{|s_{12}s_{21}| - |N^*|}{|s_{22}|^2 - |D|^2} \right| > 1 \tag{3-149}$$

where

$$D = s_{11}s_{22} - s_{12}s_{21} \tag{3-150}$$

$$M = s_{11} - Ds_{22}^* \tag{3-151}$$

$$N = s_{22} - Ds_{11}^* \tag{3-152}$$

Condition that a two-port network can be simultaneously matched with a positive real source and load:

$$K > 1 \tag{3-153}$$

$$K = \frac{1 + |D|^2 - |s_{11}|^2 - |s_{22}|^2}{2|s_{12}s_{21}|} \tag{3-154}$$

Source and load for simultaneous match

$$\Gamma_{mS} = M^* \left[\frac{B_1 \pm \sqrt{B_1^2 - 4|M|^2}}{2|M|^2} \right] \tag{3-155}$$

$$\Gamma_{mL} = N^* \left[\frac{B_2 \pm \sqrt{B_2^2 - 4|N|^2}}{2|N|^2} \right] \tag{3-156}$$

where

$$B_1 = 1 + |s_{11}|^2 - |s_{22}|^2 - |D|^2 \tag{3-157}$$

$$B_2 = 1 + |s_{22}|^2 - |s_{11}|^2 - |D|^2 \tag{3-158}$$

Maximum available power gain. If $K > 1$,

$$G_{A\max} = \left| \frac{s_{21}}{s_{12}} (K \pm \sqrt{K^2 - 1}) \right| \tag{3-159}$$

In conclusion, *a two-port network (active or passive) is unconditionally stable* when

$$K > 1$$

$$|S_{12}S_{21}| < 1 - |S_{11}|^2$$

and

$$|S_{12}S_{21}| < 1 - |S_{22}|^2$$

3-2-2 Low-Noise Amplifiers

A low-noise amplifier combines a low noise figure, reasonable gain, and stability without oscillation over its entire useful frequency range. These amplifiers are typically operated in Class A, which is characterized by a bias point more or less at the center of maximum current and voltage capability of the device used, and by RF current and voltages that are sufficiently small relative to the bias point that the bias point does not shift. This means that an amplifier operating at, say, 3 V and 1 mA for low-noise, high-gain operation will not be Class A if the RF current is larger than 0.1% of the current or the voltage swing is more than 0.1% of the current swing. These amplifiers can be designed using standard linear *S* parameters, which are (hopefully) available from the device manufacturer for the bias point of interest. If suitable *S* parameters are not available, we must resort to a CAD tool that allows us to determine a device's bias point as a function of its bias network and to generate small-signal *S* parameters for the device. With suitable *S* parameters in hand, and assuming the amplifier will be frequency selective, we then connect tuned circuits to the input and output to provide matching.

In designing amplifiers, the "power gain" of a circuit is mentioned frequently. This is actually deceptive because power gain is somewhat associated with dc power dissipation, and in considering it, we are on the verge of discussing large-signal performance. Therefore it would be better to define the power gain as a function of operating mode, such as "Class A power gain." We have also outlined that there is a difference between noise matching and gain matching. If we can be allowed to step back into the vacuum-tube era for a moment, the reader should be reminded that the difference between power gain and noise matching had to do with the grid-plate feedback capacitance which, via the Miller effect, showed up in parallel with the input. This actually detunes the input slightly off-center relative to the best gain. The reason for this is that as the feedback increases, Y_{12} or S_{12} starts playing a role, and the frequently assumed unilateral case no longer exists. The difference between noise matching and gain matching increases as a function of frequency. It can be brought closer together if appropriate reactive feedback is used. The two choices are: (1) increase the emitter or source inductance or (2) slightly increase the base-to-collector/gate-to-drain capacitance. Needless to say, the penalty for this is potential instability, and a sweep must be done from a few megahertz to several gigahertz to make sure the stage doesn't "take off."

The most efficiency this circuit provides is 50%; even this is a theoretical limit if one tries to keep distortion under control.

While the equation $P_{\text{OUT}} = (V_{\text{bat}} - V_{\text{sat}})^2/2R_L$ is valid, it really applies only to power amplifiers in the sense that for Class A, high-gain operation R_L can be made significantly higher as the absolute power output is not the important item. While the supply voltage V_{bat} is always known, the saturation voltage V_{sat} has to be obtained by the load line dc analysis. Note: The RF saturation voltage is always higher than the dc one. When looking at Figure 3-63, it becomes obvious that the conducting angle is 180°, which implies that the entire sine wave at the input is reproduced as an essentially undistorted sine wave at the output.

We can now calculate the components of the distorted current in separating them into the dc component, the fundamental frequency, and its first and second harmonics:

$$\text{Fundamental} \quad i_{c1} = i_{\text{peak}} \cdot f_1(\Theta) \tag{3-160}$$

$$\text{Second harmonic} \quad i_{c2} = i_{\text{peak}} \cdot f_2(\Theta) \tag{3-161}$$

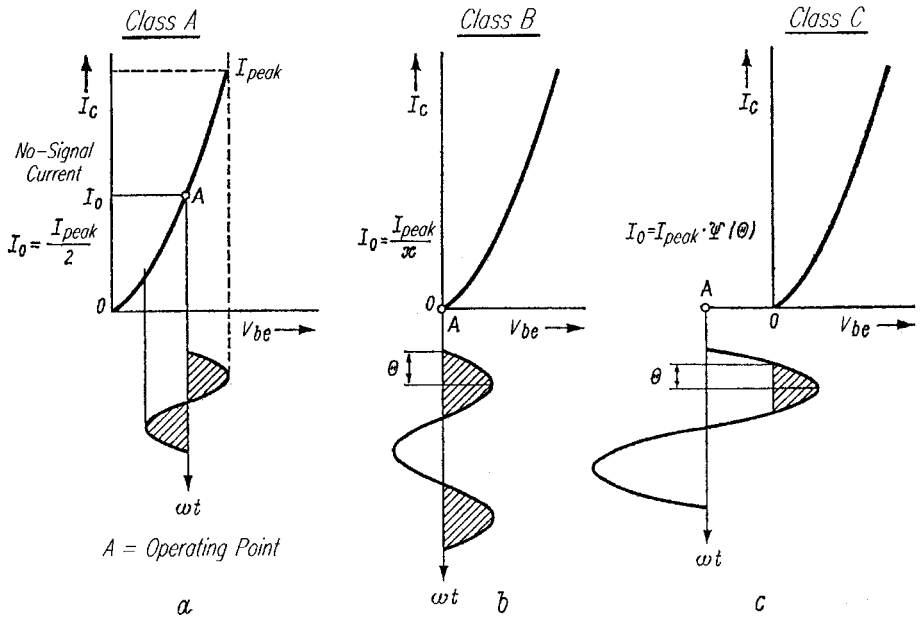


Figure 3-63 Definitions of Class A, B, and C operation, including conduction angle (Θ). Figure 3-64 shows the determination of Θ.

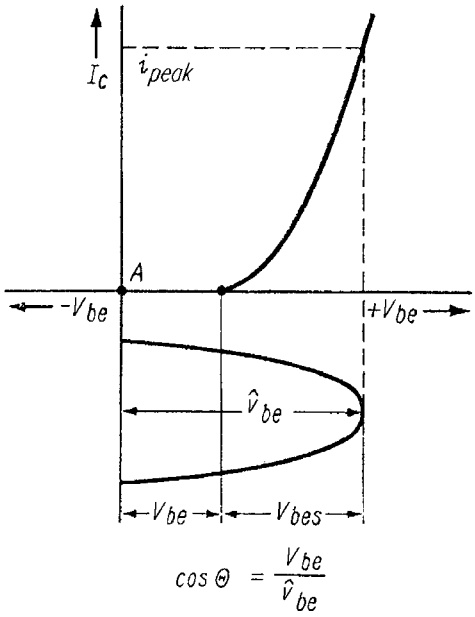


Figure 3-64 Determination of the conduction angle Θ from the base threshold voltage V_{be} and the base RF voltage $\hat{v}_{be}$.

$$\text{Third harmonic} \quad i_{c3} = i_{\text{peak}} \cdot f_3(\Theta) \tag{3-162}$$

For the collector dc current we obtain

$$I_c = i_{\text{peak}} \cdot \Psi(\Theta) \tag{3-163}$$

The frequency-dependent currents are a function of the conduction angle Θ. Figure 3-65 shows these transfer characteristics as a function of the conduction angle for a square-law dynamic transfer function; for higher-order transfer functions, the values don't change much. Tables for this can be found in various sources, including Clark and Hess [47]. Since the transistors exhibit a mixture of an exponential and a linear transfer characteristic (due to the unwanted parasitic emitter/source and collector/drain resistances when building the device), the plot in figure 3-65 is sufficiently valid for practical purposes. This function plot allows us to determine the amplitude of the current at various frequencies once the conducting angle has been set. For example, for Θ = 70°, the maximum collector current is $i_{\text{peak}} = I_c/0.253$, meaning that the peak current of the transistor is approximately four times larger than the dc value one would measure. Once the peak collector current is known, the fundamental and harmonic currents can be calculated:

$$i_{c1} = \frac{I_c}{0.253} f_1(\Theta)$$

$$i_{c1} = \frac{I_c}{0.253} 0.436$$

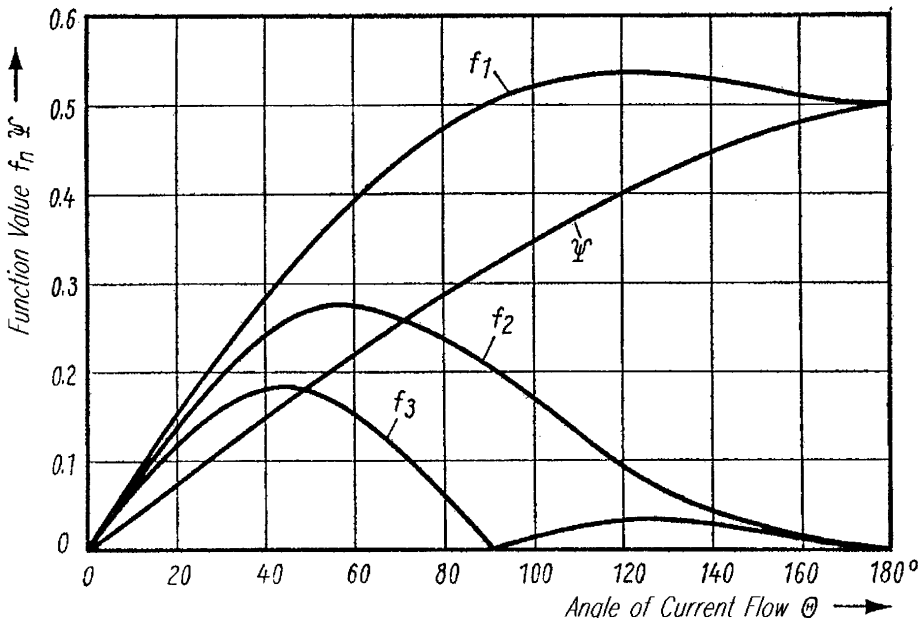


Figure 3-65 Function values for the fundamental, second harmonic, third harmonic, and dc components (f_1, f_2, f_3, and Ψ, respectively) versus device conduction angle.

In order to obtain i_{c2} and i_{c3} one needs to do the same calculations as above with the appropriate functions for f_2 and f_3.

It becomes obvious that the distorted collector current generates a lot of harmonics, which definitely requires sufficient low-pass filtering at the output, even for Class A operation. We already stated that for Class A operation the conduction angle is 180° while it is 90° for Class B operation. The conduction angle for Class C operation can be solved from the ratio of the knee voltage V_{be} and the peak value of the ac voltage at the base. We find

$$\cos \Theta = \frac{V_{be}}{\hat{V}_{be}} \tag{3-164}$$

where $\Theta = 90°$ and $V_{be} = 0$. Given the voltage V_{bes}, which determines the peak collector current (up to i_c peak), we can calculate the effective RF voltage at the base as

$$\hat{v}_{be} = \frac{v_{bes}}{1 - \cos \Theta} \tag{3-165}$$

Therefore, the peak value of v_b becomes v_{bes} for $\Theta = 90°$.

The necessary base voltage V_{be} for maximum current is obtained by the transistor model. Then one finds the threshold voltage at which the transistor becomes conductive and by using Eqs. (3-164) and (3-165) it can be checked whether this results in the appropriate conducting angle. An example for this will be shown in the high-power stage section.

Design Guidelines. The design of a one-stage amplifier consists of finding

- An input lossless matching network M_1
- An input lossless matching network M_2

so that the maximum or desired transistor gain is achieved over the operating bandwidth of the amplifier. Usually, the common-emitter or common-source configuration is chosen for highest gain per stage. If the stability factor K is greater than unity, these two networks can be found to give the maximum available gain $G_{\max}$. If the stability factor is less than or equal to unity, the amplifier could be terminated in a matching structure that causes oscillation; that is, $G_{\max}$ is infinite. This should be avoided by locating the regions of instability in the Γ_G and Γ_L planes. The input and output terminations (Γ_G and Γ_L) must be designed to avoid the instability regions. Usually, these unstable regions are near the conjugate match for S_{11} and S_{22}. Thus, a stable amplifier will require some input and/or output mismatch if K is less than or equal to unity.

There are at least two alternative approaches for potentially unstable amplifiers:

1. Add resistive matching elements to make $K \geq 1$ and $G_{\max} \approx G_{ms}$.
2. Add feedback to make $K \geq 1$ and $G_{\max \approx Gms}$.

For narrowband amplifiers, it is usually recommended that one accept a transistor with $K < 1$, design the amplifier for a gain approaching G_{ms}, and ensure that the Γ_G and Γ_L terminations provide stability at all frequencies, both inside and outside the amplifier passband.

The design of an amplifier would usually have the following specifications:

Gain and gain flatness
Bandwidth and center frequency ($f_2 - f_1, f_0$)
Noise figure
Linear output power
Input reflection coefficient ($VSWR_{in}$)
Output reflection coefficient ($VSWR_{out}$)
Bias voltage and current

For small-signal amplifiers, the small-signal S parameters are sufficient to complete the design. After selecting an appropriate transistor based on these specifications, a one-stage amplifier design should be considered if sufficient gain can be achieved; otherwise, a two-stage amplifier should be designed.

The circuit topology should be chosen to allow dc bias for the transistor. Usually, RF short-circuited stubs are placed near the transistor to allow dc biasing. If the topology does not allow dc bias, a broadband bias choke or high-resistance bias circuit that does not affect the amplifier performance must be used.

The following steps can be tabulated for the design of a one-stage amplifier (see Figure 3.66):

1. Select a transistor based on a datasheet description of the S parameters, noise figure, and linear output power.
2. Calculate K and G_{max} or G_{ms} versus frequency.
3. For $K > 1$, select the topologies that match the input and output (and allow dc biasing) at the upper band edge f_2. Ideally, this will give G_{max} and $S'_{11} = S'_{22} = 0$. Usually, $S_{12} = 0$ is assumed for the initial design. Next, the topology may be varied to flatten gain versus frequency at the expense of S'_{11} and S'_{22}.
4. For $K < 1$, plot the regions of instability on the Γ_G and Γ_L planes and select topologies that partially match the input and output at the upper band edge and avoid the unstable regions. The gain will approach G_{ms} as an upper limit. Next, the topology may be varied to flatten gain versus frequency.
5. After finding initial M_1 and M_2, plot the amplifier S parameters versus frequency; make adjustments in topology until the specifications for gain, input reflection coefficient, and output reflection are satisfied. Also plot Γ_G and Γ_L versus frequency to verify amplifier stability.
6. Design the dc bias circuit. Lay out the elements of the complete amplifier and check realizability.

As a beginning point in an amplifier design, the circuit topologies may be designed with the assumption that S_{12} is zero. Later, an exact design procedure will be described that includes the topologies with terminations in the stable region.

Design Examples. Throughout this book we always address the issue that we have to separate the battery-operated handheld applications and the more stationary designs like those for base stations. In the case of the handheld or low-voltage applications, the dynamic range requirements are significantly less than the front-end of a base station requires.

NE68133 Matched Amplifier. We will first consider the design of a low-voltage, low-current, low-noise amplifier based on the NEC BJT NE68133. This device was chosen

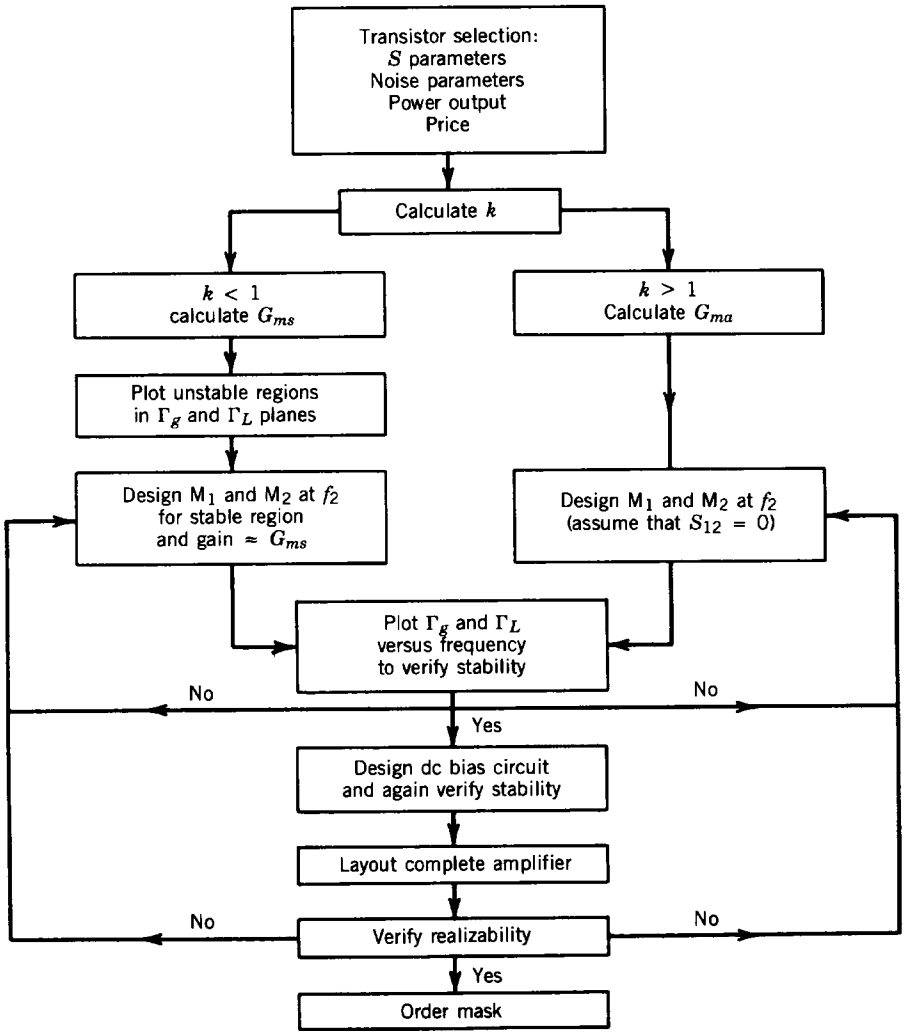


Figure 3-66 Simplified amplifier design procedure.

because it is unconditionally stable, has a minimum noise figure of about 1.2 dB at 7 mA, has 15.5 dB gain, and has an f_T of 8 GHz. While it is possible to start the design based on the published S and noise parameters, we will choose also to involve the nonlinear model, as this allows us more freedom in selecting the bias point including the collector voltage. The actual topology of the amplifier is shown in Figure 3-67.

Determining appropriate values for the input and output matching networks is essential to finalizing the design. The next several figures illustrate matching-network syntheses undertaken with the interactive Smith Tool in Ansoft's Serenade Design Environment. We evaluate the point of tangency as well as stability circles in the source and load planes, indicating that our impedance match will result in a stable amplifier.

In Figure 3-68, the center circle (with 1.1 marker) refers to the noise circle for which the noise figure of 1.1 dB can be obtained. The left circle (with the 14.3 marker) corresponds to

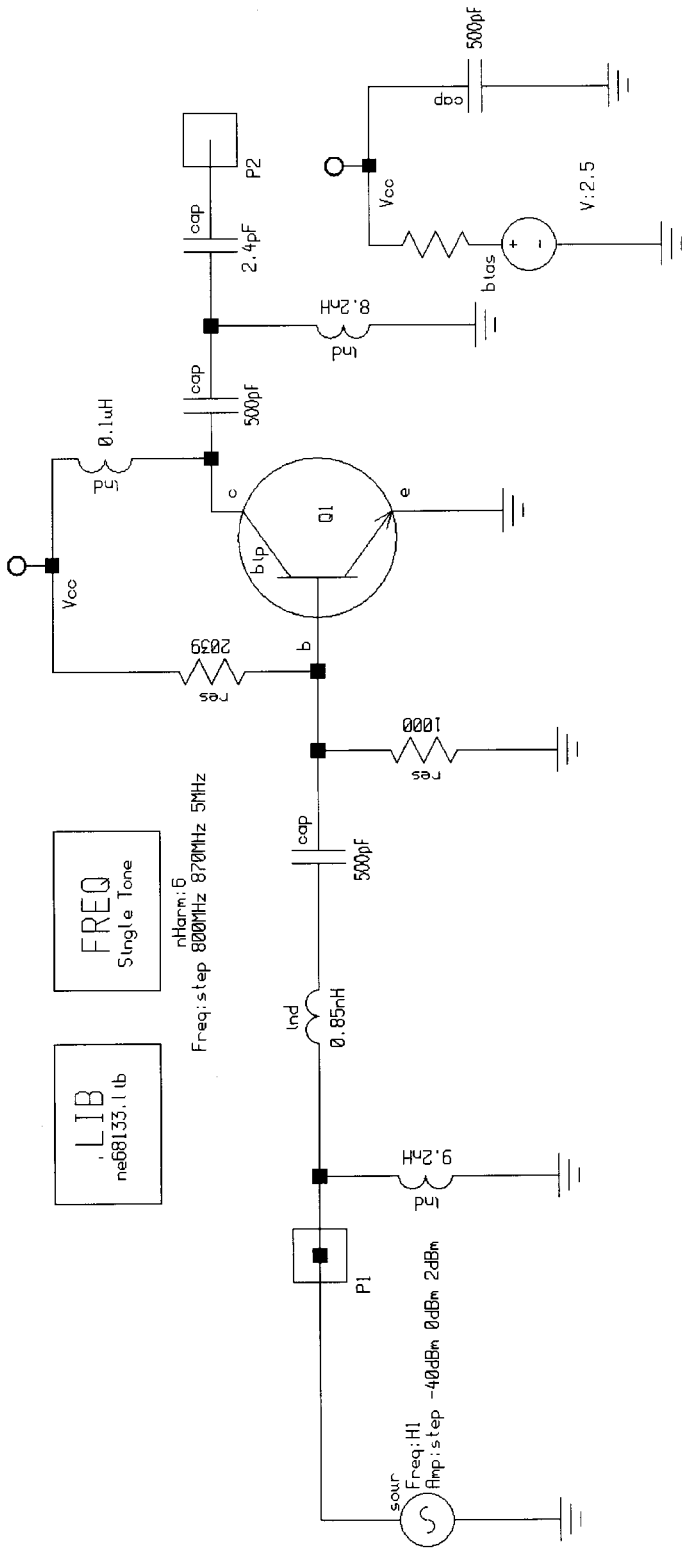


Figure 3-67 The NE68133 low-noise amplifier.

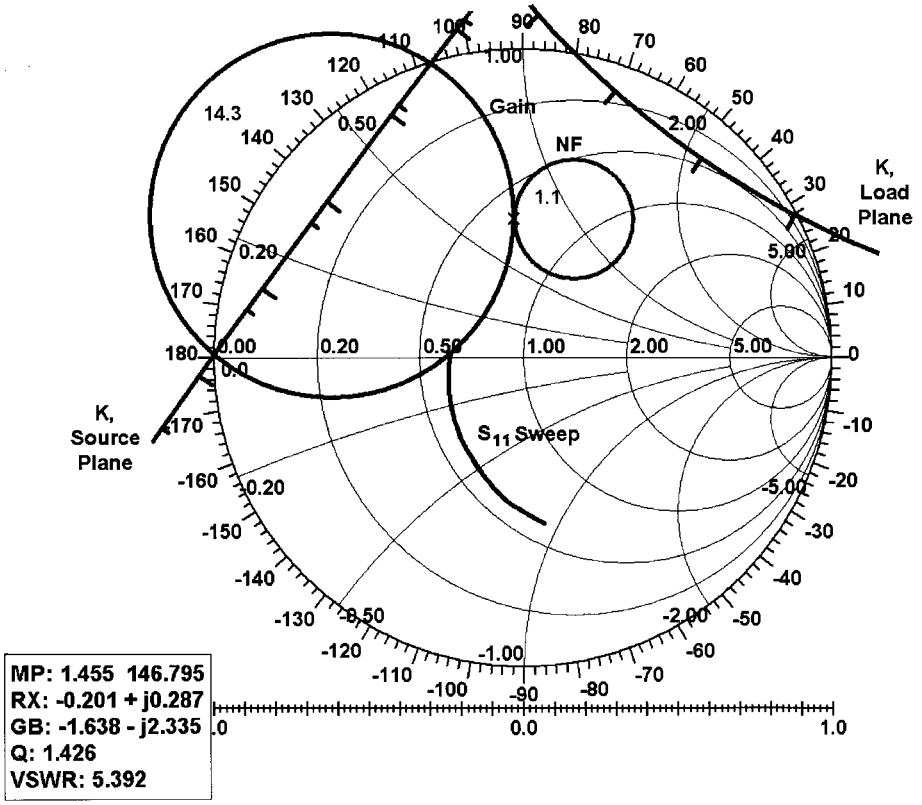


Figure 3-68 Display from the interactive Smith Tool, showing circles for gain, noise figure, and source- and load-plane stability, as well as a sweep of S_{11} versus frequency. In synthesizing an input matching circuit, we choose a single working frequency and work from the chart's origin to the point of tangency between the gain and noise figure circles. See Figure 3-69.

source-plane terminations that correspond to 14.3 dB of available gain. In order to get the required noise figure *and* gain, the input termination must be at the point of tangency of both circles. The arc touching the X axis close to 0.50 is the amplifier's S_{11} response from 800 to 875 MHz. The shallow arc that nearly bisects the 14.3-dB gain circle is actually a portion of the source-plane stability circle; the arc at the top right is a portion of the load-plane stability circle. Because we will be terminating the amplifier input with a value inside the source-plane stability circle, and terminating the amplifier output with a value outside the load-plane stability circle, the amplifier will be unconditionally stable at the frequency at which the match is obtained.

The next task is to obtain the values for the network circuit that matches the input for the 1.1-dB noise figure. A noise matching will not result in the best possible S_{11} matching (Figure 3-69). This translates into an inductor of 9.2 nH to ground and a series inductor of 0.85 nH based on the large-signal model of the NE68133 at the chosen bias point.

After we have determined our input matching network, it is time to find the output matching (Figure 3-70). First, find the termination point on the output plane that corresponds to 14.3-dB gain and 1.1-dB NF at the input plane. The output gain circle goes through the

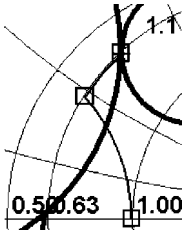


Figure 3-69 Input matching-network extraction, assuming lossless matching components and no parasitic effects.

1.1-dB NF circle at two points, and for reasons of easy matching we selected the lower point. The next step is to find the conjugate reflection coefficient $\Gamma_L = S_{22}^*$. This is the lower square marker in Figure 3-70. The remaining task is to find the matching network that brings this value to 50 Ω. By doing so, we first obtain a shunt inductance $L_{\text{shunt}} = 8.2$ nH and a series capacitor $C_{\text{series}} = 2.4$ pF. Again, these are the values applicable to the large-signal equivalent circuit.

The careful reader will have noticed that we have been avoiding the issue of small-signal exact matching using the manufacturer-supplied parameters. If this is done, the narrow band response shown in Figure 3-71 will be obtained. Therefore, we must optimize the circuit for a better frequency response to obtain the one shown in Figure 3-72.

After this optimization, we were well able to meet the specification of a center band gain of 14 dB and a corresponding NF of 1.1 dB. To show the flatness of the gain, Figure 3-73 shows a gain variation in the range of 0.25 dB.

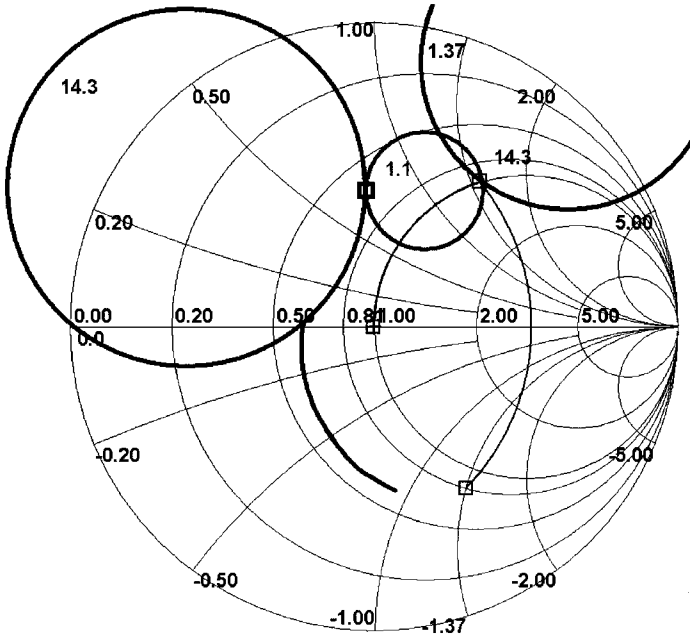


Figure 3-70 Output matching-network extraction.

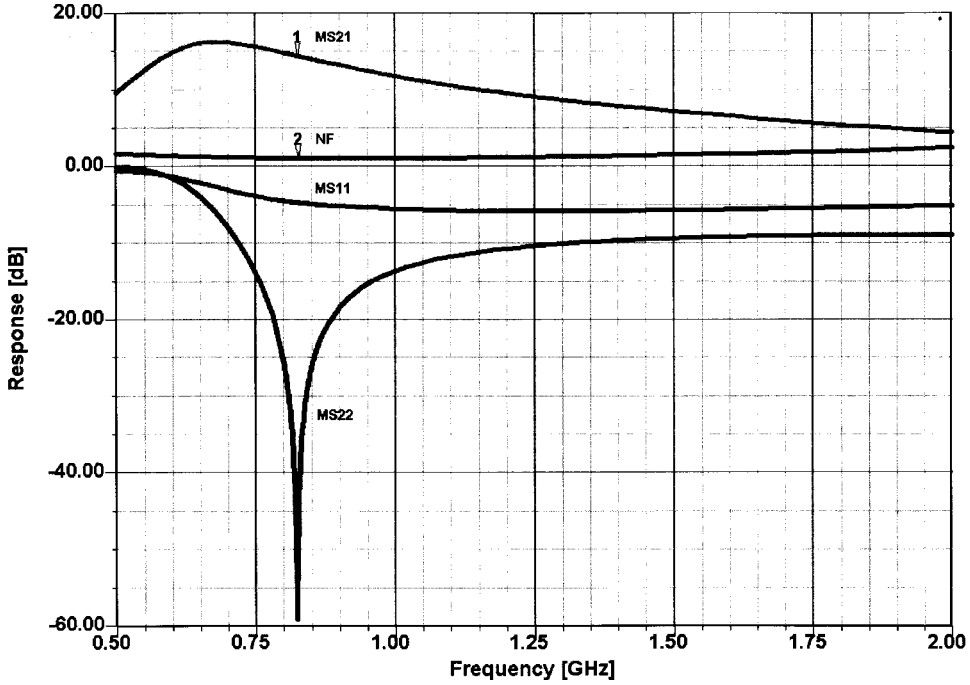


Figure 3-71 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-67.

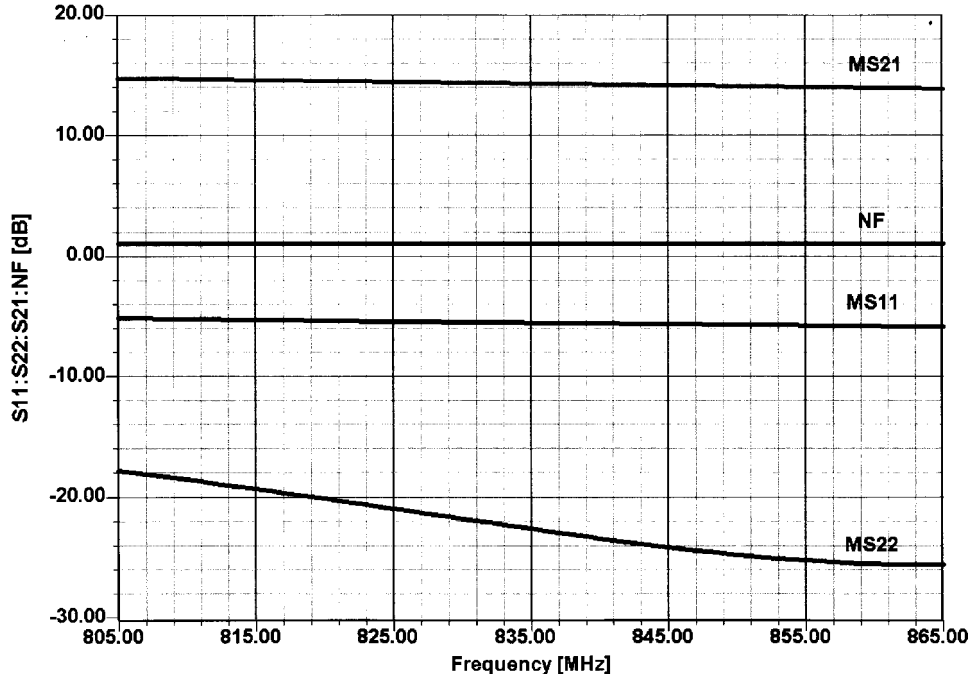


Figure 3-72 Optimized performance of input/output reflection, gain, and noise figure.

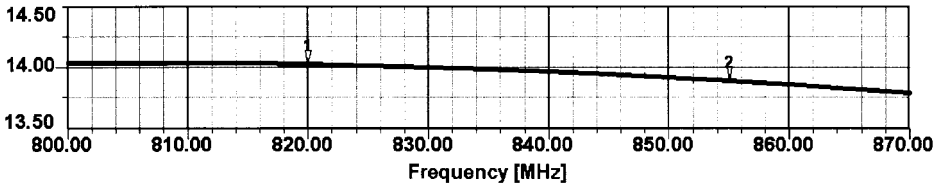


Figure 3-73 High-resolution display of gain as a function of frequency.

The next step is to look into the large-signal performance. The first order of business is to determine the third-order intercept point, which is based on a two-tone analysis. Given the fact that we run the device at low voltage and at a just-reasonable current, the actual intercept point of +13.5 dBm is quite a good number. Figure 3-74 shows the fundamental and third-order outputs. The crossover point between the two, as shown in Chapter 1, is defined as the third-order intercept point. If we had chosen the third harmonic output rather than the fundamental, we would have shown the fifth-order intercept point. The second-order IMD product ($f_1 \pm f_2$) typically can be minimized by input selectivity.

How an amplifier's linearity affects its ability to preserve the characteristics of digital signals is a major concern in wireless design. Figures 3-75 and 3-76 show constellation diagrams with the amplifier handling a PSK signal at drive levels of –40 dBm and +10 dBm, respectively. As expected, we see significant distortion at the higher drive levels due to circuit nonlinearities.

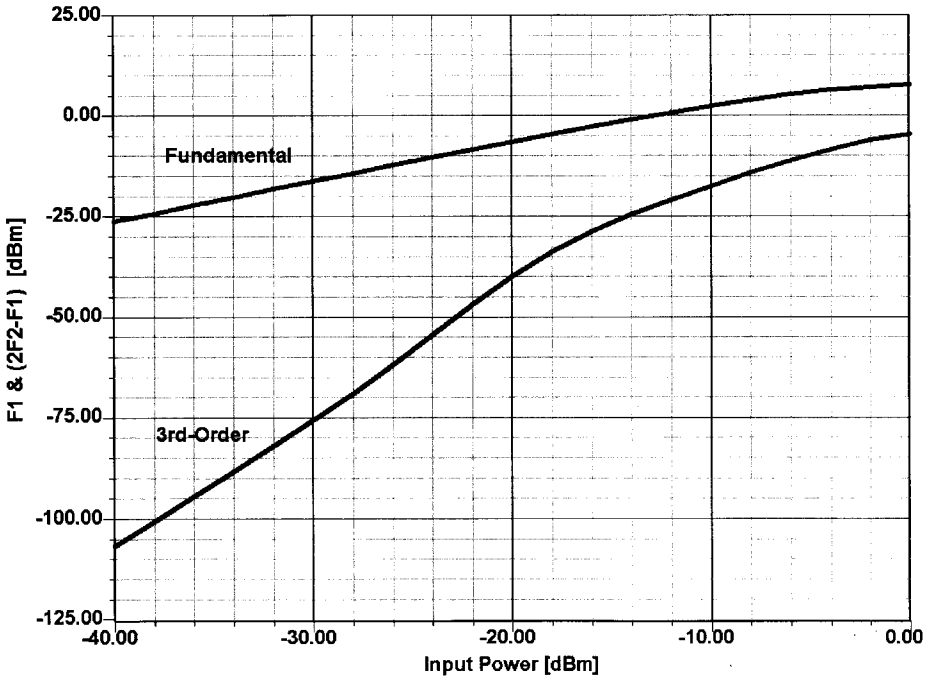


Figure 3-74 Simulated fundamental and third-order IMD outputs of the NE68133 amplifier.

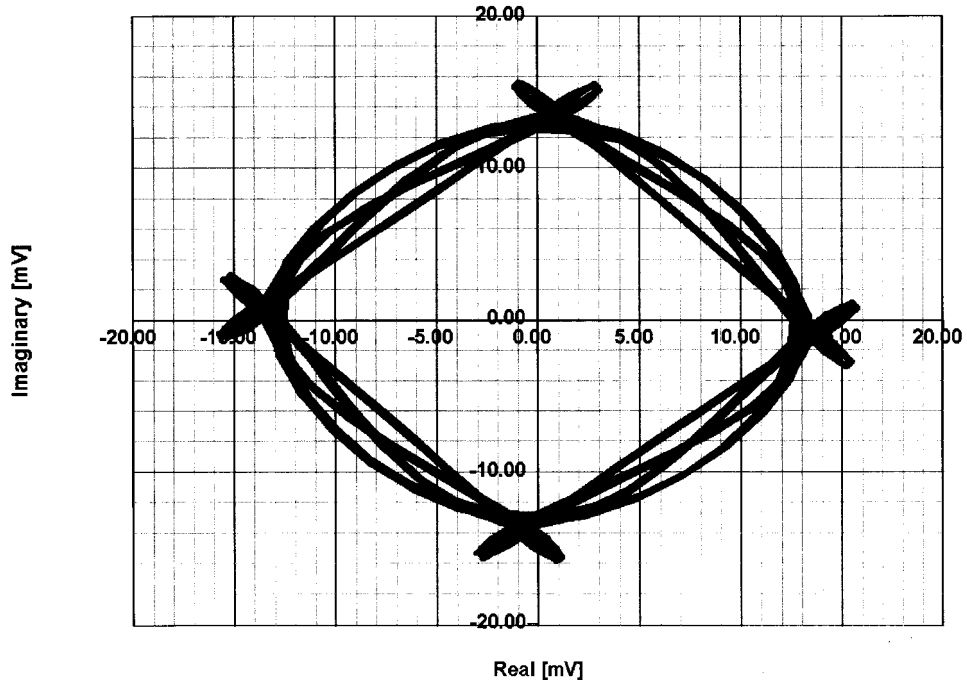


Figure 3-75 Output constellation at –40-dBm input.

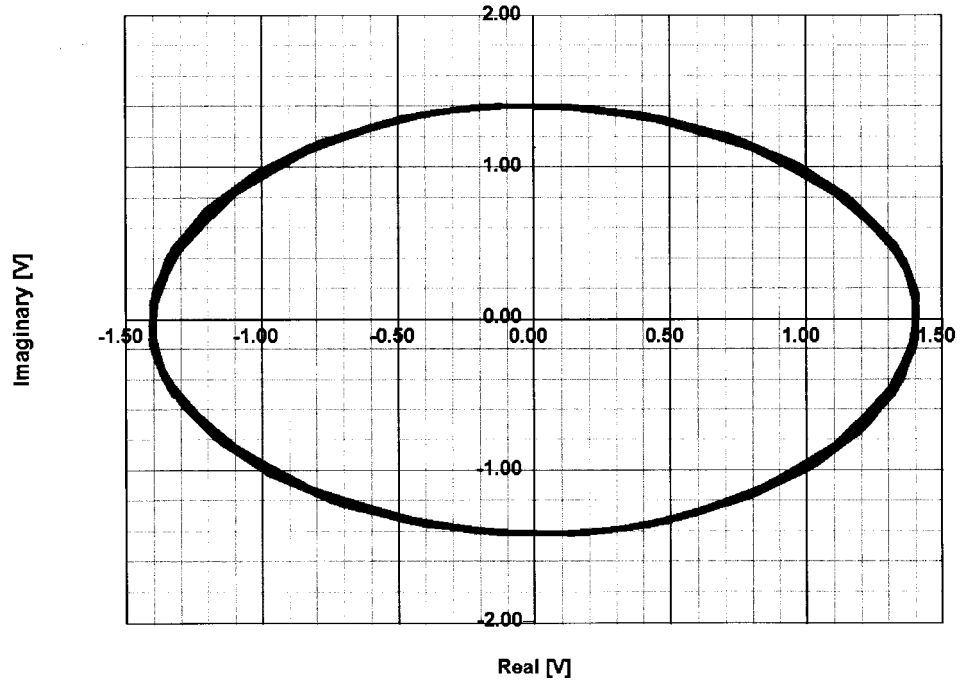


Figure 3-76 Output constellation at 10-dBm input.

BFP420 Matched Amplifier. The NE68133 has an attractive f_T of 8 GHz, but if more gain is needed, one may play with the concept of using a much "hotter" device, such as the BFP420. Recall that the NEC device was unconditionally stable, and our first worry is how the Siemens BFP420 will behave.

A study of Figure 3-77 shows that indeed the NEC device had to be stable all over and therefore was a good basis for a risk-free design. Things change dramatically with the BFP420, for which the K factor is essentially less than 1 up to a corner frequency of 2.75 GHz, where the transistor becomes unconditionally stable. Around 1.75 GHz, the K factor for the NEC transistor barely touches 1. We explained earlier that $G_{\max}$ deviating from the maximum stable gain is a more realistic number, since the conditions for maximum stable gain involve feedback and conjugate matching.

To improve the stability of the BFP420, we will now set out to use the same topology as the previous amplifier but with a resistive feedback between base and collector (Figure 3-78). A good value for this is always somewhere around 200–500 Ω. It must be remembered that this voltage feedback also stabilizes the input and output impedance and reduces the Miller effect because of the feedback being resistive and more dominant.

While the first example covered a fairly narrow frequency band, we are now attempting to cover a much wider frequency range; specifically, from 1.5 to 3 GHz reasonable gain and noise. This, by definition, requires a much higher gain-bandwidth product, and the simple matching network at the input and output provides a nonoptimal input and output matching (Figure 3-79). The feedback introduces some additional noise. This application shows the limits of single-stage UHF/SHF amplifiers with feedback.

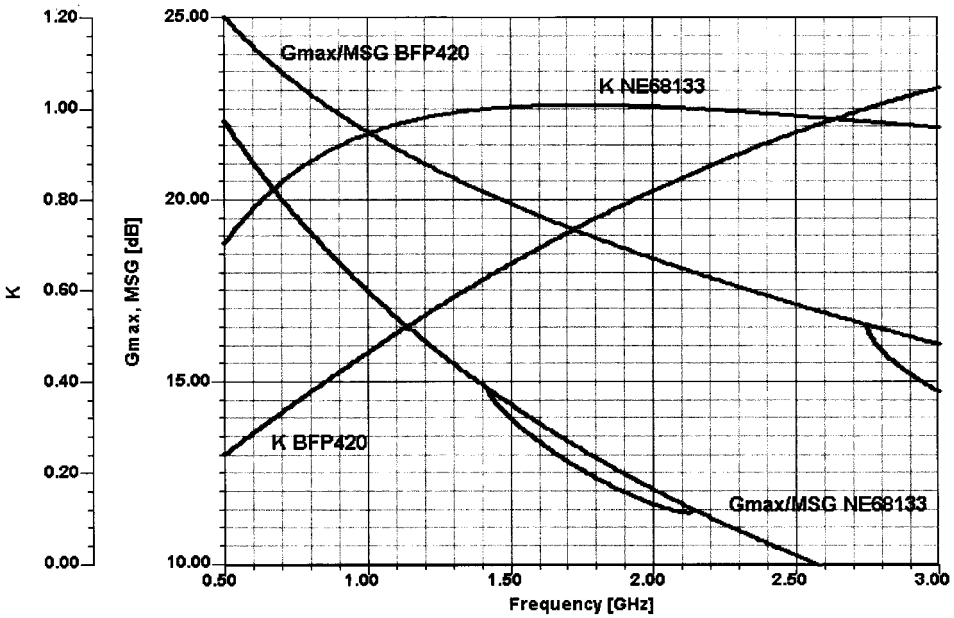


Figure 3-77 Comparison of K, $G_{\max}$, and maximum stable gain (MSG) for the NE68133 and BPF420 BJTs.

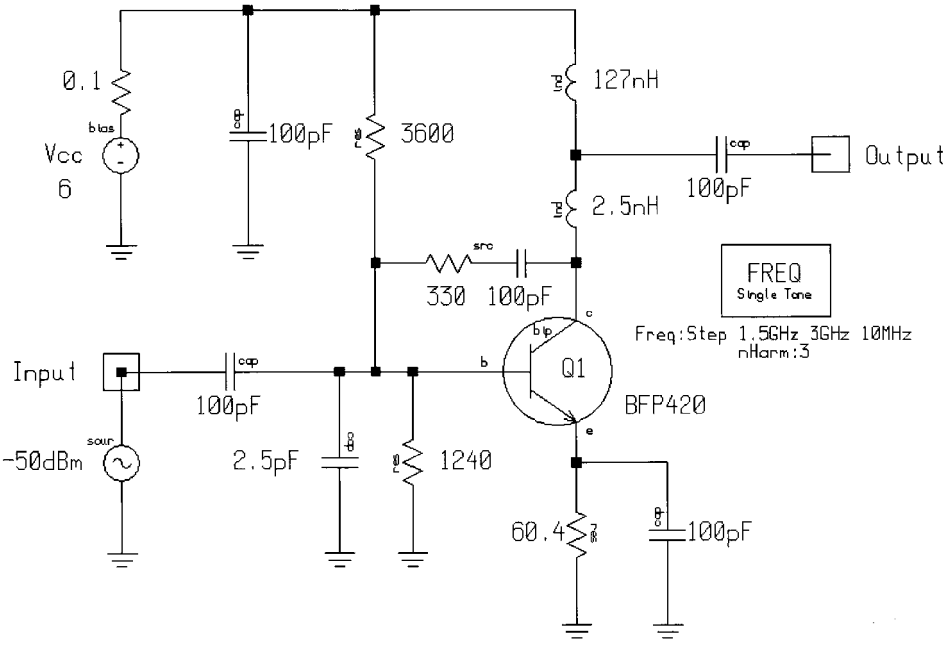


Figure 3-78 BFP420 amplifier with resistive voltage feedback from collector to base.

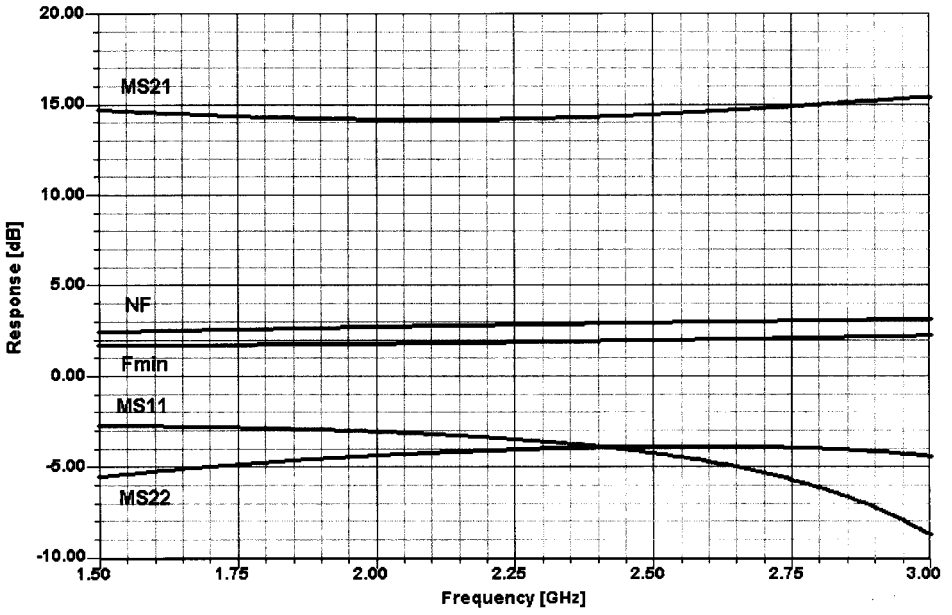


Figure 3-79 Frequency-dependent gain, matching, and noise performance of the BFP420 feedback amplifier.

Narrowband BFP420 Amplifier. Our next attempt will be to design a narrowband input stage that combines very high selectivity with good noise figure. This can only be achieved by using two tuned single-stage resonators with capacitive/magnetic coupling. The following amplifier is tailored to have a center frequency of 900 MHz. The transistor will operate at 12 V/10 mA to have a sufficient dynamic range. The input selectivity is achieved by the use of two top-coupled tuned circuits, followed by a BFP420 (Figure 3-80). The first order of business is to determine the input tuned circuit. The tuning capacitance for the two tuned circuits is arbitrarily set to a manageable 2 pF (small values have too much tolerances and larger values will require smaller inductances, resulting in lower Q). The reactance of the first tuned circuit is

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi \times 9\text{E}8 \times 2\text{E}{-}12} = 88.42\ \Omega$$

The necessary inductance for resonance will be

$$\omega L = 88.42$$

or

$$L = 88.42/\omega = 88.42/(2\pi \times 9\text{E}8) = 15.64\ \text{nH}$$

We will make both tuned circuits identical, but with different loaded Qs. By multiplying 88.42 times the loaded Q (20 for the input resonator and 50 for the output resonator), the resulting parallel-resonant resistance is $88.42 \times 20 = 1768\ \Omega$ at 900 MHz (for the input resonator) and $88.42 \times 50 = 4421\ \Omega$ (for the output resonator). These values are achieved by loading the input with the 50-Ω source and by loading the output with the input of the transistor. Therefore, the input transformation ratio $m_1 = 1768.4/50 = 35.37$. At the output, the equivalent number $m_2 = 4421/50 = 88.42$. For the purpose of simulation, and because it is extremely difficult to find the physical location of the appropriate tap on an air-wound coil for the necessary impedance, we do this "on paper" by incorporating an ideal transformer. We are also going to show later the actual transmission-line-based filter and its modeling. This filter will have a higher insertion loss based on the PC-board material used.

Using our nonlinear simulator, in small-signal ac mode, we obtain the frequency response and noise figure shown in Figure 3-81. In actually building this circuit, there is another problem besides these transformers—the top-coupling capacitor. We have used a grounded-T configuration to obtain a practically realizable value of 36 pF, whose lead inductance will not influence the circuit. (The actual coupling capacitance value for top coupling without the T configuration would have 0.1 pF. We would like to see this value repeatably realized.)

If the actual noise would be calculated relative to the output tuned circuit, the equivalent noise resistor R_n of 5 Ω would now be multiplied by m_2, resulting in 442 Ω. Fortunately, the Serenade simulator has a highly accurate BJT noise model that is both bias and temperature dependent, so the simulator will give us the answer directly.

We have not put any matching circuit at the output and leave this task to the reader. In practice, the slight tilt in the MS_{21} curve in Figure 3-81 would be tuned out. Also note that the minimum noise figure ($F_{\min}$) and the actual 50-Ω noise figure (NF) agree somewhat off center. This has to do with the Miller effect, which results in the well-known difference between noise matching and input-gain matching. By varying the coupling capacitor to move from the overcoupled response, which gives a low noise figure, to a critically coupled

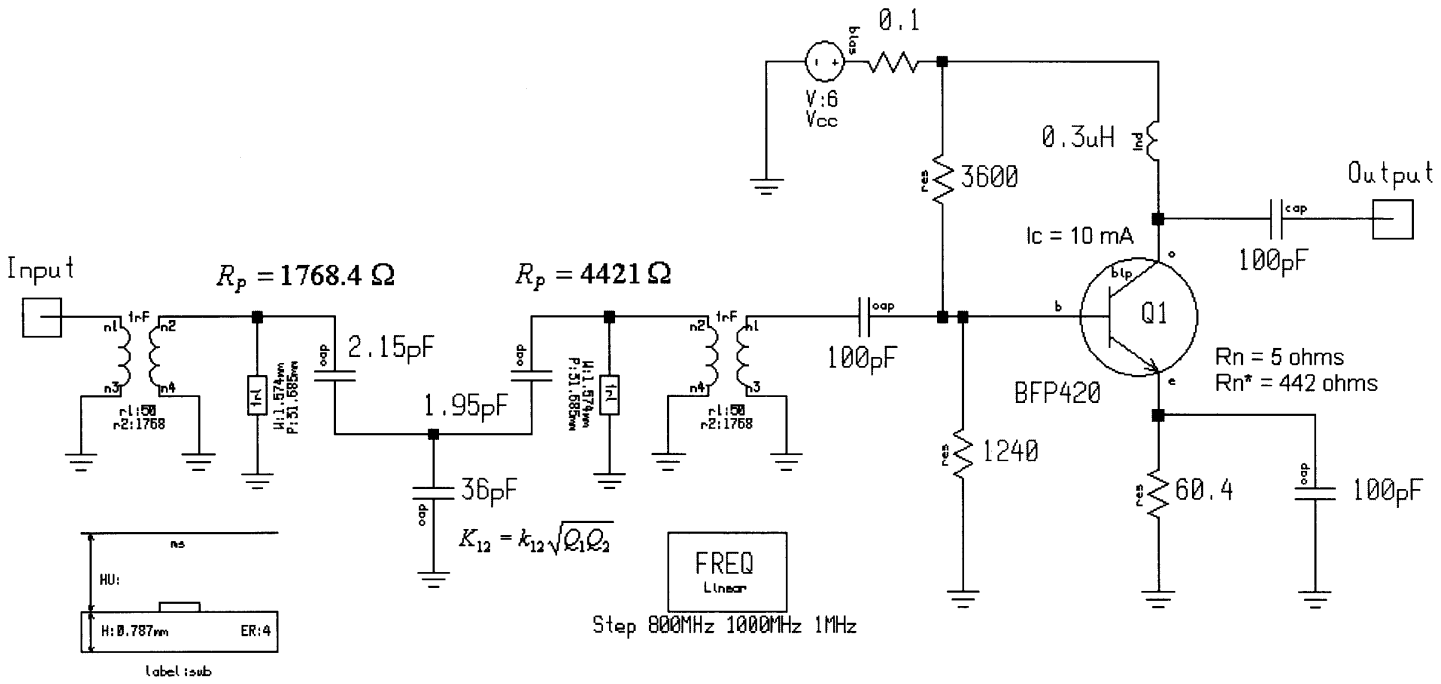


Figure 3-80 BFP420 amplifier with narrowband input filtering. Ideal transformers are used to match the filter's input and output to the input port and the BFP420 base, respectively.

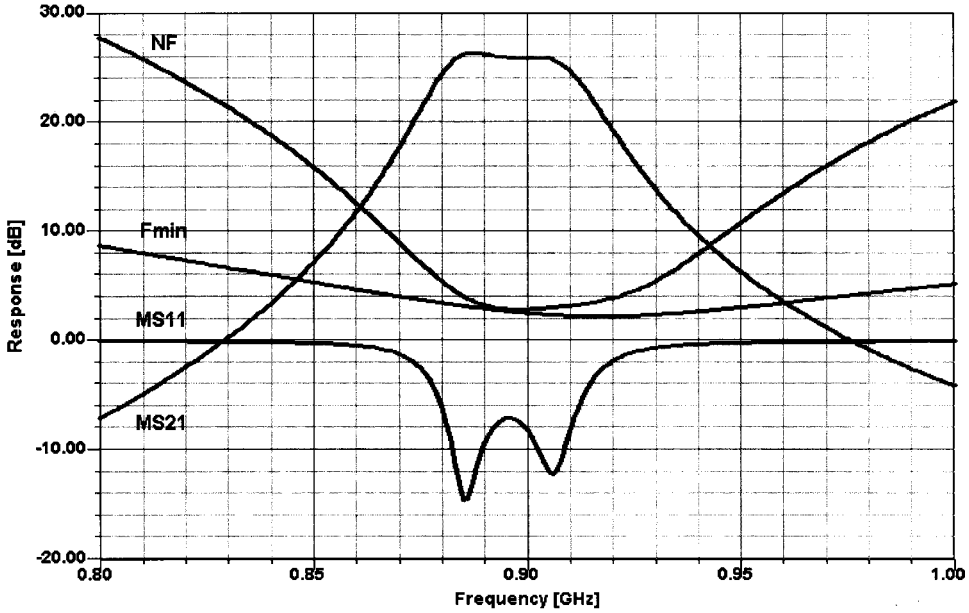


Figure 3-81 Frequency-dependent gain, matching, and noise performance of the narrowband BFP420 amplifier.

response ($K = 1$), we get significantly better selectivity and narrow bandwidth (at the obvious expense of the noise figure) as shown in Figure 3-82.

Since we somewhat compromised by not providing the reader with an actual input and output coupling that can be built, and furthermore, because the inductance in question (15.26 nH) is hard to realize with a high Q, we now show the printed circuit approach for the input filter. By using a Teflon-based material for the PC board, the losses would be less; however, the radiation would be more because ε would go down to 2.1 from the value we used (4). It is also convenient to introduce this filter to show that at higher frequencies, where lumped elements are really no longer available, one has to switch to distributed elements. A list of the commonly used distributed elements will be shown later; the transmission line and T junction are probably the most frequently used elements, followed by spiral and rectangular inductors, and bends.

The actual implementation of the input filter on a printed circuit board requires distributed elements (Figure 3-83). In practice, one has two parallel coupled lines with a tap somewhere toward the ground connection and, for modeling accuracy, we have to introduce a T junction with the appropriate impedance of 50 Ω for the input and a different impedance (if necessary) for the output. The coupling mechanism now is magnetic coupling, since no one can physically change the distance between the transmission lines by tuning some mechanical elements, one has to get it right the first time. This can only be achieved with a high-precision simulator. The reason why these simulators are so expensive is the investment of research time and personnel to produce valid models over a huge frequency and impedance range. Since we are now curious about the frequency response, we have plotted in Figure 3-84 the frequency response, which shows a higher (expected) insertion loss. The PC-board material is partially responsible for this.

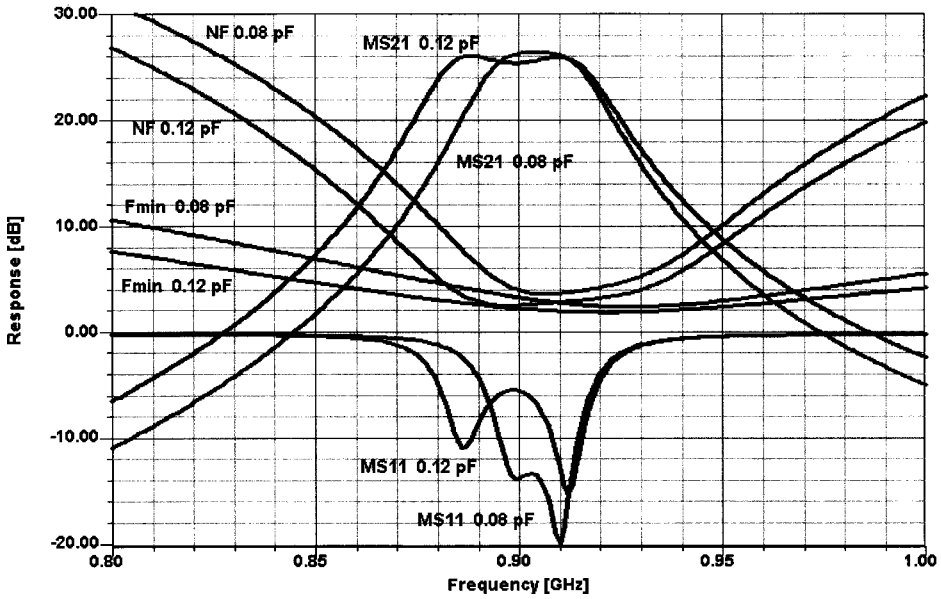


Figure 3-82 Comparison of filter responses with critical coupling (0.08-pF response) and overcoupling (0.12-pF response).

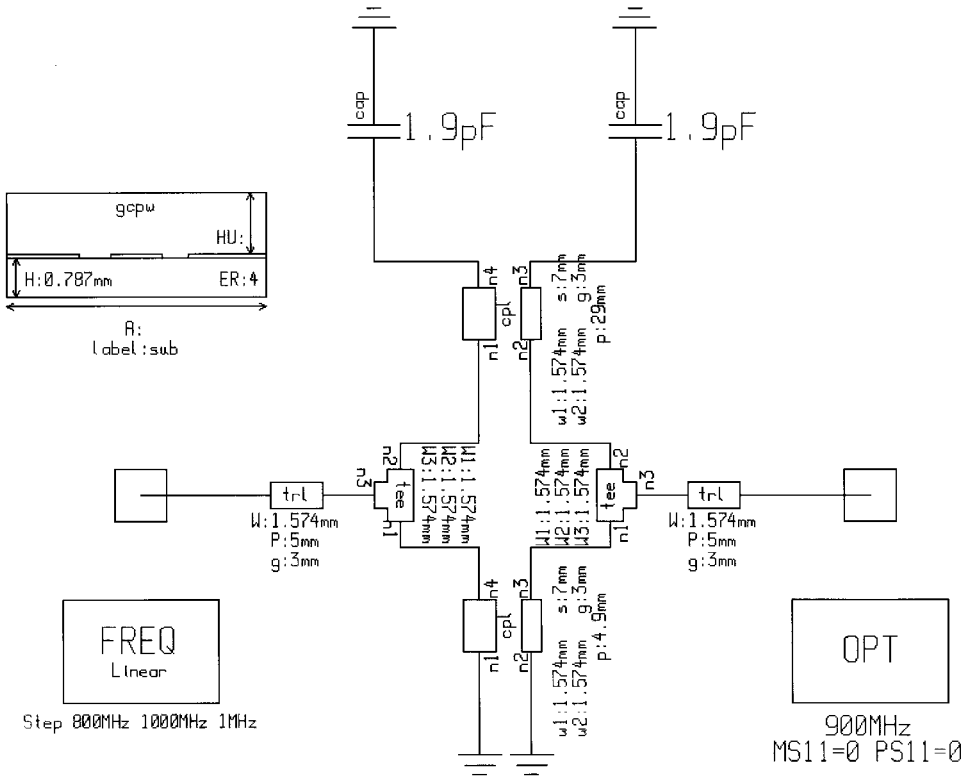


Figure 3-83 Input filter using distributed-element resonators and matching.

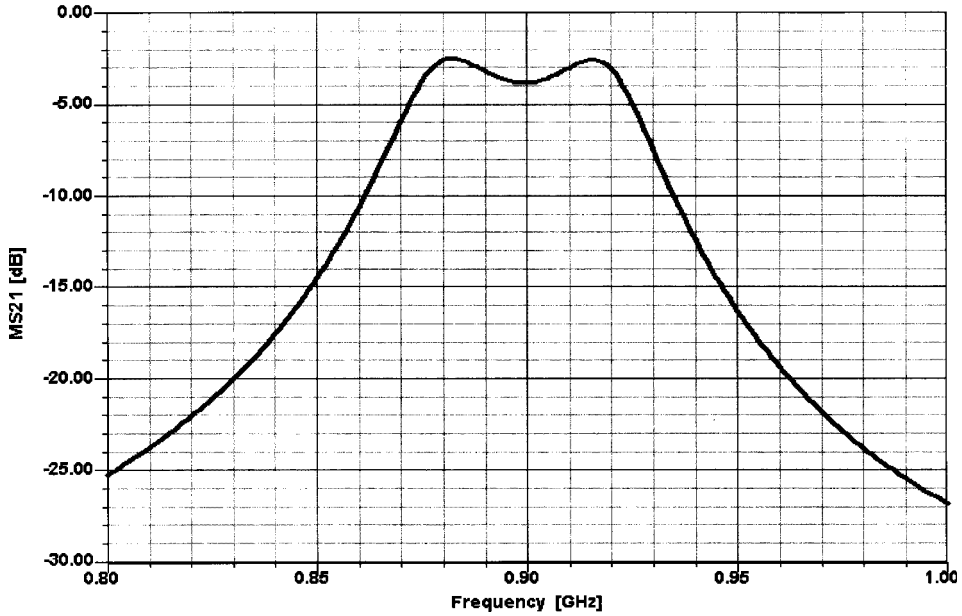


Figure 3-84 Gain versus frequency response for the distributed-element filter.

At the end of this chapter we will have a chance to discuss passive elements as needed for microwave frequencies, and one should consider frequencies above 1500 MHz as microwave frequencies.

GaAsFET Feedback Amplifier. The implementations we have shown so far have been based on bipolar transistors. Figure 3-85 shows a single-stage feedback amplifier using a GaAsFET. The feedback is based around an *LRC* network and some attempts at matching at the input and output are made. Figure 3-86 shows its frequency-dependent gain, matching, and noise performance.

3-2-3 High-Gain Amplifiers

In most cases, high-gain amplifiers receive an already amplified signal, which means we are less concerned about the signal-to-noise ratio but remain concerned about distortion. To reduce distortion, we have two options: (1) increase the dc current in the active device, which always improves the current-related distortion; and (2) add some voltage feedback, which reduces the voltage distortion. In simple terms again, only current feedback can reduce the current-related distortion of a square-law or exponential transfer characteristic and only voltage feedback can reduce the effect of overdrive as a function of the load, but only to some degree. Its main purpose remains to reduce the input impedance to useful values, specifically for FETs, since current feedback increases the input and output impedances. Some ultrasensitive circuits still require Class A operation, but now at higher current levels; these are typically CATV applications, and there are several transistors available that can be operated at 5–15 V but at 60 mA or higher. Because of the higher power consumption, this excludes handheld cordless or cellular telephones. The same circuit as used before (Figure 3-85) for

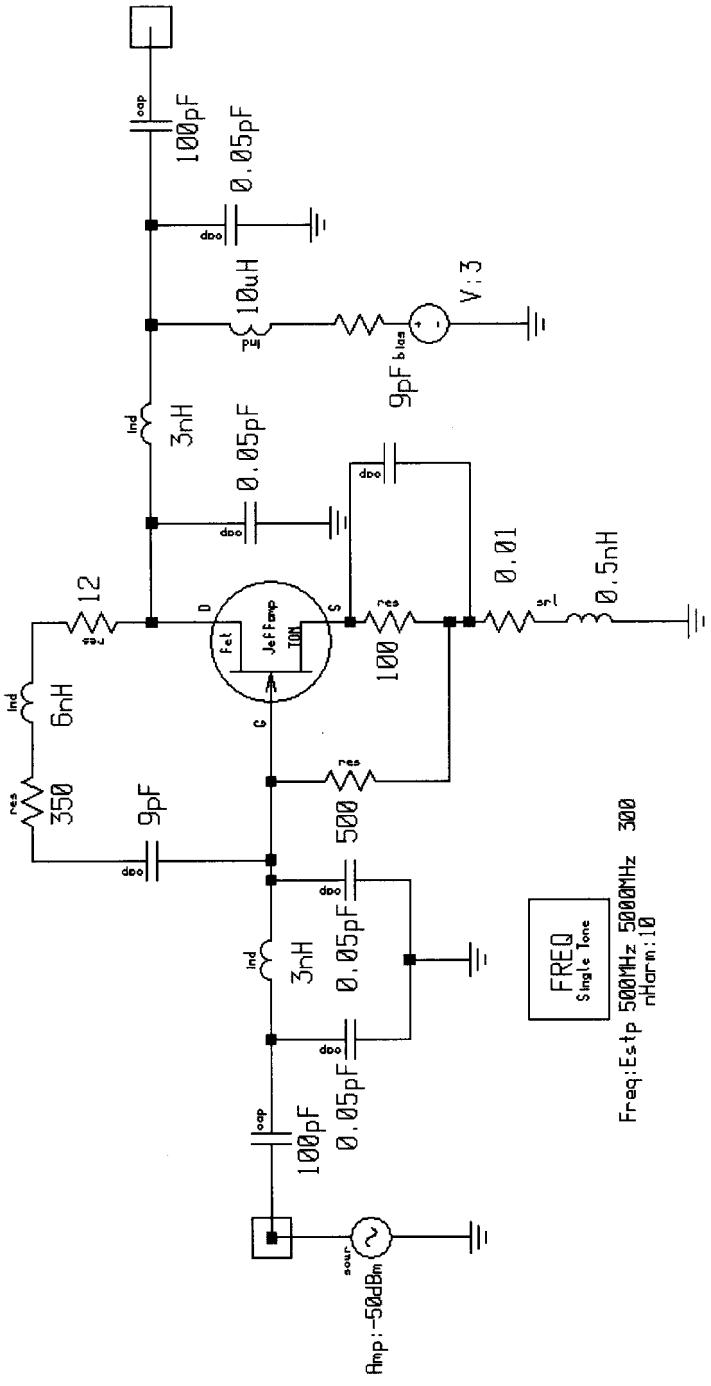


Figure 3-85 GaAsFET feedback amplifier.

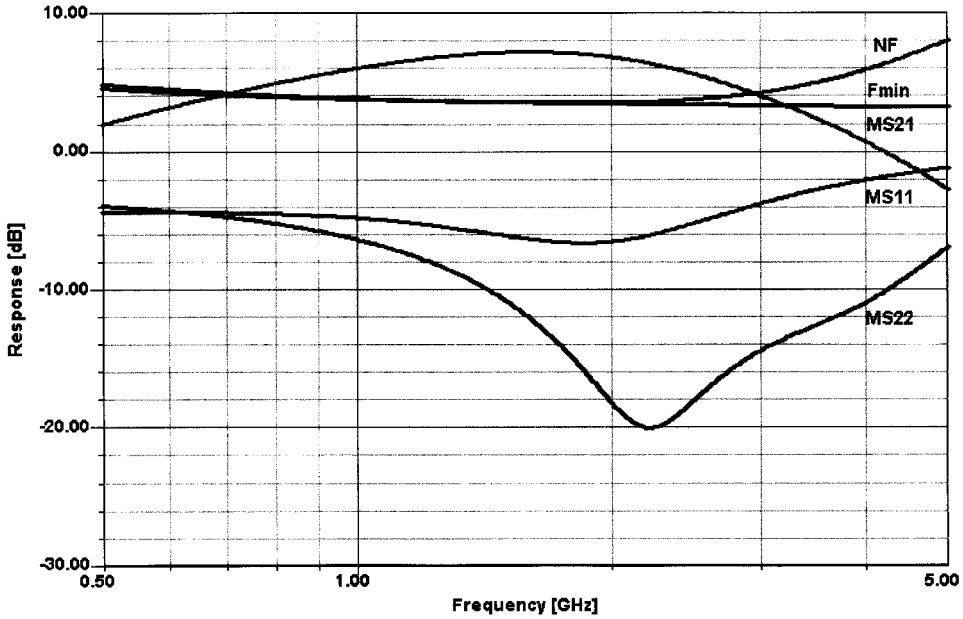


Figure 3-86 Frequency-dependent gain, matching, and noise responses for the GaAsFET amplifier.

Class A operation now can be operated at a higher current, which automatically increases the gain and the noise figure and reduces distortion. This is based on the fact that the transconductance (g_m) increases at higher currents. The second transistor in the TI circuit achieves this by being in reality a larger device, which could be modeled by several transistors, like the first stage, in parallel. This automatically increases the dc current by the scaling factor. The same is even more true for the output stage. Likewise, bipolar transistors can also be scaled. A visual inspection of some transistors in their packages will reveal that they are really not just one transistor, but several operating in parallel, each with its own ballast resistors (for bipolar transistors) or other means (for FETs).

Using tuned circuits rather than wideband matching, it is possible to go to Class AB or B operation. Class AB is really a hybrid between A and B, and Class B is probably better known. Figure 3-63 shows the Class B operating point. Most Class B stages are operated in push–pull because, together with the tuned circuit, one-half of the cycle is supplied by the other transistor. The definition of Class B operation is a conducting angle of 90° (see Figure 3-63). At this point it should be noted that some older texts define the conducting angle as being twice the value we have defined in this chapter.

Without going into further details, the highest efficiency in Class B operation is 78.5%. Based on the fact that modern wireless applications generally use a hybrid of amplitude and angle modulation, pure Class C operation finds little application today because of the severe nonlinear distortion it introduces. A Class B stage has less "power gain" than a Class A stage but provides much more output power. We have already mentioned the problem associated with characterizing a Class A stage in terms of power gain. That's the only time Kirchhoff's equations are valid, meaning that the sums of all currents and voltages have to be zero. Because Kirchhoff's equations do not consider nonlinearities and harmonic contributions,

the very moment we move into considering medium- and higher-power amplifiers, the nonlinearities will dominate and the performance of the stage is best analyzed with a high-quality circuit simulator with the appropriate dynamic range and good modeling capability. Without question, this applies to Class C operation also. As pointed out previously, the high-gain amplifier is optimized for gain rather than noise figure and typically requires more collector or drain current. This can easily be seen by rebiasing the FET amplifier as shown above to a significantly higher value of 33 mA. This is accomplished by changing the source resistor down to 3 Ω. As a result of this, we obtain the gain, matching, and noise figure responses shown in Figure 3-87.

It may be a surprise to the reader that the actual noise figure now is less than that shown in Figure 3-86. This device is now operating at 33 mA; the reason for this has to do with the fact that if the transconductance is too low, the noise figure also increases, and the NF curve for the FET has a similar-looking current-dependent minimum—a curve that is frequently not supplied by the manufacturer. This amplifier, now operating at 33 mA, should have quite good large-signal capabilities; we will examine this shortly.

Another interesting piece of information is the display of the dc I–V curves with the "load line." Since we have reactances in the circuit, we do not see a line but rather an ellipse, indicating the storage of energy. (The actual load line would be a line through the two foci of the ellipse.) This curve allows us to see immediately that at this operating point the transistor is neither going into current saturation nor becoming voltage limiting. The input power can be increased further until such a condition exists. Figure 3-88 shows the dc I–V curve for the amplifier shown in Figure 3-85.

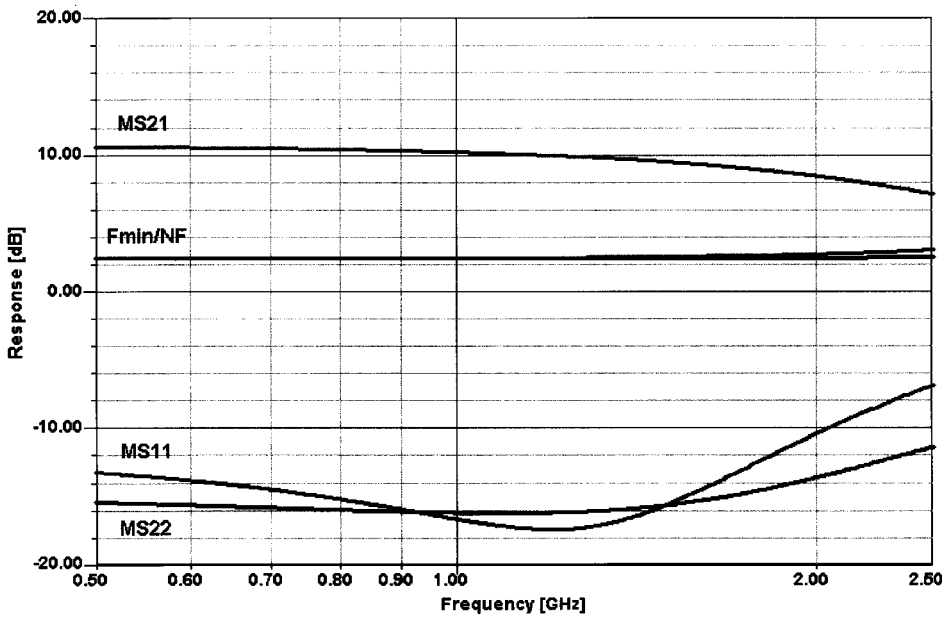


Figure 3-87 Frequency-dependent gain, matching, and noise performance for the GaAsFET amplifier of Figure 3-85 rebiased for a drain current of 33 mA.

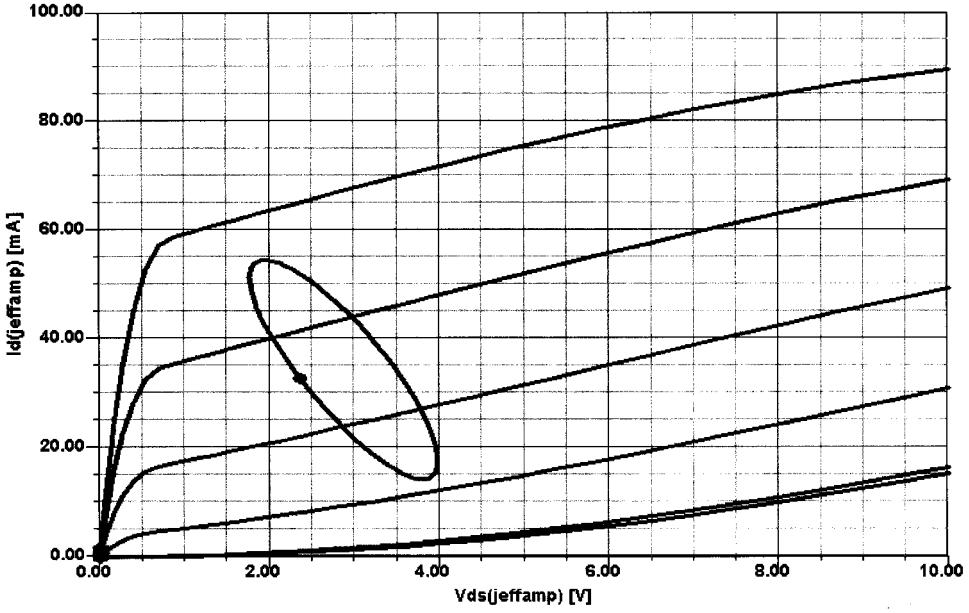


Figure 3-88 The dc *I–V* curves for the GaAsFET operating at I_D = 33 mA, including the ac load line. The load line is an ellipse rather than a straight line for the reason described in the text.

Besides looking for the adjacent-channel power ratio (ACPR), the other traditionally important points are third-order intercept point and the equivalent of a multitone simulation, as used for CATV application but relevant to the multisignal environment of wireless applications. Figure 3-89 shows the determination of the amplifier's third-order intercept point. It is most important for the reader to understand that very low values for input and output level have to be used to determine the crossover point. If this advice is not followed, totally erroneous numbers can occur.

Since an antenna supplies a large number of signals (this is true of handheld radios as well as base-station sites), a three-tone standard has been adopted as a good approximation of multisignal operation. This standard is in accordance with the German DIN 45004E,3.3 three-tone measurement. The requirement is that the intermodulation distortion products at the output are better than 60 dB suppressed relative to the largest of the three tones. One tone is used as the reference, and the other two tones are 6 dB less. This measurement has been derived from television, whereby the video transmitter, based on the vestigial sideband power, is 6 dB above the two sidebands produced by the FM sound transmitter using a modulation index of 5. The result of a three-tone test can be seen in Figure 3-90. This amplifier clearly meets the DIN specification for linearity at input levels up to –9/–15 dBm at the input.

Bipolar transistors are also quite popular for this purpose. Figure 3-91 shows a single-stage amplifier with an intrinsic transistor and all the other "hidden" intrinsics, normally not visible, added. The three diodes belong to those intrinsic elements, which are needed for accurate microwave modeling but which are not provided by the standard nonlinear models. The reader will see that we went into great detail to model all the necessary parts. This feedback amplifier, whose frequency response is shown in Figure 3-92, has been quite

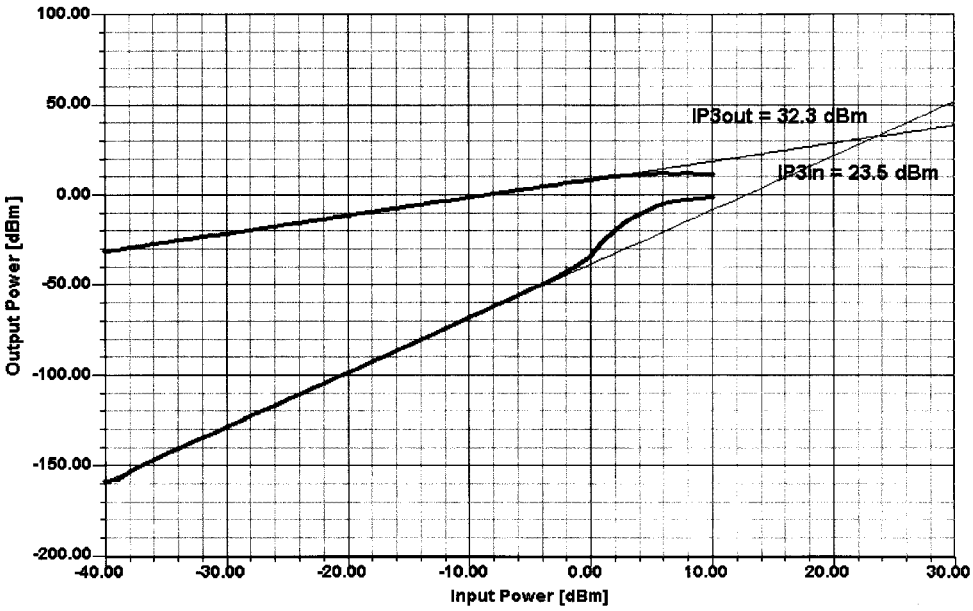


Figure 3-89 Determination of the third-order intercept point based on a two-tone measurement using tones spaced at 5% of the operating frequency (1.8 GHz).

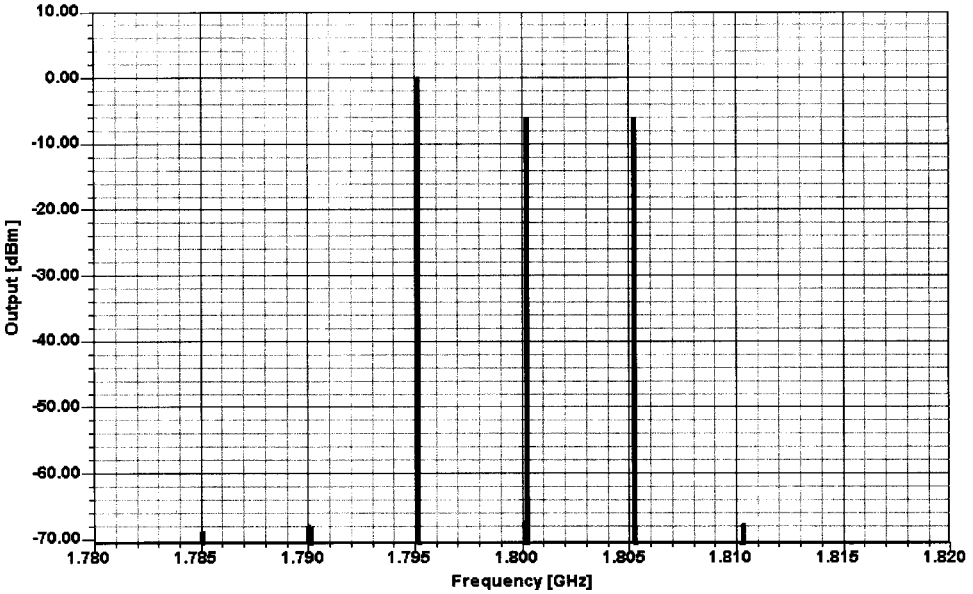


Figure 3-90 Simulated three-tone analysis of the GaAsFET amplifier with $I_D = 33$ mA.

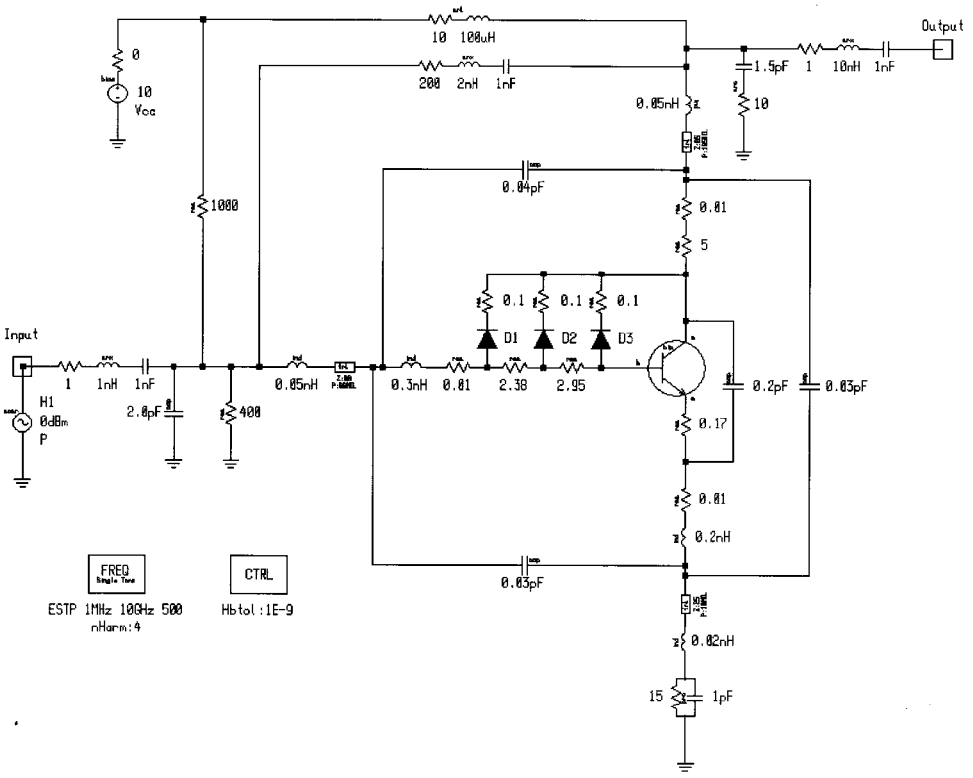


Figure 3-91 A high-linearity single-transistor amplifier modeled in great detail.

difficult to design, but there are monolithic circuits available where the higher gain–bandwidth product is being used to advantage.

Figure 3-92 shows the predicted frequency-dependent gain, matching, and noise figure at a bias point of 108 mA at 10 V. Using the same DIN three-tone test procedure, Figure 3-93 shows the expected good result.

Since this test requires two equal signals and one larger signal, we need to expect one IMD product larger than the rest. We have set our window for a dynamic range of 70 dB, and yes, in the lower left corner, we can see the product between the first two signals, while the other products are below the –50 dB level.

As a little task, Figure 3-94 shows a simple amplifier, based on a modern Siemens transistor, which is a feedback amplifier with fewer modeling details surrounding the transistor. Its associated frequency-dependent performance is shown in Figure 3-95. Interested readers are invited to improve its performance by using their skills and intuition.

Permitting higher currents and actually combining two transistors leads to our first MMIC. The following simulation is based on the MGA64135. Needless to say, there are some compensation networks inside the MMIC and distributed elements that affect the frequency response. However, with this arrangement, it is possible to obtain an appreciable gain at our wireless frequencies. The noise figure, based on the feedback, however, it not too impressive. This is a problem that occurs with most feedback amplifiers, which try to minimize the distortion and yet would like to keep the noise figure down. Figure 3-96 shows the schematic

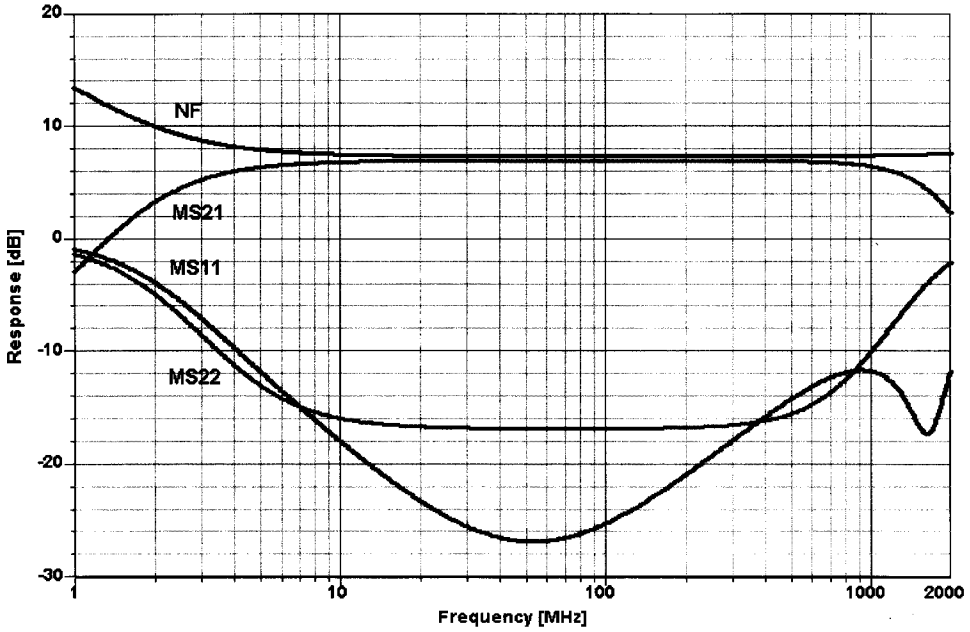


Figure 3-92 Simulated frequency-dependent gain, matching, and noise figure performance of the circuit in Figure 3-91.

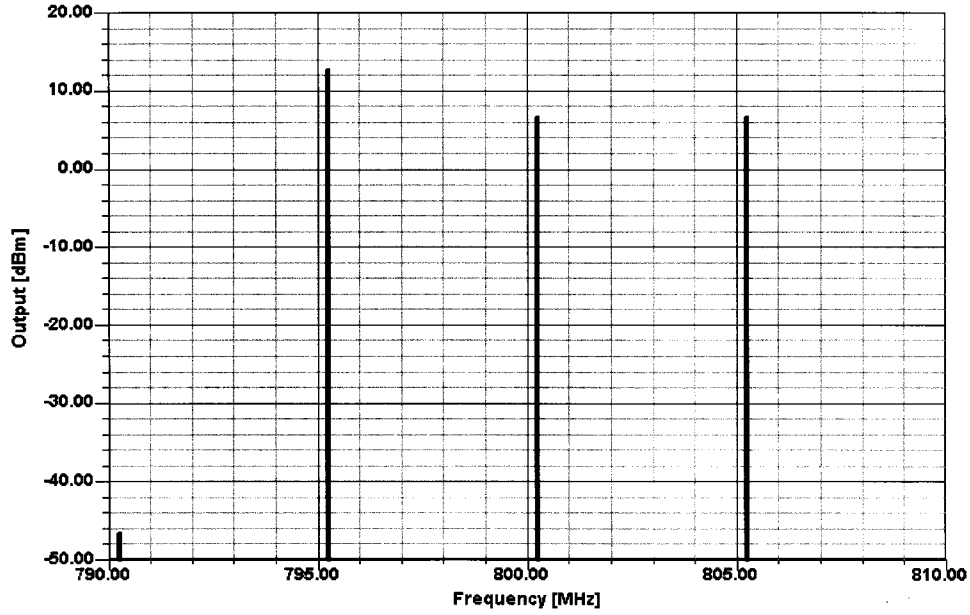


Figure 3-93 Simulated three-tone output spectrum of the amplifier in Figure 3-91.

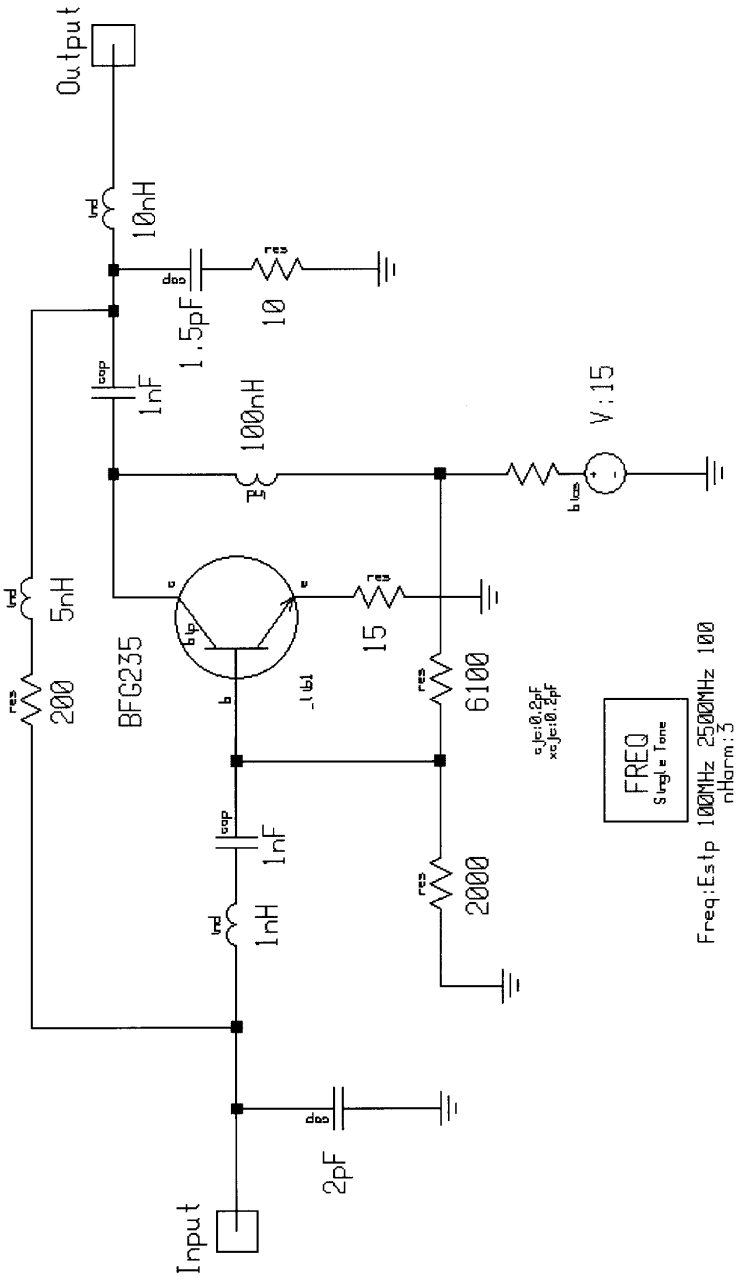


Figure 3-94 A high-linearity amplifier using the Siemens BFG235.

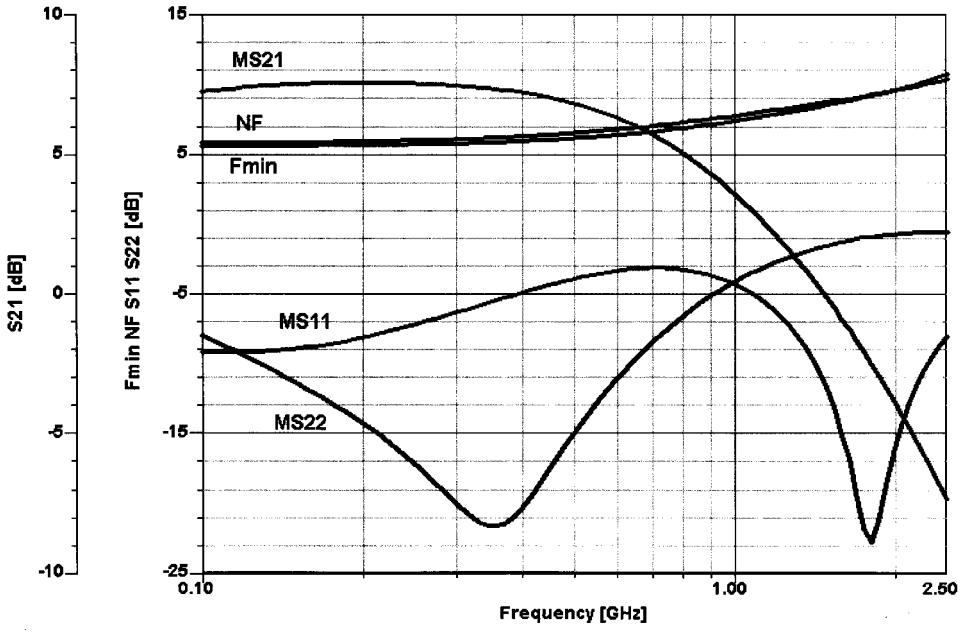


Figure 3-95 Frequency-dependent gain, matching, and noise performance of the amplifier in Figure 3-94.

of the GaAs MMIC; Figure 3-97 shows its frequency-dependent gain, matching, and noise performance; Figure 3-98 shows its two-tone response, which is used to obtain the intercept point ($IP_{3,in}$ = 8.5 dBm, $IP_{3,out}$ = 22 dBm); and, finally, Figure 3-99 shows the three-tone test results at an input RF level of –21/–27 dBm. While the gain–bandwidth product is considerably larger than the single-stage amplifier we examined before, the intercept point and three-tone performance do not reach the same values. As a general rule, the dynamic

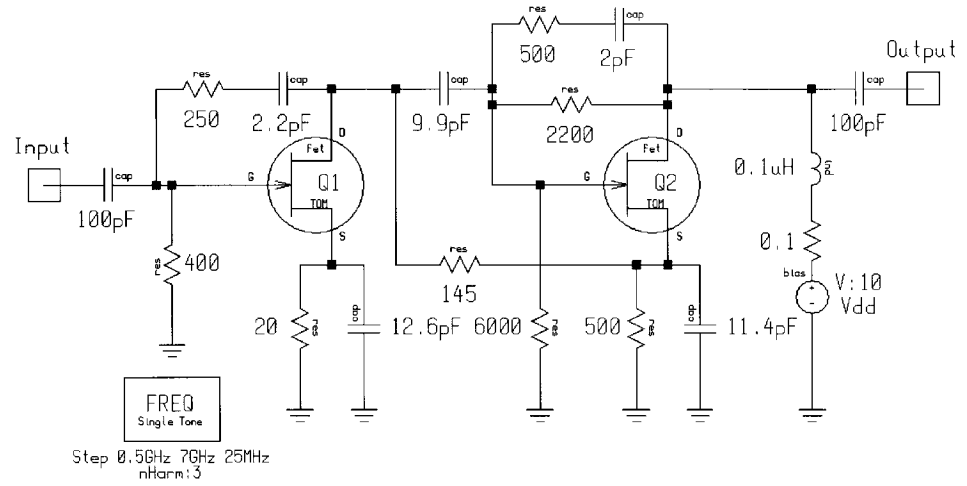


Figure 3-96 Simulation schematic for the MGA64135 MMIC amplifier. Q1 runs at 17.4 mA; Q2 at 32.2 mA.

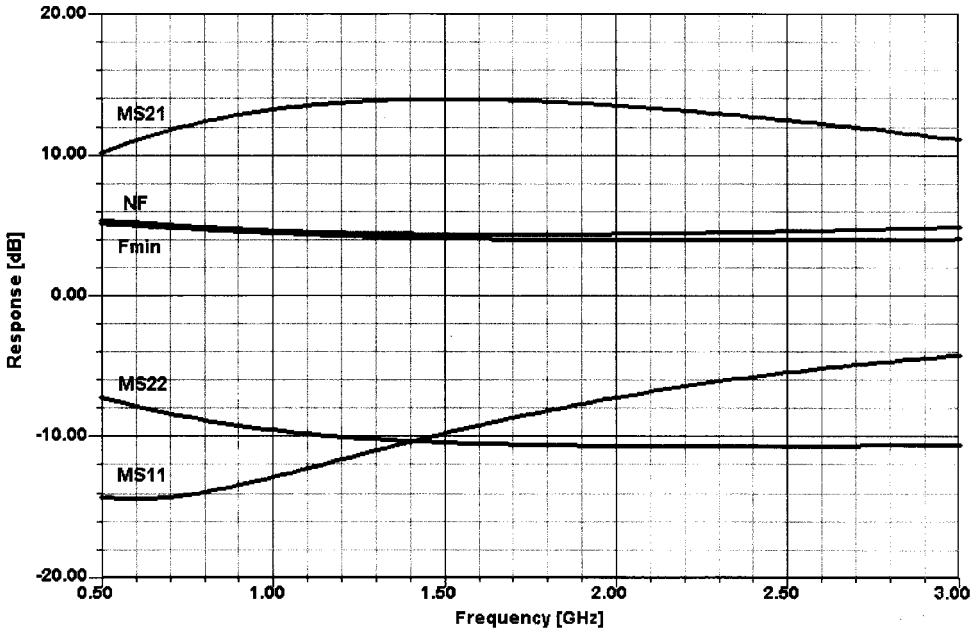


Figure 3-97 Frequency-dependent gain, matching, and noise figure performance for the MGA64135 MMIC amplifier.

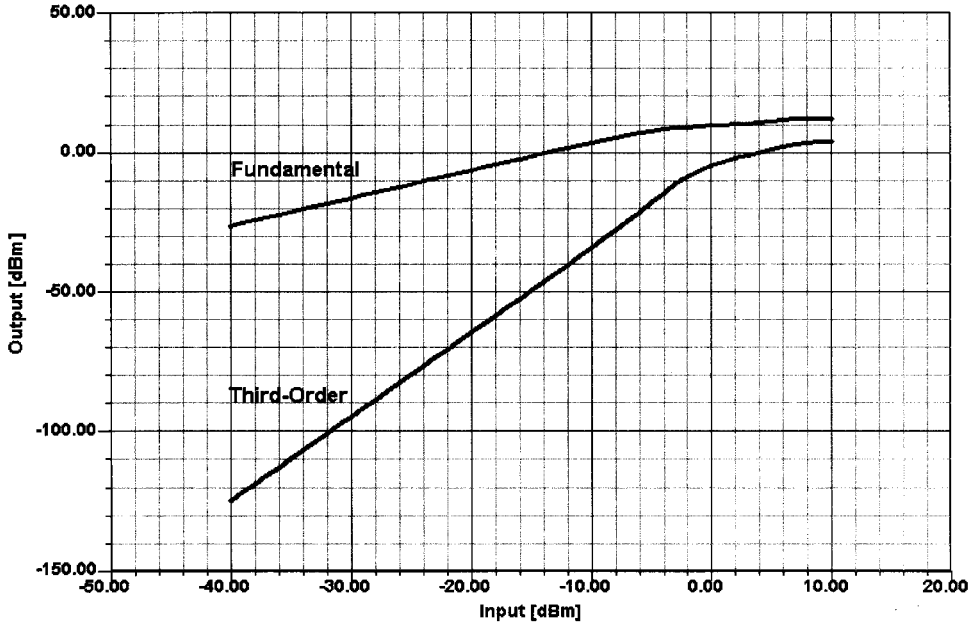


Figure 3-98 Simulated two-tone testing of the MGA64135 MMIC amplifier. $IP_{3,in}$ = 8.5 dBm and $IP_{3,out}$ = 22 dBm.

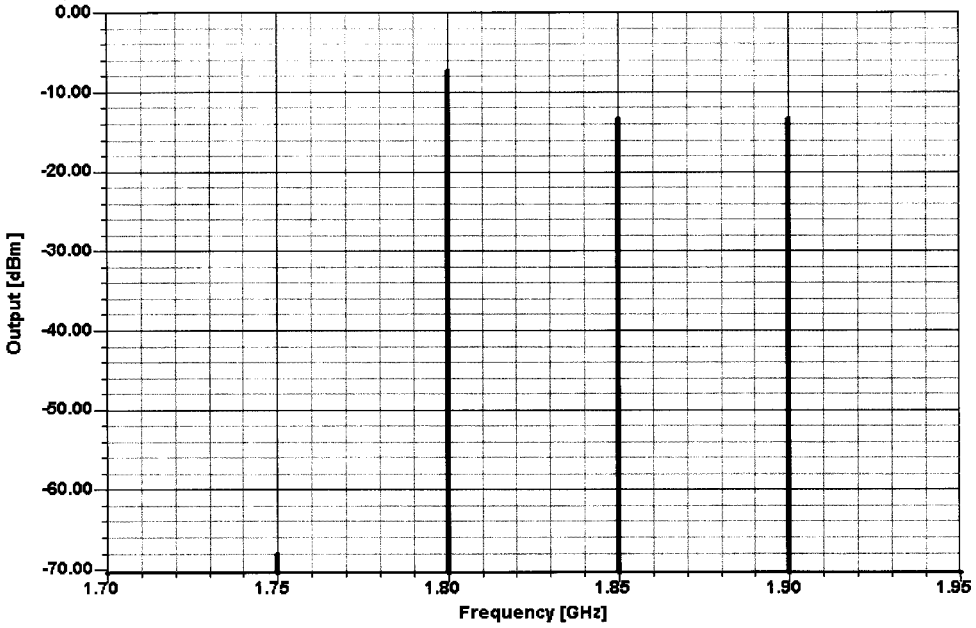


Figure 3-99 Simulated three-tone test of the MGA64135 MMIC. The test tones were 1.80 GHz (–21 dBm), 1.85 GHz (–27 dBm), and 1.90 GHz (–27 dBm). The 1.75-GHz IMD product is down 60.5 dB relative to the 1.80-GHz signal.

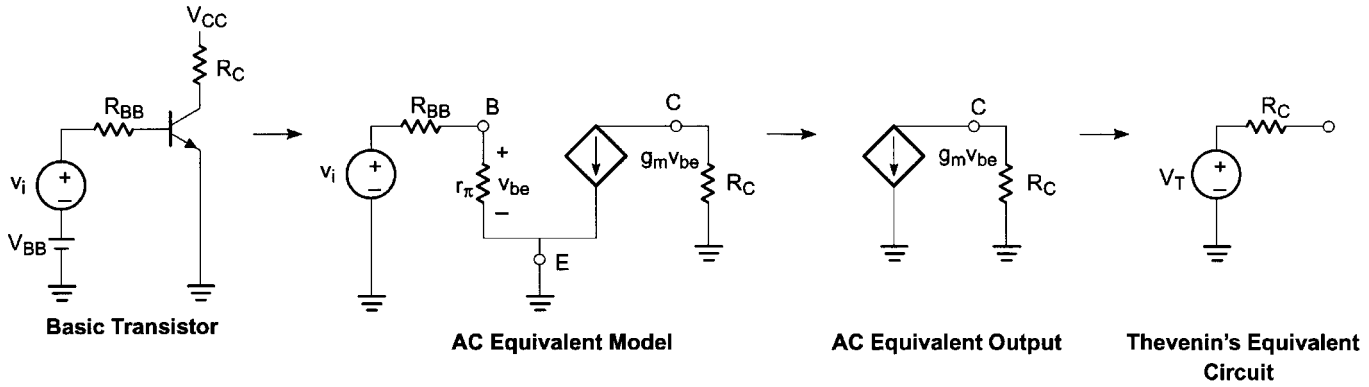
range is directly proportional to the power dissipated in the devices: specifically, the current. As explained earlier several times, the first-order nonlinearity comes from the voltage-to-current transfer characteristic, such as the transconductance itself, and can only be improved by a combination of resistive feedback in the emitter/source and by high currents. At these dc levels, the base–emitter (gate–source) junction gets linearized because the differential values are getting smaller compared to the parasitic loss resistances, including external dc/RF feedback. Once again, the voltage feedback reduces the voltage clipping but is initially more responsible for the input and output matching.

3-2-4 Low-Voltage Open-Collector Design [22–25]*

Why R_C Acts Like a Source Resistor. An open-collector output allows a designer the flexibility to choose the value of the R_C resistor. Choice of this resistor value not only sets the dc bias point of the device but also defines the source impedance value. Figure 3-100 shows the ac model of the transistor. Converting the output structure by applying a Norton and Thevenin transformation, one can conclude that R_C becomes the source resistance. Thus, by choosing R_C to be equal to the load, maximum power transfer will then occur.

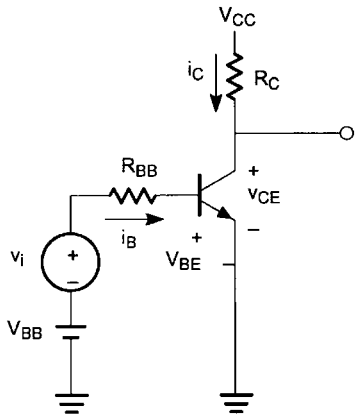
Figure 3-101 shows an active transistor with collector and base resistors. From basic transistor theory, Eq. A in Figure 3-101 is generated and has the same form as the general equation for a straight line ($y = mx + b$).

*Based on portions of the Philips Semiconductors/Signetics RF Communications Products Application Note AN1777, "Low-Voltage Front End Circuits: SA601, SA620," August 20, 1997. Used with permission.



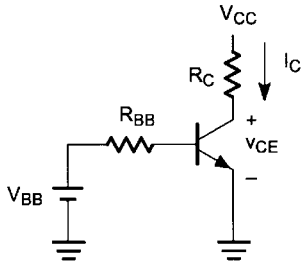
Symbol Convention:
DC: represented by uppercase symbols with uppercase subscripts, i.e., I_C
AC: represented by lowercase symbols with lowercase subscripts, i.e., i_C
Combined AC & DC: represented by lowercase symbol with uppercase subscript, i.e., i_C

Figure 3-100 *Rc* as source resistor.



DC Model

- Kill AC sources
- Open capacitors
- Short inductors

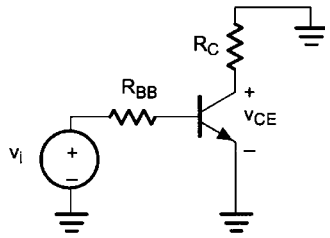


$$I_C = \left(\frac{-1}{R_C}\right) V_{CE} + \frac{V_{CC}}{R_C}$$

Eq. A

AC Model

- Kill DC sources
- Short capacitors
- Open inductors



$$i_C = \frac{-1}{R_C} (v_{CE})$$

Eq. B

Figure 3-101 Basic transistor analysis.

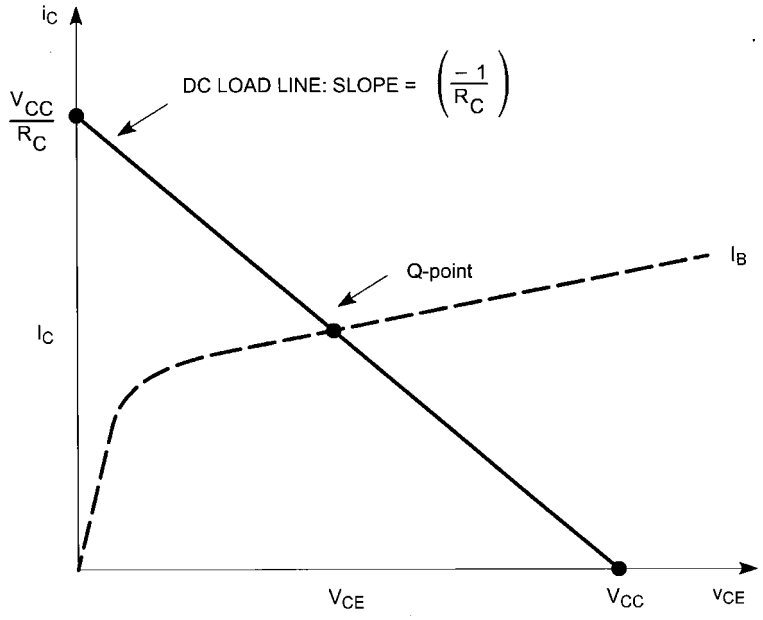


Figure 3-102 Load line and Q-point graph.

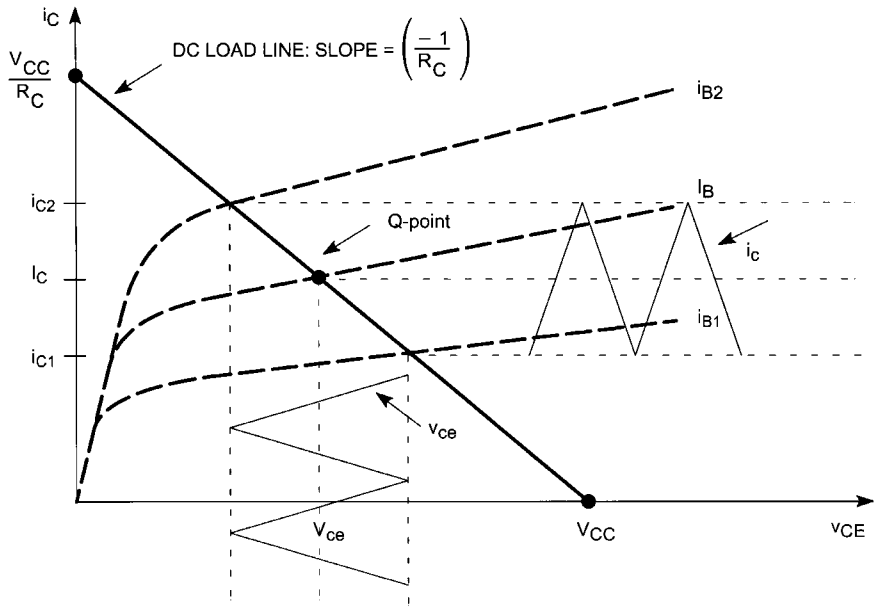


Figure 3-103 Graphical analysis for the circuitry in Figure 3-101.

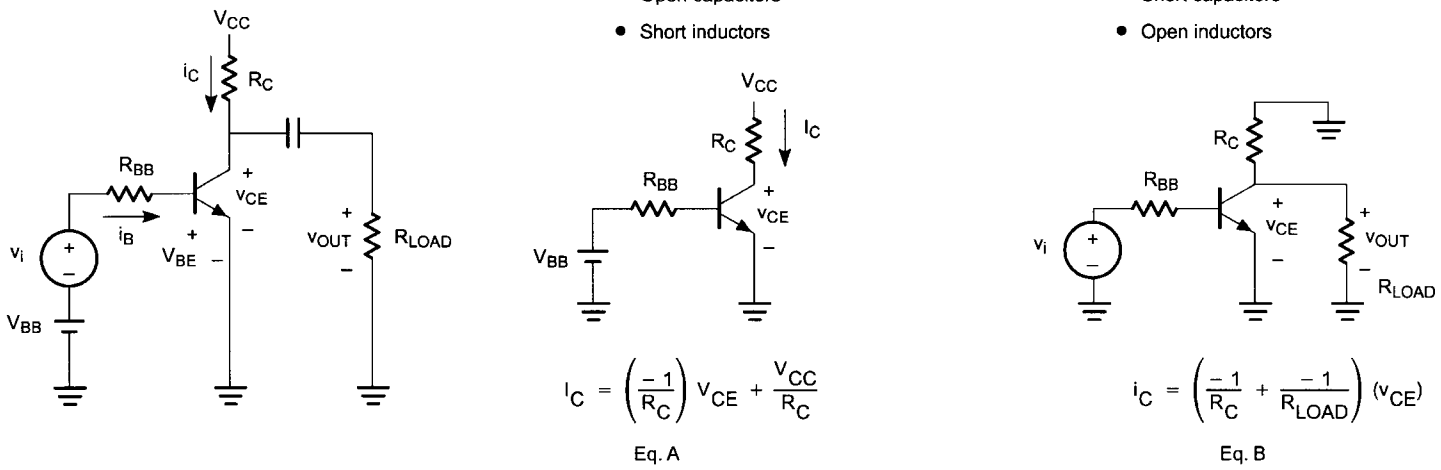


Figure 3-104 Basic transistor analysis with R_{LOAD}.

The slope of the dc load line is generated by the value of the collector resistor ($m = (1/R_C)$ and is shown in Figure 3-102. For a given small-signal base current, the collector current is shown by the dotted curve. The intersection of the dotted curve and the dc load line is called the *quiescent point* (Q-point) and dc bias determinant. The location of the Q-point is important because it determines whether the transistor is operated in the cutoff, active, or saturation region(s). In most cases, the Q-point should be in the active region because this is where the transistor acts like an amplifier.

Figure 3-103 shows the ac collector–emitter voltage (V_{CE}) output swing with respect to an ac collector current (I_C). Collector current is determined by the ac voltage presented to the input transistor's base (v_i) because it affects the base current (i_b), which then affects i_c. This is how v_i is amplified and seen at the output. Recall that this is with no external load (R_{LOAD}) present at the collector. Since no load is present, the slope of the ac load line is identical to that of the dc load line as seen in Eqs. A and B of Figure 3-101 ($m = -1/R_C$).

Open Collector with R_{LOAD}. A filter with some known input impedance is a typical load for the output of the transistor. For simplicity, we will assume a resistive load (R_{LOAD}) and neglect any reactance. Since a resistive load is used (see Figure 3-104), the ac output swing is measured at V_{OUT} or V_{CE}.

A dc blocking capacitor is used between the R_{LOAD} and the V_{CE} output to assure that the Q-point is not influenced by R_{LOAD}. It is also necessary to avoid passing dc to the load in applications where the load is a SAW filter. However, R_{LOAD} will affect the ac load line, which is seen in Eq. B in Figure 3-104. Note that the V_{CE} voltage swing is reduced; thus, the V_{OUT} signal is reduced (see Figure 3-105).

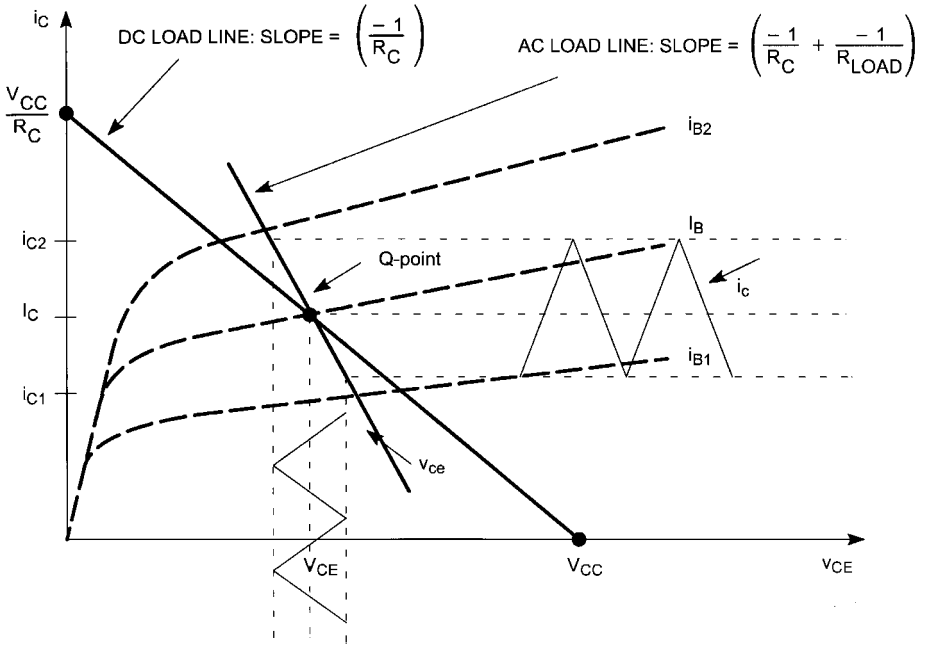


Figure 3-105 Graphical analysis for the circuitry in Figure 3-104. The ac load line intersects the Q point, and the slope is steeper than the dc load line due to R_{LOAD}, which makes V_{CE} smaller.

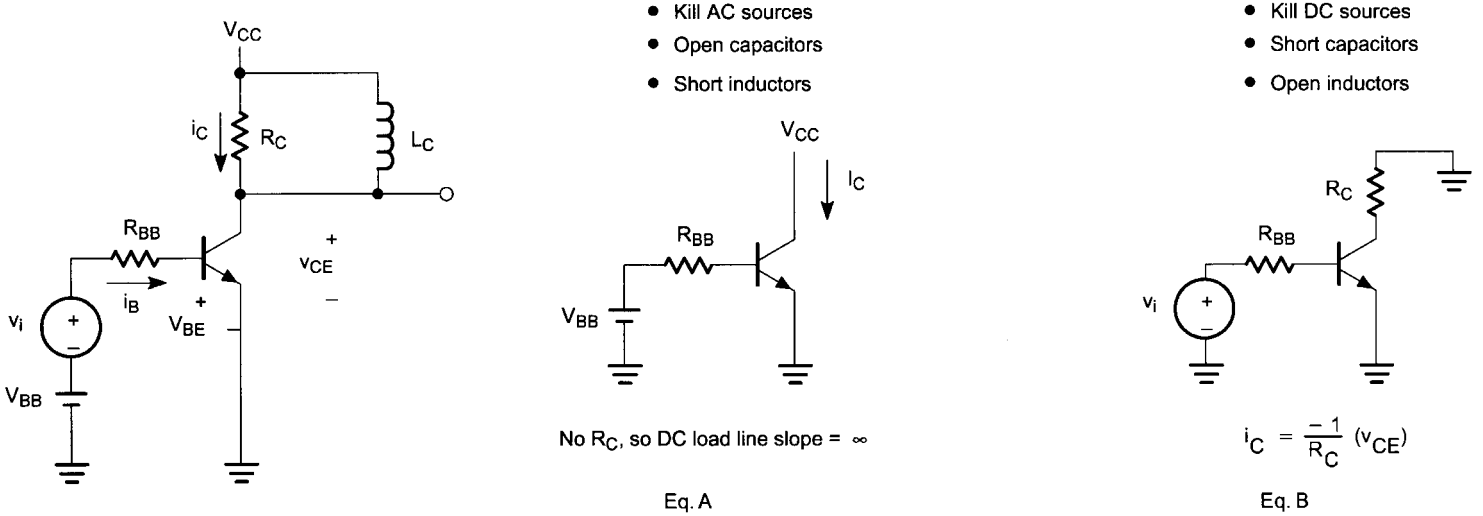


Figure 3-106 Basic transistor analysis with inductor added to collector.

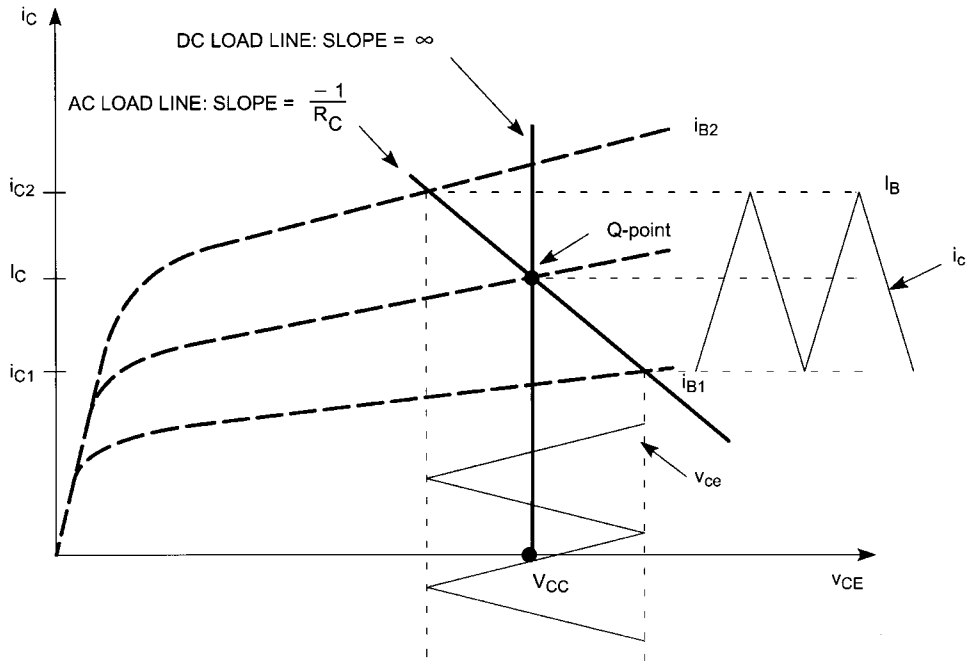


Figure 3-107 Graphical analysis for the circuitry in Figure 3-106.

Since the values of R_C and R_{LOAD} affect the ac load line slope, the value chosen is important. The higher the impedance of R_{LOAD} and R_C, the greater the ac output swing will be at the output, which means more conversion gain in a mixer. This is due to the slope getting flatter, thus allowing for more output swing.

Open Collector with Inductor (L_C). Adding an inductor in parallel with R_C can increase the ac output signal V_{CE}. Figure 3-106 shows the dc and ac analyses of this circuit configuration. In Eq. A in Figure 3-106, there is no R_C influence because the inductor acts like a short in the dc condition. This means the slope of the dc load is infinite and causes the Q-point to be centered around $V_{CC,}$ thus moving it to the right of the curve. The ac load line slope is set only by R_C because no load is present. Note that it has the same ac load line slope as the first condition in Eq. B of Figure 3-101.

Referring to Figure 3-107, one might notice that the base current (ac and dc) curves spread open as V_{CE} increases. This is caused by a noninfinite Early voltage (see Figure 3-108), which causes the collector current to be dependent on V_{CE}. Taking advantage of this nonideal condition, the peak-to-peak ac output swing V_{CE} can thereby be increased by moving the dc Q-point to the right due to the wider spreading between the curves corresponding to different base currents. Figure 3-109 combines Figures 3-103 and 3-107 to show the different ac output signals with different Q-points.

Looking at the ac output level, one might ask how the V_{CE} peak voltage can exceed the supply voltage V_{CC}. Recall that the inductor is an energy-storing device ($v = L\ di/dt$). Therefore, total instantaneous voltage is V_{CC} plus the voltage contribution of the inductor.

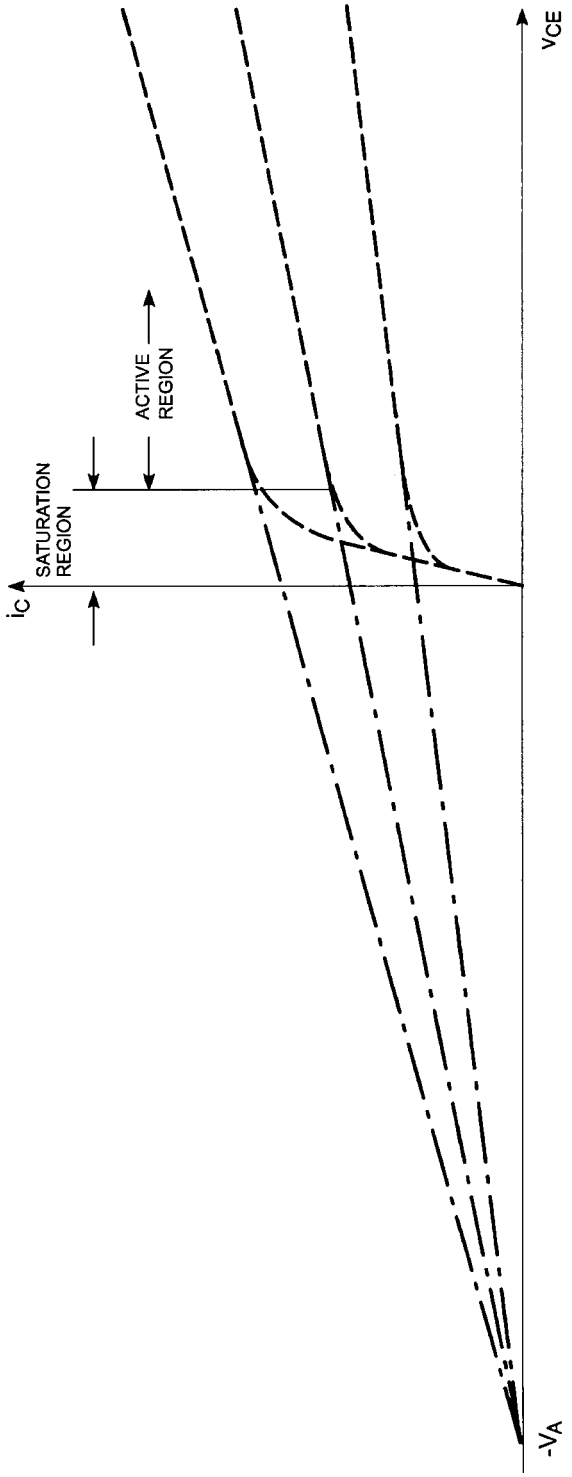


Figure 3-108 Graphical representation of Early voltage effect.

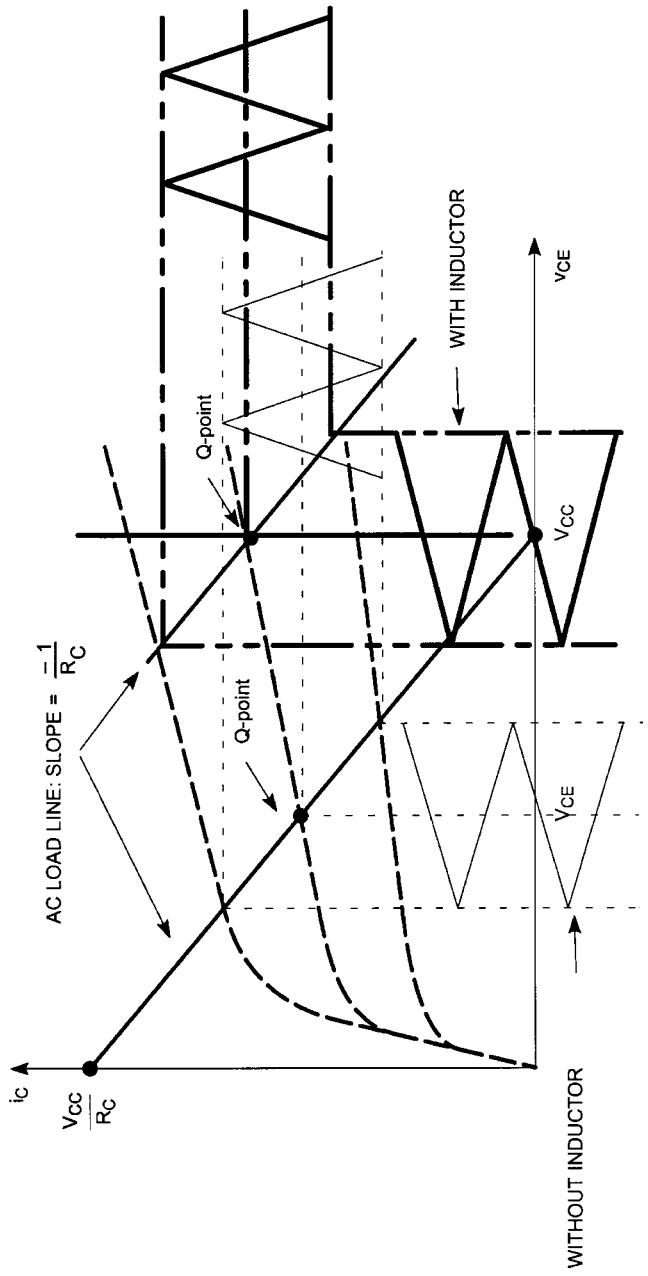


Figure 3-109 Comparison of open-collector circuit with inductor versus without inductor.

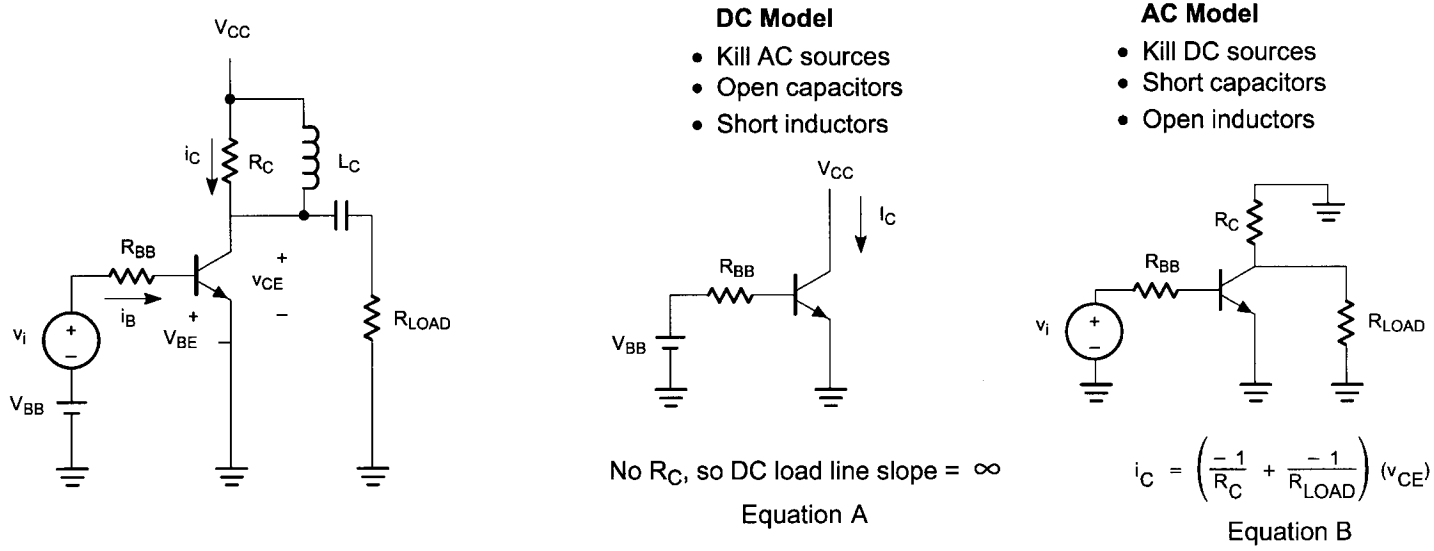


Figure 3-110 Basic transistor analysis with inductor and R_{LOAD}.

Open Collector with Inductor (L_C) and R_{LOAD}. Figure 3-110 shows the dc and ac analyses with the inductor and load resistor. Again, from the dc analysis, the inductor causes R_C to be nonexistent so the dc load line is vertical. In the ac analysis, the ac load line slope is influenced by both R_C and R_{LOAD} (see Eq. B in Figure 3-110). The ac load line slope is the same as the example of the open collector without the inductor. Figure 3-111 shows the response for the open collector with L_C and R_{LOAD}. Figure 3-112 combines Figures 3-105 and 3-111 showing the increase in ac output swing.

In conclusion, for the load line, R_C plays a role in setting up the bias-determined Q-point as well as the ac source impedance. However, when an inductor is placed in parallel with R_C, a different Q-point is set and the ac source impedance is altered. Moving the Q-point takes advantage of the transistor's nonideal i_c dependence on V_{CE} to get more signal output without having to change the base current. Since R_C is in parallel with R_{LOAD} in the ac condition, it influences the ac load line slope.

Flexible Matching Circuit [25]. A useful variation of the open-collector matching concepts previously outlined provided the capability of delivering equal power to two unequal resistive loads. This allows the power delivered to the load to be measured indirectly at another test point in the circuit where the impedance can be defined arbitrarily. If this impedance is defined to be 50 Ω, a spectrum analyzer can easily be placed directly into the circuit. This is an excellent troubleshooting technique and a valuable option to have available in high production environments.

Figure 3-113 shows the schematic for this flexible matching circuit. In this circuit, C_B functions only as a dc blocking capacitor and presents a negligible impedance at the frequency of interest. Recall from the previous open-collector matching discussions that,

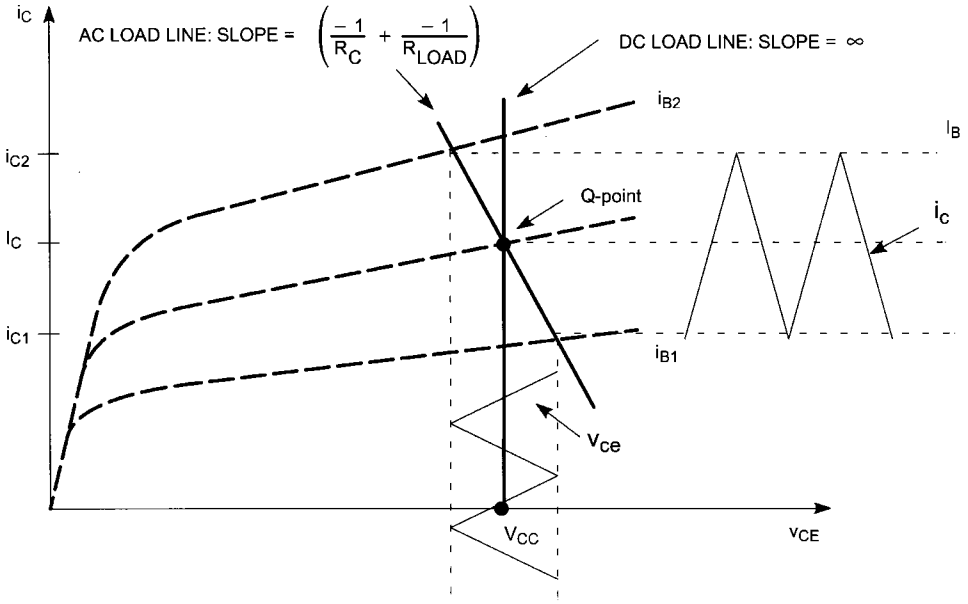


Figure 3-111 Graphical analysis for the circuitry in Figure 3-110.

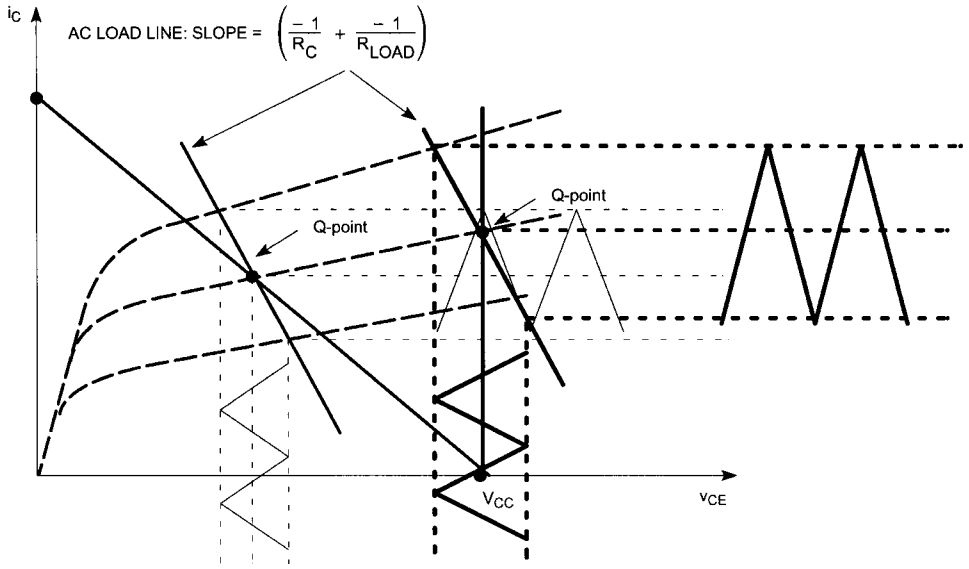


Figure 3-112 Comparison of open-collector R_{LOAD} circuit with inductor versus without inductor.

when R_C is placed in parallel with an inductor, it has no effect on the Q-point but does influence the slope of the ac load line as

$$\text{Slope} = \frac{-1}{R_C} + \frac{-1}{R_{\text{LOAD}}} \tag{3-166}$$

The capacitor C_S not only functions as a dc blocking capacitor but is also chosen such that the impedance presented by the combination of L, R_C, and C_S is equal to R_{LOAD} for optimum power transfer. The analysis is done in the following manner. First, note that inductor L is connected to V_{CC}, which is an affective ac ground. So L can be redrawn to ground. Next, R_C and C_S are converted to their parallel equivalent values as shown in Figure 3-114.

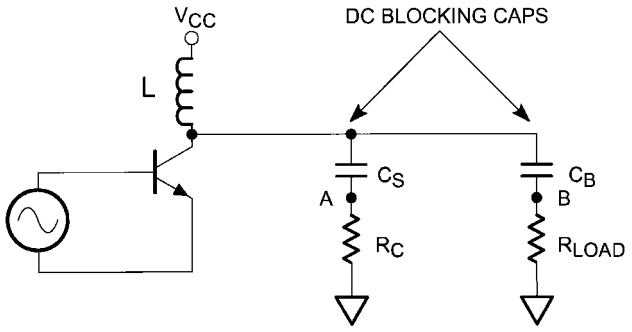


Figure 3-113 Flexible matching circuit.

Figure 3-114 Converting from series to parallel configuration.

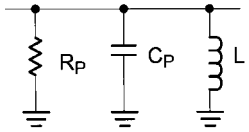


Figure 3-115 Converting from series to parallel configuration. Goal: Find L_C and C_p such that R_p looks like the R_{LOAD} value at the resonant frequency.

The resulting parallel *LCR* circuit is shown in Figure 3-115. At resonance, the parallel L–C_P combination will be an effective open circuit leaving only R_P. R_P is then simply chosen to be equal to R_{LOAD}.

Starting with the most simple transistor configuration and using *RC* types of dc-coupled amplifier circuits, we next present an overview of possible configurations for this application.

3-3 SINGLE-STAGE FEEDBACK AMPLIFIERS

This amplifier (Figure 3-116) is a standard minimal configuration amplifier with reduction of current distortion because of the emitter feedback. The voltage gain is

$$\text{VG} = \frac{R_L}{(26\text{mV})/I_C + R_e} \tag{3-167}$$

where R_L is the total load seen at the collector.

The feedback amplifier in Figure 3-117 uses voltage feedback to "correct" the input and output impedance to be close to 50 Ω but does not improve the current distortion. The voltage gain of this circuit is

$$\text{VG} = \frac{R_F}{R_g} \tag{3-168}$$

$$R_{\text{IN}} = R_{\text{OUT}}$$

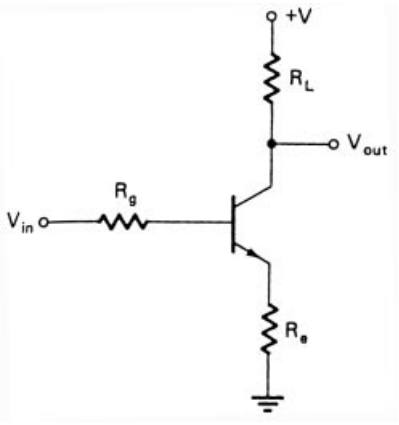


Figure 3-116 Single-stage amplifier with current feedback (provided by R_e).

$$= R_F R_D \quad \text{with} \quad R_D = \frac{26\text{mV}}{I_C} \tag{3-169}$$

Example. For $I_C = 10$ mA $\rightarrow R_D = 2.6\ \Omega$ and assuming $R_F = 200\ \Omega$, the input resistance $R_{\text{IN}} = 22.63\ \Omega$. The gain is calculated by getting the inverse of $R_D = g_m = 0.39$ A/V. Once this is multiplied with the output impedance of 22.63 Ω, the resulting gain becomes 10 log (8.6) or 9.3 dB. The gain can be increased by either using a larger collector current or a larger feedback resistor. It should be reminded that this circuit does not improve the distortion caused by the base–emitter junction.

To reduce the current distortion as well, we use the circuit shown in Figure 3-118, which contains both types of feedback. In reality, the collector–base resistor is frequently split into an unbypassed section responsible for the feedback and a larger resistor responsible for the appropriate biasing. The example circuit in Figure 3-119 shows this.

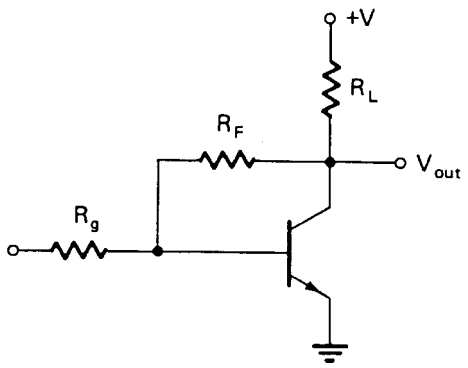


Figure 3-117 Single-stage amplifier with voltage feedback (provided by R_F).

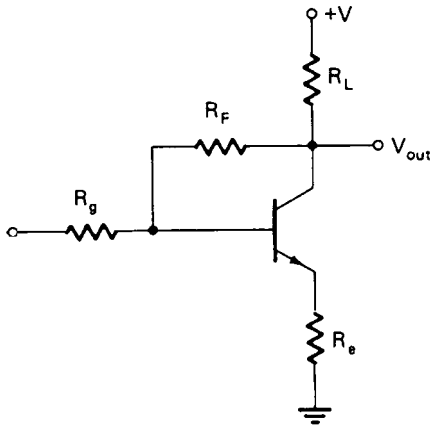


Figure 3-118 Single-stage amplifier with voltage and current feedback.

By using two of the fairly recent BFP420 transistors in parallel, and balancing the transistors by using the appropriate emitter resistors (which in total will result in 10 Ω) plus the bypassed value, we obtain the frequency-dependent gain, matching, and noise responses shown in Figure 3-120. To determine the appropriate feedback resistor R_F, we recommend the formula

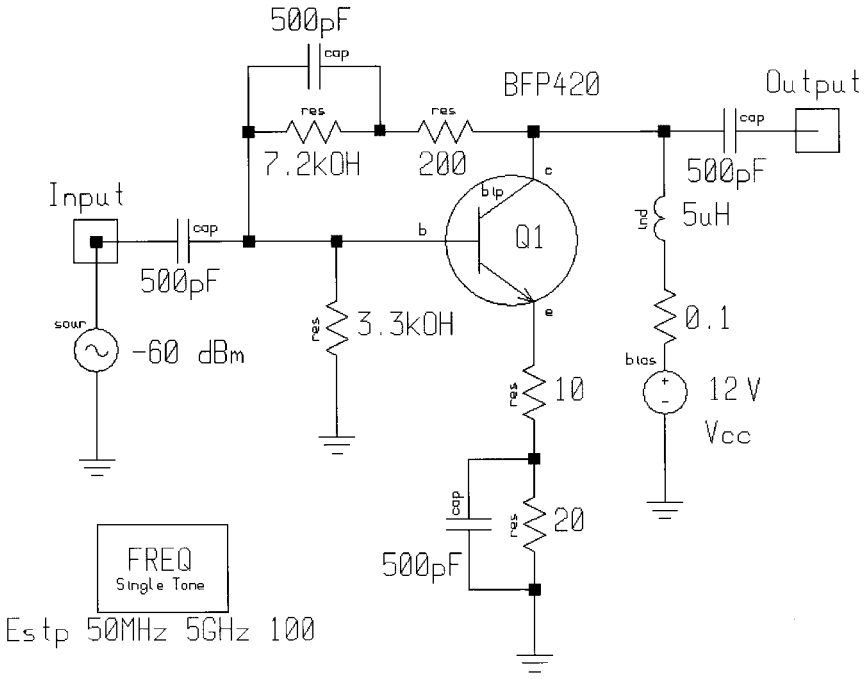


Figure 3-119 Practical amplifier circuit showing voltage and current feedback, and illustrating how reactances (in this case, 500-pF bypass capacitors) and multiple resistors can be used to provide different ac and dc resistances for the paths from collector to base and emitter to ground.

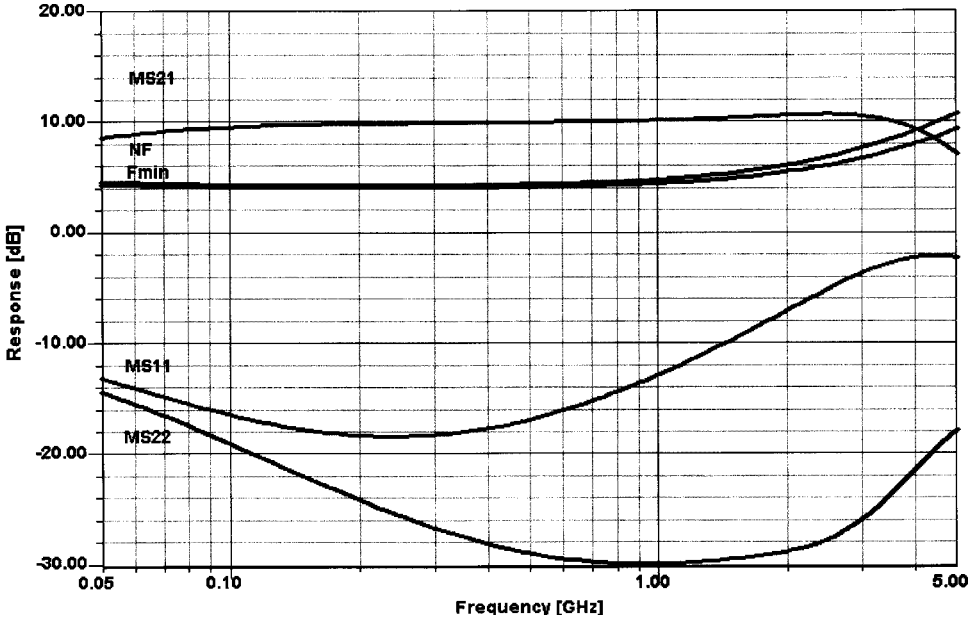


Figure 3-120 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-119.

$$R_F = \frac{(Z_{\text{in}}\, Z_{\text{out}})}{R_E} \tag{3-170}$$

This equation requires some additional comments. The combined shunt and series feedback only work as long as a sufficient transconductance is provided. In other words, there must be enough open-loop gain that will be reduced to obtain more bandwidth. A necessary condition for this approach, which really comes from the frequency range of less than 30 MHz, is that

$$Z_{\text{in}} = \frac{26\ \text{mV}}{I_C}\,\beta >> R_F \tag{3-171}$$

This requires a minimum transconductance of

$$g_m = \frac{1 + S_{21}}{Z_{\text{in}}} \tag{3-172}$$

where S_{21} must be a negative number that reflects the 180° phase shift. Of course, life becomes interesting when S_{21} goes through 1, but it (hopefully) never becomes negative and larger than 1 at the same time.

Example. The BFP420 has been used here frequently. A 10-dB gain would require an S_{21} of 3.16 ($\sqrt{10}$); therefore, we need to inspect the datasheet, which shows that the magnitude

of S_{21} is 3.16 around 4.5 GHz, yet the phase shift has changed from 180° to around 20°. In order to design an amplifier, we now set

$$R_F = Z_{\text{in}}(1 + |S_{21}|) \tag{3-173}$$

which is valid until S_{21} becomes too small. For a given 10 dB or 3.16 for S_{21}, the resulting RF is 4.16×50, or 208 Ω. The transconductance under these conditions becomes

$$g_m = \frac{1 + |S_{21}|}{Z_0} \tag{3-174}$$

where $Z_0 = 50$ Ω (the desired input impedance). g_m now becomes 83.2 mS. We now test the input and find

$$\frac{\beta}{g_m} = Z_{\text{in}} = \frac{50}{0.0832} = 600\ \Omega \tag{3-175}$$

Therefore, the emitter feedback resistor is not necessary to meet the previously mentioned condition whereby the input impedance needs to be higher than RF. Assuming the transconductance now is 0.00832, or 10% of the previous value, the input impedance would be 60 Ω, no longer meeting the prescribed requirement. Therefore, we need an unbypassed emitter resistor that fulfills the equation

$$R_E = \frac{Z_0^2}{R_F} - \frac{1}{g_{m_{\text{dc}}}} \tag{3-176}$$

where

$$g_{m_{\text{dc}}} = \frac{1 - S_{21_{\text{max}}}}{Z_0} \tag{3-177}$$

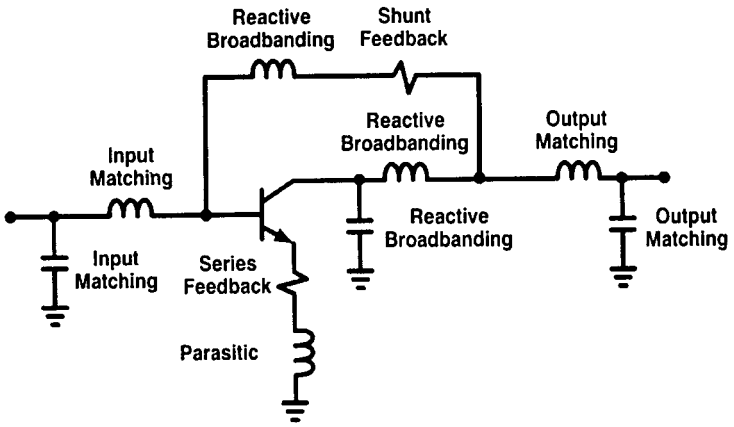


Figure 3-121 Elements of a feedback amplifier.

Again, the reader should be cautioned that in the common-emitter circuit, there is a 180° phase shift, and therefore $S_{21_{\max}}$ is a negative quantity, resulting in reality in

$$g_{m_{\text{dc}}} = \frac{1 + S_{21_{\max}}}{Z_0}$$

In our case,

$$g_{m_{\text{dc}}} = \frac{-S_{21}}{2Z_0} = \frac{38}{2 \times 50} = \frac{38}{100} = 0.38$$

so

$$R_E = \frac{50^2}{208} - \frac{1}{0.38} = 12 - 2.6 = 9.4\ \Omega \tag{3-178}$$

In practice, the designer would use a 10-Ω resistor.

Since the gain of the transistor continues to fall as a function of frequency, one can compensate to a degree for the cutoff frequency by putting an inductance in series that follows the equation

$$L_F = \frac{R_F}{2\pi f_{3\text{dB}}} \tag{3-179}$$

Assuming $f_{3\text{dB}}$ = 2 GHz, the resulting series inductance L_F = 16.6 nH. Basically, the same applies to circuits where the second transistor is used as an emitter follower to feed the signal back from the collector to the base of the first transistor.

In summary (see Figure 3-121):

- Resistive negative feedback at lower frequencies reduces open-loop gain and simultaneously matches input and output.
- Reactive elements set the gain and matching at high frequencies. Feedback is reduced at high frequencies.
- Emitter bypass capacitors are not used, as they would destroy the impedance match.

This is equally valid for FETs.

The next level of complexity can be reached by using two transistors with dc feedback.

3-3-1 Lossless or Noiseless Feedback

All the amplifiers shown above are based on a combination of resistive and reactive feedback. The very moment resistive feedback is used, the noise figure increases, as does the intercept point. As long as the intercept point increases overproportionally compared to the noise figure, this is an acceptable method. However, there are many cases in which an extremely low noise figure and a higher than normally achievable intercept point are required. To avoid this degradation of performance, Norton of Anzac developed a negative feedback structure based on the capabilities of the ferrite transformer. This patented technique (U.S. Patent Nos. 3,624,536 and 3,891,934, since expired), known as *lossless feedback*, provides lower noise figure and higher linear output than can be achieved using resistive feedback with the same transistor. Transformer feedback is used in which the transformer in the feedback network acts as a directional coupler. The coupling ratio between the input and output of the coupler

determines the magnitude of the feedback, and hence the gain. Across the frequency band over which the transformer operates as intended, the feedback is negative due to the in-phase coupling and the inversion of the active device.

This technique was used in amplifiers ranging from 5 to 500 MHz, 20 to 1000 MHz, and 300 to 1800 MHz. The difficulty associated with this type of amplifier is the fact that it is based on transformers that use twisted pairs of wires as transmission lines. In some implementations used at higher frequencies, actual directional couplers were used at the expense of bandwidth; in these circuits, the coupling was determined by the spacing of the coupler lines rather than transformer turns ratio. These microstrip couplers have much narrower bandwidths than transformers and are fairly large.

Another headache with the lossless feedback design is the fact that many of these transistor configurations are used in either grounded-base or grounded-emitter with transistors that have f_T values of 25 GHz and higher. Such devices are very hot, and achieving freedom from oscillation over a wide frequency range is a continuing problem, even for the circuits we have shown so far. A similar clever technique has been implemented in the termination-insensitive mixer, which we will examine in Chapter 4.

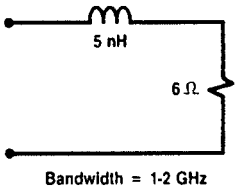
3-3-2 Broadband Matching

The circuitry above has enabled us to extend the frequency range of the amplifier and avoid the normal sharp cutoff. According to Fano [26], there is a limit to the bandwidth of matching. The concept is that impedance plus bandwidth determines the ability to match, since there is a Q of the circuit and there is a Q of the device. The following equations are valid:

$$|\rho_{\text{Bode}}| = e^{-(\pi Q_T / Q_{BP})} \tag{3-180}$$

$$Q_T = \frac{\sqrt{f_1 f_2}}{f_2 - f_1} \tag{3-181}$$

$$Q_{BP} = \frac{\text{Im}(Z)}{\text{Re}(Z)} = \frac{X_s}{R_s} \tag{3-182}$$



$$Q_{BP} = \frac{2\pi\sqrt{(2\text{ GHz})(1\text{ GHz})(5\text{ nH})}}{6\,\Omega} = 7.40$$

$$Q_T = \frac{\sqrt{(2\text{ GHz})(1\text{ GHz})}}{(2\text{ GHz} - 1\text{ GHz})} = 1.414$$

$$|\rho_{\text{Bode}}| = e^{-\left(\frac{\pi(1.414)}{7.40}\right)} = .549$$

$$\text{VSWR}_{\text{Bode}} = \frac{1 + .549}{1 - .549} = 3.43 : 1$$

Figure 3-122 Broadband matching example showing the Bode limit.

Figure 3-122 shows an example. This combination results in a VSWR of 3.43, which for practical purposes is already too high. The workaround is special circuitry (such as exponentially staggered impedance jumps in the matching network like 50 Ω > 20 Ω > 7 Ω > 3 Ω). This means we first transform the 50-Ω source of load impedance down to 3 Ω, and then look for a matching network that matches the transistor output/input, which is within a few ohms of 3 Ω. A detailed discussion of such matching can be found in the "Numerical Design" portion of Section 3-12-2 under the title "Broadband Matching Using Bandpass-Filter Networks—High-Q Case."

3-4 TWO-STAGE AMPLIFIERS

There are two popular configurations that have been used mostly for low-frequency applications. The first one (Figure 3-123, implemented in the circuit shown in Figure 3-124) shows a nice high-frequency performance assuming the emitter and collector currents are essentially the same. As for distortion products, this circuit behaves equivalently to the shunt impedance example from Figures 3-118 and 3-119, with the additional gain of the second transistor. Its frequency matching and noise performance are also shown, in Figure 3-125.

The next two-stage circuit (Figure 3-126, implemented in Figure 3-127) uses the collector current of the second stage as the feedback to the first. Proper biasing of this circuit is quite difficult and interactive. Figure 3-128 shows the frequency-dependent gain, matching, and noise figure. We believe that this circuit can be optimized further, but we leave such investigations to the reader.

A *cascode* amplifier is a special form of two-stage circuit. Although a cascaded pair acts like a single transistor, its merits are in its reduced feedback and resulting higher gain–bandwidth product. Figure 3-129 shows a simple cascode application that assumes the presence

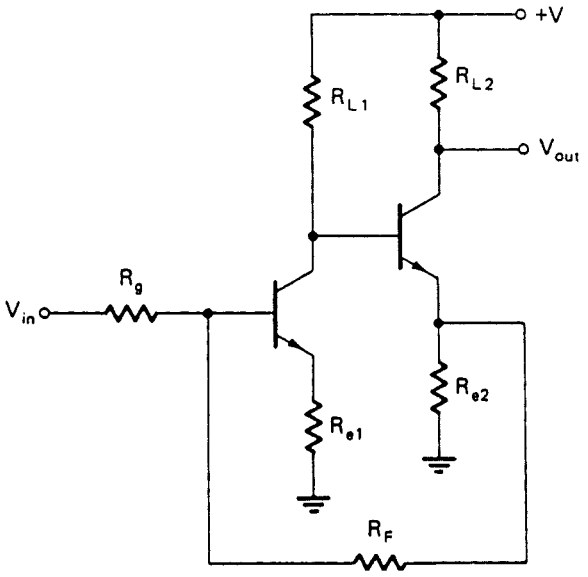


Figure 3-123 Two-stage amplifier with voltage feedback.

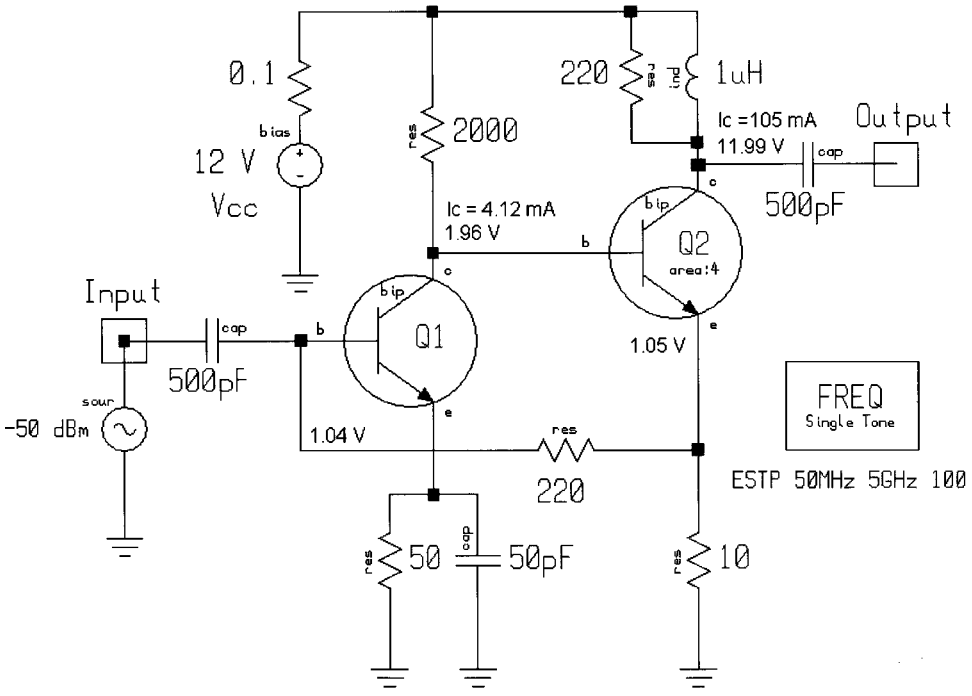


Figure 3-124 Implementation of the circuit in Figure 3-123.

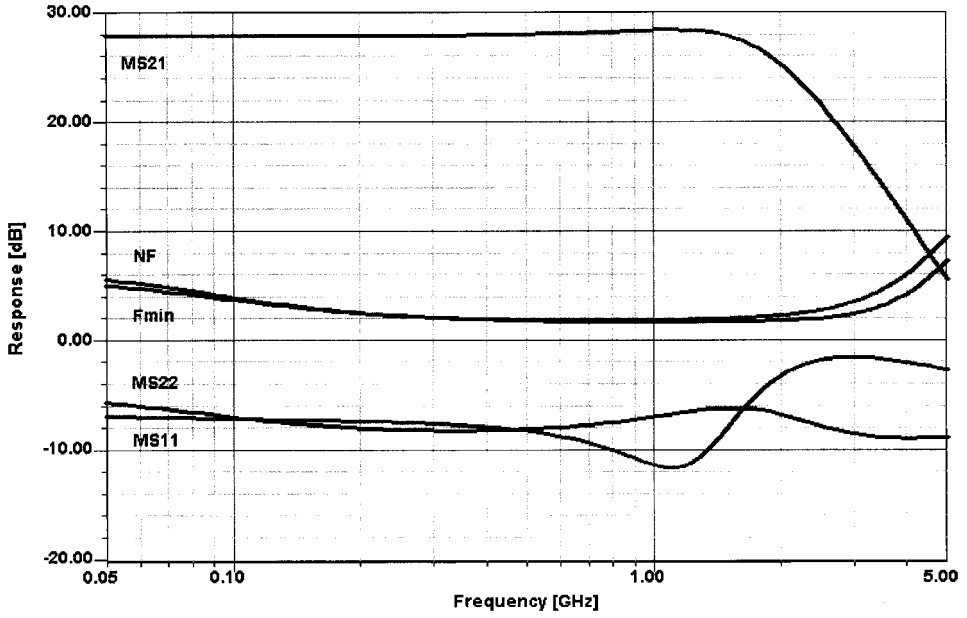


Figure 3-125 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-124.

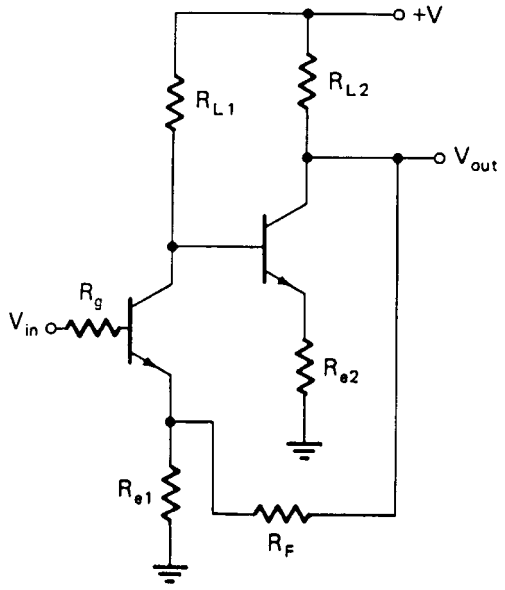


Figure 3-126 Two-stage amplifier with voltage feedback from Stage 2 to Stage 1, and current feedback in both stages.

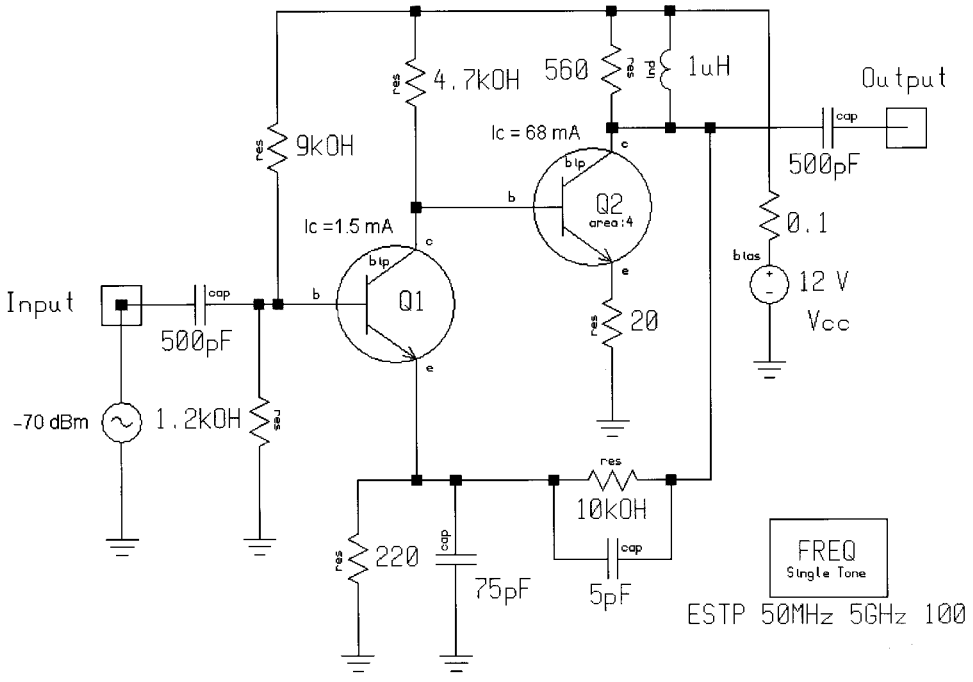


Figure 3-127 Simulated two-stage amplifier with feedback as shown in Figure 3-126.

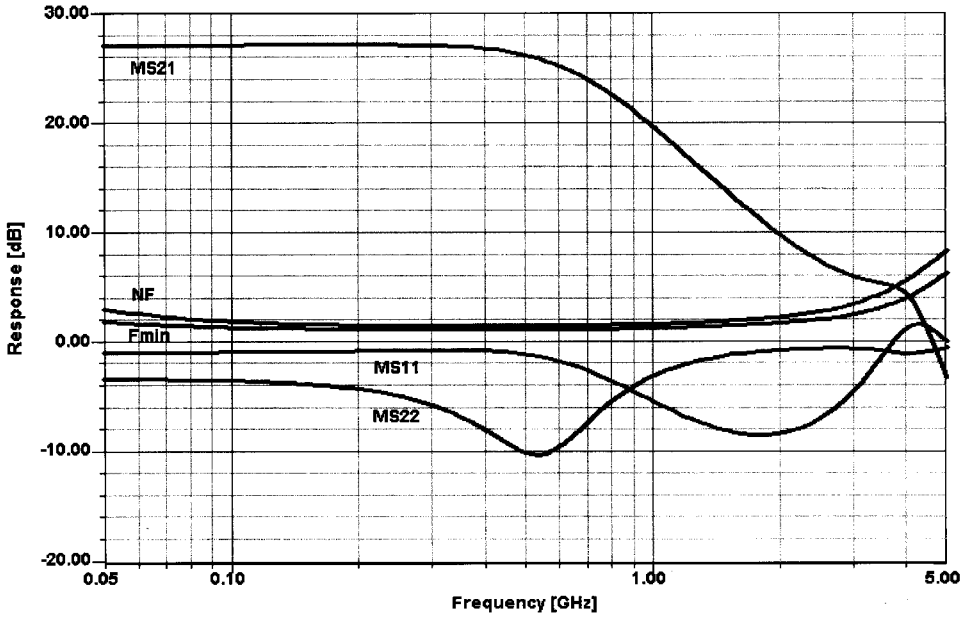


Figure 3-128 Gain, matching, and noise performance of the amplifier in Figure 3-127.

of an output transformer that absorbs the output capacitance. Figure 3-130 shows this circuit's frequency-dependent gain, matching, and noise performance.

It has been obvious to us that most published wireless application circuits are evaluated on a no-input-selectivity basis; that is, their reported performance does not take into account the filtering that is usually present in practical implementations. Evaluated in this way, a low-noise design may achieve a noise figure of less than 2 dB—1.2 dB or so is typical—but this is misleading because filters and connectors introduce insertion loss, resulting in a higher noise figure.

The cascode is a nice preamplifier combination. It combines (depending on the transistor) a low noise figure with essentially no feedback. Since the output of a grounded-base transistor provides a high-impedance source, this configuration is ideal to work into the SAW filters frequently used in wireless applications at the operating frequency. Special transistor pairs are available for use in cascode; as can be seen from this example, not every transistor combination is ideal. This one gives a higher noise figure than wanted and at frequencies above 2 GHz it tends to show possible instabilities. Both Siemens and Motorola are now offering transistors appropriate for these cascodes. Mostly, one would resort to transistors of less-exotic cutoff frequencies to guarantee stability. Typically, grounded-base/gate transistors tend to become oscillators, specifically because this configuration, like the emitter follower, allows a higher operating frequency while maintaining the same gain–bandwidth product; the actual gain for the last two circuit configurations is less. A grounded-base/gate stage has a high voltage gain and the emitter/source follower has power gain at low impedance levels. We invite our readers to experiment with the capacitor from the base of Q1 to ground in Figure 3-129 to get an interesting education about these transistors and their willingness to oscillate.

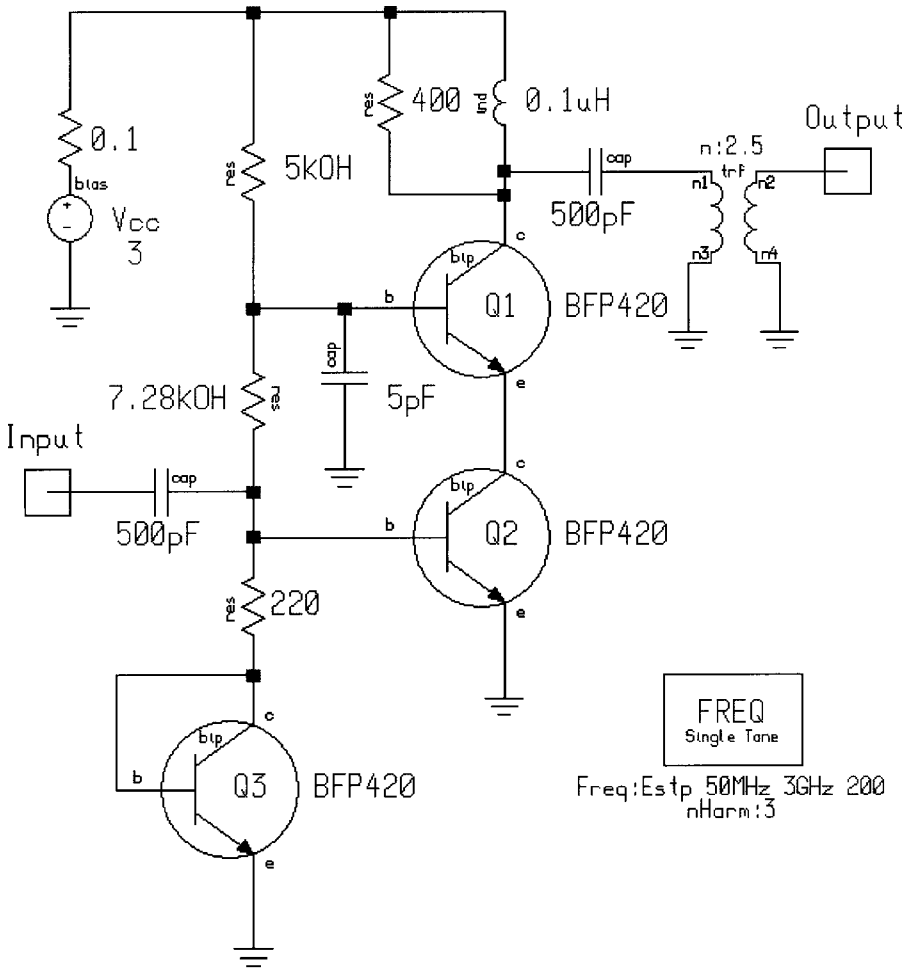


Figure 3-129 Schematic of the cascode amplifier.

The MC13144 is an example of a commercially available VHF to 2-GHz low-noise amplifier with programmable bias. As can be seen from Table 3-11, it is ideal for the 900-MHz range and its programmable gain/bias makes it very attractive. This design (Figure 3-131) is based on the cascode mentioned above and, because of high integration, uses many transistors for bias and level shifting.

The 900-MHz application circuit in Figure 3-132 shows the required external components for a successful design. Input and output are made tunable. The 5.6- and 39-nH inductances can be made from small air-wound coils. The output ceramic filter provides the necessary selectivity ahead of the mixer. With such a highly selective device in the input, the noise figure would deteriorate too much. The antenna is rarely in the vicinity of sufficiently strong transmitters to require such strong input filtering.

There are two more interesting applications for two-stage amplifiers: The one implemented in the Hewlett-Packard MSA-0735 MMIC has become a quite common one, and its usefulness up to higher frequencies depends mostly on the cutoff frequency of the transistors,

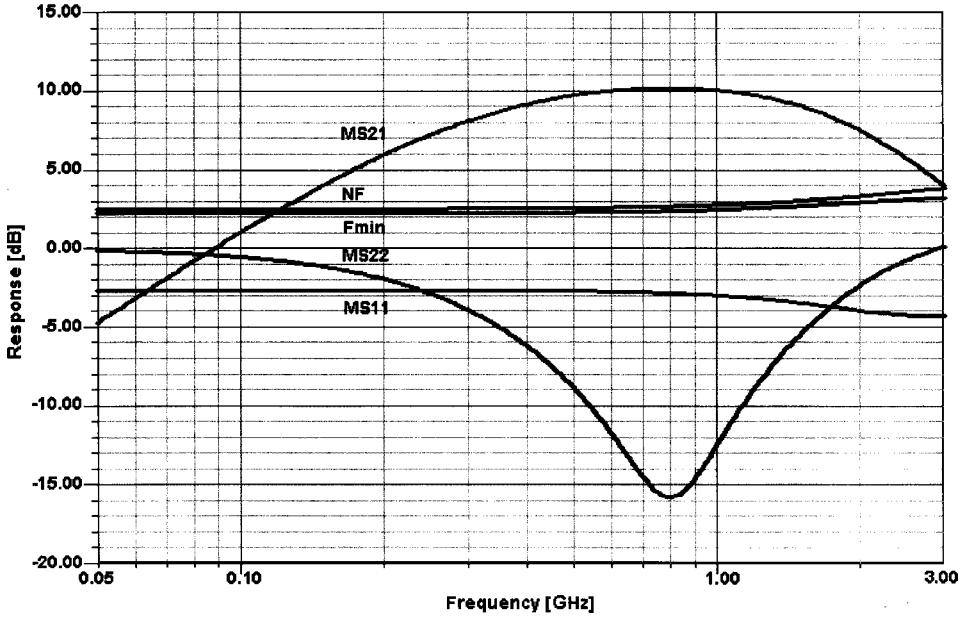


Figure 3-130 Frequency-dependent gain, matching, and noise performance of the cascode amplifier.

their parasitics, and the phase shift of the gain. The typical configuration is shown in Figure 3-133, and its frequency-dependent gain, match, and noise figure in Figure 3-134. There is an increase of S_{22} at higher frequencies; this indicates unwanted phase shift in the circuit. Some manufacturers hide the compensation circuits to overcome such peaking; most of the schematics shown by the manufacturers are really intended only to show operating principles, with proprietary details omitted.

Thanks to the properties of GaAs transistors, we can extend the frequency range significantly. The amplifier we are about to look at goes back several years and is described in great detail in Goyal [27]. Essentially what we have, seen in Figure 3-135, is an FET version of Figure 3-127. The main goal of this circuit was to have an all-monolithic circuit, avoiding all ac coupling and accomplishing all matching and interaction with dc coupling. The transistor Q1 is the main gain stage, which has the transistor Q2 as an active load. The output impedance of Q2 is $1/g_m$—typically in the area of 50–200 Ω, depending on the biasing,

Table 3-11 MC13144 features

Gain at 900 MHz, 17 dB
Noise figure at 900 MHz, 1.4 dB
P_{-1dB}, −7.0 dBm
$IP_{3,in}$, −5.0 dBm
Low operating supply voltage (1.8–6.0 Vdc)
Programmable bias with Enable 1 and Enable 2. [Enable 1 and Enable 2 programmed high for optimal noise figure and gain associated with NF. Enable can be overridden for external bias programming (up to 15 mA.)]

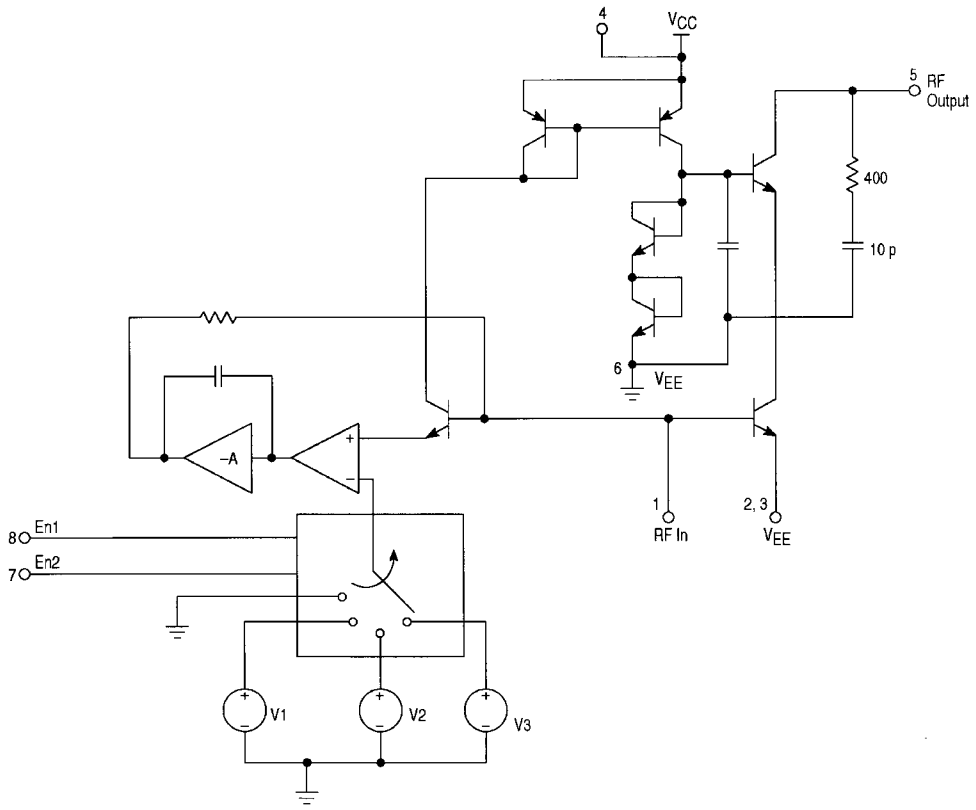


Figure 3-131 Schematic of the Motorola MC13144 cascode LNA.

which affects the transconductance. Transistor Q3 has the same function as an emitter/source follower; again, its dc bias determines its transconductance and therefore its output impedance. (Given this, a device that has a transconductance of 20 mS, or close to it, would establish a very good output match to 50 Ω.) The drains of Q1 and Q5 are at approximately 3.5 V. The diodes at the source of Q3 are level shifters that must shift the dc voltage at the gate of Q5 and at the drain of Q4 to approximately −1.5 V relative to ground—the necessary bias condition for Q1 and Q5. The standard shunt feedback circuit would show resistive feedback between the drain of Q1 (gate of Q2) and the gate of Q1; in the previous examples, we calculated the magic value of this part to be about 200 Ω. Adding in parallel to Q1 the transistor Q5 allows the designer to use a feedback scheme that isolates the feedback loop from the input. Dc-wise, the two transistors are tied together via a 10-kΩ resistor, which can be included in the actual packaged device. Insufficient information was published to allow us to duplicate the design, considering all the intricacies involved (we somewhat object to the fact that even in the *IEEE MTT-S Transactions*, so little information is given about the actual implementation of this design—using space constraints as an excuse—that its educational value is usually minimal); it shows a new principle without revealing the design steps necessary to obtain a practical circuit. Several variations of this type of circuit are available; the literature is full of discussions about them. The magnitude of the feedback is set by the ratio of the width of Q5 to that of Q1. Typical values for the Q5/Q1 ratio range

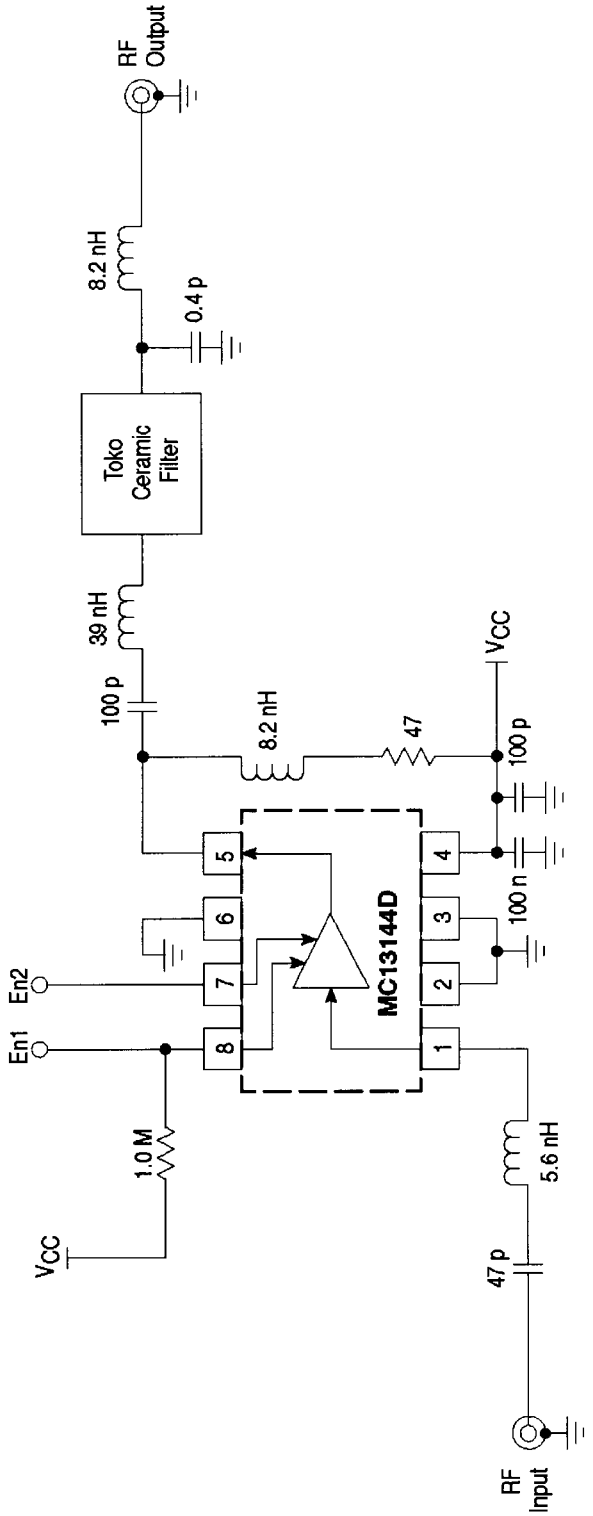


Figure 3-132 An MC13144 application circuit for 926.5 MHz.

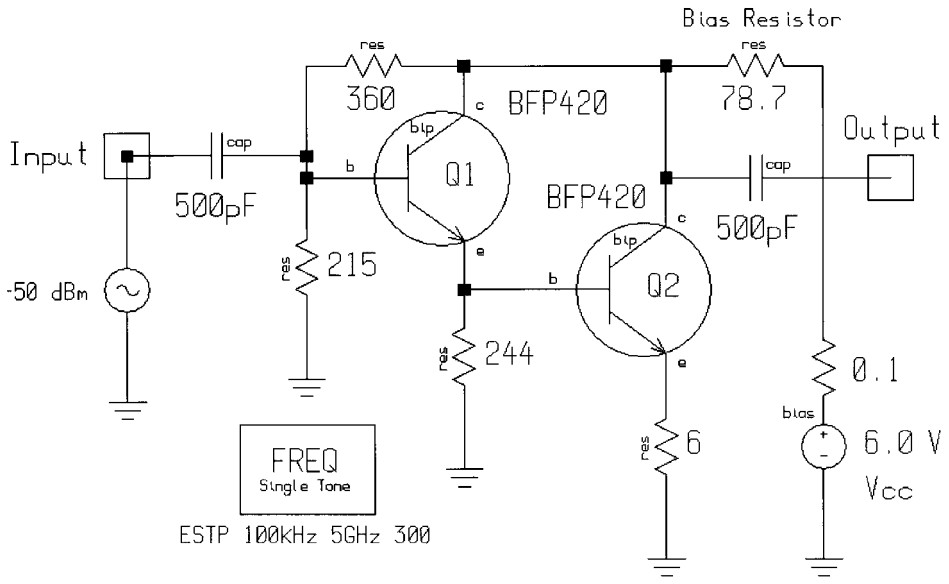


Figure 3-133 Schematic of the MSA-0375 MMIC amplifier.

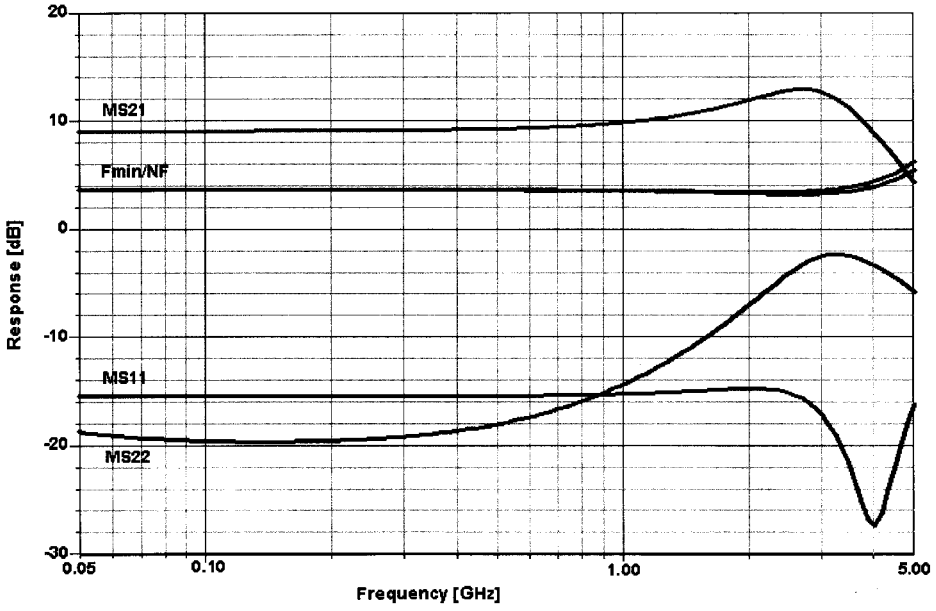


Figure 3-134 Frequency-dependent gain, matching, and noise performance of the MSA-0735 MMIC amplifier.

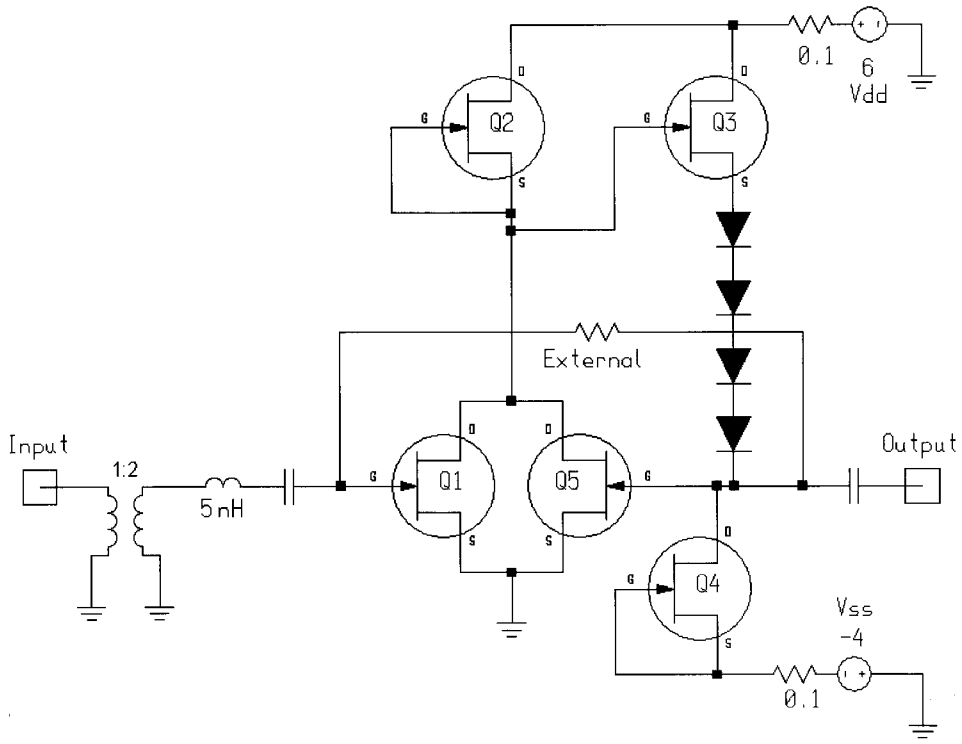


Figure 3-135 Schematic of the dc-coupled GaAsFET amplifier.

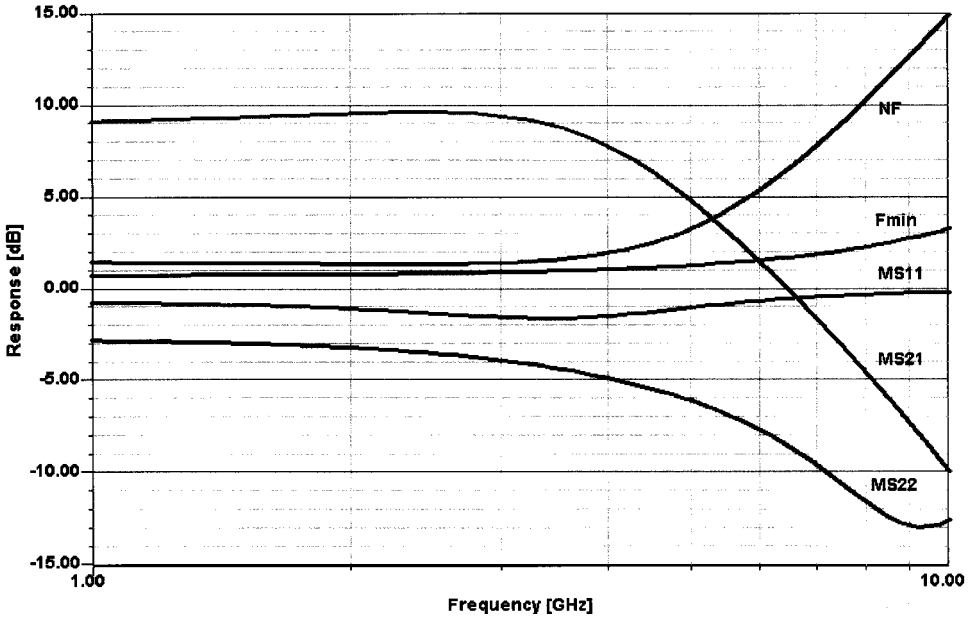


Figure 3-136 Frequency-dependent gain, matching, and noise performance of the dc-coupled GaAsFET amplifier.

from 0.15 to 0.30. In simulation, sizing the devices can best be accomplished by using the scaling factor for the FET model.

One of our questions was: "What would the perfect input termination for Q1 be?" By using an ideal transformer and playing with its turns ratio (needless to say, we did this with CAD), we determined experimentally that, relative to the gate, 200 Ω and a series inductor of 5–8 nH—the sort of details that designers often avoid telling their audiences—gave the ideal frequency response. Overall, we believe that the inventor of this circuit can be quite proud of the overall performance. The FETs we chose were taken from the Texas Instruments nonlinear foundry library; however, any similar device from another company would have worked equally well. The feedback loop used in this IC is an example of an approach frequently referred to as *active feedback*. The performance of this amplifier can be evaluated from Figure 3-136, which shows the frequency-dependent gain, matching, and noise figures (even including F_{min}).

3-5 AMPLIFIERS WITH THREE OR MORE STAGES

Most three-stage amplifiers are obtained by adding another buffer or power-gain stage to the previously described principal designs. As an example, we decided to evaluate the recent NEC UPC2749 IC, a silicon MMIC intended for application as a low-noise 1900-MHz amplifier operating at 3 V. Table 3-12 briefly summarizes its electrical specifications.

Table 3-12 NEC UPC2749 electrical characteristics

(T_A = 25 °C, $Z_L = Z_S$ = 50 Ω, V_{CC} = 3.0 V)			
	Part Number Package Outline		UPC2749T TO6
Symbols	Parameters and Conditions	Units	Typical
I_{CC}	Circuit current (no signal)	mA	6
G_S	Small-signal gain, f = 900 MHz dB	14.5	
	f = 1900 MHz	dB	16
f_{corner}	–3-dB Gain corner frequency	GHz	2.9
P_{1dB}	1 dB Compressed output power at 1900 MHz	dBm	–12.5
P_{SAT}	Saturated output power, f = 1900 MHz	dBm	–6
NF	Noise figure, f = 900 MHz	dB	3.2
	f = 1900 MHz	dB	4.0
RL_{IN}	Input return loss, f = 1900 MHz	dB	10
RL_{OUT}	Output return loss, f = 1900 MHz	dB	12.5
ISOL	Isolation, f = 1900 MHz	dB	30
OIP_3	SSB output third-order intercept, f_1 = 500 MHz, f_2 = 510 MHz	dBm	–3
	f_1 = 1000 MHz, f_2 = 1010 MHz	dBm	–3
	f_1 = 1900 MHz, f_2 = 1902 MHz	dBm	–3.5
	f_1 = 2000 MHz, f_2 = 2010 MHz	dBm	–4
$R_{TH(J-A)}$	Thermal resistance (junction to ambient)		
	Free air	°C/W	
	Mounted on a 50 × 50 × 1.6 mm epoxy glass PWB	°C/W	

In modeling the UPC2749's three transistors, we chose to use our favorite, the BFP420, which is not without its headaches at the higher frequencies. Figure 3-137 shows the circuit diagram, which closely resembles the one published by NEC; again, the only hard facts we had were the IC's basic circuit topology, gain, and dc current. Given those numbers, we staggered the current in the three devices by setting them at 0.64 mA for the first stage, 1.85 mA for the second stage, and 3 mA for the third stage. It should be noted that the two output transistors are not unlike the approach used for the Hewlett-Packard MSA-0735 (Figure 3-133). Taking (hopefully) all the right assumptions, we end up with a design that closely resembles the manufacturer's specifications for both power consumption, gain, and noise. This can be verified by looking at the results plotted in Figure 3-138.

The final example is a more complex derivative of the Philips NE/SA5204A amplifier, which was intended for standard 6-V operation and has a fairly "normal" circuit arrangement. This can be seen from Figure 3-139. A much more elaborate cousin of this is the low-voltage NEC UPC2710 amplifier. According to its specifications, it operates at 5 V but has significantly more gain than the NE/SA5204. Table 3-13 shows its electrical specifications and Figure 3-140 shows its measured gain as published by the manufacturer.

The actual interior of the circuit is shown in Figure 3-141. It consists of the standard preamplifier and then a Darlington configuration with the collectors tied together. The middle section of the circuit acts as a power supply responsible for the bias relative to the emitter of the first transistor influencing the rest of the circuit as well. Since NEC mentioned that they use their NESAT III process (f_T of 20 GHz at V_{ce} = 3) in producing the UPC2710T and we

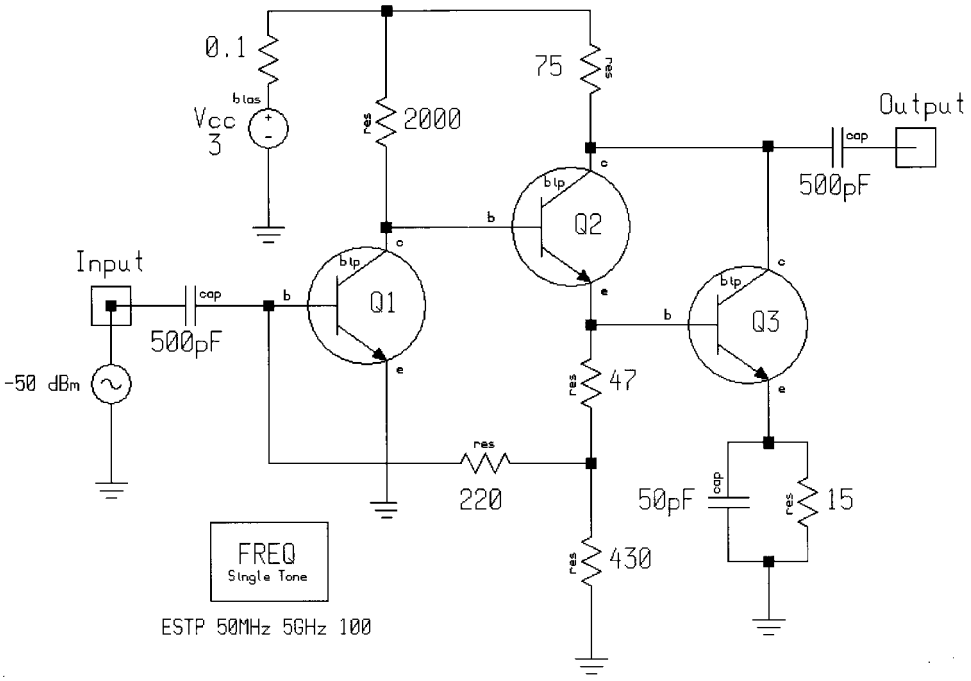


Figure 3-137 Circuit of the UPC2749 MMIC. The component values were not supplied by the manufacturer.

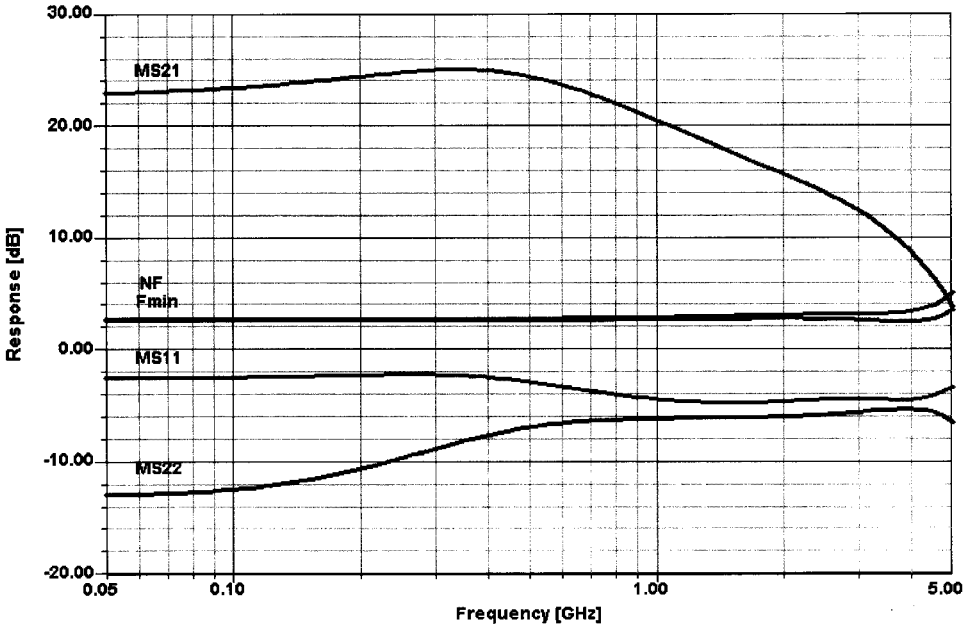


Figure 3-138 Simulated gain, matching, and noise performance of the NEC UP2749 MMIC.

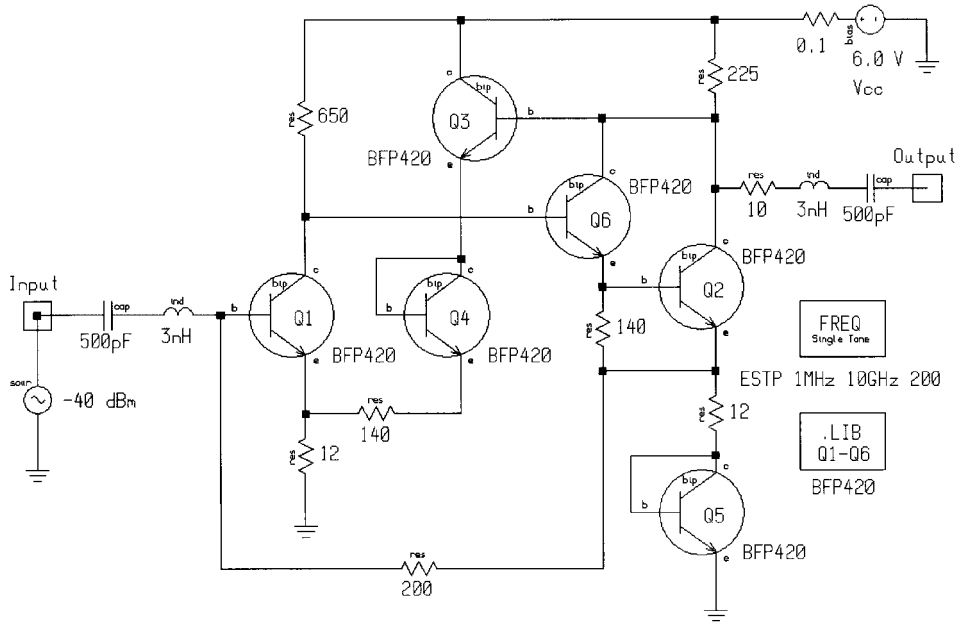


Figure 3-139 Schematic of the Philips NE/SA5204A amplifier IC entered for simulation.

Table 3-13 UPC2710 electrical specifications

	T_A = 25 °C, f = 0.5 GHz, V_{CC} = 5 V)		
	Part Number Package Outline		UPC2710T TO6
Symbols	Parameters and Conditions	Units	Typical
I_{CC}	Circuit current (no signal)	mA	22
G_S	Small-signal gain	dB	33
f_{corner}	–3-dB Gain corner frequency	GHz	1.0
ΔG_S	Gain flatness, f = 0.1–0.6 GHz f = 0.1–0.8 GHz	dB	±0.8
P_{SAT}	Saturated output power	dBm	13.5
P_{1dB}	Output power at 1dB compression point	dBm	7.5
NF	Noise figure	dB	3.5
RL_{IN}	Input return loss	dB	6
RL_{OUT}	Output return loss	dB	12
ISOL	Isolation	dB	39
ΔG_T	Gain–temperature coefficient	dB/°C	–0.006
R_{TH}	Thermal resistance (junction to ambient)	°C/W	

obtained some information regarding it, we decided to use our standard BFP420 (f_T of 25 GHz; see the datasheet in Chapter 2) for the purpose of simulation.

In our effort to simulate this circuit, we chose a combination of values that pretty much complies with the manufacturer's published specifications, but S_{22} is not close, as shown in Figure 3-142. We stopped at this point and invite our readers with CAD capabilities to provide the final touch in meeting all the published specifications for the circuit. We are particularly

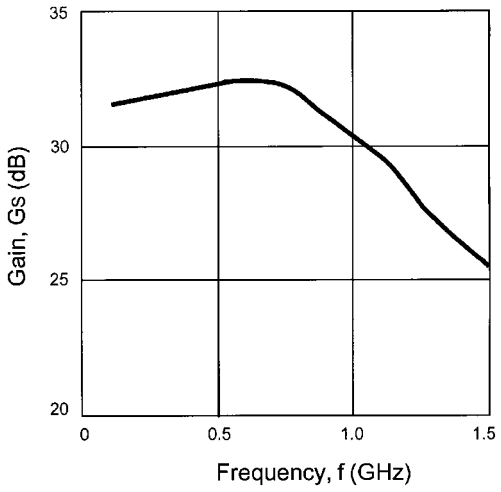


Figure 3-140 Manufacturer-supplied gain versus frequency for the NEC UPC2710T MMIC.

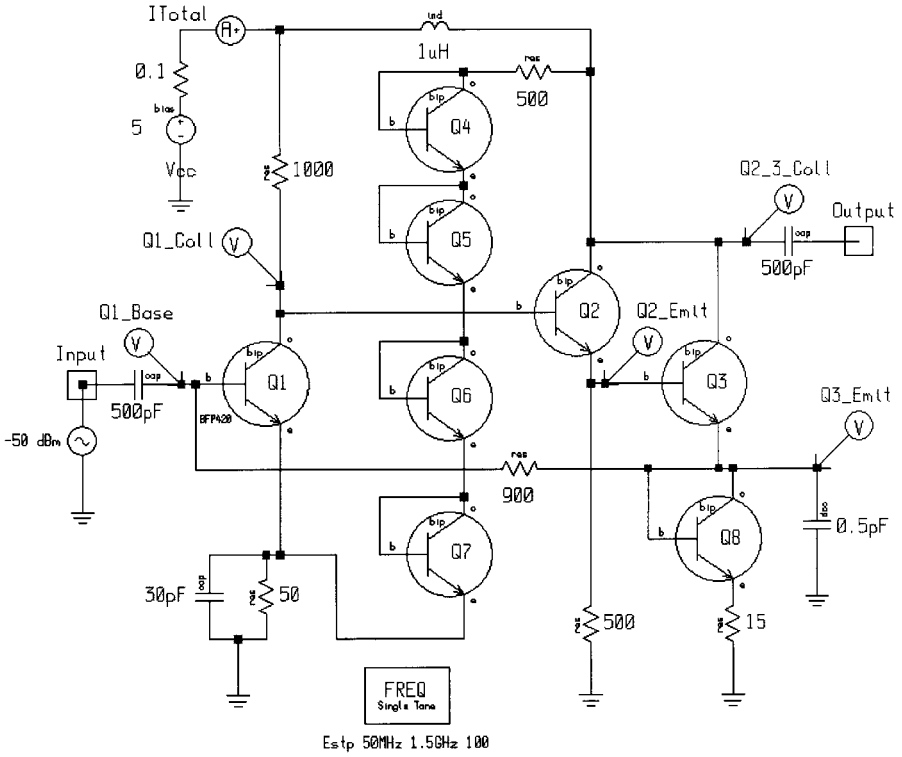


Figure 3-141 Schematic of the UPC2710T silicon MMIC entered for simulation.

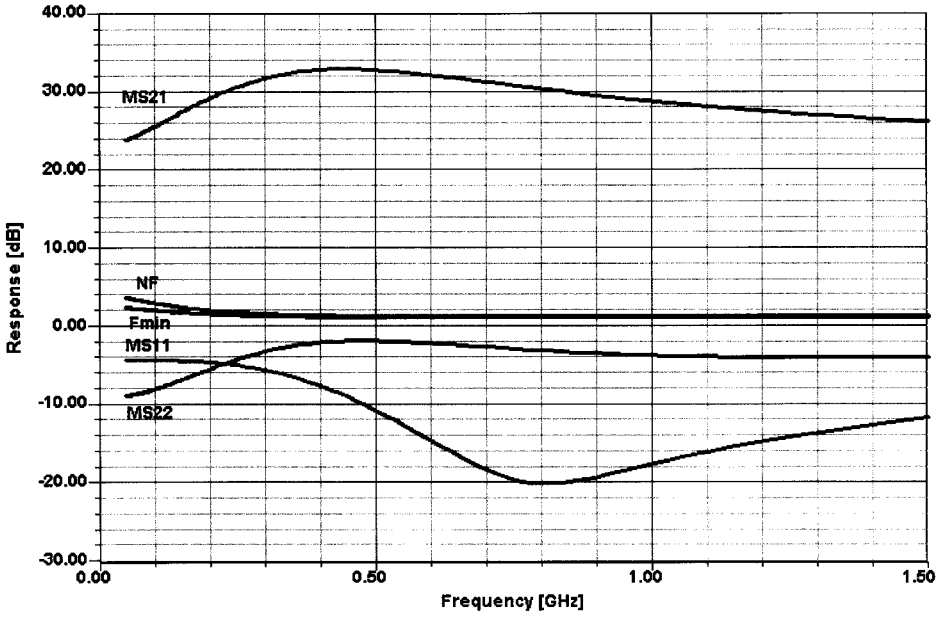


Figure 3-142 Simulated gain, match, and noise performance of the UPC2710T.

recommending such an effort because it shows very nicely the interaction of all the circuit elements in a fairly drastic way. Because of this, we admire the IC design even more, since the manufacturer has to sell a pretty consistent amplifier as a function of different production runs.

3-5-1 Stability of Multistage Amplifiers

Stability, needless to say, is a big issue, and the delay of the various stages adds to the phase shift, resulting in possible instabilities. The fact that S_{11} and/or S_{22} can become larger than 0 (in dB) means there is the potential of oscillating, depending on the reactances available. If a cable of inappropriate length is added, hell will break loose. Figure 3-143 shows the dangerous peaking of our simulation of the Philips NE5204A, in which the internal compensation was intentionally not assumed correctly. We will all agree that the high-end performance of this amplifier can generate high anxiety. As the production process of the aging device becomes more modern—meaning that transistors with much higher f_T will be the basis for the production—it will become quite difficult to maintain a previously achieved stability. This, by the way, applies also to discrete transistors, which for reasons of economics and multiple manufacturers will be improved over time. Such evolution results in generally smaller base-spreading resistances and capacitances, and other parameter changes, which, overall, cause headaches in manufacturing. We will leave the topic of three-stage amplifiers on this note.

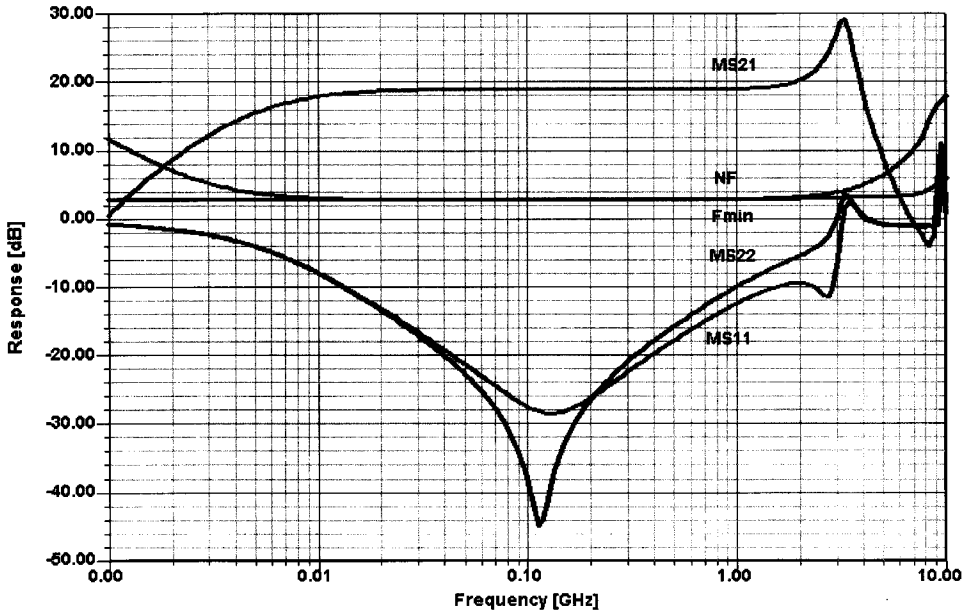


Figure 3-143 Simulated gain, match, and noise performance of the NE5204A IC.

3-6 A NOVEL APPROACH TO VOLTAGE-CONTROLLED TUNED FILTERS INCLUDING CAD VALIDATION [28]

Modern receivers control input stages as well as the oscillator band and frequency by electrical rather than mechanical means. Tuning is accomplished by voltage-sensitive capacitors (varactor diodes), and band switching by diodes with low forward conductance. Since the wireless band (essentially 400 MHz to 2.4 GHz) is so full of strong signals, the use of a tracking filter is desired as a solution to improve the performance and prevent second-order IMD products or other undesired overload effects. The dc control voltage needed for the filter can easily be derived from the VCO control voltage. There may be a small dc offset, depending on the IF used.

3-6-1 Diode Performance

The capacitance versus voltage curves of a varactor diode depend on the variation of the impurity density with the distance from the junction. When the distribution is constant, there is an "abrupt junction" and capacitance follows the law

$$C = \frac{K}{(V_d + V)^{1/2}} \tag{3-183}$$

where V_d is the contact potential of the diode and V is applied voltage.

Such a junction is well approximated by an alloyed junction diode. Other impurity distribution profiles give rise to other variations, and the above equation is usually modified to

$$C = \frac{K}{(V_d + V)^n} \tag{3-184}$$

where n depends on the diffusion profile and $C_0 = K/V_d^n$.

A so-called graded junction, having a linear decrease in impurity density with the distance from the junction, has a value of n. This is approximated in a diffused junction.

In all cases these are theoretical equations, and limitations on the control of the impurity pattern can result in a curve that does not have such a simple expression. In this case, the coefficient n is thought of as varying with voltage. If the impurity density increases away from the junction, a value of n higher than 0.5 can be obtained. Such junctions are called hyperabrupt. A typical n value for a hyperabrupt junction is about 0.75. Such capacitors are used primarily to achieve a large tuning range for a given voltage change. Figure 3-144 shows the capacitance–voltage variation for the abrupt and graded junctions as well as for a particular hyperabrupt-junction diode. Varactor diodes are available from a number of manufacturers, such as Motorola, Siemens, and Philips. Maximum values range from a few to several hundred picofarads, and useful capacitance ratios range from about 5 to 15.

Figure 3-145 shows three typical circuits that are used with varactor tuning diodes. In all cases, the voltage is applied through a large resistor R_e or, better, an RF choke in series with a small resistor. The resistance is shunted across the lower diode and may be converted to a shunt load resistor across the inductance to estimate Q. The diode also has losses that may result in lowering the circuit Q at high capacitance, when the frequency is sufficiently high. This must be considered in the circuit design.

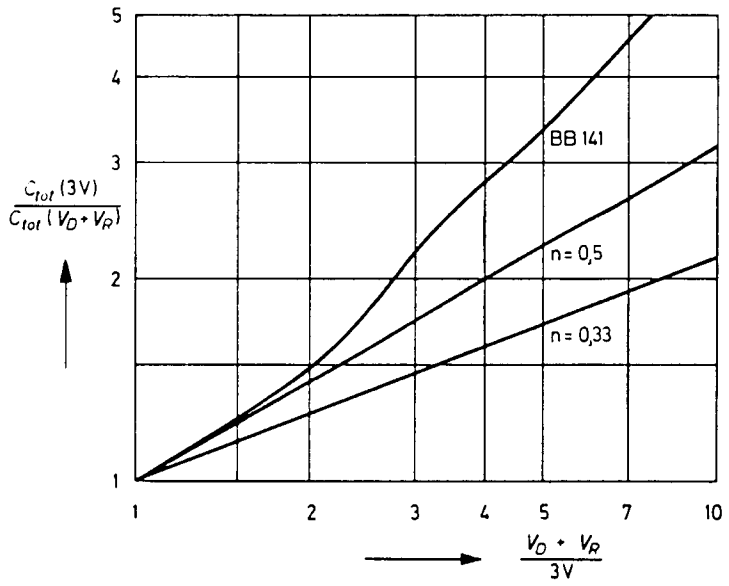


Figure 3-144 Voltage-dependent change of capacitance of different types of diodes. The BB141 is a hyperabrupt-junction diode with $n = 0.75$.

The frequency-dependent performance is not determined solely by applying the dc tuning voltage to Eq. (3-184). If the RF voltage is sufficient to drive the diode into conduction on peaks, an average current will flow in the circuits of Figure 3-145, which will increase the bias voltage. The current is impulsive, giving rise to various harmonics of the circuit. Even in the absence of conduction, Eq. (3-184) deals only with the small-signal capacitance. When the RF voltage varies over a relatively large range, the capacitance changes. In this case, Eq. (3-184) must be changed to

$$\frac{dQ}{dV} = \frac{K}{(V + V_d)^n} \tag{3-185}$$

Here Q is the charge on the capacitor. When this relation is substituted in the circuit differential equation, it produces a nonlinear differential equation, dependent on the parameter n. Thus, the varactor my generate direct current and harmonics of the fundamental frequency. Unless the diodes are driven into conduction at some point in the cycle, the direct current must remain zero.

The current of Figure 3-145c can be shown to eliminate the even harmonics, and it permits a substantially larger RF voltage without conduction than either circuit in Figure 3-145a or 3-145b. When $n = 0.5$, only the second harmonic is generated by the capacitor, and this can be eliminated by the back-to-back connection of the diode pair. Integrating Eq. (3-185), we find

$$Q + Q_A = \frac{K}{1 - n}(V + V_d)^{1-n}$$

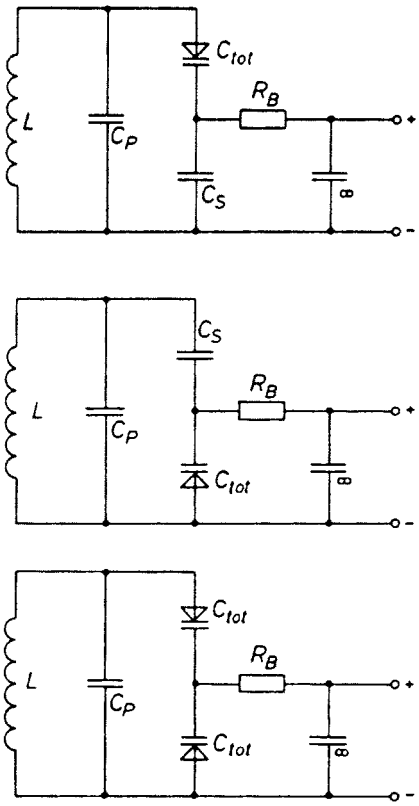


Figure 3-145 Various configurations to apply tuning diodes in a tuned circuit. The bottom version shows the lowest distortion.

$$= \frac{C_v}{1-n}(V + V_d) \tag{3-186}$$

C_v is the value of Eq. (3-184) for applied voltage V, and Q_A is a constant of integration. By letting $V = V_1 + v$ and $Q = Q_1 + q$, where the lowercase letters represent the varying RF and the uppercase letters indicate the values of bias when RF is absent, it follows that

$$q + Q_1 + Q_A = \frac{K}{1-n}[v + (V_1 + V_d)]^{1-n} \tag{3-187}$$

$$1 + \frac{v}{V'} = \left(1 + \frac{q}{Q'}\right)^{1/(1-n)} \tag{3-188}$$

where $V' = V_1 + V_d$ and $Q' = Q_1 + Q_A$. For the back-to-back connection of identical diodes, $K_{11} = K_{12} = K_1$, $V_1' = V_2' = V'$, $Q_1' = Q_2' = Q'$, $q = q_1 = -q_2$, and $v = v_1 - v_2$. Here the new subscripts 1 and 2 refer to the top and bottom diodes, respectively; v is the RF voltage across

the pair in series, and q is the charge transferred through the pair in series. This notation results in

$$\frac{v}{V'} \equiv \frac{v_1 - v_2}{V'} = \left(1 - \frac{q}{Q'}\right)^{1/(1-n)} - \left(1 - \frac{q}{Q'}\right)^{1/(1-n)} \tag{3-189}$$

For all n, this eliminates the even powers of q, and hence the even harmonics. This can be shown by expanding Eq. (3-189) in a series expansion and performing term by term combination of the equal powers of q. In the particular case $n = \frac{1}{2}$, $v/V' = 4q/Q'$, and the circuit becomes linear.

The equations hold as long as the absolute value of v_1/V' is less than unity, so that there is no conduction. At the point of conduction, the total value of v/V' may be calculated by noting that when $v_1/V' = 1$, $q/Q' = -1$, so $q_2/Q' = 1$, $v_2/V = 3$, and $v/V = -4$. The single-diode circuits conduct at $v/V' = -1$, so the peak RF voltage should not exceed this. The back-to-back configuration can provide a fourfold increase in RF voltage handling over the single diode. For all values of n, the back-to-back configuration allows an increase in the peak-to-peak voltage without conduction. For some hyperabrupt values of n, such that $1/(1 - n)$ is an integer, many of the higher-order odd harmonics are eliminated, although only $n = \frac{1}{2}$ provides elimination of the third harmonic. For example, $n = \frac{2}{3}$ results in $1/(1 - n) = 3$. The fifth harmonic and higher odd harmonics are eliminated, and the peak-to-peak RF without conduction is increased eightfold; for $n = \frac{3}{4}$ the harmonics 7 and above are eliminated, and the RF peak is increased 16 times. It must be noted in these cases that the RF peak at the fundamental may not increase so much, since the RF voltage includes the harmonic voltages.

Since the equations are only approximate, not all harmonics are eliminated, and the RF voltage at conduction, for the back-to-back circuit, may be different from that predicted. For example, abrupt-junction diodes tend to have n of about 0.46–0.48 rather than exactly 0.5. Hyperabrupt junctions tend to have substantial changes in n with voltage. The diode illustrated in Figure 3-144 shows a variation from about 0.6 at low bias to about 0.9 at higher voltages, with wiggles from 0.67 to 1.1 in the midrange. The value of V_d for varactor diodes tends to be in the vicinity of 0.7 V.

3-6-2 A VHF Example

The application of tuning diodes in double-tuned circuits in TV tuners has been common practice for many years. Figure 3-146 shows the circuit diagram. The input impedance of 50 Ω gets transformed up to 10 kΩ. The tuned circuits consist of the 0.3-μH inductor and two sets of antiparallel diodes. By dividing the RF current in the tuned circuit and using several diodes instead of just one pair, intermodulation distortion is reduced.

The coupling between the two tuned circuits is tuned via the 6-nH inductor that is common to both circuits. This type of inductance is usually printed on the circuit board. The diode parameters used for this application were equivalent to the Siemens BB515 diode. The frequency response of this circuit is shown in Figure 3-147.

The coupling is less than critical. This results in an insertion loss of about 2 dB and relatively steep passband sides. Once the circuit's large-signal performance (Figure 3-148) is seen, a third-order intercept point of about –2 dBm is not so unexpected. The reason for this poor performance is the high impedance (high L/C ratio), which provides a large RF voltage swing across the diodes. A better approach appears to be using even more diodes and at the same time changing the impedance ratio (L/C ratio).

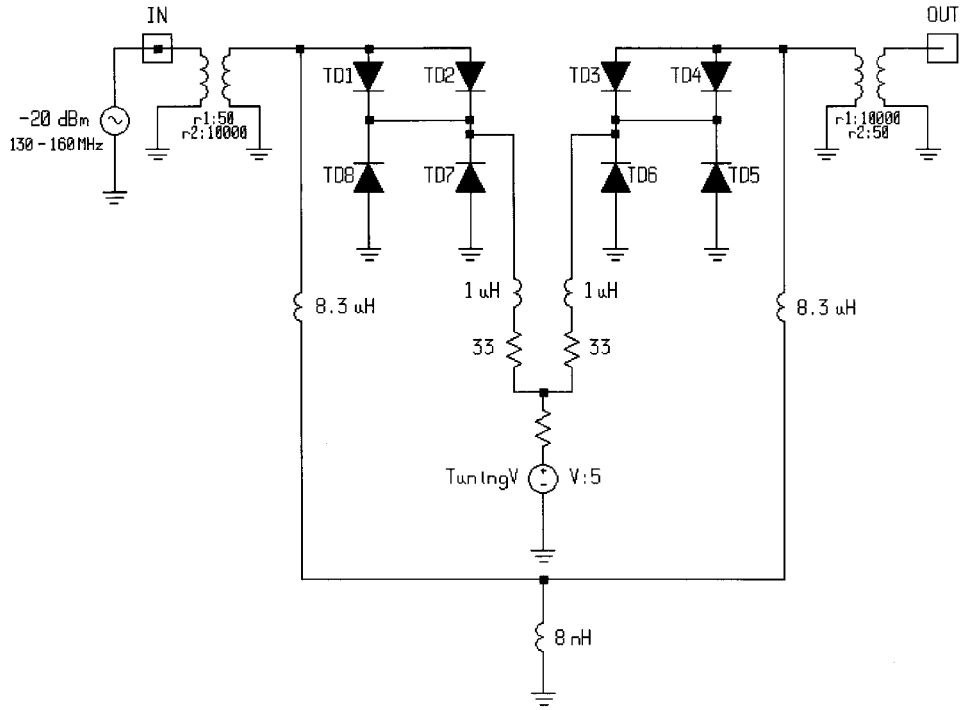


Figure 3-146 Double-tuned filter at 161 MHz using hyperabrupt-junction tuning diodes. By using several parallel diodes, the IMD performance improves.

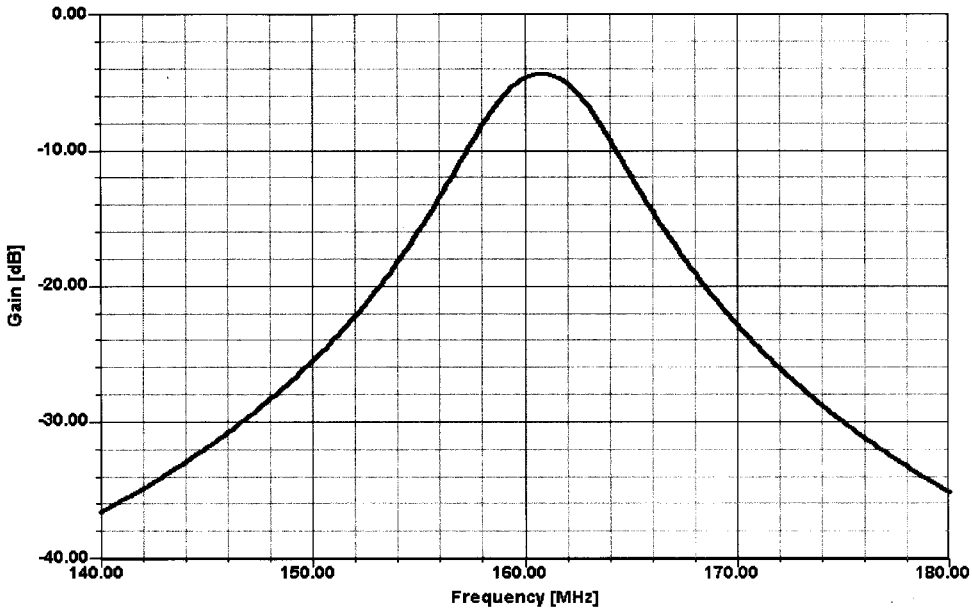


Figure 3-147 Frequency response of the tuned filter shown in Figure 3-146. The circuit is undercoupled (less than transitional coupling); $Q \times k < 1$.

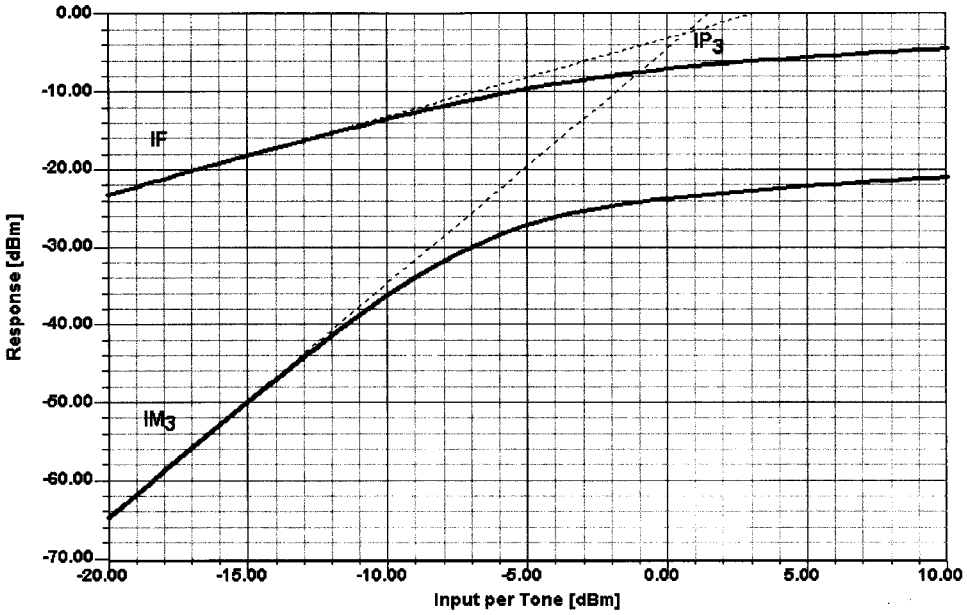


Figure 3-148 Prediction of intercept point of the double-tuned circuit shown in Figure 3-146. Note the compression of both the input signal and the IMD product.

3-6-3 An HF/VHF Voltage-Controlled Filter

The above example used a step-up procedure typically done by a tap at the input and output inductance. This method allows an impedance transformation; however, if one desires to change it into a series-tuned arrangement, it has to be done with a transformer. The large-signal conditions in the frequency range from 10 to 30 MHz on a medium- to large-sized antenna are equivalent to, if not worse than, the conditions for VHF operation. The only exception would be line of sight into the transmitter such as a tower in the middle of the city. Examples of such hostile conditions would be any large city, such as Munich, New York, Miami, Chicago, and San Francisco, where the authors had significant experience with intermodulation distortion problems.

By translating the circuit into a low-impedance-drive arrangement, building it symmetrically, and reducing the tuning range somewhat by capacitors and many series diode pairs, the filter's large-signal performance was significantly improved. The tuning diodes had about 125 pF at 1 Vdc per unit. Initial tests trying to use a high-capacitance diode, such as the BB112, resulted in much higher IMD products. The B112 has a 1-V capacitance of 470 pF down to about 20 pF at 8 V. Standard diodes show about 30 pF at 1 V, while the diodes used in the experiment had a 125-pF capacitance at 1 V. These experimental diodes were made available from a well-known manufacturer for this evaluation. Figure 3-149 shows the circuit diagram.

A step-down transformer drives the tuned circuit with a source impedance of 12.5 Ω. The circuit is symmetrical. There are two 0.8-μH inductors and 2×5 diodes in the loop. The output of the circuit is transformed back up to 50 Ω. The tuning voltage is supplied via a heavily filtered arrangement. This circuit has actually been implemented in Rohde & Schwarz field strength meter equipment. They guarantee an intercept point of about +20 dBm. The selectivity of these circuits is quite reasonable and mostly intended to reduce

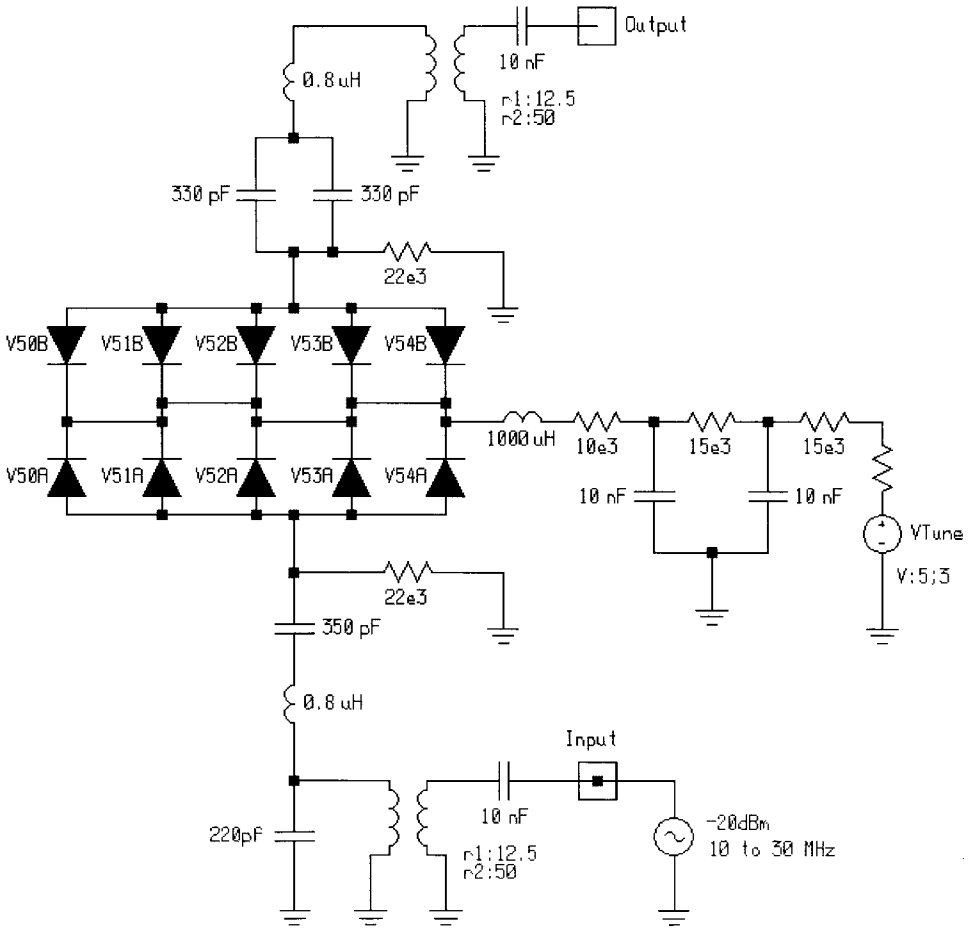


Figure 3-149 High dynamic range of a 10–20-MHz filter.

second-order intermodulation distortion products by about 10 dB, at the same time having a high third-order intercept point. Figure 3-150 shows the selectivity curves at two different tuning voltages.

The most interesting number, however, is the third-order intercept point (already determined as about 20 dBm), which still had to be simulated. This sometimes sounds like a contradiction: Once the measured values are available and are acceptable, why would one want—or, rather, need—to do a simulation besides the necessity to develop a high-input intercept filter? It was desirable to validate the nonlinear models and prove that the above-mentioned equations will hold true. This type of simulation is now more difficult because the number of nonlinear elements went from four to ten, and some numerical problems, such as convergence difficulties, can be expected, and the effect of harmonic frequency cancellation (compensation) can also be seen. By using diode combinations that result in $1(1 - n)$, n being an integer number, these IMD products can drastically be reduced. This implies that each of the five diode pairs has a selected value for n to meet this condition. Later, a final attempt will be made to improve the first VHF filter with the proper diode combinations.

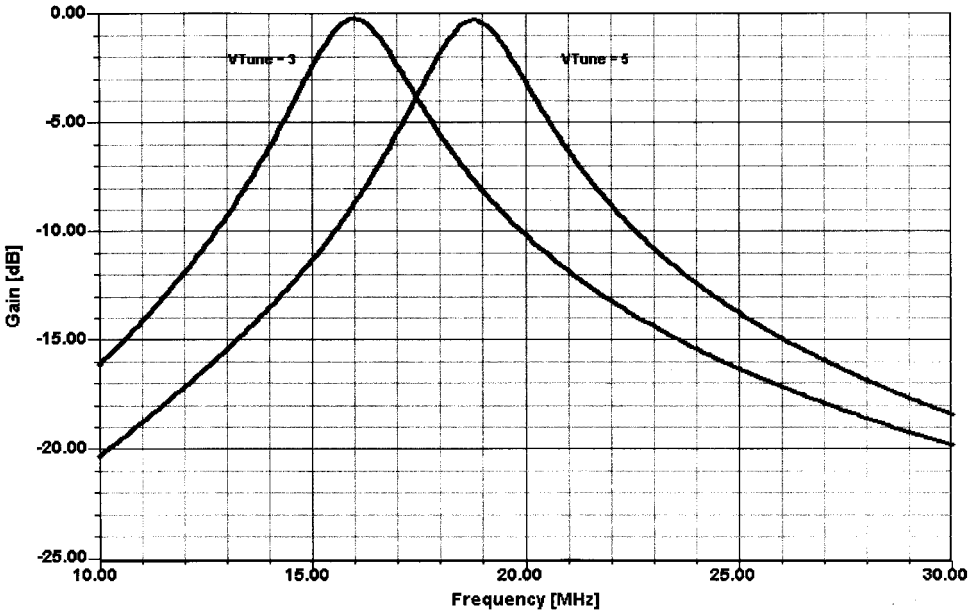


Figure 3-150 Frequency response of the filter shown in Figure 3-149. Filters of this type are intended to reduce second-order IMD by providing about 20-dB suppression at half the center frequency.

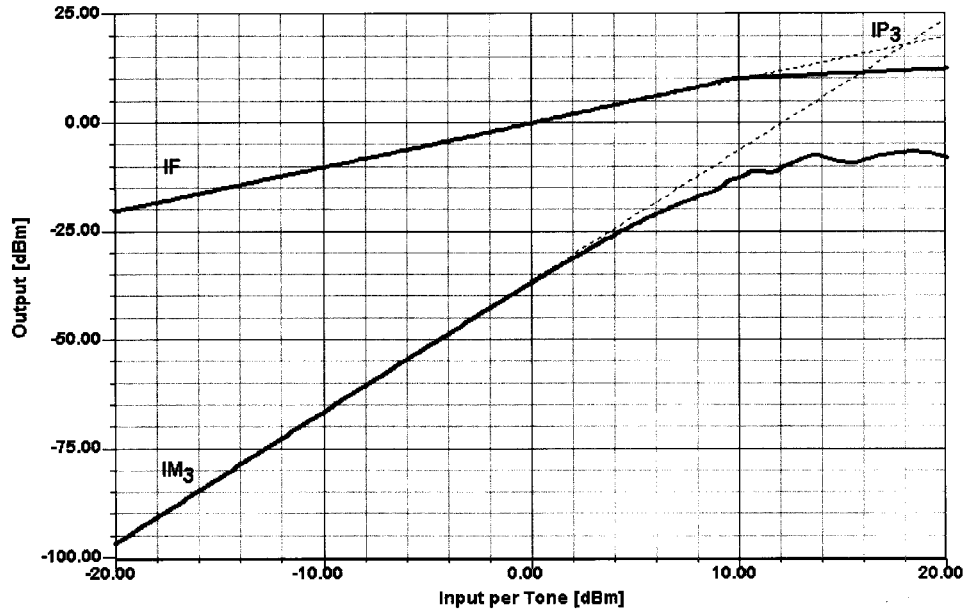


Figure 3-151 Prediction of third-order intercept point of the filter shown in Figure 3-149. The reason for the curves on the IMD plot lies in the interaction between the ten nonlinear devices.

Figure 3-151 shows the calculation of third-order intercept point for the arrangement shown in Figure 3-149. This number has now increased to +18 dBm. The difference between selection and measurement is approximately 2 dB. The reason for this more pessimistic value compared to the measured value probably has to do with slight variations of the exponent of the diodes' voltage-dependent capacitance. Therefore, the authors consider both the circuit performance and the simulation accuracy to be extremely good. This type of circuit, as mentioned earlier, has wide application in oscillator circuits. Danzeisen of Rohde & Schwarz was probably the first to have used this type of circuit by paralleling many diodes for improved performance.

3-6-4 Improving the VHF Filter

By selecting the appropriate diode combinations and circuit modifications as shown in Figure 3-152, a significant IMD improvement of the circuit shown in Figure 3-146 is obtained. A special shunt arrangement of several diodes with different exponents has been developed, which allows its value to be "adjusted." This circuit, for which a patent has been applied, showed an improvement from –2 dBm to 32 dBm. The same high operating Q was maintained. Figure 3-153 shows the frequency response after the modification. Note that a coupling slightly greater than critical ($Q \times k = 1.1$) has been selected.

3-6-5 Conclusion

After explaining some of the nonlinearities in mathematical terms, we have given some examples of voltage-tuned circuits and discussed their large-signal performance. This section has also shown that modern CAD tools can accurately predict the performance of such circuits.

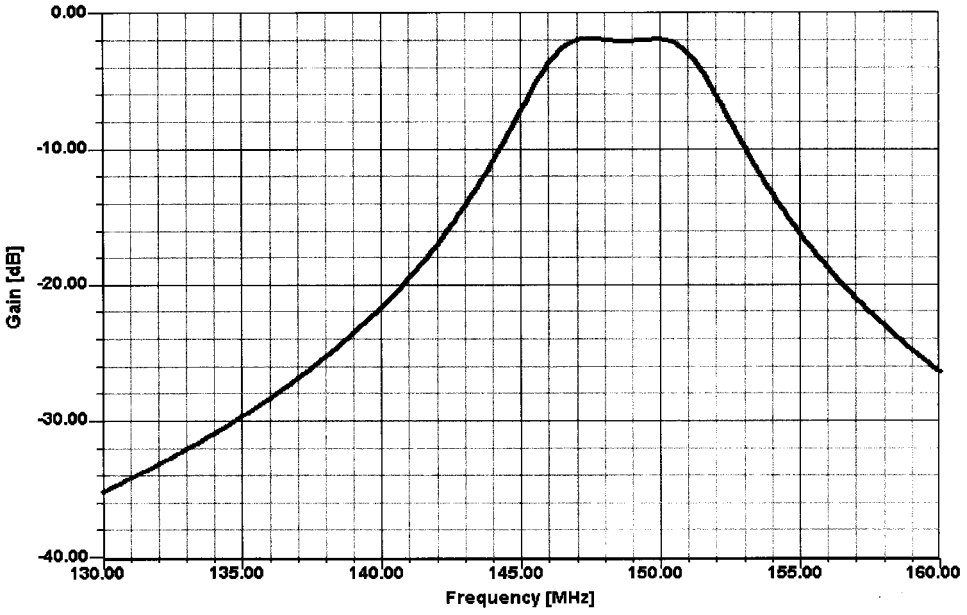


Figure 3-152 The frequency response of the improved circuit derived from Figure 3-146.

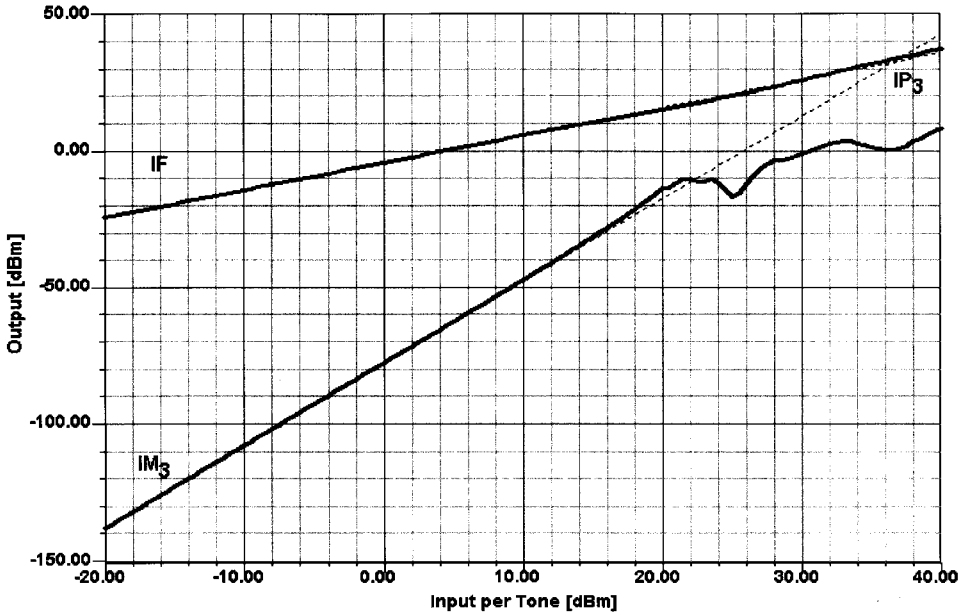


Figure 3-153 Predicted IMD performance of the improved version of the circuit shown in Figure 3-146. This type of improvement is significant for all applications. A patent is in the process of being obtained for this; therefore, the circuit cannot be disclosed at present.

3-7 DIFFERENTIAL AMPLIFIERS

The differential amplifier (Figure 3-154) goes back many years and also was the first step from TTL to ECL. The differential amplifier consists of an emitter follower (Q1) and a grounded-base stage (Q2). The constant-current generator (Q3) between ground and the emitters of Q1 and Q2 determines the current for these stages; the dc operating currents of Q1 and Q1 are the same. Most modern integrated circuits use this differential amplifier in one way or another.

If the circuit is operated about its point of symmetry, the dc component of the current through either Q1 or Q2 remains half of the emitter constant current through Q3 for all symmetrical driving signals. In addition, for input signals that are aperiodic, no even harmonics are generated. This assumes that the stage is driven symmetrically at both inputs. In our case, the differential amplifier has its second stage grounded, meaning that there is a base-to-ground capacitor, and the stage is asymmetrically fed. As the input voltage is increased, the differential stage becomes a limiter, which limits the current between plus and minus I_S, with I_S being the source current. The resulting output will become a square wave. This stage can also be used as a line receiver, translating a sine wave to an output logic suitable for driving emitter-coupled logic (ECL) stages.

One can assume that the differential amplifier has a transconductance of

$$g_m = \frac{\alpha g_{\text{in}}}{2} \tag{3-190}$$

where α is the ratio of collector to emitter current. Within the limits of the output being a sine wave, this stage has very low distortion. The output current now is determined by

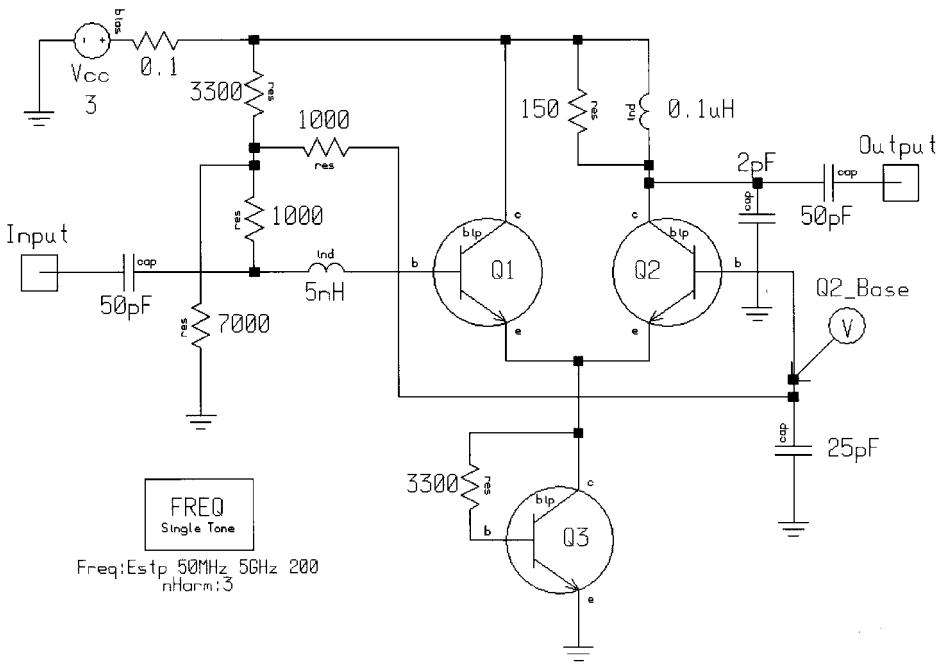


Figure 3-154 The differential amplifier schematic. The transistors are Siemens BFP420s. Figure 3-155 shows the circuit's frequency-dependent gain, matching, and noise performance.

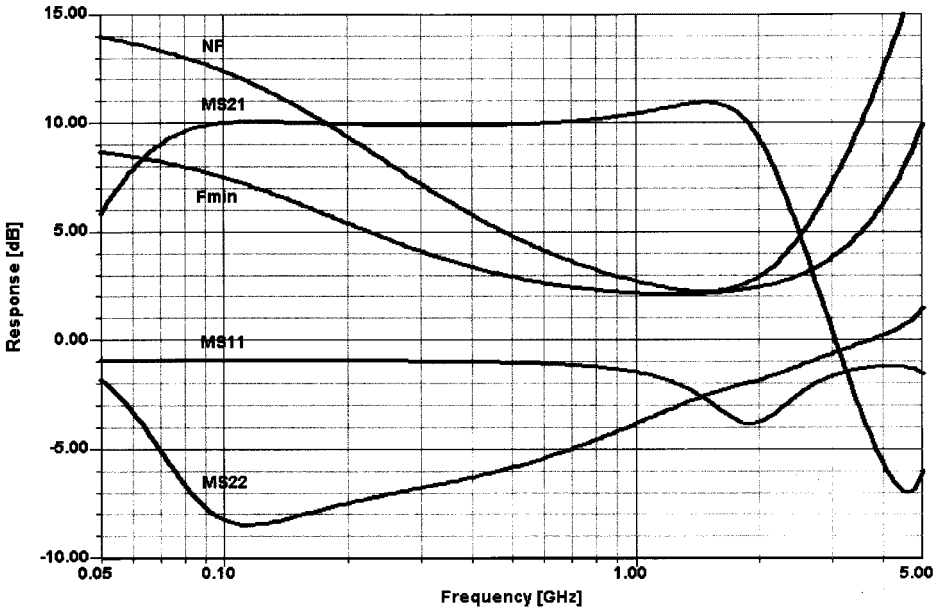


Figure 3-155 Frequency-dependent gain, matching, and noise performance of the differential amplifier.

$$i = \frac{I_{\text{in}}}{2} \tanh\left(\frac{x}{2} \cos \omega t\right) \tag{3-191}$$

where ωt is the input frequency and I_{in} is the current forced by Q3. The reason why the transconductance has to be divided by 2 comes from the fact that the current is split into two equal components for Q1 and Q2. Multiplying the new transconductance with the output load gives the voltage gain of this amplifier; its power gain can be obtained by multiplying this with $\sqrt{\text{output resistance} \div \text{load resistance}}$.

Having done the linear analysis and having learned that because of the hyperbolic tangent functions this amplifier can also be used as a line receiver, and converting sinusoidal waves to square waves, it may be useful to take a closer look at these issues. If we take a CAD oscilloscope, we see the resulting output waveform being a nonequal duty cycle output voltage, which results in high harmonic content. The amount of harmonic content can be controlled by the drive power and by the actual bias. Figure 3-156 shows the resulting output waveform.

If we now switch to a CAD spectrum analyzer, we can evaluate the harmonic content (Figure 3-157). Figure 3-158 shows the output transistor's dc I–V curves and ac load line under these conditions. Based on the previously defined conducting angle of the differential amplifier, one can optimize for harmonic content. Therefore, it is logical at this point to look at frequency multipliers, as they are sometimes needed. Finally, we are curious to see the so-called load line of the output transistor, which is, based on the interaction of all the stages, not looking the same way as the expected one, which we know from previous amplifier examples. Part of the reason for this is that for a certain voltage range, the collector voltage is in the negative region, and all kinds of saturation effects result in this surprising waveform. By reducing the input power, one would obtain a more "expected" load line. We invite readers with CAD capabilities to play with this example, as it gives a lot of insight into the operation of ECL stages.

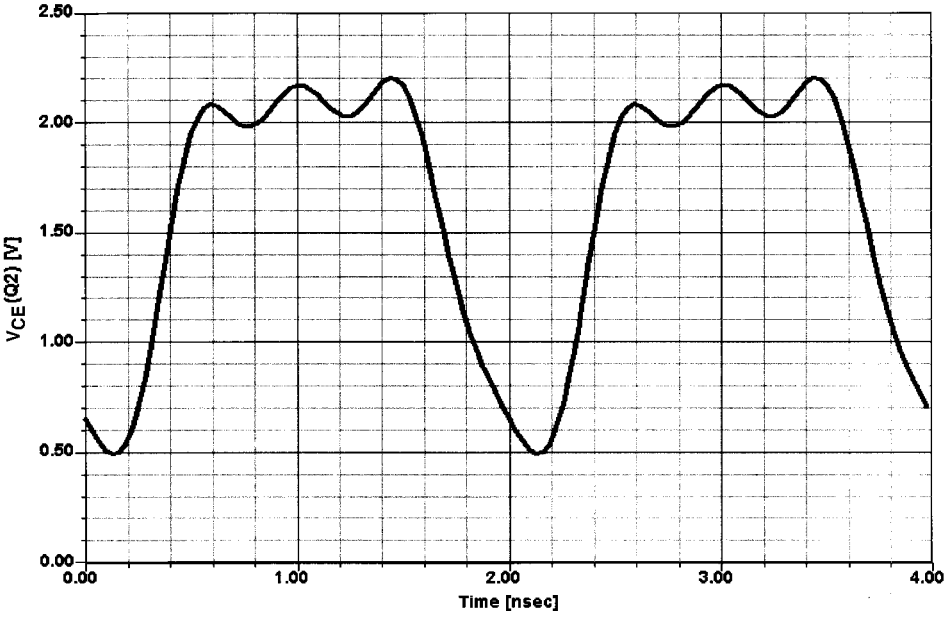


Figure 3-156 V_{CE} versus time for output transistor Q2 with a drive of 10 dBm.

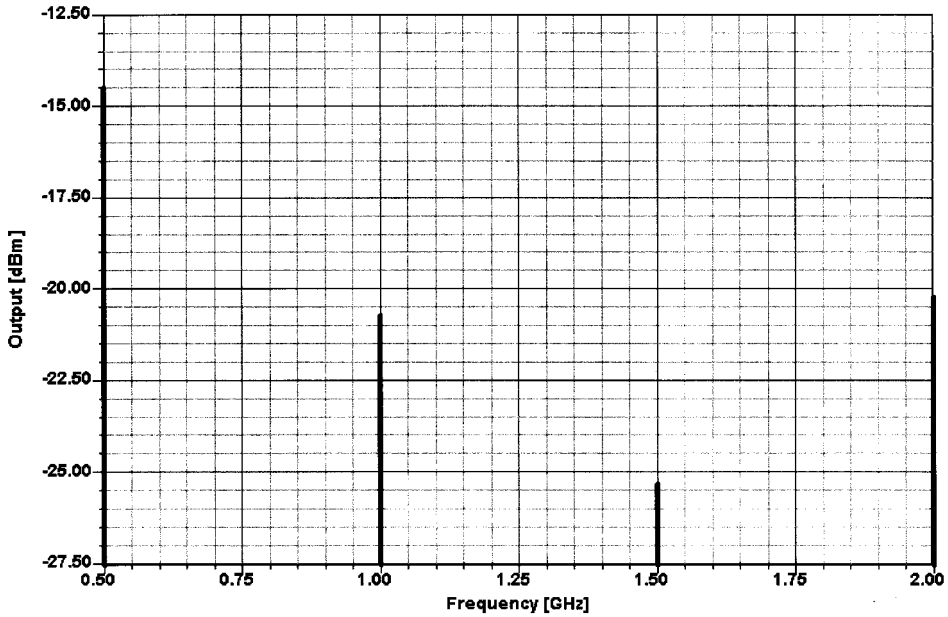


Figure 3-157 Output spectrum of the differential amplifier with a drive of 10 dBm. The biasing is not optimal for frequency multiplication.

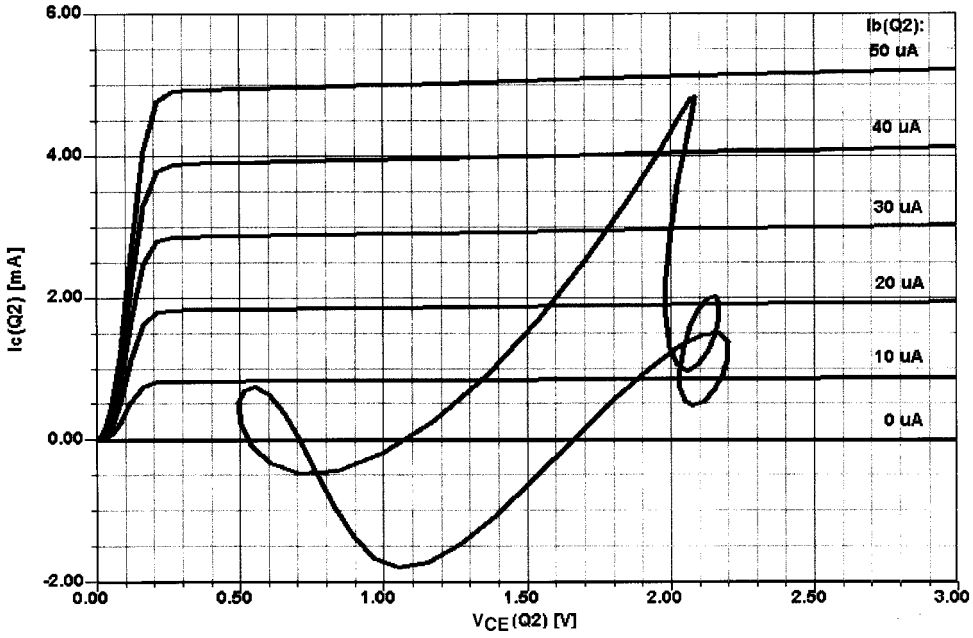


Figure 3-158 The dc *I–V* curves and ac load line for output transistor Q2 with a drive of 10 dBm.

3-8 FREQUENCY DOUBLERS

Having seen that the harmonic content invites us to bias a stage from Class A into Class B or C if our goal is a harmonically rich current, we will now evaluate a single stage that will be biased close to cutoff and driven by a relatively low frequency (500 MHz). To obtain the best results, we will use a parallel notch filter in the output, followed by a high-pass matching circuit to "catch" the desired harmonic. The input frequency is heavily suppressed by the notch filter. Based on its finite Q, there is a limited range over which this frequency can be varied with constant subharmonic attenuation. Figure 3-159 will be our test circuit.

This circuit consists of a bias transistor, Q2. This method has been discussed before and solves the temperature-dependent bias emitter thermal runaway problems through "temperature compensation." The purists among our readers may point out that the currents drawn by Q1 and Q2 are not quite identical (which would give 100% cancellation), but the approach shown is sufficient to do its job. Also, we have chosen a Class AB operation point, which is defined by Q1's 1-mA dc standing current. The conducting angle has not been optimized for the second harmonic—a task that the interested reader can do by using the conducting angle relationship shown at the beginning of this chapter. To give readers with access to nonlinear CAD some homework, we have not set this at the optimum point.

The second transistor, here called Q1 (because of its importance at RF), has a matching network at the input. This is recommended for optimum energy transfer. The matching condition can be validated by examining Figure 3-160. If we look at the higher frequency response, the input shows a possible trend of oscillation (negative loss); this could actually be eliminated by using a low-pass filter at the input instead of the high-pass filter used now. Findings of this type always make us nervous, and we hope that the nonlinear models are sufficiently accurate at higher frequencies to predict such behavior correctly. In the case of the FET doubler mentioned in Chapter 2, we found that the device modeling done by Compact (now Ansoft) gave the best results, the university answers were fairly close, and the device manufacturer data were worst. We have since become highly skeptical about device modeling unless we can validate it ourselves. Synergy Microwave has a complete set of test equipment in the laboratory, including network and spectrum analyzers with special software created by Rohde & Schwarz that can be used in addition to the Scout program to do in-house validation. This was achieved by the use of nonlinear optimization at a drive level of 0 dBm. One could say that this is similar to using a load-pull approach at the input, and then resorting to the large-signal S parameters obtained. Another note of caution: Capacitors at these low values tend to have a very large percentage error; for instance, a 1-pF capacitor, if not selected carefully, can vary between 0.7 and 1.3 pF—a 50% error. Needless to say, this is not acceptable for useful production.

The output consists of a high-pass filter and a notch filter with the purpose of reducing the subharmonic significantly. We have again done this on purpose, leaving room for the reader to come up with a more inventive output circuit like a Cauer-type (same as elliptical) low-pass filter with discrete notches at higher harmonics of the input frequency, such as 1.5 GHz, 2 GHz, and so on. In addition, we have shown three ways how to implement the inductance for the notch filter. This is equally applicable for other inductors in this value range. As frequency increases, this becomes more interesting but one needs to remember that many materials like silicon and gallium arsenide are extremely lossy. The same inductor approach has been used in Figure 5-61 of Chapter 5. For silicon, even the manufacturer becomes an issue. Figure 3-161 shows a top view of a "rectangular" inductor that actually approaches the spiral inductor; its implementation can be seen in Figure 3-162. Modeling this was a

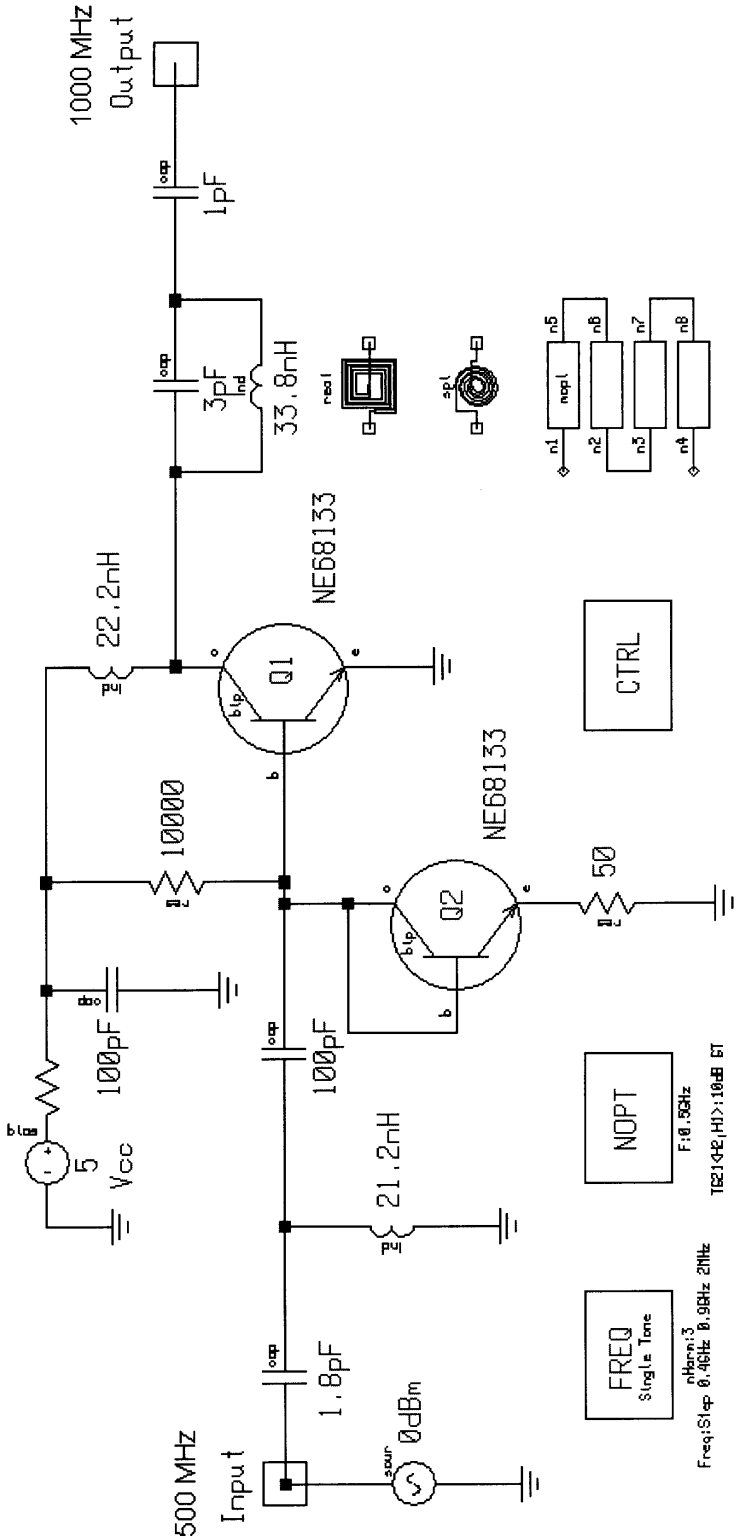


Figure 3-159 BJT frequency doubler schematic. A parallel-tuned trap in Q1's collector attenuates fundamental feedthrough.

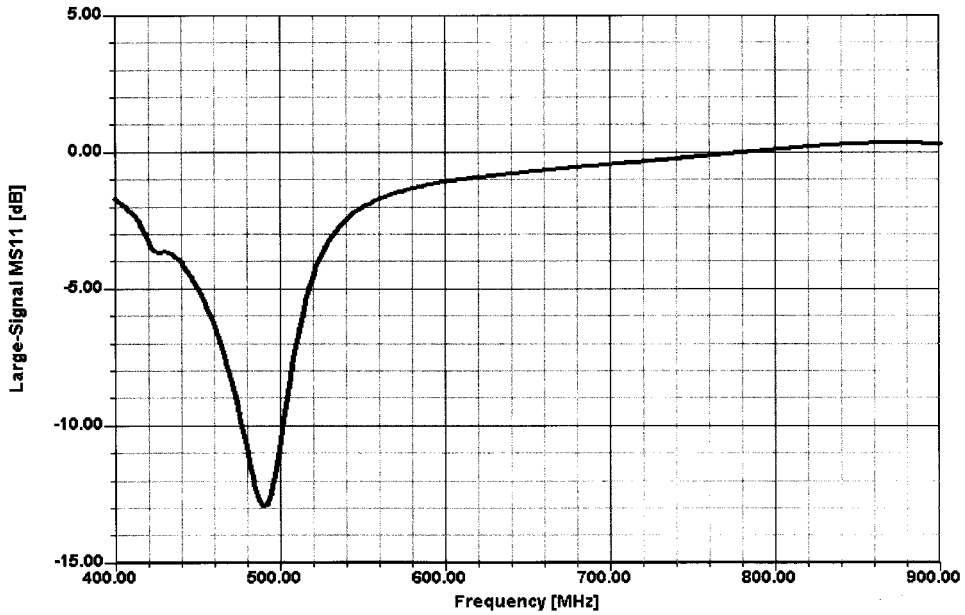


Figure 3-160 Large-signal MS_{11} for the BJT doubler. Note the positive region above 800 MHz.

nightmare, and during our Compact days, we actually had to modify the electromagnetic simulator to accommodate this and get results where measurements and simulation agreed. Figure 3-163 shows the measured and simulated Q as a function of frequency of a silicon-based inductor. In the case of other materials, such as PC board or GaAs, the models in the CAD tools are sufficient to predict the losses of the inductors. Following this initial exercise, Motorola then developed an improved silicon spiral inductor; its frequency-dependent Q is shown in Figure 3-164. Note that this Q was measured and not simulated.

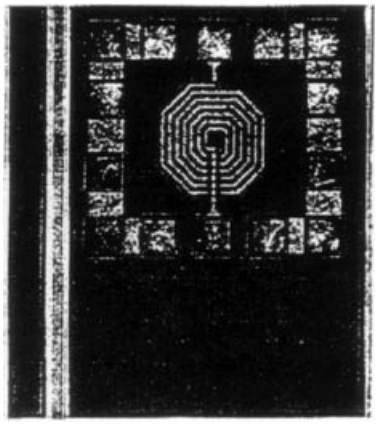


Figure 3-161 A 6.5-turn spiral inductor in silicon.

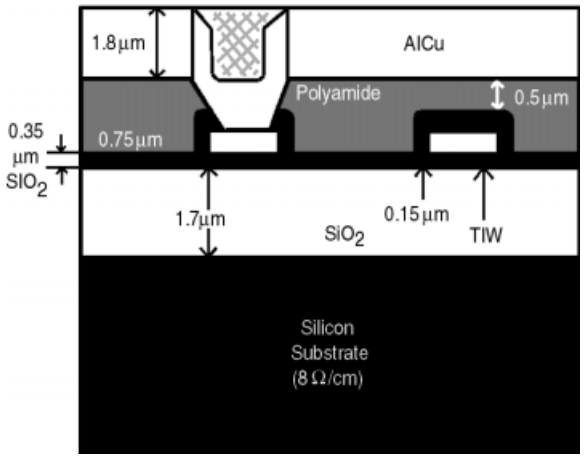


Figure 3-162 Cross section of the silicon inductor.

One of the reasons why we use a transistor multiplier is that we expect gain (because of the square-law characteristic, an FET would have been better tailored for a doubler); therefore, we have "measured" the gain of this circuit while driving it with 0 dBm. Another important factor is that, because of the way the CAD tool displays it, the gain is referenced to the 500-MHz input. Mentioning this is important; otherwise, the reader may look for the 1-GHz gain, which does not exist. Figure 3-165 shows the doubler gain.

By now we are curious to see the doubler in action, and by reviewing Figure 3-166 we can see the output spectrum. The output spectrum validates the statement that the output circuit, while doing a marvelous job of removing the 500-MHz input frequency, does not cure the problem of unwanted multiples of the input frequency. Again, a Cauer low-pass filter, while introducing some possible additional loss, should be considered for this circuit to have a bandpass response. On the other hand, many oscillators have very little harmonic suppression, and the same type of postfiltering needs to be considered. There are two-transistor solutions available also, such as push–push for frequency doubling (push–pull would

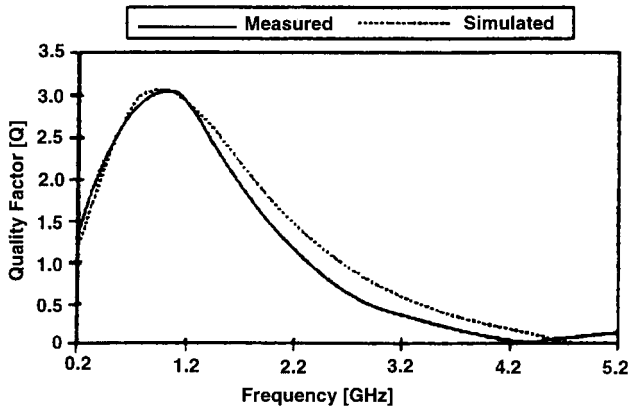


Figure 3-163 Measured and simulated *Q* versus frequency for the simulated inductor.

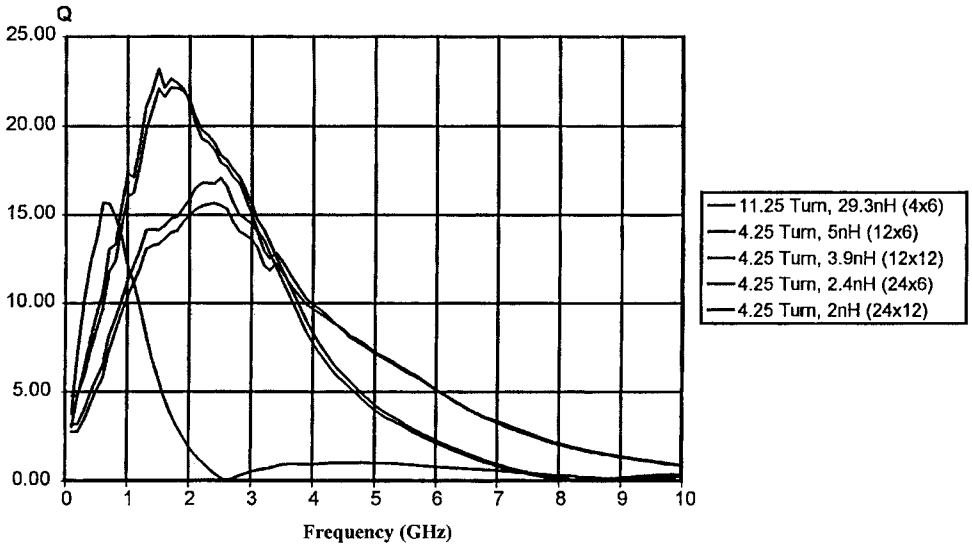


Figure 3-164 *Q* versus frequency for Motorola improved inductor.

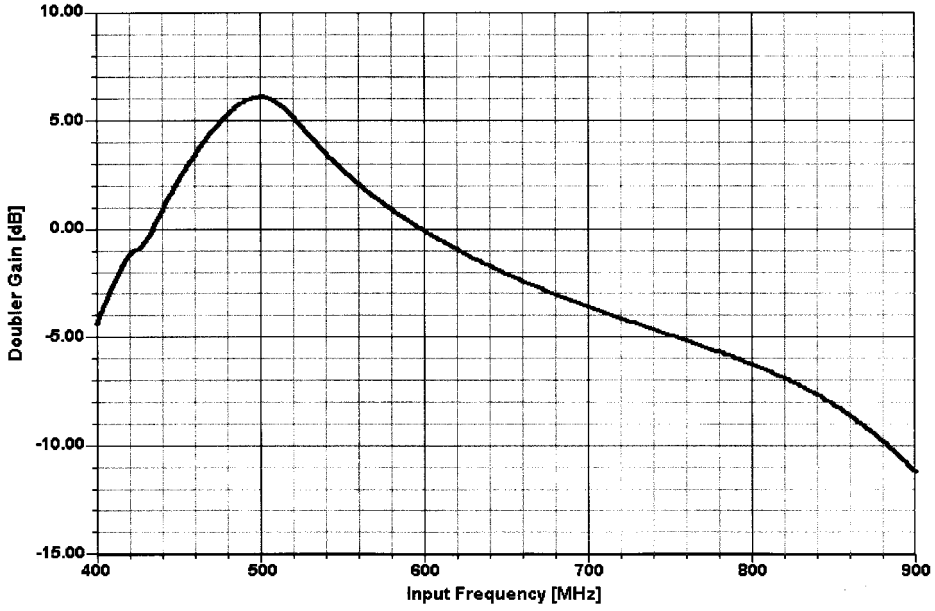


Figure 3-165 Frequency-dependent gain of the BJT doubler.

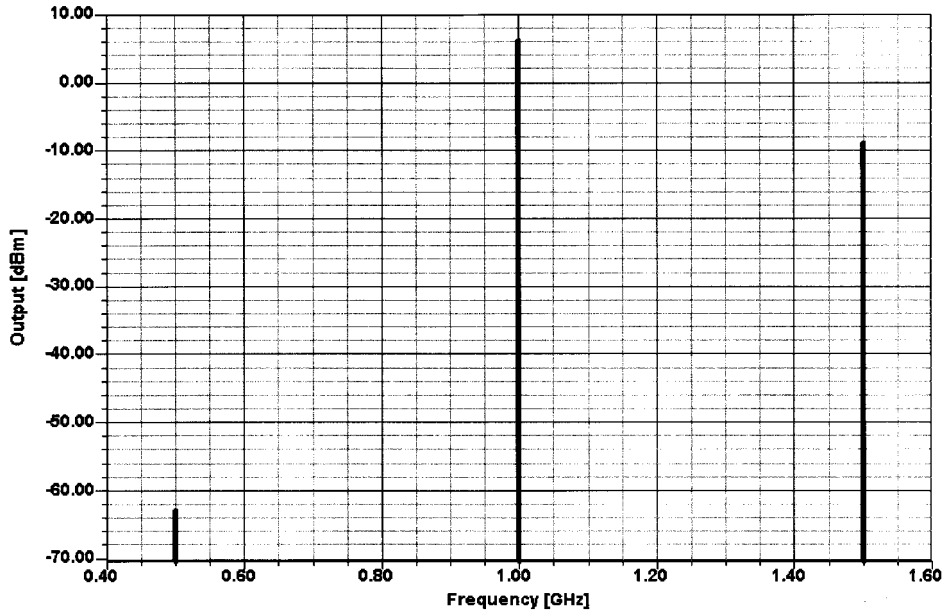


Figure 3-166 Output spectrum of the BJT doubler.

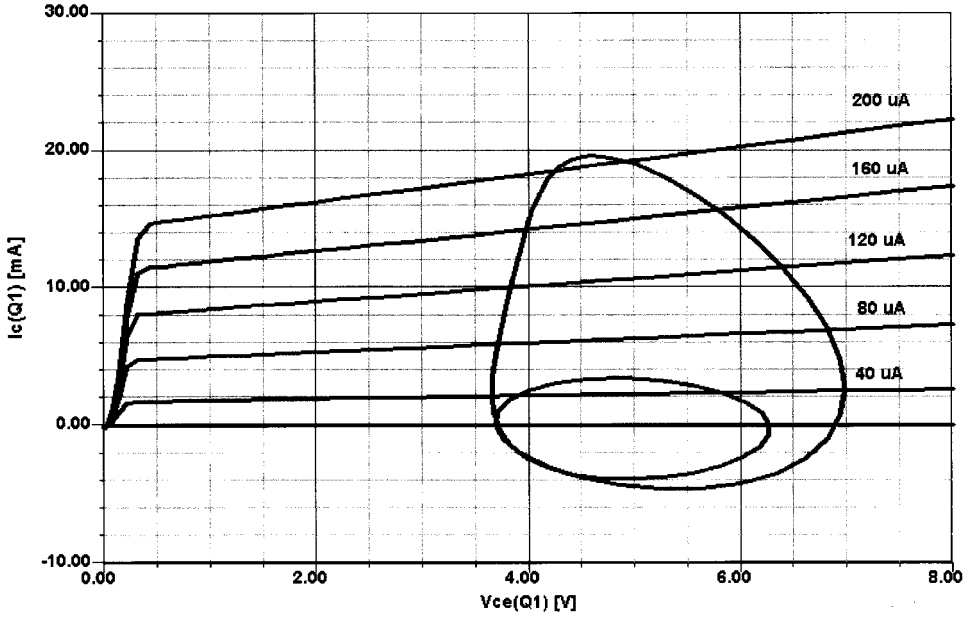


Figure 3-167 The dc *I–V* curves and ac load line for RF transistor Q1 in the BJT frequency doubler.

eliminate the desired second harmonic quite well). Actually, two diodes could also have been used for doubling, but since this requires transformers and a postamplifier, we have shown the preferred application. The only advantage of the diode application is that it is wideband, while this BJT circuit, as we have just learned, must be made narrowband for reasons of spectral purity.

Finally, we are curious to see the actual operating state of this amplifier, and this can be done by analyzing the load line (Figure 3-167). This is as complicated as the analysis for the differential amplifier. Because of the stored energy in the various tuned circuits, the load line looks quite impressive and would probably be difficult to predict without a nonlinear CAD tool.

3-9 MULTISTAGE AMPLIFIERS WITH AUTOMATIC GAIN CONTROL (AGC)

Having discussed the basic differential amplifier, we will now evaluate the performance of the MC1350/1490 video amplifier chip from Motorola. Most of the circuitry in the chip is self-explanatory. Figure 3-168 shows its inner circuit, and Figure 3-169 show a recommended application. By grounding one side of the input, we effectively get the same performance as the differential amplifier just discussed, with one difference being that the input is now fed to the constant-current generator. To minimize the distortion at the two lower transistors, they are decoupled by a 66-Ω resistor. Figure 3-170 shows the frequency-dependent gain, matching, and noise figure. Since this particular device has been specifically designed for AGC purposes, we also show in Figure 3-171 its gain reduction versus applied AGC voltage.

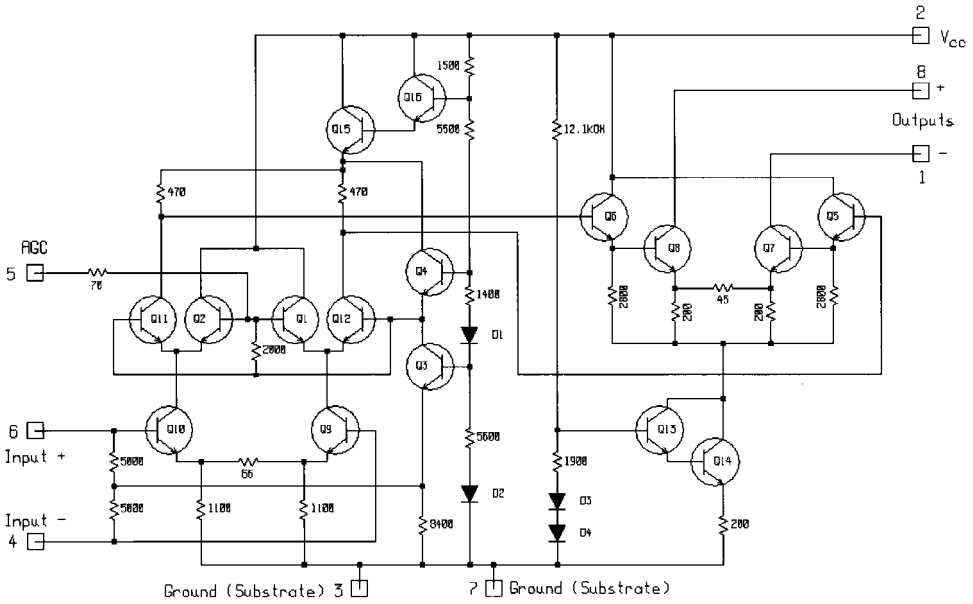


Figure 3-168 Schematic of the Motorola MC1350/1490 entered for simulation.

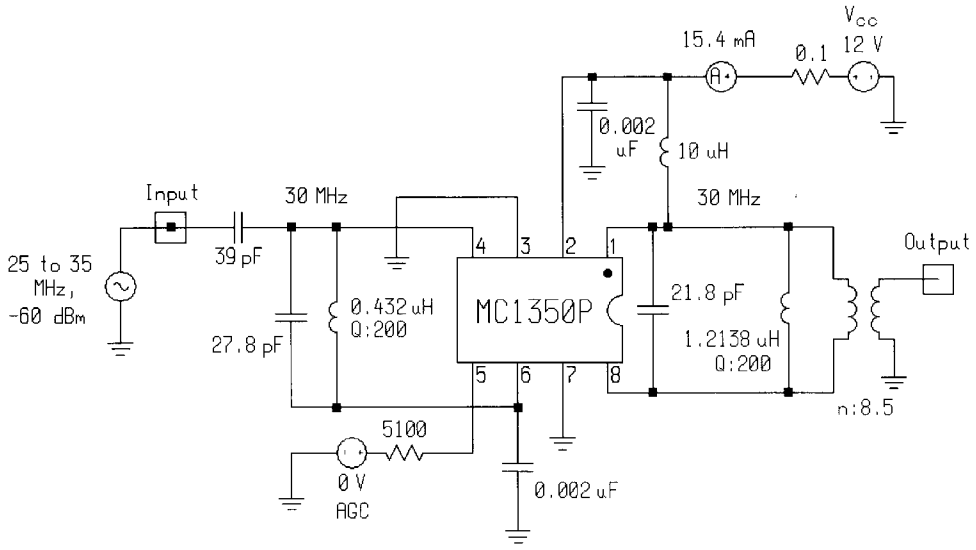


Figure 3-169 MC1350/1490 30-MHz application circuit.

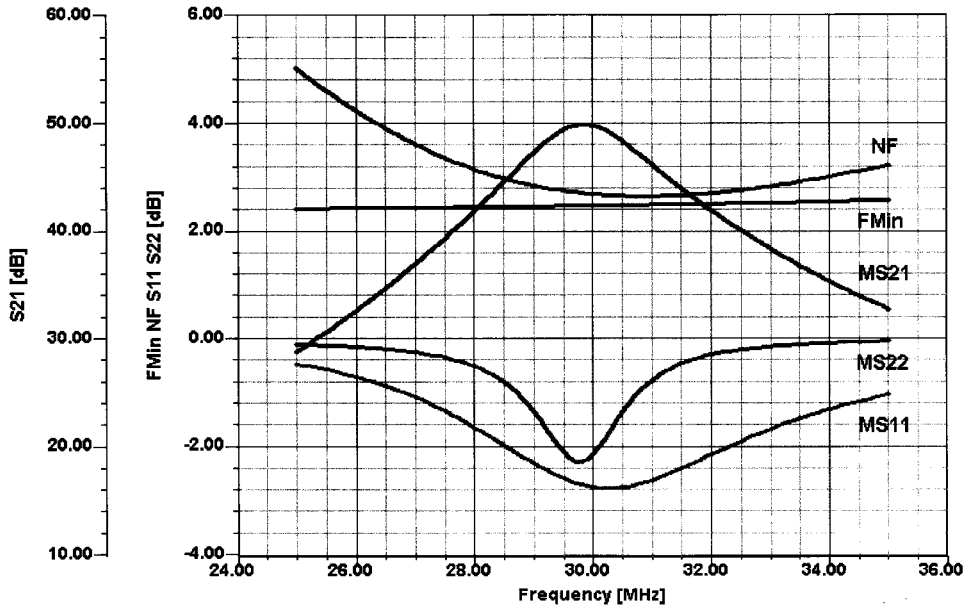


Figure 3-170 Frequency-dependent gain, matching, and noise performance of the MC1350/1490 in the circuit of Figure 3-169.

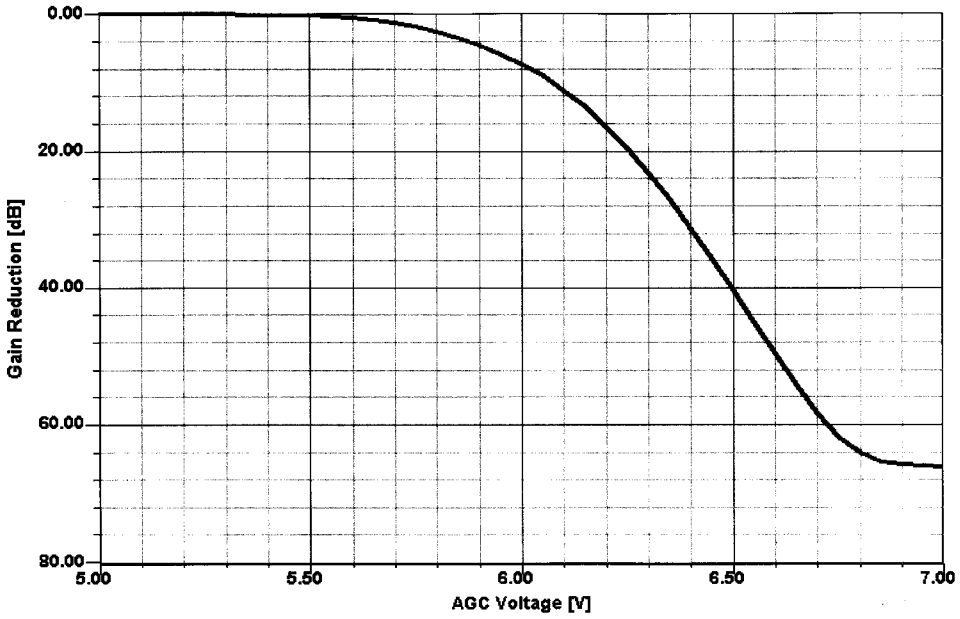


Figure 3-171 Gain reduction versus applied AGC voltage for the MC1350/1490 IC.

3-10 BIASING

A lot has been written about biasing, and since the purpose of this book is not to focus too much on dc, the following examples including RF consequences have been selected and considered sufficient. Early application reports by Motorola and others have provided a wealth of data. As a side effect of the introduction of the wireless era, we find that the engineers are so pressed for time that the number of complete application reports has approached zero. This is one of the reasons why we have tried to minimize the mathematical basis of this book (some is still needed), while still providing practical assistance. Figure 3-172 shows a textbook type of approach for RF.

In order to decouple the transistor from the biasing, we use RF chokes and dc decoupling (bypass) capacitors. This is the technique used in previous examples, so one might ask, "What's new?" As frequency increases, these inductors are either not manufacturable or have such a low Q that their use becomes questionable. This is the point where one may introduce the so-called distributed elements.

Figure 3-173 shows the very same circuit but resorting to distributed rather than lumped elements. The newly introduced elements (part of any good, up-to-date CAD tool; see element library for how to use their physical element descriptions) are described next.

- *Transmission Line.* Any printed connection between two points on a circuit board is a transmission line (Figure 3-174). Its characteristic electrical impedance is a function of the square root of the dielectric constant (ε_r), the width, metallization, thickness and height above substrate of the line, and the loss tangent of the substrate. Since lines frequently have to be laid out in the form of curved connections or have a bend in their direction, we have to add elements capable of describing the high-frequency conse-

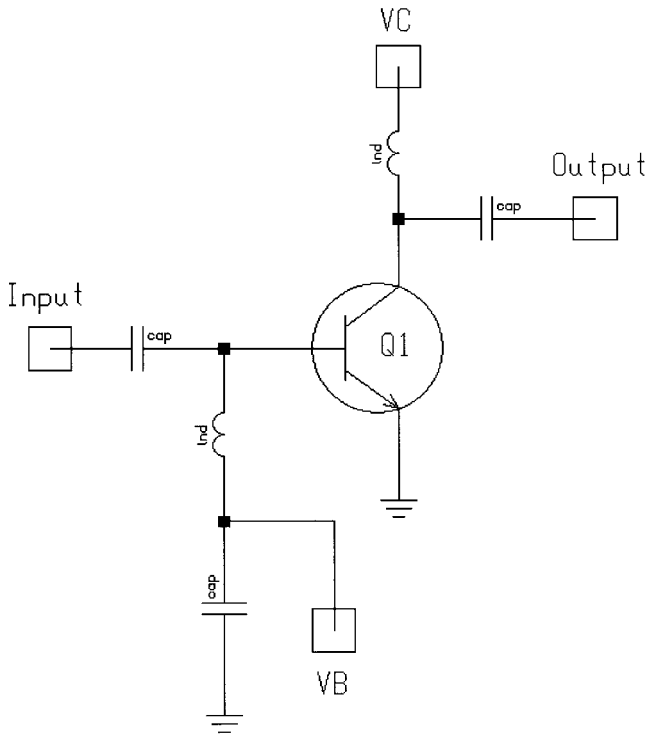


Figure 3-172 Simple BJT RF amplifier with lumped elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

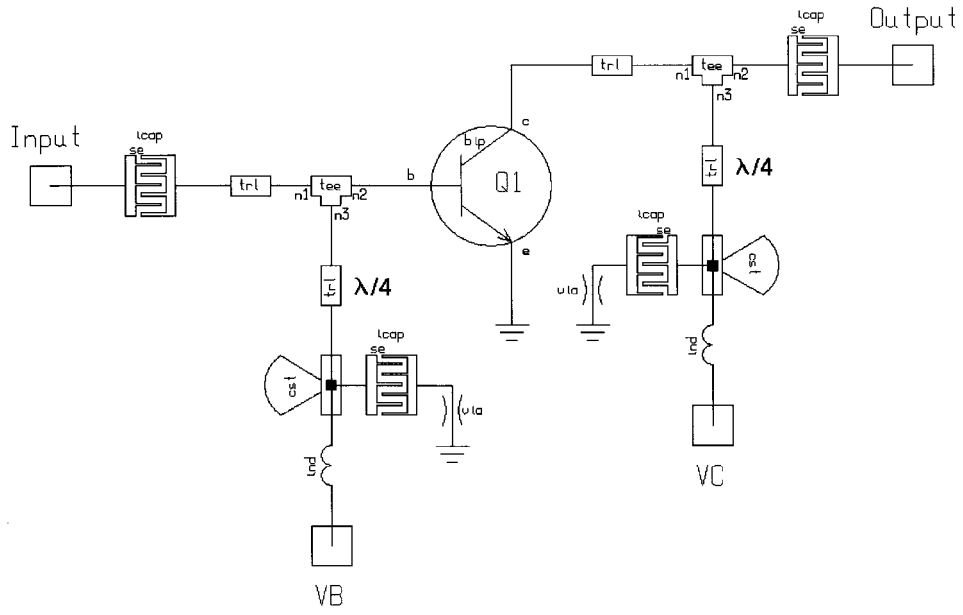


Figure 3-173 Simple BJT RF amplifier with distributed elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

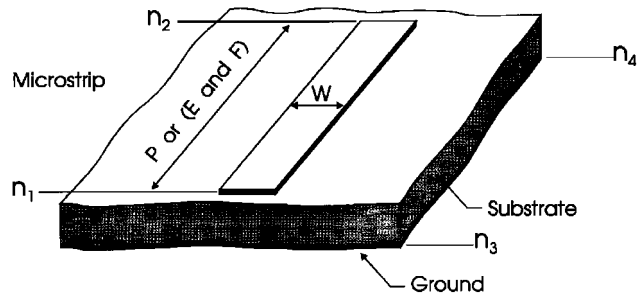


Figure 3-174 Transmission line in microstrip.

quences of such connections. Figures 3-175 and 3-176 show mitered and radial bend elements that perform this function.

- T *Junction, Cross, and* Y *Junction.* By the time a point like a collector or base, or its FET equivalent, spreads out into connections with other elements, we need additional modeling capability to describe ⊤ connections, crossings, and Y junctions. Figures 3-177, 3-178, and 3-179 show the way in which these connections must be modeled.

If the need exists, the standard inductances must be replaced with a transmission line whose length is λ/8 at the operating frequency. At higher frequencies, these transmission lines, however, then go into λ/4 resonant mode and later become capacitive. This type of design makes it fairly narrowband. A way around this is the use of printed inductors, as shown in Figures 3-180 and 3-181.

These inductors have a self-resonant frequency similar to the transmission line mentioned above, but the safety margin is significantly higher.

Talking about printed inductors, a logical extension of this is the printed transmission-line-based transformer, as shown in Figure 3-182. One can consider this as two interlaced rectangular inductors and, based on the substrate material, they are useful over a wide frequency range. Besides being used as a transformer, they can also be used to transit from

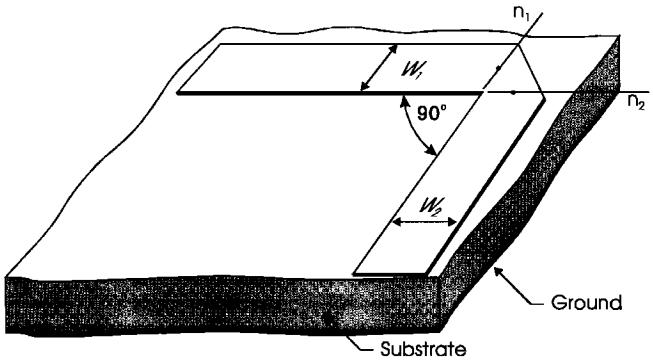


Figure 3-175 Mitered bend.

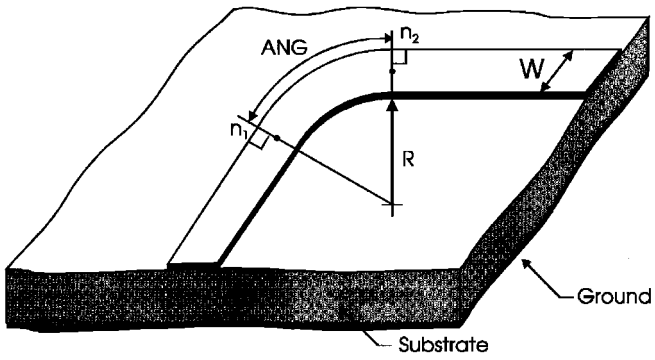


Figure 3-176 Radial bend.

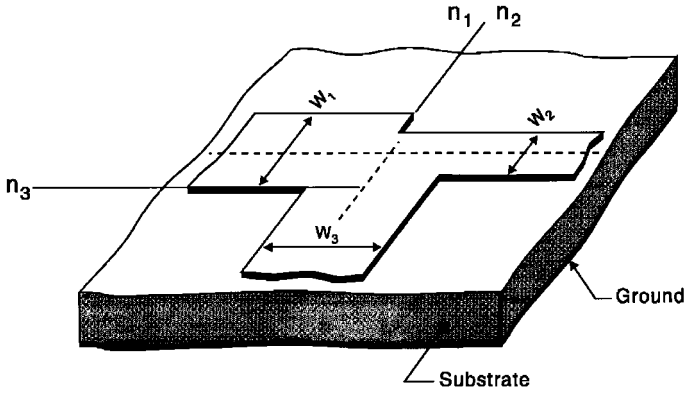


Figure 3-177 T junction.

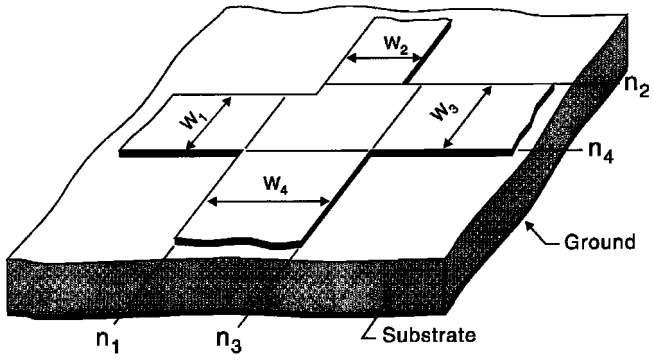


Figure 3-178 Cross.

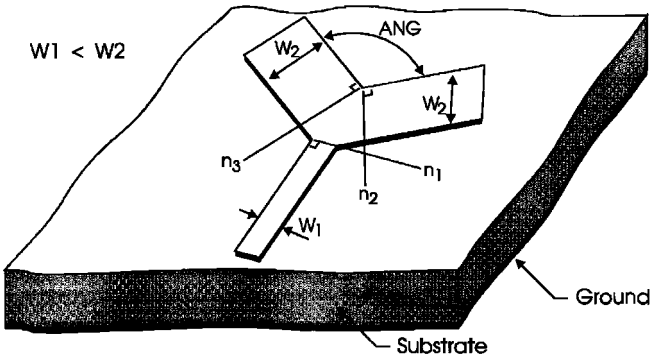


Figure 3-179 Y junction.

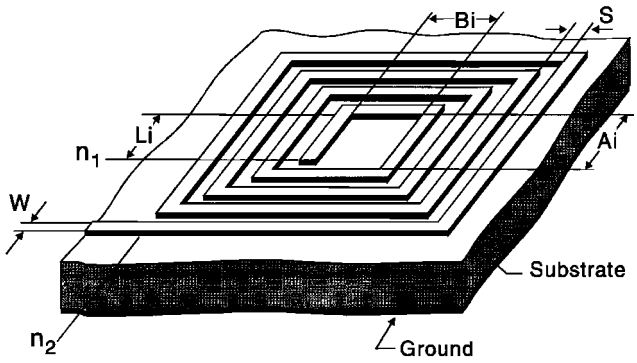


Figure 3-180 Rectangular inductor.

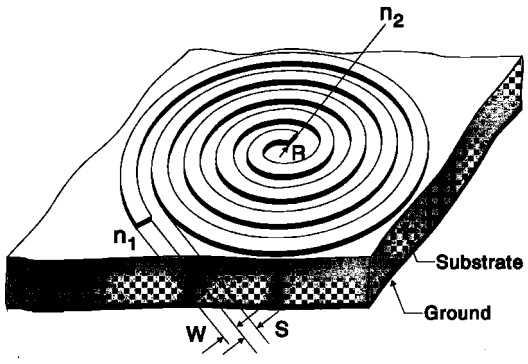


Figure 3-181 Spiral inductor.

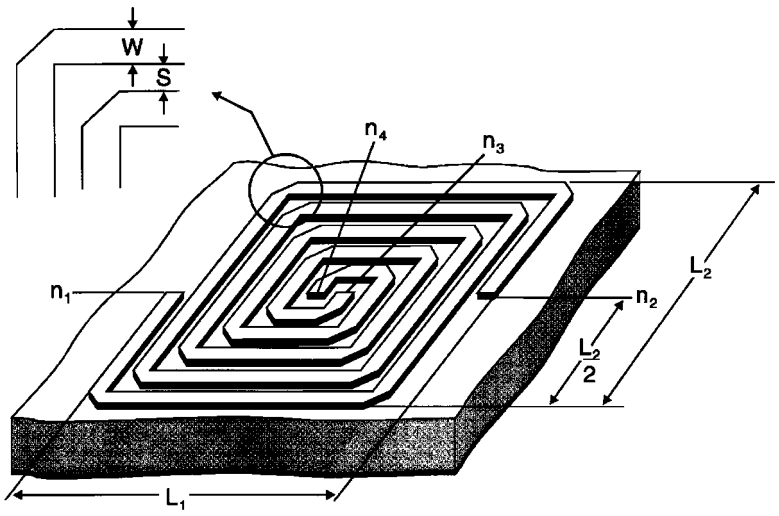


Figure 3-182 Transformer in microstrip.

unbalanced to balanced transmission provided that the difference in length from a connection point of view does not cause any problems (this is a layout issue).

A popular form of combining stages is the so-called Lange coupler (Figure 3-183) invented by the German Julius Lange. It is one of the major contributions in wideband applications.

Useful application of the Lange coupler probably starts at 4 GHz. It consists of parallel transmission lines with the appropriate connections as shown. Lange couplers are typically built with four, six, and eight fingers. The Ansoft Serenade product has a Lange coupler synthesis program that can be used to gain more insight into this coupler's application. We assume that other modern software has similar capabilities.

Where meander-type inductors are necessary, a neat way to implement and simulate them is to use the multiple coupled lines element (Figure 3-184) of the Serenade product, which both quickly and accurately calculates the behavior of the meander, including self-resonances and losses. We made use of this arrangement in our previous examples.

- *Interdigital Capacitors.* The issue of tolerances of small capacitors already has been brought up. The interdigital capacitor can be made on printed circuit board material as

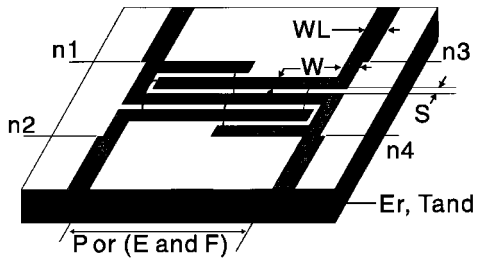


Figure 3-183 Four-strip Lange coupler.

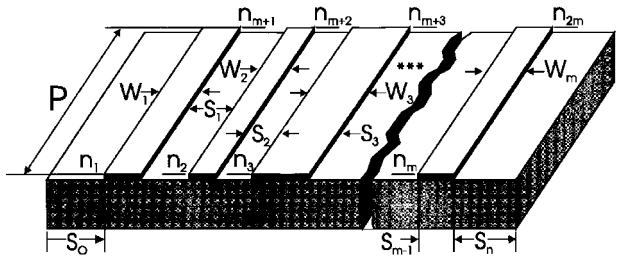


Figure 3-184 Multiple coupled lines element in microstrip.

well as gallium arsenide, and if its dimensions are continuous with the transmission-line width it does not cause any abrupt changes in impedance. This type of capacitor permits one to obtain very small values. An alternative to this is the use of transmission lines with a length of $\frac{3}{8}\lambda$. We learned earlier that a transmission line is inductive at frequencies below its fundamental resonance, is resistive at below resonance, and becomes capacitive above resonance. Again, bandwidth is also an issue. An interdigital capacitor consists of a number of parallel fingers as shown in Figure 3-185, and its capacitance can be varied by adjusting the number of fingers and their spacing. The advantage of the interdigital capacitor compared to discrete components is its low variation in value.

- *Radial Stubs.* The radial stub (Figure 3-186) is not much different from a $\lambda/4$ resonator, but its bandwidth is much greater than a simple transmission line. This is another way to ground the "cold" side of a transmission line or part of a circuit that needs to be grounded for RF. Of course, the interdigital capacitor comes in a version that is a combination of a capacitor and a via hole.
- *Via Holes.* The "cold" end of the transmission line, being considered either an inductor or a capacitor, needs to be connected to a sufficiently large copper backplane. One very efficient way to do this, especially if there is not enough copper left on the top of the board, is the use of via holes (Figure 3-187). One could theoretically generate

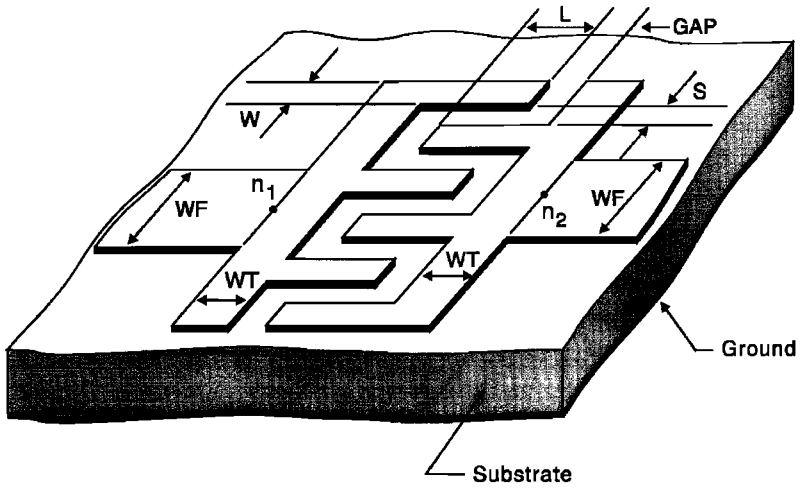


Figure 3-185 Interdigital capacitor.

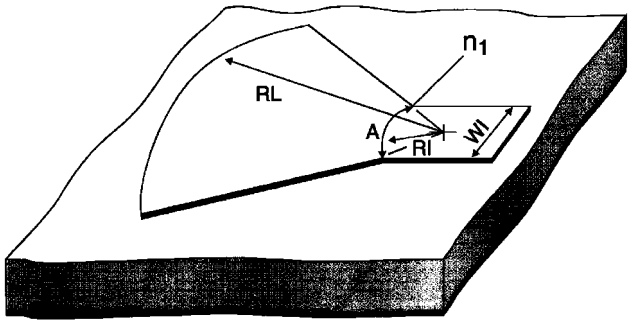


Figure 3-186 Radial stub.

a via hole with a rivet, but most manufacturing processes don't allow this; the normal solution is to use plated-through holes left open. In PC boards, via holes are typically cylindrical; on substrates like GaAs, they may be conical.

- *Correction Elements.* Although the behavior of an actual circuit proceeds regardless of our ability to measure and describe it, we do not enjoy this luxury in simulating circuit behavior in software. In a simulator, effects that are insufficiently described will be inaccurately simulated. For instance, segments of high-impedance transmission line (e.g., 120 Ω) are frequently used for dc feeds and RF chokes. In predicting the effect of a transition from 120-Ω line to 50-Ω line, a simulator must be alerted to the discontinuity so it can do the necessary mathematical corrections to account for the impedance jump. To do this, a specific circuit element, the STEP (Figure 3-188), must be inserted between the 120-Ω and 50-Ω line elements in the simulation circuit file. In addition to substrate data, we characterize a STEP by providing the widths of its input and output lines. The element itself has no physical length.

A similar correction is necessary if a transmission line is used as a resonator or just "left open" at one end. Such a transmission line tends to radiate and, because of its high-impedance properties, it reacts differently as far as its electrical length is concerned. A zero-length one-port element, the OPEN (Figure 3-189), must be added to such a line for mathematically correct calculation.

We end this excursion into distributed elements here. Interested readers may want to obtain a CD containing the element library of their favorite CAD tools that combine capabilities in the microwave area with a full set of distributed circuit elements. It is most important to keep

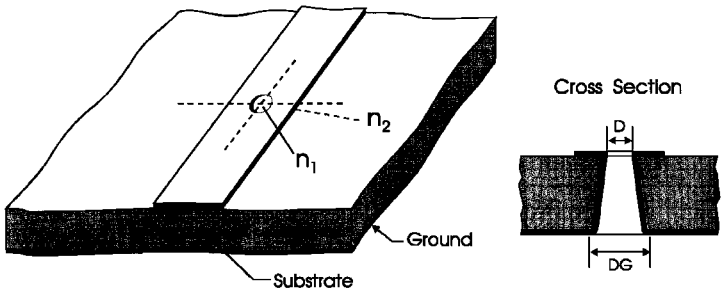


Figure 3-187 Via hole.

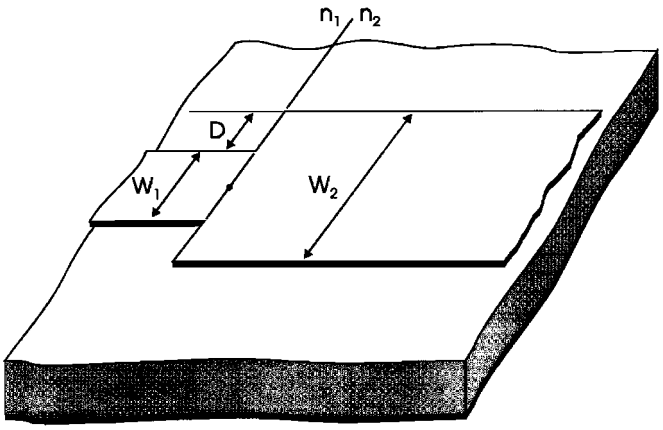


Figure 3-188 The STEP element tells the simulator to calculate the effects of joining transmission lines of differing characteristics.

in mind that, as frequency increases, we rapidly move into the area where we must consider all these distributed elements to achieve accurate simulations—even if doing so makes simulation a painful and time-consuming effort.

Finally, anyone who adventures into this area *must* obtain a foundry manual from the company that will build the integrated circuit or hybrid under design. There are basically two foundries, one applicable for MOS technology and one for GaAs. In the case of the bipolar transistor, the foundry service is not yet well-established. While we were able to interest one German company and one U.S. company in generating a custom bipolar IC within the activities of this book, the only "open" foundry we know of for bipolar (actually HBT) technology is the one operated by TRW. As of 1999, this seems to be the leading company for BJT-related products, and many currently available ICs have been developed with this foundry service.

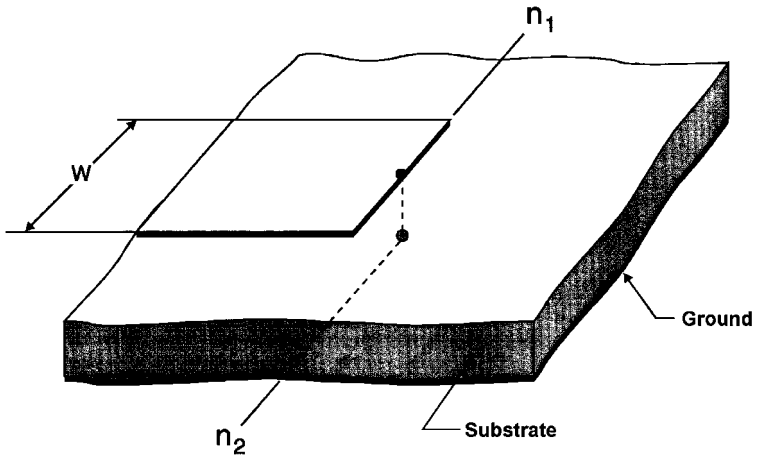


Figure 3-189 The OPEN element tells the simulator to calculate the effects of leaving the end of a transmission line unconnected.

We mention the issue of foundries here again because each foundry has its own proprietary approach to modeling discontinuities. The availability of a foundry service somewhat eases the requirement that a designer be fully up to speed on the nuances of discontinuities, because a foundry's designer service will help customers account for all relevant parasitics or discontinuities in their designs. In addition, there are tables of *S* parameters for standard cells of either capacitors, resistors, or inductors. The designer may then be forced to adjust the circuit so that it will work with a particular inductance value or value of another component within the resolution of the table that describes these elements. Information on the active parts, such as diodes and transistors, was given in Chapter 2.

3-10-1 RF Biasing

Applying the knowledge we have just acquired about distributed elements, Figure 3-190 shows a somewhat exaggerated case of using the various elements. While it is consistent with the simple RF case with which we started, modeling of this type is certainly necessary at higher frequencies. Many but not all of the elements we have just described are used, but we have not solved the question of the actual dc biasing.

3-10-2 dc Biasing

The following discussion is the recommended procedure for biasing a transistor. We will start with the collector current (in our example, Figure 3-191, 10 mA). To separate RF and dc, we have a combination of a load resistor (R_L) and an inductance (L_C). Information about this operation was given already in Section 3-2-6. The emitter resistor should have a voltage

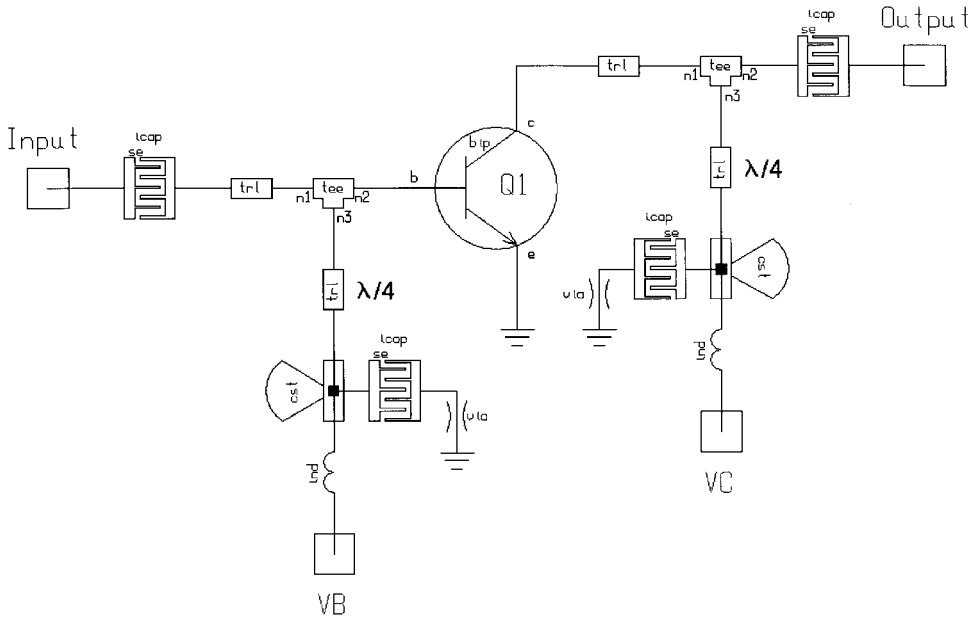


Figure 3-190 Simple BJT RF amplifier with distributed elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

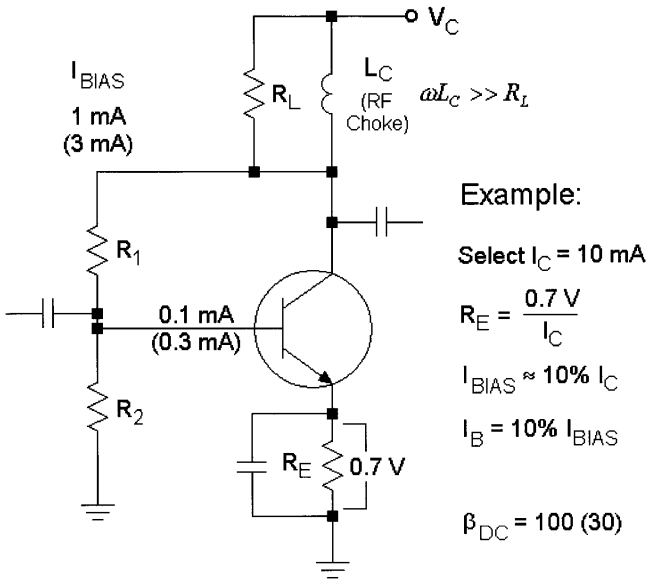


Figure 3-191 A dc biasing example.

drop of about 0.7 V, and therefore the emitter resistor is $0.7/I_C$. The next assumption is that the transistor has a dc current gain of 100, and therefore we decide to have the bias resistor chain (R_1 and R_2) draw 10% of I_C, or 1 mA. Assuming a collector voltage of 5 V, the value of $R_1 + R_2$ has to be 5 kΩ to result in the required 1 mA. Because the base current now is 10% of the divider current, the base will draw 0.1 mA. The voltage across R_2 has to be the voltage drop across R_E (0.7 V) + V_{BE}, which is also 0.7 V, or 1.4 V. Given the 1 mA, R_2 = 1.4/1E–3 = 1.4 kΩ. R_1 therefore is 5 kΩ – 1.4 kΩ = 3.6 kΩ. This concludes the calculation.

In the case of a dc beta of only 30, we need to raise I_{BIAS} from 1 mA to 3 mA since the base current now will be 0.3 mA. This changes the resistor values. $R_1 + R_2$ = 5 V/3 mA = (rounded) 1.6 kΩ. Following the same approach, we need 1.4/3E–3 = (rounded) 470 Ω. This makes R_1 = 1600 Ω – 470 Ω = (rounded) 1100 Ω. These roundings were necessary because resistors can only be bought in certain resistance steps.

However, for low battery voltages, the 0.7 V between the emitter and ground cannot be wasted. Therefore, a dc stabilization circuit, to which we have already referred, will be considered next. Figure 3-192 shows a combination of an RF transistor, Q1, and a dc stabilization transistor, Q2. The most important thing about this circuit is the fact that the difference between the supply voltage and the collector of Q1 is 0.3 V, which is approximately half of the voltage we had to "waste" across R_E in the above-mentioned circuit. Plus, the voltage drop is now in the collector, and given the fact that we feed Q1's collector through the inductor of a quasi-tuned circuit rather than through only a resistor, we get a much higher voltage swing. The base–collector diode temperature dependency of Q2 is compensated for with the silicon diode feeding the base. It is also fascinating to see that, because of the low currents, both the voltage drop of the diode and V_{BE} of Q2 are about the same, and yet Q1, operating at a reasonable current, shows the expected 0.7–0.8 V V_{BE}. We have supplied all the necessary dc voltages to give the reader an incentive to calculate the mechanism of this stabilization. The key equation to this is to determine the current in the voltage divider feeding

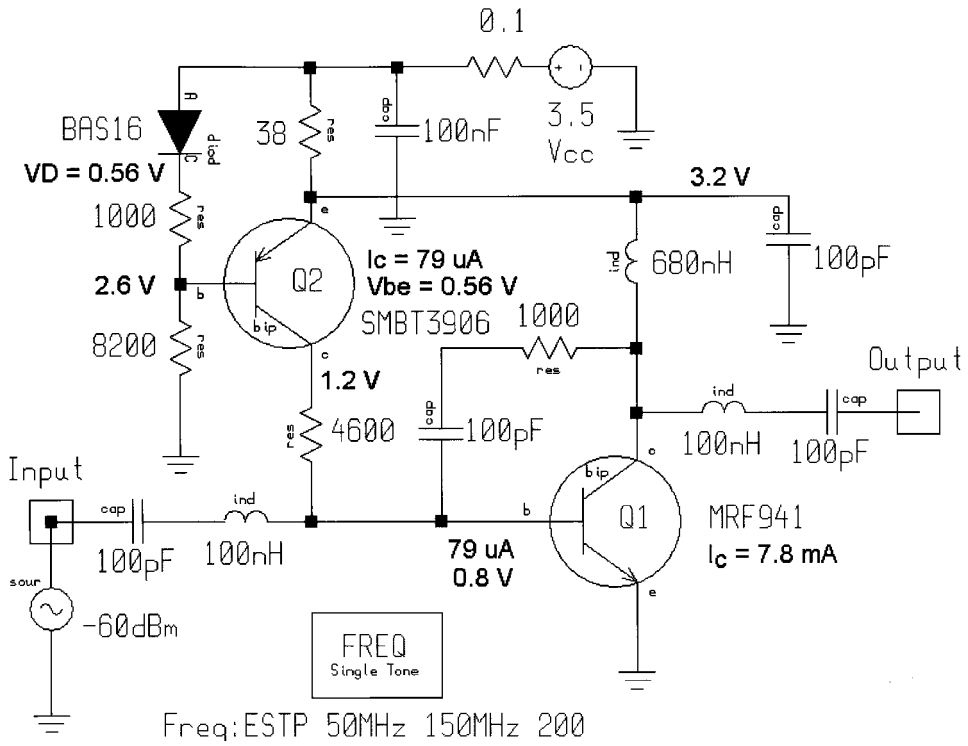


Figure 3-192 An RF amplifier with active biasing. (After the low-noise amplifier circuit in [29].)

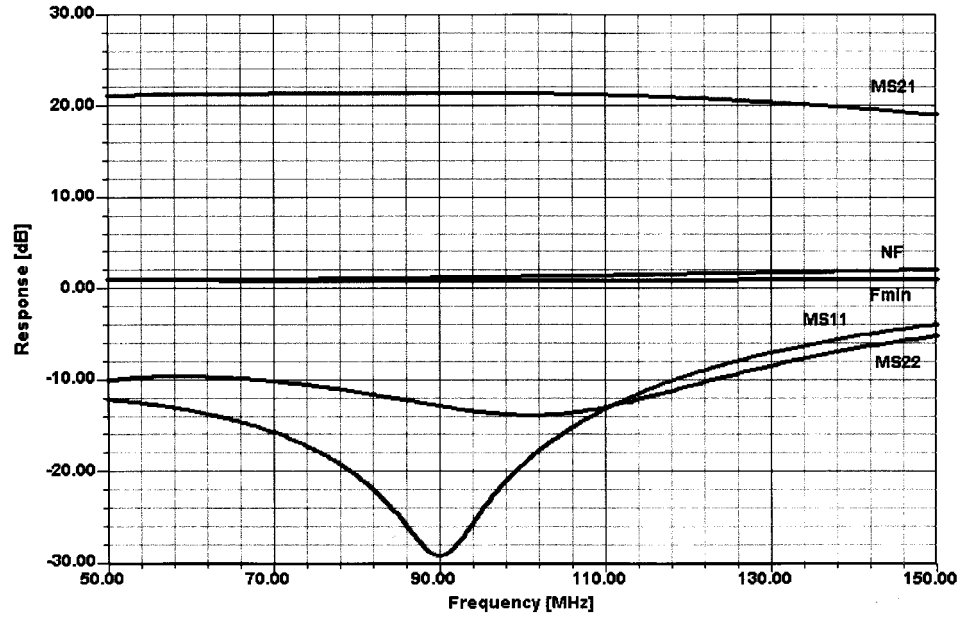


Figure 3-193 Frequency-dependent gain, matching, and noise responses for the amplifier in Figure 3-192.

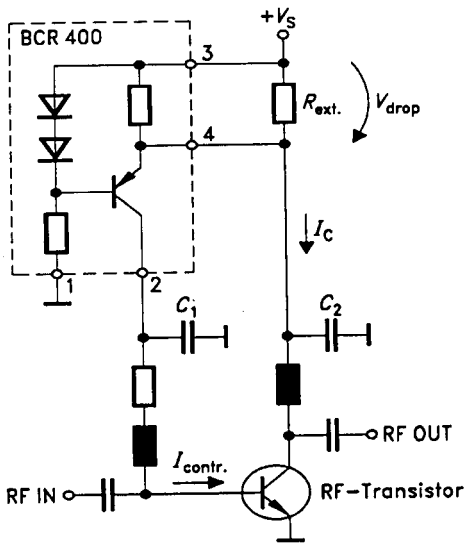


Figure 3-194 The Siemens BCR400 active bias controller can be applied to FETs as well as BJTs.

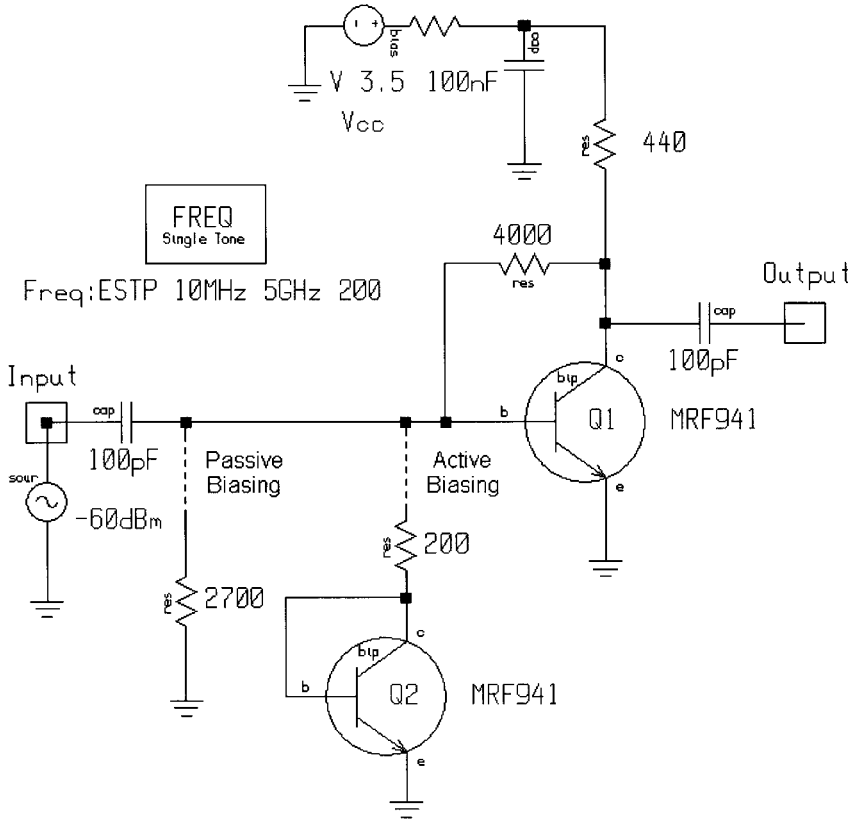


Figure 3-195 Amplifier stage (I_C = 3 mA) showing active and passive bias alternatives.

Q2, and proceeding from there. On the other hand, having a nice CAD tool that provides insight into the dc voltages and allows the addition of dc voltage probes makes life much easier. To prove that this circuit actually works, Figure 3-193 shows its frequency-dependent gain, matching, and noise figure performance. It is based on a Motorola application, which had excessive voltage drop for Q1 and which we have improved to be suitable for low-voltage operation [29].

Integrated active bias solutions are also available. Figure 3-194 shows the Siemens BCR400 active bias controller applied to a BJT; it can also be applied to FETs and TR switching diodes, as we saw in Figures 3-57 and 3-59, respectively.

3-10-3 dc Biasing of IC-Type Amplifiers

In integrated circuits, we do not have the unlimited flexibility described earlier. This results frequently in constant-current sources or transistors being used as diodes for biasing purposes. Our first suspicion is that this is going to cause a lot of noise in the IC; Figure 3-195 shows a test example.

As suspected, choosing the same bias point of 3 mA, the noise figure using active bias is quite a bit higher than the passive one, as shown in Figure 3-196.

3-11 PUSH–PULL/PARALLEL AMPLIFIERS

Since it is practically impossible to obtain $S_{11} = S_{22} = 0$, using a combiner at the input and output reduces the problem. Figure 3-197 shows the 3-dB hybrid approach. It has a practical

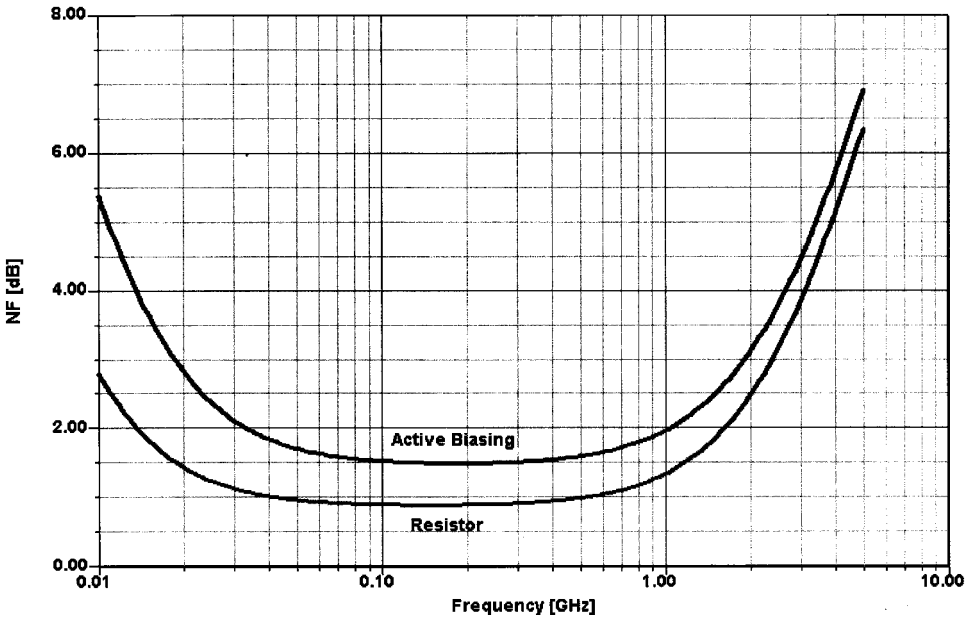


Figure 3-196 Frequency-dependent noise figure of the amplifier circuit shown in Figure 3-195. The active biasing curve is considerably noisier.

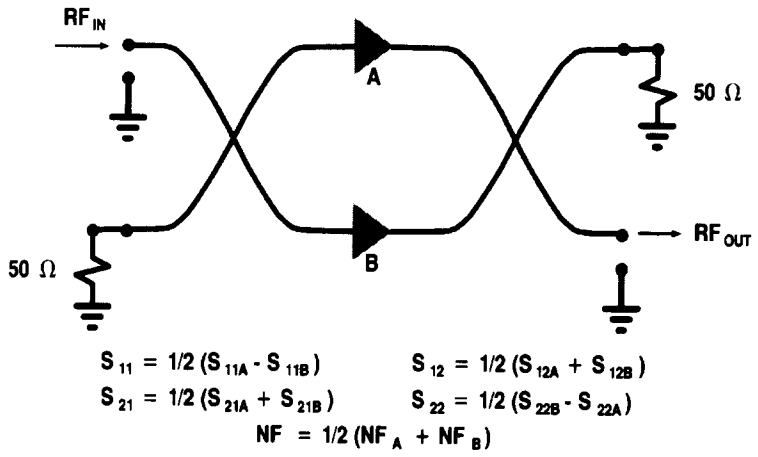


Figure 3-197 Combining power with 3-dB hybrid couplers.

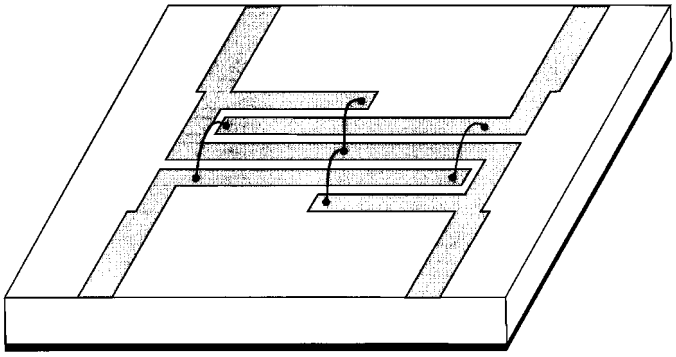


Figure 3-198 Four-strip version of the Lange coupler, a broadband 90° hybrid.

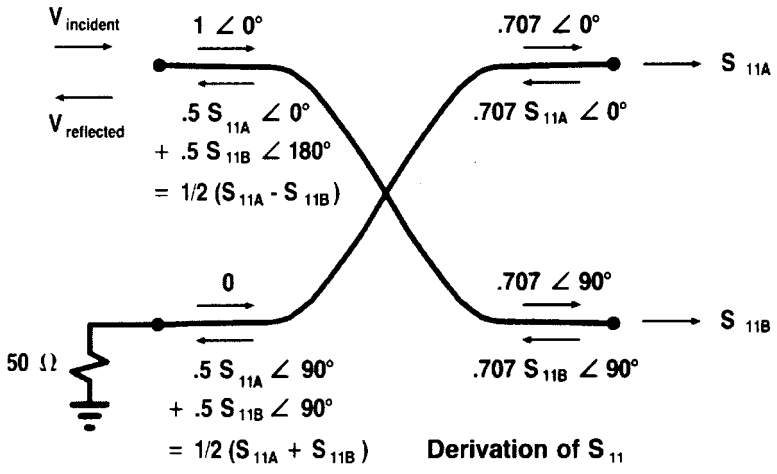


Figure 3-199 3-dB hybrids. This analysis method may be applied to any coupler.

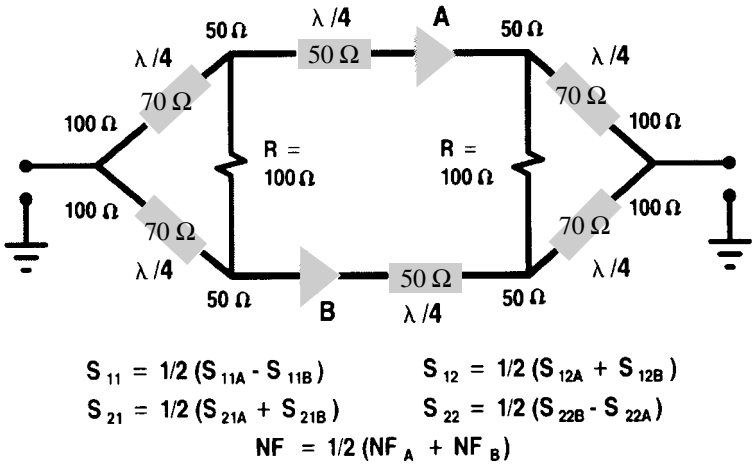


Figure 3-200 Wilkinson divider/combiners.

bandwidth limit of about 4:1; its advantage is that it is compact and ideal for cascading. If one stage opens, P_{OUT} and gain drop only 6 dB; however, it does not protect against load mismatch. A recommended distributed form is the Lange coupler (Figure 3-198). To analyze this approach, Figure 3-199 shows the resulting impedances and necessary phase shifts. If a Wilkinson coupler [30] is used instead of the 3-dB hybrid, the bandwidth is reduced. The Wilkinson coupler is larger in size than the 3-dB hybrid but more easily realized, leading to lower cost. Figure 3-200 shows the actual impedances necessary for matching, with A1 and A2 being the actual amplifiers combined.

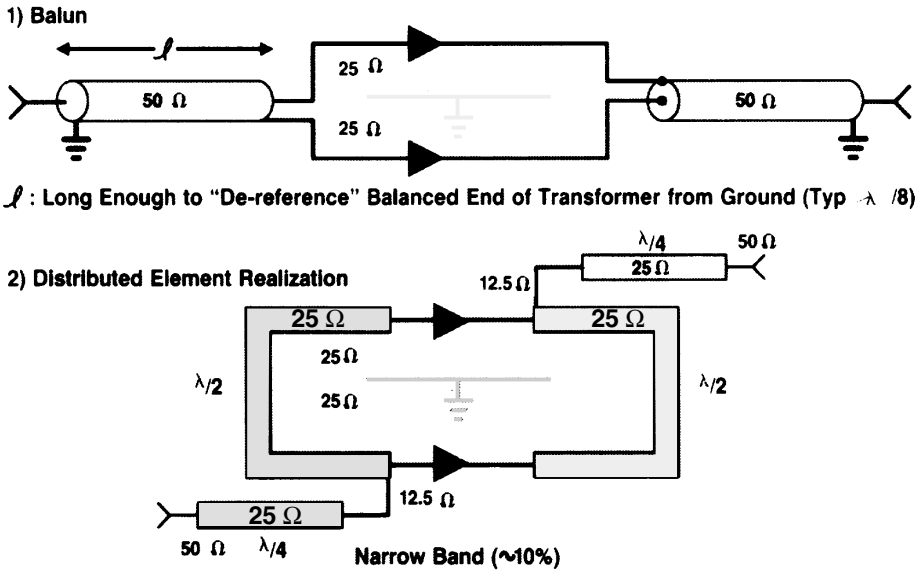


Figure 3-201 Push–pull amplifiers.

The solution for push–pull amplifiers resorts to a balun with the following results obtained: Bandwidths of greater than one decade are achievable; the balun structure gives 4:1 impedance advantage; there is no isolation; $S_{11} \neq 0$, $S_{22} \neq 0$; and cancellation of even-order products is provided.

Since most amplifiers have 50-Ω rather than 25-Ω impedances, Figure 3-201 shows a circuit that matches 25 Ω to 50 Ω on both sides. Since the standard amplifier operates at 50 Ω rather than 25 Ω, an additional network may be necessary to transform the impedance up to 50 Ω again. A similar problem occurs at the amplifier output.

3-12 POWER AMPLIFIERS

Power amplifiers are devices that are tasked to transform as much dc power into RF/microwave power as possible. Depending on the technology, devices are available that can produce at least 100 W up to 2 GHz. Since the transconductance of FETs at conducting angles of less than 180° becomes very small, the actual power gain sharply decreases compared to Class A operation while the power-added efficiency is optimized to values of 55% (Class B) or so. The dissipation of these devices is, of course, quite high, while in Class B operation this is even more of an issue. Power devices in Class C are dying to have big heat sinks to survive. One of the nice things about FETs is a much reduced thermal runaway effect compared to their bipolar brothers.

Power amplifier classes are defined in terms of the device conduction angle and/or the type of device operation:

Class A	"Linear" operation; 180° conduction angle
Class AB	"Linear" operation; conduction angle between 90° and 180°
Class B	"Linear" operation; 90° conduction angle
Class C	Fixed drive; less than 90° conduction angle
Class D	Switched operation, conduction range may vary with time from 0° to 180° or may be fixed.

There are also Classes E and F, which are currently under evaluation and showing some promising results.

Figure 3-202 shows Class C operation. The output of the amplifier is a current pulse train, which frequently leads to misunderstanding as to the calculation of its load. Given the fact that the amplifier is a current generator, it is silly to assume one has to do a conjugate matching looking backward into the amplifier. This is the condition for Class A operation. Unfortunately, this is still not understood by some engineers, but we hope we can clear this up today. The theoretical efficiency is 89.7% at a conducting angle of less than 180°; in practice, efficiencies more than 85% are not obtainable. (All of these efficiencies are those of the device itself; losses in matching networks are not taken into account in these efficiency values.) Considering optimal gain and optimal output power, the conducting angle actually settles somewhere between 100° and 140°. Again, the power gain in Class C operation is less than in Class A and Class B operation.

3-12-1 Example 1: 7-W Class C BJT Amplifier for 1.6 GHz

The following demonstrates how to design a 1.6-GHz, 7-W amplifier operating in Class C. We need to point out again that all modulation forms that contain the information in both

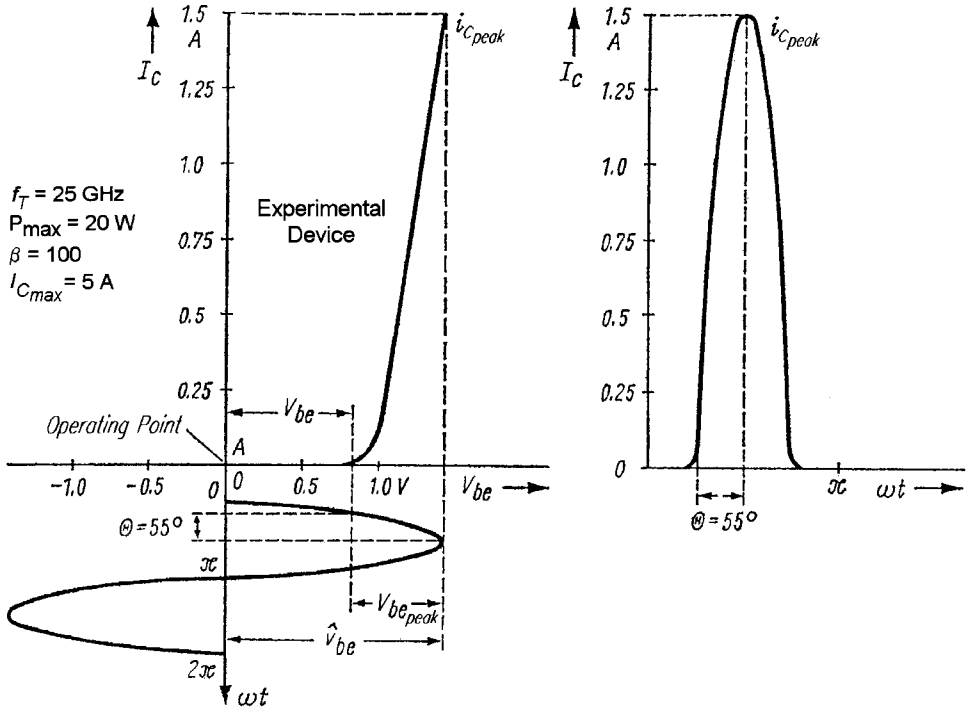


Figure 3-202 Output current as a function of drive voltage, taking into consideration the knee voltage of the base–emitter *pn* junction. The output is a series of current pulses with a duty cycle determined by the conduction angle.

amplitude and phase require more linearity than Class C affords. Needless to say, this affects the efficiency; therefore, at a later point we will also show a Class AB power amplifier. This is a tutorial, so we will present a UHF example and walk the reader through all the problems involved in the design. The selection of the transistor is done by first looking at the f_T, which in our case should be at least 4–5 GHz. The first order of business is to define the point of operation for the output amplifier. To do this, we need to determine the maximum voltage and current values supplied by the manufacturer. We have chosen a device that operates at 28 V and can sustain a peak collector current of 1.5 A. Figure 3-202 lets us determine the conducting angle Θ from the I_C/V_{BE} plane. This figure shows on the left side the I_C/V_{BE} plane and the necessary voltage to drive the amplifier.

Initially, we assume that the base–emitter junction is operated without any bias ($V_{BE} = 0$). The dc I–V curve shows that the collector current starts at about 650 mV. The amplifier itself, therefore, is already in Class C operation. For full output power, the drive voltage $V_{BE,\text{peak}}$ has to be 1.14 V. Therefore, the conducting angle can be determined by

$$\cos \Theta = \frac{V_{BE}}{\hat{\boldsymbol{v}}_{BE}} = \frac{0.65}{1.14} = 0.57$$

$$\cos 0.57 = 55^\circ$$

V_{BE} is the dc voltage at which the transistor starts conducting. From Figure 3-65 and Eqs. (3-160), (3-161), and (3-162), the fundamental current i_{c1} can be determined from

$$i_{c1} = i_{\text{peak}} \cdot f_1(\Theta)$$

or

$$i_{c1} = 1.5 \text{ A} \times 0.36 = 0.54 \text{ A}$$

The average collector dc current can be determined from

$$i_{c2} = i_{\text{peak}} \cdot f_2(\Theta)$$

or

$$i_{c2} = 1.5 \text{ A} \times 0.21 = 0.315 \text{ A}$$

The collector saturation voltage $V_{CE(\text{sat})}$, unfortunately, is around 2–3 V; for the purpose of the calculation, we assume that the saturation voltage is 3 V, and, therefore, we can assume that the maximum collector ac voltage $\hat{v}_{CE} \approx (V_B - 3) = 25$.

The appropriate collector load has to be

$$R_C = \frac{\hat{v}_{CE}}{i_{C1}} = \frac{25}{0.54} = 46.3 \ \Omega$$

The maximum output power is

$$P_{\text{out}} = \frac{(V_{CE} - V_{CE(\text{sat})})^2}{2R_C} = \frac{625}{92.6} = 6.75 \text{ W}$$

The only way to increase the power would be to have a smaller collector load at a higher current. In this family of Motorola transistors, there was no 10-W transistor; that's why we have chosen a 20-W experimental device, which for this purpose is overkill. The theoretical efficiency of the amplifier is defined as

$$\eta = \frac{P_{\text{out}}}{P_{\text{in}}} = \frac{6.75 \text{ W}}{0.315 \times 28} = 85\%$$

If the saturation voltage would be less, our efficiency would automatically increase, since for the given load, the actual output power would increase (6.75 W $\rightarrow$ 7.5 W). Assuming that this stage would operate at 1900 MHz with a gain of roughly 22 dB, the required drive would be significantly less than the output power, or roughly 40 mW. For the purpose of the following calculations, we will assume that the drive power is 0.7 W. The total dissipation of the device now adds up to be

$$P_T = P_{\text{in(dc)}} - P_{\text{out}} + P_{\text{drive}}$$

$$P_T = (28 \times 0.315) - 6.75 + 0.7 = 2.77$$

Power-added efficiency is derived from this. Knowing the power dissipation, one can calculate a thermal resistance R_{therm} of the heat sink and the transistor:

$$R_{\text{therm}} = \frac{T_J - T_A}{P_T}$$

Where T_J is the device junction temperature and T_A is the ambient temperature. For a maximum ambient temperature of 60 °C and a maximum junction temperature of 150 °C, the thermal resistance is

$$R_{\text{therm}} = \frac{120 - 60}{2.77} = 21.7\ ^\circ\text{C/W}$$

R_{therm} is the sum of thermal resistance of the transistor and its package, and the thermal resistance between the transistor package and the heat sink, and the thermal resistance between the heat sink and the ambient air. Therefore,

$$R_{\text{therm}} = R_{\text{th(case)}} + R_{\text{th(case/heat sink)}} + R_{\text{th(heat sink)}}$$

The required thermal resistance of the heat sink is calculated from

$$R_{\text{th(heat sink)}} = 21.7 - 10 - 1.4 = 15.6\ ^\circ\text{C/W}$$

Assuming an aluminum heat sink (a 2-mm-thick plate), the required heat-sink surface area A (cm^2) is

$$A = \frac{1}{\alpha R_{\text{th(heat sink)}}}; \quad \alpha \approx 1.5\ \frac{\text{mW}}{^\circ\text{Ccm}^2}$$

Therefore, the required area is

$$A = \frac{1}{1.5 \times 0.0156} = 42.7\ \text{cm}^2$$

The next task is to match both the input and output to the (most likely) 50-Ω source and termination. The following section will guide us through the procedure. The input impedance Z_{11} (after some lengthy search) was found to be $1 + j10\ \Omega$. This means that the input consists of the base-spreading resistance (approximately 1 Ω) and an input capacitance, which because of phase shift now turns out to be inductive. Since the popular method for input matching these days is a pi configuration, we are going to use a simple pi filter both for input and output matching. There are several equations in the literature that give the values for the two capacitors and inductors; most of them result in an incorrect resonance frequency. Since we have access to software tools, we decided to take three approximate values and use the optimizer to give us the appropriate matching. The only remaining variable in such an approach is the bandwidth. If we specify $S_{21} = 1$ and $S_{11} = 0$ (100% energy transfer and perfect return loss), we will realize that this is possible at only one frequency. We therefore need to identify a frequency band or bandwidth over which an acceptable match is needed. As the bandwidth increases, the match for such a simple circuit will get poorer; there are not enough elements available to do better. In this case, we took the center frequency ±100 MHz.

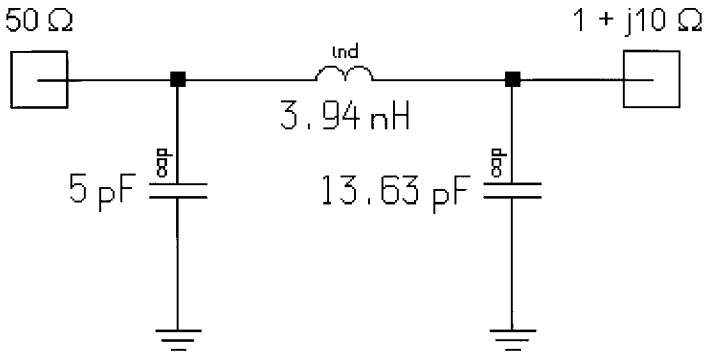


Figure 3-203 Input matching network for the amplifier used in this tutorial.

Figures 3-203 and 3-204 show the actual circuit with the element values and the resulting input match. Likewise, we did the same thing at the output, but the actual resulting dynamic output capacitance is not known *a priori*. For low frequencies, the integral equation gives a value of $2C_{CE}$; however, practical tests in frequencies above 500 MHz show that the actual increase is only 1.3 times the dc output capacitance. Therefore, we decided to design the filter with no output capacitance assumed, and the input capacitance of about 5 pF must be reduced for proper reactive matching. Mathematically, this could result in a negative number, specifically if the resulting output capacitance exceeds the 5-pF value. To compensate for this, one needs to use an RF choke that tunes out the resulting dynamic output capacitance and then one can revert to the original 5 pF as calculated.

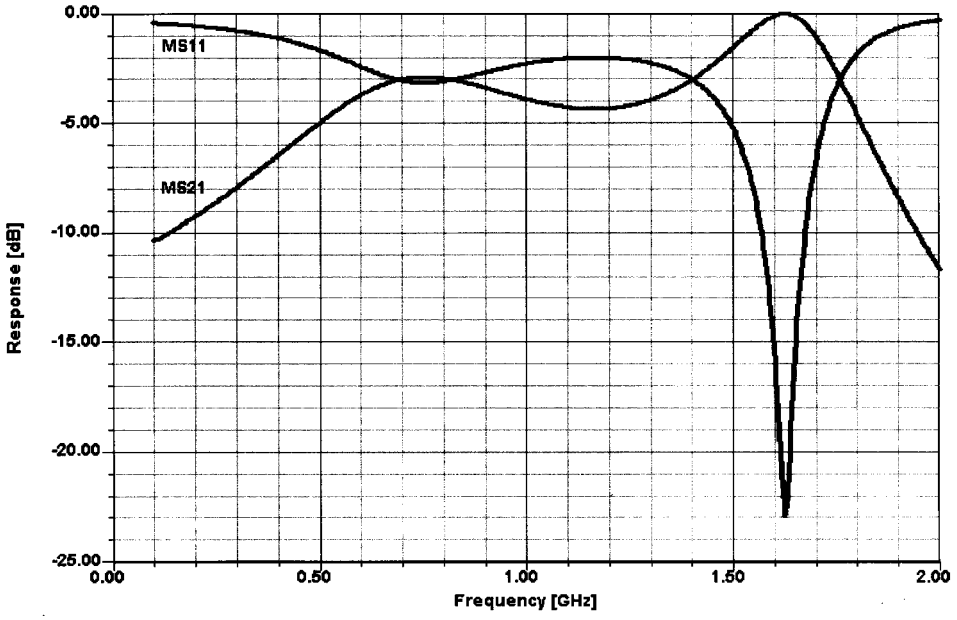


Figure 3-204 Frequency response of the input matching network.

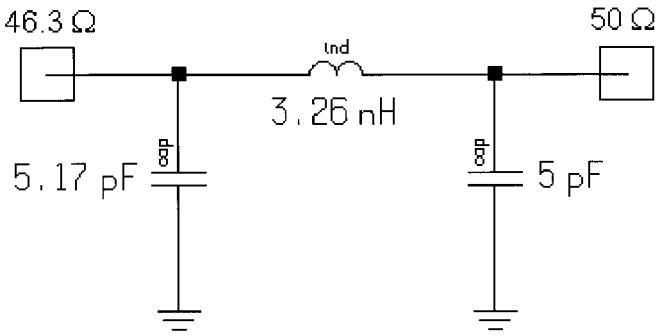


Figure 3-205 Output matching network for the amplifier.

The selection of the actual component is another issue, because these capacitors need to handle several amperes of RF current. It is not always easy to find capacitors with the appropriate value that can handle the current. The inductors can be printed. The paragraph following this section therefore will deal with the issue of how to translate the lumped values into distributed values, which will make the current handling much easier (we hope).

For this, we did the same optimizing approach by using a 50-Ω source and a 46.3-Ω load. Figures 3-205 and 3-206 show the circuit with the appropriate values and its simulated frequency-dependent responses.

The next step is to connect the input and output matching network to an "actual" transistor. This is a point where difficulties arise, since very few, if any, companies provide SPICE-type

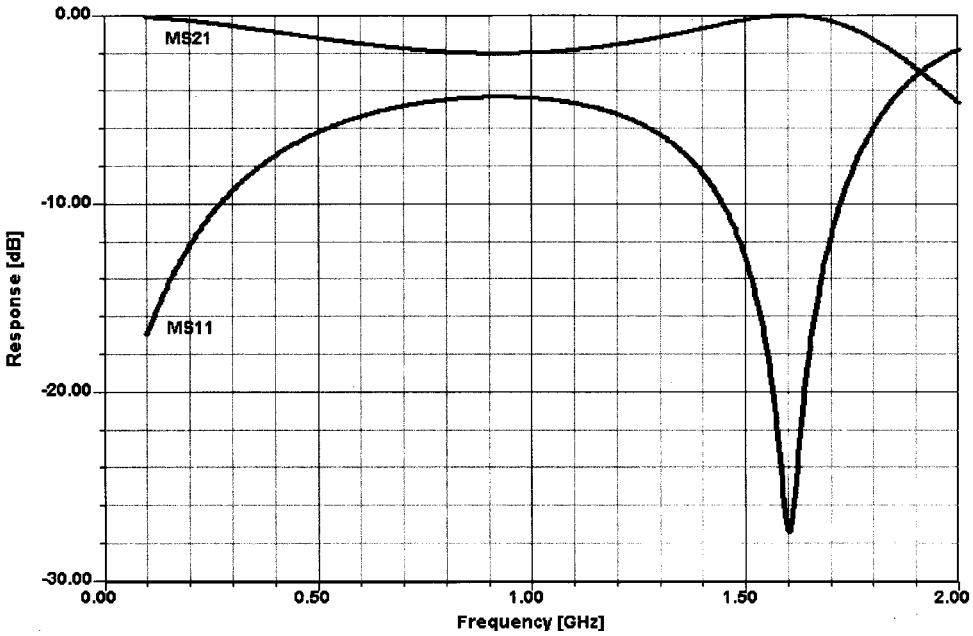


Figure 3-206 Frequency response of the output matching network.

parameters for transistors operating at currents of more than 200 mA. Figures 3-207 and 3-208 show the overall schematic and frequency response of the actual amplifier with all filters connected. Figure 3-209 details the amplifier's peak gain response. If more bandwidth is needed, the matching circuit must be increased in complexity, as will be shown in the following section.

Now is a good time to look at the final results of our simulation, which is based on the various assumptions as outlined above. Figure 3-210 shows the dc *I–V* curves with a straight load line and no reactances. This was accomplished by using a wideband transformer instead of an output matching circuit. Figure 3-211, however, shows the "load line" that results when the transformer is replaced with a real matching network. Based on the harmonics content and its improper termination, one can see how the output load line opens and shows ringing in the low-current area. If we expect to see the saturation voltage, we will be disappointed because the saturation voltage is frequency dependent (increases with frequency) and this figure gives an erroneous impression. Hence, we show dc *I–V* curves. To display the conduction angle, which we showed previously, we examine the output current as a function of time at a drive level of +16.5 dBm (Figure 3-212). During the negative half, the transistor is not conducting as predicted, so this validates our assumptions. Next, we are interested in determining the power gain. Therefore, we look at the output spectrum for a drive level of +16.5 dBm (Figure 3-213). It shows an output level of 38.6 dBm. If we subtract the drive level (16.5 dBm) from this, we see that the power gain is 22.1 dB at this drive level. Since we are used to thinking in watts, Figure 3-214 shows the absolute output power and the harmonics prior to filtering. Figure 3-215 shows the output spectrum after replacing the transformer with a simple pi output network.

During the conduction period, there will be a voltage across the transistor 180° out of phase, as shown in Figure 3-216. Finally, to highlight the nonlinearity of a Class C amplifier, Figure 3-217 shows the amplifier's output power versus drive, and Figure 3-218 shows the amplifier's gain versus drive. When the transistor turns on is clearly evident.

A particular difficulty arises in the case of FETs, where the charge energy makes for a difference between dc and pulsed measurements. Figure 3-219 shows this. Some of the more modern transistors, such as SiGe types, may need some pulsing also; as can be seen in Figure 3-219, even depending on the pulsing, one may get different dc *I–V* curves. This area is still under investigation, and even the "experts" do not agree on all aspects of this. In real life, it turns out that a lot of experimentation is the answer. Anyone who thinks that a pure CAD approach will give the right design answers immediately will encounter many surprises! This is not a fault of the CAD tools—assuming that they are properly engineered—but rather the availability of appropriately "tweaked" models for high-power applications, and the willingness of device suppliers to standardize on certain models and support them.

Because of the difficulties involved in cooling transistors under test and simulating RF pulse and CW conditions, depending on the application, manufacturers have so far stayed away from investing money in this area. The designer must therefore resort to application notes published by device manufacturers. Because even slight layer changes can have a large impact on the frequency response of a given device, application-note-based designs can run into difficulty if a device's fabrication has significantly evolved since the publication of its application notes.

Given the multitude of models outlined in Chapter 2, the CAD user is sometimes left hanging as to which is the best model to use. The latest indications are that the MEXTRAM (*m*ost *e*xquisite *t*ransistor *m*odel), developed by the University of Delft, Holland, with Philips of Holland, shows the most promising result, specifically in the area of IMD products. While

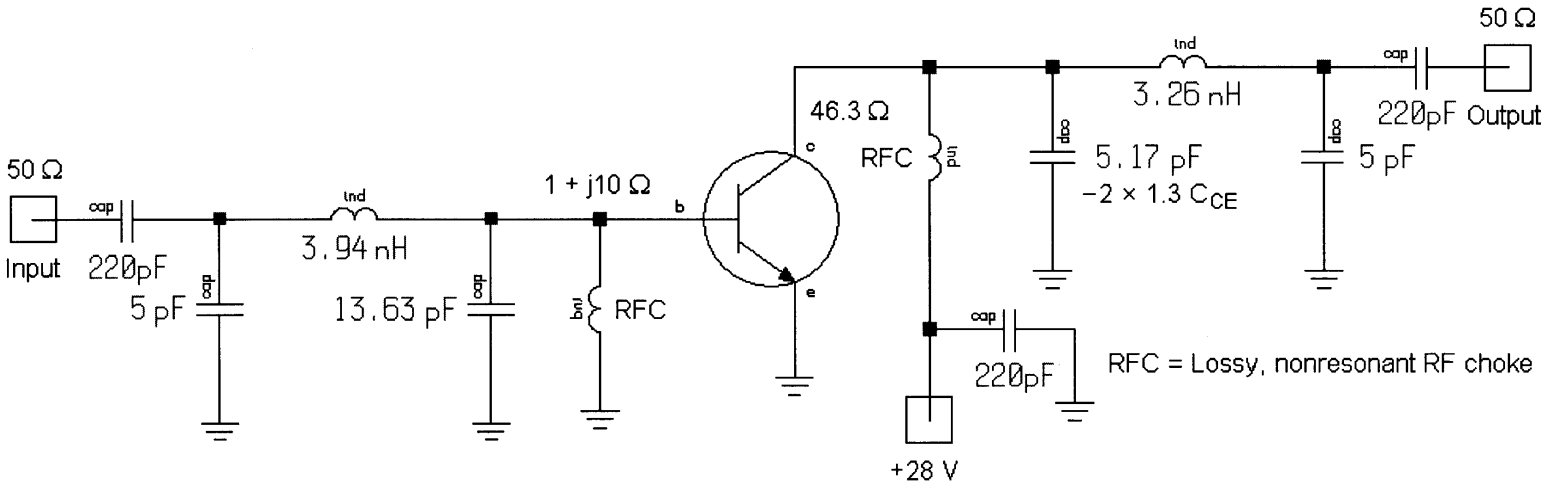


Figure 3-207 Actual schematic of our experimental 1.6-GHz amplifier. The RF chokes have enough RF losses not to be resonant and yet don't load the circuit. This is frequently accomplished with either ferrite beads or resistors in parallel with the chokes.

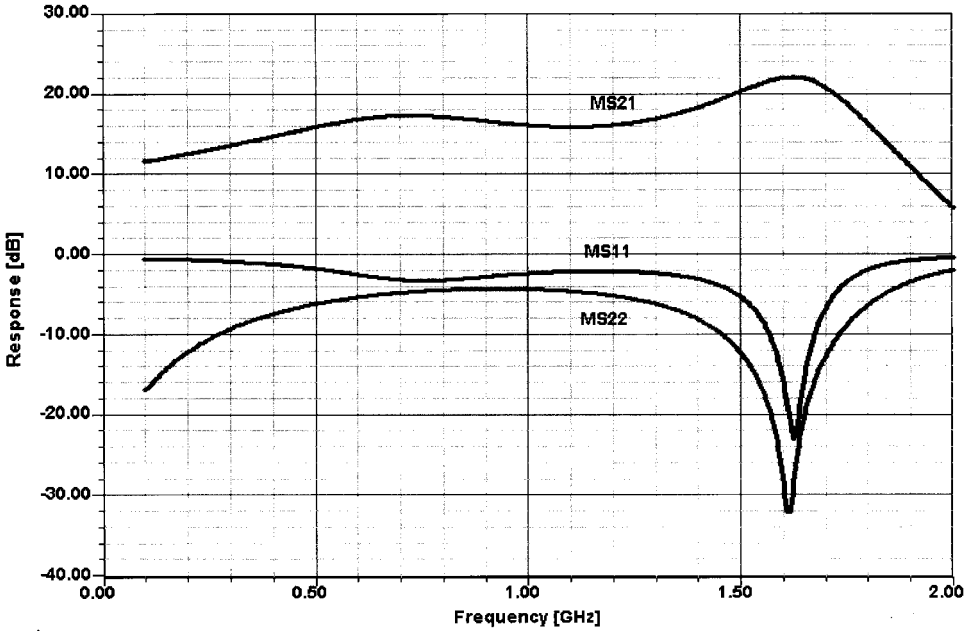


Figure 3-208 Overall frequency response of our experimental amplifier. Input and output matching, and gain, are shown as a function of frequency.

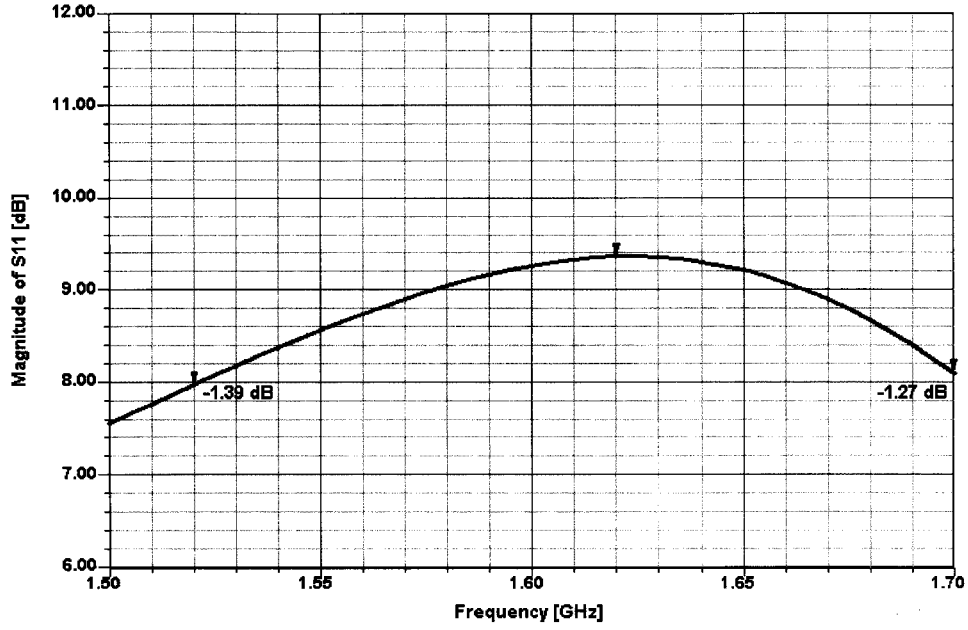


Figure 3-209 Close-in display of the frequency-dependent gain.

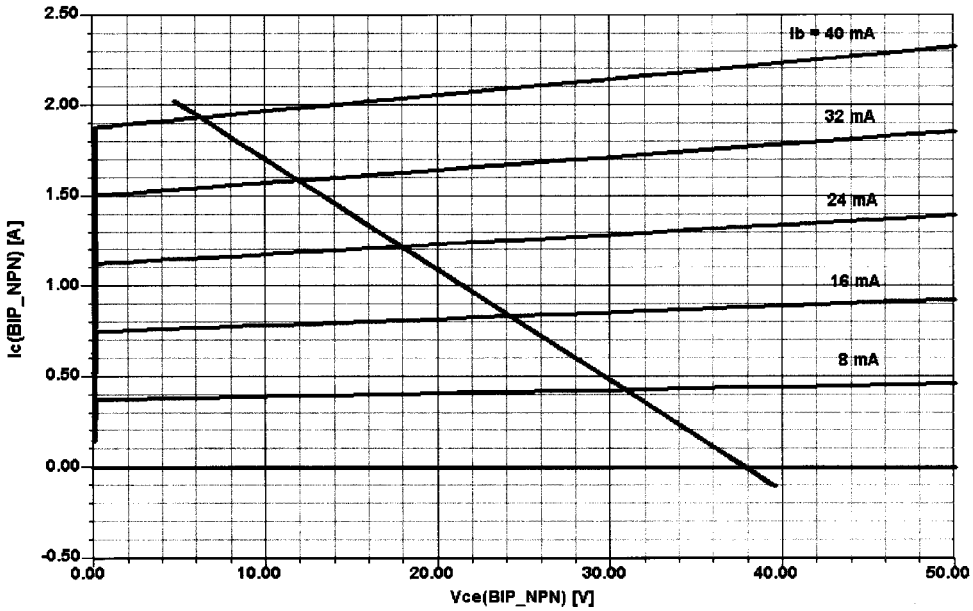


Figure 3-210 The dc *I–V* curves assuming a perfect resistive load. No RF saturation voltage or other RF saturation effects are evident.

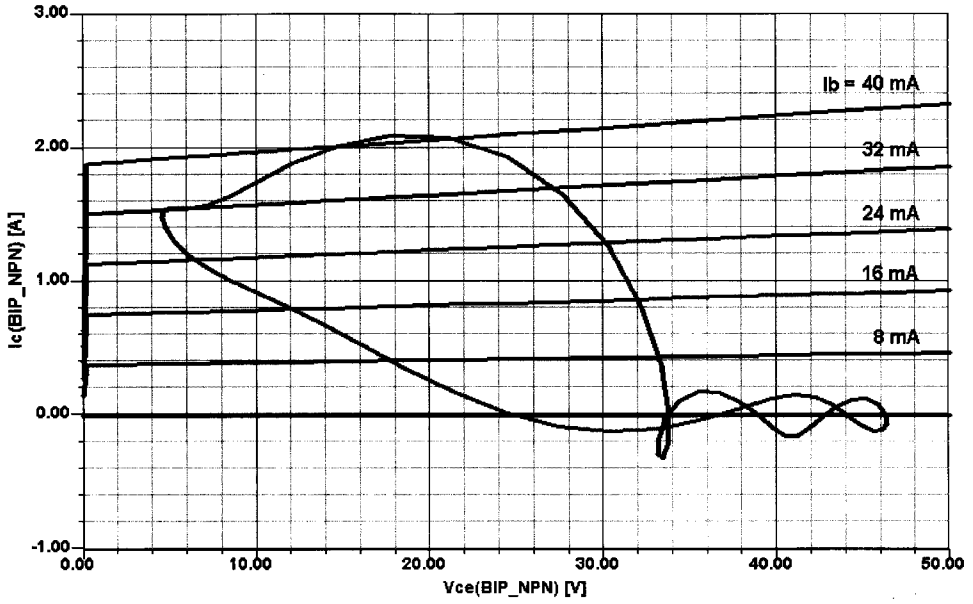


Figure 3-211 Load line with the output transformer replaced with a real matching circuit.

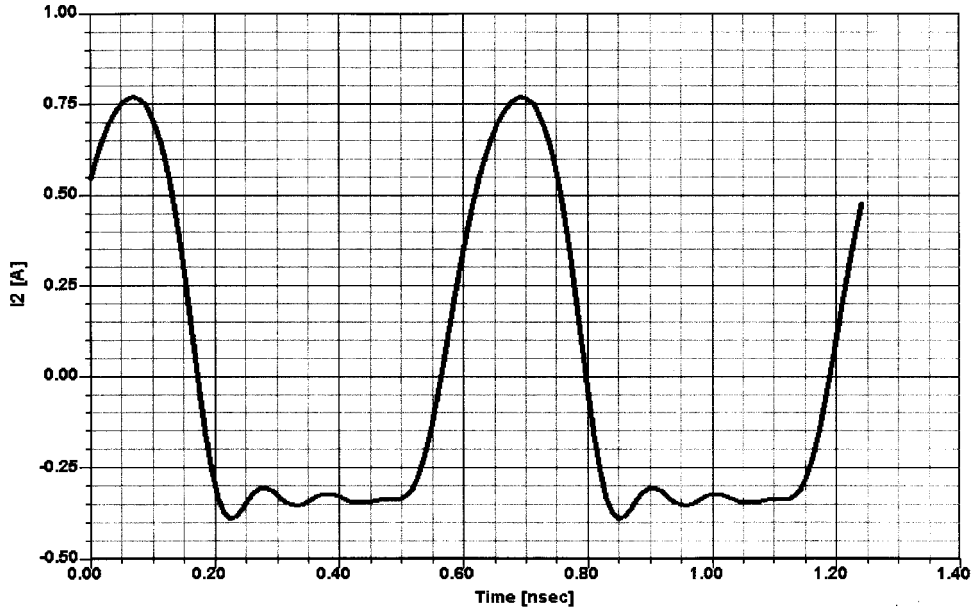


Figure 3-212 Output current as a function of time. The duty cycle is determined by the device conduction angle.

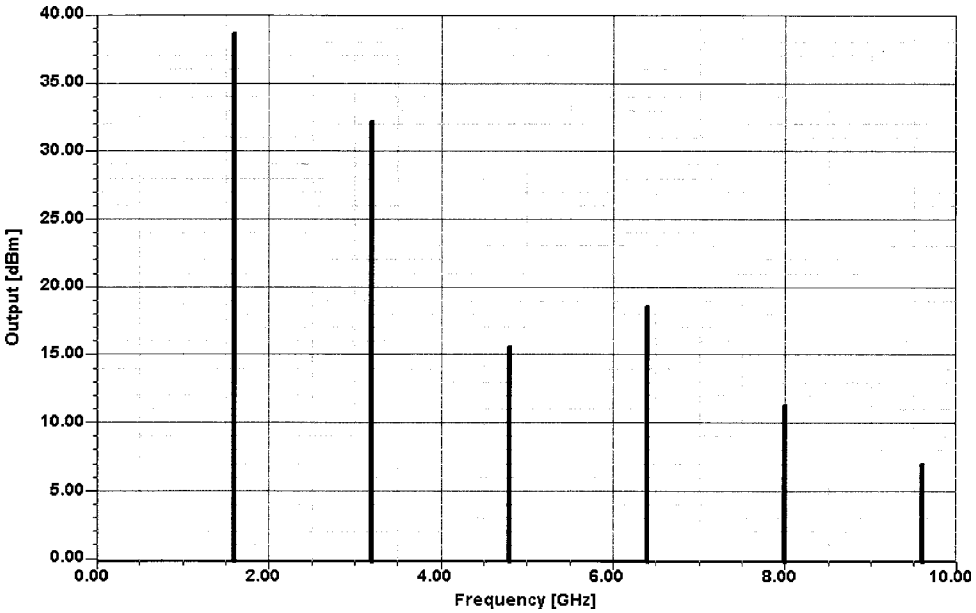


Figure 3-213 Power output at the collector prior to any filtering, assuming a real output termination.

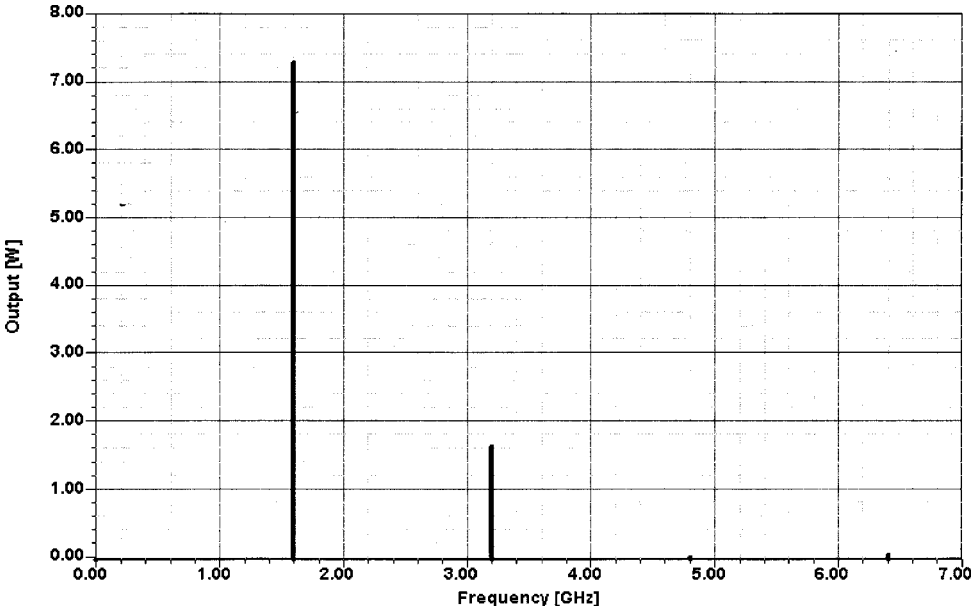


Figure 3-214 Available output power in watts, prior to any filtering, assuming a real output termination.

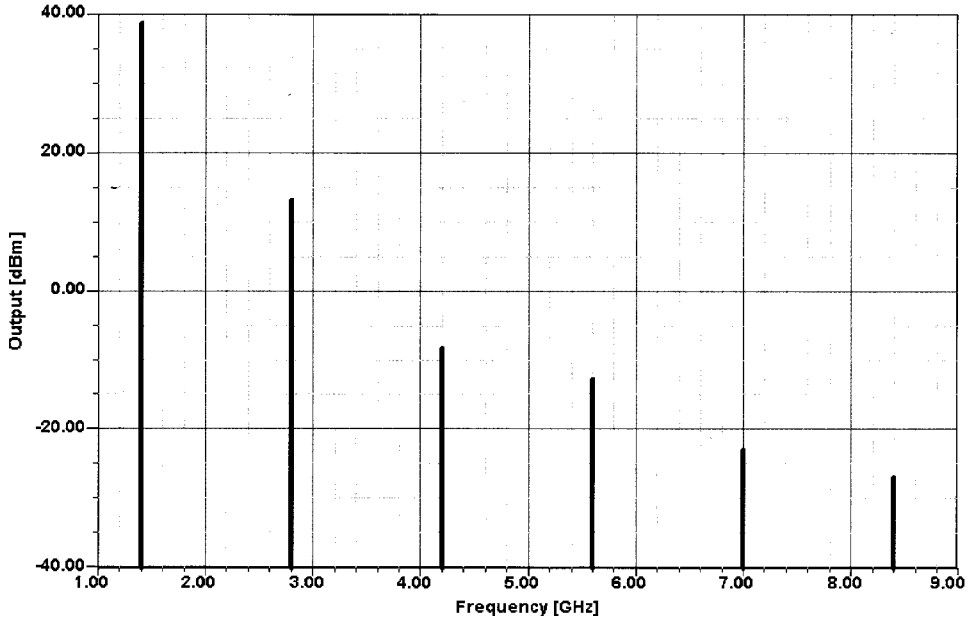


Figure 3-215 The output spectrum after replacing the transformer with a simple pi output network.

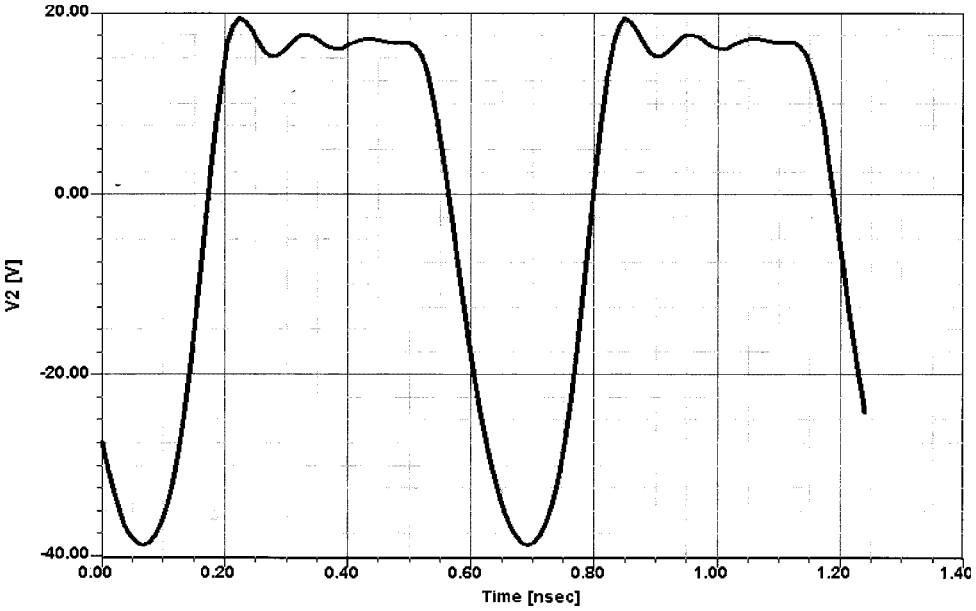


Figure 3-216 Output voltage at the collector. The voltage is asymmetrical compared to 0; this is important in the selection of a device capable of handling these voltages.

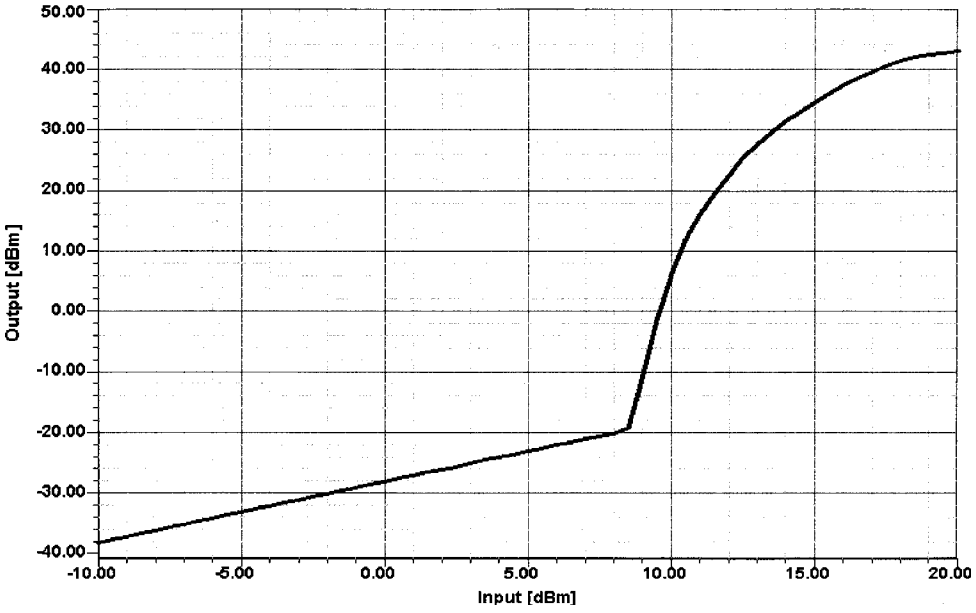


Figure 3-217 Output power versus input power for the Class C amplifier. Until the transistor turns on (at a drive level of about 8.5 dBm), only fedthrough power appears at the output.

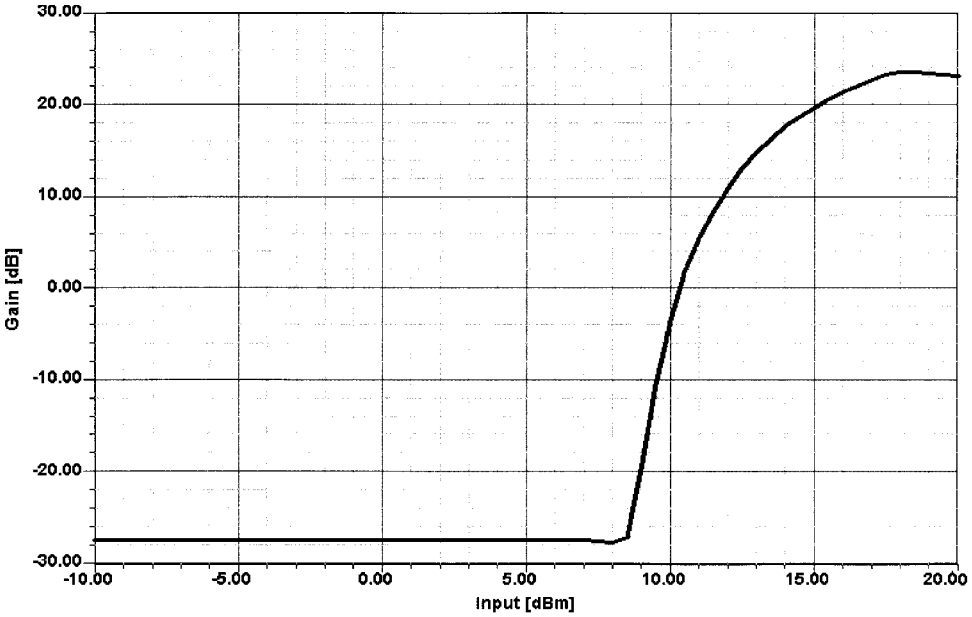


Figure 3-218 Amplifier gain as a function of drive. It is evident that one needs to cross the 0.65-V point for the transistor to become conductive and show gain. The initial amplification as a function of drive shows a very steep curve that then levels off and drops as the amplifier moves into saturation.

other later models have up to seven internal nodes, MEXTRAM fortunately uses only five. (The MEXTRAM model actually consists of three transistors, which makes modeling fun.) Power-added efficiency simulation is not available, since the MEXTRAM model was developed for small-signal applications, meaning circuits that operate at less than 100 mW of RF output power.

As if the absence of trustworthy high-power device models is not challenge enough, in developing this 1.6-GHz Class C amplifier example, we ran into another complication: Using a CAD optimizer to simultaneously match the circuit's input and output networks failed because of the device's input-impedance variation with drive. Every matching adjustment that improved the drive to the device resulted in an input-impedance change that reduced the drive to the device! In the end, the optimizer couldn't keep up. In actually building such a circuit, one would make the matching-network values adjustable on the initial breadboard and then, with the necessary input drive level applied, iteratively adjust the input and output matching for the desired output level. It appears that the CAD program could not handle such a degree of nonlinearity.

We then went to the approach of using large-signal *S* parameters at a drive level slightly less than 16 dBm, but this detuned the input filter, resulting in a poorer input match than the one initially predicted. Likewise, the output matching, which in a power stage must result in maximum output power rather than a conjugate match, was less impressive. The large-signal *S* parameters at the output for a power amplifier are somewhat dubious because Class C operation generates so many harmonics that the output termination will reflect a lot of energy and, depending on the output filter topology, different results can be expected. We firmly believe that there will be further mathematical improvement in the CAD tools, and we very much invite

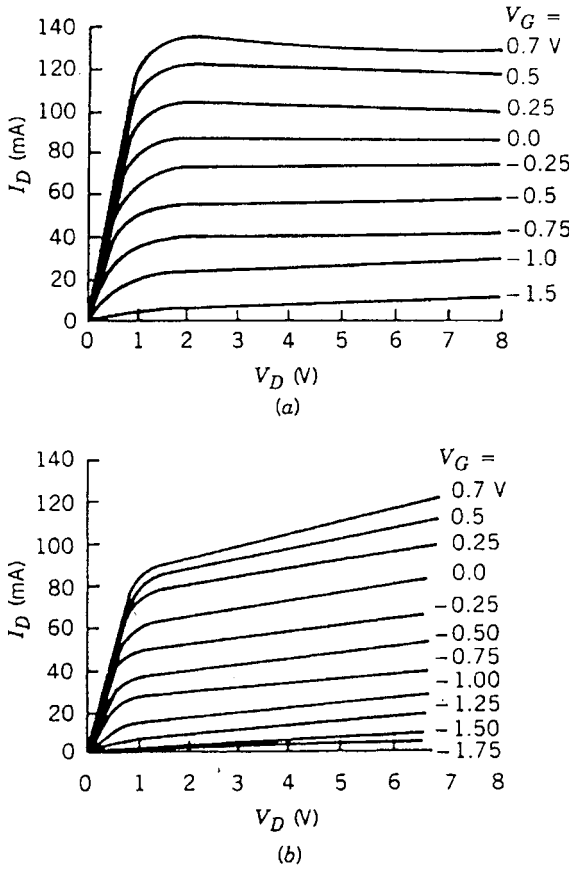


Figure 3-219 Comparison of FET drain current characteristics obtained from (a) pulsed *I–V* and (b) rectified sine-wave evaluation. The rectified sine-wave curves can be considered equivalent to curves obtained at dc.

our readers to use such a nonlinear example to test their tools, provided that a high-power model becomes available in this power range.

It has been rumored that the load-pull technique, rather than the application of large-signal *S* parameters, is a vehicle to solve such matching problems. However, the load-pull technique is really applied only to the fundamental frequency and does not deal with the issue of harmonics of the output current. More details about the load-pull technique and its capabilities are found in Appendix B.

Actual measurements of amplifiers of this type show that our simulated efficiency and drive level come quite close to reality, but at a frequency somewhat offset from that chosen for the design. As an example, the input and output filters tuned at 1.4 GHz instead of 1.6 GHz using the large-signal *S* parameters. Again, an iterative process, similar to manually trimming a prototype, will give a similar result. This is consistent with our earlier statement that most final designs are really hand-tweaked, and that CAD can only bring us close, but not all the way, to a good solution. For passive problems, the CAD does a significantly better job; the nonlinear cases are the ones that still cause headaches.

3-12-2 Impedance Matching Networks Applied to RF Power Transistors*

Introduction. Some graphical and numerical methods of impedance matching will be reviewed here. The examples given refer to high-frequency power amplifiers.

Although matching networks normally take the form of filters and therefore are also useful to provide frequency discrimination, this aspect will be considered only as a corollary of the matching circuit.

Matching is necessary for the best possible energy transfer from stage to stage. In RF power transistors, the input impedance is of low value, decreasing as the power increases or as the chip size becomes larger. This impedance must be matched either to a generator—of generally 50-Ω internal impedance—or to a preceding stage. Impedance matching has to be made between two complex impedances, which makes the design still more difficult, especially if matching must be accomplished over a wide frequency band.

Device Parameters

Input Impedance. The general shape of the input impedance of RF power transistors is as shown in Figure 3-220. It is a large-signal parameter, expressed here by the parallel combination of a resistance R_p and a reactance X_p [31].

The equivalent circuit shown in Figure 3-221 accounts for the behavior illustrated in Figure 3-220. With the presently used stripline or flange packaging, most of the power devices for VHF low band will have their R_p and X_p values below the series resonant point f_s. The input impedance will be essentially capacitive.

Most of the VHF high-band transistors will have the series-resonant frequency within their operating range; that is, they will be purely resistive at one single frequency f_s, which the parallel-resonant frequency f_p will be outside.

Parameters for 1- or 2-GHz transistors will be beyond f_s and approach f_p. They show a high value of R_p and X_p with inductive character.

A parameter that is very often used to judge on the broadband capabilities of a device is the input Q or Q_{IN}, defined simply as the ratio R_p/X_p. Practically, Q_{IN} ranges around 1 or less for VHF devices and around 5 or more for microwave transistors.

Q_{IN} is an important parameter to consider for broadband matching. Matching networks normally are low-pass or pseudo-low-pass filters. If Q_{IN} is high, it can be necessary to use bandpass filter-type matching networks and to allow insertion losses. But broadband matching is still possible. This will be discussed later.

Output Impedance. The output impedance of RF power transistors, as given by all manufacturers' datasheets, generally consists of only a capacitance C_{OUT}. The internal resistance of the transistor is supposed to be much higher than the load and is normally neglected. In the case of a relatively low internal resistance, the efficiency of the device would decrease by the factor

$$1 + R_L/R_T \tag{3-192}$$

*Based on Motorola application note AN-721 ("Impedance Matching Networks Applied to RF Power Transistors"). Copyright of Motorola; used by permission.

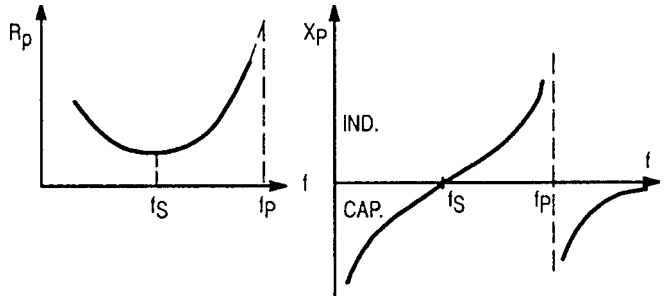


Figure 3-220 Input impedance of RF power transistors as a function of frequency.

where R_L is the load resistance, seen at the collector–emitter terminals, and R_T is the internal transistor resistance equal to

$$\frac{1}{\omega_T (C_{TC} + C_{DC})} \tag{3-193}$$

defined as a single parameter, where ω_T is the transit angular frequency and C_{TC} and C_{DC} are the transition and diffusion capacitances, respectively, at the collector junction.

The output capacitance, C_{OUT}, which is a large-signal parameter, is related to the small-signal parameter C_{CB}, the collector–base transition capacitance.

Since a junction capacitance varies with the applied voltage, C_{OUT} differs from C_{CB} in that it has to be averaged over the total voltage swing. For an abrupt junction and assuming certain simplifications, $C_{\text{OUT}} = 2C_{CB}$.

Figure 3-222 shows the variation of C_{OUT} with frequency. C_{OUT} decreases partly due to the presence of the collector lead inductance, but mainly because of the fact that the base–emitter diode does not shut off anymore when the operating frequency approaches the transit frequency, f_T.

Output Load. In the absence of a more precise indication, the output load R_L is taken as

$$R_L = \frac{[V_{CC} - V_{CE(\text{sat})}]^2}{2P_{\text{OUT}}} \tag{3-194}$$

with $V_{CE(\text{sat})}$ equal to 2 or 3 V, increasing with frequency.

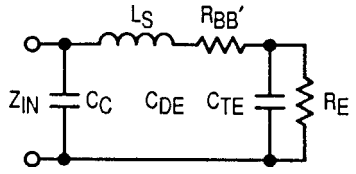


Figure 3-221 Equivalent circuit for the input impedance. R_E is the emitter diffusion resistance; C_{DE} and C_{TE} are the diffusion and transition capacitances, respectively, of the emitter junction; $R_{BB'}$ is the base-spreading resistance; C_C is the package capacitance; and L_S is the base lead inductance.

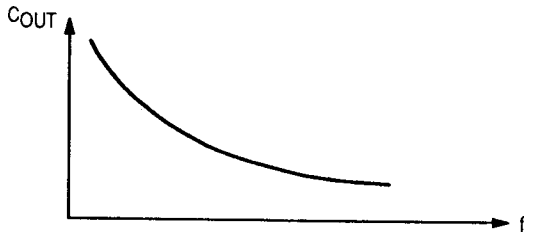


Figure 3-222 Output capacitance C_{OUT} as a function of frequency.

The above equation just expresses a well-known relation, but it also shows that the load, in first approximation, is not related to the device, except for $V_{CE(sat)}$. The load value is primarily dictated by the required output power and the peak voltage; it is not matched to the output impedance of the device.

At high frequencies this approximation becomes less exact, and for microwave devices the load that must be presented to the device is indicated on the datasheet.

Strictly speaking, impedance matching is accomplished only at the input. Interstage and load matching are more impedance transformations of the device input impedance and of the load into a value R_L (sometimes with additional reactive component) that depends essentially on the power demanded and the supply voltage.

Matching Networks. In the following, matching networks will be described by the order of complexity. These are ladder-type reactance networks. The different reactance values will be calculated and determined graphically. Increasing the number of reactances broadens the bandwidth. However, matching networks consisting of more than four reactances are rare. Above four reactances, the improvement is small.

Numerical Design

TWO-RESISTANCE NETWORKS. Resistance terminations will first be considered. Figure 3-223 shows the reactive L section and the terminations to be matched.

Matching or exact transformation from R_2 into R_1 occurs at a single frequency f_0. At f_0, X_1 and X_2 are equal to

$$X_1 = \pm R_1 \sqrt{\frac{R_2}{R_1 - R_2}} = R_1 \frac{1}{\sqrt{n-1}} \tag{3-195}$$

$$X_2 = \mp \sqrt{R_2(R_1 - R_{2)}} = R_1 \frac{\sqrt{n-1}}{1} \tag{3-196}$$

At f_0, $X_1X_2 = R_1R_2$. X_1 and X_2 must be of opposite sign. The shunt reactance is in parallel with the larger resistance.

The frequency response of the L section is shown in Figure 3-224, where the normalized current is plotted as a function of the normalized frequency.

If X_1 is capacitive and consequently X_2 is inductive, then

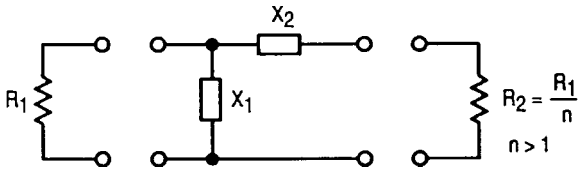


Figure 3-223 Two-reactance matching network.

$$X_1 = -\frac{f_0}{f} R_1 \sqrt{\frac{R_2}{R_1 - R_2}} = -\frac{f_0}{f} R_1 \frac{1}{\sqrt{n-1}} \tag{3-197}$$

and

$$X_2 = \frac{f}{f_0} \sqrt{R_2(R_1 - R_2)} = \frac{f}{f_0} R_1 \frac{\sqrt{n-1}}{1} \tag{3-198}$$

The normalized current absolute value is equal to

$$\left|\frac{I_2}{I_0}\right| = \frac{2\sqrt{n}}{\sqrt{(n-1)^2 \cdot (f/f_0)^4 - 2\,(f/f_0)^2 + (n+1)^2}} \tag{3-199}$$

where $I_0 = \sqrt{n}E/2R_1$ and is plotted in Figure 3-224 [32].

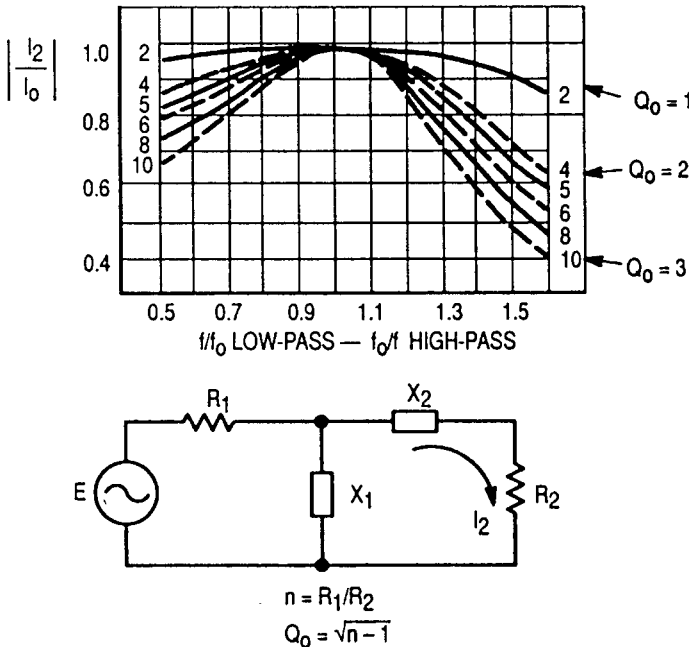


Figure 3-224 Normalized frequency response for the L section in low-pass or high-pass form.

If X_1 is inductive and consequently X_2 is capacitive, the only change required is a replacement of f by f_0 and vice versa. The L section has low-pass form in the first case and high-pass form in the second case.

The Q of the circuit at f_0 is equal to

$$Q_0 = \frac{X_2}{R_2} = \frac{R_1}{X_1} = \sqrt{n-1} \tag{3-200}$$

For a given transformation ratio n, there is only one possible value of Q. On the other hand, there are two symmetrical solutions for the network that can be either a low-pass filter or a high-pass filter.

The frequency f_0 does not need to be the center frequency, $(f_1 + f_2) / 2$, of the desired band limited by f_1 and f_2. In fact, as can be seen from the low-pass configuration of Figure 3-224, it may be interesting to shift f_0 toward the high-band edge frequency f_2 to obtain a larger bandwidth w, where

$$w = \frac{2(f_1 + f_2)}{f_2 - f_1} \tag{3-201}$$

This will, however, be at the expense of poor harmonic rejection.

Example. For a transformation ratio of $n = 4$, we can determine the following values from the above relations:

Bandwidth w	0.1	0.3
Maximum insertion losses	0.025	0.2
X_1/R_1	1.730	1.712

If the terminations R_1 and R_2 have a reactive component X, the latter may be taken as part of the external resistance as shown in Figure 3-225. This compensation is applicable as long as

$$Q_{int} = \frac{X_{\text{int}}}{R_2} \quad \text{or} \quad \frac{R_1}{X_{\text{int}}} < n - 1 \tag{3-202}$$

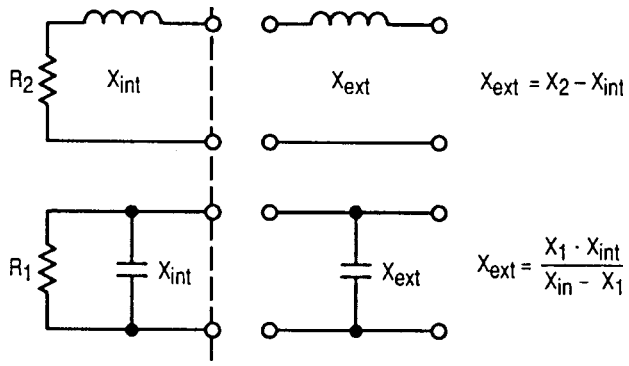


Figure 3-225 Termination reactance compensation.

Tables giving reactance values can be found in Matthaei et al. [33, 34].

USE OF TRANSMISSION LINES AND INDUCTORS. In the preceding section, the inductance was expected to be realized by a lumped element. A transmission line can be used instead (Figure 3-226).

As can be seen from the computed selectivity curves (Figure 3-227) for the two configurations, transmission lines result in a larger bandwidth. The gain is important for a transmission line having a length $L = \lambda/4$ ($\Theta = 90°$) and a characteristic impedance $Z_0 = \sqrt{R_1 R_2}$. It is not significant for lines short with respect to $\lambda/4$. One will notice that there is an infinity of solutions, one for each value of C, when using transmission lines.

THREE-REACTANCE MATCHING NETWORKS. The networks that will be investigated are shown in Figure 3-228. They are made of three reactances alternately connected in series and shunt.

A three-reactance configuration allows the designer to make the quality factor (Q) of the circuit and the transformation ratio $n = R_2/R_1$ independent of each other and consequently to choose the selectivity between certain limits.

For narrowband designs, one can use the following formulas (see [35], where tables are given):

Network (a)

$$X_{C1} = R_1/Q \quad Q \text{ must be selected first} \tag{3-203}$$

$$X_{C2} = R_2 \sqrt{\frac{R_1 R_2}{(Q^2 + 1) - R_1/R_2}} \tag{3-204}$$

$$X_L = \frac{QR_1 + (R_1 R_2/X_{C2})}{Q^2 + 1} \tag{3-205}$$

Network (b)

$$X_{L1} = R_1 Q \quad Q \text{ must be selected first} \tag{3-206}$$

$$X_{L2} = R_2 B \tag{3-207}$$

$$A = R_1(1 + Q^2) \tag{3-208}$$

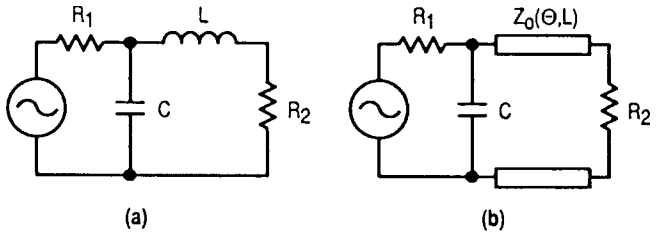


Figure 3-226 Use of a transmission line in the L section.

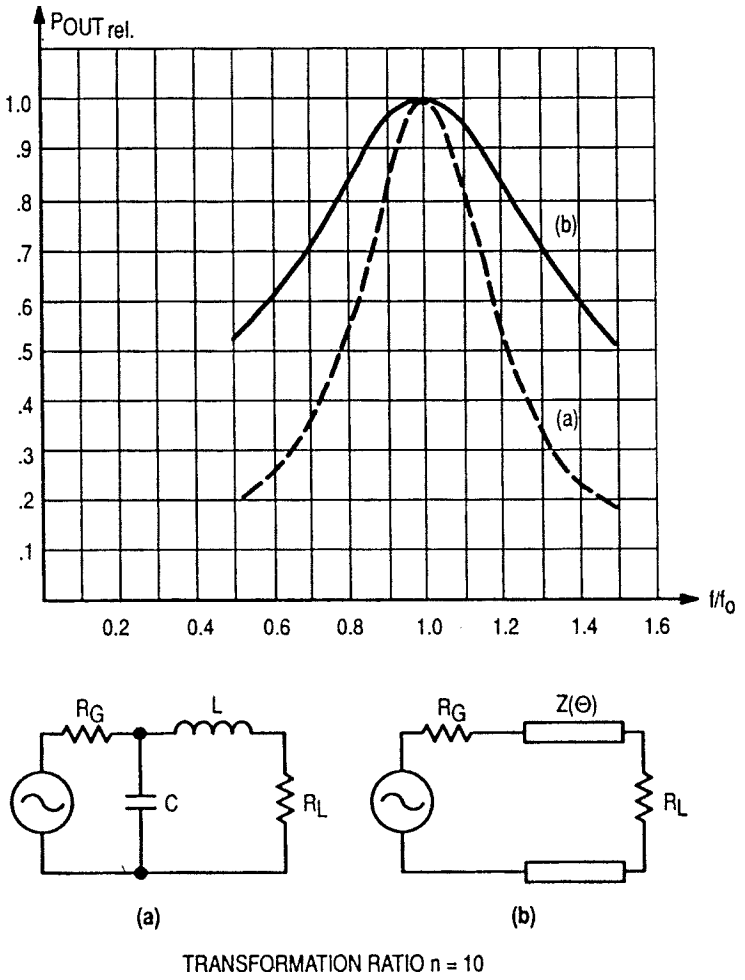


Figure 3-227 Bandwidth of the L section for $n = 10$ (a) with lumped constants and (b) with transmission line (λ/4).

$$X_{C1} = \frac{A}{Q + B} \tag{3-209}$$

$$B = \sqrt{\frac{A}{R_2} - 1} \tag{3-210}$$

Network (c)

$$X_{L1} = QR_1 \quad Q \text{ must be selected first} \tag{3-211}$$

$$X_{C2} = AR_2 \tag{3-212}$$

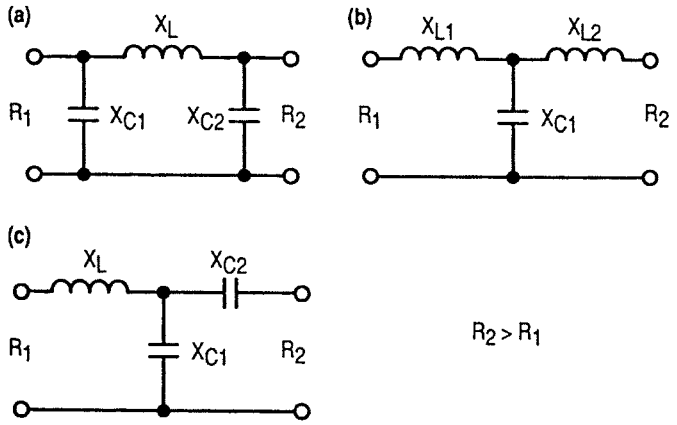


Figure 3-228 Three-reactance matching networks.

$$A = \sqrt{\frac{R_1(1+Q^2)}{R_2} - 1} \tag{3-213}$$

$$X_{C1} = \frac{B}{Q - A} \tag{3-214}$$

$$B = R_1(1+Q^2) \tag{3-215}$$

The network that yields the most-practical component values should be selected for a given application.

The three-reactance networks can be thought of as being formed of an L section (two reactances) and of a compensation reactance. The L section essentially performs the impedance transformation, while the additional reactance compensates for the reactive part of the transformed impedance over a certain frequency band.

Figure 3-229 shows a representation in the Z plane of the circuit of Figure 3-228a split into two parts: $R_1 - C_1 - L_1$ and $C_2 - R_2$. Exact transformation from R_1 into R_2 occurs at the points of intersection M and N. Impedances are then conjugate, or $Z' = R' + jX'$ and $Z'' = R'' + jX''$ with $R' = R''$ and $X' = -X''$.

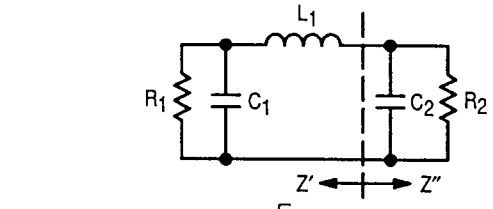
The only possible solution is obtained when X' and $-X''$ are tangential to each other. For the dashed curve, representing another value of L_1 or C_1, a wider frequency band could be expected at the expense of some ripple inside the band. However, this can only be reached with four reactances, as will be shown below.

With a three-reactance configuration, there are not enough degrees of freedom to permit $X' = -X''$ and simultaneously obtain the same variation of frequency on both curves from point M' to point N'. Exact transformation can, therefore, be obtained at only one frequency.

The values of the three reactances can be calculated by making

$$X' = -X'', \quad R' = R'', \quad \text{and} \frac{dX'}{dR'} = -\frac{dX''}{dR''} \tag{3-216}$$

The general solution of these equations leads to complicated calculations. Therefore, computed tables should be used.



$$Z' = \frac{R_1}{1 + \omega^2 R_1^2 C_1^2} + j\omega \left[L1 - \frac{R_1^2 C_1}{1 + \omega^2 R_1^2 C_1^2} \right] = R' + jX'$$

$$Z'' = \frac{R_2}{1 + \omega^2 C_2^2 R_2^2} - \frac{j\omega^2 C_2 R_2^2}{1 + \omega^2 C_2^2 R_2^2} = R'' + jX''$$

$$|x'| = \left(\frac{L_1}{C_1 R_1^2} - \frac{R'}{R_1} \right) \sqrt{\frac{R_1}{R'} - 1}$$

$$|x''| = \sqrt{R' R_2 - R'^2}$$

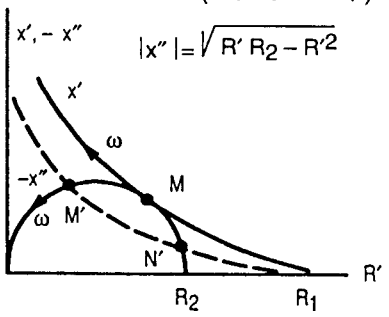


Figure 3-229 *Z*-plane representation of the circuit of Figure 3-228a.

One will note from Figure 3-229 that the compensation reactance contributes somewhat to the impedance transformation; that is, R' varies when going from M to R_2.

The circuit of Figure 3-228b is dual with respect to the first one and gives exactly the same results in a *Y*-plane representation. The circuit of Figure 3-228c is somewhat different since only one intersection M exists, as shown in Figure 3-230. Narrower frequency bands must be expected from this configuration. The widest band is obtained for $C_1 = \infty$.

Again, if one of the terminations has a reactive component, the latter can be taken as a part of the matching network, provided that it is not too large (see Figure 3-227).

FOUR-REACTANCE NETWORKS. Four-reactance networks are used essentially for broadband matching. The networks that will be considered in the following consist of two two-reactance sections in cascade. Some networks have pseudo-low-pass filter character; others, bandpass filter character. In principle, the former show narrower bandwidth since they extend the impedance transformation to very low frequencies unnecessarily, while the latter ensure good matching over a wide frequency band around the center frequency only (see Figure 3-233).

The two-reactance sections used in the networks in Figure 3-231 have either transformation properties or compensation properties. Impedance transformation is obtained with one series reactance and one shunt reactance. Compensation is made with both reactances in series or in shunt. If two cascaded transformation networks are used, transformation is accomplished partly by each one.

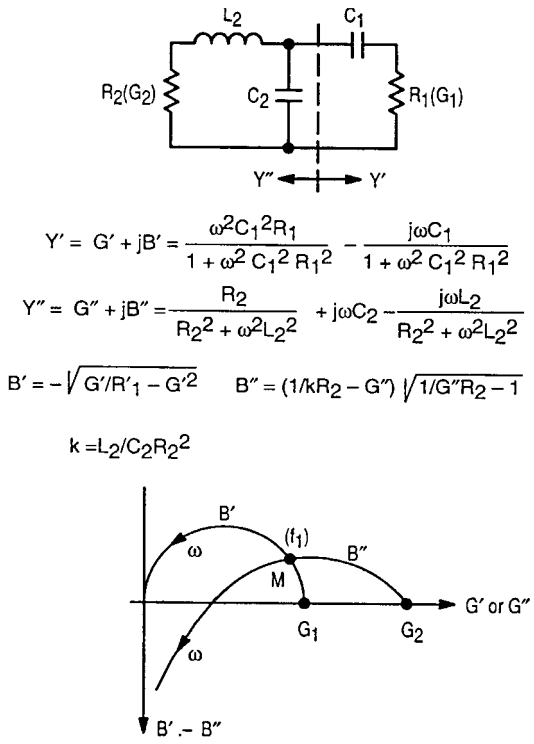


Figure 3-230 *Y*-plane representation of the circuit of Figure 3-228c.

With four-reactance networks there are two frequencies, f_1 and f_2, at which the transformation from R_1 into R_2 is exact. These frequencies may also coincide. For the network in Figure 3-231b, for instance, at point M, R_1 or R_2 is transformed into $\sqrt{R_1R_2}$ when both frequencies fall together. At all points (M), Z_1 and Z_2 are conjugate if the transformation is exact.

In the case of Figure 3-231b, the reactances are easily calculated for equal frequencies:

$$X_1 = \frac{R_1}{\sqrt{(n-1)^{1/2}}} \tag{3-217}$$

$$X_2 = R_1\sqrt{\frac{n^{1/2}-1}{n}} \tag{3-218}$$

$$X_1X_4 = R_1R_2 = X_2X_3 \tag{3-219}$$

$$X_3 = \frac{R_1}{\sqrt{n(n^{1/2}-1)}} \tag{3-220}$$

$$X_4 = \frac{R_1}{n}\sqrt{n^{1/2}-1} \tag{3-221}$$

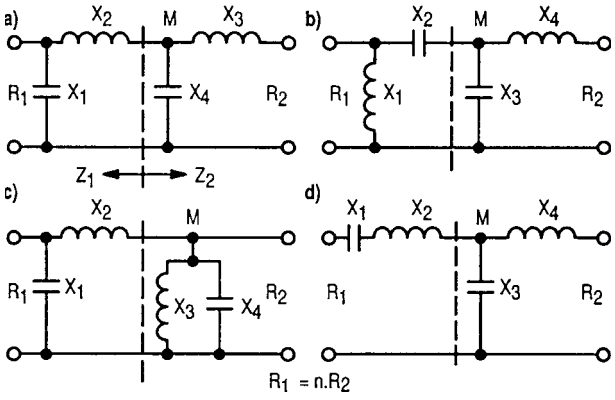


Figure 3-231 Four-reactance networks.

For the network in Figure 3-231a, normally at point M, Z_1 and Z_2 are complex. This pseudo-low-pass filter has been computed elsewhere [33]. Many tables can be found in the literature for networks of four and more reactances having Chebyshev character or maximally flat response [33, 34, 36].

Figure 3-232 shows the transformation path from R_1 to R_2 for networks (a) and (b) in Figure 3-231 on a Smith Chart.

Case (a) of Figure 3-232 has been calculated using tables in Matthaei [34]. Case (b) has been obtained from the relationship given above for $X_1, \ldots, X_4$. Both apply for a transformation ratio equal to 10 and for $R_1 = 1$. There is no simple relationship for $X_1', \ldots, X_4'$ of network (b) if f_1 is made different from f_2 for larger bandwidth. Figure 3-233 shows the respective bandwidths of networks (a) and (b) for the circuits shown in Figure 3-232.

If the terminations include a reactive component, the computed values for X_1 or X_4 may be adjusted to compensate for this.

For configuration (a), it can be seen from Figure 3-232 that, in the considered case, the Q values are equal to 1.6. For configuration (b), Q_1', which is equal to Q_2', is fixed for each transformation ratio

n	2	4	8	10	16
$Q'_1 = Q'_2$	0.65	1	1.35	1.46	1.73

$$Q' = \sqrt{n^{1/2} - 1} \qquad (3\text{-}222)$$

The maximum value of reactance that the terminations may have for use in this configuration can be determined from the above values of Q'.

If R_1 is the load resistance of a transistor, the internal transistor resistance may not be equal to R_1. In this case, the selectivity curve will be different from the curves given in Figure 3-233. Figure 3-234 shows the selectivity for networks (a) and (b) when the source resistance R_1 is infinite. From Figure 3-234 it can be seen that network (a) is more sensitive to changes in R_1 than network (b).

As mentioned earlier, the four-reactance network can also be thought of as two cascaded two-reactance sections: one used for transformation, the other for compensation. Figure 3-235 shows commonly used compensation networks, together with the associated L section.

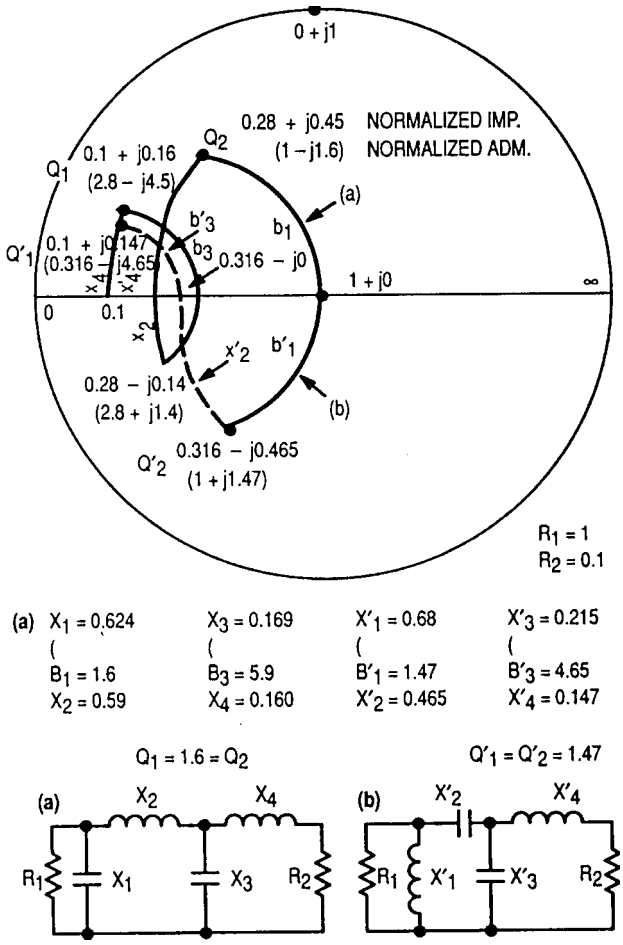


Figure 3-232 Transformation paths for networks (a) and (b).

The circuit of Figure 3-235a can be compared to the three-reactance network shown in Figure 3-228c. The difference is that capacitor C_2 of that circuit has been replaced by an *LC* circuit. The resulting improvement can be seen by comparing Figure 3-236 with Figure 3-230.

By adding one reactance, exact impedance transformation is achieved at two frequencies. It is now possible to choose component values such that the point of intersection M' occurs at the same frequency f_1 on both curves and simultaneously that N' occurs at the same frequency f_2 on both curves. Among the infinite number of possible intersections, only one allows the achievement of this.

When M' and N' coincide in M, the new $dX'/df = dX''/df$ condition can be added to the condition $X' = -X''$ (for three-reactance networks) and similarly $R' = R''$ and $dR'/df = dR''/df$.

Again, a general solution of the above equations leads to still more complicated calculations than in the case of three-reactance networks. Therefore, tables are preferable [33, 34, 36].

The circuit of Figure 3-235b is the dual of the circuit of Figure 3-233a and does not need to be treated separately. It gives exactly the same results in the *Z* plane. Figure 3-235c shows a higher-order compensation requiring six reactive elements.

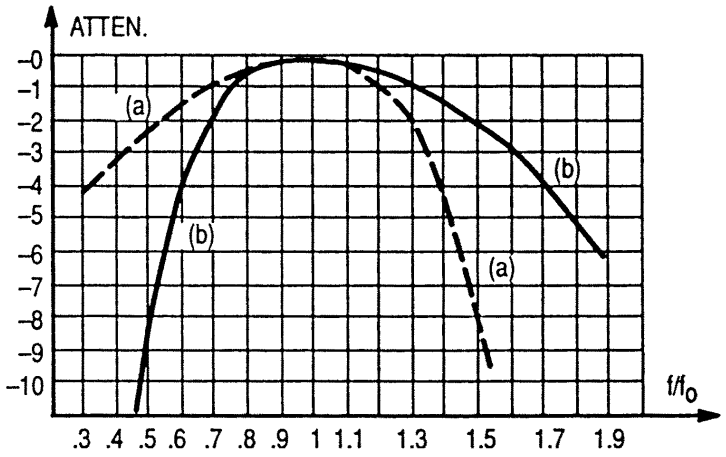


Figure 3-233 Selectivity curves for networks (a) and (b) of Figure 3-232.

The above-discussed matching networks employing compensation circuits result in narrower bandwidths than the former solutions (see p. 573) using two transformation sections. A matching arrangement with higher-order compensation, such as in Figure 3-235c, is not recommended. Better use can be made of the large number of reactive elements by using them all for transformation.

When the above configurations are realized using short portions of transmission lines, the equations or the usual tables no longer apply. The calculations must be carried out on a computer because of their complexity. However, a graphical method can be used (see the next section) that will consist essentially in tracing a transformation path on the Z–Y chart using the computed lumped element values and replacing it by the closest path obtained with distributed constants. The bandwidth change is not significant as long as short portions of lines are used.

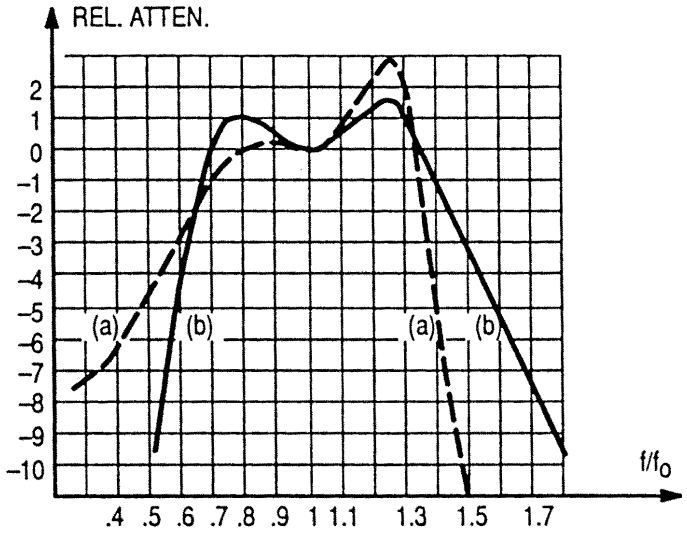


Figure 3-234 Selectivity curves for networks (a) and (b) of Figure 3-232 with infinite R_1.

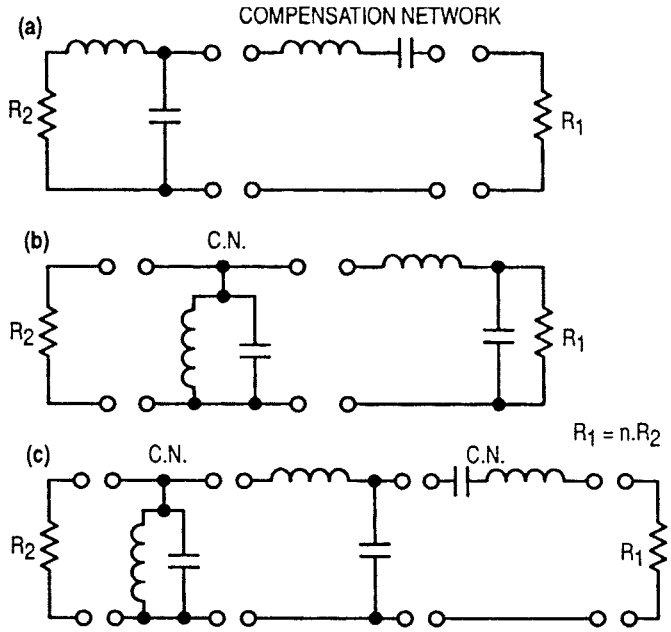


Figure 3-235 Compensation networks used with an L section.

MATCHING NETWORKS USING QUARTER-WAVE TRANSFORMERS. At sufficiently high frequencies, where λ/4-long lines of practical size can be realized, broadband transformations can easily be accomplished by the use of one or more λ/4 sections. Figure 3-237 summarizes the main relations for (a) one-section and (b) two-section transformations. A compensation network can be realized using a λ/2-long transmission line. Figures 3-238 and 3-239 show the selectivity curves for different transformation ratios and section numbers.

EXPONENTIAL LINES. Exponential lines have largely frequency-independent transformation properties. The characteristic impedance of such lines varies exponentially with their length l:

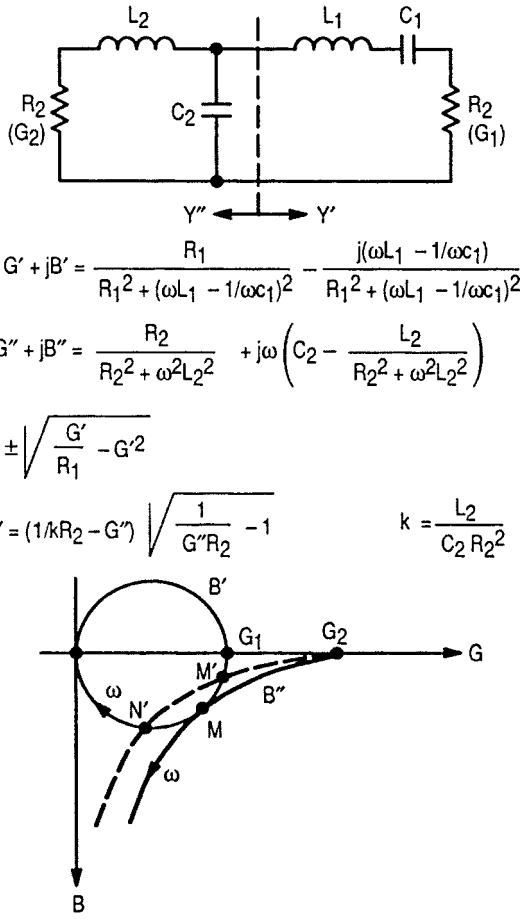
$$Z = Z_0 e^{kl} \tag{3-223}$$

where k is a constant, but these properties are preserved only if k is small.

BROADBAND MATCHING USING BANDPASS FILTER NETWORKS—HIGHQ CASE. The above circuits are applicable to devices having low input or output Q, if broadband matching is required. Generally, if the impedances to be matched can be represented, for instance, by a resistor R in series with an inductor L (sometimes a capacitor C) within the band of interest and if L is sufficiently low, the latter can be incorporated into the first inductor in the matching network. This is also valid if the representation consists of a shunt combination of a resistor and a reactance.

Practically, this is feasible for Q values around 1 or 2. For higher Q values or for input impedances consisting of a series or parallel resonant circuit (see Figure 3-221), as it appears to be for large bandwidths, a different treatment must be followed.

Let us first recall that, as shown by Bode [37] and Fano [38], limitations exist on the impedance matching of a complex load. In the example of Figure 3-240, the load to be matched consists of a capacitor C and a resistor R in shunt.



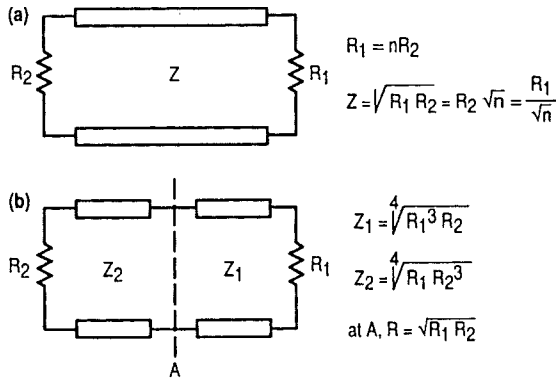
$$Y' = G' + jB' = \frac{R_1}{R_1^2 + (\omega L_1 - 1/\omega c_1)^2} - \frac{j(\omega L_1 - 1/\omega c_1)}{R_1^2 + (\omega L_1 - 1/\omega c_1)^2}$$

$$Y'' = G'' + jB'' = \frac{R_2}{R_2^2 + \omega^2 L_2^2} + j\omega\left(C_2 - \frac{L_2}{R_2^2 + \omega^2 L_2^2}\right)$$

$$B' = \pm\sqrt{\frac{G'}{R_1} - G'^2}$$

$$B'' = (1/kR_2 - G'')\sqrt{\frac{1}{G''R_2} - 1} \qquad k = \frac{L_2}{C_2 R_2^2}$$

Figure 3-236 *Y*-plane representation of the circuit of Figure 3-235a.



$$R_1 = nR_2$$

$$Z = \sqrt{R_1 R_2} = R_2\sqrt{n} = \frac{R_1}{\sqrt{n}}$$

$$Z_1 = \sqrt[4]{R_1^3 R_2}$$

$$Z_2 = \sqrt[4]{R_1 R_2^3}$$

$$\text{at A, } R = \sqrt{R_1 R_2}$$

Figure 3-237 Transformation networks using λ/4-long transmission lines.

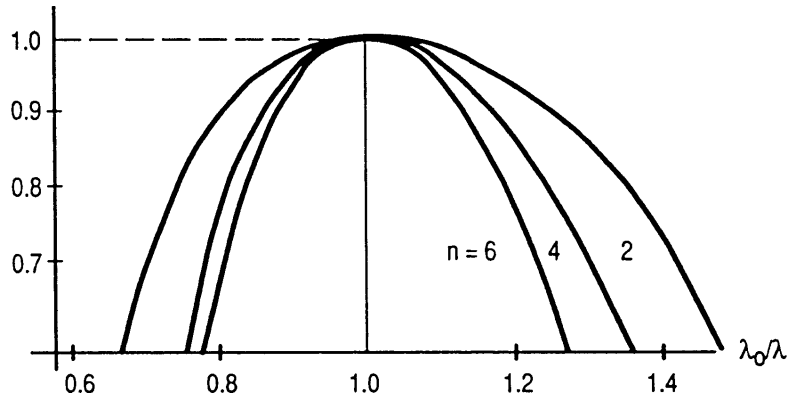


Figure 3-238 Selectivity curves for two λ/4-section networks at different transformation ratios.

The reflection coefficient between the transformed load and generator is equal to

$$\Gamma = \frac{Z_{\text{IN}} - R_G}{Z_{\text{IN}} + R_G} \tag{3-224}$$

When $\Gamma = 0$, there is perfect matching; when $\Gamma = 1$, there is total reflection. The ratio of reflected to incident power is

$$\frac{P_r}{P_i} = |\Gamma|^2 \tag{3-225}$$

The fundamental limitation on the matching takes the form

$$\int_{\omega=0}^{\infty} \ln\left(\frac{1}{|\Gamma|}\right) d\omega \le \frac{\pi}{RC} \qquad \text{Bode equation} \tag{3-226}$$

and is represented in Figure 3-241.

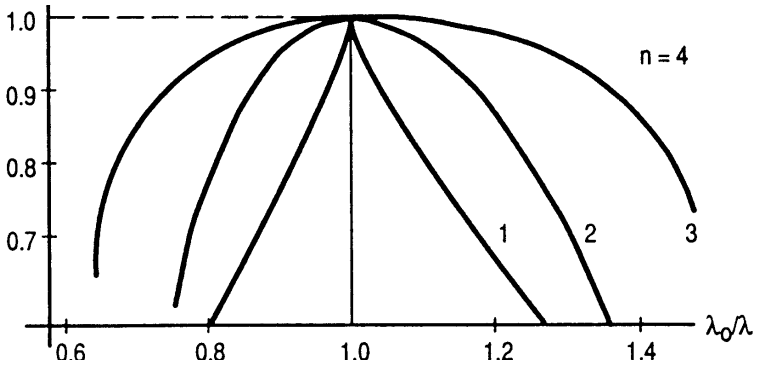


Figure 3-239 Selectivity curves for one, two, and three λ/4 sections.

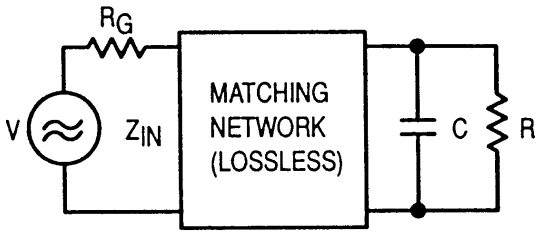


Figure 3-240 General matching conditions.

The meaning of the Bode equation is that the area S under the curve cannot be greater than π/RC and, therefore, if matching is required over a certain bandwidth, this can only be done at the expense of less power transfer within the band. Thus, power transfer and bandwidth appear as interchangeable quantities.

It is evident that the best utilization of the area S is obtained when $|\Gamma|$ is kept constant over the desired band ω_c and made equal to 1 over the rest of the spectrum. Then

$$|\Gamma| = e^{-\pi/\omega_c RC} \tag{3-227}$$

within the band and no power transfer happens outside. A network fulfilling this requirement cannot be obtained in practice because an infinite number of reactive elements would be necessary.

If the attenuation is plotted versus frequency for practical cases, one may expect to have curves like the ones shown in Figure 3-242 for a low-pass filter having Chebyshev character.

For a given complex load, an extension of the bandwidth from ω_1 to ω_2 is possible only with a simultaneous increase of the attenuation a. This is especially noticeable for Q values exceeding 1 or 2 (see Figure 3-243). Thus, devices having relatively high input Q values are usable for broadband operation, provided the consequent higher attenuation or reflection introduced is acceptable.

For average insertion losses or attenuation a (neglecting the ripple), the attenuation decreases if the number n of the network element increases. But above $n = 4$, the improvement is small. For a given attenuation a and bandwidth, the larger the n the smaller the ripple; for a given attenuation, the larger the n the larger the bandwidth.

Computations show that for $Q < 1$ and $n \leq 3$, the attenuation is below 0.1 dB (approximately). The impedance transformation ratio is not free here. The network is a

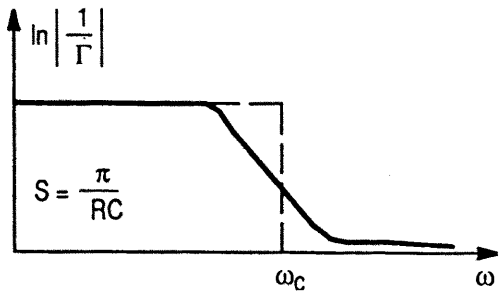


Figure 3-241 Representation of the Bode equation.

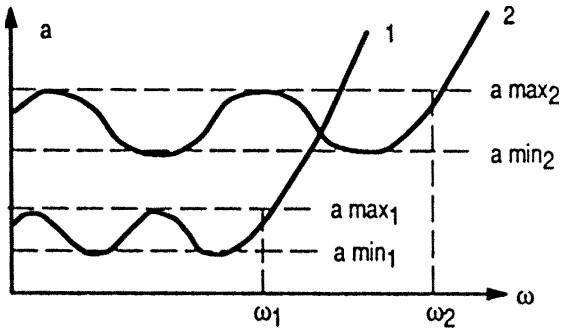


Figure 3-242 Attenuation versus angular frequency for different bandwidths with the same load.

true low-pass filter. For a given load, the optimum generator impedance will result from the computation.

Before impedance transformation is introduced, a conversion of the low-pass prototype into a bandpass filter network must be made. Figure 3-244 summarizes the main relations for this conversion. r is the conversion factor. For the bandpass filter, $Q_{\text{IN max}}$, or the maximum possible input Q of the device to be matched, has been increased by the factor r (from Figure 3-244, $Q'_{\text{IN max}} = rQ_{\text{IN max}}$).

Impedance inverters will be used for impedance transformation. These networks are suitable for insertion into a bandpass filter without affecting the transmission characteristics.

Figure 3-245 shows four impedance inverters. Note that one of the reactances is negative and must be combined in the bandpass network with a reactance of at least equal positive value. Insertion of the inverter can be made at any convenient place [32, 39].

When using the bandpass filter for matching the input impedance of a transistor, reactances $L'_1C'_1$ should be made to resonate at ω_0 by addition of a convenient series reactance.

As stated earlier, the series combination of R_0, L'_1 and C'_1 normally constitutes the equivalent input network of a transistor when considered over a large bandwidth. This is a good approximation up to about 500 MHz.

In practice, the normal procedure for using a bandpass filter matching network will be the following:

1. For a given bandwidth center frequency and input impedance of a device to be matched—for example, 50 Ω—first determine Q'_{IN} from the datasheet as

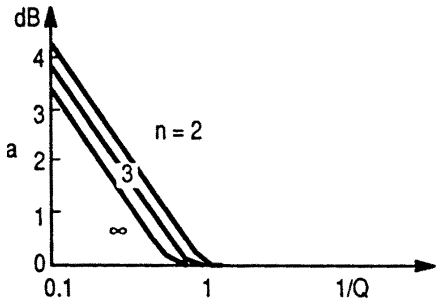


Figure 3-243 Insertion losses as a function of 1/*Q*.

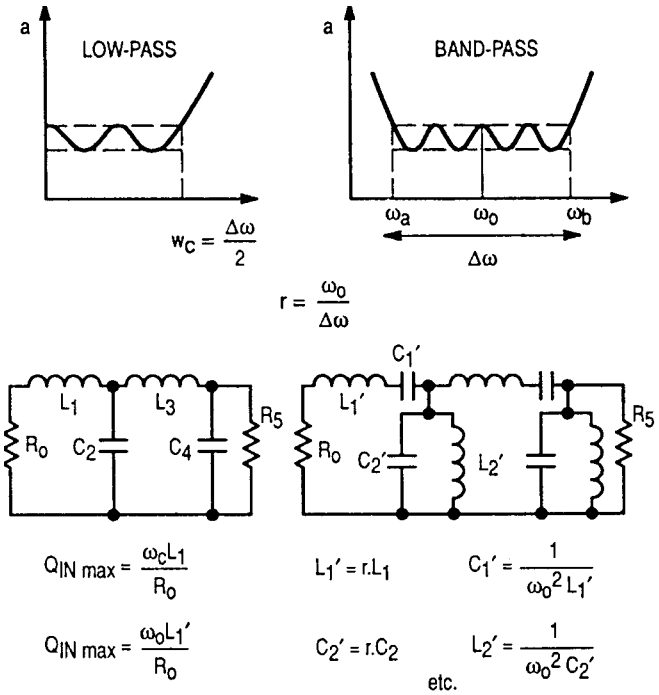


Figure 3-244 Conversion from low-pass filter into bandpass filter.

$$\frac{\omega_0 L_1'}{R_0} \tag{3-228}$$

after having eventually added a series reactor for centering.

2. Convert the equivalent circuit $R_0 L_1' C_1'$ into a low-pass prototype $R_0 L_1$ and calculate Q_{IN} using the formulas of Figure 3-244.
3. Determine the other reactance values from tables [33] for the desired bandwidth.
4. Convert the element values found by Step 3 into series- or parallel-resonant circuit parameters.
5. Insert the impedance inverter in any convenient place.

In the above discussions, the gain roll-off has not been taken into account. This is of normal use for moderate bandwidths (e.g., 30%). However, several methods can be employed to obtain a constant gain within the band despite the intrinsic gain decrease of a transistor with frequency. Tables have been computed elsewhere [40] for matching networks approximating 6 dB/octave attenuation versus frequency.

Another method consists of using the above-mentioned network and then adding a compensation circuit, as shown in Figure 3-246.

Resonance ω_b is placed at the high edge of the frequency band. Choosing Q correctly, roll-off can be made 6 dB/octave. The response of the circuit of Figure 3-246 is expressed by

$$\frac{1}{1 + Q^2(\omega/\omega_b - \omega_b/\omega)^2} \tag{3-229}$$

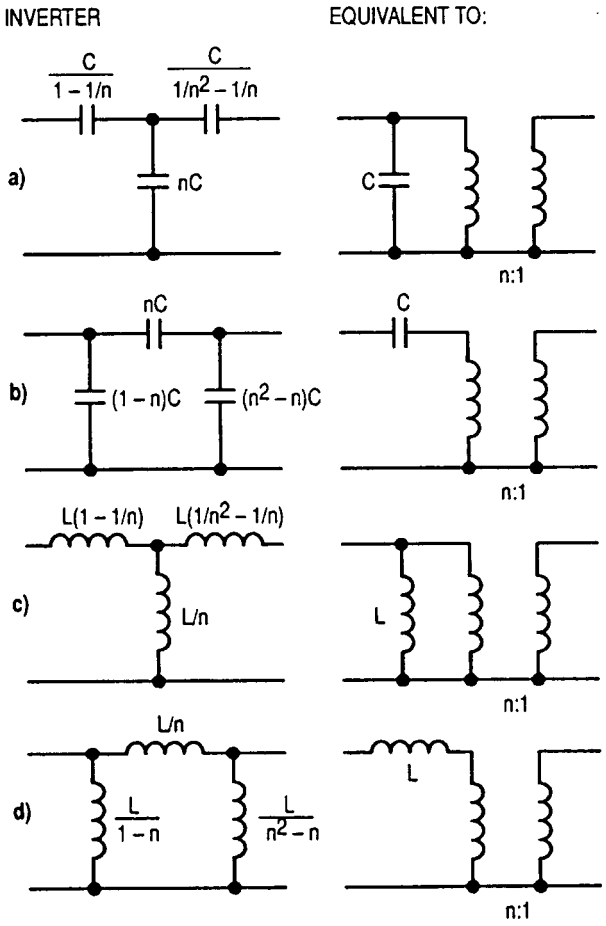


Figure 3-245 Impedance inverters.

where $\omega < \omega_b$. This must be equal to ω/ω_b for 6 dB/octave attenuation.

At the other band edge a, exact compensation can be obtained if

$$Q = \frac{(\omega_b/\omega_a)^2 - 1}{\omega_a/\omega_b - (\omega_b/\omega_a)^2} \tag{3-230}$$

There are several methods available for synthesis of matching networks. They should not be confused with filters because they can be either high-pass, low-pass, or bandpass, and their main purpose is to match one complex impedance to another. In many cases, one of the terminations happens to be 50 Ω and real, but this is just a "lucky" case of the general approach. An excellent paper on this [41] refers to the "real frequency" technique. Another equally important publication is the one by Cuthbert [42], since followed by his book [43]. Two more good references are Vendelin et al. [44] and Potter [45].

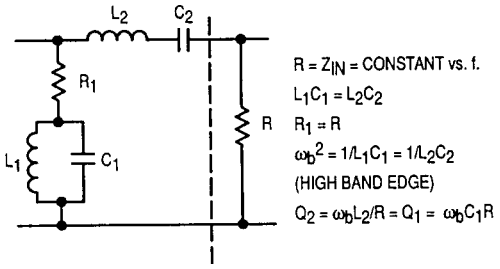


Figure 3-246 Roll-off compensation network.

3-12-3 Example 2: Low-Noise Amplifier Using Distributed Elements

In our previous examples, the amplifiers were built around lumped elements. Since a transmission line length of less than $\lambda/4$ is an inductor and $\frac{3}{4}\lambda$ is a capacitor, the whole amplifier can be built around microstrip technology.

If anyone wonders why we are going to show a low-noise amplifier in the middle of the power section, the reason is that we will select a low-power Siemens transistor, which, being operated in Class A, has an intercept point of +30 dBm and can easily achieve a noise figure of less than 2 dB. Such an amplifier is attractive when feeding an input stage from an antenna that sees many signals and curtails the bandwidth with an input filter of good selectivity that needs a good output termination, which in many cases works against the achievement of a good noise figure because when designing for best noise figure the input impedance of the transistor stage is far from 50 Ω.

First, we will start with the textbook approach and use the Smith diagram to obtain the best noise figure of which the device is capable. Figure 3-247 shows the actual approach to input and output matching.

We start our design by querying the Smith Tool about the minimum noise figure and the gain that comes with it. We mark the point of the lowest noise figure (labeled Γ_{opt}). The small

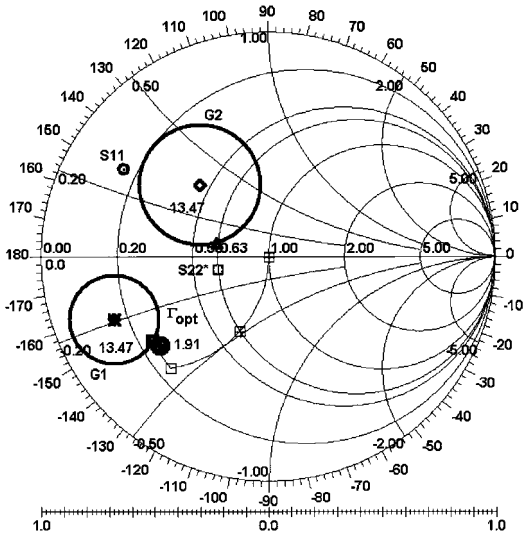


Figure 3-247 Input and output matches for 13.47 dB of gain and NF = 1.91 dB as displayed by the interactive Smith Tool.

noise circle yields the location for Γ_{opt} and tells us the minimum noise figure is going to be 1.91 dB. (By the way, this assumes no further losses, even in soldering the transistor to the PC board.) Then we construct the input gain circle (G1), which touches Γ_{opt} and turns out to represent 13.47 dB of available gain.

The next step is to find a matching network between Γ_{opt} and the 50-Ω source impedance. Immediately there is a conflict: If we look at S_{11}, which is the input reflection coefficient shown in the Smith diagram, it is on the other side of the complex plane and sufficiently far away, indicating that there is going to be quite a difference between noise matching and power-gain matching. We have two options for matching the input: Either we start from the origin and move toward the position of Γ_{opt}, or we start with Γ_{opt} and go to 50 Ω. Depending on the direction in which we do this, the result can be different topologies, such as a high-pass or low-pass filter. The high-pass filter is obtained by starting with Γ_{opt}; the low-pass topology is obtained by starting at the 50-Ω point. The Smith Tool then provides us with the corresponding values of capacitors and inductors after the transformation has been accomplished.

The final step is to do the same type of matching at the output. We start with S_{22}, and either select S_{22}^* (the complex conjugate for highest gain) or S_{22load}. As a rule, the second one is always a lower impedance (in terms of the real part) than the conjugate match. This graphic procedure is somewhat trial and error. As to the topology, a high-pass match should be preferred at the input and a low-pass match at the output. The reason for this is that the high-pass filter will tend to protect the stage from the enormous number of signals present below 1 GHz; using a low-pass filter at the output suppresses unwanted harmonics that the amplifier may generate. Figure 3-248 shows one of the approaches we like best simply because it gives a good input termination and we were trying for a simultaneous match for best noise figure and power gain.

The way we were able to achieve this was to take advantage of the existing feedback in the transistor and select a less-than-optimal output termination but gaining at the input. Taking advantage of the Miller effect, we rotate the phase angle of the feedback with the load to move the optimum noise and gain matching points closer. The danger in this is that the circuit is now highly sensitive to both input and output termination. The output *must* see 50 Ω to maintain this good input match, but since the second transistor stage can now be optimized for best input match, a two-stage combination can be configured to meet this

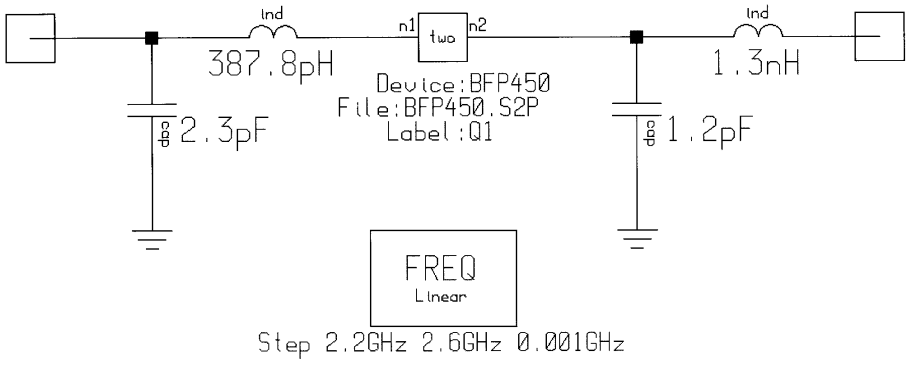


Figure 3-248 A 2.4-GHz amplifier using a BFP450 transistor and lumped-element input and output matching.

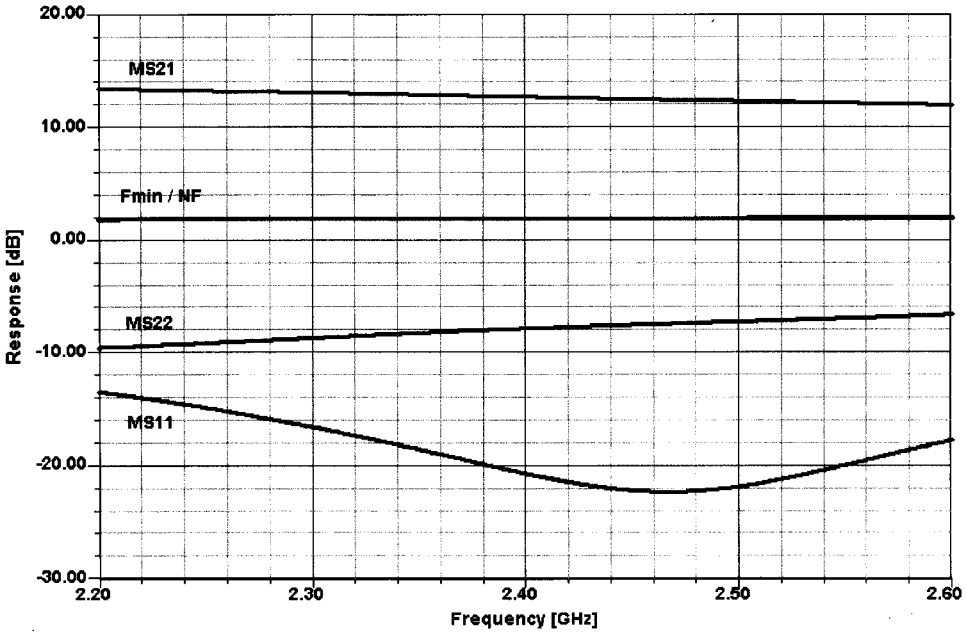


Figure 3-249 Frequency-dependent gain, matching, and noise characteristics of the matched high-dynamic-range BFP450 2.4-GHz amplifier.

requirement. Also, let us not forget that the device is running at 50 mA and as far as bipolar transistors are concerned, this is probably the best amplifier combination we have seen for getting the highest dynamic range and lowest noise figure. The resulting gain, matching, and noise figure characteristics are shown in Figure 3-249.

Now we turn to using distributed elements and at the same time use a shunt feedback at the output; to be specific, we will load it with 150 Ω. The reason for this is to be able to obtain a good compromise at the input that will depend less on the output and at the same

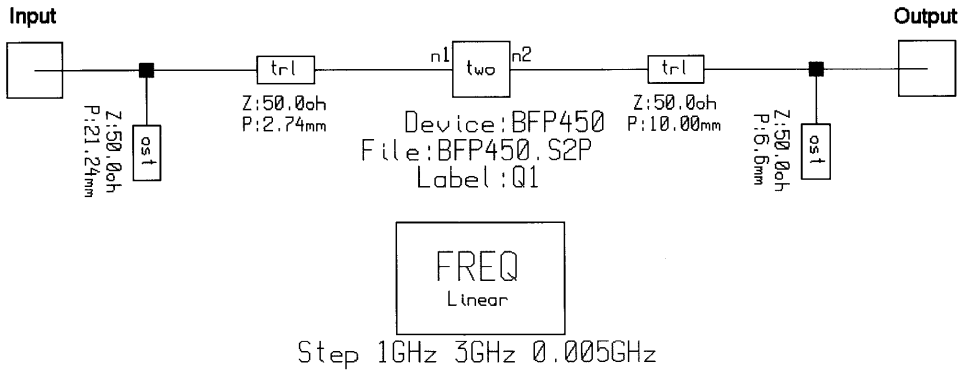


Figure 3-250 Distributed-element input and output matching for the 2.4-GHz BFP450 amplifier.

time improve the S_{22} matching at the output. The procedure for doing the matching with a transmission line offers only certain degrees of freedom, while doing the matching with lumped elements allows the choice of series or parallel inductors or capacitors. In the case of a transmission line, we have open stubs, shorted stubs, and transmission lines. This may result in somewhat more complex matching schemes. On the positive side, they can easily be trimmed, are reproducible, and certainly cost less than exotic capacitors, which need to have small tolerances and to be able to handle high RF currents.

The resulting input and output matching is shown in Figure 3-250, and its frequency response in Figure 3-251. We were able to maintain a good input and output match with an insignificant deterioration of the noise figure.

Since this is a high-dynamic-range amplifier, we also show its predicted two-tone response in Figure 3-252. The manufacturer specifies an $IP_{3,in}$ of 29 dB for a bias point of $V_{CE} = 3$ V and $I_C = 50$ mA with $Z_S = Z_{S,opt}$ and $Z_L = Z_{L,opt}$ at 1.8 GHz. Significantly, however, the manufacturer-supplied nonlinear BFP450 library we used exhibits considerably different characteristics at $V_{CE} = 4$ V and $I_C = 50$ mA than those reflected in the corresponding S parameters supplied by the same manufacturer. In our circuit, the S parameters resulted in about 13.5 dB of gain (based on a compromise between matching for optimum noise and maximum gain) versus about 9.5 dB of gain for the nonlinear library (match optimized for maximum gain). The difference between the 2.4-GHz F_{min} values returned for the two modeling approaches is similarly striking: 1.90 dB for the bare S parameter set and 7.76 dB for the bare nonlinear library biased to $V_{CE} = 4$ V and $I_C = 50$ mA. Clearly, the nonlinear modeling of the BFP450 needs work!

Having fully evaluated the amplifier at 2.4 GHz as a working example, we note that this frequency is used by radio amateurs for some applications. We will therefore supply

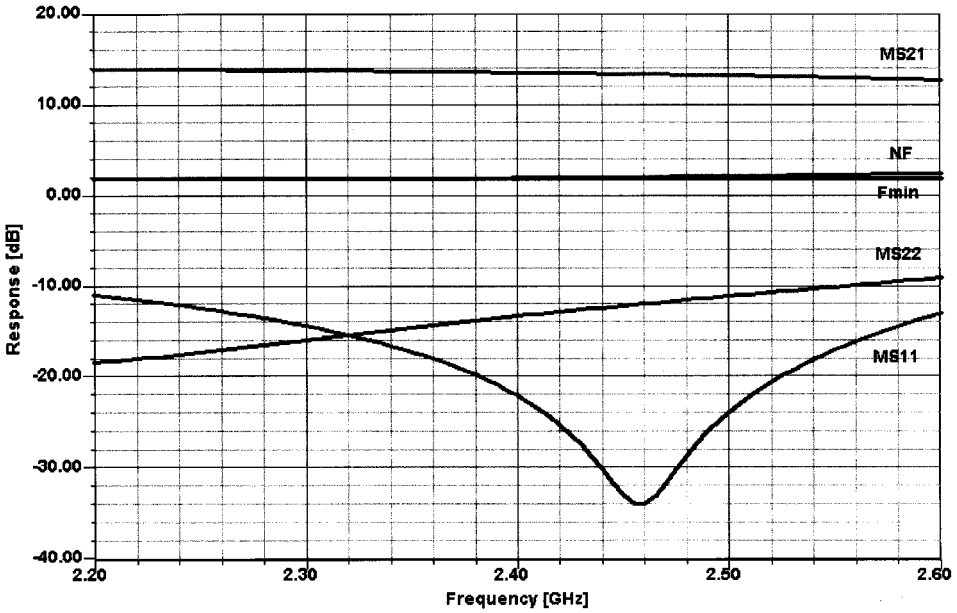


Figure 3-251 Predicted frequency-dependent gain, matching, and noise performance for the BFP450 amplifier with distributed-element matching.

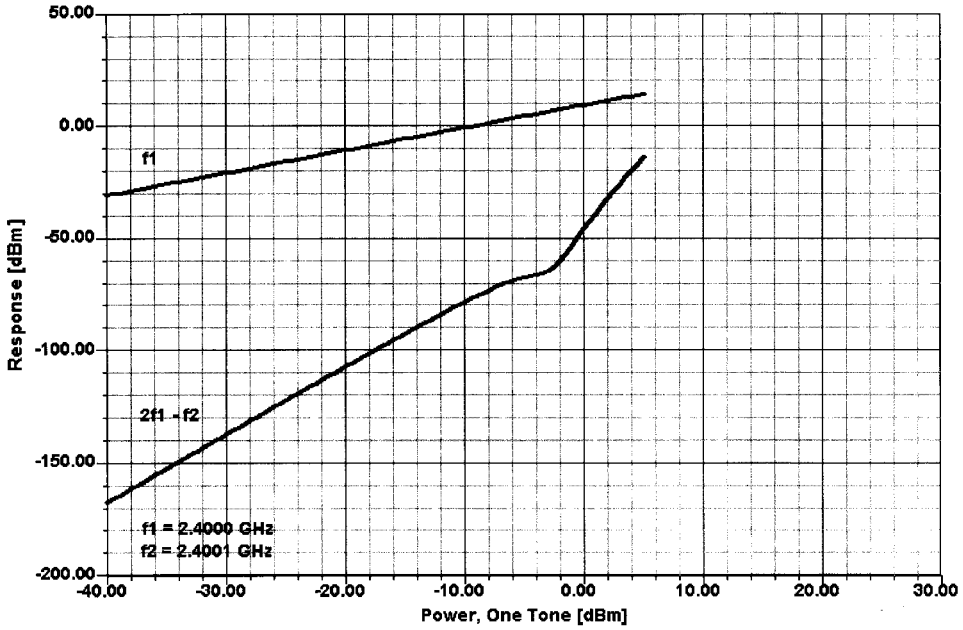


Figure 3-252 Predicted two-tone response of the BFP450 amplifier with the circuit adjusted for maximum gain (9.5 dB): $IP_{3,in} \approx 28.5$ dBm and $IP_{3,out} \approx 38$ dBm. That this version's gain is significantly lower than that of the *S*-parameter-based equivalent illustrates the need for better accuracy in parameter extraction.

examples for 432 and 1296 MHz, popular spot frequencies in the amateur 70- and 23-cm bands, respectively. Figures 3-253 and 3-254 show the 432-MHz version's schematic and frequency-dependent responses. Figures 3-255 and 3-256 show the schematic and frequency-dependent responses for the 1296-MHz version.

Using a bipolar transistor that operates at much less current, the noise figure can be improved at the expense of the dynamic range/intercept point. It is possible to reduce the noise figure by about 1 dB or so, but the reduction in intercept point will be much more dramatic. Another choice is the use of GaAsFETs for this application, but their matching in comparable circuits has been described as a nightmare simply because the devices have too much gain and very high input impedances (essentially 0.2 pF in series with 3 Ω) combined with reasonable output impedances in the vicinity of 240 Ω. Here is a wide field in which engineers can experiment and examine price/performance and performance/real estate issues. GaAsFETs also tend to require more than 30–40 mA to be alive, and even the low operating voltage of 4–5 V makes the device consume a lot of dc power.

3-12-4 Example 3: 1-W Amplifier Using the CLY15

The following is an application showing a fairly wideband amplifier with +30 dBm (1 W) output. We need to point out that in accordance with the various digital modulation modes used with this amplifier, this is not continuous operation, but a much less than 1:1 duty cycle operation. The device would probably not survive operating at full output under CW conditions. This is subject to a discussion with the manufacturer, Siemens, who has not

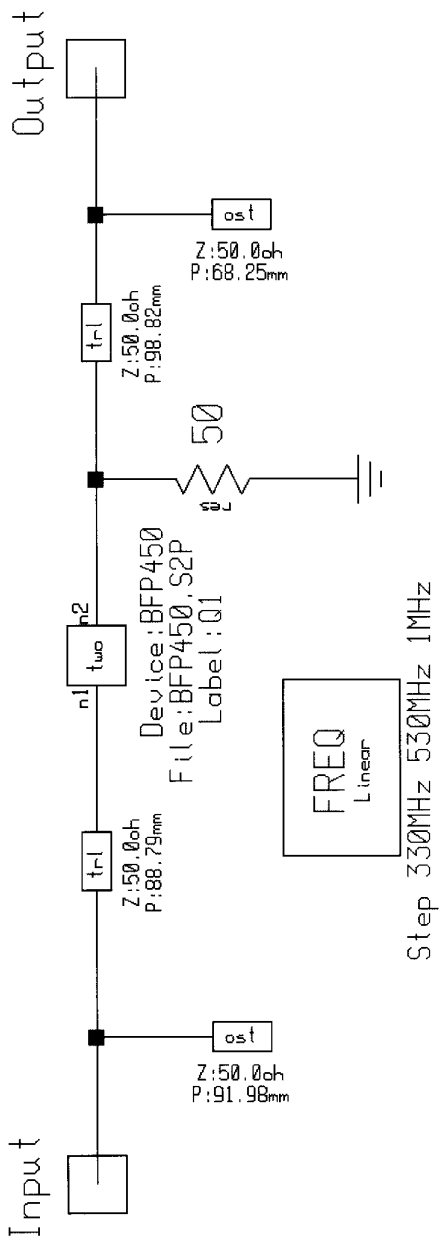


Figure 3-253 Schematic of BPF450 amplifier for 432 MHz. Because f_T = 25 GHz for the BFP450, at this lower frequency we must shunt the device output with a 50-Ω resistor to maintain stability.

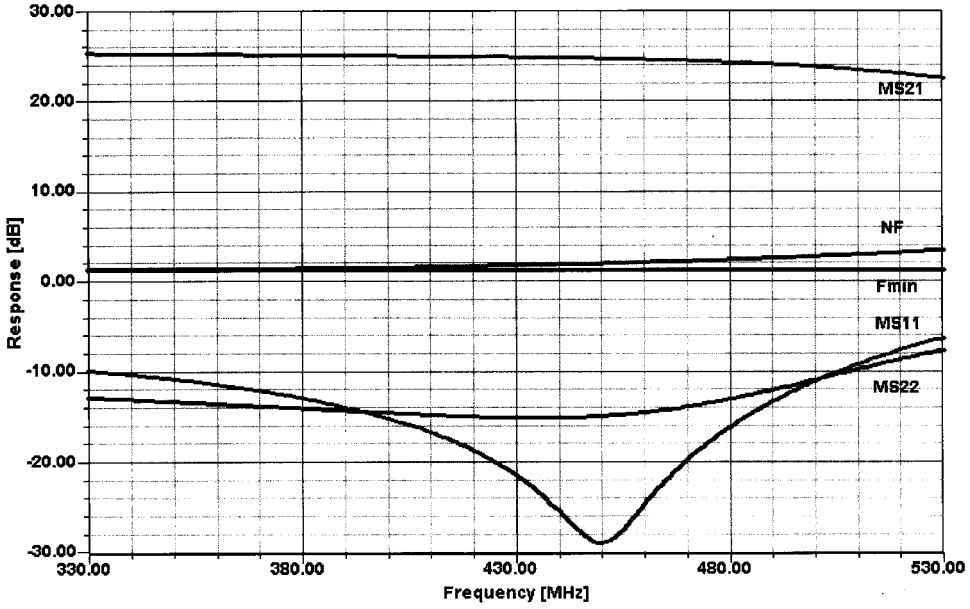


Figure 3-254 Frequency-dependent gain, matching, and noise performance for the BPF450 amplifier at 432 MHz. As with the circuit in Figure 3-253, the input and output network values have been chosen to achieve a low NF and good input matching at the expense of the output match.

released sufficient data for us to make a judgment as to the CLY15's continuous power capability. Since there is no useful large-signal model available at the time of this writing, we will design this Class A amplifier using large dc bias *S* parameters. In the actual design there will be some cut-and-try, but we find that in this type of application this is still the best way of handling the design.

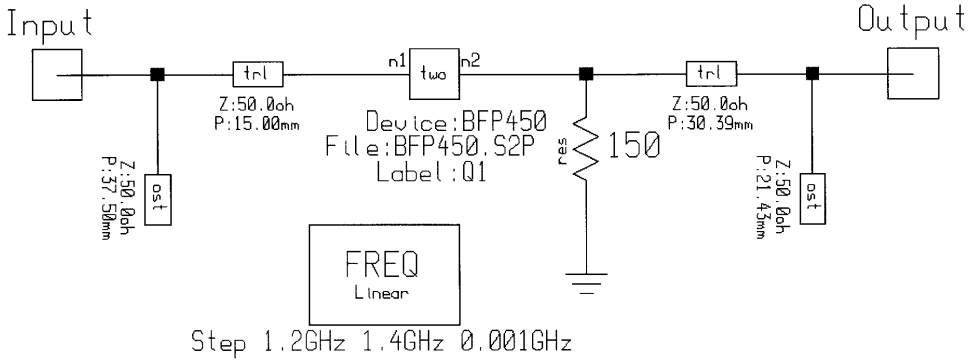


Figure 3-255 Schematic of the BPF450 amplifier for 1296 MHz. A loading resistor is still necessary to ensure stability, but because the transistor exhibits less gain at this frequency than at 432 MHz, we can increase the loading resistor to 150 Ω.

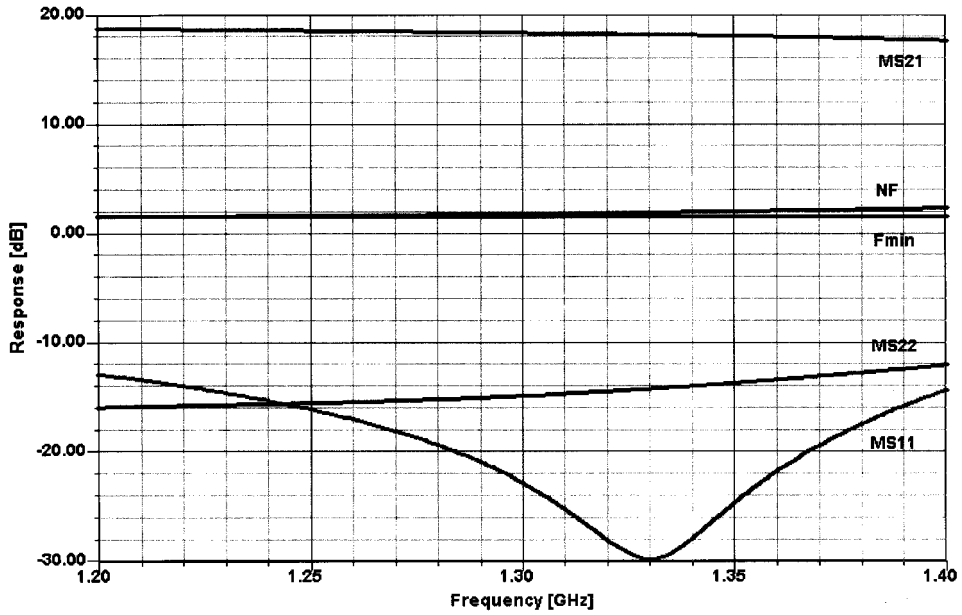


Figure 3-256 Frequency-dependent gain, matching, and noise responses of the BFP450 amplifier at 1296 MHz. As with the two previous versions, the input and output network values have been chosen to achieve a low NF and good input matching at the expense of the output match.

The CLY15 (gate width 16 mm) is the largest of a series of Siemens GaAsFETs; it was already mentioned in Chapter 2, where its datasheet is reproduced. Its device family also includes the CLY2 (gate width 2 mm), CLY5 (4 mm), and CLY10 (8 mm). It is possible that bigger devices will be offered in the future. One of the most critical parameters will be the saturation voltage; from the datasheet of the CLY15, we can see that at 1.5 A the saturation voltage is approximately 1 V. This drop of 1 V includes losses that result from other parasitic elements; the dc *I–V* curve (Figure 3-257a) indicates a saturation voltage of 0.75. Since we aim for 1-W output power, the resistive portion of the load will be $(5 - 1)^2 / (1 \times 2) = 8\ \Omega$. This assumes a supply voltage of 5 V.

The next step is to design the input and output matching network. Previewing the *S* parameters, it will be noticed that they are essentially inductive, which means that the input match, using lumped elements, will result in a combination of capacitors at both the input and the output. Figures 3-258 and 3-259 show the ways in which the input and output matches were done on the Smith diagram resulting in lumped matching elements. Inspecting the lumped input matching, we find the following. We start with S_{11}, which is in the lower part of the Smith diagram, and determine its conjugate match, S_{11}^*. We already pointed out that this device is inductive, which means that matching it from the conjugate point will require two capacitors. The first one is a series capacitor, as identified in the Smith diagram, and followed by a shunt capacitor to transform the impedance to 50 Ω. A close inspection of the Smith diagram shows that the resulting point is not quite on S_{11}; it is an iterative process to exactly find the crossover point, yet the input match of better than 20 dB is more than acceptable.

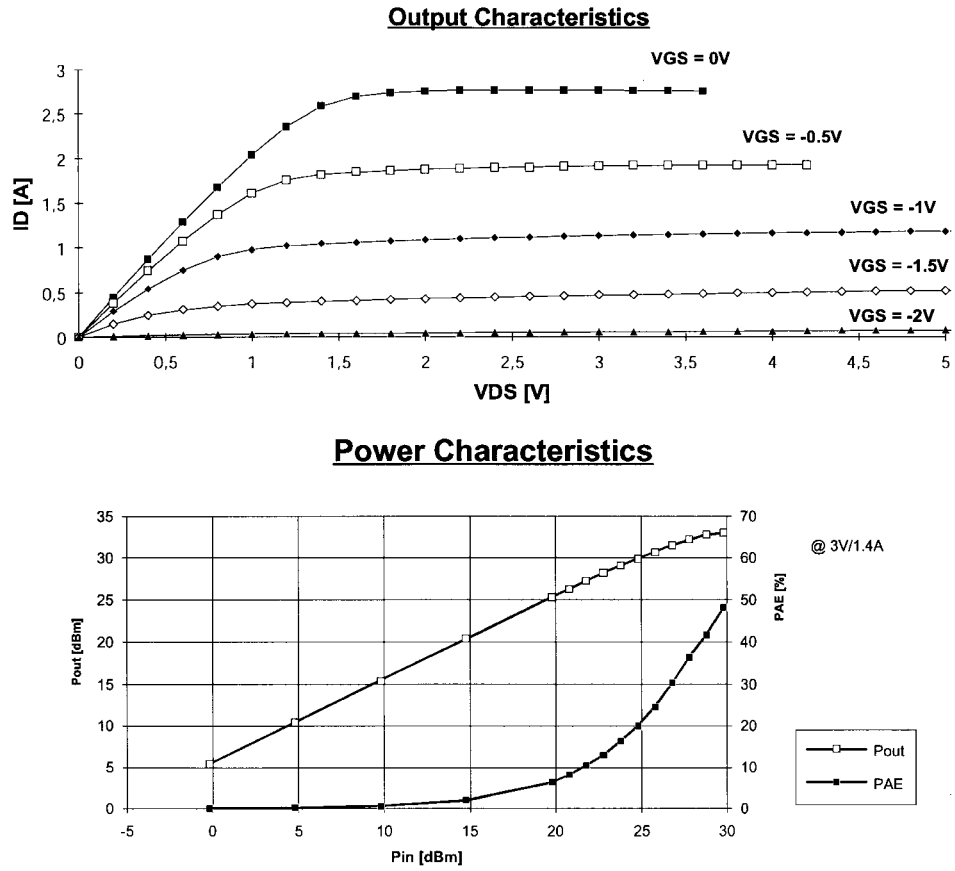


Figure 3-257 Output and power characteristics for the CLY15 medium-power GaAsFET.

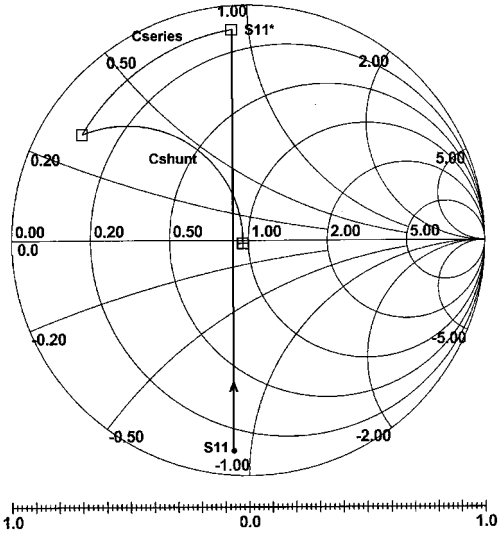


Figure 3-258 Input match for S_{11}^{*} using lumped elements.

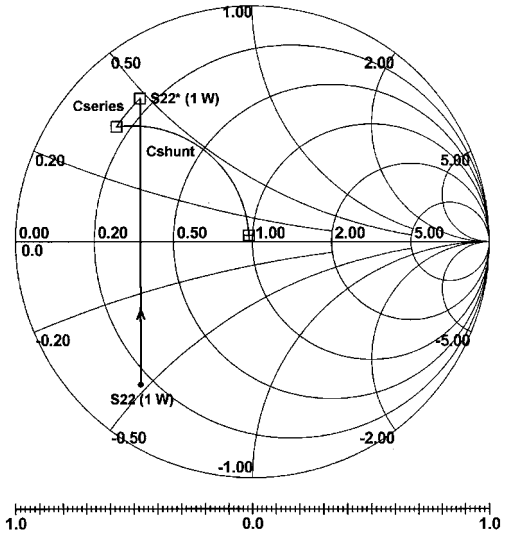


Figure 3-259 Output match for S_{22}^* (1 W) using ideal lumped elements.

As far as the output is concerned, we use the same technique by moving from S_{22} to S_{22}^*. S_{22} has been determined from the required output power; as we showed above, the real portion is 8 Ω. Using the lumped elements, the matching technique is identical; we first need to use a series capacitor and then a shunt capacitor for the transformation. The arc for the series capacitor this time is much longer, resulting in a series capacitance of 11.2 pF instead of the 2.1-pF capacitance used at the input. The actual schematic is shown Figure 3-260. The drawback of lumped elements is that it is not practical to buy components that can handle the heavy RF currents that flow in the input and output networks. Taking the values obtained

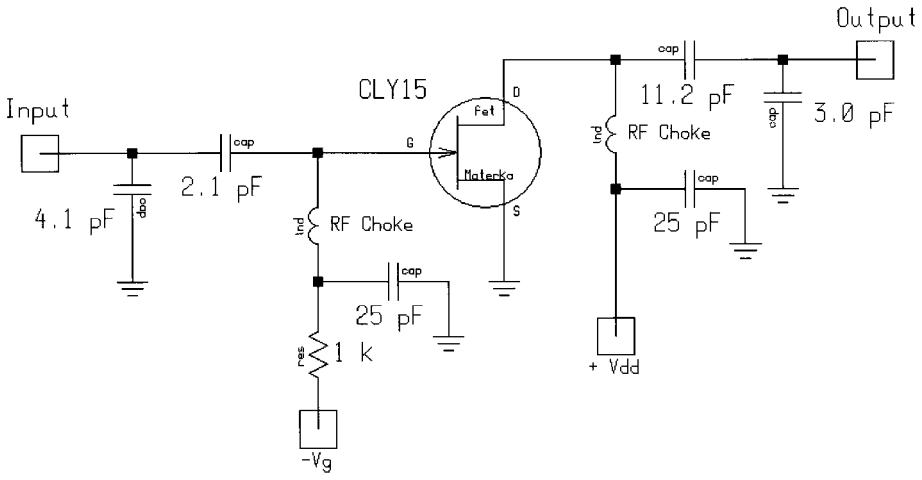


Figure 3-260 The complete amplifier using the CLY15 medium-power GaAsFET, with lumped-element input and output matching.

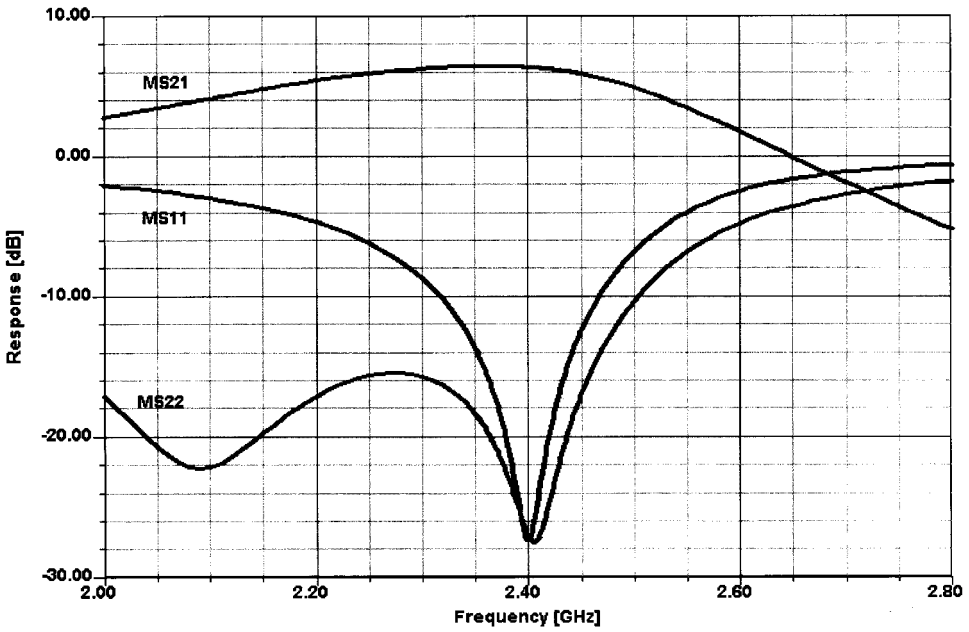


Figure 3-261 Frequency-dependent gain and matching responses of the CLY15 amplifier using ideal lumped elements.

from the lumped input and output matching approach, Figure 3-261 shows the resulting frequency response.

The next step is to provide the same matching condition using distributed elements, which of course have the advantage that they can be printed on the PC board. Instead of using ideal elements, we used a loss tangent of 2E–3, or a Q of 500. The manufacturer of the R4000 material (Rogers) provided this specification. Now we start with the same approach by moving from S_{11} to S_{11}^*. From there, we use an open stub with a transmission-line impedance of 50 Ω, followed by a transmission line to the point in the left lower corner, and, finally, a shorted transmission-line stub allows the 50-Ω matching. For this amplifier, we are not looking for noise matching but for input and output power matching. Figure 3-262 shows this match on the Smith Chart.

At the output, the procedure is again similar: We move from the S_{11} for 1 W to a conjugate value, then use an open stub that acts like a capacitor until we cross the real axis; finally, a λ/4 transmission line is used to match from approximately 75 Ω to 50 Ω. Figure 3-263 shows this match on the Smith Chart. This concludes the distributed matching design for the CLY15 amplifier.

Finally, the complete circuit using distributed elements is shown in Figure 3-264. From a dc point of view, we took advantage of the incoming shorted stub by using a transmission line of the appropriate length and a bypass capacitor to ground. This allows us to feed the gate with the required negative voltage without interfering with the frequency response of the amplifier. When evaluating the total frequency response (Figure 3-265), we noticed that this circuit seems to have a higher Q than the lumped equivalent. Again, to obtain a wider bandwidth more transformation stages will be needed. We leave this task to the interested reader.

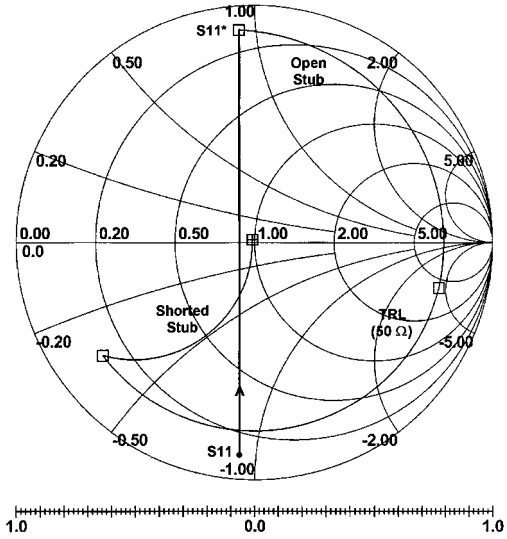


Figure 3-262 Input match for S_{11}^* (1 W) using distributed elements.

The 1-dB bandwidth of the amplifier using distributed, lossy elements is 120 MHz. If a wider bandwidth is required, matching networks with more elements are required. Because of the gain–bandwidth product, the peak gain for such a wider-bandwidth arrangement will be less than this narrowband example.

Having used a lumped to distributed transformation, we next present a power amplifier that uses coaxial lines, baluns, and other transformation steps capable of handling the increased power requirement.

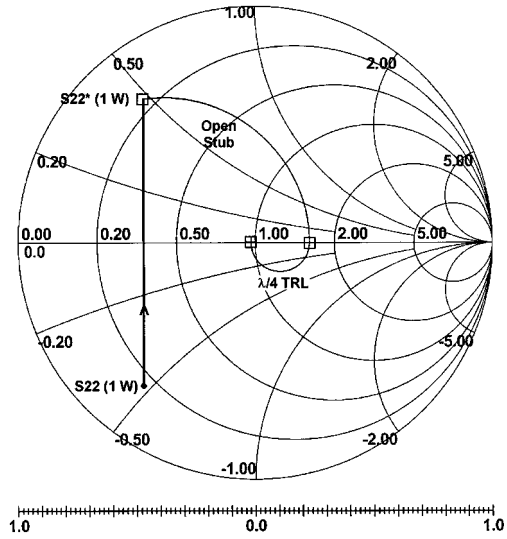


Figure 3-263 Output match for S_{22}^* (1 W) using distributed elements.

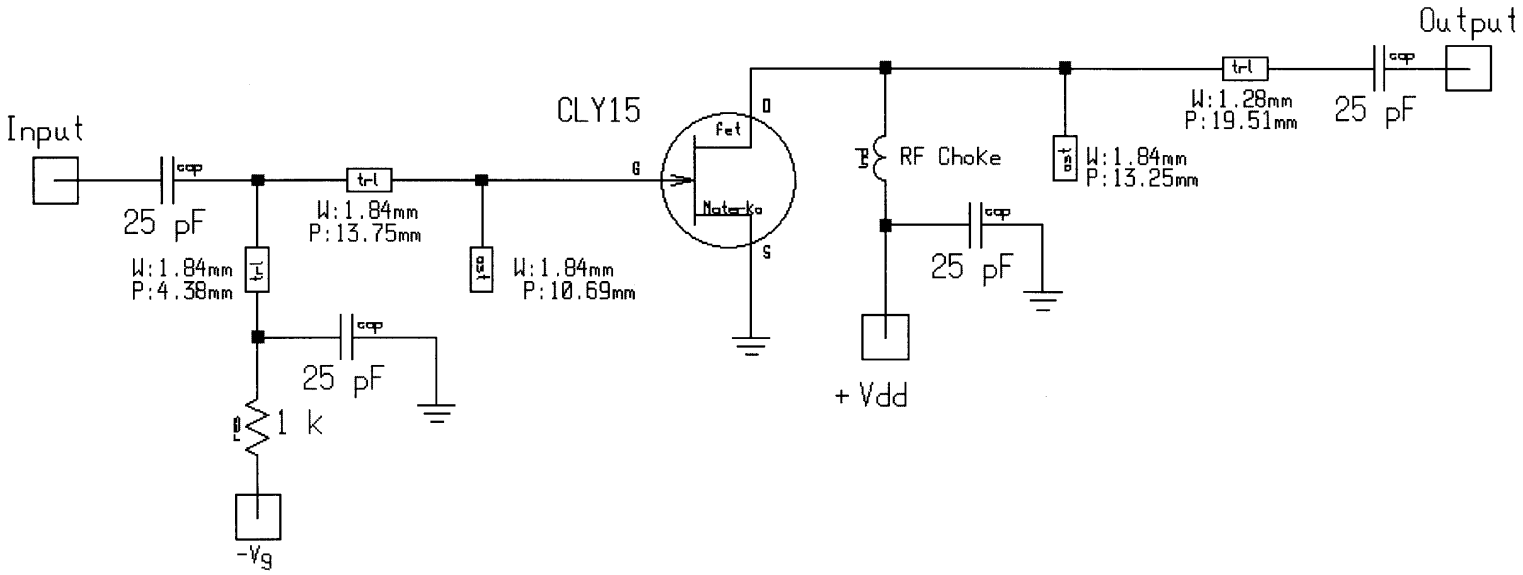


Figure 3-264 The complete amplifier using the CLY15 medium-power GaAsFET, with distributed-element input and output matching.

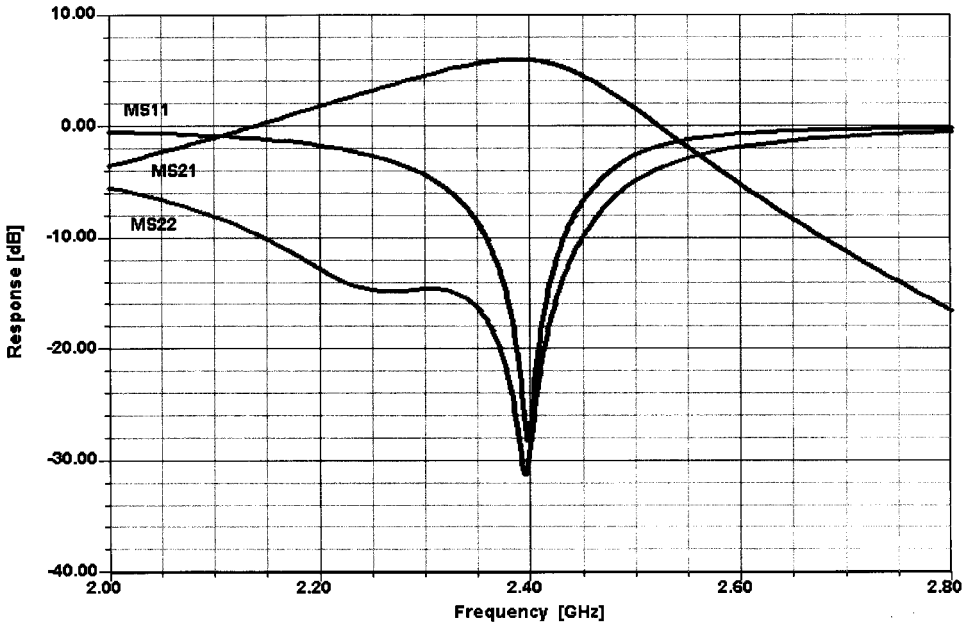


Figure 3-265 Frequency-dependent gain and matching using distributed, lossy elements.

3-12-5 Example 4: 90-W Push–Pull BJT Amplifier at 430 MHz

The last complete application we want to examine is a 90-W output, push–pull amplifier using two TRW transistors (that are no longer made—an all-too-frequent occurrence after a design is complete, resulting in a scramble for replacements). Figure 3-266 shows the circuit diagram of this UHF transistor amplifier, and Figure 3-267 shows its mechanical design. The amplifier requires 28 Vdc and will draw approximately 6 A, requiring 15 W of drive at 438 MHz. This amplifier operates in Class AB, which means it is not limited to amplifying FM and other angle-modulated signals, and can accommodate even single-sideband (SSB) operation.

The drive power coming from the source connected to the input N connector (15 W) is fed through a short piece of coaxial cable to the PC board sitting on top of the heat sink. To change from single ended to push–pull, a 50-Ω coaxial cable is coiled to provide sufficient electrical length to act as a balun that provides two symmetrical paths through the amplifier. The RC section of 10 Ω in parallel with a 220-pF capacitor stabilizes the amplifier at lower frequencies, while at higher frequencies the 10-Ω resistor has no contribution. There are two independent 4:1 transformers, one for each section; they are built using 25-Ω coaxial cable with a set of two binocular ferrite cores to force the transformation. The dc supply is also fed to the transistors at this point. The stabilization circuit consists of a temperature monitor using the *pn* junction of a transistor, side-mounted on the heat sink, and a combination of operational amplifiers and adjustments. This guarantees proper, temperature-independent biasing. The 4:1 input baluns result in 12.5 Ω impedance, or 3.125 Ω per transistor. Since the input impedance at this drive level is somewhere in the vicinity of $[(1 \ldots 2) + j(3 \ldots 5)]$ Ω; the final matching is accomplished by an *RC* section in which the inductance is actually printed

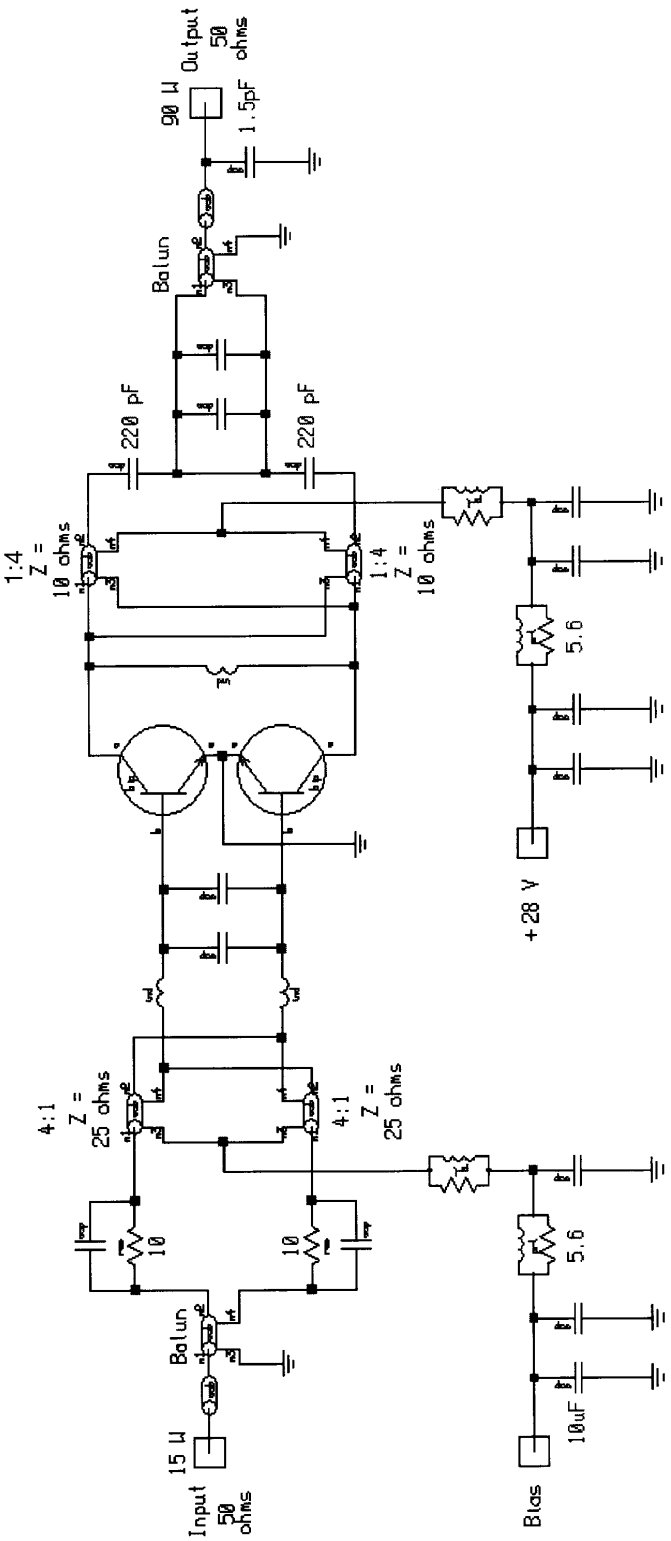


Figure 3-266 Circuit diagram of the 90-W, push–pull, 438-MHz linear amplifier.

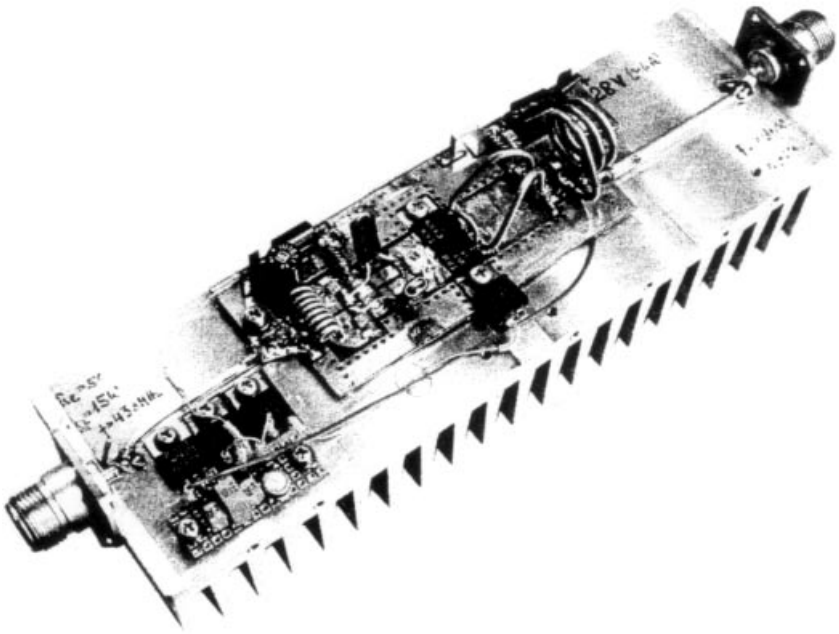


Figure 3-267 Mechanical assembly of the 90-W amplifier shown in Figure 3-266.

and the shunt capacitor across is soldered from base to base as close to the transistor bodies as possible. A ceramic-dielectric variable capacitor in parallel with this allows fine tuning.

The output consists of an inductor from collector to collector that tunes out the capacitive reactance and makes the transistor output impedance real. We then follow the same pattern by taking a 1:4 transformer, but this time built using 10-Ω rigid line instead of 25-Ω rigid line. Together with the output capacitance, prior to the output balun, this is a medium-bandwidth matching arrangement that transforms up to 50 Ω. The collector impedance per collector had to be $(28 - 2)^2/2 \times 90$, resulting in roughly 3.8 Ω—not that different from the input match at the balun point. This way, the matching networks at the input and output can be identical. At the output of the matching network, another coiled 50-Ω coaxial cable serves as balun; because it must operate at the 90-W level, larger-diameter line was used than in the input balun. Finally, another section of rigid line is used to connect the output balun on the PC board to the output N connector. The output capacitor again helps in the transformation.

Both the base and collector voltage supplies are heavily filtered using ferrite-bead-based RF chokes and heavy wire inductors wound on low-impedance resistors. A fairly large number of these amplifiers were actually built and are still in use.

3-12-6 Quasiparallel Transistors for Improved Linearity

Some applications require really low intermodulation distortion. One of the authors (Rohde), during his tenure at RCA Government Systems Division, and his team developed a parallel-device topology that is somewhat similar to the conventional parallel approach but shows vastly increased linearity. In creating higher-power devices, manufacturers parallel several transistors in one package, connecting them in a clever way so the overall drive, dissipation, and output power are equally divided between all the devices. Special nickel–chromium

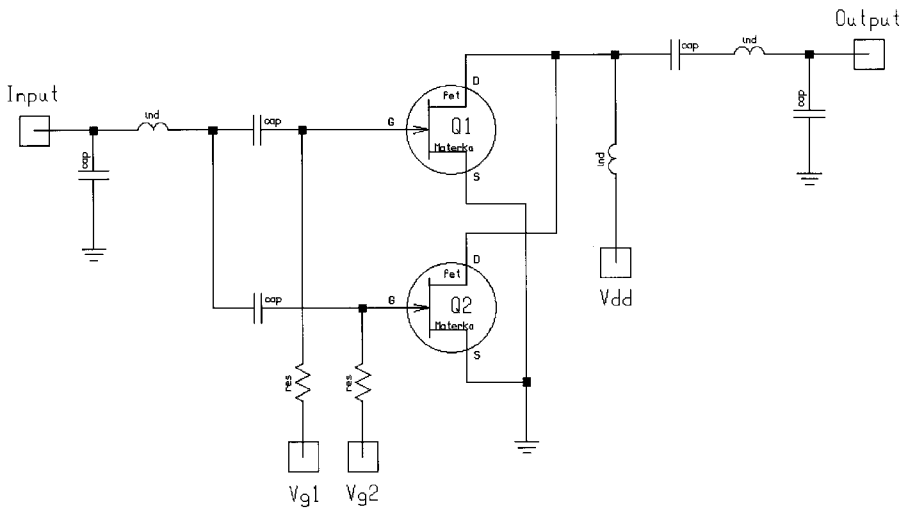


Figure 3-268 The quasiparallel amplifier topology looks like a standard parallel amplifier. Biasing the devices to different operating points nets an improvement in IMD performance.

ballasting resistors in the many fingers of the emitter connections of the multiple internal devices equalize any differences among the transistors while increasing the saturation voltage by only a small amount. In our high-linearity design, we used two transistors in parallel, one being biased somewhere between Class A and B, and the second one somewhere between Class B and C (Figure 3-268). The exact bias voltages and operating points had to be found through experiment. In practice, Q1, the "initially on" transistor, is driven to a point close to that at which its IMD products would increase. Q2 is biased such that at this same drive level it adds to the output power.

We frequently referred to this configuration as "turbocharged." The IMD products at full output power were down 40–45 dB, compared to 26–30 dB for a single transistor. While such an arrangement shows superior IMD performance, because of the somewhat abrupt kick-in of the second transistor, the transconductance response is semismooth, a characteristic that might make alert readers think that IMD will occur at this point. On the other hand, the transconductance for Q1, in effect, saturates or stays constant, so when the transconductance responses of both transistors are overlaid, the result is a smooth response overall.

This technique has been validated up to 100 W and up to at least 50 MHz. Its drawbacks include the need for two transistors with their associated costs, additional real estate, and higher input and output capacitances. Nonetheless, considering that MOS transistors capable of operation to 1 GHz were unavailable when this technique was developed, we believe that it should be revisited.

The Class D and higher operating modes require a large number of additional components and will make the efficiency higher (selective in frequency), but not necessarily reduce the IMD products. A further point of consideration is that it is incorrect to assume that reducing the input level to a given Class A will necessarily reduce its IMD products. As with tubes, the closer one gets to the knee, even in Class A operation, the worse the IMD. (This phenomenon was particularly troublesome in systems using 1-kW amplifiers up to 150 MHz;

underdrive with the intent of getting better IMD performance produced just the opposite result.) If a wide power range is considered, we recommend using a bias scheme that includes sufficient intelligence to find the best operating point.

Another hot amplifier issue is that of IMD as a function of load impedance. If an amplifier is used at output power considerably below its design value, its load impedance will be too low and, depending on the circuit configuration, the result may be additional IMD.

3-12-7 Distribution Amplifiers

It is not uncommon that a signal from a particular source like an antenna has to be distributed to several places. A condominium is a typical situation in which energy from a main signal source must be split for distribution to a larger number of users—there must be enough voltage or power to produce a noise-free picture. An amplifer capable of doing this is a *distribution* amplifier, not to be confused with a *distributed* (or traveling-wave) amplifier, such as that shown earlier in this chapter in Figure 3-6. An interesting means of achieving this is to use multiple Wilkinson couplers following an amplifier. Figure 3-269 shows an arrangement that makes heavy use of multiple power dividers to provide multiple outputs. Theoretically, the reverse is also possible: If we want to collect energy from a number of sources and combine them, as on a cable, we can use Wilkinson couplers at the *input* of a gain block. Note, however, that depending on the level and number of the combined signals, this may put enormous strain on the linearity of the amplifier. What we want to avoid is a lot of distortion, which, in the case of television, will show up as color confetti in the noise and signal.

3-12-8 Stability Analysis of a Power Amplifier

One of the hidden pitfalls in narrowband circuit design is that a circuit may be unstable outside the frequency range of interest. Circuits containing structures built from distributed elements, which exhibit multiple resonances, can be particularly susceptible to such instabilities. In this example, we analyze the stability of a simple FET power amplifier designed using small-signal concepts and the interactive Smith Tool utility in Ansoft's Serenade

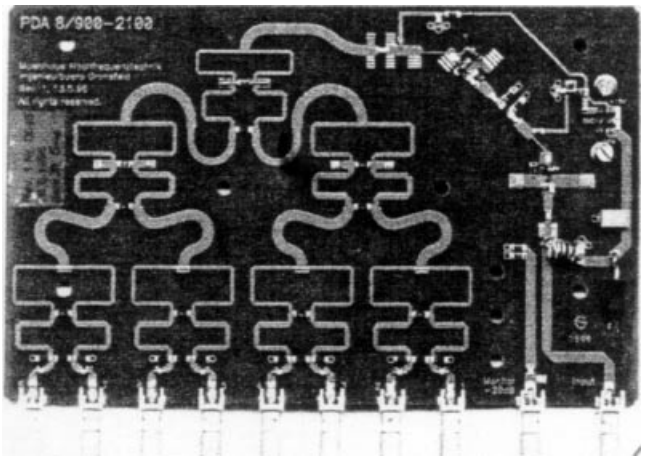


Figure 3-269 Distribution amplifier showing the use of multiple Wilkinson power dividers.

Design Environment. The evaluation shows how K-factor analysis alone is not sufficient to check for stability of a design.

Small-Signal ac Analysis. We begin by taking the Figure 3-270 amplifier, which was designed for about 16 dB of gain at 2.2 GHz. Due to the very low, capacitive input impedance of the FET and the proximity of the source-plane gain circles and stability circles to the edge of the Smith Chart, the matching network attempts to transform 50 Ω to an impedance of approximately $(2.5 + j15)$ Ω. The impedance lies close to the stability circles at 2 GHz but remains in a stable region. The bias point selected (–3.5 V) is chosen for Class AB operation. Higher linear gain can be achieved by reducing the gate voltage to –2.0 V; at this bias point, the amplifier gain will increase to about 19 dB.

First, we will do a small-signal ac analysis and examine the S parameters and K and B_1 factors [see Eqs. (3-154) and (3-157), respectively] from 1 GHz to 6 GHz—well beyond the bandwidth of the design. Figure 3-271 shows the circuit's frequency-dependent gain and matching characteristics.

If we examine B_1 and the stability factor K, we see that the amplifier is not unconditionally stable in the 2.2-GHz passband as shown in Figure 3-272. K dips slightly below 1, although B_1 remains above zero. In practice, a slight change in design would be required to bring K above 1 and remove this potential instability. However, for our example, this is of no consequence, as we shall see below.

Nyquist Stability Analysis. Now we will utilize a Nyquist stability analysis to check the stability of the amplifier. To fully realize the advantages of a Nyquist stability analysis, we must sweep frequency over a very wide range. We want to be able to pick up instabilities at *any* frequency, not just near the design frequency. In this example we sweep from 1 kHz to 20 GHz. A start frequency of 1 kHz is chosen so we can determine the starting point of the Nyquist plot and we ensure that any low-frequency resonances are picked up. The stop frequency is chosen past the f_{max} of the transistor (where we know it cannot oscillate).

The analysis results are shown in polar plot form in Figure 3-273 and in magnitude–angle form in Figure 3-274. The polar graph shows the real and imaginary parts of the system determinant as frequency is swept. The magnitude–angle graph shows the magnitude and cumulative angle (no cut at ±180°) of the system determinant as frequency is swept. The key to interpreting these graphs lies in the following statement drawn from the Nyquist criterion:

The circuit will be unstable if the Nyquist plot encircles the origin in a clockwise direction.

The term "unstable" refers to the existence of natural frequencies lying in the right-half of the complex frequency plane (RHP). These natural frequencies on the RHP will cause oscillations. To encircle the origin in a clockwise direction means that the Nyquist plot will travel in a clockwise path, cross the negative real axis, and continue to completely surround the origin. If the path appears to encircle the origin but then "unravels" itself in a counter-clockwise direction, the origin is not actually encircled. Figure 3-273 shows that the origin is not encircled and therefore this circuit is stable.

The magnitude–angle plot can assist in determining if a clockwise encirclement has actually been made with the help of the following:

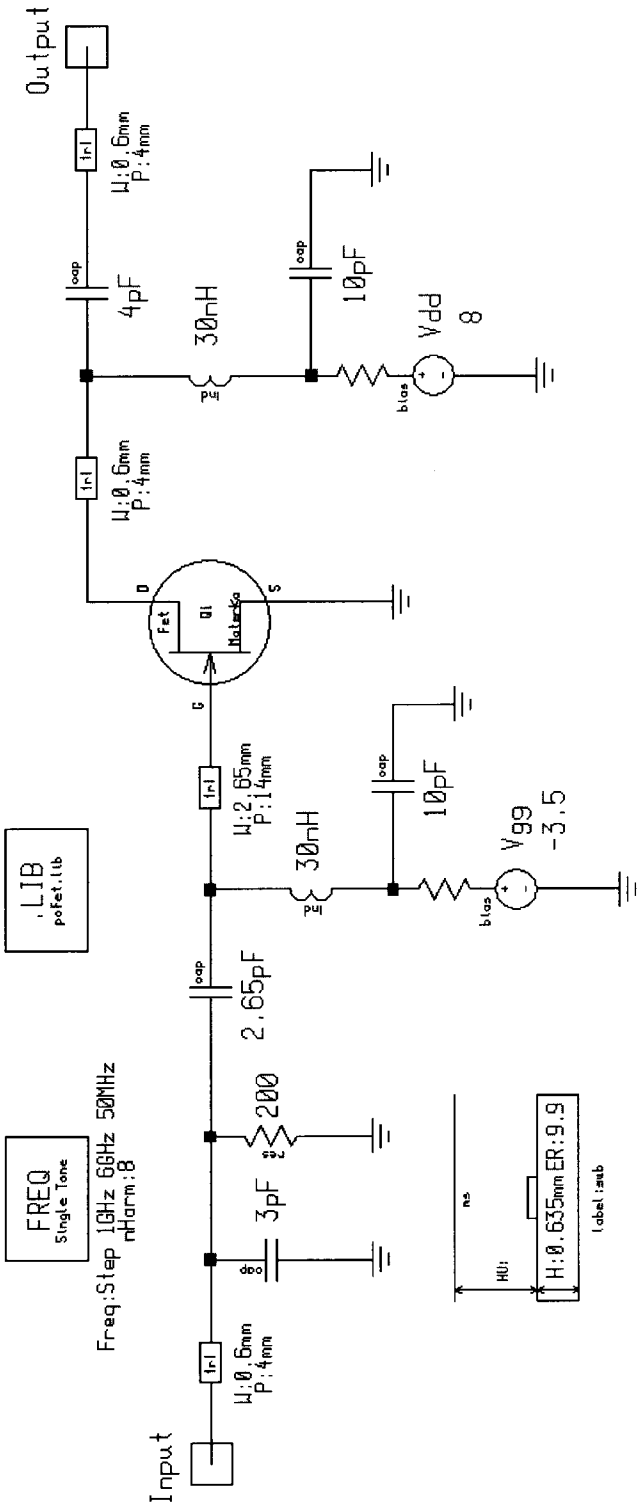


Figure 3-270 A 2.2-GHz amplifier design for stability analysis.

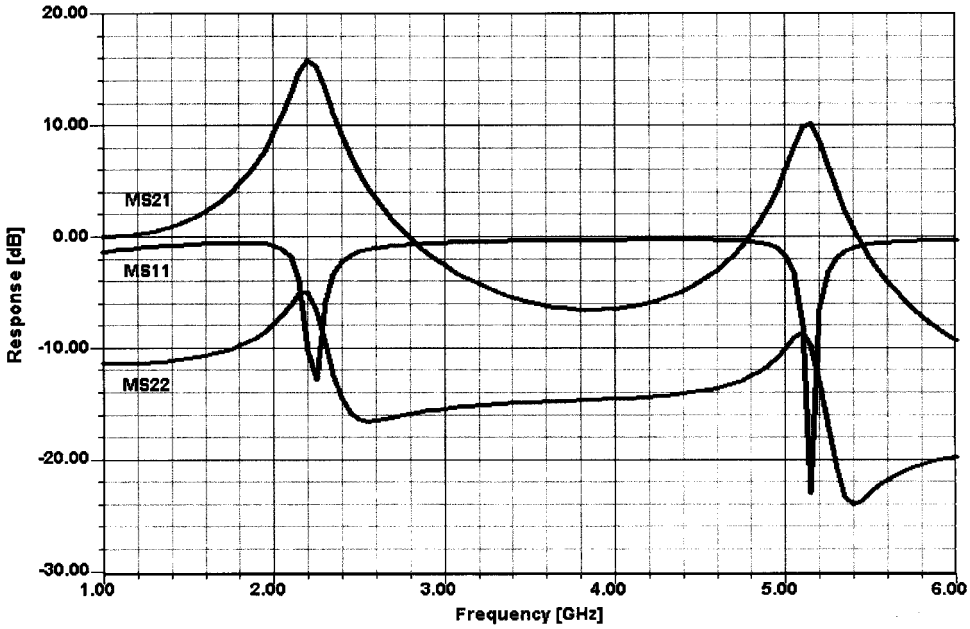


Figure 3-271 Frequency-dependent gain and matching characteristics of the amplifier. Although the circuit is designed for 2.2-GHz operation, the periodic responses of its distributed-element-based input and output matching networks result in a second gain peak near 5.1 GHz. As we shall see, however, instabilities may result at frequencies that cannot be intuited from this graph.

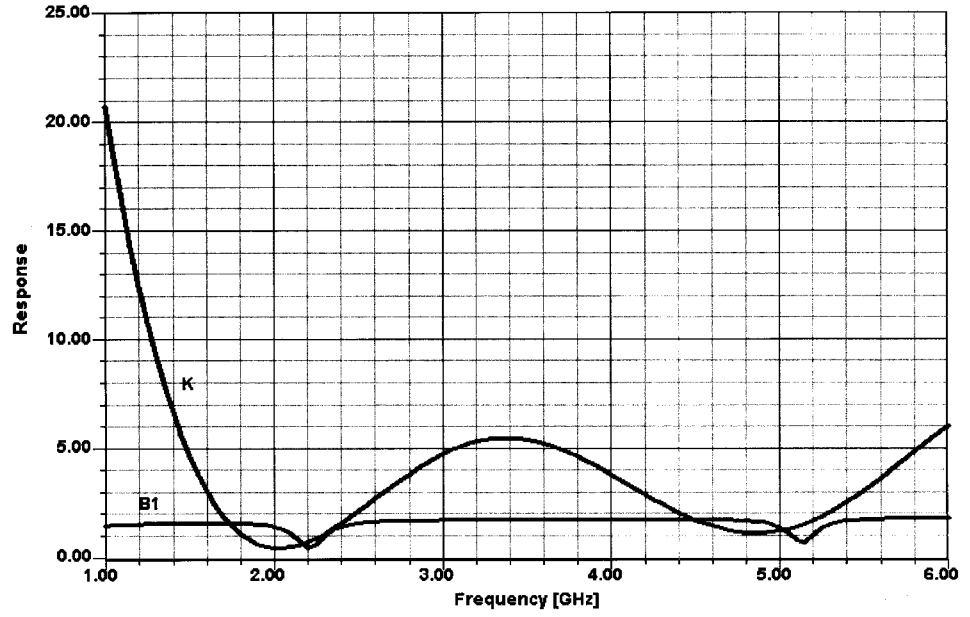


Figure 3-272 K and B_1 for the 2.2-GHz amplifier.

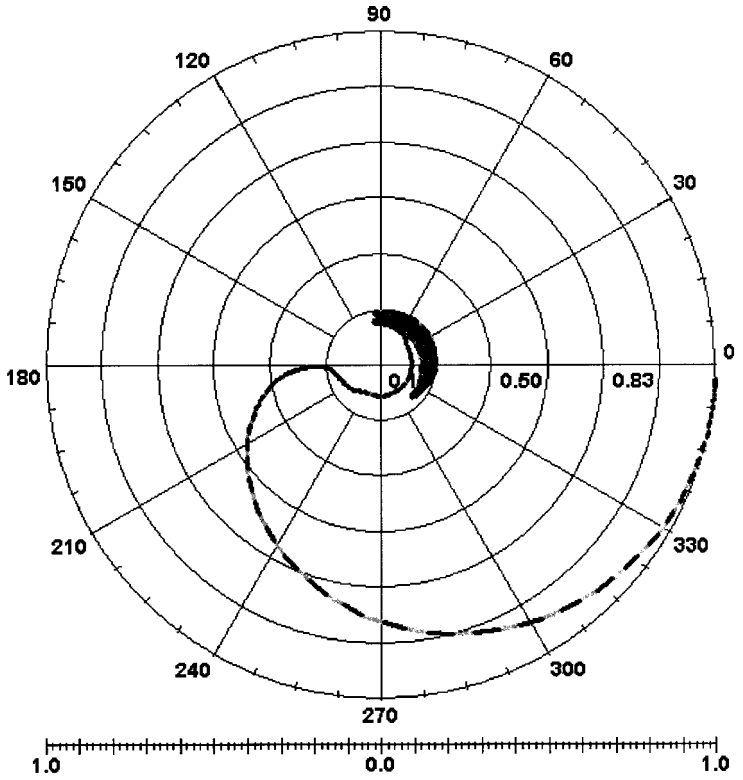


Figure 3-273 Polar Nyquist plot for the 2.2-GHz amplifier covering 1 kHz to 20 GHz.

- To travel in a clockwise direction on a polar chart means to decrease phase angle. On a magnitude–angle plot, the phase will become more negative as frequency is increased (although it may not be monotonic).
- If the path crosses the negative real axis on the polar chart, it will cross –180° on the magnitude–angle plot.
- Encirclement will occur if the phase angle continues to decrease to, say, –360°. If "unraveling" occurs, the phase angle will not stay below –360°. Figure 3-274 clearly shows that the phase angle decreases to about –180° but does not decrease any more. Rather, the angle returns to a positive value and oscillates between about –40° and +100°. Therefore an encirclement did not occur, and the circuit is stable.

An Unstable Case. Let us now change the gate bias to –2 V—a value that increases the amplifier gain enough to cause instability. Figures 3-275 and 3-276 show the results. From Figure 3-275 we can see that the origin is encircled in a clockwise direction. Figure 3-276 confirms this since the phase angle goes through –180° and continues without unraveling.

Approximate Frequency of Oscillation. The frequency where the Nyquist plot crosses –180° is often the approximate frequency of the unstable right-half-plane zeros and provides an estimate of the potential oscillation. For this example, there is a sharp crossing at 200 MHz that is characteristic of a circuit resonance. Note that the frequency is not related

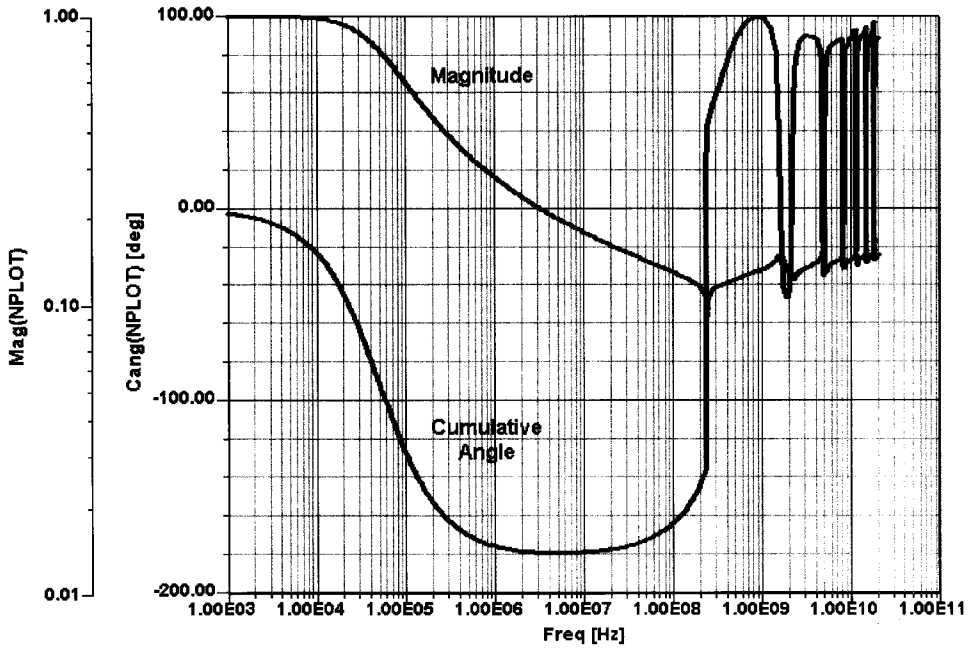


Figure 3-274 Magnitude and cumulative angle Nyquist plot for the 2.2-GHz amplifier covering 1 kHz to 20 GHz.

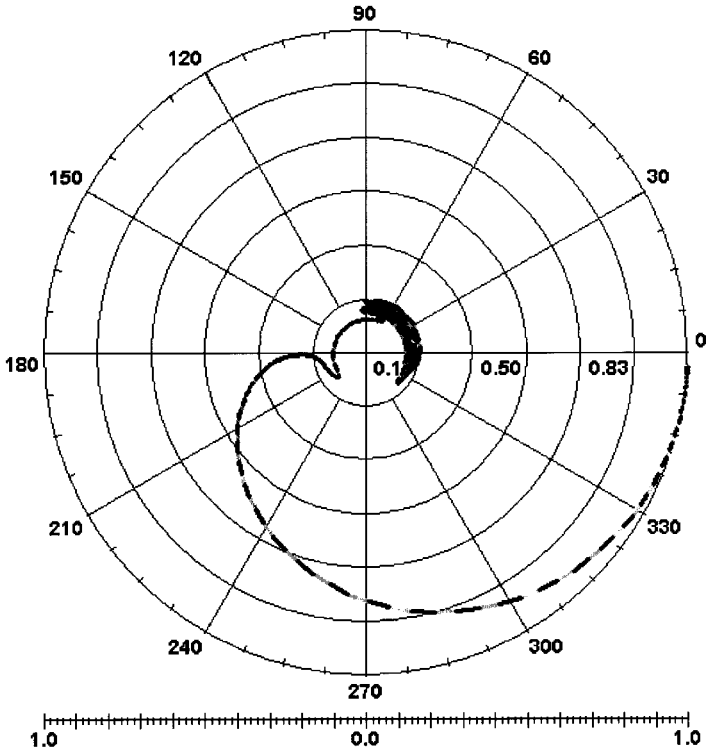


Figure 3-275 Polar Nyquist plot of the unstable amplifier.

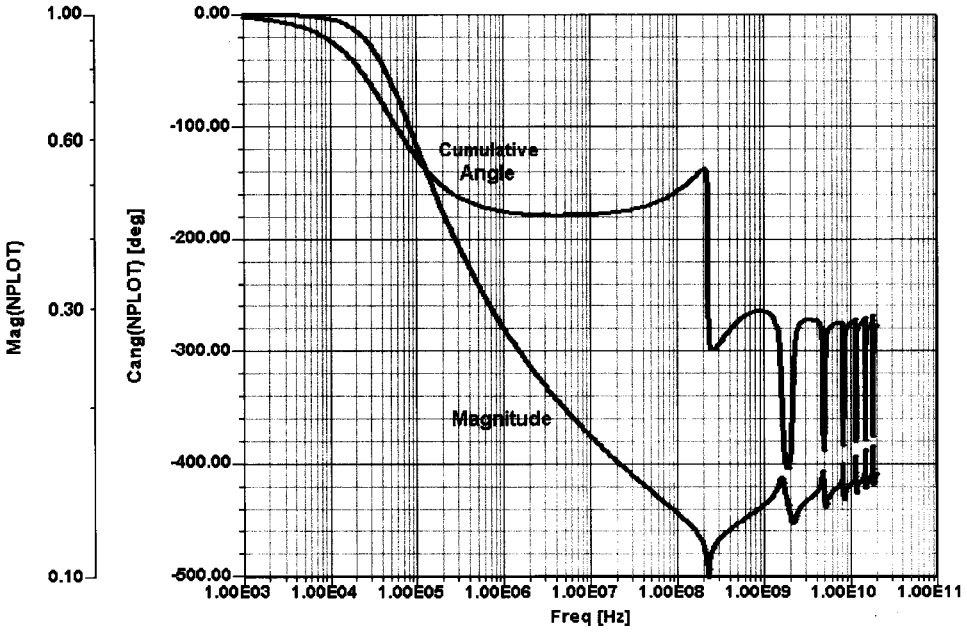


Figure 3-276 Magnitude–cumulative angle Nyquist plot of the unstable amplifier.

to a resonant frequency in a rigorous mathematical fashion, but it has been our experience that this frequency approximates the resonant frequency. To confirm the circuit's readiness for oscillation in the vicinity of this frequency, we evaluate it with Serenade's Oscillator Design Aid from 10 MHz to 1 GHz. The Design Aid finds an initial resonant frequency near 230.6 MHz—not too far from the –180° crossing frequency of 200 MHz of the Nyquist plot. Figure 3-277 shows this point.

Oscillator analysis finds a final oscillation frequency of 228.48 MHz. Figure 3-278 shows the resulting output spectrum at Port 2. At this point we decided to do a quick jump into Chapter 5 and take advantage of the CAD capability to predict the resulting phase noise, which is shown in Figure 3-279. As we compare this with other measured results, the approximately 228-MHz oscillator has a poor phase noise since the same phase noise is achievable at 1 GHz and higher. We remind the reader that this is an unwanted effect, but all of us who have built amplifiers will have experienced either low-frequency oscillations (motorboating) or very-high-frequency oscillation, sometimes caused by the components in the immediate vicinity of the transistor itself and found somewhere above 3 GHz. This case is really a demonstration case; a closer look reveals that the transconductance required for this condition is unrealistically high; but nonetheless it is a good example to demonstrate the stability issue.

Where Does Oscillation Begin? An interesting question to solve now is, "What gate bias voltage is needed for oscillation to begin?" Ansoft's Serenade Design Environment includes oscillator synchronous stability analysis, which uses a modification of the harmonic-balance technique to determine critical points, such as the voltage needed to "turn on" an oscillator. Mathematically, this is called a *Hopf bifurcation* because the behavior of

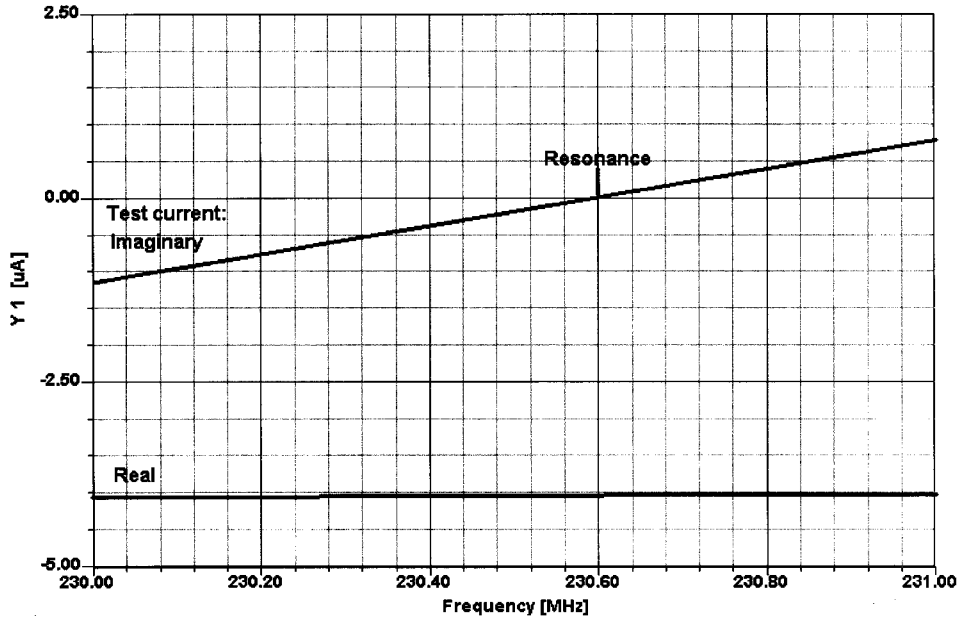


Figure 3-277 Graph from Serenade Oscillator Design Aid analysis of the amplifier. A resonant frequency is indicated where the imaginary part of the test current equals zero and the real part is negative.

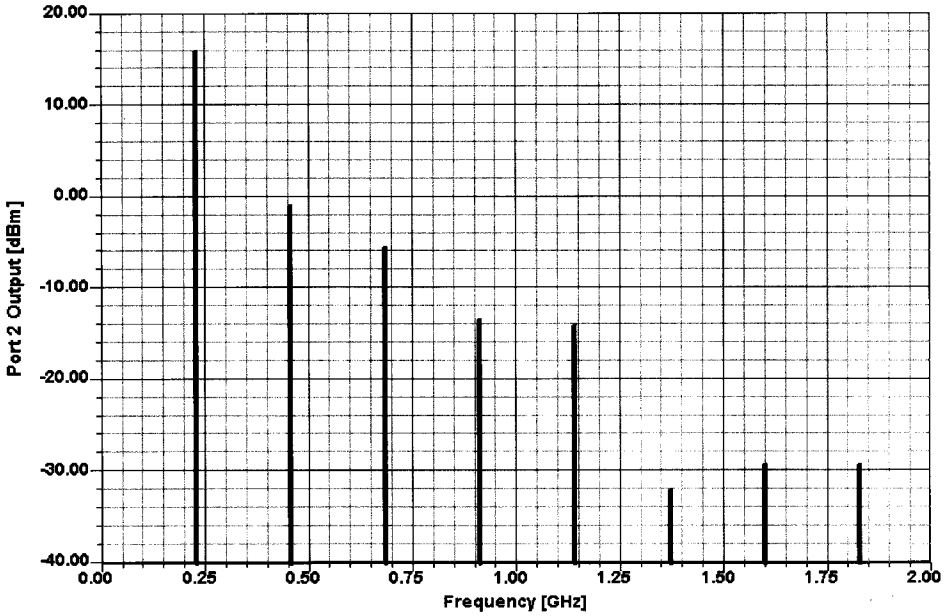


Figure 3-278 Predicted output spectrum of the unintended oscillator.

the circuit changes dramatically from a dc steady state to an oscillatory response. Figure 3-280 shows this effect by plotting output power versus gate bias. The turning point at –3.2 V (22 mW) is clearly visible, as is the bifurcation at –3.1 V (0 mW output).

Figure 3-280 can be interpreted as follows:

1. If the bias voltage is increased from, say, –4.0 V, the circuit will begin to oscillate at the bifurcation voltage and the output power will jump to 26 mW.
2. If the circuit is oscillating and the voltage is decreased from –2.0 V, the circuit will cease oscillating at the turning voltage of –3.2 V.
3. The branch between the turning point and the bifurcation point is unstable and cannot be realized in practice. This forms a hysteresis loop around points (–3.1 V, 0 mW), (–3.1 V, 26 mW), (–3.2 V, 22 mW), and (–3.2 V, 0 mW).

This amplifier stability analysis section is based on the stability analysis examples provided by Ansoft for the Serenade Design Environment.

An interesting example that combines both CAD and the feed-forward amplifier technique can be found in the Serenade Design Environment system simulator examples manual. This technique has been under evaluation for the last two years, but our feeling is that not enough measured data at frequencies above 500 MHz with reliable performance information have yet been made available in the literature. Along these lines, work done at Ansoft under a U.S. Air Force contract has provided very good insight into high-efficiency amplifiers at microwave frequencies [46].

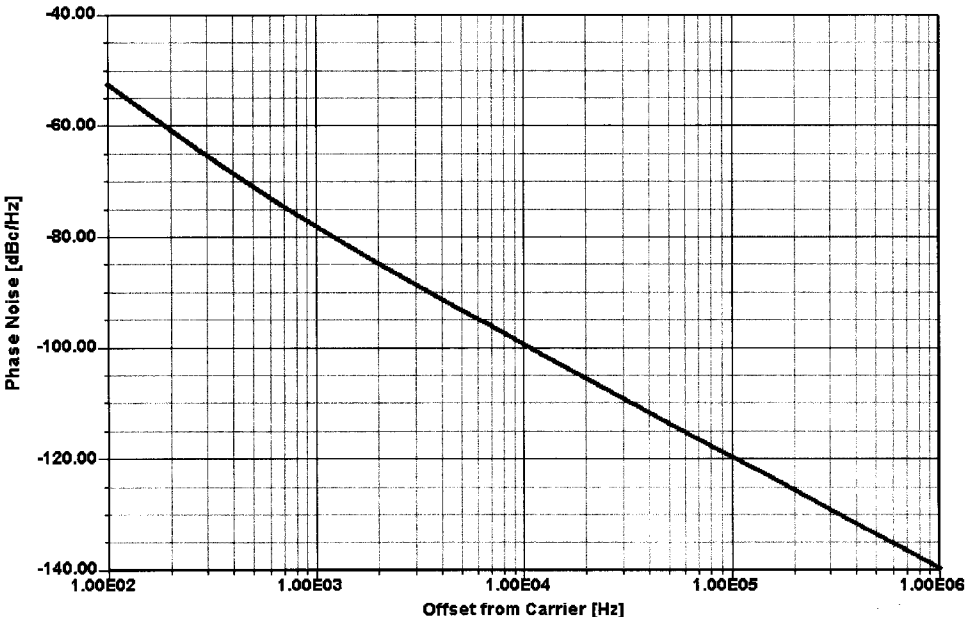


Figure 3-279 Predicted phase noise of the oscillating amplifier.

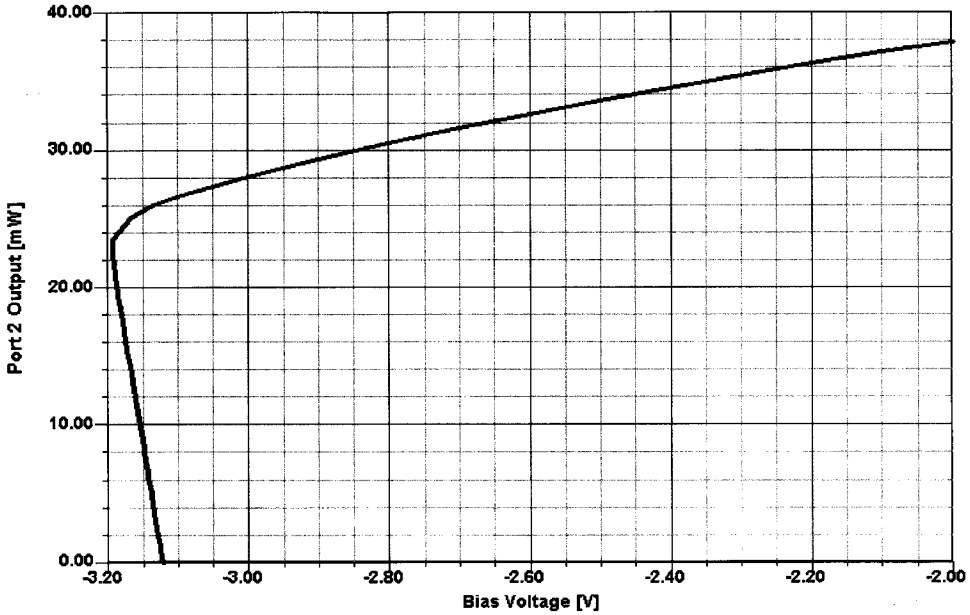


Figure 3-280 Output power versus gate bias for the unstable amplifier.

3-13 POWER AMPLIFIER DATASHEETS AND MANUFACTURER-RECOMMENDED APPLICATIONS

Having discussed the various aspects of BJTs and FETs, we have decided to use MOS examples from Ericsson and Motorola, and a bipolar example from Motorola, as good examples to show the technology required to build high-power transistor amplifiers in this frequency range. In the following pages, we reproduce, with permission, datasheets for Ericsson's PTF 10009 FET and Motorola's MRF186 FET, followed by Motorola's MRF899 BJT.

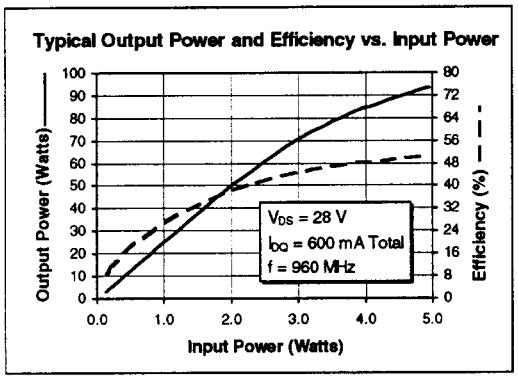
ERICSSON

PTF 10009
85 Watts, 1.0 GHz
LDMOS Field Effect Transistor

Description

The 10009 is a common source N-channel enhancement-mode lateral MOSFET intended for large signal amplifier applications to 1.0 GHz. It is rated at 85 watts minimum output power. Nitride surface passivation and gold metallization are used to ensure excellent device lifetime and reliability. 100% lot traceability is standard.

- Performance at 960 MHz, 28 Volts
  - Output Power = 85 Watts
  - Power Gain = 13.0 dB Typ
  - Efficiency = 50% Typ
- Gold Metallized
- Silicon Nitride Passivated
- Excellent Thermal Stability



Package 20230

Maximum Ratings

Parameter	Symbol	Value	Unit
Drain-Source Voltage	V_{DSS}	65	Vdc
Gate-Source Voltage	V_{GS}	±20	Vdc
Operating Junction Temperature	T_J	200	°C
Total Device Dissipation at T_{flange} = 25°C Above 25°C derate by	P_D	270 1.54	Watts W/°C
Storage Temperature Range	T_{STG}	-65 to 150	°C
Thermal Resistance (T_{flange} = 70°C)	$R_{\theta JC}$	0.65	°C/W

PTF 10009

ERICSSON

Electrical Characteristics (100% Tested)

Characteristic (per side)	Conditions	Symbol	Min	Typ	Max	Units
Drain-Source Breakdown Voltage	V_{GS} = 0 V, I_D = 25 mA	$V_{(BR)DSS}$	65	—	—	Volts
Drain-Source Leakage Current	V_{DS} = 28 V, V_{GS} = 0 V	I_{DSS}	—	—	1.0	mA
Gate Threshold Voltage	V_{DS} = 10 V, I_D = 75 mA	$V_{GS(th)}$	—	2.0	—	Volts
Forward Transconductance	V_{DS} = 10 V, I_D = 3 A	g_{fs}	—	2.8	—	Siemens

Dynamic Characteristics

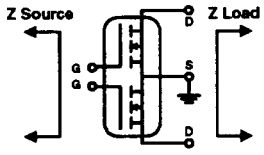
Characteristic (per side)	Symbol	Min	Typ	Max	Units
Input Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{iss}	—	90	—	pF
Output Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{oss}	—	36	—	pF
Reverse Transfer Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{rss}	—	1.9	—	pF

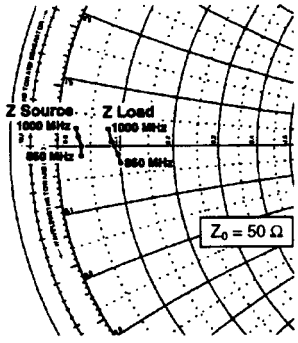
RF Specifications (100% Tested)

Characteristic	Symbol	Min	Typ	Max	Units
Gain (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz)	G_{ps}	12.0	13.0	—	dB
Drain Efficiency (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz)	η	47	50	—	%
Load Mismatch Tolerance (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz— all phase angles at frequency of test)	Ψ	—	—	5:1	—

Impedance Data (data shown for fixed-tuned broadband circuit)

(V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA)



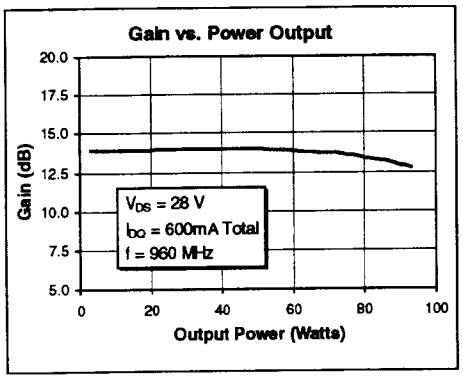


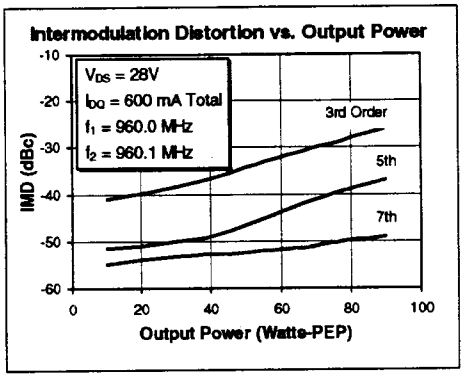
Frequency	Z Source		Z Load	
MHz	R	jX	R	jX
860	1.76	-0.78	5.00	-1.50
900	1.80	-0.05	4.80	-0.78
960	1.58	0.69	4.24	0.36
1000	1.39	1.35	3.95	1.41

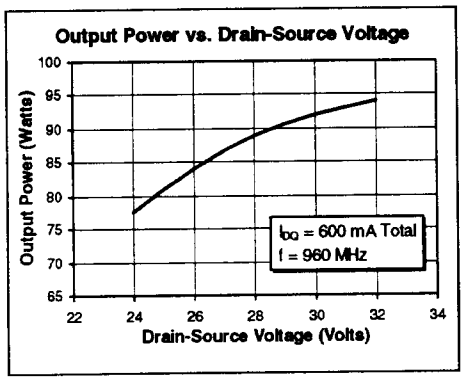
ERICSSON

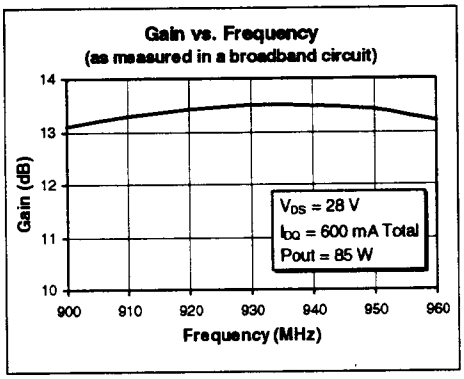
PTF 10009

Typical Performance





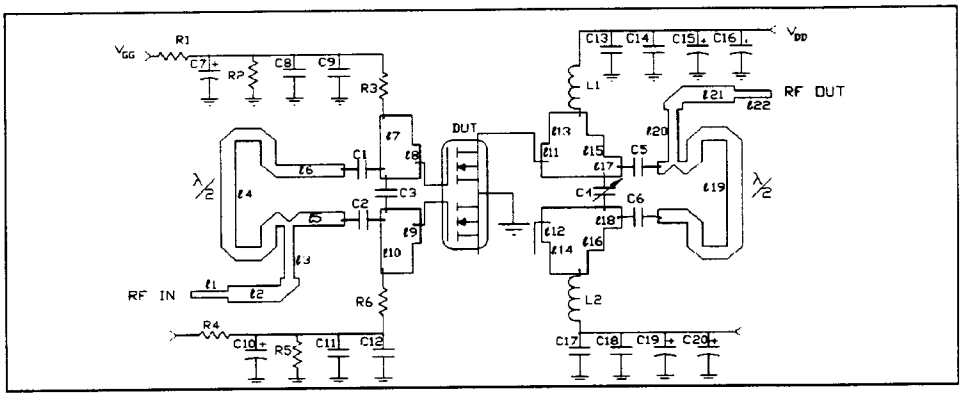




PTF 10009

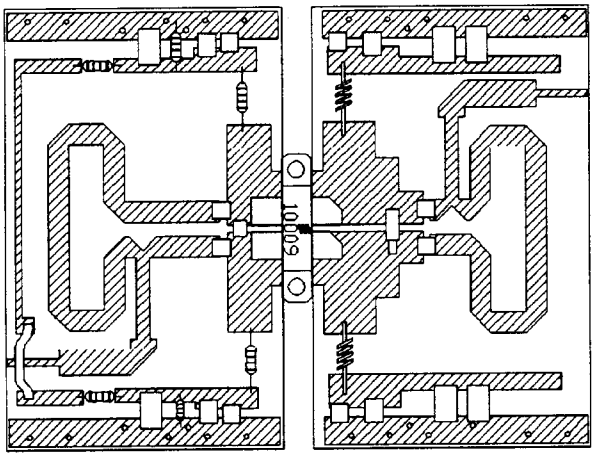
ERICSSON

Test Circuit



Schematic for f = 960 MHz

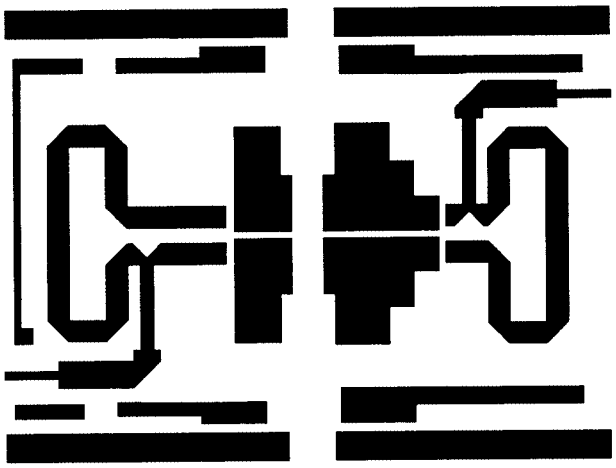
DUT	10009
C1-2, C5-6, C9, C12-13, C17	33 pF, Capacitor ATC 100 B
C3	11 pF, Capacitor ATC 100 B
C4	6.0 pF, Variable Capacitor, JMC 5701
C7, C10	10 µF, +10 V Electrolytic Capacitor
C8, C11, C14, C18	0.01 µF, Capacitor ATC 100 B
C15, C16, C19, C20	10 µF, +30 V Electrolytic Capacitor
L1. L2	4 Turn, #20 AWG, .120" I.D.
R1, R2, R4, R5	1.0 K, Ω Resistor
R3, R6	5.1 K, 1/4 Ω Resistor
ℓ1, ℓ22	50 Ω, .030 λ
ℓ2, ℓ21	20 Ω, .080 λ
ℓ3, ℓ20	32 Ω, .191 λ
ℓ4, ℓ19	25 Ω, .500 λ
ℓ5, ℓ6	25 Ω, .091 λ
ℓ7, ℓ10	7 Ω, .056 λ
ℓ8, ℓ9	13.0 Ω, .017 λ
ℓ11, ℓ12	13.0 Ω, .017 λ
ℓ13, ℓ14	7.0 Ω, .064 λ
ℓ15, ℓ16	10.0 Ω, .029 λ
ℓ17, ℓ18	19.0 Ω, .028 λ
Circuit Board	.031" Thick, ε_r = 4.0, AlliedSignal, G200



Parts Layout (not to scale)

ERICSSON

PTF 10009



Artwork (1 inch ——— *)*

Ericsson Components
RF Power Products
675 Jarvis Drive
Morgan Hill, CA 95037 USA
Telephone: 408-778-9434

1-877-GOLDMOS
(1-877-465-3667)
e-mail: rfpower@ericsson.com
www.ericsson.com/rfpower

Specifications subject to change
without notice.
L1
© 1997 Ericsson Inc.
EUS/KR 1301-PTF 10009 Uen Rev. B 09-29-98

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF186/D

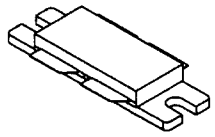
The RF MOSFET Line
RF Power Field-Effect Transistor
N–Channel Enhancement–Mode Lateral MOSFET

MRF186

120 W, 1.0 GHz
28 VOLTS
LATERAL N–CHANNEL
BROADBAND
RF POWER MOSFET

Designed for broadband commercial and industrial applications at frequencies from 800 MHz to 1.0 GHz. The high gain and broadband performance of this device makes it ideal for large–signal, common source amplifier applications in 28 volt base station equipment.

- Guaranteed Performance @ 960 MHz, 28 Volts
  Output Power — 120 Watts (PEP)
  Power Gain — 11 dB
  Efficiency — 30%
  Intermodulation Distortion — –28 dBc
- Excellent Thermal Stability
- 100% Tested for Load Mismatch Stress at all Phase Angles with 5:1 VSWR @ 28 Vdc, 960 MHz, 120 Watts CW



CASE 375B–02, STYLE 2

MAXIMUM RATINGS (2)

Rating	Symbol	Value	Unit
Drain–Source Voltage	V_{DSS}	65	Vdc
Drain–Gate Voltage (R_{GS} = 1 MΩ)	V_{DGR}	65	Vdc
Gate–Source Voltage	V_{GS}	±20	Vdc
Drain Current — Continuous	I_D	14	Adc
Total Device Dissipation @ T_C = 70°C Derate above 70°C	P_D	162.5 1.25	Watts W/°C
Storage Temperature Range	T_{stg}	– 65 to +150	°C
Operating Junction Temperature	T_J	200	°C

THERMAL CHARACTERISTICS (2)

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	0.8	°C/W

NOTE – **CAUTION** – MOS devices are susceptible to damage from electrostatic charge. Reasonable precautions in handling and packaging MOS devices should be observed.

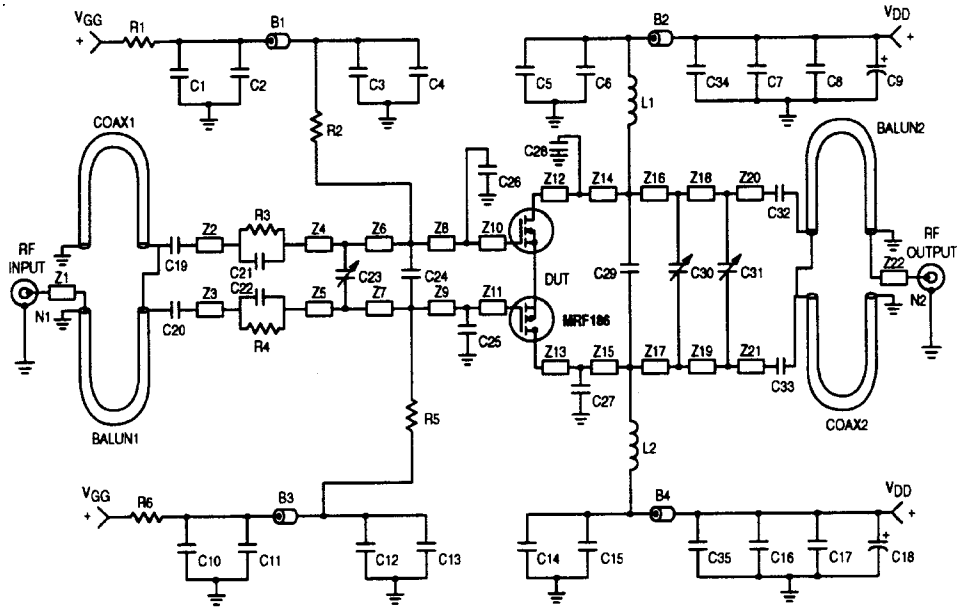
© Motorola, Inc. 1999

MOTOROLA

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted)

Characteristic	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS (1)					
Drain–Source Breakdown Voltage (V_{GS} = 0 Vdc, I_D = 50 μAdc)	$V_{(BR)DSS}$	65	—	—	Vdc
Zero Gate Voltage Drain Current (V_{DS} = 28 Vdc, V_{GS} = 0)	I_{DSS}	—	—	1	μAdc
Gate–Source Leakage Current (V_{GS} = 20 Vdc, V_{DS} = 0)	I_{GSS}	—	—	1	μAdc
ON CHARACTERISTICS (1)					
Gate Quiescent Voltage (V_{DS} = 26 Vdc, I_D = 300 μAdc Per Side)	$V_{GS(th)}$	2.5	3	4	Vdc
Gate Quiescent Voltage (V_{DS} = 26 Vdc, I_D = 300 mAdc Per Side)	$V_{GS(Q)}$	3.3	4.2	5	Vdc
Delta Gate Threshold Voltage (Side to Side) (V_{DS} = 28 V, I_D = 300 mA Per Side)	$\Delta V_{GS(Q)}$	—	—	0.3	Vdc
Drain–Source On–Voltage (V_{GS} = 10 Vdc, I_D = 3 Adc Per Side)	$V_{DS(on)}$	—	0.58	0.7	Vdc
Forward Transconductance (V_{DS} = 10 Vdc, I_D = 3 Adc Per Side)	g_{fs}	2.4	2.8	—	S
DYNAMIC CHARACTERISTICS (1)					
Input Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{iss}	—	177	—	pF
Output Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{oss}	—	45	—	pF
Reverse Transfer Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{rss}	—	3.4	—	pF
FUNCTIONAL CHARACTERISTICS (In Motorola Test Fixture) (2)					
Two–Tone Common Source Amplifier Power Gain (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	G_{ps}	11	12.2	—	dB
Two–Tone Drain Efficiency (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	η	30	35	—	%
3rd Order Intermodulation Distortion (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	IMD	—	–32	–28	dBc
Input Return Loss (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	IRL	9	16	—	dB
Two–Tone Common Source Amplifier Power Gain (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	G_{ps}	—	12	—	dB
Two–Tone Drain Efficiency (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	η	—	33	—	%
3rd Order Intermodulation Distortion (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	IMD	—	–32	—	dBc
Input Return Loss (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	IRL	—	16	—	dB
Output Mismatch Stress (V_{DD} = 28 Vdc, P_{out} = 120 W CW, I_{DQ} = 2 x 400 mA, f = 960 MHz, VSWR = 5:1, All Phase Angles at Frequency of Tests)	Ψ	No Degradation In Output Power Before and After Test			

(1) Each side of device measured separately.
(2) Device measured in push–pull configuration.



B1 – B4	Fair Rite Products Short Ferrit Bead, 2743021446	C31	0.8 – 8.0 pF, Variable Capacitor, Johanson Gigatrim
C1, C7, C8, C10, C16, C17	10 µF, 50 V, Tantalum	L1, L2	3 Turns, #20 AWG, IDIA 0.126", 24.7 nH
C2, C11, C34, C35	0.1 µF, Chip Capacitor	N1, N2	Type N Connectors
C3, C6, C12, C15	330 pF, Chip Capacitor	R1, R6	1 kΩ, 1/4 W, Carbon Resistor
C4, C5, C13, C14, C19, C20, C32, C33	47 pF, Chip Capacitor	R2, R5	1.2 kΩ, 0.1 W, Chip Resistor
C9, C18	250 µF, 50 V, Electrolytic Capacitor	R3, R4	75 Ω, 0.1 W, Chip Resistor
C21, C22	12 pF, Chip Capacitor	Z1 – Z22	Microstrip (See Component Placement)
C23, C30	0.6 – 4.5 pF, Variable Capacitor, Johanson Gigatrim	Balun1, Balun2, Coax1, Coax2	2.20" 50 Ω, 0.086" OD Semi–Rigid Coax
C24, C25, C26	5.1 pF, Chip Capacitor	Board	1/32" Glass Teflon®, ε_r = 2.55
C27, C28	3.9 pF, Chip Capacitor		

Figure 1. 930 – 960 MHz Test Circuit Schematic

TYPICAL CHARACTERISTICS

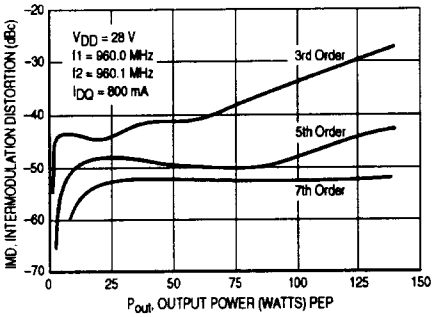


Figure 2. Intermodulation Distortion versus Output Power

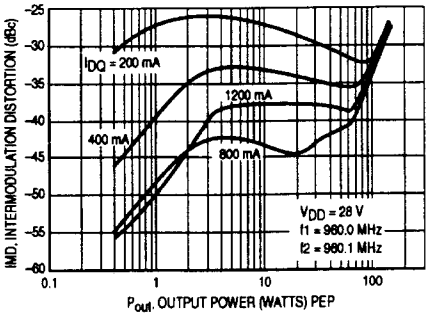


Figure 3. Intermodulation Distortion versus Output Power

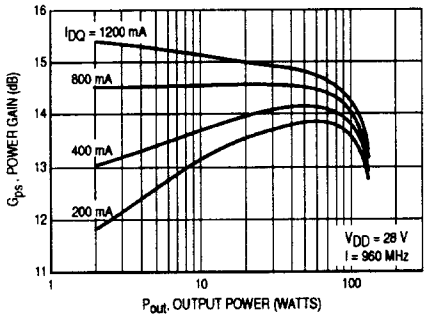


Figure 4. Power Gain versus Output Power

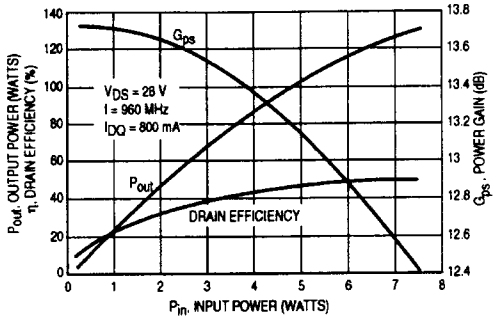


Figure 5. Output Power versus Input Power

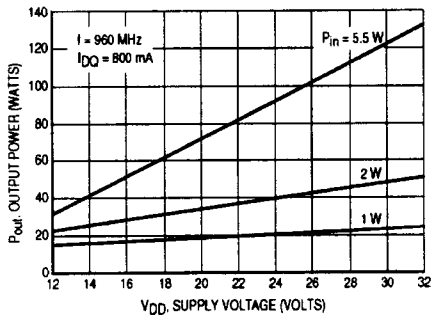


Figure 6. Output Power versus Supply Voltage

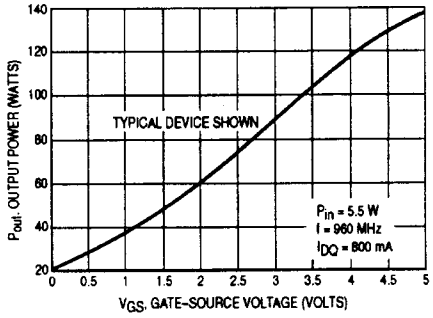


Figure 7. Output Power versus Gate Voltage

TYPICAL CHARACTERISTICS

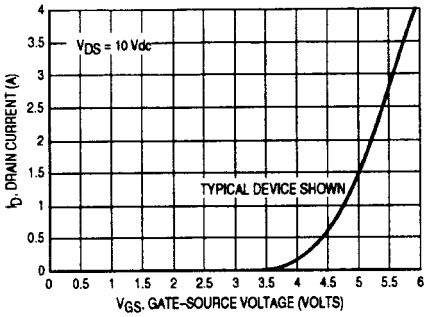


Figure 8. Drain Current versus Gate Voltage

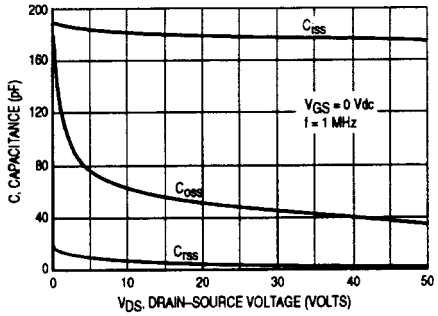


Figure 9. Capacitance versus Voltage

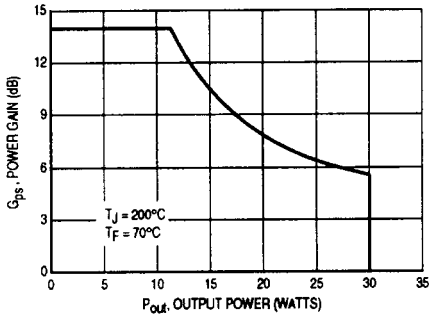


Figure 10. DC Safe Operating Area

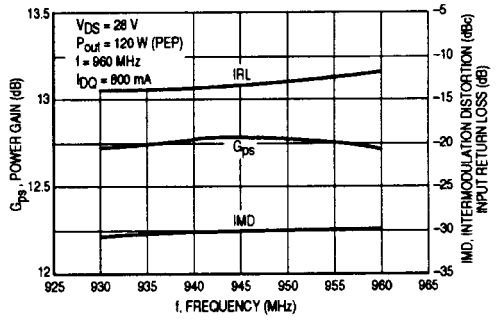
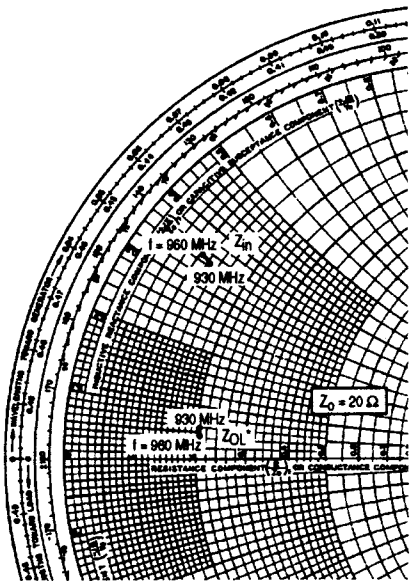


Figure 11. Broadband Circuit Performance



V_{CC} = 28 V, I_{DQ} = 2 × 400 mA, P_{out} = 120 Watts (PEP)

f MHz	Z_{in} Ω	Z_{OL}^* Ω
930	2.5 + j6.9	4.3 + j1.2
945	2.5 + j7.0	4.3 + j1.0
960	2.2 + j7.1	4.3 + j0.9

Z_{in} = Conjugate of impedance.

Z_{OL}^* = Conjugate of the optimum load impedance at a given output power, voltage, IMD, bias current, efficiency and frequency.

Note: Z_{OL}^* was chosen based on tradeoffs between gain, output power, drain efficiency and intermodulation performance. Impedances shown represent a single channel (1/2 of MRF186) impedance measurement.

Figure 12. Series Equivalent Input and Output Impedance

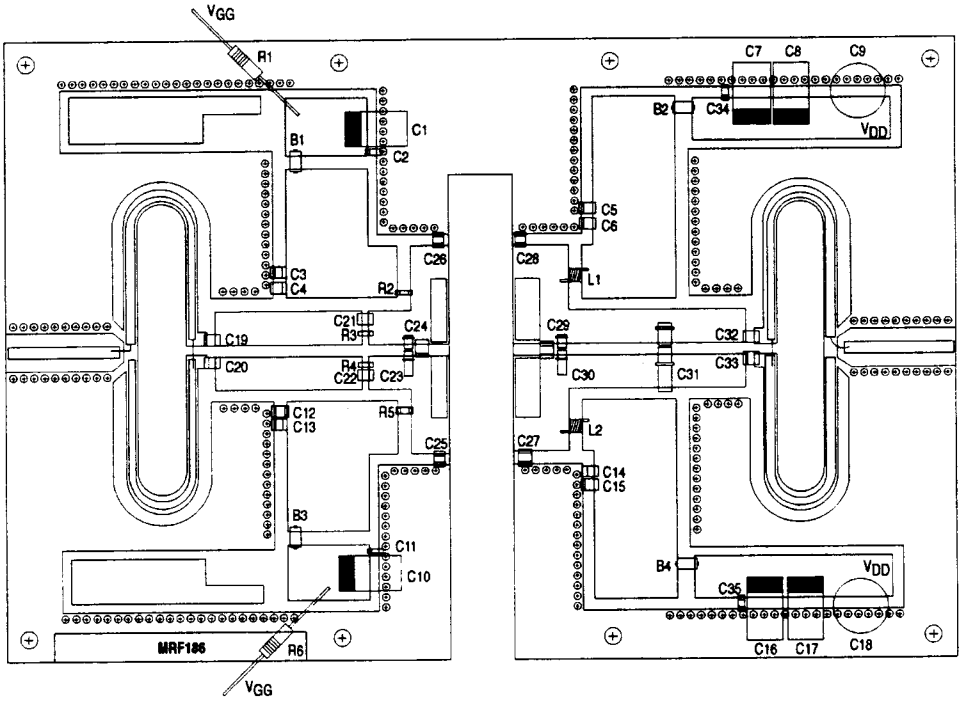
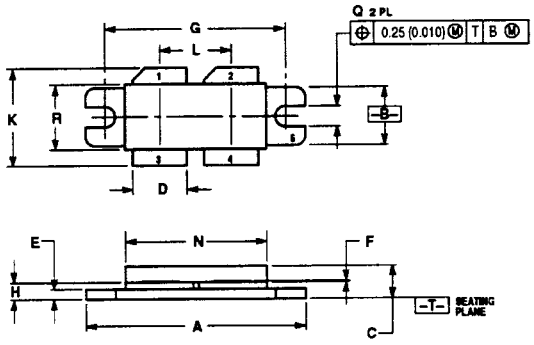


Figure 13. Component Placement Diagram of 930 – 960 MHz Broadband Test Fixture

PACKAGE DIMENSIONS



NOTES:
1. DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2. CONTROLLING DIMENSION: INCH.

DIM	INCHES MIN	INCHES MAX	MILLIMETERS MIN	MILLIMETERS MAX
A	1.330	1.350	33.78	34.29
B	0.375	0.395	9.52	10.03
C	0.180	0.205	4.57	5.21
D	0.320	0.340	8.13	8.64
E	0.060	0.070	1.52	1.77
F	0.004	0.006	0.11	0.15
G	1.100 BSC		27.94 BSC	
H	0.082	0.097	2.09	2.46
K	0.580	0.620	14.73	15.75
L	0.435 BSC		11.05 BSC	
N	0.845	0.875	21.46	22.23
Q	0.118	0.130	3.00	3.30
R	0.390	0.410	9.91	10.41

STYLE 1:
PIN 1. COLLECTOR
2. COLLECTOR
3. BASE
4. BASE
5. EMITTER

CASE 375A–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Motorola data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and Ⓜ are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

Mfax is a trademark of Motorola, Inc.

How to reach us:
USA/EUROPE/Locations Not Listed: Motorola Literature Distribution; P.O. Box 5405, Denver, Colorado 80217. 303–675–2140 or 1–800–441–2447

Mfax™: RMFAX0@email.sps.mot.com – TOUCHTONE 602–244–6609 – US & Canada ONLY 1–800–774–1848

INTERNET: http://motorola.com/sps

JAPAN: Nippon Motorola Ltd.: SPD, Strategic Planning Office, 4–32–1, Nishi–Gotanda, Shinagawa–ku, Tokyo 141, Japan. 81–3–5487–8488

ASIA/PACIFIC: Motorola Semiconductors H.K. Ltd.; 8B Tai Ping Industrial Park, 51 Ting Kok Road, Tai Po, N.T., Hong Kong. 852–26629298

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF899/D

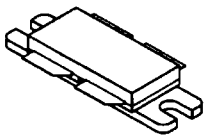
The RF Line
NPN Silicon RF Power Transistor

MRF899

150 W, 900 MHz
RF POWER
TRANSISTOR
NPN SILICON

Designed for 26 Volt UHF large–signal, common emitter, Class AB linear amplifier applications in industrial and commercial FM/AM equipment operating in the range 800–960 MHz.

- Specified 26 Volt, 900 MHz Characteristics
  Output Power = 150 Watts (PEP)
  Minimum Gain = 8.0 dB @ 900 MHz, Class AB
  Minimum Efficiency = 35% @ 900 MHz, 150 Watts (PEP)
  Maximum Intermodulation Distortion –28 dBc @ 150 Watts (PEP)
- Characterized with Series Equivalent Large–Signal Parameters from 800 to 960 MHz
- Silicon Nitride Passivated
- 100% Tested for Load Mismatch Stress at all Phase Angles with 5:1 VSWR @ 26 Vdc, and Rated Output Power
- Gold Metallized, Emitter Ballasted for Long Life and Resistance to Metal Migration
- Circuit board photomaster available upon request by contacting RF Tactical Marketing in Phoenix, AZ.



CASE 375A–01, STYLE 1

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Collector–Emitter Voltage	V_{CEO}	28	Vdc
Collector–Emitter Voltage	V_{CES}	60	Vdc
Emitter–Base Voltage	V_{EBO}	4.0	Vdc
Collector–Current — Continuous	I_C	25	Adc
Total Device Dissipation @ T_C = 25°C Derate above 25°C	P_D	230 1.33	Watts W/°C
Storage Temperature Range	T_{stg}	–65 to +150	°C

THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	0.75	°C/W

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS					
Collector–Emitter Breakdown Voltage (I_C = 100 mAdc, I_B = 0)	$V_{(BR)CEO}$	28	37	—	Vdc
Collector–Emitter Breakdown Voltage (I_C = 50 mAdc, V_{BE} = 0)	$V_{(BR)CES}$	60	85	—	Vdc
Emitter–Base Breakdown Voltage (I_E = 10 mAdc, I_C = 0)	$V_{(BR)EBO}$	4.0	4.9	—	Vdc
Collector Cutoff Current (V_{CE} = 30 Vdc, V_{BE} = 0)	I_{CES}	—	—	10	mAdc
ON CHARACTERISTICS					
DC Current Gain (I_{CE} = 1.0 Adc, V_{CE} = 5.0 Vdc)	h_{FE}	30	75	120	—
DYNAMIC CHARACTERISTICS					
Output Capacitance (V_{CB} = 26 Vdc, I_E = 0, f = 1.0 MHz) (1)	C_{ob}	—	75	—	pF

(1) For information only. This part is collector matched.

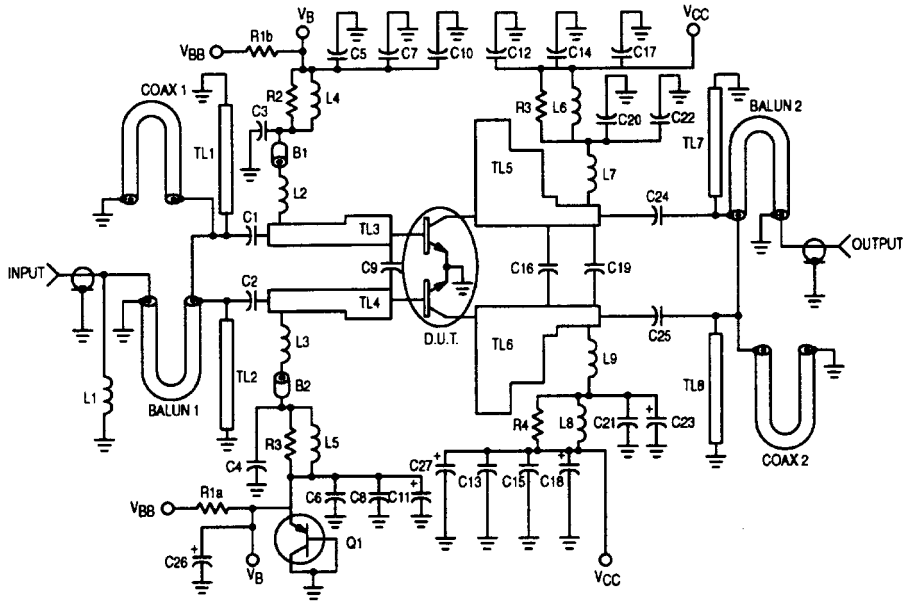
(continued)

© Motorola, Inc. 1997

MOTOROLA

ELECTRICAL CHARACTERISTICS — continued (T_C = 25°C unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
FUNCTIONAL CHARACTERISTICS					
Common–Emitter Amplifier Power Gain V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	G_{pe}	8.0	9.0	—	dB
Collector Efficiency V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	η	30	40	—	%
3rd Order Intermodulation Distortion V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	IMD	—	–32	–28	dBc
Output Mismatch Stress V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz, VSWR = 5:1 (all phase angles)	Ψ	No Degradation in Output Power Before and After Test			



B1, B2 — Ferrite Bead, Ferroxcube #56–590–65–3B
C1, C2, C24, C25 — 43 pF, B Case, ATC Chip Capacitor
C3, C4, C20, C21 — 100 pF, B Case, ATC Chip Capacitor
C5, C6, C12, C13 — 1000 pF, B Case, ATC Chip Capacitor
C7, C8, C14, C15 — 1800 pF, AVX Chip Capacitor
C9 — 9.1 pF, A Case, ATC Chip Capacitor
C10, C11, C17, C18, C22, C23 — 10 µF, Electrolytic Capacitor Panasonic
C16 — 3.9 pF, B Case, ATC Chip Capacitor
C19 — 0.8 pF, B Case, ATC Chip Capacitor
C26 — 200 µF, Electrolytic Capacitor Mallory Sprague
C27 — 500 µF Electrolytic Capacitor

L1 — 5 Turns 24 AWG IDIA 0.059″ Choke, 19.8 nH
L2, L3, L7, L9 — 4 Turns 20 AWG IDIA 0.163″ Choke
L4, L5, L6, L8 — 12 Turns 22 AWG IDIA 0.140″ Choke
N1, N2 — Type N Flange Mount, Omni Spectra
Q1 — Bias Transistor BD136 PNP
R2, R3, R4, R5 — 4.0 x 39 Ohm 1/8 W Chips in Parallel
R1a, R1b — 56 Ohm 1.0 W
TL1–TL8 — See Photomaster
Balun1, Balun2, Coax 1, Coax 2 — 2.20″ 50 Ohm 0.086″ o.d. Semi–rigid Coax, Micro Coax
Board — 1/32″ Glass Teflon, ε_r = 2.55″ Arlon (GX–0300–55–22)

Figure 1. 900 MHz Power Gain Test Circuit

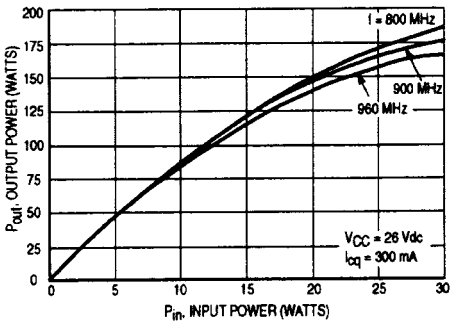


Figure 2. Output Power versus Input Power

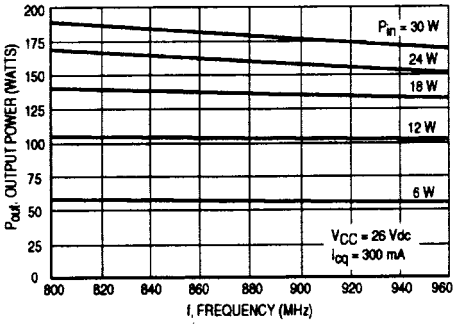


Figure 3. Output Power versus Frequency

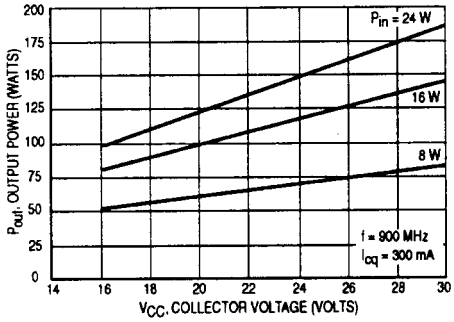


Figure 4. Output Power versus Supply Voltage

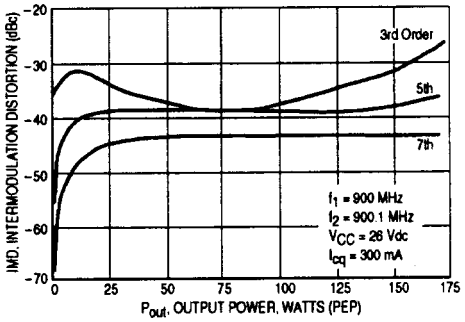


Figure 5. Intermodulation versus Output Power

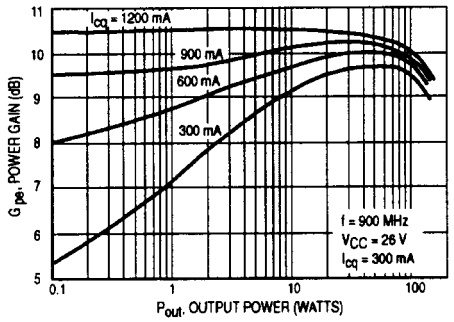


Figure 6. Power Gain versus Output Power

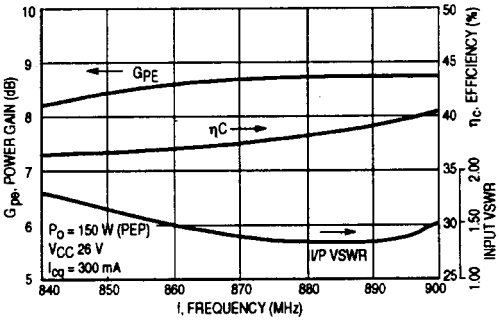
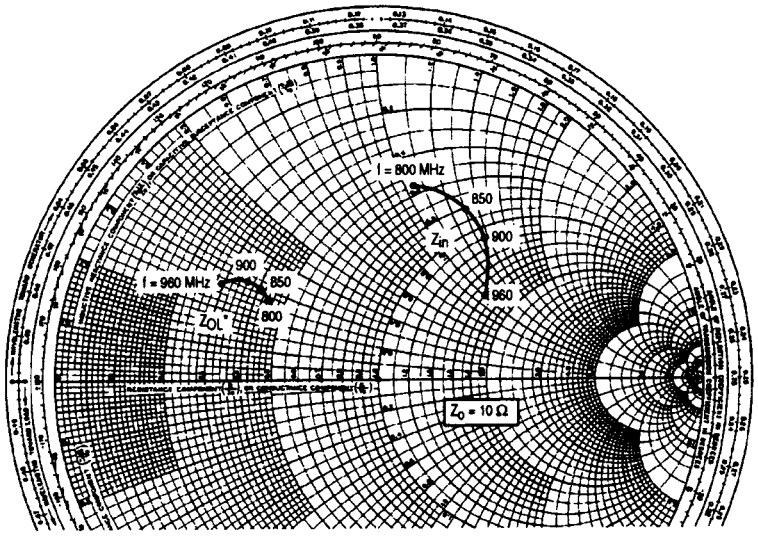


Figure 7. Broadband Test Fixture Performance



f MHz	Z_{in} Ohms	Z_{OL}^* Ohms
800	5.51 + j10.6	4.52 + j2.64
850	8.17 + j13.2	4.21 + j2.98
900	11.2 + j13.8	3.68 + j2.97
960	16.8 + j10.1	2.98 + j2.71

Z_{OL}^* = Conjugate of optimum load impedance into which the device operates at a given output power, voltage and frequency.

NOTE: Z_{in} & Z_{OL}^* are given from base–to–base and collector–to–collector respectively

Figure 8. Input and Output Impedances with Circuit Tuned for Maximum Gain @ P_O = 150 W (PEP), V_{CC} = 26 V

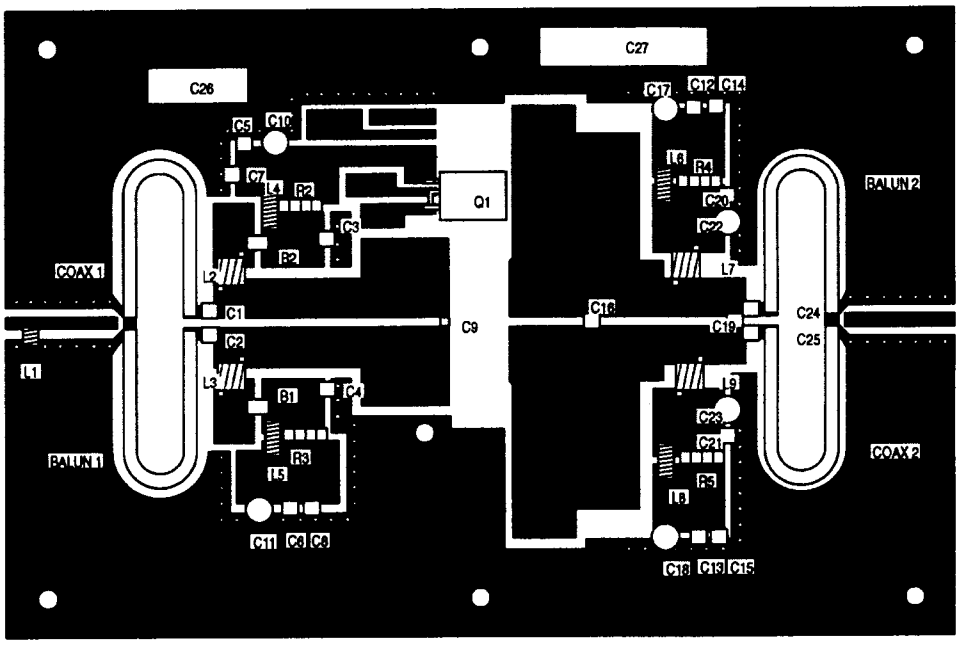
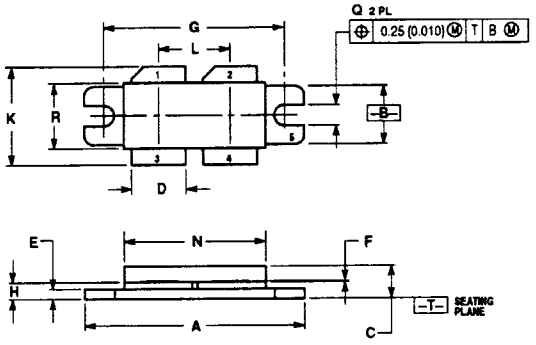


Figure 9. MRF899 Test Fixture Component Layout

PACKAGE DIMENSIONS



NOTES:
1. DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2. CONTROLLING DIMENSION: INCH.

DIM	INCHES MIN	INCHES MAX	MILLIMETERS MIN	MILLIMETERS MAX
A	1.330	1.350	33.79	34.29
B	0.375	0.395	9.52	10.03
C	0.160	0.205	4.57	5.21
D	0.320	0.340	8.13	8.64
E	0.060	0.070	1.52	1.77
F	0.004	0.006	0.11	0.15
G	1.100 BSC		27.94 BSC	
H	0.062	0.097	2.08	2.46
K	0.580	0.620	14.73	15.75
L	0.435 BSC		11.05 BSC	
N	0.845	0.875	21.46	22.23
Q	0.116	0.130	3.00	3.30
R	0.390	0.410	9.91	10.41

STYLE 1:
PIN 1. COLLECTOR
2. COLLECTOR
3. BASE
4. BASE
5. EMITTER

CASE 375A–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Motorola data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and M are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

Mfax is a trademark of Motorola, Inc.

How to reach us:
USA/EUROPE/Locations Not Listed: Motorola Literature Distribution;
P.O. Box 5405, Denver, Colorado 80217. 303–675–2140 or 1–800–441–2447

Mfax™: RMFAX0@email.sps.mot.com – TOUCHTONE 602–244–6609
– US & Canada ONLY 1–800–774–1848

INTERNET: http://motorola.com/sps

JAPAN: Nippon Motorola Ltd.: SPD, Strategic Planning Office, 4–32–1, Nishi–Gotanda, Shinagawa–ku, Tokyo 141, Japan. 81–3–5487–8488

ASIA/PACIFIC: Motorola Semiconductors H.K. Ltd.; 8B Tai Ping Industrial Park, 51 Ting Kok Road, Tai Po, N.T., Hong Kong. 852–26629298

REFERENCES

1. Ulrich L. Rohde, Anthony M. Pavio, and Robert A. Pucel, "Accurate Noise Simulation of Microwave Amplifiers Using CAD," *Microwave J.*, **31**(12), 130–141, December 1988.
2. H. T. Friis, "Noise Figures of Radio Receivers," *Proc. IRE*, **32**, 419–422, July 1944.
3. "IRE Standards on Methods of Measuring Noise in Linear Twoports," *Proc. IRE*, **48**, 60–68, January 1960.
4. H. Rothe and W. Dhalke, "Theory of Noisy Fourpoles," *Proc. IRE*, **44**, 811–818, June 1956.
5. "Representation of Noise in Linear Twoports," *Proc. IRE*, **48**, 69–74, January 1960.
6. H. A. Haus and R. B. Adler, *Circuit Theory of Linear Noisy Networks*, Cambridge, MA: MIT Press, 1959/New York: Wiley, 1959.
7. H. Fukui, "Noise Performance of Microwave Transistors," *IEEE Trans. Electron Devices*, **ED-13**, 329–341, March 1966.
8. S. Iversen, "The Effect of Feedback on Noise Figure," *Proc. IEEE*, **63**, 540–542, March 1975.
9. L. Besser, "Stability Considerations of Low-Noise Transistor Amplifiers with Simultaneous Noise and Power Match," *IEEE MTT-S International Microwave Symposium Digest*, 1975, pp. 327–329.
10. G. D. Vendelin, "Feedback Effects on the Noise Performance of GaAs Mesfets," *IEEE MTT-S International Microwave Symposium Digest*, 1975, pp. 324–326.
11. W. A. Suter, "Feedback and Parasitic Effects on Noise," *Microwave J.* **26**(2), 123–129, February 1983.
12. K. Hartmann and M. J. O. Strutt, "Changes in Four Noise Parameters Due to General Changes in Linear Two-Port Circuits," *IEEE Trans. Electron Devices*, **ED-20**(10), 874–877, October 1973.
13. H. Fukui, "Available Power Gain, Noise Figure, and Noise Measure of Two-Ports and Their Graphical Representation," *IEEE Trans. Circuit Theory*, **CT-13**, 137–142, June 1966.
14. C. R. Poole and D. K. Paul, "Optimum Noise Measure Terminations for Microwave Transistor Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**(11), 1254–1257, November 1985.
15. H. Hillbrand and P. H. Russer, "An Efficient Method for Computer Aided Noise Analysis of Linear Amplifier Networks," *IEEE Trans. Circuits Syst.*, **CAS-23**(4), 235–238, April 1976.
16. W. E. Pastori, *Topics on Noise*, Eaton Electronics Instrumentation Division (EID) Seminar Notes, Los Angeles, 1983.
17. W. E. Pastori, "Image and Second Stage Corrections Resolve Noise Figure Measurement Confusion," *Microwave Syst. News*, **13**(5), 67–86, May 1983.
18. R. Q. Lane, "The Determination of Device Noise Parameters," *Proc. IEEE*, **57**, 1461–1462, August 1969.
19. R. Q. Lane, "A Microwave Noise and Gain Parameter Test Set," *IEEE International Solid-State Circuits Conference*, 1978, pp. 172–173.
20. N. Kuhn, "Accurate and Automatic Noise Figure Measurements with Standard Equipment," *IEEE MTT-S International Microwave Symposium Digest*, 1980, pp. 425–427.
21. E. W. Strid, "Measurement of Losses in Noise-Matching Networks," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**, 247–252, March 1981.
22. Chris Bowick, *RF Circuit Design*. Indianapolis: Howard W. Sams & Co., 1982.
23. Adel S. Sedra and Kenneth C. Smith, *Microelectronic Circuits*. Chicago: Saunders College Publishing, 1991.
24. Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed. New York: Wiley, 1984.

25. S. H. Lee, A. K. Wong, and M. G. Wong, *A Current-Combiner Circuit for Better Mixer Conversion Gain*, Philips Semiconductors applications document.
26. F. M. Fano, "Theoretical Limitations in the Broadband Matching Arbitrary Impedances," *J. Franklin Inst.*, **249**, 57–85, 139–155, 1950.
27. Ravender Goyal, ed., *High-Frequency Analog Integrated-Circuit Design*. New York: Wiley, 1995, Chap. 6.
28. Stefan Georgi and Ulrich L. Rohde, "A Novel Approach to Voltage Controlled Tuned Filters Using CAD Validation," *Microwave Eng. Eur.*, 35–42, October 1997.
29. *Analog/Interface ICs Device Data*, Vol. II, Q4/96, DL128, Rev 6, Motorola, Inc., MC13158 Wideband FM IF Subsystem specifications, pp. 8-308 to 8-329.
30. E. J. Wilkinson, "An *N*-Way Power Divider," *IRE Trans. Microwave Theory Tech.*, **MTT-8**, 116, January 1960.
31. Motorola Application Note AN-282A, *Systemizing RF Power Amplifier Design*.
32. W. E. Everitt and G. E. Anner, *Communication Engineering*. New York: McGraw-Hill, 1956.
33. G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks and Coupling Structures*. New York: McGraw-Hill, 1980.
34. G. L. Matthaei, "Tables of Chebyshev Impedance-Transforming Networks of Low-Pass Filter Form," *Proc. IEEE*, August 1964.
35. Motorola Application Note AN-267, *Matching Network Designs with Computer Solutions*.
36. E. G. Cristal, "Tables of Maximally Flat Impedance-Transforming Networks of Low-Pass Filter Form," *IEEE Trans. Microwave Theory Tech.*, **MTT 3**(5), September 1965.
37. H. W. Bode, *Network Analysis and Feedback Amplifier Design*. New York: Van Nostrand, 1943.
38. R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," *J. Franklin Inst.*, 139–154, January-February 1950.
39. J. H. Horwitz, "Design Wideband UHF-Power Amplifiers," *Electron. Des.* **11**, x–x, May 24, 1969.
40. O. Pitzalis and R. A. Gilson, "Tables of Impedance Matching Networks Which Approximate Prescribed Attenuation Versus Frequency Slopes," *IEEE Trans. Microwave Theory Tech.*, **MTT-9**(4), April 1971.
41. Anthony N. Gerkis, "Broadband Impedance Matching Using the 'Real Frequency' Network Synthesis Technique," *Appl. Microwave & Wireless*, **10**(6), 26–36, July/August 1998.
42. T. Cuthbert, "Broadband Impedance Matching Methods," *RF Des.*, **17**(8), 64–71, August 1994.
43. T. Cuthbert, *Broadband Direct-Coupled and Matching RF Networks*. Greenwood, AR, T. R. Cuthbert, 1999.
44. George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990, pp. 279–283.
45. Andy Potter, "HP RF Compiler Automates Schematic Capture and Extends Capabilities of Circuit Synthesis." *Appl. Microwave & Wireless*, **11**(6), 106–118, June 1999.
46. Aron Kain, *Final Report for an Integrated CAD Tool for High-Power, High-Efficiency RF and Microwave Amplifiers*, STTR Phase II, Contract no. F33615-95-C-174, Dept. of the Air Force, Small Business Technology Transfer (STTR) Program.
47. Kenneth K. Clarke and Donald T. Hess, *Communication Circuits: Analysis and Design.* Reading, MA: Addison-Wesley, 1971.

FURTHER READING

P. L. D. Abrie, *The Design of Impedance-Matching Networks for Radio-Frequency and Microwave Amplifiers*. Norwood, MA: Artech House, 1985.

P. L. D. Abrie and P. Rademeyer, "A Method for Evaluating and the Evaluation of the Influence of the Reverse Transfer Gain on the Transducer Power Gain of Some Microwave Transistors," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**(8), 711–713, August 1985.

A. Ambrózy, *Electronic Noise*. New York: McGraw-Hill, 1982.

W. Baechtold and M. J. O. Strutt, "Noise in Microwave Transistors," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**(9), 578–585, September 1968.

A. Cappy, "Noise Modeling and Measurement Techniques," *IEEE Trans. Microwave Theory Tech.*, **MTT-36**, 1–10, January 1988.

R. S. Carson, *High Frequency Amplifiers*. New York: Wiley, 1975.

Cezar A. A. Carioca and Paulo H. P. de Carvalho, "Computer-Aided Design of Diode Frequency Multipliers," *Appl. Microwave & Wireless*, **11**(4), 38–46, April 1999.

H. J. Carlin, "A New Approach to Gain-Bandwidth Problems," *IEEE Trans. Circuits Syst.*, **CAS-23**, 170–175, April 1977.

H. J. Carlin and P. Amstutz, "On Optimum Broad-Band Matching," *IEEE Trans. Circuits Syst.*, **CAS-23**, 401–405, April 1977.

H. J. Carlin and P. Amstutz, "On Optimum Broadband Matching," *IEEE Trans. Circuits Syst.*, **CAS-28**(5), May 1981.

H. J. Carlin and J. J. Komiak, "A New Method of Broad-Band Equalization Applied to Microwave Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**, 93–99, February 1979.

H. J. Carlin and B. Yarman, "A Simplified Real Frequency Technique Applied to Broad-Band Multistage Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-30**(12), December 1982.

H. J. Carlin and B. S. Yarman, "The Double Matching Problem: Analytic and Real-Frequency Solutions," *IEEE Trans. Circuits Syst.*, **CAS-30**(1), 15–28, January 1983.

G. Caruso and M. Sannino, "Computer-Aided Determination of Microwave Two-Port Noise Parameters," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 639–642, September 1978.

Complex Math User's Guide, Compact Software, Paterson, NJ, 1988.

Steve C. Cripps, *RF Power Amplifiers for Wireless Communications*. Norwood, MA: Artech House, 1999.

T. Cuthbert, *Circuit Design Using Personal Computers*. New York: Wiley, 1983.

J. A. Eisenberg, "Designing Amplifiers for Optimum Noise Figure," *Microwaves*, **13**, 36–41, April 1974.

R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," *J. Franklin Inst.*, **249**, 57–83, January 1960/139–155, February 1960.

H. Fukui, *Low-Noise Microwave Transistors and Amplifiers*. New York: IEEE Press, 1981.

C. Gentili, *Microwave Amplifiers and Oscillators*. New York: McGraw-Hill, 1987.

Stefan Georgi and Ulrich L. Rohde, "A Novel Approach to Voltage-Controlled Tuned Filters Using CAD Validation," *Microwave Eng. Eur.*, 35–42, October 1997.

G. Gonzales, *Microwave Transistor Amplifier Analysis and Design*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

H. A. Haus and R. B. Adler, "Optimum Noise Performance of Linear Amplifiers," *Proc. IRE*, **46**, 1517–1533, August 1958.

Aron Kain, *Final Report for Bias Dependence Noise Modeling of Heterojunction Bipolar Transistors*, USAF SBIR Phase II (PIIN), F33615-95-C-1707, November 1997. Issued by USAF/AFMC/ASC, Wright Laboratory WL/AAKE BLD 7, 2530 C ST, Wright-Patterson AFB, OH 45433-7607.

M. J. Köppen, "A Single-Stage Power Amplifier for 17 watts Output in the 470-MHz Band Using the 266BLY, Operating from 12.5 volt Supply Voltage," C.A.B. Communications, Eindhoven, Holland, Electronic Application Laboratory Report, August 8, 1970.

William B. Kuhn and Naveen K. Yanduru, "Spiral Inductor Substrate Loss Modeling in SiliconRFICs," *Microwave J.*, **11**(3), 66–81, March 1999.

J. Lange, "Noise Characterization of Linear Two-Ports of Invariant Parameters," *IEEE J. Solid-State Circuits*, **SC-2**, 37–40, June 1967.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

David Lovelace, Natalino Camilleri, and George Kannell, "Silicon MMIC Inductor Modeling for High Volume, Low Cost Applications," *Microwave J.*, **39**(8), 60–71, August 1994.

Jarek Lucek and Robbin Damen, "Designing an LNA for a CDMA Front End," *RF Des.*, **22**(2), 20–30, February 1999.

Stephen A. Maas, "FET Models for Volterra-Series Analysis," *Microwave J.*, **42**(5), 260–270, May 1999.

Donald A. McClure, KB2Z, "Transmission-Line Transformers," *Commun. Q.*, 45–58, Summer 1997.

D. J. Mellor and J. G. Linvill, "Synthesis of Interstage Networks and Prescribed Gain Versus Frequency Slopes," *IEEE Trans. Microwave Theory Tech.*, **MTT-23**(12), 1013–1020, December 1975.

Sean Mercer, "Exact Simulation of LNAs Reduces Design Cycle Time," *Appl. Microwave & Wireless*, 74–82, January 1999.

R. P. Meys, "A Wave Approach to the Noise Properties of Linear Microwave Devices," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 34–37, January 1978.

C. K. S. Miller, W. C. Daywitt, and M. G. Arthur, "Noise Standards, Measurements and ReceiverNoise Definitions," *Proc. IEEE*, **55**, June 1967.

Brian J. Minnis, *Designing Microwave Circuits by Exact Synthesis* . Boston: Artech House, 1996.

R. Minton, *Design of Large Signal VHF Transistor Power Amplifiers* , RCA Application Note SMA-36, July 1964.

R. Minton and H. C. Lee, *Frequency Multiplication Using Overlay Transistors* , RCA Application Note SMA-40, September 1965.

M. Mitama and H. Katoh, "An Improved Computational Method for Noise Parameter Measurement," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**, 612–615, June 1979.

Motorola tutorial, "The Cascode Amplifier: A Useful RF Building Block," *Microwave Prod. Digest*, 39, 59, March 1999.

T. Mukaihata, "Applications and Analysis of Noise Generation in *N*-Cascaded Mismatched Two-Port Networks," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 699–708, September 1968.

W. W. Mumford and E. H. Scheibe, *Noise Performance Factors in Communications Systems* . Dedham, MA: Horizon House, 1968.

D. E. Norton, "High Dynamic Range Transistor Amplifier Using Lossless Feedback," *Microwave J.*, **19**(5), 53–57, May 1976.

D. E. Norton, "Strong Signal Survivability Design for Lossless Feedback Amplifiers," *Proceedings of the Eleventh Annual European Microwave Conference* , 1981.

H. S. Okean and P. O. Lombardo, "Noise Performance of M/W and MM-Wave Receivers," *Microwave J.*, **16**(1), January 1973.

T. Y. Otoshi, "The Effect of Mismatched Components on Microwave Noise-Temperature Calibrations," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 675–686, September 1968.

P. Penfeld, "Wave Representations of Amplifier Noise," *IRE Trans. Circuit Theory*, 84–86, March 1962.

Ulrich L. Rohde, Jerry C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*, 2nd ed. New York: McGraw-Hill, 1997, Chap. 3.

C. Rorabaugh, *Circuit Design and Analysis—Featuring C Routines* . New York: McGraw-Hill, 1992.

M. Sannino, "On the Determination of Device Noise and Gain Parameters," *Proc. IEEE*, **67**, 1364–1366, September 1979.

SYNMAT, matching synthesis program by Radom, Inc., developed by Douglas Mellor.

A. van der Ziel, *Noise in Solid State Devices and Circuits* . New York: Wiley, 1986.

Dave Vondran, "Noise Figure Measurement: Corrections Related to Match and Gain," *Microwave J.*, **42**(3), 22–38, March 1999.

D. F. Wait, "Thermal Noise from a Passive Linear Multiport," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 687–691, September 1968.

B. S. Yarman and H. J. Carlin, "A Simplified 'Real Frequency' Technique Applied to Broad-Band Multistage Microwave Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-30**(12), 2216–2222, December 1982.

G. P. Young and S. O. Scanlan, "Matching Network Design Studies for Microwave Transistor Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**(10), 1027–1035, October 1981.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

3

AMPLIFIER DESIGN WITH BJTs AND FETs

3-1 PROPERTIES OF AMPLIFIERS

3-1-1 Introduction

The goal of this chapter is to move from the details on the semiconductor devices themselves to our first practical application. One of the more interesting features of amplifiers is that they can easily become what will be covered in Chapter 5: *oscillators*. This is due to the high gain of the devices, unaccounted for parasitic elements, and other design flaws. The amplifiers we need will fall into three categories:

- Low-noise amplifiers
- High-gain amplifiers
- Medium- to high-power amplifiers

See Figure 3-1. The low-noise amplifier always operates in Class A, typically at 15–20% of its maximum useful current. The high-gain amplifier can operate in Class A, as well as Class B (mostly push–pull). The higher dc current for the same device results in a higher noise figure and more gain, and ultimately more output power. Class C operating mode is really reserved for either FM transmissions or constant-carrier modes like CW. Some of the modern digital modulation types are sensitive to phase distortion rather than amplitude changes and because of the resulting output spectrum designers have stayed away from Class C operation.

Some of these amplifiers, such as bipolar versions, can be dc-coupled with very few difficulties; others, like those in the FET families, will cause more headaches. An interesting example, although much too high in frequency for the purpose of this book, is shown in Figures 3-2, 3-3, 3-4, and 3-5. It shows a three-stage amplifier that has noise feedback for the first two stages and resistive feedback for the output stage. These are the three types of amplifiers, regardless of technology, we will be evaluating. The achievement of accurate noise analysis of this circuit, reported in Rohde et al. [1], marked the first time a complete

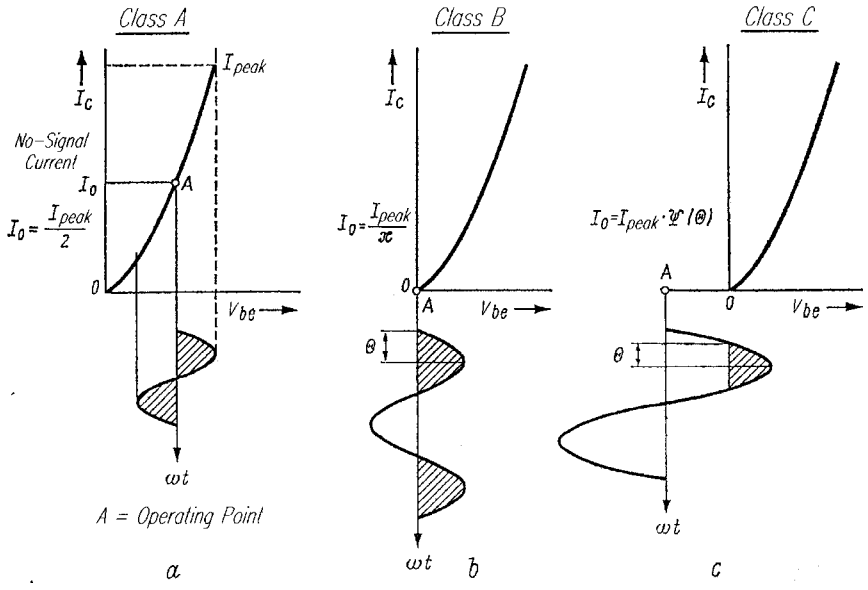


Figure 3-1 Definition of Class A, B, and C operation and resulting bias, including conduction angle (Θ, to be discussed in Section 3-2-2). The transfer characteristic can be either quadratic or exponential. This results in different distortion but does not change the basic operating mode.

linear noise model with essentially no frequency limitation had been developed. Its accuracy depends solely on the accuracy of the element values of the linear equivalent circuit, which can be obtained from measured *S*-parameters by an optimization process. Needless to say, this technique is also applicable at lower frequencies, such as 1.5–3 GHz, and its deviation from measured values has never been worse than 0.2 dB or, relative to the noise figure, 10% expressed in dB.

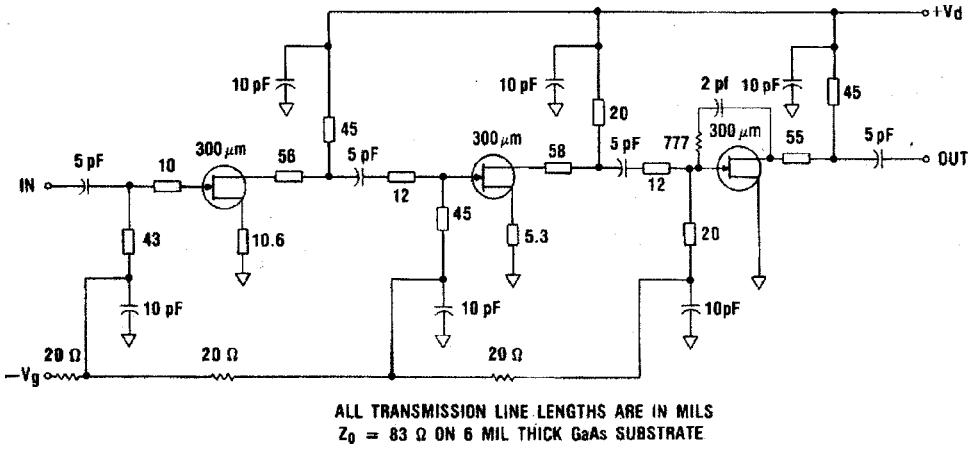


Figure 3-2 Schematic of the X-band GaAs monolithic low-noise amplifier (Texas Instruments EG8021). For reasons of linearity, all three stages operate in Class A.

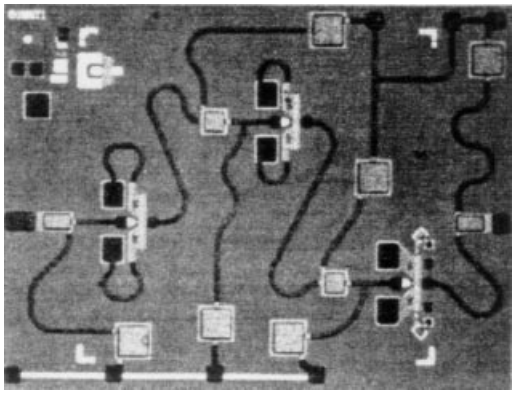


Figure 3-3 Photograph of the EG8021 monolithic amplifier chip. The area pictured is 0.09 inch by 0.12 inch in size.

The results of Figure 3-4 were obtained by using the linear FET model as supplied by Texas Instruments. It becomes somewhat obvious that a match for S_{11} of –60 dB is not likely. We then replaced the linear transistor with the SPICE-type nonlinear model, resulting in the responses shown in Figure 3-5. S_{11} now looks much more realistic in the center of the range, but at the same time exhibits a tendency toward instability at frequencies around 8 GHz. The actual circuit, when measured, did not exhibit this potential instability; on the other hand,

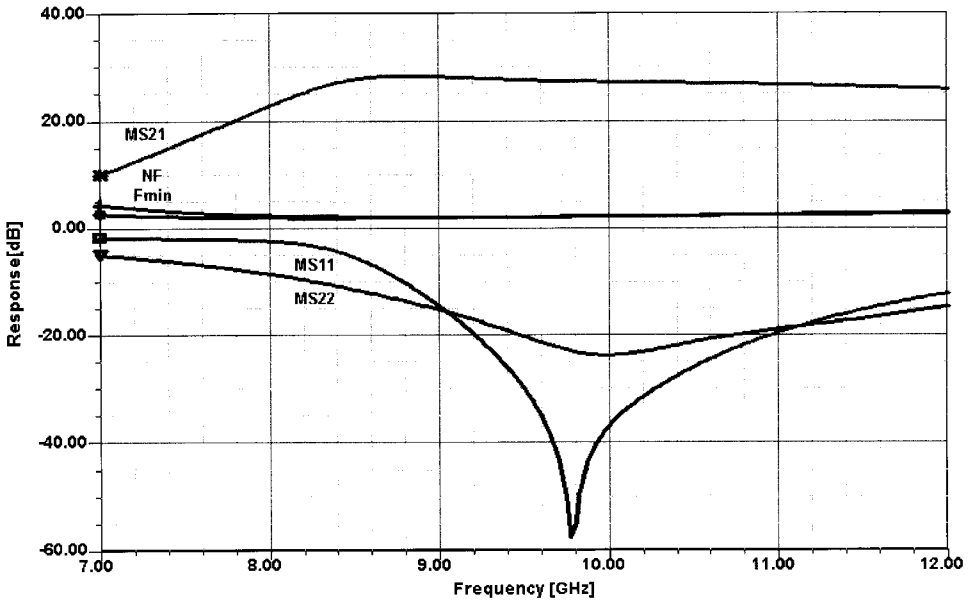


Figure 3-4 Simulated F_{min}, noise figure, S_{11}, S_{21}, and S_{22} responses for the three-stage GaAsFET amplifier using the TI linear FET model. The values at 10 GHz are: F_{min}, 2.20 dB; NF, 2.21 dB; S_{11}, –36.8 dB; S_{21}, 27.2 dB; and S_{22}, –23.7 dB.

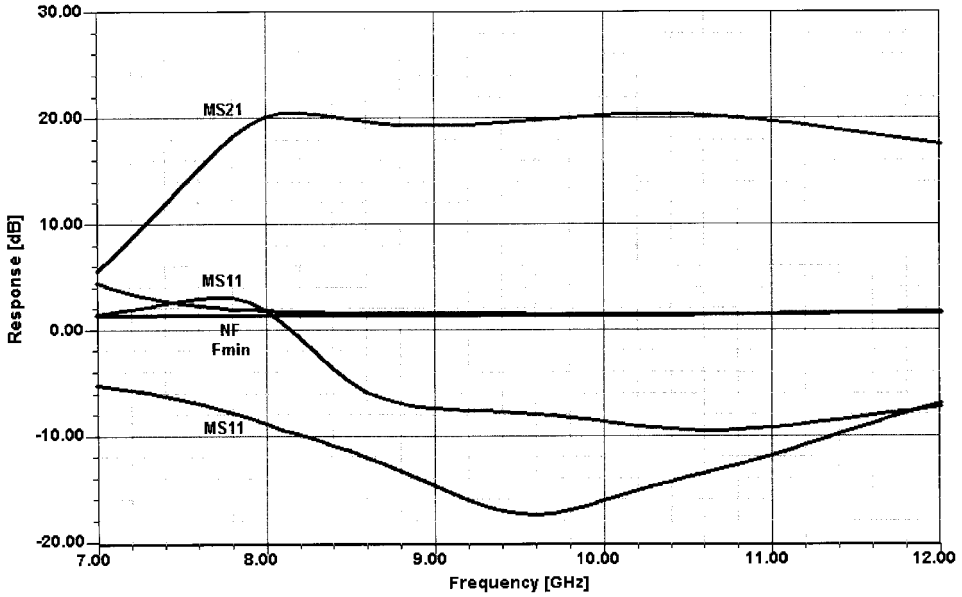


Figure 3-5 Simulated F_{min}, noise figure, S_{11}, S_{21}, and S_{22} responses for the three-stage GaAsFET amplifier using the nonlinear FET model. The values at 10 GHz are: F_{min}, 1.53 dB; NF, 1.65 dB; S_{11}, –8.5 dB; S_{21}, 20.3 dB; and S_{22}, –15.9 dB.

the noise figure and gain agreed quite well. The discrepancies between the results with the linear and nonlinear models emphasize the importance of accurate device modeling.

Figure 3-6 shows a wideband amplifier, specifically a distributed amplifier, that covers 1–20 GHz. Figure 3-7 shows its simulated frequency-dependent gain, matching, and noise performance. The example given here as the upper frequency for wireless systems is rapidly expanding and the technology to accommodate this is already in place. Modern CAD tools can handle such high frequencies accurately, which means one can trust them at the significantly lower wireless frequencies of 500 MHz to 3 GHz. Above 1500 MHz, one really should model using distributed elements. We will come to this later in this chapter.

Input/output and interstage matching will become another challenging task, specifically if one requires sufficient bandwidth. As circuits become more complex and as the operating frequency increases, passive components play an increasingly important role; the use of

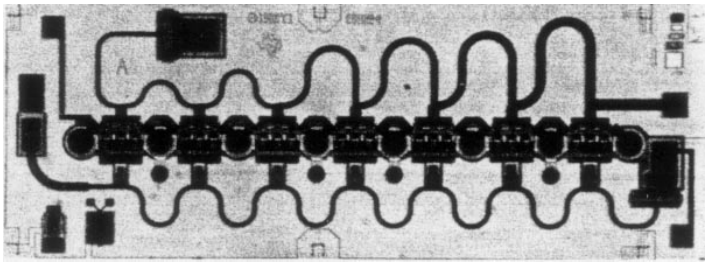


Figure 3-6 Photo of the 1–20-GHz distributed amplifier.

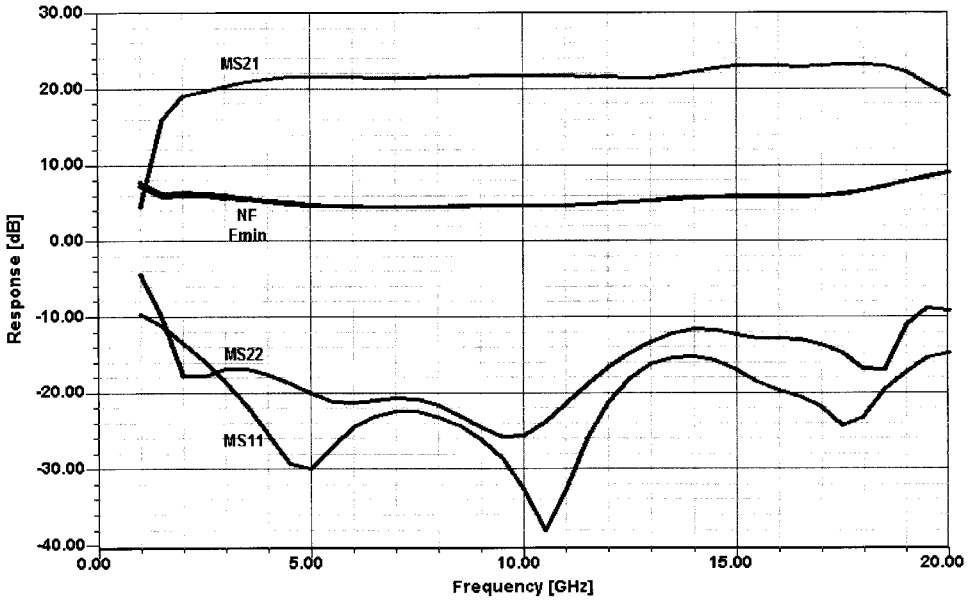


Figure 3-7 Simulated F_{min}, NF, S_{11}, S_{21}, and S_{22} responses. At the center of the range, the following numbers were found: F_{min}, 4.80 dB; NF, 4.80 dB; S_{11}, –32.8 dB; S_{21}, 21.8 dB; and S_{11}, –25.2 dB.

distributed elements is practically a necessity above 1500 MHz. For this reason, we begin this chapter by covering the application of transistors as amplifiers, following this with a discussion of the surrounding external circuitry, such as transmission lines, inductors, capacitors, and resistors. Monolithic ICs on silicon are a totally different ball game, and neither of the authors has sufficient expertise in this area. For this reason, we refer readers who will concentrate on MMIC design to the literature mentioned at the end of this chapter.

The goal of the applications presented here is to address, where possible, high-performance wireless stages that are less subject to constraints of space, power consumption, and cost, than portable/mobile applications. This is consistent with our preface, in which we have clearly stated our goal. The designers of handsets typically cannot consider the full variety of circuits we will present here. These types of stages (low noise, high gain and power) frequently will be used with filtering, and we will give some examples of input filters, including a tracking high-performance design.

A key topic besides the frequently mentioned gain is the whole noise issue. We will give the reader a complete introduction to this important area, followed by a similar introduction to the gain and matching of transistors. *Caution*: The majority of applications published go through the exercise of using *S*-parameters manually to design low-noise and small-signal amplifiers. While we have been tempted to repeat the information found in several textbooks, we feel that today's engineers will have access to CAD tools one way or another. It is highly unlikely that readers will undertake the inefficient process of going through all the design steps manually (a procedure that can take days); hopefully, engineers will resort to a linear/nonlinear CAD tool to do so. Only for the purpose of summarizing the necessary steps will we provide the relevant information regarding gain, stability, and related topics, demonstrating their usefulness or possible misuse with a modern CAD tool—in our case,

the Serenade Design Environment product from Ansoft. At this time, several additional companies beside Ansoft and Hewlett-Packard have emerged, but it is dangerous to list all of them as they come and go.

3-1-2 Gain

In practical terms, when we talk about gain we mean the output power relative to the input power. The industry standard now uses 50 Ω; however, at the point at which we have to match high-impedance devices such as monolithic crystal filters where the impedance jumps, things are not so obvious. The transformation equation that allows us to transform from one impedance to another is

$$m = \sqrt{\frac{R_2}{R_1}} \tag{3-1}$$

This correction factor, m, is required to do the transformation from R_1 (50 Ω) to R_2 (e.g., 1.2 kΩ). It needs to be understood that while the power gain remains the same, the voltage gain requires this correction factor. Since the power gain is expressed as 10 log A, A being the loop gain, 20 log A is the voltage gain, both expressed in dB. The reason for 10 versus 20 is that the power gain is always the voltage squared, which accounts for the 2. As a reminder of these relationships we present the following:

Distinction: Power Ratios–Voltage Ratios

$$n \text{ dB} = 10 \log_{10}(P_2/P_1) \quad \xrightarrow{PP = V^2/Z} \quad n = 10 \log_{10}\left(\frac{V_2^2/Z_2}{V_1^2/Z_1}\right)\text{dB}$$

$$Z_2 = Z_1 \rightarrow n \text{ dB} = 10 \log_{10}(V_2^2/V_1^2) = 10 \log_{10}(V_2/V_1)^2$$

$$\Longrightarrow \boldsymbol{n \text{ dB} = 20 \log_{10}(V_2/V_1) \text{ dB}}$$

The gain is typically defined under the condition of real resistive source and load terminations, which means that all the available power is being supplied to the amplifier, and likewise through proper matching at the output all the available power is fed to the termination. This gain figure is typically measured at a single spot frequency, but since most amplifiers are not totally flat, one needs to specify the gain tolerances or gain flatness. A good number is x dB ± 0.5 dB gain variation. In the case of a narrowband amplifier, the gain typically is shaped like a tuned circuit or like a critically coupled bandpass filter frequency response, while in the case of a real wideband application, the desired gain has some kind of ripple, which is due to the various compensation components in the circuit. An extreme case of a wideband amplifier is a 6–18-GHz single-stage amplifier that uses some very clever matching. This is shown in the following example. Figure 3-8 shows the circuit diagram, and Figure 3-9 shows its frequency-dependent performance.

While this is a cute approach for generating a single-stage amplifier over such a wide bandwidth, in real life it is not acceptable. The reason for this is that neither S_{11} nor S_{22} meets any useful specifications over the entire range. Typically, S_{11}/S_{22} values are only a few dB, while in reality more than 10 dB is required. We have chosen this example to show that gain

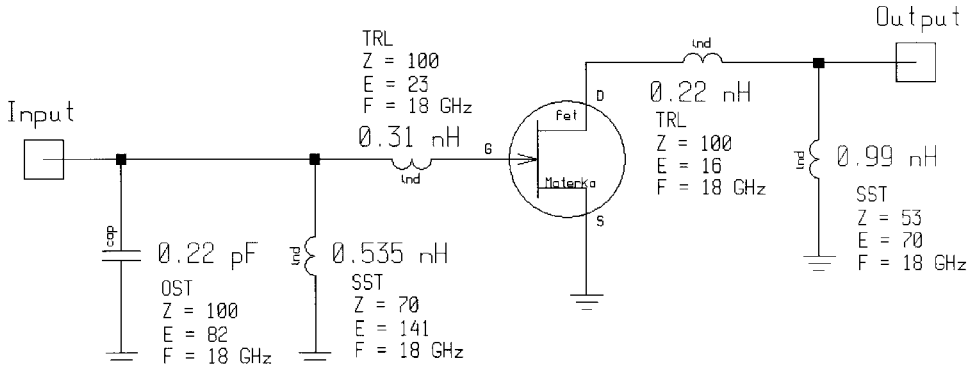


Figure 3-8 A 6–18-GHz FET amplifier.

and matching, depending on the circuit topology, are connected. We get similar results for the noise figure, whereby the best achievable noise figure ($F_{\min}$) and the actual 50-Ω noise figure are only a few dB apart; however, its frequency dependence in a wideband stage is certainly not acceptable.

Another parameter associated with gain is the stability factor K of the circuit. A value of $K > 1$ is mandatory to guarantee stability together with the S_{11} and S_{22} magnitudes being less than 1. Figure 3-10 plots the K stability factor, which for absolute stability must be >1, and maximum stable gain (MSG) as well as $G_{\max}$. It turns out that MSG is a highly artificial

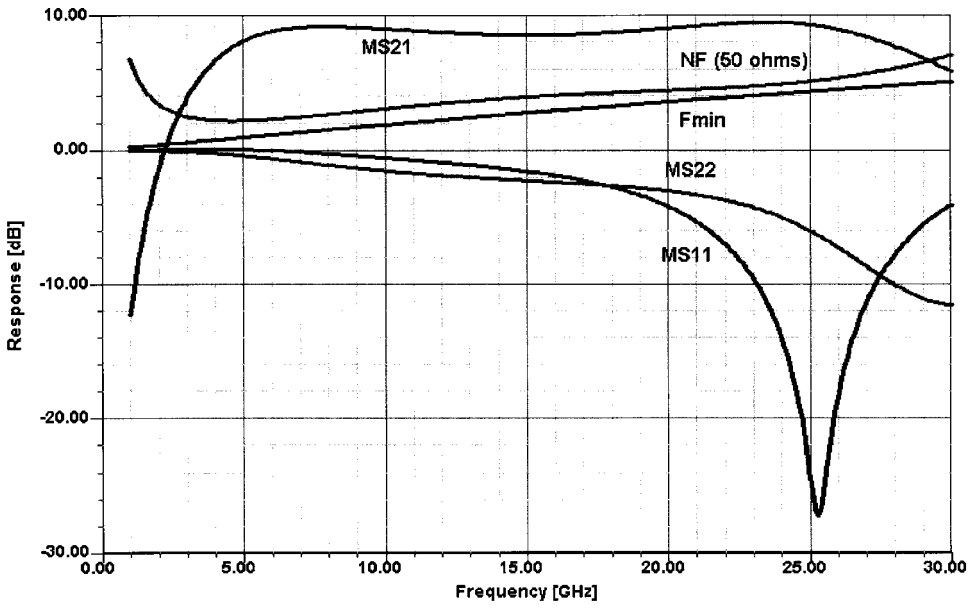


Figure 3-9 Frequency-dependent gain, matching, and noise performance of the amplifier shown in Figure 3-8.

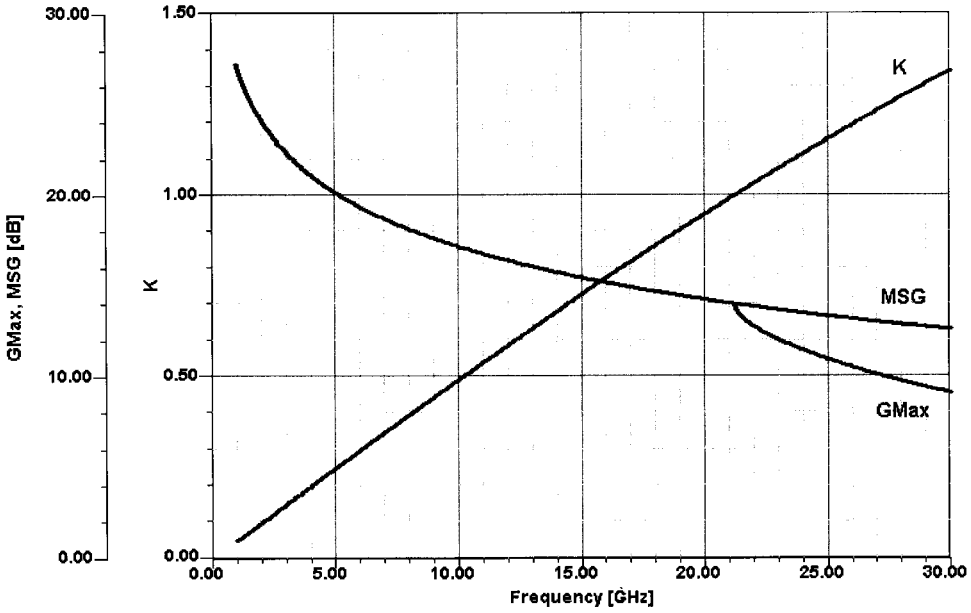


Figure 3-10 G_{max}, maximum stable gain (MSG), and K performance of the amplifier in Figure 3-8.

definition of stability that implies that input and output are terminated with pure resistances and then conjugately matched. There are four ways of achieving this resistive loading, either in the emitter as a shunt resistor between gate/drain or base/collector, or in parallel to the input and output to ground. Needless to say, this is not a real operating mode because it deteriorates both noise and gain. In our case, the entity G_{max} is a function of S_{12} of the circuit, not just the transistor, and depends highly on S_{12}. If we bring the system close to oscillation, then G_{max} will increase at the expense of bandwidth. Unfortunately, some of these microwave definitions of gain (see Table 3-1) have become widely used in industry without always being useful. Evaluating these equations for S_{12} is not legitimate because other quantities like K become infinite, and the entire system of equations falls apart. The power gain with input and output conjugately matched, or the transducer power gain for arbitrary source and load matching, still give more insight into the system.

In the above example, we specifically stated that the input and output VSWR was totally unacceptable, and this was due to a one-stage design. Figure 3-11 shows an amplifier chip built around two transistors with standard feedback techniques applied; Figures 3-12 and 3-13 show its frequency-dependent gain, matching, stability, and noise performance. This amplifier was one of the early examples of GaAsFET wideband amplifiers, but it is interesting because while it has wild (and wide) gain swings, it exhibits a fairly constant input and output impedance or reflection coefficient. A further evaluation of the circuit shows that not only is the circuit unconditionally stable, but that the maximum stable gain has an extremely high value based on essentially minimal S_{12}, the K factor is always significantly higher than 1, and G_{max} follows, to some degree, the actual gain response of the amplifier. (Ultimately, designers gave up on simple *RCL* feedback amplifiers for such a wide bandwidth, replacing them with distributed amplifiers previously called traveling wave amplifiers.) The problem of gain and stability starts at audiofrequency amplifiers, where the

Table 3-1 Nine power gains

Transducer power gain in 50-Ω system	$G_T = \lvert S_{21}\rvert^2$
Transducer power gain for arbitrary Γ_G and Γ_L	$G_T = \dfrac{(1-\lvert\Gamma_G\rvert^2)\,\lvert S_{21}\rvert^2\,(1-\lvert\Gamma_L\rvert^2)}{\lvert(1-S_{11}\Gamma_G)\,(1-S_{22}\Gamma_L)-S_{12}S_{21}\Gamma_G\Gamma_L\rvert^2}$
Unilateral transducer power gain	$G_{\mathrm{TU}} = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_G\rvert^2)\,(1-\lvert\Gamma_L\rvert^2)}{\lvert 1-S_{11}\Gamma_G\rvert^2\lvert 1-S_{22}\Gamma_L\rvert^2}$
Power gain with input conjugate matched	$G = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_L\rvert^2)}{\lvert 1-S_{22}\Gamma_L\rvert^2(1-\lvert S'_{11}\rvert^2)} = \dfrac{\lvert S_{21}\rvert^2}{1-\lvert S_{11}\rvert^2}$ (for $\Gamma_L = 0$)
Available power gain with output conjugate matched	$G_A = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_G\rvert^2)}{\lvert 1-S_{11}\Gamma_G\rvert^2(1-\lvert S'_{22}\rvert^2)} = \dfrac{\lvert S_{21}\rvert^2}{1-\lvert S_{22}\rvert^2}$ (for $\Gamma_G = 0$)
Maximum available power gain[a]	$G_{ma} = \left\lvert\dfrac{S_{21}}{S_{12}}\right\rvert\,(k-\sqrt{k^2-1})$
Maximum unilateral transducer power gain	$G_{\mathrm{TUmax}} = \dfrac{\lvert S_{21}\rvert^2}{(1-\lvert S_{11}\rvert^2)(1-\lvert S_{22}\rvert^2)}$
Maximum stable power gain	$G_{ms} = \dfrac{\lvert S_{21}\rvert}{\lvert S_{12}\rvert}$
Unilateral power gain	$U = \dfrac{1/2\lvert S_{21}/S_{12}-1\rvert^2}{k\lvert S_{21}/S_{12}\rvert-\mathrm{Re}(S_{21}/S_{12})}$

[a]Definition for $G_{\max}$: If $k > 1$, it is called G_{ma}, and if $k < 1$, it is called G_{ms}.

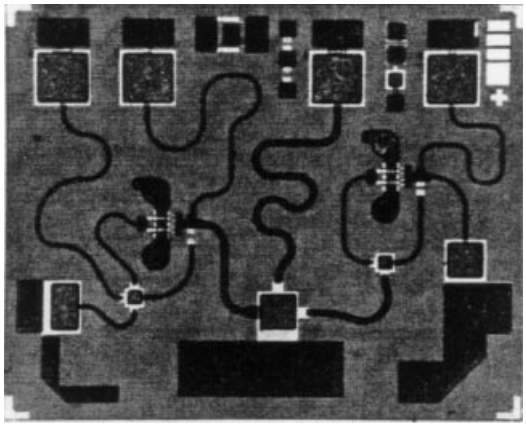


Figure 3-11 A two-stage, wideband feedback amplifier.

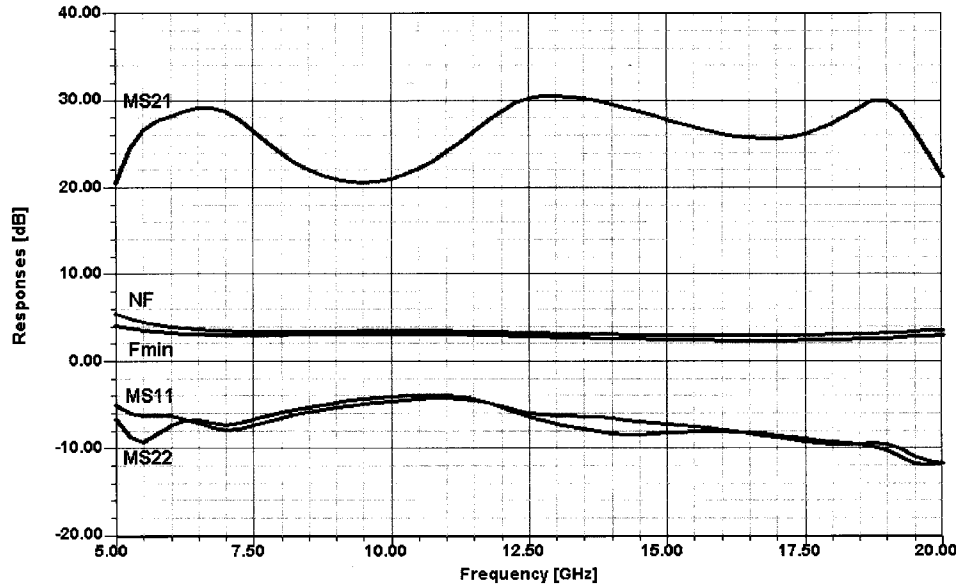


Figure 3-12 Frequency-dependent gain, matching, and noise performance for the amplifier in Figure 3-11.

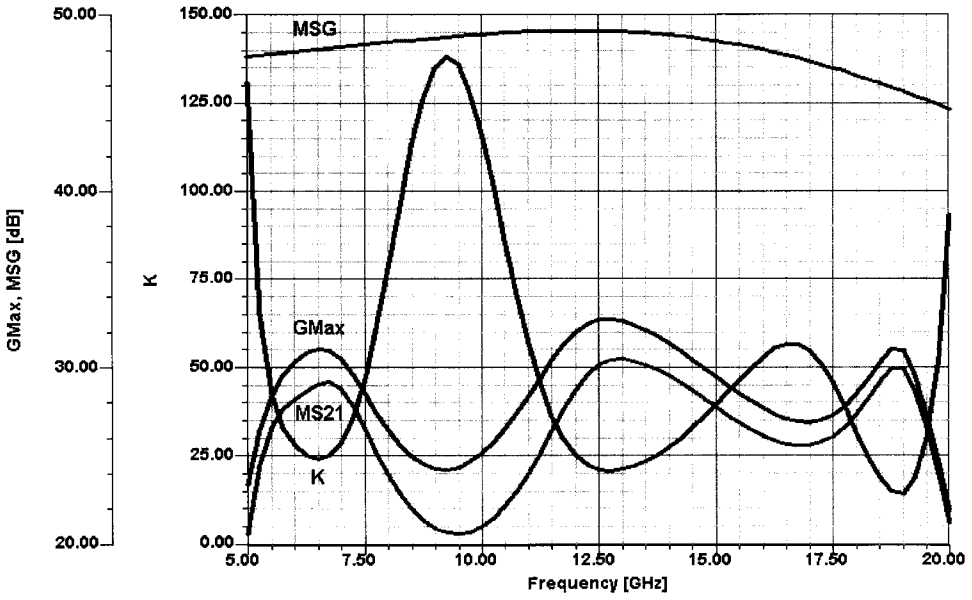


Figure 3-13 Frequency-dependent G_{max}, K, maximum stable gain, and gain performance for the amplifier in Figure 3-11.

motorboating phenomenon has been known for years, and the problem moves up in frequency with more modern technology and is largely dependent not only on the device and the circuit, but also on the layout. The motorboating effect has to do with ground loops and the fact that if the printed circuit traces from the power supply connection to the output power stage present too high a resistance, then at high currents the voltage will drop to a point below useful operation and then, as the volume is lowered, become stable again. In this case the amplitude of the signal through the amplifier controls the stability of the output stage. Needless to say, this most unpleasant effect applies to high-power stages for all output transistors, where we use Class B or C operation, meaning that the dc operating current as a function of output does not remain constant.

In terms of gain, there is also an important gain characteristic called *differential gain* and *differential phase*. These expressions were first used in television circuits, where linearity in both areas is an absolute must. The amplitude of the signal determines the black and white contrast, and the phase determines the color. As an example, transmitting an image of one of our presidents playing golf in California to the New York area is done by many microwave hops. Any multipath reception, such as that caused by reflection from airplanes, may have little effect on the signal amplitude, but the resulting phase shift can transform a green lawn to a red lawn if a total shift of 180° occurs. Given a particular operating point, if the 1-dB compression point of multiple amplifiers in the chain is approached or exceeded, their differential gain, which is defined by the IEEE as "the difference between (a) the ratio of the output of a small high-frequency sine-wave signal at two different levels of a low-frequency signal on which it is superimposed, and (b) unity," may increase. This mode of saturation charges all the capacitances of the amplifier transistors, causing huge phase shifts that result in the known problem of changing colors. The same applies, of course, in test equipment amplifiers, in which such changes cannot be tolerated. A somewhat similar, but not quite the same, problem is spectral regrowth [evaluated in terms of adjacent-channel power ratio (ACPR)], which is also due to nonlinearities. The question of linearity has already been covered in detail in Chapter 1. Needless to say, any overdrive condition also changes an amplifier's input and output SWR.

3-1-3 Noise Figure (NF)*

Introduction. In this section we look at the best way to describe noise in active devices (black boxes). Even when a two-port is linear, the output waveform will differ from the input, because of the failure to transmit all spectral components with equal gain (or attenuation) and delay. By careful design of the two-port, or by limitation of the bandwidth of the input waveform, such distortions can largely be avoided. However, noise generated within the two-port can still change the waveform of the output signal. In a linear passive two-port, noise arises only from the losses in the two-port; thermodynamic considerations indicate that such losses result in the random changes that we call noise. When the two-port contains active devices, such as transistors, there are other noise mechanisms present. A very important consideration in a system is the amount of noise that it adds to the transmitted signal. This is often judged by the ratio of the output signal power to the output noise power

*Section 3-1-3 is reprinted from Chapter 2 of George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*, copyright 1990 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

(S/N). The ratio of signal plus noise power to noise power [$(S + N)/N$] is generally easier to measure and approaches S/N when the signal is large.

In the evaluation of a two-port it is important to know the amount of noise added to a signal passing through it. An important parameter for expressing this characteristic is the noise factor. The signal energy coming from a generator or antenna is amplified or attenuated in passing from the input to the output of a two-port, as is the noise that accompanies the input signal energy. A system generally includes a cascade of two-port networks that constitute one overall two-port that amplifies the signal to a high-enough power level for its intended use. The *noise factor* of a system is defined as the ratio of signal-to-noise ratios available at input and output:

$$F = \frac{(S/N)_{\text{input}}}{(S/N)_{\text{output}}} \geq 1 \tag{3-2}$$

The noise figure (or factor) of a receiver is an easily measured quantity that describes the signal-to-noise ratio reduction of that receiver.

When this ratio of powers is converted to decibels, it is generally referred to as the *noise figure* rather than noise factor. Various conventions are used to distinguish the symbols used for noise factor and noise figure. Here we use F to represent the noise factor and NF to represent the noise figure, although the terms are usually used interchangeably.

For an amplifier with the power gain G, the noise factor can be rearranged as

$$F = \frac{S_i/N_i}{GS_i/G(N_i + N_a)} \tag{3-3}$$

when N_a is the additional noise power added by the amplifier referred to the input. This can be computed to be

$$F = 1 + N_a/N_i \tag{3-4}$$

The noise factor is often replaced by the noise figure (NF), which is defined in decibels as

$$\text{NF} = 10 \log_{10} F \tag{3-5}$$

In applications like satellite receivers, the noise factor becomes such a small number that it is inconvenient to work with. Many people have adopted the use of an equivalent noise temperature for a circuit to remedy this situation. Since the thermal noise power available from a resistor at temperature T_e is

$$N = kT_eB \tag{3-6}$$

where k is Boltzmann's constant (1.38×10^{-23} J/K), T_e is the effective temperature in kelvins, and B is the bandwidth in hertz. The equation above may be used to associate an effective noise temperature with circuits containing more than just thermal noise sources. This allows Eq. (3-4) to be written as

$$F = 1 + \frac{kT_eB}{kT_0B} = 1 + \frac{T_e}{T_0} \tag{3-7}$$

where T_e is the effective noise temperature of the circuit and T_0 is the temperature of the generator resistor in kelvins. The noise temperature T_e now characterizes our circuit noise contribution and can directly be related to the noise factor.

Assuming a reference noise temperature of 290 K (–273 + 290 = 17 °C), let us determine the noise temperature of the system with a noise factor of 2.6 (4.15 dB):

$$T_e = (2.6 - 1)\ (290) = 464 \text{ K}$$

This temperature T_e should not be confused with the environmental operating temperature T_0. It is quite common to operate low-noise amplifiers with T_e below 100 K at an ambient temperature of 290 K.

Signal-to-Noise Ratio. Let us consider the signal-to-noise ratio of power delivered from a generator to a load as shown in Figure 3-14. The signal power delivered to the input is given by

$$S_{\text{in}} = P_{\text{in}} = \frac{E_g^2 \text{Re}(Z_{\text{in}})}{|Z_g + Z_{\text{in}}|^2} \tag{3-8}$$

where E_g is the rms voltage of the input signal supplied to the system, and the noise power supplied to the input is expressed by

$$N_{\text{in}} = \frac{\overline{v_n^2} \text{Re}(Z_{\text{in}})}{|Z_g + Z_{\text{in}}|^2} \tag{3-9}$$

where the noise power at the input is provided by the noise energy of the real part of Z_g. The input impedance Z of the system in the form $Z = R_{\text{in}} + jX_{\text{in}}$ is assumed to be complex.

The Johnson noise of a resistor [here $\text{Re}(Z_g)$], given by the mean-square voltage

$$\overline{v_n^2} = 4kTRB \tag{3-10}$$

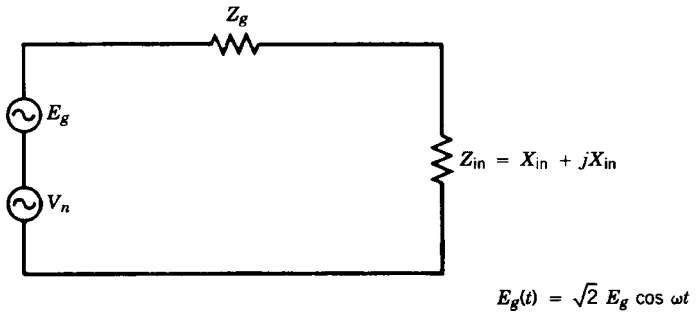


Figure 3-14 Combination of signal and noise voltages supplied to a complex termination.

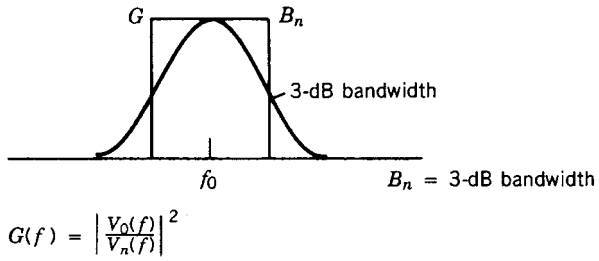
where k (Boltzmann's constant) = 1.38×10^{-23} J/K, T is the absolute temperature of the resistor, and B is the bandwidth, is sufficiently small that the resistive component of impedance does not change. The available signal power from the generator has a lower limit, even if the signal is attenuated by the highest possible attenuation. The generator resistor acts as a Johnson noise generator, its power being

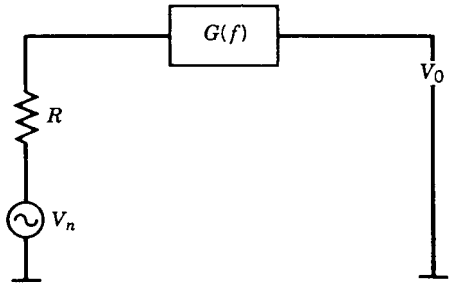
$$P_A = \frac{4kTRB}{4R} = kTB \tag{3-11}$$

where k is Boltzmann's constant, T is the absolute temperature, and B is the bandwidth. This power is the maximum available output power.

For an ambient temperature of 290 K, $kT = 4 \times 10^{-21}$ W/Hz. This expression is also given as kT = –204 dBW/Hz = –174 dBm/Hz = –114 dBm/MHz. We can combine Eqs. (3-8) to (3-10) to obtain

$$\left(\frac{S}{N}\right)_{\text{in}} = \frac{E_g^2}{4kT\text{Re}(Z_g)B} \tag{3-12}$$





$$V_0^2 = \int_0^\infty 4kT_0RG(f)\,dF$$
$$= 4kT_0R \int_0^\infty G(f)\,dF$$
$$B_n = \frac{1}{G} \int_0^\infty G(f)\,dF$$
B_n = noise bandwidth

Figure 3-15 Graphical and mathematical explanations of the noise bandwidth from a comparison of the Gaussian-shaped bandwidth to the rectangular filter response.

This is the value of *S/N* contributed by the generator, which does not include the noise generated by the load, in this case $\mathrm{Re}(Z_{\mathrm{in}})$, which would need to be included in the measurement of the total *S/N* across the input impedance.

A critical parameter is the noise bandwidth, B_n, which is defined as the equivalent bandwidth, as shown in Figure 3-15. For reasons of group delay correction, most practical filters have round rather than sharp corners. The noise figure measurements shown later can be used to determine the "integrated" bandwidth, which is B_n.

An active system such as a combination of amplifiers and mixers will add noise to the input signals, and the noise factor that describes this is defined as the *S/N* ratio at the input to the *S/N* ratio at the output, which is always greater than unity [2]. In practice, a certain minimum signal-to-noise ratio is required for operation. For example, in a communication system such a minimum is required for intelligible transmission, either voice or data. For high-performance TV reception, to provide a noise-free picture to the eye, a typical requirement is for a 60-dB *S/N*. In the case of a TV system, a large dynamic range is required, as well as a very large bandwidth to reproduce all colors truthfully and all shades from high-intensity white to black. Good systems will have a bandwidth of 8 MHz or more.

Noise Figure Measurements. Some of the noise equations are based on mathematical models and physics. To understand some of these expressions, it is useful to look at a practical system, with amplifiers, that has to be evaluated.

Let us look at Figure 3-16, which consists of a signal generator, the system or device under test (DUT), and a selective receiver with a built-in root-mean-square (rms) voltmeter to determine the signal and noise voltages. It is necessary that the system have enough gain so that the noise voltage supplied by the generator will be indicated [3].

If we assume that our selective receiver is a video noise meter calibrated in rms voltage levels, we can perform two measurements. With an input termination connected to the TV system (typically, 75 Ω for cable TV, 50 Ω for satellite TV), the noise receiver/meter will read a value for proper termination that can easily be calculated. Since one-half of the mean-square noise voltage appears across the input,

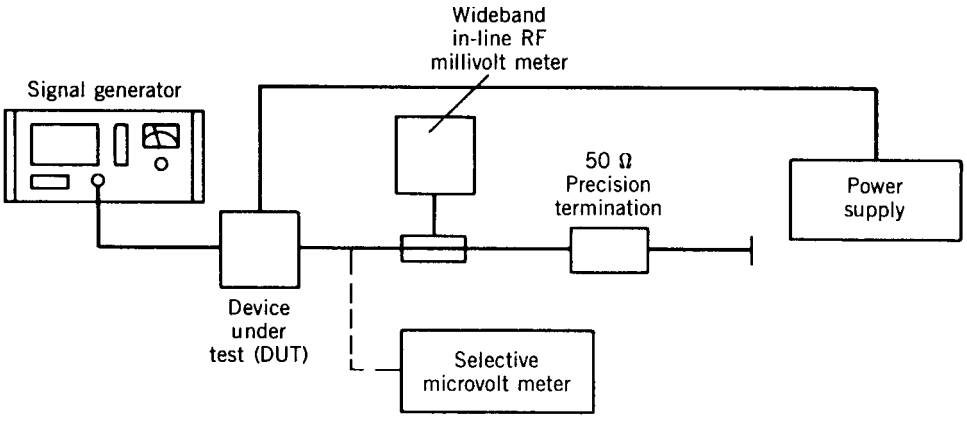


Figure 3-16 Test setup to measure signal-to-noise ratio.

$$v_{\text{in}} = \frac{v_n}{2} = \frac{\sqrt{4kTRB}}{2} \tag{3-13}$$

With $B = 10$ MHz, $T = 290$ K (T is always expressed in absolute temperature; $T_0 = -273$ °C), and $k = 1.38 \times 10^{-23}$ J/K, then for $R = 75\ \Omega$,

$$v_{\text{in}} = \frac{v_n}{2} = 1.73\ \mu\text{V} \tag{3-14}$$

where the rms noise voltage has been referred to the input port. We can verify this with our first measurement.

Now we increase the input voltage of the signal generator to a value that indicates a 60-dB *S/N* ratio at the output port. This should be about

$$E_g = \frac{v_n}{2}\sqrt{F} \times 1000 = 1.73\sqrt{F} \quad \text{mV}$$

where F is the noise factor of the receiver. For a receiver noise factor of 10, we would obtain $E_g = 5.48$ mV (rms value). If the noise energy equivalent to a noise factor of F is assumed, we need $\sqrt{F}$ times more voltage. For a 60-dB ratio, this means that $E_g = 1000 \times (v_n/2) \times \sqrt{F}$.

As they are done here, over a power range of 60 dB, the measurements can be performed over such a wide range only if special equipment is available. In cases where the internal detector of a piece of communication equipment is used, the signal-to-noise ratio measurements are performed over much smaller power ranges.

Let us assume that for the above-mentioned case ($F = 10$) we find a *S/N* ratio of 10 dB at the output for an input signal of 5.47 μV. By rewriting Eq. (3-7) as

$$E_g = \frac{v_n}{2}\sqrt{F} = \sqrt{KTRBF} \tag{3-15}$$

with F being the noise factor, we can solve for F with

$$F = \frac{P_s}{P_n} = \frac{E_g^2/R}{kTB} \tag{3-16}$$

While the input power from the thermal energy of the input termination resistor was $kTB = 4 \times 10^{-4}$ W, the input power required for the 10-dB *S/N* ratio was

$$P_s = \frac{(5.47 \times 10^{-6})^2}{75} = 3.98 \times 10^{-13}\ \text{W} \tag{3-17}$$

The noise factor is defined as the ratio P_s/P_n:

$$F = \frac{3.98 \times 10^{-13}}{4 \times 10^{-14}} = 10 \tag{3-18}$$

which is the proof.

This method is used more frequently at the 3-dB point, or double the input power if the dynamic range of the detector is small or only a linear indicator is available. Because of hum and other pickup, this is not an easy measurement. Using a signal generator is very expensive because in a laboratory or production environment a wide frequency range requires several generators.

Another method is the use of a wideband noise generator. Modern gas discharge diodes or avalanche diodes are available that provide essentially white noise energy over a large frequency range. These microwave diodes typically have an output of 30 dB above kT when switched on and kT when switched off. To provide good matching at microwave frequencies, a 15-dB attenuator is cascaded. This means that the noise power of the source in the ON condition is about 15 dB above kT.

In the early 1960s, low-cost noise figure test equipment was built around vacuum diodes whose operating range was limited to 1200 MHz due to the resonance effects of the structure. Today, automatic noise gain analyzers use calibrated solid-state noise sources up to 26.5 GHz. It appears that the upper frequency limit has to do with matching, and the lower-frequency limit with $1/f$ noise.

Noisy Two-Port Description. Based on the convention by Rothe and Dalke [4], any linear two-port can be in the form shown in Figure 3-17. This general case of a noisy two-port can be redrawn showing noise sources at the input and at the output. Figure 3-17b shows this in admittance form and Figure 3-17c in impedance form. The internal noise sources are assumed to produce very small currents and voltages, and we assume that linear two-port equations are valid. The internal noise contributions have been expressed by using external noise sources:

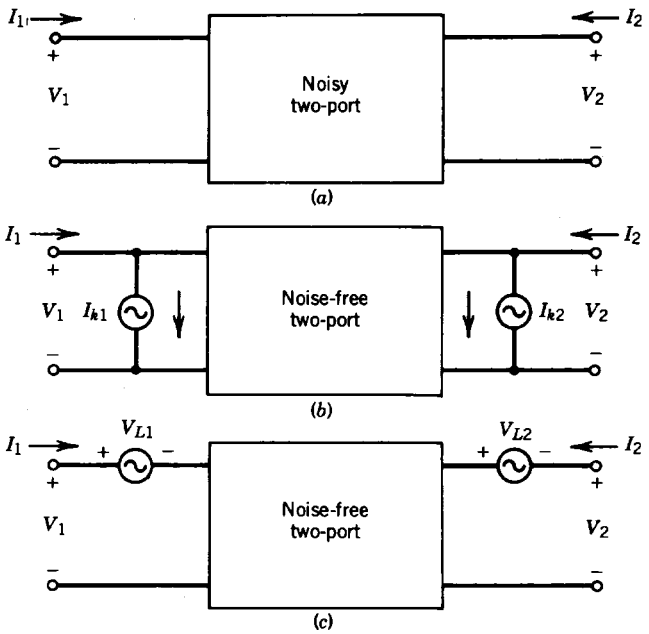


Figure 3-17 Noisy linear two-ports: (a) general form, (b) admittance form, and (c) impedance form.

$$I_1 = y_{11}V_1 + y_{12}V_2 + I_{K1}$$
$$I_2 = y_{21}V_1 + y_{22}V_2 + I_{K2} \tag{3-19}$$

$$V_1 = z_{11}I_1 + z_{12}I_2 + V_{L1}$$
$$V_2 = z_{21}I_1 + z_{22}I_2 + V_{L2} \tag{3-20}$$

where the external noise sources are I_{K1}, I_{K2}, V_{L1}, and V_{L2}.

Since we want to describe our noisy circuit in terms of the noise figure, the *ABCD*-matrix description will be more convenient since it refers both noise sources to the input of the two-port [5]. This representation is given below (note the change in direction of I_2):

$$V_1 = AV_2 + BI_2 + V_A$$
$$I_1 = CV_2 + DI_2 + I_A \tag{3-21}$$

where V_A and I_A are the external noise sources.

It is important to remember that all of these matrix representations are interrelated. For example, the noise sources for the *ABCD*-matrix description can be obtained from the *z*-matrix representation shown in Eq. (3-20). This transformation is

$$V_A = -\frac{I_{K2}}{y_{21}} = V_{L1} - \frac{V_{L2}z_{11}}{z_{21}} \tag{3-22}$$

$$I_A = I_{K1} - \frac{I_{K2}y_{11}}{y_{21}} = -\frac{V_{L2}}{z_{21}} \tag{3-23}$$

The *ABCD* representation is particularly useful because it allows us to define a noise temperature for the two-port referenced to the input. The two-port itself (shown in Figure 3-18) is assumed to be noise-free.

In the past, *z* and *y* parameters have been used, but in microwave applications it has become common to use *S*-parameter definitions. This is shown in Figure 3-19. The previous equations can be rewritten in their new form using *S*-parameters:

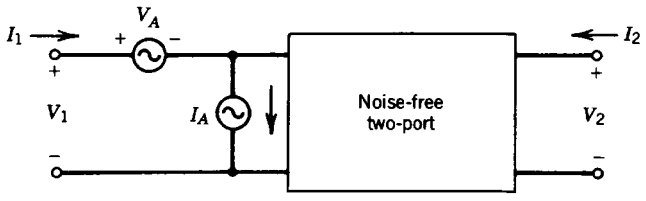


Figure 3-18 Chain-matrix form of linear noisy two-ports.

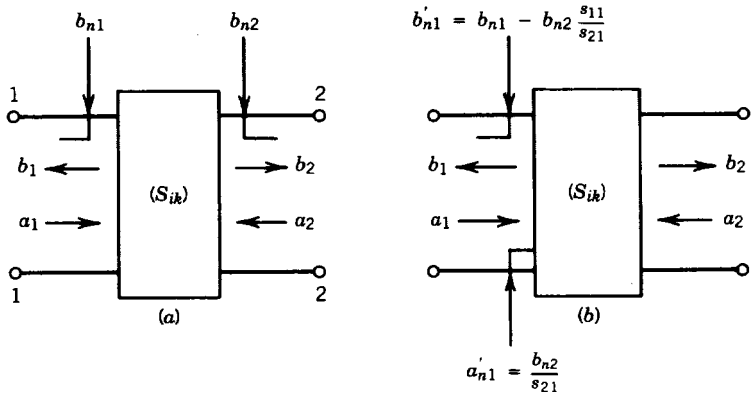


Figure 3-19 *S*-parameter form of linear noisy two-ports.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_{n1} \\ b_{n2} \end{bmatrix} \tag{3-24}$$

There are different physical origins for the various sources of noise. Typically, thermal noise is generated by resistances and loss in the circuit or transistor, whereas shot noise is generated by current flowing through semiconductor junctions and vacuum tubes. Since these many sources of noise are represented by only two noise sources at the device input, the two equivalent noise sources are often a complicated combination of the circuit internal noise sources. Often, some fraction of V_A and I_A is related to the same noise source. This means that V_A and I_A are not independent in general. Before we can use V_A and I_A to calculate the noise figure of the two-port, we must calculate the correlation between the V_A and I_A shown in Figure 3-18.

The noise source V_A represents all the device noise referred to the input when the generator impedance is zero; that is, when the input is short-circuited. The noise source I_A represents all the device noise referred to the input when the generator admittance is zero; that is, the input is open-circuited.

The correlation of these two noise sources considerably complicates analysis. By defining a correlation admittance, we can simplify the mathematics and get some physical intuition for the relationship between noise figure and generator admittance. Since some fraction of I_A will be correlated with V_A, we split I_A into correlated and uncorrelated parts as follows:

$$I_A = I_n + I_u \tag{3-25}$$

I_u is the part of I_A uncorrelated with V_A. Since I_n is correlated with V_A, we can say that I_n is proportional to V_A and the constant of proportionality is the correlation admittance:

$$I_n = Y_{\text{cor}} V_A \tag{3-26}$$

This leads us to

$$I_A = Y_{\text{cor}} V_A + I_u \tag{3-27}$$

The following derivation of noise figure will use the correlation admittance. Y_{cor} is not a physical component located somewhere in the circuit. Y_{cor} is a complex number derived by correlating the random variables I_A and V_A. To calculate Y_{cor}, we multiply each side of Eq. (3-27) by V_A^* and average the result. This gives

$$\overline{V_A^* I_A} = Y_{\text{cor}} \overline{V_A^2} \tag{3-28}$$

where the I_u term averaged to zero since it was uncorrelated with V_A. The correlation admittance is thus given by

$$Y_{\text{cor}} = \frac{\overline{V_A^* I_A}}{V_A^2} \tag{3-29}$$

Often, people use the term "correlation coefficient." This normalized quantity is defined as

$$c = \frac{\overline{V_A^* I_A}}{\sqrt{\overline{V_A^2} I_A^2}} = Y_{\text{cor}} \sqrt{\frac{\overline{V_A^2}}{I_A^2}} \tag{3-30}$$

Note that the dual of this admittance description is the impedance description. Thus, the impedance representation has the same equations as above with Y replaced by Z, I replaced by V, and V replaced by I.

V_A and I_A represent internal noise sources in the form of a voltage source acting in series with the input voltage and a source of current flowing in parallel with the input current. This representation conveniently leads to the four noise parameters needed to describe the noise performance of the two-port. Again using the Nyquist formula, the open-circuit voltage of a resistor at the temperature T is

$$\overline{V_A^2} = 4kTRB \tag{3-31}$$

This voltage is a mean-square fluctuation (or spectral density). It is the method used to calculate the noise identity. We could also define a noise equivalent resistance for a noise voltage as

$$R_n = \frac{\overline{V_A^2}}{4kTB} \tag{3-32}$$

The resistor R_n is not a physical resistor but can be used to simulate different portions of the noise equivalent circuit.

In a similar manner a mean-square current fluctuation can be represented in terms of an equivalent noise conductance G_n, which is defined by

$$G_n = \frac{\overline{I_A^2}}{4kTB} \tag{3-33}$$

and

$$G_u = \frac{\overline{I_u^2}}{4kTB} \tag{3-34}$$

for the case of the uncorrelated noise component. The input generator to the two-port has a similar contribution:

$$G_G = \frac{I_G^2}{4kTB} \tag{3-35}$$

with Y_G being the generator admittance and G_G being the real part. With the definition of F above, we can write

$$F = 1 + \left| \frac{I_A + Y_G V_A}{I_G} \right|^2 \tag{3-36}$$

The use of the voltage V_A and the current I_A has allowed us to combine all the effects of the internal noise sources.

We can use the previously defined [Eq. (3-29)] correlation admittance, $Y_{\text{cor}} = G_{\text{cor}} + jB_{\text{cor}}$, to simplify Eq. (3-36). First, we determine the total noise current:

$$\overline{I_A^2} = 4kT(|Y_{\text{cor}}|^2 R_n + G_u)B \tag{3-37}$$

where R_n and G_u are defined in Eqs. (3-32) and (3-34). The noise factor can now be determined:

$$F = 1 + \frac{G_u}{G_g} + \frac{R_n}{G_g}[(G_G + G_{\text{cor}})^2 + (B_G + B_{\text{cor}})^2] \tag{3-38}$$

$$F = 1 + \frac{R_u}{R_g} + \frac{G_n}{R_g}[(R_G + R_{\text{cor}})^2 + (X_G + X_{\text{cor}})^2] \tag{3-39}$$

The noise factor is a function of various elements, and the optimum impedance for the best noise figure can be determined by minimizing F with respect to generator reactance and resistance. This gives

$$R_{0n} = \sqrt{\frac{R_u}{G_n} + R_{\text{cor}}^2} \tag{3-40}$$

$$X_{0n} = -X_{\text{cor}} \tag{3-41}$$

and

$$F_{\min} = 1 + 2G_n R_{\text{cor}} + 2\sqrt{R_u G_n + (G_n R_{\text{cor}})^2} \tag{3-42}$$

(To distinguish between optimum noise and optimum power, we have introduced the convention 0*n* instead of the more familiar abbreviation *opt*.) At this point we see that the optimum condition for minimum noise figure is not a conjugate power match at the input port. We can explain this by recognizing that the noise source V_A and I_A represent all the

two-port noise, not just the thermal noise of the input port. We should observe that the optimum generator susceptance, $-X_{corr}$, will minimize the noise contribution of the two noise generators.

In rearranging for conversion to S parameters, we write

$$F = F_{\min} + \frac{g_n}{R_G}|Z_G - Z_{0n}|^2 \tag{3-43}$$

$$F = F_{\min} + \frac{R_n}{G_G}|Y_G - Y_{0n}|^2 \tag{3-44}$$

From the definition of the reflection coefficient,

$$\Gamma_G = \frac{Y_0 - Y_G}{Y_0 + Y_G} \tag{3-45}$$

and with

$$g_G = \frac{G_G}{Y_0} \tag{3-46}$$

$$r_n = \frac{R_n}{Z_0} \tag{3-47}$$

the normalized equivalent noise resistance,

$$F = F_{\min} + \frac{4r_n|\Gamma_G - \Gamma_{0n}|^2}{g_G(1 - |\Gamma|^2)|1 + \Gamma_{0n}|^2} \tag{3-48}$$

$$r_n = (F_{50} - F_{\min})\frac{|1 + \Gamma_{0n}|^2}{4|\Gamma_{0n}|^2} \tag{3-49}$$

$$\Gamma_{0n} = \frac{Z_{0n} - Z_0}{Z_{0n} + Z_0} \tag{3-50}$$

The noise performance of any linear two-port can now be determined if the values of the four noise parameters, $F_{\min}$, r_n (= $R_n/50$), and Γ_{0n} (magnitude and phase), are known.

Figure 3-20 shows the noise factor of a high-frequency transistor as a function of B_g for G_g = constant and as a function of G_g for $B_g = B_{opt}$.

Noise Figure of Cascaded Networks. In a system with many circuits connected in cascade (Figure 3-21), we must consider the contributions of the various circuits. In considering the equivalent noise resistor R_n in series with the input circuit,

$$F = \frac{R_G + R_n}{R_G} \tag{3-51}$$

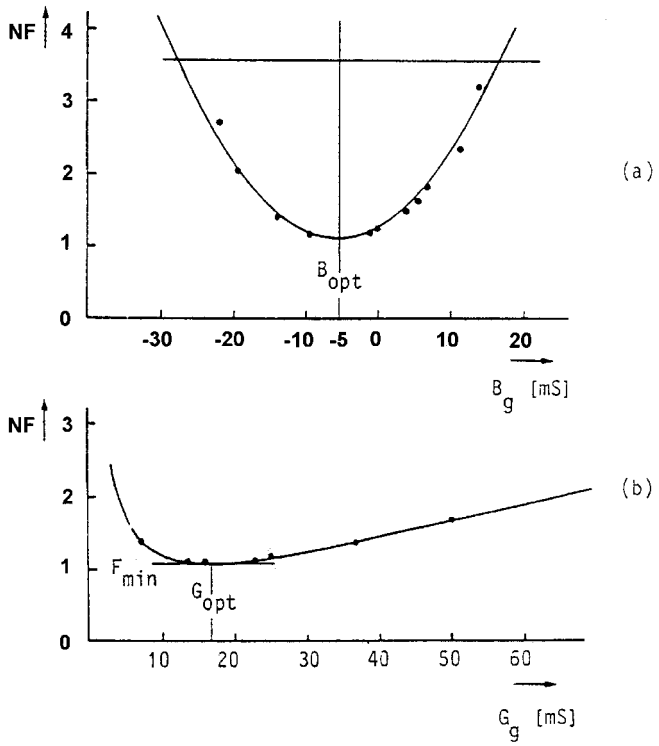


Figure 3-20 Noise factor in high-frequency BJTs for $f = 600$ MHz: (a) as a function of B_g for G_g = constant and (b) as a function of G_g for $B_g = B_{opt}$.

$$F = 1 + \frac{R_n}{R_G} \tag{3-52}$$

The excess noise added by the circuit is R_n/R_G.

In considering two cascaded circuits a and b, by definition the available noise at the output of b is

$$N_{ab} = F_{ab} G_{ab} kTB \tag{3-53}$$

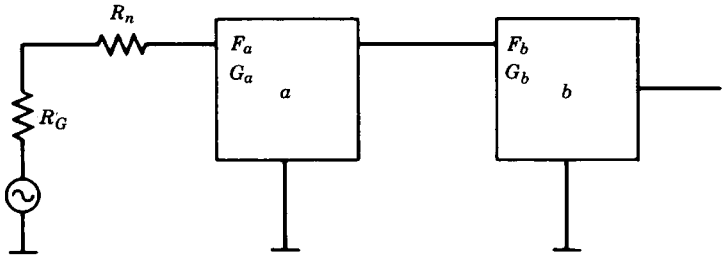


Figure 3-21 Cascaded noisy two-ports with the noise factors F_a and F_b and the gain figures G_a and G_b.

with B the equivalent noise bandwidth in which the noise is measured. The total available gain G is the product of the individual available gains, so

$$N_{ab} = F_{ab} G_a G_b kTB \tag{3-54}$$

The available noise from network a at the output of network b is

$$N_{a/b} = N_a G_b = F_a G_a G_b kTB \tag{3-55}$$

The available noise added by network b (its excess noise) is

$$N_{b/b} = (F_b - 1) G_b kTB \tag{3-56}$$

The total available noise N_{ab} is the sum of the available noise contributed by the two networks:

$$\begin{aligned} N_{ab} &= N_{a/b} + N_{b/b} = F_a G_a G_b kTB + (F_b - 1) G_b kTB \\ &= \left(F_a + \frac{F_b - 1}{G_a} \right) G_a G_b kTB \end{aligned} \tag{3-57}$$

$$F_{ab} = F_a + \frac{F_b - 1}{G_a} \tag{3-58}$$

For any number of circuits, this can be extended to

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \cdots \tag{3-59}$$

When considering a long chain of cascaded amplifiers, there will be a minimum noise figure achievable for this chain. This is a figure of merit and was proposed by Haus and Adler [6]. It is calculated by rearranging Eq. (3-59):

$$(F_{\text{tot}})_{\min} = (F_{\min} - 1) + \frac{F_{\min} - 1}{G_A} + \frac{F_{\min} - 1}{G_A^2} + \cdots + 1 \tag{3-60}$$

where $F_{\min}$ is the minimum noise figure for each stage and G_A is the available power gain of the identical stages. Using

$$\frac{1}{1 - X} = 1 + X + X^2 + \cdots \tag{3-61}$$

we find a quantity $(F_{\text{tot}} - 1)$, which is defined as the noise measure M. The minimum noise measure

$$(F_{\text{tot}})_{\min} - 1 = \frac{F_{\min} - 1}{1 - 1/G_A} = M_{\min} \tag{3-62}$$

refers to the noise of an infinite chain of optimally tuned, low-noise stages, so it represents a lower limit on the noise of an amplifier.

The minimum noise measure $M_{\min}$ is an invariant parameter and is not affected by feedback. It is somewhat similar to a gain–bandwidth product, in its use as a system invariant. The minimum noise measure is achieved when the amplifier is tuned for the available power gain and $\Gamma_G = \Gamma_{0n}$, given by Eq. (3-50).

Influence of External Parasitic Elements. Mounting an active two-port such as a transistor usually adds stray capacitance and lead inductance to the device, as shown in Figure 3-22. These external components consisting of transmission lines and parasitic reactances modify the noise parameters and the gain. Some researchers have published the results of these parasitic effects and have made manual computations or used some limited computer programs.

In a paper by Fukui [7], an attempt was made to determine the necessary equations, but the formulas are too involved even for pocket calculators. A more generic study by Iversen [8] is also very involved because of the various matrix manipulations and is more suitable for a computer. Besser's paper [9] in the *IEEE MTT-S* in 1975 and Vendelin's paper [10] in the same issue have shown for the first time some practical results using computers and even optimization methods, using an early version of COMPACT. The intention of these investigations was to find feedback that modifies the device noise and scattering parameters such that a noise match could also provide a low-input VSWR. It can be seen from these discussions that some feedback, besides resulting in some gain reduction, may improve the noise matching at the input for a limited frequency range.

A more recent paper by Suter [11] based on a report by Hartmann and Strutt [12] has given a simple transformation starting from the *S*-parameters and the noise parameters from common-source (or common-emitter) measurements. The noise parameters for the "packaged" device are calculated. This means that the parameters for the "new" device, including

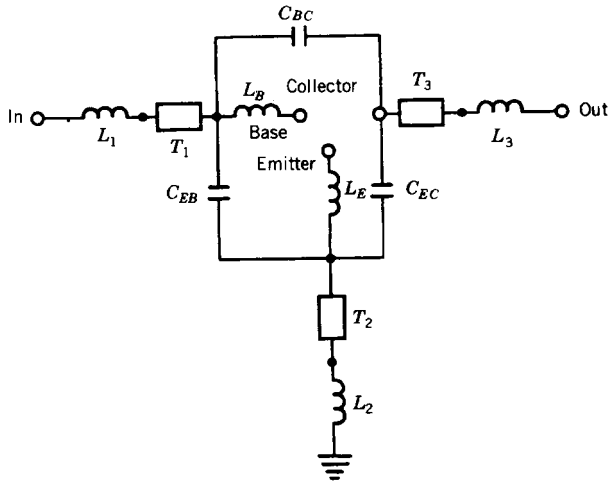
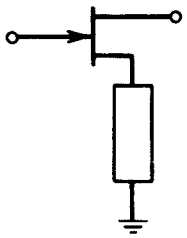


Figure 3-22 Equivalent circuit of the transistor package.

Table 3-2 Transformation matrix parameters

Series Feedback



$$Z_S = R_S + jX_S = Z_S$$

$$[n] = \begin{bmatrix} n_{11} = 1 & n_{12} = Z_0 \dfrac{S_{21}M - S'_{21}N}{S_{21}C'_1 + S'_{21}C_1}\ \Omega \\ n_{21} = 0 & n_{22} = \dfrac{S_{21}C'_1}{S_{21}C'_1 + S'_{21}C_1} \end{bmatrix}$$

where

$$S'_{11} = S'_{22} = \frac{-1}{1 + 2Z_S}$$

$$S'_{12} = S'_{21} = \frac{2Z_S}{1 + 2Z_S}$$

$$M = (1 + S'_{11})(1 - S'_{22}) + S'_{12}S'_{21}$$

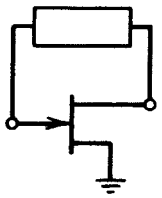
$$N = (1 + S_{11})(1 - S_{22}) + S_{12}S_{21}$$

$$C_1 = (1 - S_{11})(1 - S_{22}) - S_{12}S_{21}$$

$$C'_1 = (1 + S'_{11})(1 - S'_{22}) - S'_{12}S'_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

Shunt Feedback



$$Z_P = R_P + jX_P = Z_P$$

$$[n] = \begin{bmatrix} n_{11} = \dfrac{S_{21}C'_2}{S_{21}C'_2 + S'_{21}C_2} & n_{12} = 0\ \Omega \\ n_{21} = \dfrac{1}{Z_0}\dfrac{S_{21}P - S'_{21}Q}{S_{21}C'_2 + S'_{21}C_2}\,S & n_{22} = 1 \end{bmatrix}$$

where

$$S'_{11} = S'_{22} = \frac{Z_P}{2 + Z_P}$$

(*continued*)

Table 3-2 (*Continued*)

$$S'_{12} = S'_{21} = \frac{2}{2 + Z_P}$$

$$P = (1 - S'_{11})(1 + S'_{22}) + S'_{12}S'_{21}$$

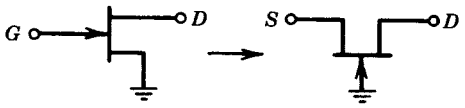
$$Q = (1 - S_{11})(1 + S_{22}) + S_{12}S_{21}$$

$$C_2 = (1 + S_{11})(1 + S_{22}) - S_{12}S_{21}$$

$$C'_2 = (1 + S'_{11})(1 + S'_{22}) - S'_{12}S'_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

Common Gate



$$[n] = \begin{bmatrix} n_{11} = \dfrac{2S_{21}}{-2S_{21} + C_4} & n_{12} = 0\ \Omega \\ n_{21} = \dfrac{1}{Z_0}\dfrac{C_3C_4 - 4S_{12}S_{21}}{V(-2S_{21} + C_4)}S & n_{22} = -1 \end{bmatrix}$$

where

$$V = (1 + S_{11})(1 + S_{22}) - S_{12}S_{21}$$

$$C_3 = (1 - S_{11})(1 + S_{22}) + S_{12}S_{21}$$

$$C_4 = (1 + S_{11})(1 - S_{22}) + S_{12}S_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \text{common-source } S \text{ parameters}$$

the common-gate (or common-base) case, are calculated. The equations are device dependent. They are valid for any active two-port.

A transformation matrix, n, may be used to combine the noise sources of the various circuit configurations. The transformation matrix parameters are given in Table 3-2 for series feedback, shunt feedback, and the common-gate (base) case, which will be important for oscillator analysis. The transformation matrix gives the new noise parameters as follows:

$$R'_n = R_n|n_{11} + n_{12}Y_{\text{cor}}|^2 + G_n|n_{12}|^2 \tag{3-63}$$

$$G'_n = \frac{G_nR_n}{R'_n}|n_{11}n_{22} - n_{12}n_{21}|^2 \tag{3-64}$$

$$Y'_{\text{cor}} = \frac{R_n}{R'_n}(n_{21} + n_{22}Y_{\text{cor}})(n^*_{11} + n^*_{12}Y^*_{\text{cor}}) + \frac{G_n}{R'_n}n_{22}n^*_{12} \tag{3-65}$$

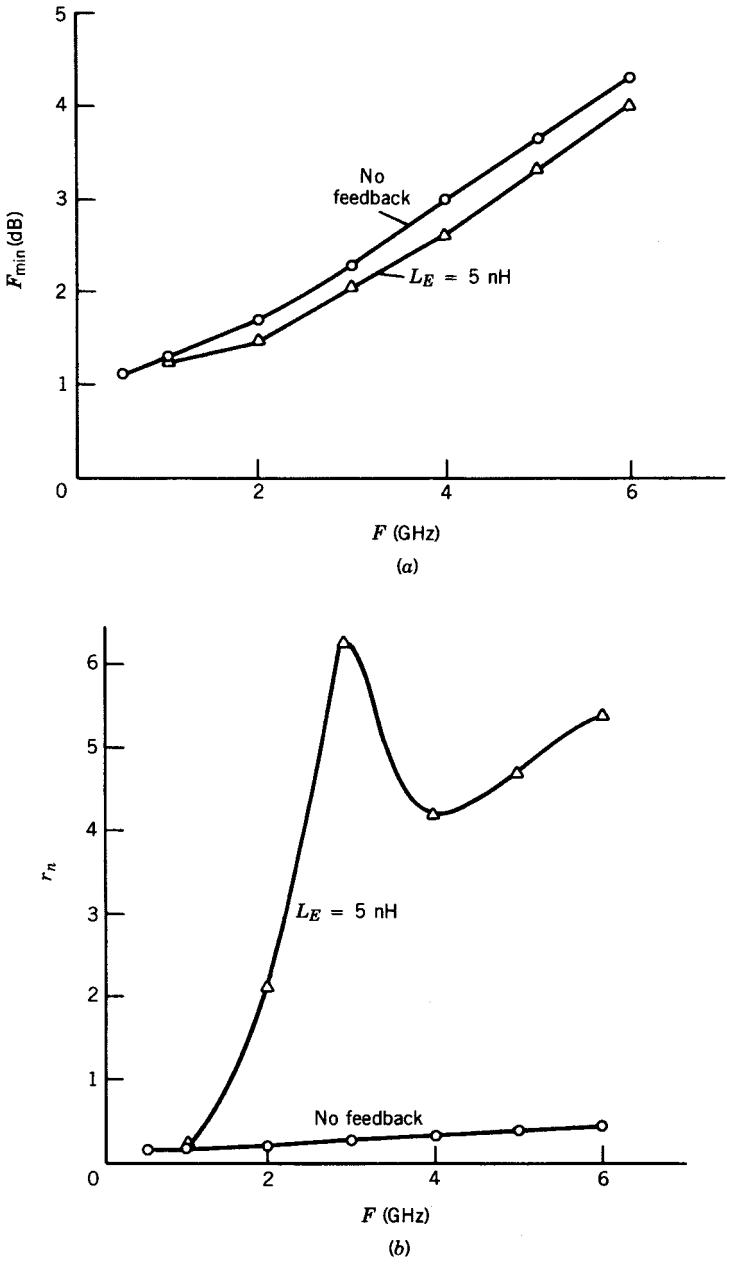


Figure 3-23 Noise parameters versus feedback for the AT-41435 silicon bipolar transistor: (a) F_{min} versus frequency and feedback; (b) r_n versus frequency and feedback; (c) Γ_{0n} versus frequency and feedback.

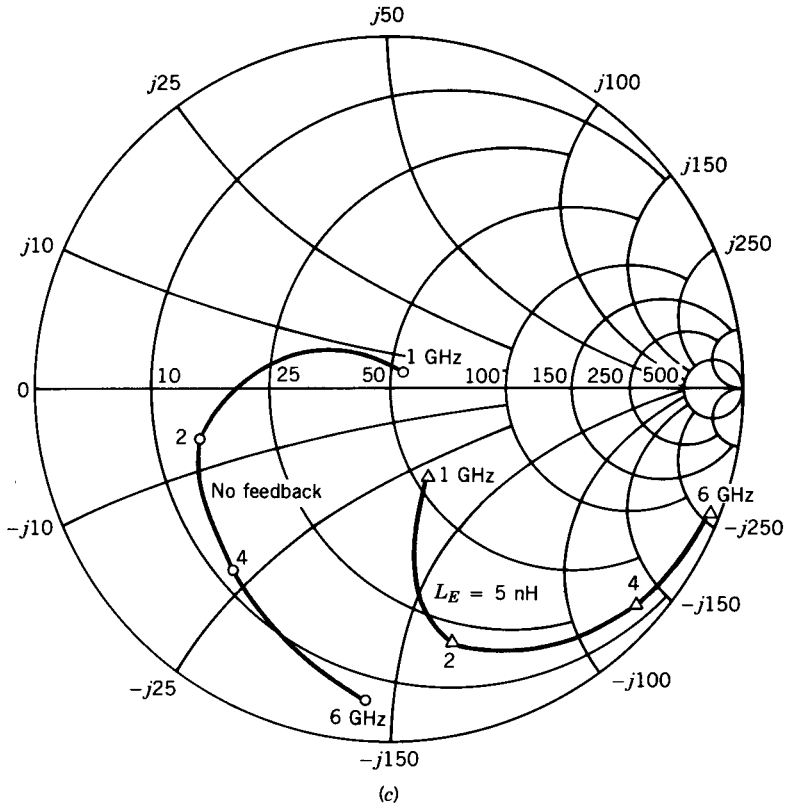


Figure 3-23 (*Continued*)

A final transformation to the more common noise-parameter format given by Eq. (3-44) is still needed [10]:

$$F_{\min} = 1 + 2R_n'(G_{\text{cor}}' + G_{0n}') \tag{3-66}$$

$$R_n = R_n' \tag{3-67}$$

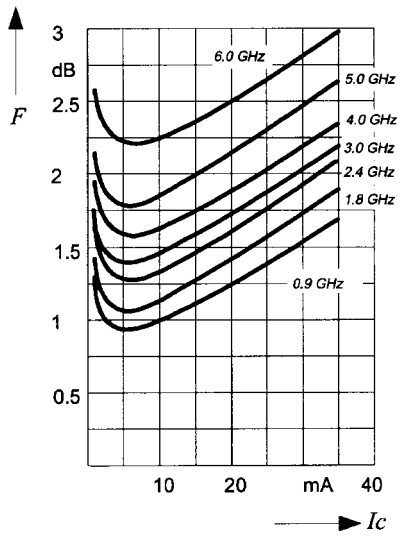
$$G_{0n} = \sqrt{\frac{G_n'}{R_n'} + G_{\text{cor}}'^2} \tag{3-68}$$

$$B_{0n} = -B_{\text{cor}}' \tag{3-69}$$

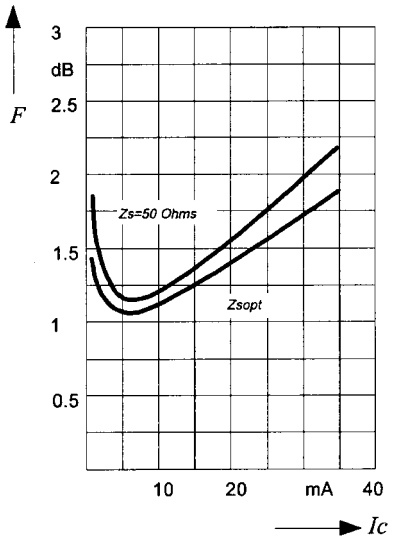
Figure 3-23 shows the noise figure as a function of external feedback for a low-noise microwave bipolar transistor, the AT-41435.

Bias-Dependent Noise Parameters. Since we have elaborated on the effect of parasitics, it is now useful to point out that noise is also a strong function of bias point and frequency. Figure 3-24 shows the noise figure expressed in dB as a function of different dc currents and frequencies for a BFP420 BJT by Siemens. A change in the collector voltage will also affect the noise parameters, not quite as strongly as shown here, but still significantly. The reason

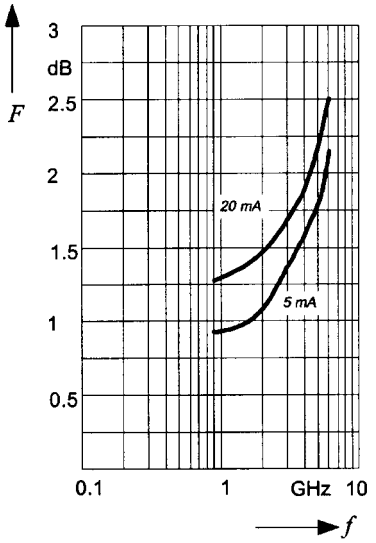
Noise Figure
versus Collector Current
V_{CE} = 2 V, $Z_S = Z_{Sopt}$



Noise Figure
versus Collector Current
VCE = 2 V, f = 1.8 GHz



Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA,
$Z_S = Z_{Sopt}$



Source Impedance for min.
Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA

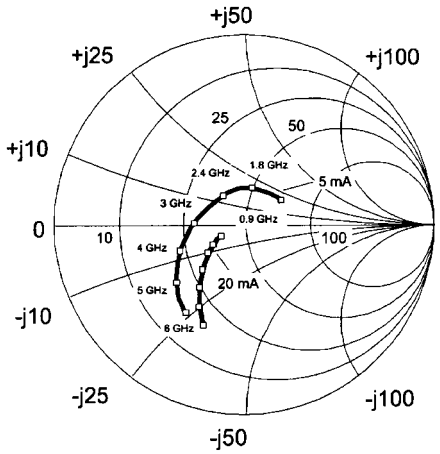


Figure 3-24 Noise figure versus collector current, noise figure versus frequency, and source impedance for minimum noise figure versus frequency for the Siemens BFP420 transistor.

Table 3-3 BFP420 common-emitter noise parameters

f	$F_{\min}{}^a$	$G_a{}^a$	Γ_{opt}		R_N	r_n	$F_{50\Omega}{}^b$	$\lvert S_{21} \rvert^{2b}$
(GHz)	dB	(dB)	(Magnitude)	(Angle)	(Ω)		(dB)	(dB)
V_{CE} = 2 V, I_C = 5 mA								
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	–175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	–157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	–142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	–123.0	15.0	0.30	3.30	5.3

[a]Input matched for minimum noise figure, output for maximum gain.
[b]$Z_S = Z_L = 50\ \Omega$.

for the turning point of the noise figure has to do with the fact that for lower currents the f_T cutoff frequency has not reached its peak yet, and therefore the noise increases. The increase of the noise at higher currents is due to the Schottky noise calculated from the equation

$$P_n = 2Iq \tag{3-70}$$

where P_n is the noise power, I is the saturation, collector, or drain current, and q is the charge of an electron (1.60×10^{-19} coulomb).

Table 3-3 shows the common-emitter noise parameters for the BFP420.

Noise Circles. From the section entitled "Influence of External Parasitic Elements," we see that the noise factor is a strong function of the generator admittance (or impedance) presented to the input terminals of the noisy two-port. Noise tuning is used to change the values of the input admittance to obtain the best noise performance. There is a range of values of input reflection coefficient over which the noise figure is constant. In plotting these points of constant noise figure, we obtain the so-called noise circles, which can be drawn on the Smith chart Γ_G plane [13]. Starting with the noise equation [see Eq. (3-44)] for a 50-Ω generator impedance, we find that

$$F_{50} = F_{\min} + 4r_n \frac{|\Gamma_{0n}|^2}{|1 + \Gamma_{0n}|^2} \tag{3-71}$$

We want to find the position of the reflection coefficient on the Smith chart, as in the case of the gain circles, for which F is constant. First we rearrange Eq. (3-67) to read

$$r_n = (F_{50} - F_{\min}) \frac{|1 + \Gamma_{0n}|^2}{4|\Gamma_{0n}|^2} \tag{3-72}$$

By introducing

$$N_i = \frac{F_i - F_{\min}}{4r_n} |1 + \Gamma_{0n}|^2 \tag{3-73}$$

we can find an expression for a circle of constant noise figure as introduced by Rothe and Dahlke [4] (see also Fukui [13]). The center for the noise circle is

$$C_i = \frac{|\Gamma_{0n}|}{1 + N_i} \tag{3-74}$$

and the radius is

$$r_i = \frac{\sqrt{N_i^2 + N_i(1 - |\Gamma_{0n}|^2)}}{1 + N_i} \tag{3-75}$$

with the definition of N used previously. However, if we consider only the minimum noise figure for a given device, we will not obtain the minimum noise figure for the multistage amplifier system. This was explained when the noise measure was introduced. (See Figure 3-25.) Therefore, a better way to design the amplifier would be to use circles at constant noise measure instead of circles of constant noise figure. This was recently done by Poole

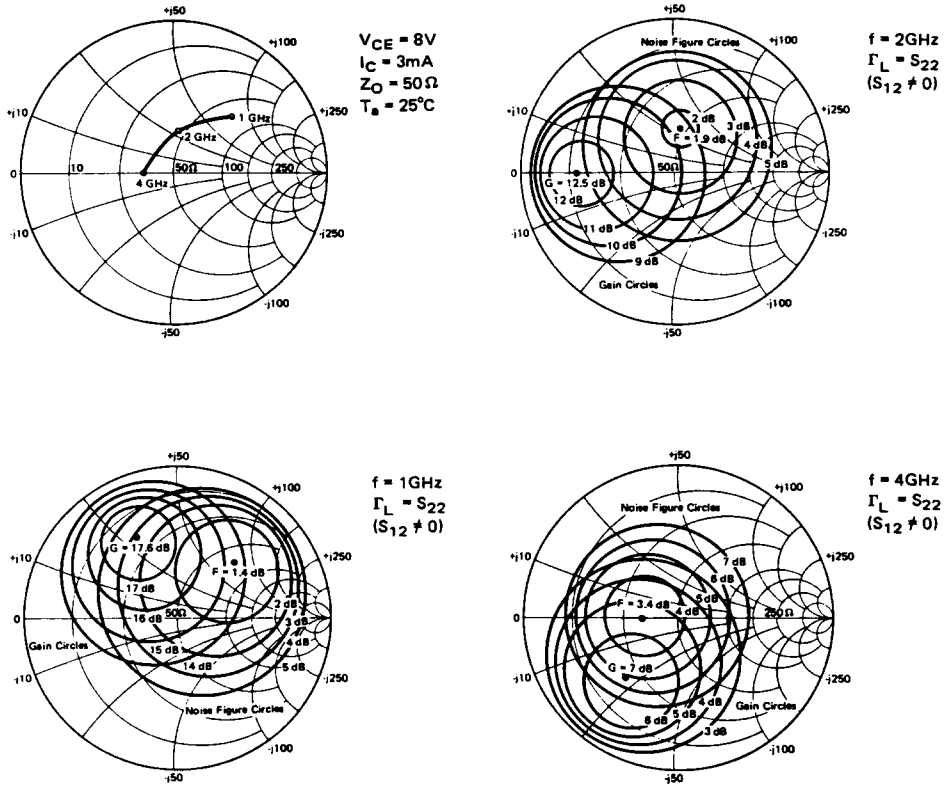


Figure 3-25 Typical noise figure circles and gain circles.

and Paul [14]. They derived the expressions for the noise measure circles as a function of S parameters, noise parameters, and Γ_G, using

$$G_A = \frac{|S_{21}|^2(1 - |\Gamma_G|^2)}{(1 - |S_{22}|^2) + |\Gamma_G|^2 (|S_{11}|^2 - |\Delta|^2) - 2\,\mathrm{Re}(\Gamma_G C_1)} \tag{3-76}$$

where

$$C_1 = S_{11} - S_{22}^* \Delta$$

$$\Delta = S_{11}S_{22} - S_{12}S_{11}$$

In terms of the generator reflection coefficient, the noise measure can be expressed as

$$M = [(F_{\min} - 1)(1 - |\Gamma_G|^2)|S_{21}|^2|1 + \Gamma_{0n}|^2 + 4r_n|S_{21}|^2|\Gamma_G - \Gamma_{0n}|^2]$$

$$\times \left\{|1 + \Gamma_{0n}|^2[|S_{21}|^2(1 - |\Gamma_G|^2) - (1 - |S_{22}|^2) - |\Gamma_G|^2(|S_{11}|^2 - |\Delta|^2) + 2\,\mathrm{Re}(\Gamma_G C_1)]\right\}^{-1} \tag{3-77}$$

Equation (3-77) can be shown to represent circles in the source reflection coefficient plane described by the following:

$$|\Gamma_G|^2 + |\Gamma_m|^2 - \Gamma_G\Gamma_m^* - \Gamma_G^*\Gamma_m = r_m^2 \tag{3-78}$$

The centers and radii of the constant-noise-measure circles are given by

$$C_m = \frac{M|1 + \Gamma_{0n}|^2 C_1^* + 4r_n|S_{21}|^2\Gamma_{0n}}{M|1 + \Gamma_{0n}|^2 P + |S_{21}|^2(4r_n - W)} \tag{3-79}$$

$$r_m = \frac{\sqrt{M^2M_a + MM_b + M_c}}{M|1 + \Gamma_{0n}|^2 P + |S_{21}|^2(4r_n - W)} \tag{3-80}$$

where

$$P = |S_{21}|^2 + |S_{11}|^2 - |\Delta|^2$$

$$Q = |S_{21}|^2 + |S_{22}|^2 - 1$$

$$W = |1 + \Gamma_{0n}|^2(F_{\min} - 1)$$

$$M_a = |1 + \Gamma_{0n}|^4(PQ + |C_1|^2)$$

$$M_b = |1 + \Gamma_{0n}|^2|S_{21}|^2[8r_n\,\mathrm{Re}(\Gamma_{0n}C_1)$$

$$- (4r_n|\Gamma_{0n}|^2 + W)P - (W - 4r_n)Q]$$

$$M_c = |S_{21}|^4[W - 4r_n(1 - |\Gamma_{0n}|^2)]$$

The value of the minimum noise measure can be found by considering the noise measure circle of zero radius; that is, set r_m equal to zero in Eq. (3-80). This results in

$$M_{\min} = \frac{-M_b \pm \sqrt{M_b^2 - 4M_a M_c}}{2M_a} \tag{3-81}$$

Equation (3-81) yields the same value of $M_{\min}$ as would have been obtained by using the immittance parameter equation given by Fukui [13] and can, therefore, be considered as the reflection coefficient plane analog of Fukui's expression. The elimination of the need for the parameter R_{eg}, however, results in a considerable simplification as compared with the earlier approach [13].

The value of the minimum noise measure is taken as the smallest nonnegative value of $M_{\min}$ given by Eq. (3-81). The source reflection coefficient that results in the minimum noise measure can now be obtained by employing Eq. (3-79):

$$\Gamma_{0m} = \frac{M_{\min}|1 + \Gamma_{0n}|^2 C_1^* + 4r_n|S_{21}|^2\Gamma_{0n}}{M_{\min}|1 + \Gamma_{0n}|^2 P + |S_{21}|^2\,(4r_n - W)} \tag{3-82}$$

The output reflection coefficient of the device, when Γ_{0m} is presented to the input port, is given by

$$S'_{22} = \frac{S_{22} - \Delta\Gamma_{0m}}{1 - S_{11}\Gamma_{0m}} \tag{3-83}$$

Noise Correlation in Linear Two-Ports Using Correlation Matrices. In the introduction to two-port noise theory, it was indicated that noise correlation matrices form a general technique for calculating noise in n-port networks. Haus and Adler [6] have described the theory behind this technique. In 1976, Hillbrand and Russer [15] published equations and transformations that aid in supplying this method to two-port CAD.

This method is useful because it forms a base from which we can rigorously calculate the noise of linear two-ports combined in arbitrary ways. For many representations, the method of combining the noise parameters is as simple as that for combining the circuit element matrices. In addition, noise correlation matrices can be used to calculate the noise in linear frequency conversion circuits. The following is an introduction to this subject.

A linear, noisy two-port can be modeled as a noise-free two-port with two additional noise sources. These noise sources must be chosen so that they add directly to the resulting vector of the representation, as shown in Eqs. (3-84) and (3-85) and Figure 3-17:

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \tag{3-84}$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{3-85}$$

where the i and v vectors indicate noise sources for the y and z representations, respectively. This two-port example can be extended to n-ports in a straightforward, obvious way.

Since the noise vector for any representation is a random variable, it is much more convenient to work with the noise correlation matrix. The correlation matrix gives us deterministic numbers with which to calculate. The correlation matrix is formed by taking the mean value of the outer product of the noise vector. This is equivalent to multiplying the noise vector by its adjoint (complex conjugate transpose) and averaging the result:

$$\langle \bar{i}\ \bar{i}^{+} \rangle = \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} [i_1^* \quad i_2^*] = \begin{bmatrix} \langle i_1 i_1^* \rangle & \langle i_1 i_2^* \rangle \\ \langle i_1^* i_2 \rangle & \langle i_2 i_2^* \rangle \end{bmatrix} + [C_y] \tag{3-86}$$

where the angular brackets denote the average value.

Note that the diagonal terms are the "power" spectrum of each noise source and the off-diagonal terms are complex conjugates of each other and represent the "cross-power" spectra of the noise sources. "Power" is used because these magnitude-squared quantities are proportional to power.

To use these correlation matrices in circuit analysis, we must know how to combine them and how to convert them between various representations. An example using y matrices will illustrate the method for combining two-ports and their correlation matrices. Given two matrices y and y', when we parallel them we have the same port voltages, and the terminal currents add (Figure 3-26):

$$\begin{aligned} I_1 &= y_{11}V_1 + y_{12}V_2 + y'_{11}V_1 + y'_{12}V_2 + i_1 + i'_1 \\ I_2 &= y_{21}V_1 + y_{22}V_2 + y'_{21}V_1 + y'_{22}V_2 + i_2 + i'_2 \end{aligned} \tag{3-87}$$

or

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} + y'_{11} & y_{12} + y'_{12} \\ y_{21} + y'_{21} & y_{22} + y'_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 + i'_1 \\ i_2 + i'_2 \end{bmatrix} \tag{3-88}$$

Here we can see that the noise current vectors add just as the y parameters add. Converting the new noise vector to a correlation matrix yields

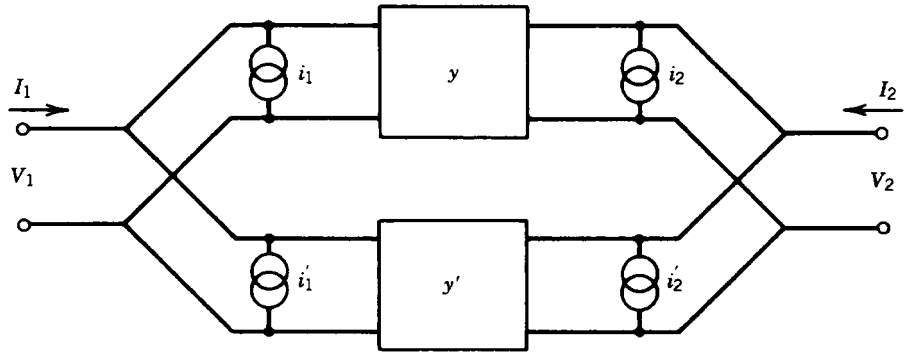


Figure 3-26 Parallel combination of two-ports using y parameters.

$$\langle \bar{i}_{\text{new}} \bar{i}^{+}_{\text{new}} \rangle = \left\langle \begin{bmatrix} i_1 + i_1' \\ i_2 + i_2' \end{bmatrix} [i_1^* + i_1'^* \quad i_2 i_2'^*] \right\rangle \tag{3-89}$$

$$= \begin{bmatrix} \langle i_1 i_1^* \rangle + \langle i_1' i_1'^* \rangle & \langle i_1 i_2^* \rangle + \langle i_1' i_2'^* \rangle \\ \langle i_2 i_1^* \rangle + \langle i_2' i_1'^* \rangle & \langle i_2 i_2^* \rangle + \langle i_2' i_2'^* \rangle \end{bmatrix} \tag{3-90}$$

The noise sources from different two-ports must be uncorrelated, so there are no cross products of different two-ports. By inspection, Eq. (3-90) is just the addition of the correlation matrices for the individual two-ports, so

$$[C_{y\text{new}}] = [C_y] + [C_y'] \tag{3-91}$$

The same holds true for g, h, and z parameters, but $ABCD$ parameters have the more complicated form shown below. If

$$[A_{\text{new}}] = [A][A'] \tag{3-92}$$

then

$$[C_{A\text{new}}] = \{C_A] + [A]\,[C_{A'}]\,[A]^{+} \tag{3-93}$$

The transformation of one representation to another is best illustrated by an example. Let us transform the correlation matrix for a Y representation to a Z representation. Starting with

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = [Y] \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \tag{3-94}$$

we can move the noise vector to the left side and invert y:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = [Y^{-1}] \begin{bmatrix} I_1 - i_1 \\ I_2 - i_2 \end{bmatrix} = [Y^{-1}] \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + [Y^{-1}] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-95}$$

Table 3-4 π Noise matrix $T_{\alpha\beta}$ transformation

		Original Form (α Form)					
		Y		Z		A	
Resulting Form (β Form)	Y	1	0	y_{11}	y_{12}	$-y_{11}$	1
		0	1	y_{21}	y_{22}	$-y_{21}$	0
	Z	z_{11}	z_{12}	1	0	1	$-z_{11}$
		z_{21}	z_{22}	0	1	0	$-z_{21}$
	A	0	A_{12}	1	$-A_{11}$	1	0
		1	A_{22}	0	$-A_{21}$	0	1

Since $(Y)^{-1} = (Z)$, we have

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = [Z] \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + [Z] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-96}$$

so

$$= [Z] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} = [T_{yz}] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-97}$$

where the signs of i_1 and i_2 are superfluous since they will cancel when the correlation matrix is formed. Here the transformation of the Y noise current vector to the Z noise voltage vector is done simply by multiplying by (Z). Other transformations are shown in Table 3-4.

To obtain the noise correlation matrix, we again form the mean of the outer product:

$$\langle vv^+ \rangle = \begin{bmatrix} \langle v_1 v_1^* \rangle & \langle v_1 v_2^* \rangle \\ \rangle v_1^* v_2 \rangle & \langle v_2 v_2^* \rangle \end{bmatrix} = [Z] \left\langle \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} [i_1^* \quad i_2^*] \right\rangle [Z]^+ \tag{3-98}$$

or

$$[C_z] = [Z]\,[C_y]\,[Z]^+ \tag{3-99}$$

where

$$v^+ = [i_1^* \quad i_2^*]\,[Z]^+ \tag{3-100}$$

This is called a congruence transformation. The key to all of these derivations is the construction of a correlation matrix from the noise vector, as shown in Eq. (3-90).

These correlation matrices may easily be derived from the circuit matrices of passive circuits with only thermal noise sources. For example,

$$[C_z] = 2kT\Delta f\,\mathrm{Re}([Z]) \tag{3-101}$$

and

$$[C_y] = 2kT\Delta f\,\mathrm{Re}([Y]) \tag{3-102}$$

The $2kT$ factor comes from the double-sided spectrum of thermal noise. The correlation matrix from the $ABCD$ matrix may be related to the noise figure, as shown by Hillbrand and Russer [15]. We have

$$F = 1 + \frac{\overline{Y}[C_a]\overline{Y}^+}{2kT\,\mathrm{Re}(Y_G)} \tag{3-103}$$

where

$$\overline{Y} = \begin{bmatrix} Y_G \\ 1 \end{bmatrix}$$

Expressing the noise factor in terms of the correlation matrix, we obtain a complete formula:

$$F = 1 + \frac{C_{22}^{A}(f) + 2\,\mathrm{Re}\{Y_g(f)\,C_{12}^{A}(f)\} + |Y_g(f)|^2\,C_{11}^{A}(f)}{2kT_0\,\mathrm{Re}\{Y_g(f)\}} \tag{3-104}$$

Once we transform this in the Y parameter form, we obtain the following equation:

$$F(f) = F_{\min}(f) + \frac{R_n(f)|Y_{\mathrm{opt}}(f) - Y_g(f)|^2}{\mathrm{Re}\{Y_g(f)\}} \tag{3-105}$$

It should be noted that all these values are frequency dependent, as expressed in this equation.

The *ABCD* correlation matrix can be written in terms of the noise figure parameters as (double-sided spectrum)

$$[C_a] = 2kT\begin{bmatrix} R_n & \dfrac{F_0 - 1}{2} - R_n Y_{0n}^{*} \\ \dfrac{F_0 - 1}{2} - R_n Y_{0n} & R_n|Y_{0n}|^2 \end{bmatrix} \tag{3-106}$$

The noise correlation matrix method forms an easy and rigorous technique for handling noise in networks. This technique allows us to calculate the total noise for complicated networks by combining the noise matrices of subcircuits. It should be remembered that although noise correlation matrices apply to n-port networks, noise figure calculations apply only to pairs of ports. The parameters of the C_a matrix can be used to give the noise parameters:

$$Y_{0n} = \sqrt{\frac{C_{ii*}}{C_{uu*}} - \left[\mathrm{Im}\left(\frac{C_{ui*}}{C_{uu*}}\right)\right]^2} + j\,\mathrm{Im}\left(\frac{C_{ui*}}{C_{uu*}}\right) \tag{3-107}$$

$$F_0 = 1 + \frac{C_{ui*} + C_{uu*}Y_{0n}^{*}}{kT} \tag{3-108}$$

$$R_u = C_{uu*} \tag{3-109}$$

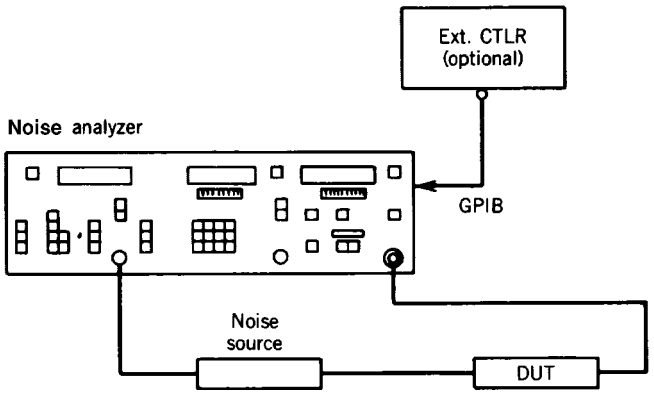


Figure 3-27 Noise figure measurement.

Noise Figure Test Equipment. Figure 3-27 shows the block diagram of a noise test setup. It includes the noise source and the other components. The metering unit has a special detector that is linear and over a certain dynamic range measures linear power. The tunable receiver covers a wide frequency range (e.g., 10–1800 MHz) and controls the noise source. The receiver is a double-conversion superheterodyne configuration with sufficient image rejection to avoid double-sideband noise measurements that would give the wrong results.

These receivers are microprocessor controlled and the measurement is a two-step procedure. The first is a calibration step that measures the noise figure of the receiver system and a reference power level. Then the device under test (DUT) is inserted and the system noise figure and total output power are measured. The noise factor is calculated by

$$F_1 = F_{\text{system}} - \frac{F_2 - 1}{G_1} \tag{3-110}$$

and the gain is given by the change in output power from the reference level [16]. The noise of the system is calculated by measuring the total noise power with the noise source on and off. With the excess noise ratio (ENR) known [16],

$$F_{\text{system}} = \frac{\text{ENR}}{Y - 1} \tag{3-111}$$

The noise bandwidth is usually set by the bandwidth of the receiver, which is assumed to be constant over the linear range. The ENR of the noise source is given by

$$\text{ENR} = \frac{T_{\text{hot}}}{T_{\text{cold}}} - 1 \tag{3-112}$$

where T_{cold} is usually room temperature (290 K). This ENR number is about 15 dB for noise sources with a 15-dB pad and 5 dB for noise sources with a 25-dB pad. Since both gain and noise were stored in the initial calibration, a noise/gain sweep can be performed.

For frequencies above 1800 MHz, we can extend the range with the help of the external signal generators, as shown in Figure 3-28. As shown, a filter ahead of the external mixer reduces the noise energy in the image band. If the DUT has a very broad frequency range

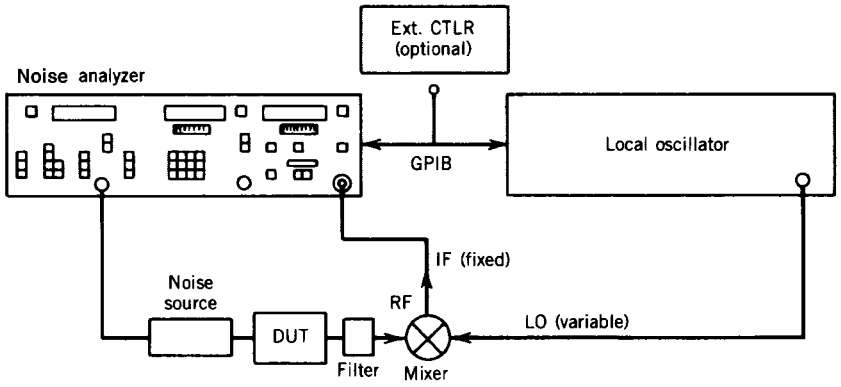


Figure 3-28 Single-sideband (SSB) noise figure measurements using an external mixer.

and has flat gain and noise over that range, a double-sideband (DSB) measurement is possible, with the image-rejection filter removed. However, a single-sideband (SSB) measurement is always more accurate [17].

How To Determine the Noise Parameters. The noise figure of a linear two-port network as a function of source admittance may be represented by

$$F = F_{\min} + \frac{R_n}{G_G}\left[(G_{0n} - G_G)^2 + (B_{0n} - B_G)^2\right] \tag{3-113}$$

where $G_G + jB_G$ = generator admittance presented to the input of the two-port
$G_{0n} + jB_{0n}$ = generator admittance at which optimum noise figure occurs
R_n = empirical constant relating sensitivity of the noise figure to generator admittance, with dimensions of resistance

It may be noted that for an arbitrary noise figure measurement with a known generator admittance, Eq. (3-113) has four unknowns—$F_{\min}$, R_n, G_{0n}, and B_{0n}. By choosing four known values of generator admittance, a set of four linear equations is formed and solutions of the four unknowns can be found [18, 19]. Eq. (3-113) may be transformed to

$$F = F_{\min} + \frac{R_n|Y_{0n}|^2}{G_G} - 2R_nG_{0n} + \frac{R_n|Y_G|^2}{G_G} - 2R_nB_{0n}\frac{B_G}{G_G} \tag{3-114}$$

or

$$F = F_{\min} + \frac{R_n}{G_G}|Y_G - Y_{0n}|^2 \tag{3-115}$$

Let

$$X_1 = F_{\min} - 2R_nG_{0n}$$

$$X_2 = R_n|Y_{0n}|^2$$

$$X_3 = R_n$$

$$X_4 = R_nB_{0n}$$

Then the generalized equation may be written

$$F_i = X_1 + \frac{1}{G_{si}}X_2 + \frac{|Y_{si}|^2}{G_{si}}X_3 - 2\frac{G_{si}}{B_{si}}X_4 \tag{3-116}$$

or, in matrix form,

$$[F] = [A]\,[X] \tag{3-117}$$

and the solution becomes

$$[X] = [A]^{-1}\,[F] \tag{3-118}$$

These parameters completely characterize the noise behavior of the linear two-port network. Direct measurement of these noise parameters by this method would be possible only if the receiver on the output of the two-port were noiseless and insensitive to its input admittance. In actual practice, the receiver itself behaves as a noisy two-port network and can be characterized in the same manner. What is actually being measured is the system noise figure of the two-port and the receiver. Thus, it becomes apparent that to do a complete two-port noise characterization, the gain of the two-port must be measured [20]. In addition, any losses in the input-matching networks must be carefully accounted for, because they add directly to the measured noise figure reading [21].

3-1-4 Linearity

Consistent with the elaborate discussion of linearity questions above, here is a quick refresher on what must be considered in amplifiers.

Dynamic Range, Compression, and IMD. An amplifier's dynamic range is related to the minimum discernible signal and the 1-dB compression point according to the equation

$$DR_n = \frac{(n-1)(IP_{n(\text{in})} - MDS)}{n} \tag{3-119}$$

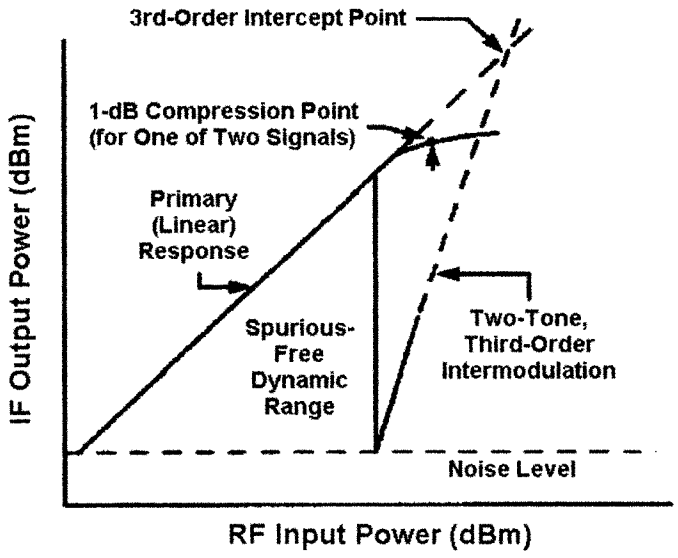


Figure 3-29 Amplifier linearity evaluation, including compression and two-tone IMD dynamic range. $P_{-1\text{dB}}$ for a single-tone cannot be read directly from this graph because the values shown are the result of two equal-power tones.

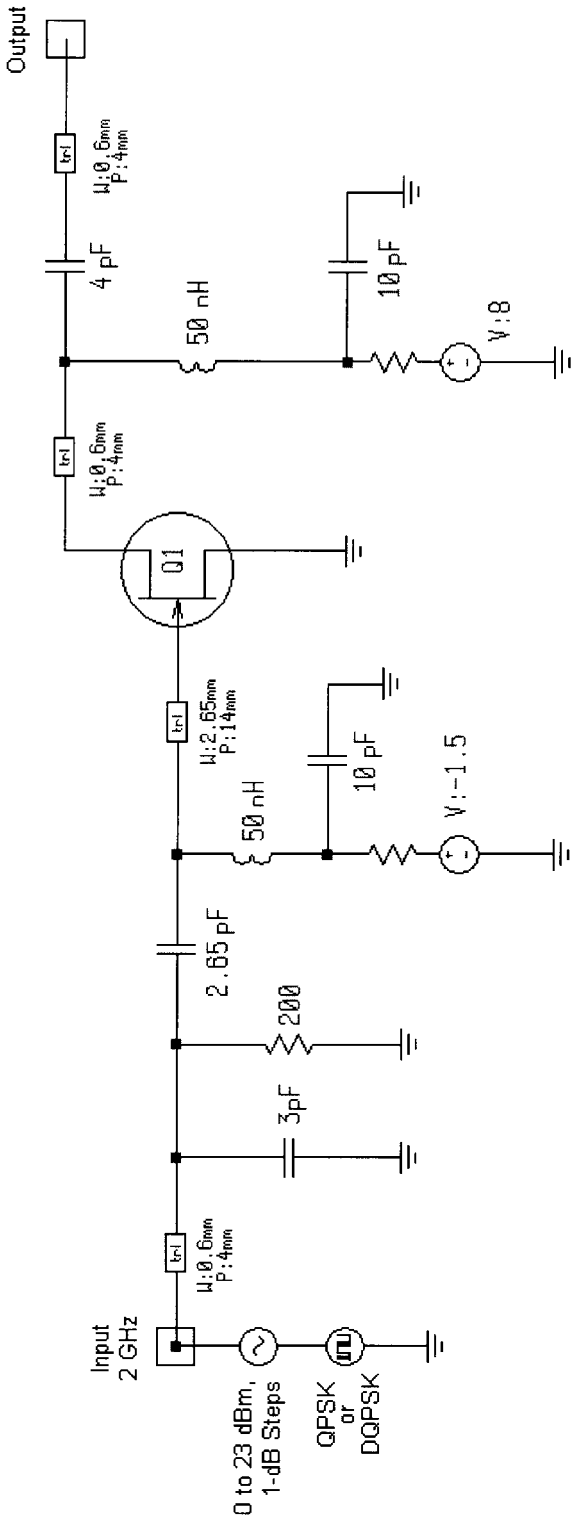


Figure 3-30 The power amplifier schematic.

where DR is the dynamic range in decibels, n is the order, $IP_{(in)}$ is the input intercept power in dBm, and MDS is the minimum detectable signal power in dBm. Once the 1-dB compression point is known, it is a fair assumption that the third-order intercept point, expressed in dBm, is 10 dB above this.

The intermodulation distortion gives the quality of an amplifier to withstand multiple signals without generating large intermodulation products or when using amplitude-modulated signals (cross-modulation). Figure 3-29 provides a quick reminder of IMD issues.

In some amplifiers, we would like to have control over the gain. While the better way to control gain while obtaining high linearity is a PIN-diode attenuator, special AGC circuits have become part of the amplifiers. Figure 3-48 shows a circuit of the SL610, made by Plessey, which has an AGC input provision. The noise figure of the PIN-diode attenuator is essentially equal to its attenuation. The AGC in multistage amplifiers is distributed over several stages and there is a lesser correlation between noise and AGC than in purely passive attenuators.

Special Case of Linearity Requirements for Digital Modulation. For cellular telephone systems we use digital modulation formats, such as QPSK and π/4-DQPSK, that combine phase and amplitude modulation. Amplifiers handling such signals must be carefully characterized and designed if adequate amplitude and phase linearity are to be maintained.

Examples of Power Amplifiers: Looking into the Effects of Distortion. This example presents modulation analysis using QPSK in a 2-GHz power amplifier (Figure 3-30). The simple amplifier was designed using the electronic Smith Tool of Serenade to determine the matching network for a narrowband design. The example demonstrates setup of modulation analysis and available results.

Siemens among others provides integrated amplifiers based on the concept depicted in Figure 3-30, the most popular ones being the CGY96 and CGY94. For the purpose of measurements, these amplifiers were actually put on a breadboard. Figure 3-31 shows the big brother, the CGY96. It is a GaAs MMIC intended as a power amplifier for GSM Class 4 phones with 3.2 W (35 dBm) at 3.5 V at an overall power added efficiency of 50%. Table 3-5 shows its electrical characteristics.

For the purpose of our experiments, we considered the CGY94 because it requires less external peripheral circuits. Table 3-6 shows its electrical characteristics. For those of us interested in seeing part of the internal circuit, Figure 3-32 shows the chip. This two-stage amplifier uses an interesting internal feedback scheme. Its schematic is shown in Figure 3-33.

For the purpose of looking at the waveforms, we will continue with the single-stage amplifier (Figure 3-30) and then compare the measured versus predicted performance of this CGY94-based amplifier.

A modulation source is connected to the RF source at the input port of the amplifier. The RF source specifies the carrier power (or voltage or current) and the modulation source specifies the modulation format and properties of the modulated signal.

A brief review of the modulation source used in this example indicates the following properties:

Property	Value	Description
NB	128	Number of bits
BR	1.2288E6	Bit rate
N	8	Number of samples per bit
BW1	590.E3	One-sided bandwidth of the main channel in the PIB or P1IB calculation
FS1	740.E3	Start baseband frequency of the fist adjacent channel
BW2	590.E3	One-sided bandwidth of the main channel in the P2IB calculation
FS2	1990.E3	Start baseband frequency of the second adjacent channel
BW3	590.E3	One-sided bandwidth of the main channel in the P3IB calculation
FS3	3240.E3	Start baseband frequency of the third adjacent channel
M	4	Order of the signal space; M = 4 is QPSK
DLY	0.0	Fractional bit delay
IASC	1.0	I channel amplitude scale; multiplier for the I waveform to model amplitude imbalance
QASC	1.0	Q channel amplitude scale; multiplier for the Q waveform to model amplitude imbalance
LPF	3	Baseband low-pass filter; 3 = root-raised cosine filter
LPFC	665.E3	–3 dB cutoff frequency of LPF
LPFN	3	Number of resonators if LPF = 1
LPFR	0.35	Roll-off factor if LPF = 3 or LPF = 4
FILE	N/A	File name of user-defined modulated signal file

The modulation source is a QPSK stream with a bit rate of 1.2288 Mbps. A total of 128 bits will be analyzed, and each bit is sampled eight times to construct an accurate analog waveform. The BW and FS parameters are defined here to specify the main (in-band) channel and adjacent channel for ACPR calculations. Imbalance of the I and Q amplitudes and phases can be described by the IASC, QASC, and DLY parameters, if desired. The modulation signal

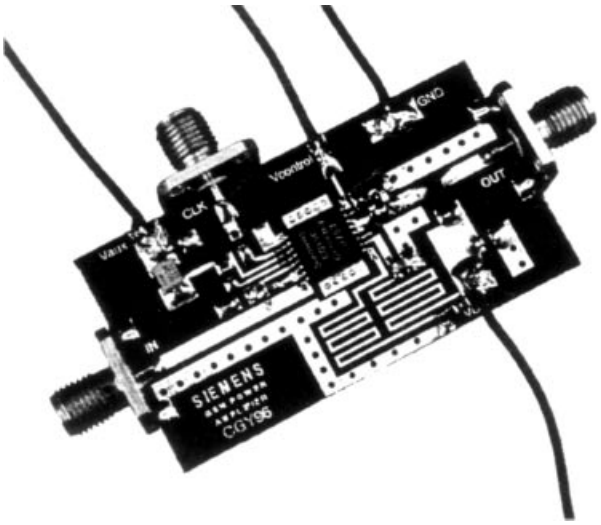


Figure 3-31 CGY96 GaAs MMIC power amplifier.

Table 3-5 CGY96 electrical characteristics

(T_A = 25 °C, V_{neg} = 5 V, $V_{control}$ = 2.2 V; duty cycle 12.5%, t_{on} = 577 µs)

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
Frequency range	f	880	—	915	MHz
Supply current P_{in} = 0 dBm	I_D	—	1.8	—	A
Supply current negative voltage generator V_{aux} = 3.5 V	I_{AUX}	—	10	—	mA
Gain (small signal)	G	—	40	—	dB
Power gain P_{in} = 0 dBm	G_P	—	35	—	dB
Output power P_{in} = 0 dBm, $V_{control}$ = 2.0–2.5 V	P_{OUT}	—	35	—	dBm
Overall power added efficiency P_{in} = 0 dBm	η	—	50	—	%
Dynamic range output power $V_{control}$ = 0.2–2.2 V		—	80	—	dB
Harmonics P_{in} = 0 dBm	$H(2f_0)$	—	–40	—	dBc
	$H(3f_0)$	—	–43	—	dBc
	$H(4f_0)$	—	–44	—	dBc
Noise power in RX (935–960 MHz) P_{in} = 0 dBm, P_{out} = 35 dBm, 100 kHz RBW	N_{RX}	—	–81	—	dBm
Stability *all spurious outputs* < –60 dBc, *VSWR load, all phase angles*		—	10:1	—	—
Input VSWR		—	1.7:1	—	—

Table 3-6 CGY94 electrical characteristics

T_A = 25 °C, f = 0.9 GHz, $Z_S = Z_L$ = 50 Ω, V_D = 3.6 V, V_G = 4 V, VTR pin connected to ground, unless otherwise specified; pulsed with a duty cycle of 10%, ton = 0.33 ms)

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
Supply current V_D = 3.0 V P_{in} = 10 dBm	I_{DD}	—	1.18	—	A
Negative supply current *(normal operation)*	I_G	—	2	—	mA
Shut-off current *VTR n.c.*	I_D	—	400	—	µA
Negative supply current (*shut off mode, VTR pin n.c.*)	I_G	—	10	—	µA
Gain P_{in} = *5* dBm	G	27.0	29.0	—	dB
Power gain V_D = 3.6 V; P_{in} = 10 dBm	G	22.8	23.6	—	dB
Output power V_D = 3.0 V*;* P_{in} = 10 dBm	P_O	31.5	32.3	—	dBm
Output power V_D = 3.6 V*;* P_{in} = 10 dBm	P_O	32.8	33.6	—	dBm
Output power V_D = 5 *V;* P_{in} = 10 dBm	P_O	34.5	35.5	—	dBm
Overall power added efficiency V_D = 3.0 V; P_{in} = 10 dBm	η	43	48	—	%
Overall power added efficiency V_D = 3.6 V; P_{in} = 10 dBm	η	42	47	—	%
Overall power added efficiency V_D = 5 V; P_{in} = 10 dBm	η	41	46	—	%
Harmonics (P_{in} = 10 dBm, *CW*) $2f_0$	—	—	–49	—	dBc
V_D = 3.6 V; (P_{out} = 33.1 dBm) $3f_0$	—	—	–45	—	dBc
Input VSWR V_D = 3.6 V	—	—	1.5:1	2.0:1	—

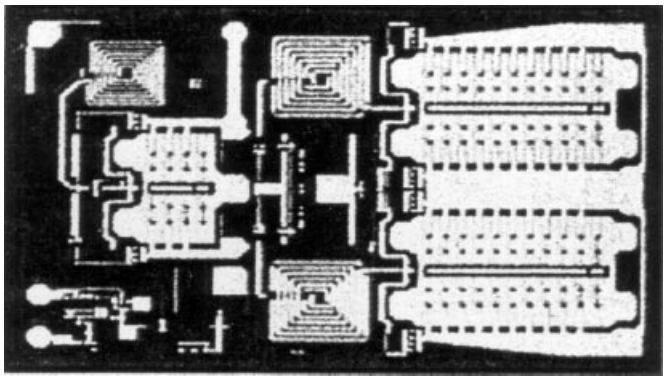


Figure 3-32 The CGY94 amplifier chip.

is often filtered, and several types of filters are available. Here a root-raised cosine filter is used with a cutoff frequency of 665 kHz and a roll-off factor of 0.35.

Analysis. Figure 3-34 shows the compression characteristics of the amplifier as the available RF source power (a sinusoid) is swept. Note that P_{-1dB} is 11.5 dBm referred to the input and 28.5 dBm referred to the output. We will look at modulation characteristics in the linear region of operation at a source power of 0 dBm, and the nonlinear region at a source power of 20 dBm.

Next, we will view the voltage across the transistor to examine its behavior under compression. Figure 3-35 shows the drain–source voltage of the FET as the source power is swept. It is clear from this graph that clipping due to pinchoff and forward-gate conduction

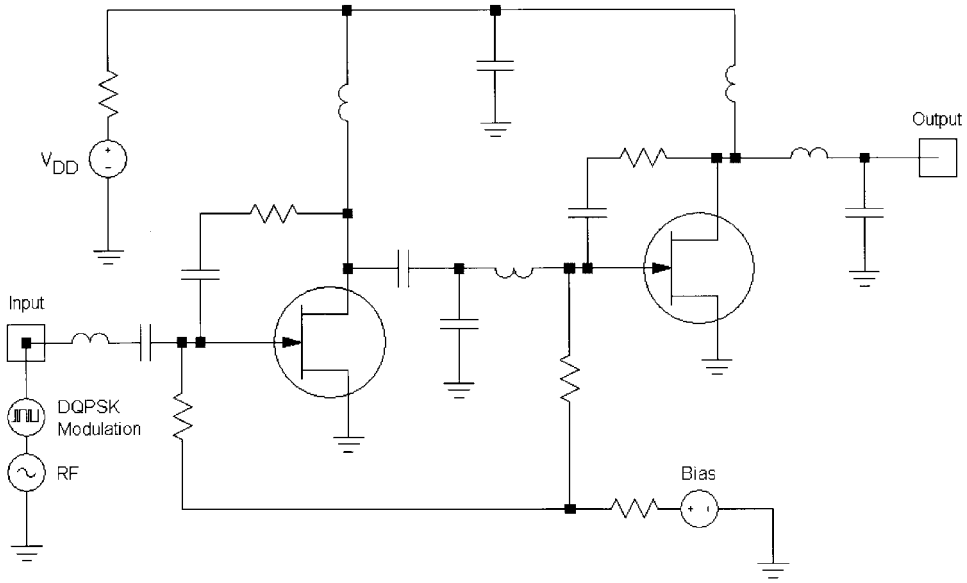


Figure 3-33 Schematic of the CGY94 GaAs MMIC power amplifier.

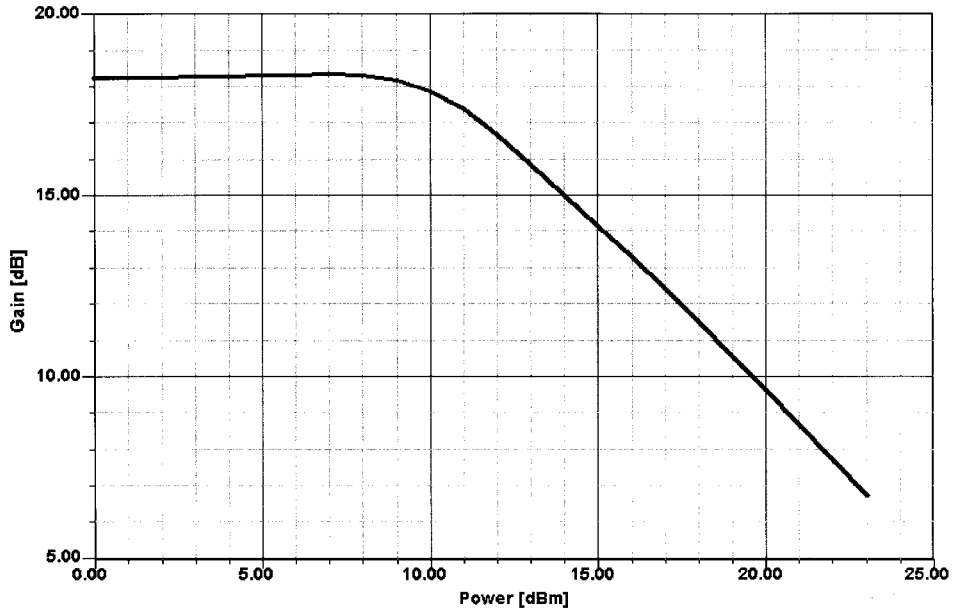


Figure 3-34 Single-tone RF power sweep analysis of the FET amplifier.

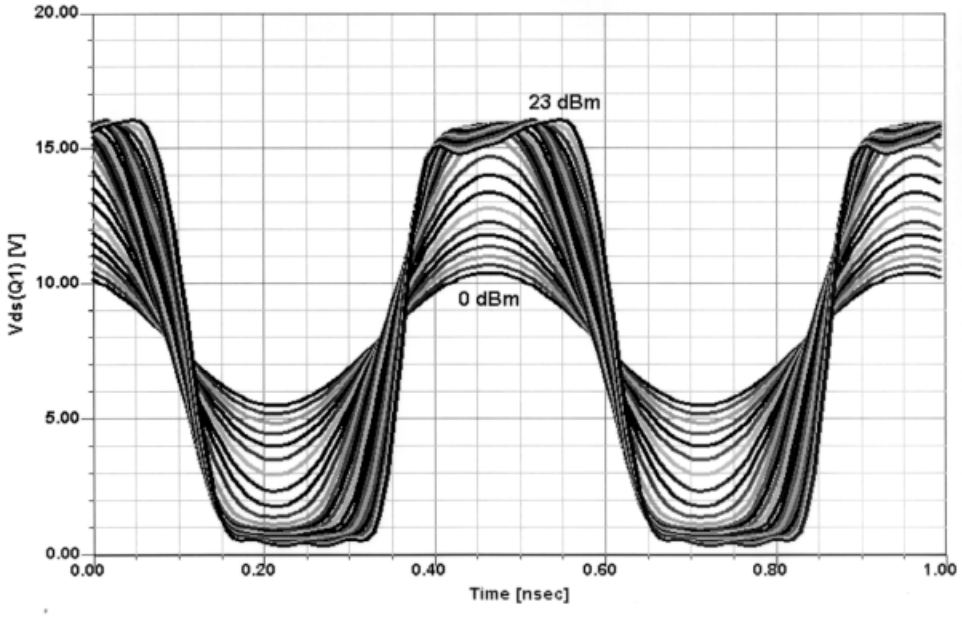


Figure 3-35 V_{DS} of the FET versus RF source power.

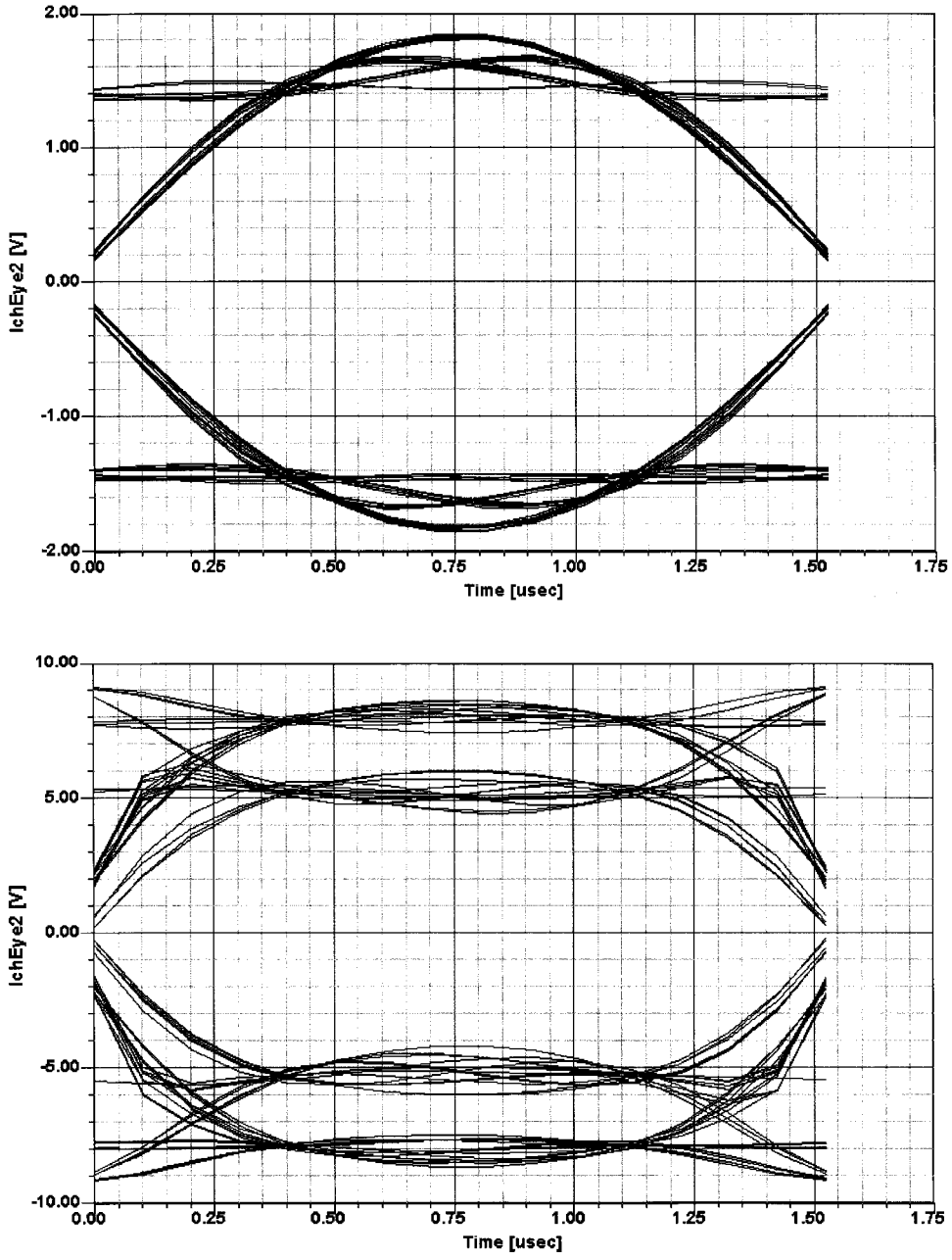


Figure 3-36 Eye diagrams of the I channel for source powers of (a) 0 dBm and (b) 20 dBm.

is limiting the performance of the device. A phase shift of the voltage waveform as power is increased is also apparent. The combination of the power compression and phase shift, or AM to PM conversion, will be used in the modulation analysis to determine the overall distortion of the modulated signal.

The output available from modulation analysis includes:

- I and Q channel waveforms
- Eye and constellation diagrams
- Spectral plots
- ACPR, in-band, and adjacent power

Figure 3-36 shows the eye diagrams for the 128-bit QPSK signal at two RF source powers, 0 and 20 dBm. Operating linearly at 0 dBm input, the amplifier doesn't distort the eye and it remains wide open. Note that in this example the filter bandwidth doesn't produce any intersymbol interference, as evidenced by the open eye. At an input power of 20 dBm, the amplifier compresses the signal, and distortion in the eye is evident.

We can also view the modulation spectrum at these two power levels to investigate the intermodulation distortion and spectral regrowth that takes place. Figure 3-37 shows the spectral plots. The lower trace corresponds to 0 dBm source power and shows almost no regrowth, while the upper trace at 20 dBm source power shows considerable regrowth.

The uncorrupted modulation source can also be shown. This corresponds to the source as a pure voltage without any circuit influence.

Now we evaluate the CGY94 board in a similar fashion. Figure 3-38 shows the waveform used for CAD purposes. Needless to say, such a signal cannot be produced by any signal

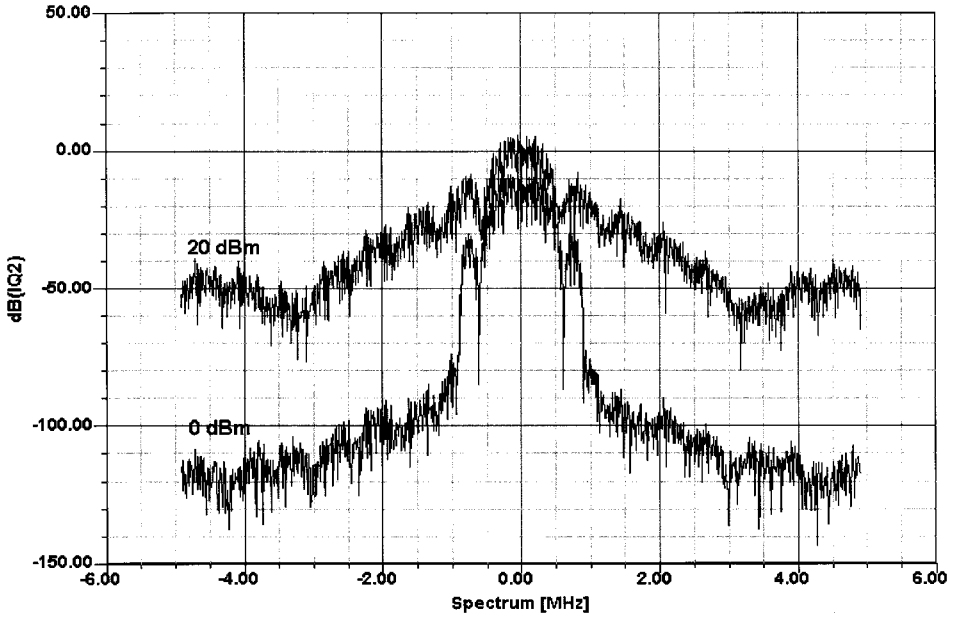


Figure 3-37 Spectrum of the modulation signal at the amplifier output. Source powers are 0 and 20 dBm.

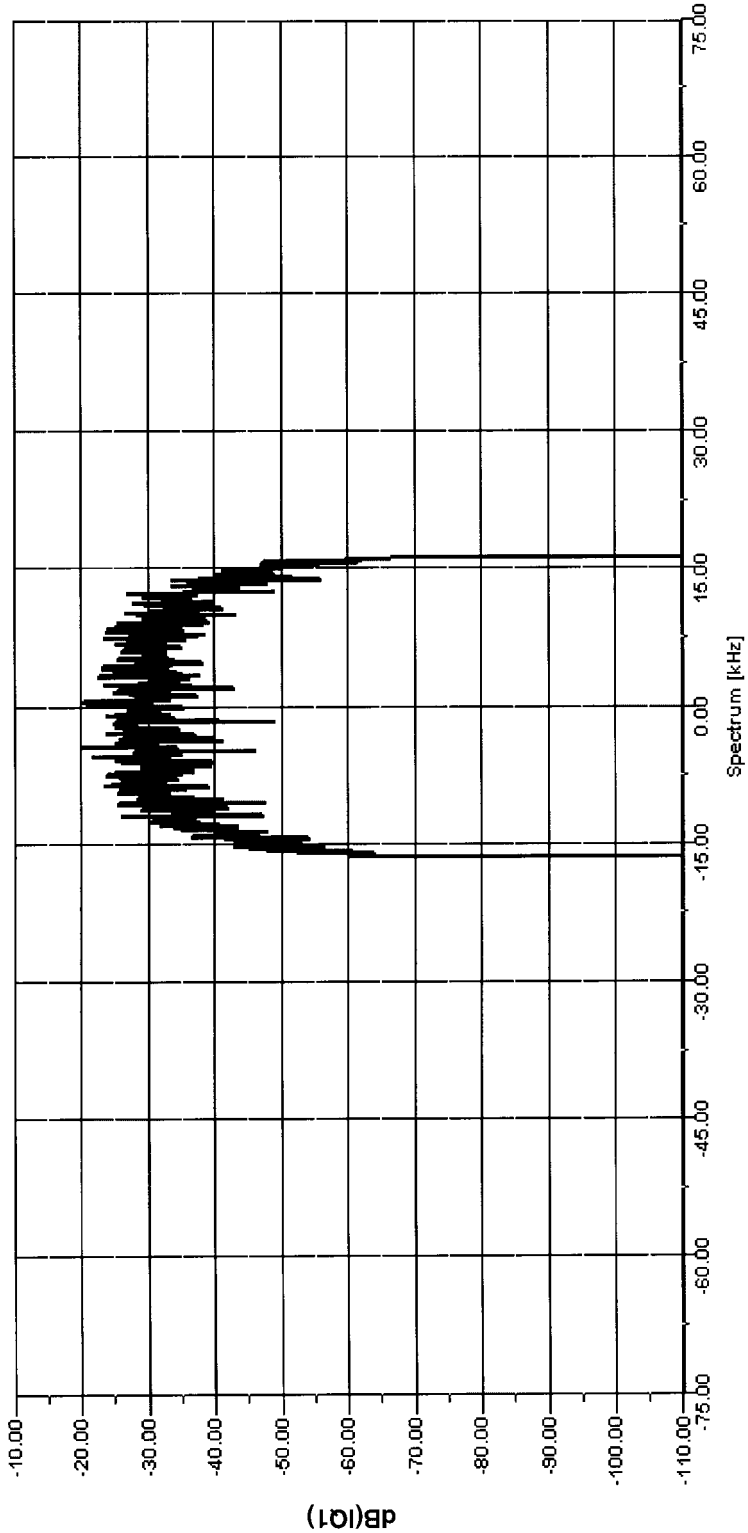


Figure 3-38 Simulated signal used for CAD analysis of the CGY94 amplifier.

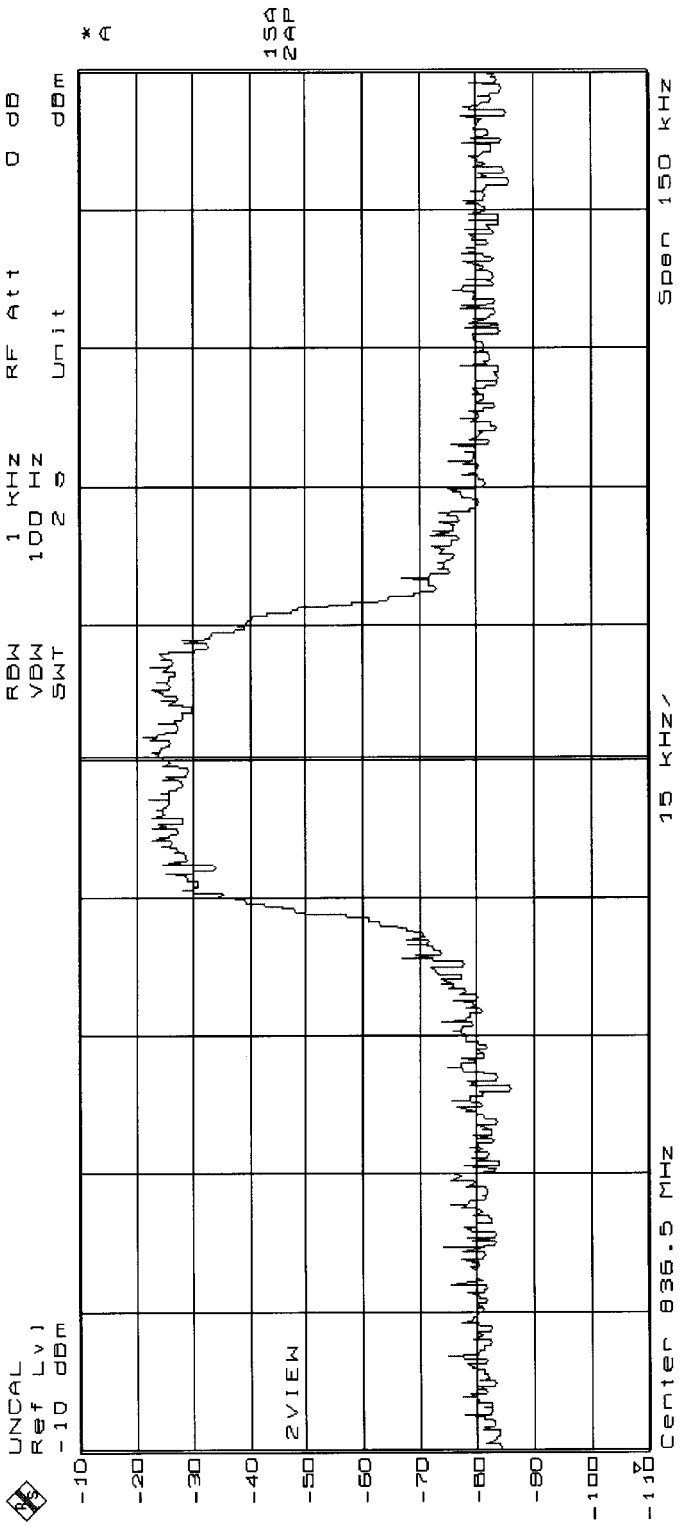


Figure 3-39 Real signal used for testing the actual amplifier.

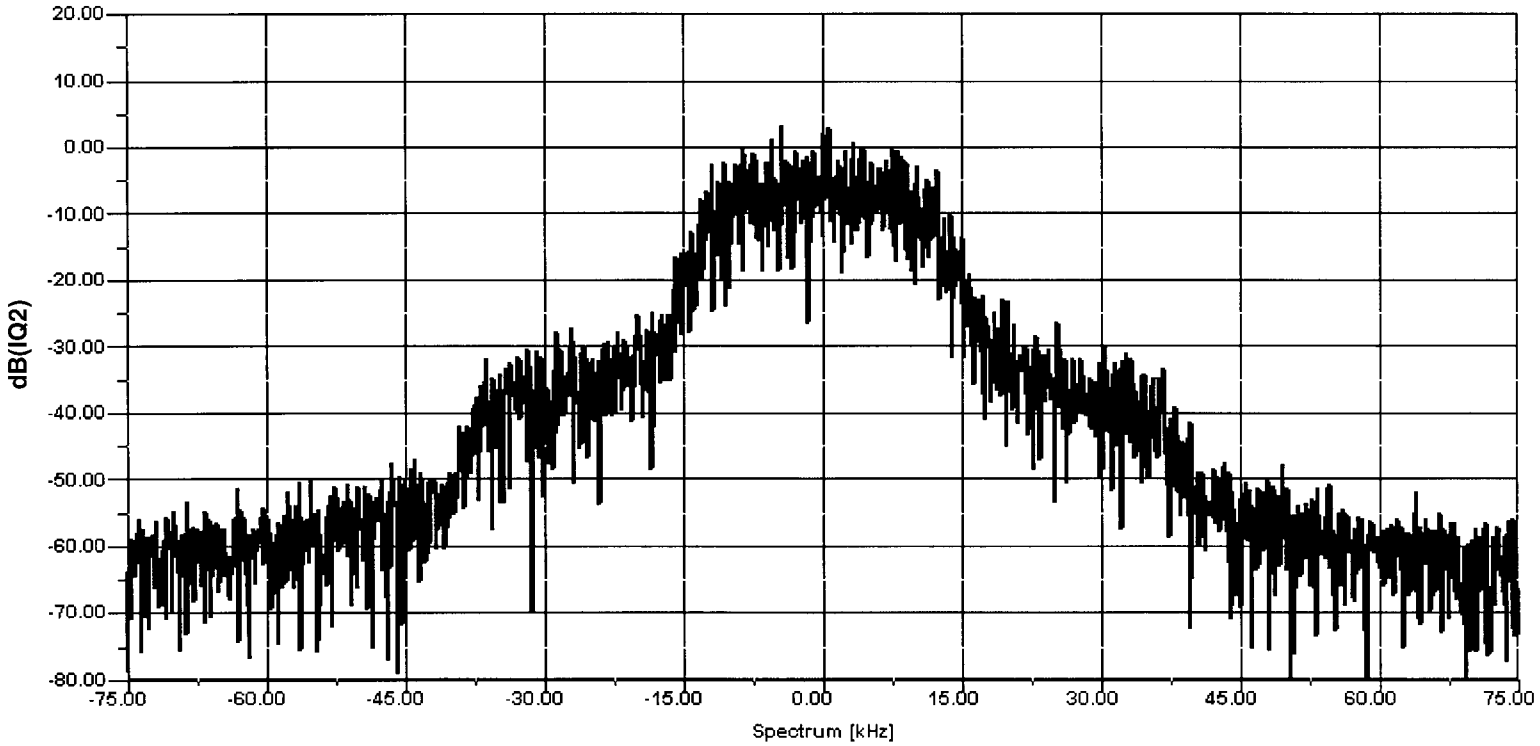


Figure 3-40 Spectral regrowth predicted by the simulation.

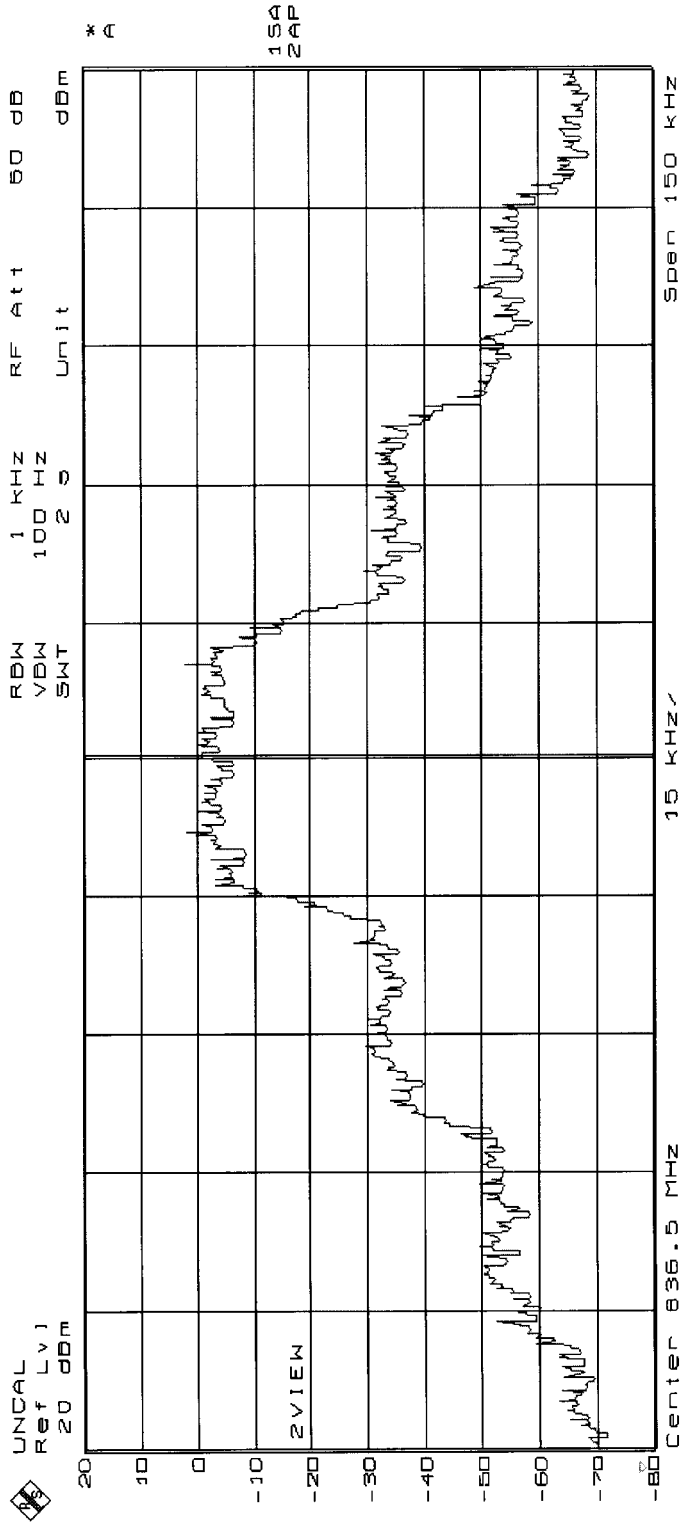


Figure 3-41 Measured spectral regrowth of the CGY94 amplifier.

generator. The one generated by an actual signal generator is shown in Figure 3-39. The wideband noise is around –80 dBm and the input level is about –25 dBm, consistent with the simulation signal. Looking at the output, Figure 3-40 shows the simulated signal with the expected increase in bandwidth. This is due to the nonlinearity of the amplifier and is consistent with Figure 3-37 for the single-stage amplifier. The actual measured output for the CGY94 shown in Figure 3-41 is quite close to its simulated output, indicating that the mathematical approach used for simulating is correct. This is only one example; in reality, many more waveforms can be analyzed and predicted using this method. Figure 3-41 shows essentially three distinct steps, which are sufficiently close to the response shown in the simulation. A dynamic range of 70 dB, while displayed in Figures 3-40 and 3-41, is not really necessary for good-quality transmission, and the discrepancy between the two pictures is in part due to the limited modeling quality of the active device as provided by the company that did the parameter extraction for this transistor.

Another useful view to aid in the understanding of the compression and distortion within the amplifier is to look at the magnitude of the complex waveform. Since the signal is composed of the in-phase and quadrature-phase components it is represented as

$$s(t) = i(t) + j \cdot q(t) \tag{3-120}$$

The time-domain magnitude is written

$$|s(t)| = \sqrt{i(t)^2 + q(t)^2} \tag{3-121}$$

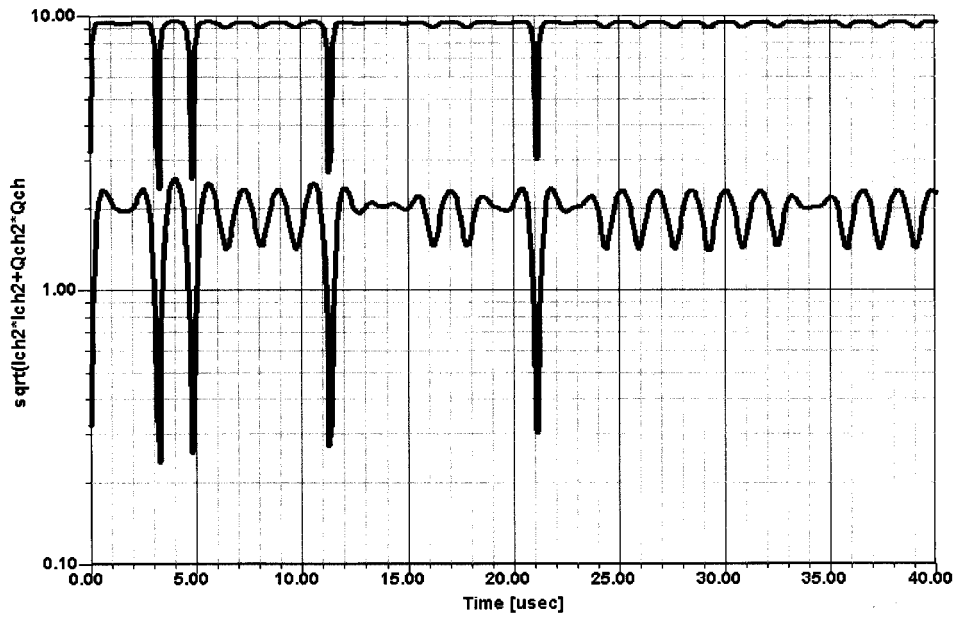


Figure 3-42 Time-domain magnitude of the complex modulation signal at source powers of 0 and 20 dBm.

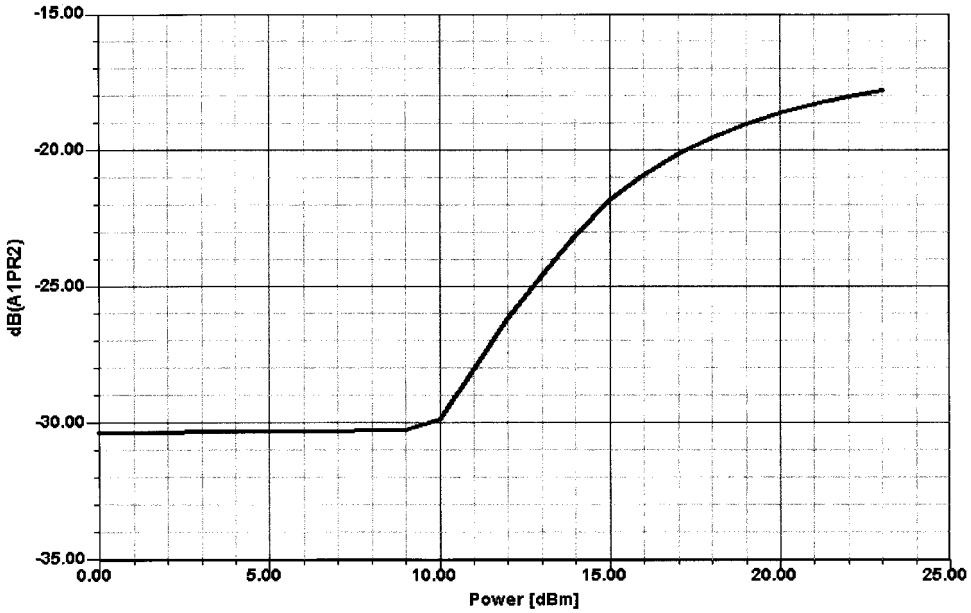


Figure 3-43 Adjacent-channel power ratio as function of RF source power.

Figure 3-42 shows the magnitude at source powers of 0 and 20 dBm. The *X* axis has been rescaled to 40 μs for improved viewing. Note the significant compression of the 20-dBm waveform at the higher signal levels. It is clear from this view that the signal is severely distorted.

We can also compute the adjacent-channel power ratio (ACPR) against the RF power sweep. ACPR is the ratio of the adjacent-channel power to the in-band channel power. The bandwidth (BW) and adjacent-channel start frequency (FS) are used for this calculation. For accurate ACPR calculation, a large number of bits are needed. This example uses 128 bits so the computation time is short, but 512 or more bits should be used for a more accurate computation.

Figure 3-43 shows the ACPR as a function of RF source power. The ACPR is nearly constant (and nonzero due to the gradual skirt of the baseband filter and inherent spillover to the adjacent channel) up to the P_{-1dB} compression point . As the intermodulation products spill into the adjacent channel and spectral regrowth occurs, the ACPR degrades.

π/4-DQPSK Circuit Analysis. Next, we will examine the same amplifier using a π/4-DQPSK modulation source. The task is identical to the previous one except for the source and the number of bits. The number of bits for this project has been increased from 128 to 256. Figure 3-44 shows the eye diagrams from this modulation format, at source powers of 0 and 20 dBm. The distortion and related intersymbol interference are clearly evident.

The constellation plot at 0 dBm is shown in Figure 3-45.

It may be interesting to take a look at the different waveforms currently in use. Figure 3-46 shows a spectrum analyzer picture of the most popular ones.

For reasons of time-division management, the European standard, GSM, which is now offered by Omnipoint at the 1.8-GHz band in the United States, uses gating to reduce the actual bandwidth of the spectrum. Figure 3-47 compares the GSM signal with and without gating.

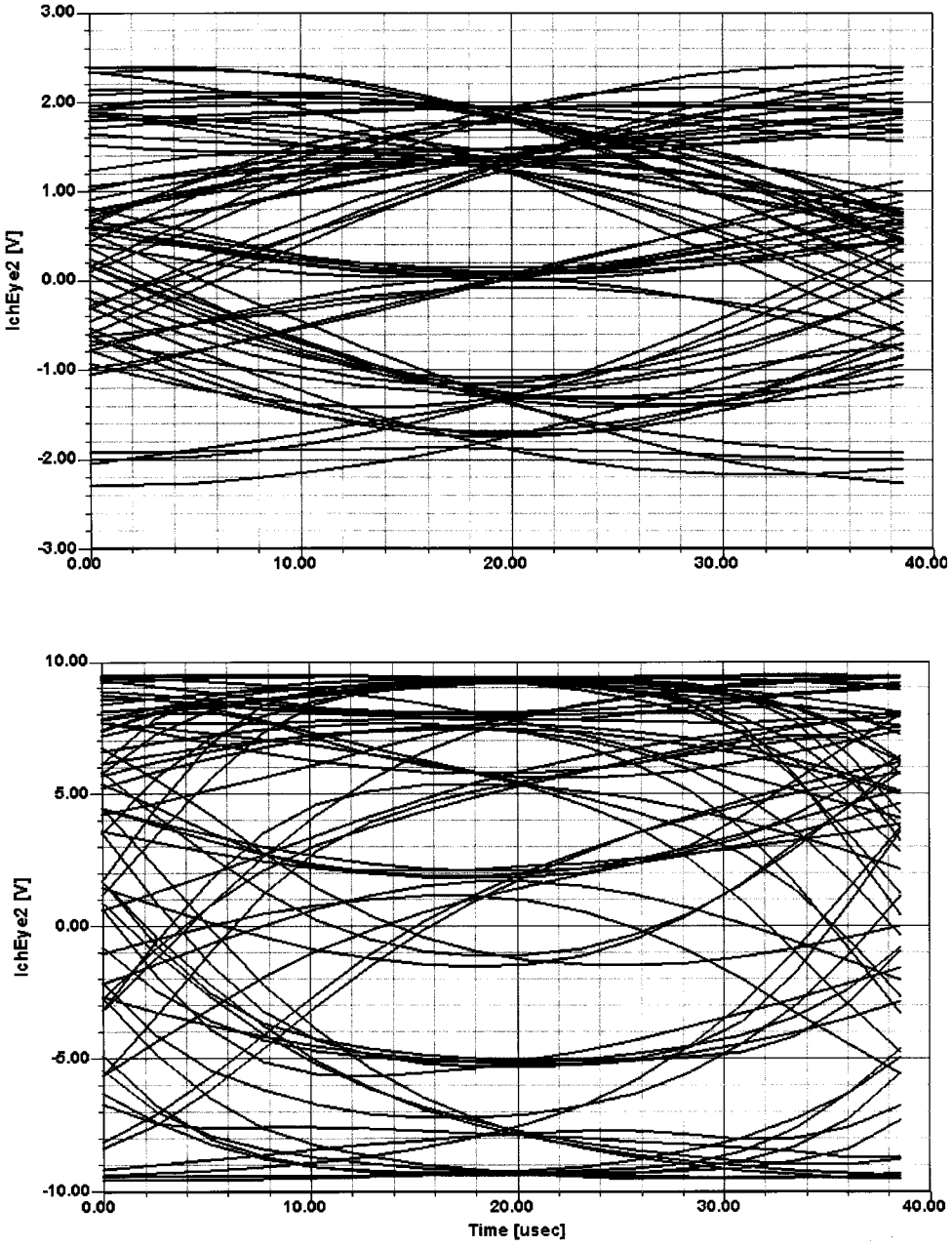


Figure 3-44 Eye diagrams for $\pi/4$-DQPSK at source powers of (a) 0 dBm and (b) 20 dBm. The distortion and related intersymbol interference are clearly visible.

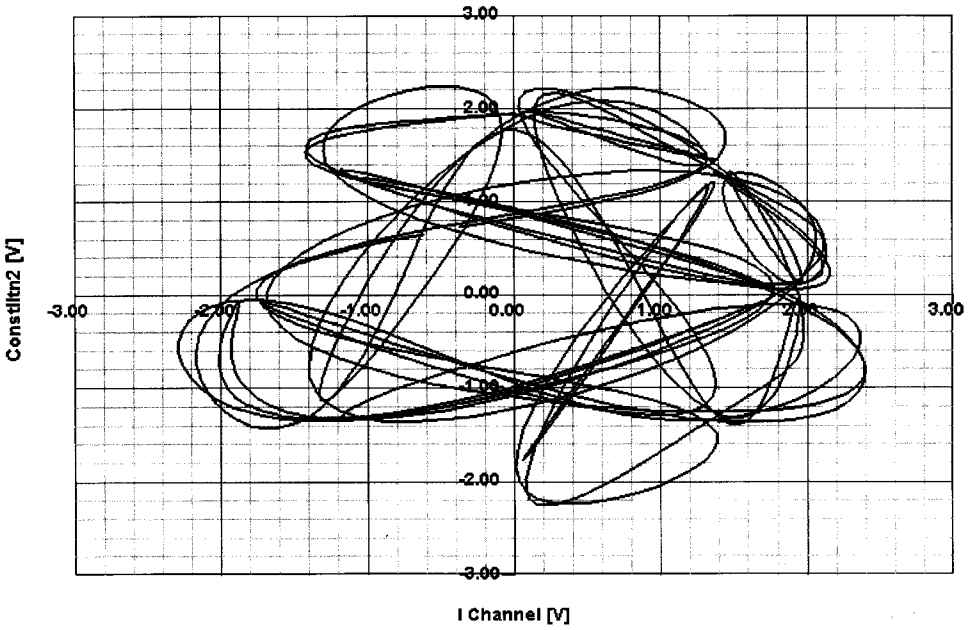


Figure 3-45 Constellation diagram for π/4-DQPSK at Port 2 and 0-dBm source power.

3-1-5 AGC

Many wireless systems consist of an up conversion or down conversion using an intermediate frequency, typically 10% or less of the receiving frequency (the FM broadcast IF has been set at 10.7 MHz for the same reason).

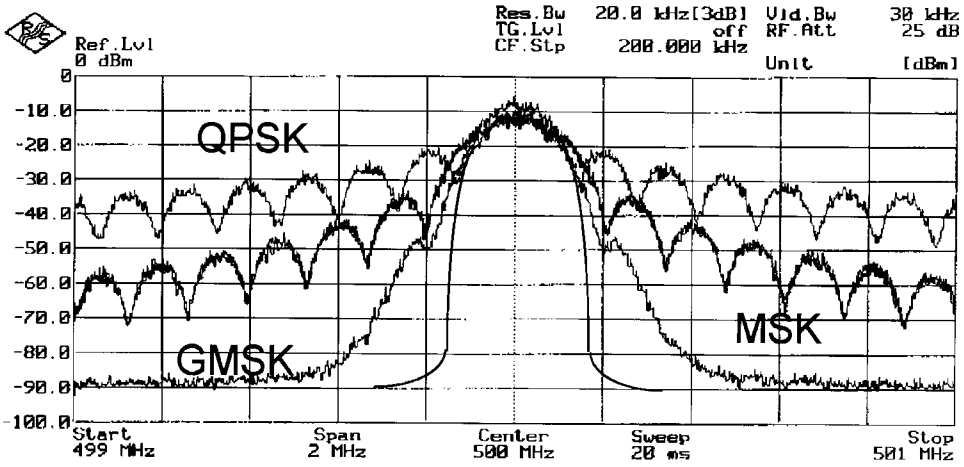


Figure 3-46 Spectrogram showing the characteristics of MSK, GMSK, and QPSK.

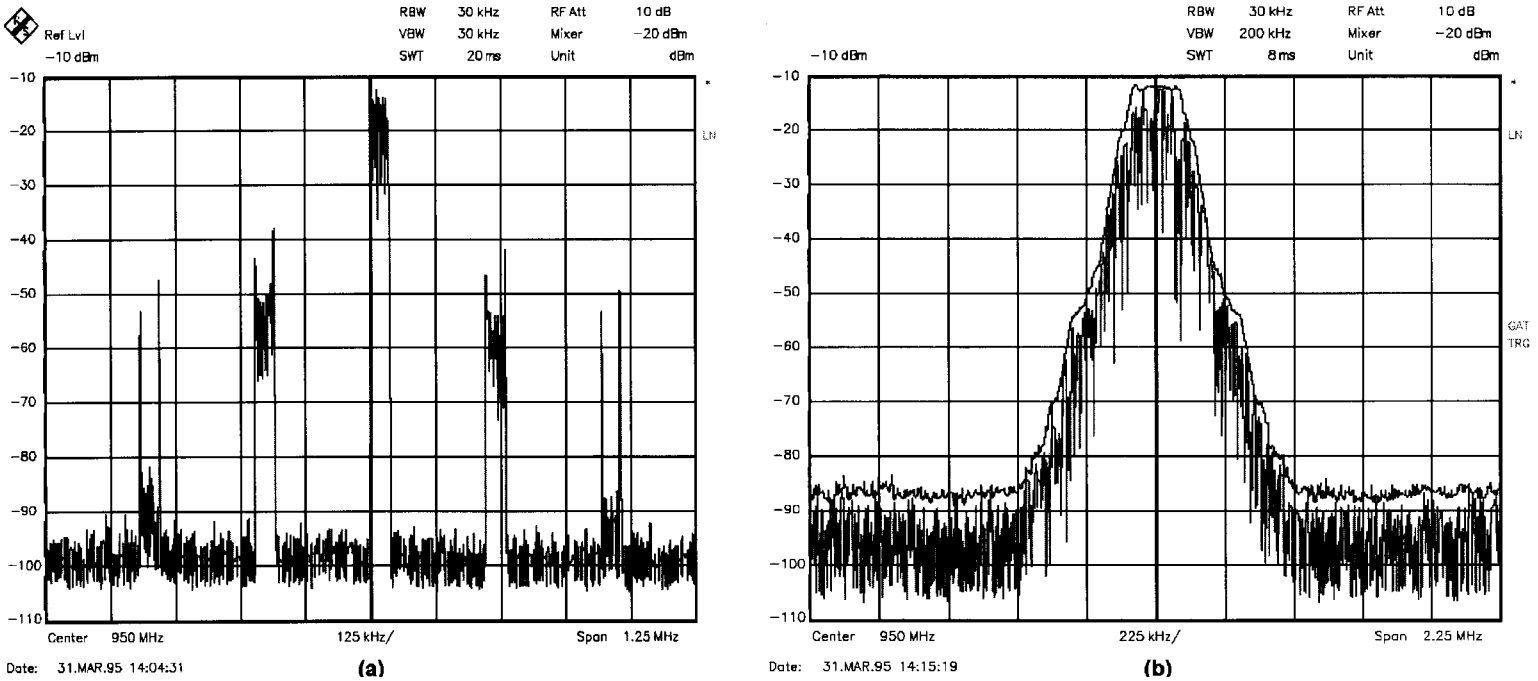


Figure 3-47 Pulsed GSM signal (a) without gating and (b) with gating.

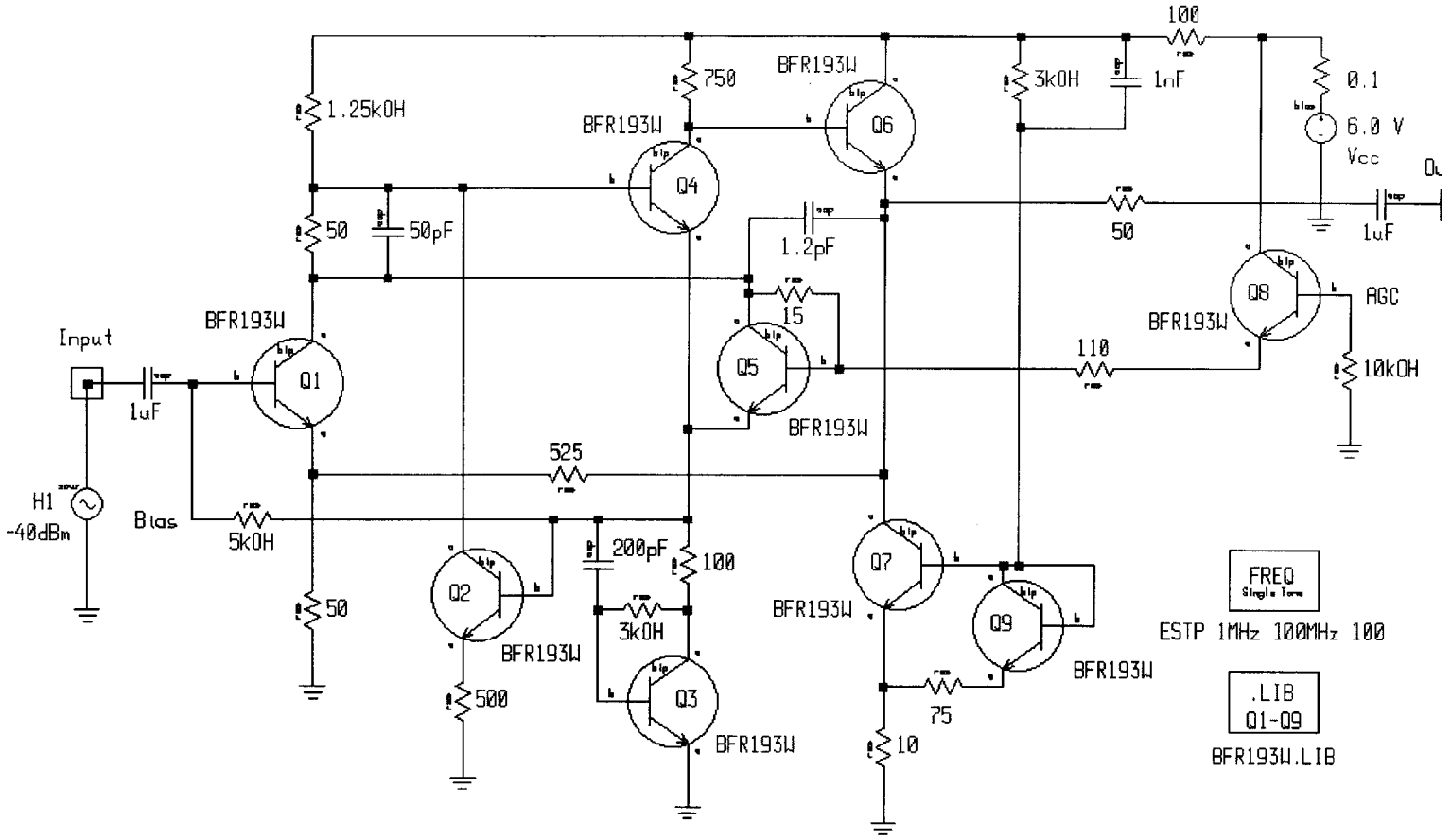


Figure 3-48 Schematic of the Plessey SL610 wideband amplifier with AGC. The 75-Ω resistor in the emitter of Q9 had to be added to obtain the appropriate dc current.

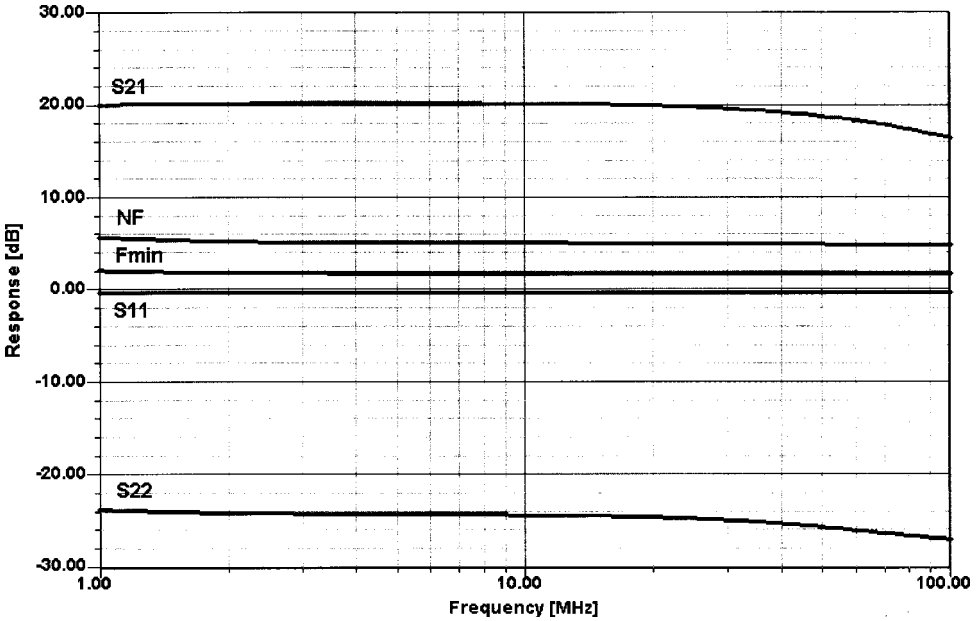


Figure 3-49 Simulated frequency response of the Plessey SL610 wideband amplifier IC with AGC = 0 V.

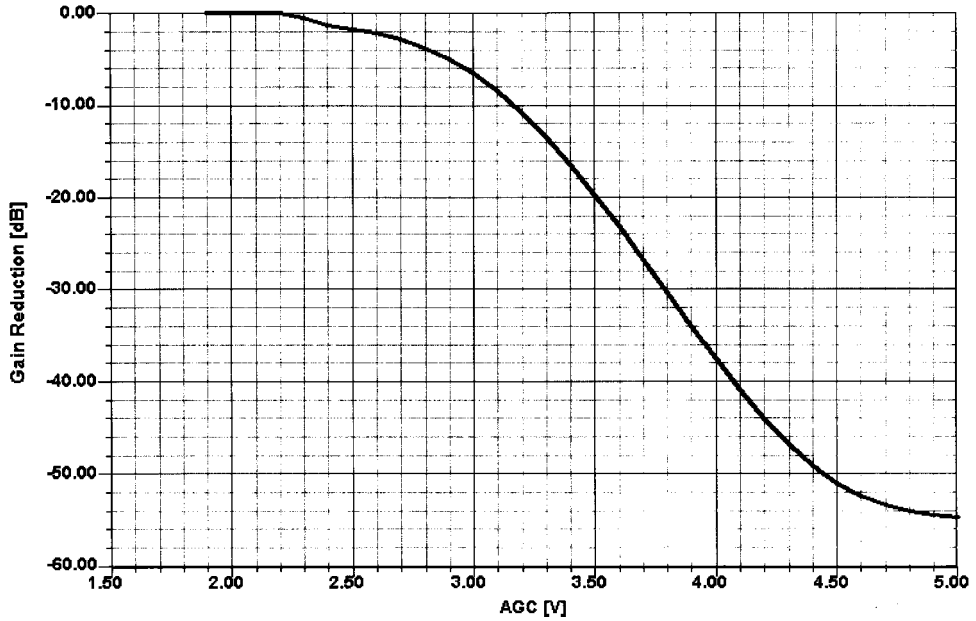


Figure 3-50 Simulated gain reduction versus AGC of the Plessey SL610 wideband amplifier IC.

Table 3-7 CGY121A features

- Variable gain amplifier (MMIC amplifier) for mobile communication
- Typical gain control range over 50 dB
- Positive control voltage
- 50-Ω input and output matched
- Low power consumption
- Operating voltage range: 2.7–6 V
- Frequency range 800 MHz to 2.5 GHz

We therefore will show two amplifiers with AGC and available data. The first one is a wideband amplifier type SL610, which has been manufactured by Plessey; its current fate is unknown. However, since we put the circuit into the simulator it provides significant insight into the operation of such AGC stages. Figure 3-48 shows the actual schematic of the SL610 entered in Ansoft's Serenade Design Environment for simulation.

The AGC action is derived from Q8 through Q5, which, together with Q4, forms a differential amplifier. Each consecutive stage (Q1, Q4, Q6) has increasing current, totaling about 20 mA; the last transistor takes about 13 mA. Based on the feedback surrounding the transistors, the actual transistor characteristics are less important. We have shown this by putting much higher f_T devices into the circuit than were available for the original design (20 years ago). The circuit is still interesting today because of its good large-signal-handling capability.

Table 3-8 CGY121A electrical characteristics at 900 MHz and 1.8 GHz

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
(T_A 25 °C, f = 900 MHz, V_g = –4 V, $RS = RL$ = 50 Ω unless otherwise specified)					
Power gain V_d = 3 V; I = 45 mA; V_{con} = 3 V	G	17	19	—	dB
Input return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{in}	—	11	—	dB
Output return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{out}	—	10	—	dB
Gain control range V_{con} = 3 V ... 0 V; V_d = 3 V; I = 45 mA	dG	48	53	—	dB
1-dB gain compression V_d = 3 V; I = 45 mA; V_{con} = 3 V	P_{1dB}	—	14	—	dBm
(T_A 25 °C, f = 1800 MHz, V_g = –4 V, $RS = RL$ = 50 Ω unless otherwise specified)					
Power gain V_d = 3 V; I = 45 mA; V_{con} = 3 V	G	15.5	17.5	—	dB
Input return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{in}	—	10	—	dB
Output return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{out}	—	8.5	—	dB
Gain control range V_{con} = 3 V ... 0 V; V_d = 3 V; I = 45 mA	dG	48	53	—	dB
1-dB gain compression V_d = 3 V; I = 45 mA; V_{con} = 3 V	P_{1dB}	—	14	—	dBm

Figure 3-49 shows the amplifier's frequency response which is extremely close to the measured data. The good simulation also can be seen from comparing the published AGC curve with the one we found in the simulation. The bias of the first transistor remains essentially constant, resulting in a noise figure more independent of the AGC than the PIN-diode attenuator, and also stabilizes the input impedance as a function of AGC. Figure 3-50 shows the SL610's predicted AGC response.

The same requirement is applicable for the wireless frequency range. Tables 3-7 and 3-8 and Figures 3-51 through 3-54 present specifications for the recently introduced CGY121A GaAs MMIC by Infineon Technologies AG, formerly the Semiconductor Group part of Siemens AG.

3-1-6 Bias and Power Voltage and Current (Power Consumption)

Biasing of transistors operating in Class A or B follows some standard rules. The only deviation from this tends to be when the operating voltage is very low, below 3 V, because then the luxury of using voltage drops for good dc stability as a function of temperature disappears. (We will discuss low-voltage design in detail in Section 3-2-6.) In our simple example (Figure 3-55), we are operating from a 12-V power supply, and the first rule of thumb is to make sure that the voltage between the emitter and ground is at least 0.7 V. This is necessary to compensate the base–emitter junction threshold voltage, which decreases by about 2 mV/°C as the temperature increases. A larger voltage drop in the emitter–ground connection greatly reduces the influence. In this amplifier the emitter–ground path consists of an emitter unbypassed resistor of 10 Ω in series with a bypassed 501-Ω resistor, resulting in an overall voltage drop of 2.6. We could reduce the supply voltage by about 1.9 V at the collector side and still maintain a safety margin of 0.7 V at the emitter–ground connection. The current 2.6 V, so to speak, is overkill.

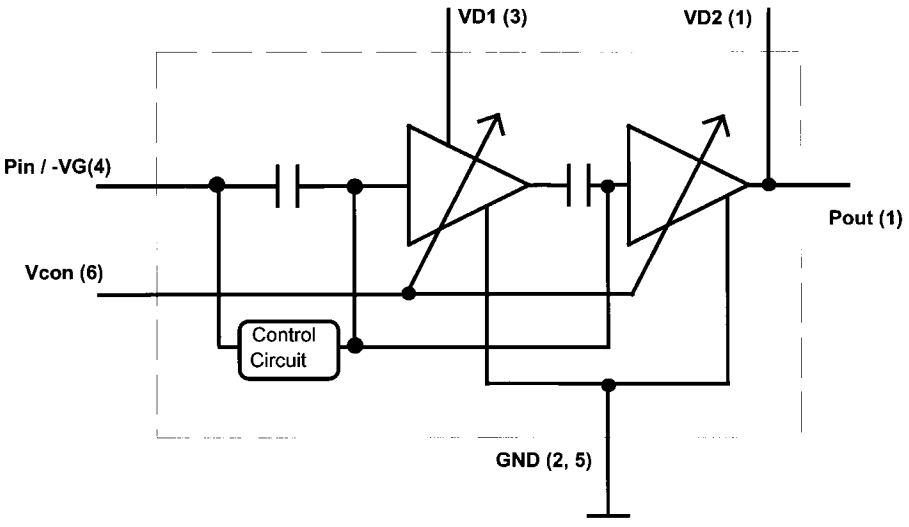
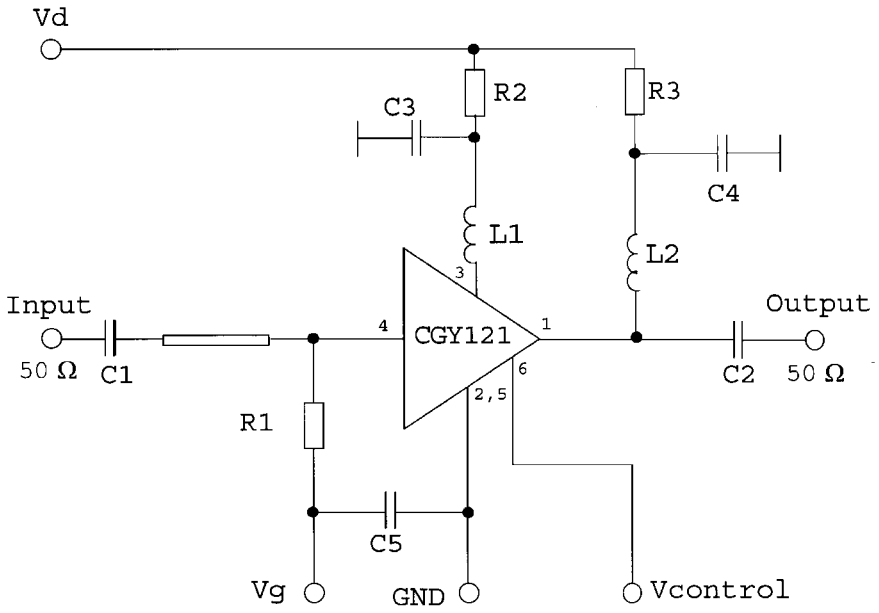


Figure 3-51 Functional block diagram of the gain-controllable CGY121A GaAs MMIC. Gain control is achieved by applying 0–3 V dc to the Vcon pin.



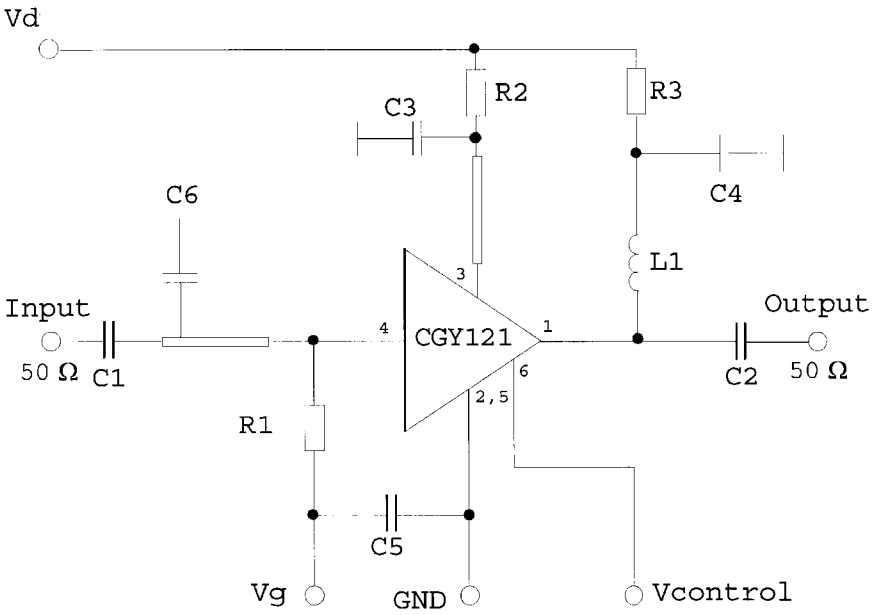
Parts List

Frequency	900 MHz
C1, C2 (Siemens Size 0603)	22 pF
C3, C4 (Siemens Size 0603)	100 nF
C5 (Siemens Size 0603)	47 nF
L1 (Coilcraft 0805CS-150XKBC)	15 nH
L2 (Coilcraft 0805CS-270XMBC)	27 nH
R1 (Siemens B 54102-A1271-J60)	270 Ω
R2 (Siemens B 54102-A1120-J60)	12 Ω
R3	6.8 Ω

Figure 3-52 CGY121A 900-MHz application circuit and parts list.

On the collector side, for an *RC* type of amplifier, it is not a bad idea to set the collector voltage at about half the supply, which allows the RF swing to be half the supply voltage. In this case, we have violated the rule because of our 2.6-V drop. Assuming we had chosen a 0.7-V drop, then the collector–emitter voltage would be 6.3 V available for the swing. The choice of 5 mA for the transistor has come from a combination of good linearity and low noise figure. While the 10-Ω unbypassed resistor will grade the noise figure somewhat, it reduces the emitter-current-dependent distortion. The base–emitter diffusion resistance equals 26 mV/5 mA, or roughly 5 Ω. The 10-Ω resistor is twofold larger and therefore adds to the linearity and to the noise.

To determine the base bias resistor it is a good assumption that the current going through the two resistors (in our case, 8.2 and 3.3 kΩ) is about 10% (or a slightly higher fraction) of



Parts List

Frequency	1.9 GHz
C1, C2 (Siemens size 0603)	12 pF
C3, C4 (Siemens size 0603)	100 nF
C5 (Siemens size 0603)	47 nF
C6 (Siemens size 0603)	1.2 pF
L1 (Coilcraft 0805CS-150XKBC)	15 nH
R1 (Siemens B 54102-A1271-J60)	270 Ω
R2 (Siemens B 54102-A1120-J60)	12 Ω
R3	2.7 Ω

Figure 3-53 CGY121A 1.9-GHz application circuit and parts list.

the collector current. Since the dc current gain of the transistor is typically between 50 and 100, it is sufficient if the current through the resistive network is between 5 and 10 times higher than the dc bias current taken off the network. This type of bias scheme is only valid for Class A to AB_1. Later, we will see that for monolithic circuits, additional transistors will be used to generate the required voltage drops. Lately, combinations of *npn* and *pnp* transistors have become popular to avoid using any resistance in the emitter for dc biasing purposes. Figure 3-56 shows such a circuit. This, of course, invites thermal runaway. To minimize the circuitry, there are bias ICs, such as the Siemens BCR400R, now available. Figures 3-57 and 3-58 show its application in biasing BJTs and FETs, respectively; it can also be used to control PIN-diode bias current for TR antenna switching, as shown in Figure 3-59.

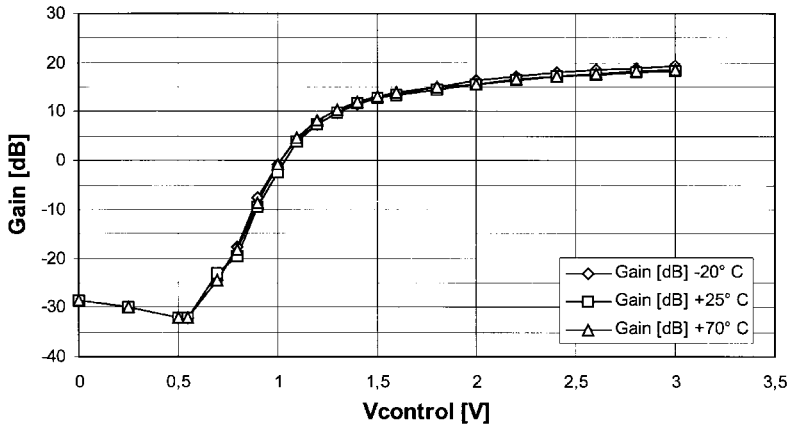


Figure 3-54 CGY121A gain versus $V_{control}$ at 1.9 GHz.

To validate this, we have used the Serenade 8.0 CAD tool, which has a built-in bias- and temperature-dependent noise model for BJTs and FETs. The SPICE-type dc analysis determines the actual currents, which were mentioned above. By maintaining the same bias and just bypassing both emitter resistors at the same time, the set of curves shown in Figure 3-60 will be obtained. As an exercise, we recommend the user to determine the optimum generator impedance for F_{min} with a 10-Ω resistor. Again, both the emitter current and collector voltage affect the noise figure, the intermodulation, and f_T. If a transistor is used

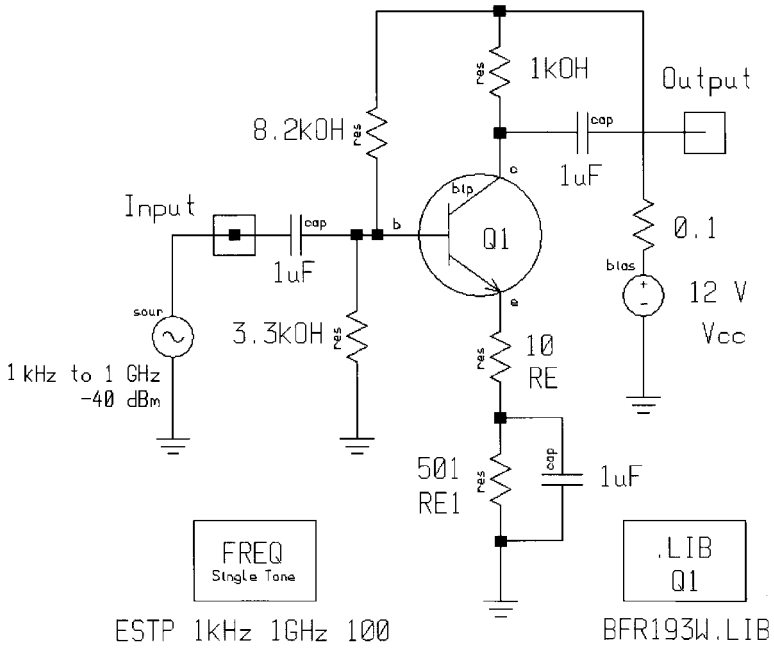


Figure 3-55 BJT amplifier example.

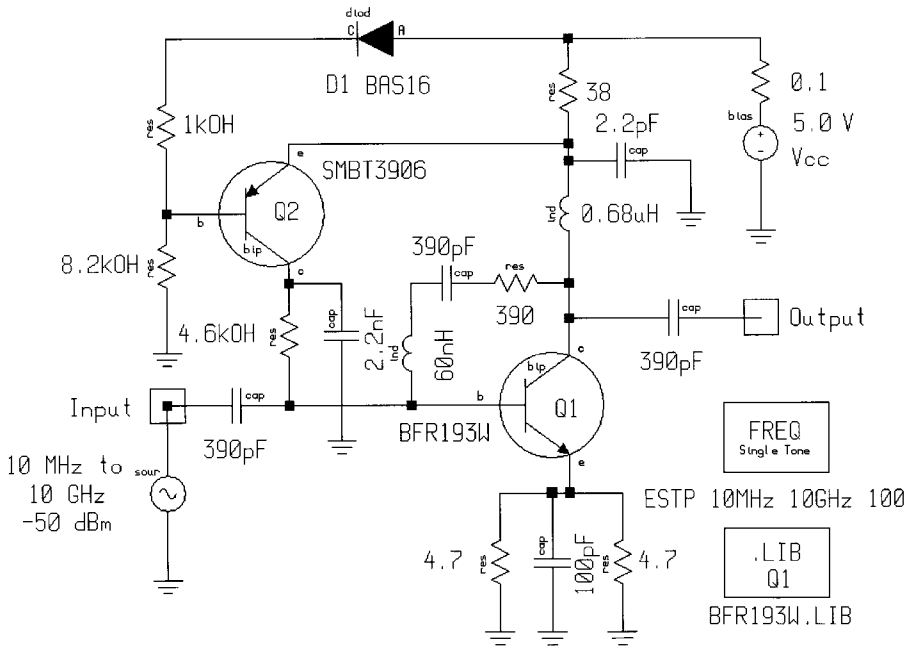


Figure 3-56 Wideband amplifier with active biasing. The 4.7-Ω resistors in the BFR193W emitter compensate the circuit's gain at low frequencies.

that is intended to be operated at significantly higher currents, then f_T will suffer while the other two parameters are less affected.

The dc input power translates into thermal dissipation, and Figure 3-60 is a plot of the noise figure as a function of temperature with and without emitter feedback. The heavy dc stabilization prevents a significant effect on the other parameters. The actual dc shift is from 5.0 to 5.2 mA for a temperature change of 70 K.

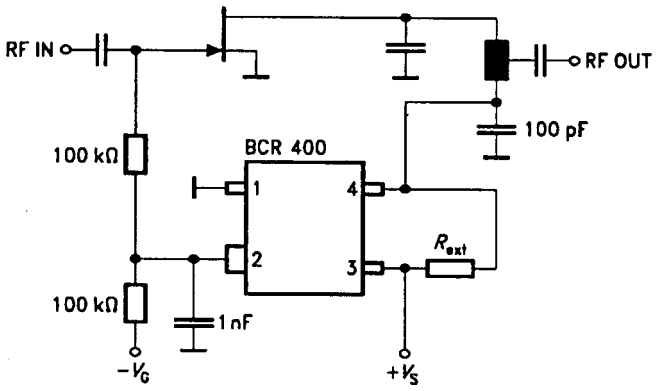


Figure 3-57 The Siemens BCR400 bias controller applied to a GaAsFET.

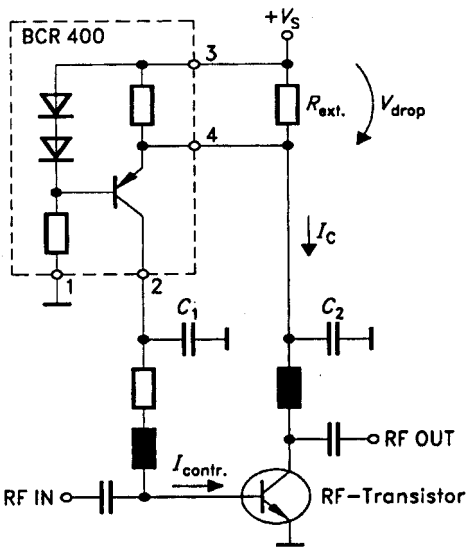


Figure 3-58 The Siemens BCR400 bias controller applied to a BJT.

As far as FETs are concerned, the enhancement types, typically PMOS, can use a similar type of bias, while the NMOS types complicate the designer's life as they require a negative gate voltage. For low-noise operation, these devices are operated somewhat close to the pinchoff voltage. Putting a source resistor in place causes mostly headaches simply because the gain is so high that the transistors are always ready to find a frequency at which they love to oscillate. The transistor will look for every possible parasitic to turn the amplifier into an oscillator.

3-2 AMPLIFIER GAIN, STABILITY, AND MATCHING

Most semiconductor manufacturers describe their devices these days by providing either their *S* parameters (small-signal) or SPICE-type parameters (large-signal). One of our headaches in writing this book has been that many of the published models don't translate

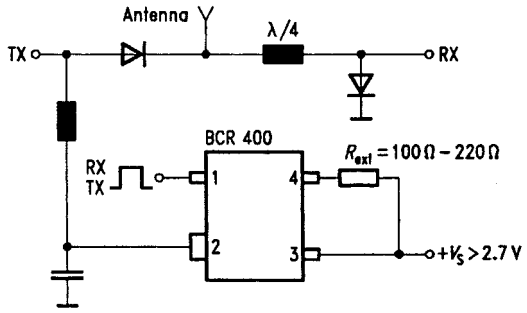


Figure 3-59 The Siemens BCR400 bias controller applied to TR switching.

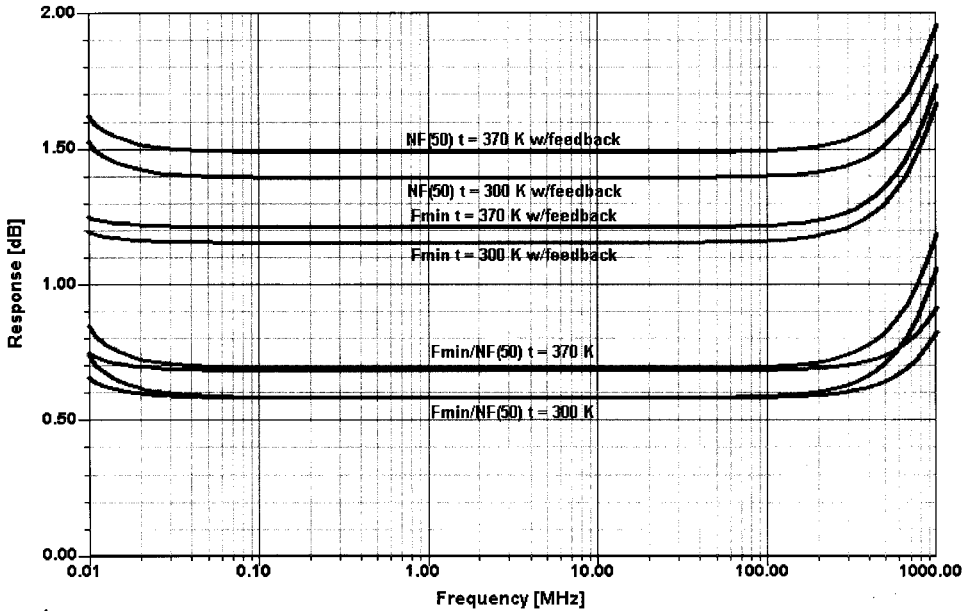


Figure 3-60 F_{min} and NF for the amplifier in Figure 3-55, with and without emitter feedback, at 300 and 370 K.

from large signal to small signal without giving unexpected or false results. Despite the availability of tools to do it right, many companies still can only provide data that are practically inadequate, and unfortunately the CAD manufacturers then get the blame for results for which they are not really responsible.

Tables 3-9 and 3-10 show typical data for a bipolar transistor (the BFP420 by Siemens). The next section presents a summary of the derivation and use of *S* parameters as published by Hewlett-Packard and others.

3-2-1 Scattering Parameter Relationships

Scattering (*S*) parameters characterize a network in terms of incident and reflected waves (Figures 3-61 and 3-62).

$$b_1 = s_{11}a_1 + s_{12}a_2 \tag{3-122}$$

$$b_2 = s_{21}a_1 + s_{22}a_2 \tag{3-123}$$

Notice that

$$s_{11} = \frac{b_1}{a_1} = \frac{V_1/I_1 - Z_0}{V_1/I_1 + Z_0} = \frac{Z_1 - Z_0}{Z_1 + Z_0} \tag{3-124}$$

and

Table 3-9 BFP420 common-emitter *S* parameters[a]

f	S_{11}		S_{21}		S_{12}		S_{22}	
(GHz)	(Magnitude)	(Angle)	(Magnitude)	(Angle)	(Magnitude)	(Angle)	(Magnitude)	(Angle)
				V_{CE} = 2 V, I_C = 20 mA				
0.01	0.452	–2.3	37.62	178.3	0.0011	94.4	0.956	–0.6
0.1	0.447	–52.1	36.30	164.7	0.0068	82.5	0.941	–12.4
0.5	0.386	–101.1	23.41	121.0	0.0262	61.7	0.632	–47.2
1.0	0.378	–146.2	13.99	96.0	0.0395	57.8	0.395	–63.9
2.0	0.405	173.5	7.18	70.8	0.0664	54.0	0.222	–87.3
3.0	0.446	149.4	4.77	52.6	0.0949	47.1	0.133	–111.3
4.0	0.501	130.0	3.52	36.8	0.1206	38.5	0.133	–158.5
6.0	0.599	104.8	2.27	8.2	0.1646	18.9	0.196	142.0
8.0	0.700	78.5	1.51	–20.8	0.1800	–2.4	0.289	99.3
9.0	0.758	67.6	1.25	–34.4	0.1820	–13.0	0.379	84.1
10.0	0.800	62.0	1.04	–43.5	0.1800	–19.3	0.465	76.6
				V_{CE} = 2 V, I_C = 5 mA				
0.01	0.790	–1.0	15.14	179.2	0.0012	83.4	0.988	–0.7
0.1	0.786	–11.6	14.98	171.8	0.0092	84.1	0.982	–6.5
0.5	0.702	–55.7	12.86	140.1	0.0398	62.8	0.857	–29.8
1.0	0.589	–99.1	9.63	112.6	0.0603	46.5	0.647	–48.6
2.0	0.507	–156.0	5.60	79.4	0.0798	34.6	0.401	–70.3
3.0	0.511	168.5	3.84	57.1	0.0957	29.8	0.267	–84.2
4.0	0.549	142.0	2.87	38.5	0.1121	25.1	0.207	–110.5
5.0	0.604	123.9	2.26	22.1	0.1285	19.4	0.150	–137.3
6.0	0.633	110.0	1.86	6.7	0.1442	13.1	0.173	–169.8

Table 3-10 BFP420 noise parameters

f	$F_{\min}$[a]	G_a[a]	Γ_{opt}		R_N	r_n	$F_{50\Omega}$[b]	$\lvert S_{21} \rvert^2$ [b]
(GHZ)	dB	(dB)	(Magnitude)	(Angle)	(Ω)		(dB)	(dB)
				V_{CE} = 2 V, I_C = 5 mA				
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	–175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	–157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	–142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	–123.0	15.0	0.30	3.30	5.3

[a]Input matched for minimum noise figure, output for maximum gain.
[b]$Z_S = Z_L = 50\ \Omega$.

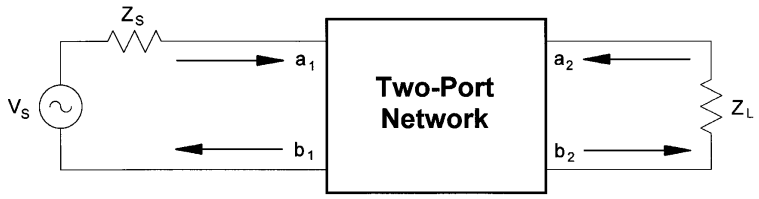


Figure 3-61 *S* parameters characterize a network (a two-port is shown) in terms of incident waves (a_1, a_2) and reflected waves (b_1, b_2). Figure 3-62 shows the flow graph for this network.

$$Z_1 = Z_0 \frac{(1 + s_{11})}{(1 - s_{11})} \tag{3-125}$$

where $Z_1 = V_1/I_1$ is the input impedance at Port 1.

This relationship between reflection coefficient and impedance is the basis of the Smith Chart transmission-line calculator. Consequently, the reflection coefficients S_{11} and S_{22} can be plotted on Smith charts, converted directly to impedance, and easily manipulated to determine matching networks for optimizing a circuit design.

The above equations show one of the important advantages of *S* parameters: They are simply gains and reflection coefficients, both familiar quantities to engineers. By comparison, some of the *Y* parameters described earlier in the article are not so familiar. For example, the *Y* parameter corresponding to insertion gain S_{21} is the "forward transadmittance," Y_{21}. Clearly, insertion gain gives by far the greater insight into the operation of the network.

S parameters are simply related to power gain and mismatch loss, quantities that are often of more interest than the corresponding voltage functions:

$$|s_{11}|^2 = \frac{\text{Power reflected from the network input}}{\text{Power incident on the network input}} \tag{3-126}$$

$$|s_{21}|^2 = \frac{\text{Power delivered to a } Z_0 \text{ load}}{\text{Power available from } Z_0 \text{ source}} \tag{3-127}$$

$$|s_{12}|^2 = \text{Reverse transducer power gain with } Z_0 \text{ load and source.} \tag{3-128}$$

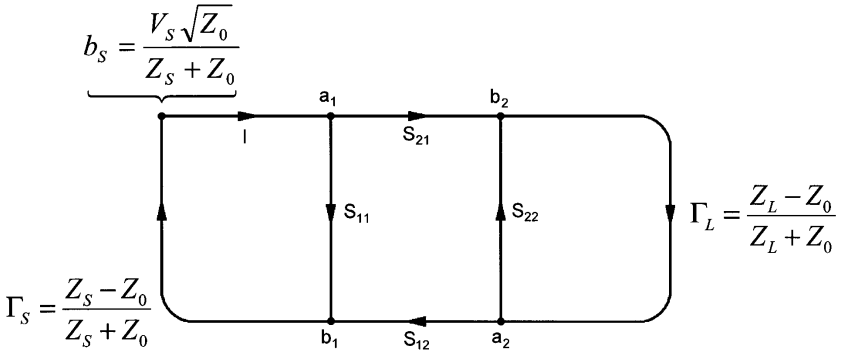


Figure 3-62 Flow graph for the network shown in Figure 3-61.

$$|s_{22}|^2 = \frac{\text{Power reflected from the network output}}{\text{Power incident on the network output}} \tag{3-129}$$

From this we obtain:
Input reflection coefficient with arbitrary Z_L:

$$s'_{11} = s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1 - s_{22}\Gamma_L} \tag{3-130}$$

Output reflection coefficient with arbitrary Z_S:

$$s'_{22} = s_{22} + \frac{s_{12}s_{21}\Gamma_S}{1 - s_{11}\Gamma_S} \tag{3-131}$$

Voltage gain with arbitrary Z_L and Z_S:

$$A_V = \frac{V_2}{V_1} = \frac{s_{21}(1 + \Gamma_L)}{(1 - s_{22}\Gamma_L)(1 + s'_{11})} \tag{3-132}$$

$$\text{Power gain} = \frac{\text{Power delivered to load}}{\text{Power input to network}}:$$

$$G = \frac{|s_{21}|^2 (1 - |\Gamma_L|^2)}{(1 - |s_{11}|^2) + |\Gamma_L|^2 (|s_{22}|^2 - |D|^2) - 2\,\text{Re}\,(\Gamma_L N)} \tag{3-133}$$

$$\text{Available power gain} = \frac{\text{Power available from network}}{\text{Power available from source}}:$$

$$G_A = \frac{|s_{21}|^2 (1 - |\Gamma_S|^2)}{(1 - |s_{22}|^2) + |\Gamma_S|^2 (|s_{11}|^2 - |D|^2) - 2\,\text{Re}\,(\Gamma_S M)} \tag{3-134}$$

$$\text{Transducer power gain} = \frac{\text{Power delivered to load}}{\text{Power available from source}}:$$

$$G_T = \frac{|s_{21}|^2 (1 - |\Gamma_S|^2)(1 - |\Gamma_L|^2)}{|(1 - s_{11}\Gamma_S)(1 - s_{22}\Gamma_L) - s_{12}s_{21}\Gamma_L\Gamma_S|^2} \tag{3-135}$$

Unilateral transducer power gain ($S_{12} = 0$):

$$G_{Tu} = \frac{|s_{21}|^2(1 - |\Gamma_S|^2)(1 - |\Gamma_L|^2)}{|1 - s_{11}\Gamma_S|^2|1 - s_{22}\Gamma_L|^2}$$

$$= G_0 G_1 G_2 \tag{3-136}$$

$$G_0 = |s_{21}|^2 \tag{3-137}$$

$$G_1 = \frac{1 - |\Gamma_S|^2}{|1 - s_{11}\Gamma_S|^2} \tag{3-138}$$

$$G_2 = \frac{1 - |\Gamma_L|^2}{|1 - s_{22}\Gamma_L|^2} \tag{3-139}$$

Maximum unilateral transducer power gain when $|S_{11}| < 1$ and $|S_{22}| < 1$:

$$G_u = \frac{|s_{21}|^2}{|(1 - |s_{11}|^2)(1 - |s_{22}|)^2|}$$

$$= G_0\, G_{1\max}\, G_{2\max} \tag{3-140}$$

$$G_{i\max} = \frac{1}{1 - |s_{ii}|^2} \quad i = 1, 2 \tag{3-141}$$

Constant-gain circles (unilateral case: $S_{12} = 0$):

- Center of constant-gain circle is on line between center of Smith Chart and point representing S_{ii}^*
- Distance of center of circle from center of Smith Chart:

$$r_i = \frac{g_i |s_{ii}|}{1 - |s_{ii}|^2(1 - g_i)} \tag{3-142}$$

- Radius of circle:

$$\rho_i = \frac{\sqrt{1 - g_i}(1 - |s_{ii}|^2)}{1 - |s_{ii}|^2(1 - g_i)} \tag{3-143}$$

where $i = 1, 2$ and

$$g_i = \frac{G_i}{G_{i\max}} = G_i(1 - |s_{ii}|^2) \tag{3-144}$$

Unilateral figure of merit:

$$u = \frac{|s_{11}s_{22}s_{12}s_{21}|}{|(1 - |s_{11}|^2)(1 - |s_{22}|^2)|} \tag{3-145}$$

Error limits of unilateral gain calculation:

$$\frac{1}{(1 + u^2)} < \frac{G_T}{G_{Tu}} < \frac{1}{(1 - u^2)} \tag{3-146}$$

Conditions for absolute stability. No passive source or load will cause a network to oscillate if a, b, and c are all satisfied:

$$\text{a.}\quad |s_{11}| < 1, |s_{22}| < 1 \tag{3-147}$$

$$\text{b.} \quad \left| \frac{|s_{12}s_{21}| - |M^*|}{|s_{11}|^2 - |D|^2} \right| > 1 \tag{3-148}$$

$$\text{c.} \quad \left| \frac{|s_{12}s_{21}| - |N^*|}{|s_{22}|^2 - |D|^2} \right| > 1 \tag{3-149}$$

where

$$D = s_{11}s_{22} - s_{12}s_{21} \tag{3-150}$$

$$M = s_{11} - Ds_{22}^* \tag{3-151}$$

$$N = s_{22} - Ds_{11}^* \tag{3-152}$$

Condition that a two-port network can be simultaneously matched with a positive real source and load:

$$K > 1 \tag{3-153}$$

$$K = \frac{1 + |D|^2 - |s_{11}|^2 - |s_{22}|^2}{2|s_{12}s_{21}|} \tag{3-154}$$

Source and load for simultaneous match

$$\Gamma_{mS} = M^* \left[\frac{B_1 \pm \sqrt{B_1^2 - 4|M|^2}}{2|M|^2} \right] \tag{3-155}$$

$$\Gamma_{mL} = N^* \left[\frac{B_2 \pm \sqrt{B_2^2 - 4|N|^2}}{2|N|^2} \right] \tag{3-156}$$

where

$$B_1 = 1 + |s_{11}|^2 - |s_{22}|^2 - |D|^2 \tag{3-157}$$

$$B_2 = 1 + |s_{22}|^2 - |s_{11}|^2 - |D|^2 \tag{3-158}$$

Maximum available power gain. If $K > 1$,

$$G_{A\max} = \left| \frac{s_{21}}{s_{12}} (K \pm \sqrt{K^2 - 1}) \right| \tag{3-159}$$

In conclusion, *a two-port network (active or passive) is unconditionally stable* when

$$K > 1$$

$$|S_{12}S_{21}| < 1 - |S_{11}|^2$$

and

$$|S_{12}S_{21}| < 1 - |S_{22}|^2$$

3-2-2 Low-Noise Amplifiers

A low-noise amplifier combines a low noise figure, reasonable gain, and stability without oscillation over its entire useful frequency range. These amplifiers are typically operated in Class A, which is characterized by a bias point more or less at the center of maximum current and voltage capability of the device used, and by RF current and voltages that are sufficiently small relative to the bias point that the bias point does not shift. This means that an amplifier operating at, say, 3 V and 1 mA for low-noise, high-gain operation will not be Class A if the RF current is larger than 0.1% of the current or the voltage swing is more than 0.1% of the current swing. These amplifiers can be designed using standard linear *S* parameters, which are (hopefully) available from the device manufacturer for the bias point of interest. If suitable *S* parameters are not available, we must resort to a CAD tool that allows us to determine a device's bias point as a function of its bias network and to generate small-signal *S* parameters for the device. With suitable *S* parameters in hand, and assuming the amplifier will be frequency selective, we then connect tuned circuits to the input and output to provide matching.

In designing amplifiers, the "power gain" of a circuit is mentioned frequently. This is actually deceptive because power gain is somewhat associated with dc power dissipation, and in considering it, we are on the verge of discussing large-signal performance. Therefore it would be better to define the power gain as a function of operating mode, such as "Class A power gain." We have also outlined that there is a difference between noise matching and gain matching. If we can be allowed to step back into the vacuum-tube era for a moment, the reader should be reminded that the difference between power gain and noise matching had to do with the grid-plate feedback capacitance which, via the Miller effect, showed up in parallel with the input. This actually detunes the input slightly off-center relative to the best gain. The reason for this is that as the feedback increases, Y_{12} or S_{12} starts playing a role, and the frequently assumed unilateral case no longer exists. The difference between noise matching and gain matching increases as a function of frequency. It can be brought closer together if appropriate reactive feedback is used. The two choices are: (1) increase the emitter or source inductance or (2) slightly increase the base-to-collector/gate-to-drain capacitance. Needless to say, the penalty for this is potential instability, and a sweep must be done from a few megahertz to several gigahertz to make sure the stage doesn't "take off."

The most efficiency this circuit provides is 50%; even this is a theoretical limit if one tries to keep distortion under control.

While the equation $P_{\text{OUT}} = (V_{\text{bat}} - V_{\text{sat}})^2/2R_L$ is valid, it really applies only to power amplifiers in the sense that for Class A, high-gain operation R_L can be made significantly higher as the absolute power output is not the important item. While the supply voltage V_{bat} is always known, the saturation voltage V_{sat} has to be obtained by the load line dc analysis. Note: The RF saturation voltage is always higher than the dc one. When looking at Figure 3-63, it becomes obvious that the conducting angle is 180°, which implies that the entire sine wave at the input is reproduced as an essentially undistorted sine wave at the output.

We can now calculate the components of the distorted current in separating them into the dc component, the fundamental frequency, and its first and second harmonics:

$$\text{Fundamental} \quad i_{c1} = i_{\text{peak}} \cdot f_1(\Theta) \tag{3-160}$$

$$\text{Second harmonic} \quad i_{c2} = i_{\text{peak}} \cdot f_2(\Theta) \tag{3-161}$$

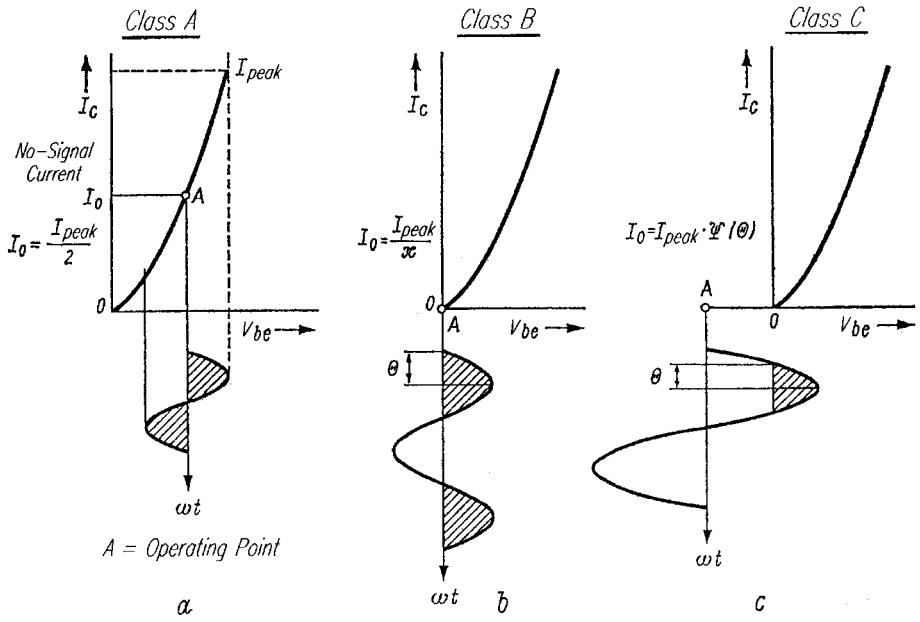


Figure 3-63 Definitions of Class A, B, and C operation, including conduction angle (Θ). Figure 3-64 shows the determination of Θ.

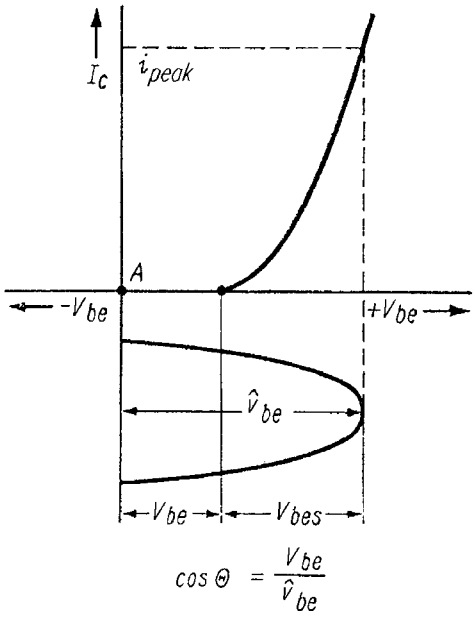


Figure 3-64 Determination of the conduction angle Θ from the base threshold voltage V_{be} and the base RF voltage $\hat{v}_{be}$.

$$\text{Third harmonic} \quad i_{c3} = i_{\text{peak}} \cdot f_3(\Theta) \tag{3-162}$$

For the collector dc current we obtain

$$I_c = i_{\text{peak}} \cdot \Psi(\Theta) \tag{3-163}$$

The frequency-dependent currents are a function of the conduction angle Θ. Figure 3-65 shows these transfer characteristics as a function of the conduction angle for a square-law dynamic transfer function; for higher-order transfer functions, the values don't change much. Tables for this can be found in various sources, including Clark and Hess [47]. Since the transistors exhibit a mixture of an exponential and a linear transfer characteristic (due to the unwanted parasitic emitter/source and collector/drain resistances when building the device), the plot in figure 3-65 is sufficiently valid for practical purposes. This function plot allows us to determine the amplitude of the current at various frequencies once the conducting angle has been set. For example, for Θ = 70°, the maximum collector current is $i_{\text{peak}} = I_c/0.253$, meaning that the peak current of the transistor is approximately four times larger than the dc value one would measure. Once the peak collector current is known, the fundamental and harmonic currents can be calculated:

$$i_{c1} = \frac{I_c}{0.253} f_1(\Theta)$$

$$i_{c1} = \frac{I_c}{0.253} 0.436$$

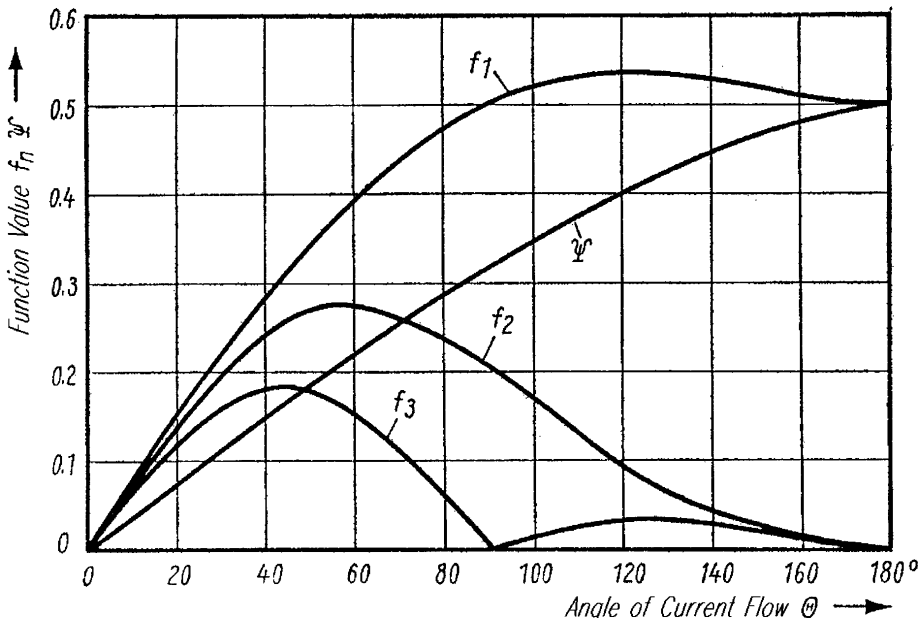


Figure 3-65 Function values for the fundamental, second harmonic, third harmonic, and dc components (f_1, f_2, f_3, and Ψ, respectively) versus device conduction angle.

In order to obtain i_{c2} and i_{c3} one needs to do the same calculations as above with the appropriate functions for f_2 and f_3.

It becomes obvious that the distorted collector current generates a lot of harmonics, which definitely requires sufficient low-pass filtering at the output, even for Class A operation. We already stated that for Class A operation the conduction angle is 180° while it is 90° for Class B operation. The conduction angle for Class C operation can be solved from the ratio of the knee voltage V_{be} and the peak value of the ac voltage at the base. We find

$$\cos \Theta = \frac{V_{be}}{\hat{V}_{be}} \tag{3-164}$$

where $\Theta = 90°$ and $V_{be} = 0$. Given the voltage V_{bes}, which determines the peak collector current (up to i_c peak), we can calculate the effective RF voltage at the base as

$$\hat{v}_{be} = \frac{v_{bes}}{1 - \cos \Theta} \tag{3-165}$$

Therefore, the peak value of v_b becomes v_{bes} for $\Theta = 90°$.

The necessary base voltage V_{be} for maximum current is obtained by the transistor model. Then one finds the threshold voltage at which the transistor becomes conductive and by using Eqs. (3-164) and (3-165) it can be checked whether this results in the appropriate conducting angle. An example for this will be shown in the high-power stage section.

Design Guidelines. The design of a one-stage amplifier consists of finding

- An input lossless matching network M_1
- An input lossless matching network M_2

so that the maximum or desired transistor gain is achieved over the operating bandwidth of the amplifier. Usually, the common-emitter or common-source configuration is chosen for highest gain per stage. If the stability factor K is greater than unity, these two networks can be found to give the maximum available gain $G_{\max}$. If the stability factor is less than or equal to unity, the amplifier could be terminated in a matching structure that causes oscillation; that is, $G_{\max}$ is infinite. This should be avoided by locating the regions of instability in the Γ_G and Γ_L planes. The input and output terminations (Γ_G and Γ_L) must be designed to avoid the instability regions. Usually, these unstable regions are near the conjugate match for S_{11} and S_{22}. Thus, a stable amplifier will require some input and/or output mismatch if K is less than or equal to unity.

There are at least two alternative approaches for potentially unstable amplifiers:

1. Add resistive matching elements to make $K \geq 1$ and $G_{\max} \approx G_{ms}$.
2. Add feedback to make $K \geq 1$ and $G_{\max \approx Gms}$.

For narrowband amplifiers, it is usually recommended that one accept a transistor with $K < 1$, design the amplifier for a gain approaching G_{ms}, and ensure that the Γ_G and Γ_L terminations provide stability at all frequencies, both inside and outside the amplifier passband.

The design of an amplifier would usually have the following specifications:

Gain and gain flatness
Bandwidth and center frequency ($f_2 - f_1, f_0$)
Noise figure
Linear output power
Input reflection coefficient ($VSWR_{in}$)
Output reflection coefficient ($VSWR_{out}$)
Bias voltage and current

For small-signal amplifiers, the small-signal S parameters are sufficient to complete the design. After selecting an appropriate transistor based on these specifications, a one-stage amplifier design should be considered if sufficient gain can be achieved; otherwise, a two-stage amplifier should be designed.

The circuit topology should be chosen to allow dc bias for the transistor. Usually, RF short-circuited stubs are placed near the transistor to allow dc biasing. If the topology does not allow dc bias, a broadband bias choke or high-resistance bias circuit that does not affect the amplifier performance must be used.

The following steps can be tabulated for the design of a one-stage amplifier (see Figure 3.66):

1. Select a transistor based on a datasheet description of the S parameters, noise figure, and linear output power.
2. Calculate K and G_{max} or G_{ms} versus frequency.
3. For $K > 1$, select the topologies that match the input and output (and allow dc biasing) at the upper band edge f_2. Ideally, this will give G_{max} and $S'_{11} = S'_{22} = 0$. Usually, $S_{12} = 0$ is assumed for the initial design. Next, the topology may be varied to flatten gain versus frequency at the expense of S'_{11} and S'_{22}.
4. For $K < 1$, plot the regions of instability on the Γ_G and Γ_L planes and select topologies that partially match the input and output at the upper band edge and avoid the unstable regions. The gain will approach G_{ms} as an upper limit. Next, the topology may be varied to flatten gain versus frequency.
5. After finding initial M_1 and M_2, plot the amplifier S parameters versus frequency; make adjustments in topology until the specifications for gain, input reflection coefficient, and output reflection are satisfied. Also plot Γ_G and Γ_L versus frequency to verify amplifier stability.
6. Design the dc bias circuit. Lay out the elements of the complete amplifier and check realizability.

As a beginning point in an amplifier design, the circuit topologies may be designed with the assumption that S_{12} is zero. Later, an exact design procedure will be described that includes the topologies with terminations in the stable region.

Design Examples. Throughout this book we always address the issue that we have to separate the battery-operated handheld applications and the more stationary designs like those for base stations. In the case of the handheld or low-voltage applications, the dynamic range requirements are significantly less than the front-end of a base station requires.

NE68133 Matched Amplifier. We will first consider the design of a low-voltage, low-current, low-noise amplifier based on the NEC BJT NE68133. This device was chosen

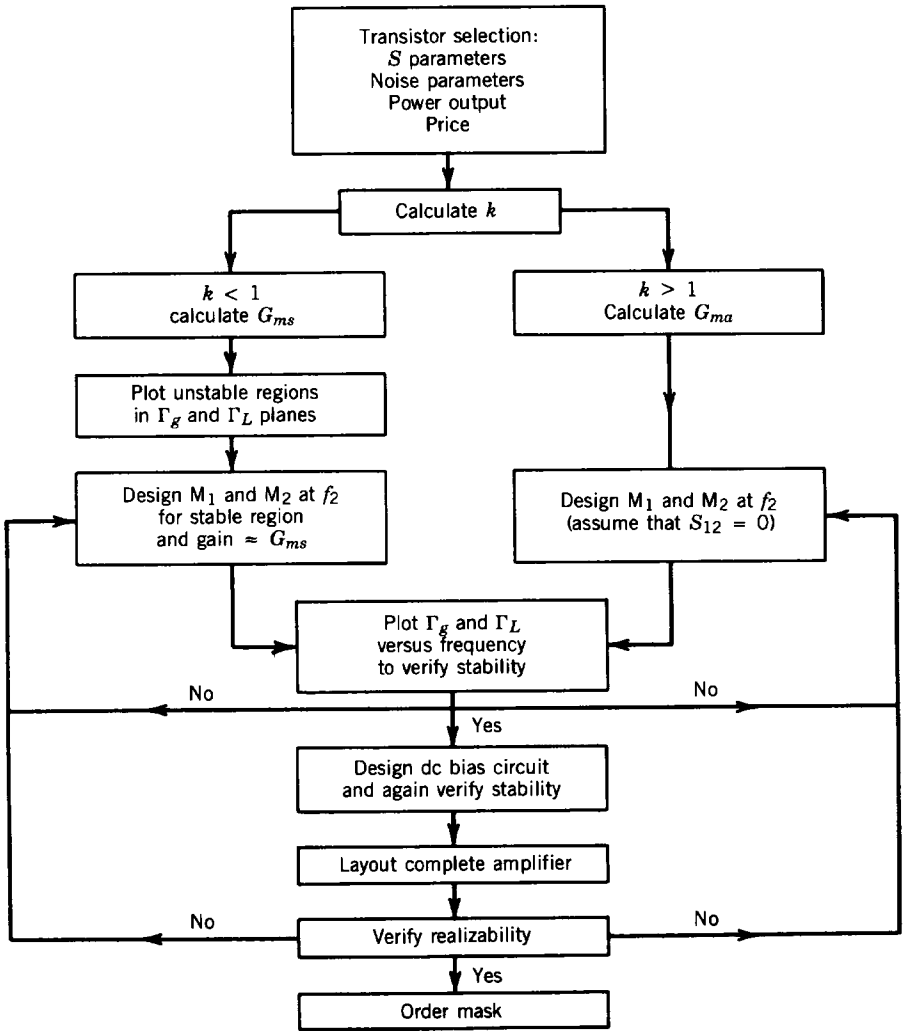


Figure 3-66 Simplified amplifier design procedure.

because it is unconditionally stable, has a minimum noise figure of about 1.2 dB at 7 mA, has 15.5 dB gain, and has an f_T of 8 GHz. While it is possible to start the design based on the published S and noise parameters, we will choose also to involve the nonlinear model, as this allows us more freedom in selecting the bias point including the collector voltage. The actual topology of the amplifier is shown in Figure 3-67.

Determining appropriate values for the input and output matching networks is essential to finalizing the design. The next several figures illustrate matching-network syntheses undertaken with the interactive Smith Tool in Ansoft's Serenade Design Environment. We evaluate the point of tangency as well as stability circles in the source and load planes, indicating that our impedance match will result in a stable amplifier.

In Figure 3-68, the center circle (with 1.1 marker) refers to the noise circle for which the noise figure of 1.1 dB can be obtained. The left circle (with the 14.3 marker) corresponds to

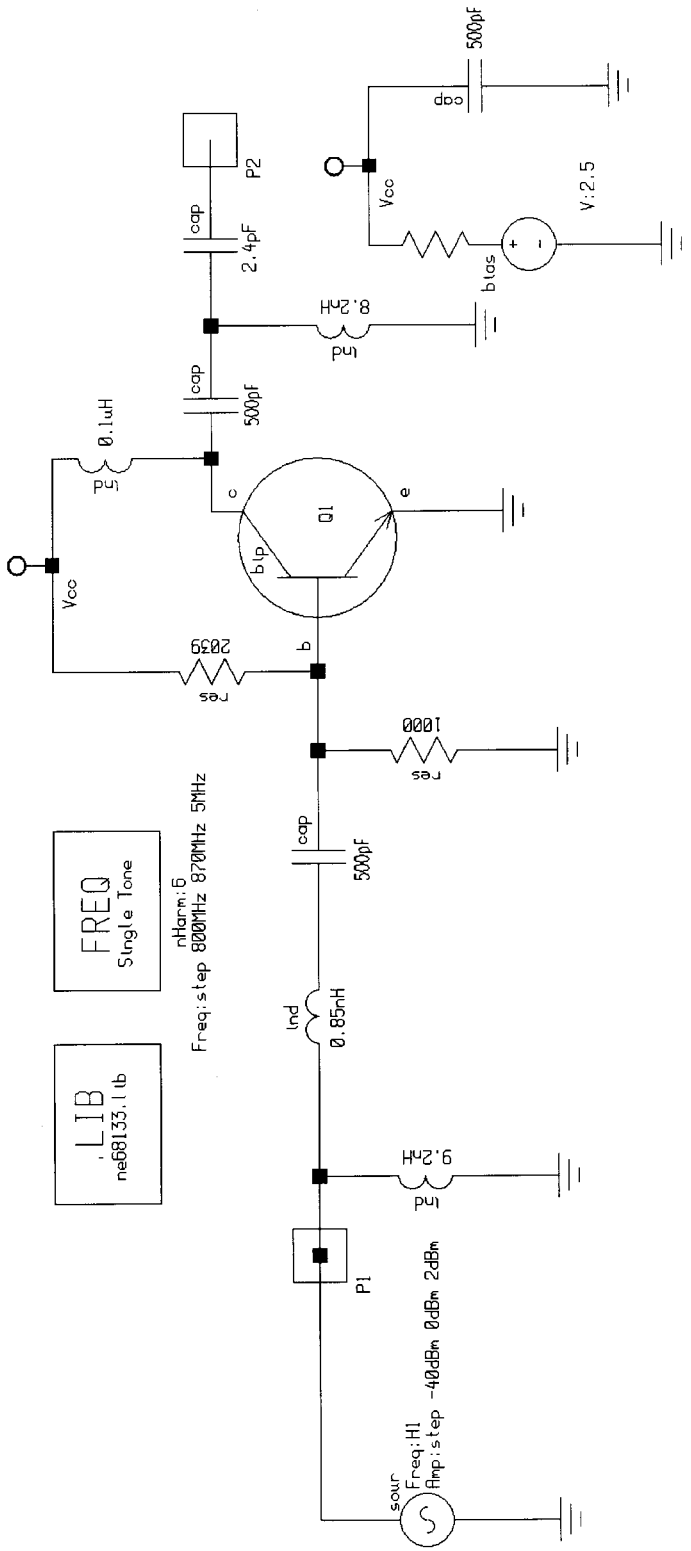


Figure 3-67 The NE68133 low-noise amplifier.

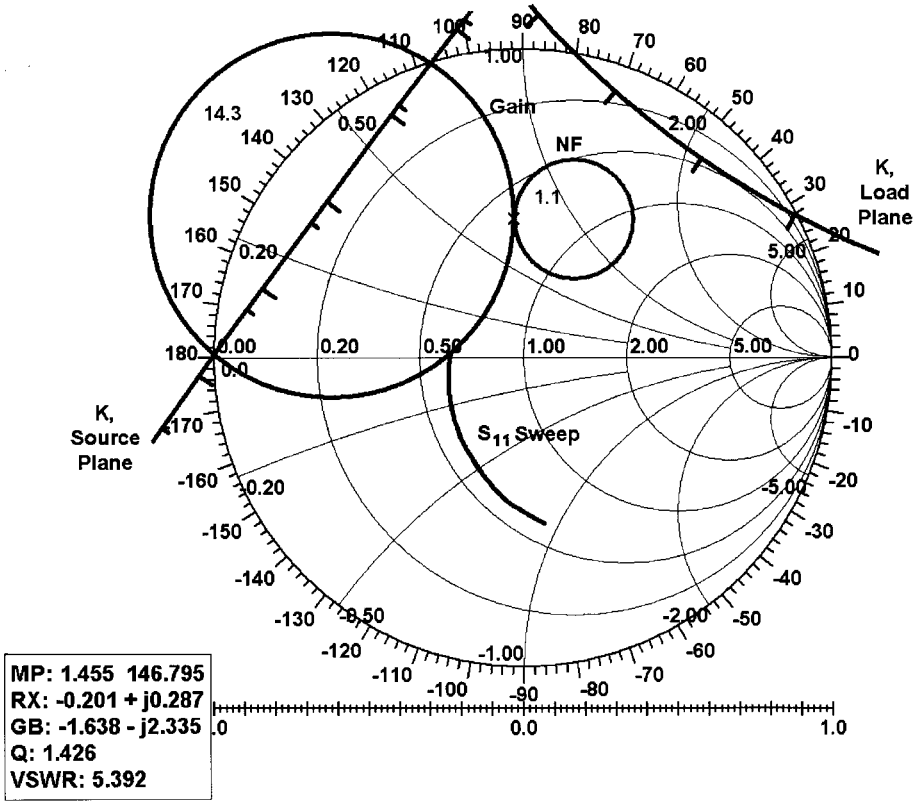


Figure 3-68 Display from the interactive Smith Tool, showing circles for gain, noise figure, and source- and load-plane stability, as well as a sweep of S_{11} versus frequency. In synthesizing an input matching circuit, we choose a single working frequency and work from the chart's origin to the point of tangency between the gain and noise figure circles. See Figure 3-69.

source-plane terminations that correspond to 14.3 dB of available gain. In order to get the required noise figure *and* gain, the input termination must be at the point of tangency of both circles. The arc touching the X axis close to 0.50 is the amplifier's S_{11} response from 800 to 875 MHz. The shallow arc that nearly bisects the 14.3-dB gain circle is actually a portion of the source-plane stability circle; the arc at the top right is a portion of the load-plane stability circle. Because we will be terminating the amplifier input with a value inside the source-plane stability circle, and terminating the amplifier output with a value outside the load-plane stability circle, the amplifier will be unconditionally stable at the frequency at which the match is obtained.

The next task is to obtain the values for the network circuit that matches the input for the 1.1-dB noise figure. A noise matching will not result in the best possible S_{11} matching (Figure 3-69). This translates into an inductor of 9.2 nH to ground and a series inductor of 0.85 nH based on the large-signal model of the NE68133 at the chosen bias point.

After we have determined our input matching network, it is time to find the output matching (Figure 3-70). First, find the termination point on the output plane that corresponds to 14.3-dB gain and 1.1-dB NF at the input plane. The output gain circle goes through the

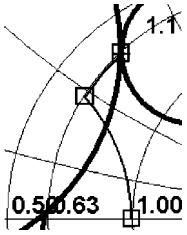


Figure 3-69 Input matching-network extraction, assuming lossless matching components and no parasitic effects.

1.1-dB NF circle at two points, and for reasons of easy matching we selected the lower point. The next step is to find the conjugate reflection coefficient $\Gamma_L = S_{22}^*$. This is the lower square marker in Figure 3-70. The remaining task is to find the matching network that brings this value to 50 Ω. By doing so, we first obtain a shunt inductance $L_{\text{shunt}} = 8.2$ nH and a series capacitor $C_{\text{series}} = 2.4$ pF. Again, these are the values applicable to the large-signal equivalent circuit.

The careful reader will have noticed that we have been avoiding the issue of small-signal exact matching using the manufacturer-supplied parameters. If this is done, the narrow band response shown in Figure 3-71 will be obtained. Therefore, we must optimize the circuit for a better frequency response to obtain the one shown in Figure 3-72.

After this optimization, we were well able to meet the specification of a center band gain of 14 dB and a corresponding NF of 1.1 dB. To show the flatness of the gain, Figure 3-73 shows a gain variation in the range of 0.25 dB.

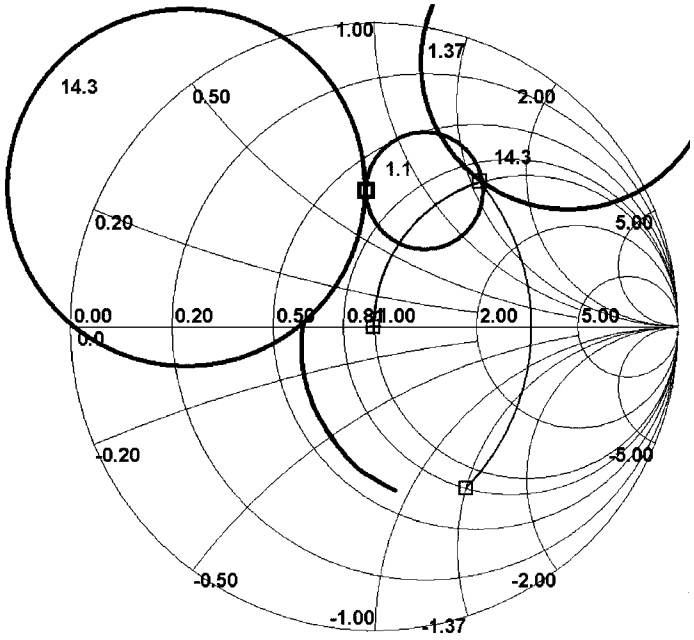


Figure 3-70 Output matching-network extraction.

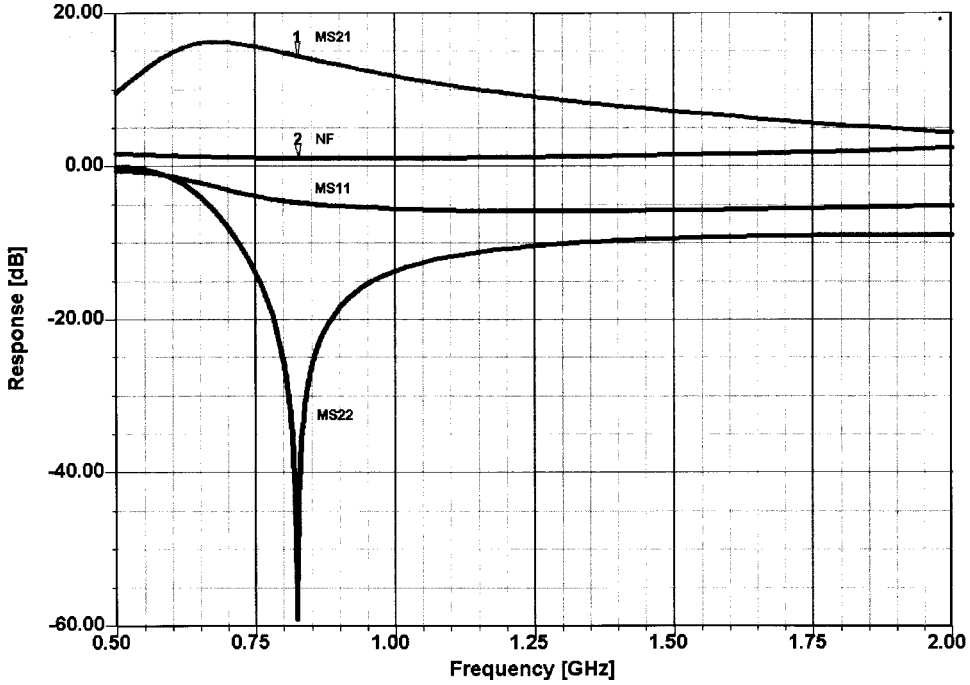


Figure 3-71 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-67.

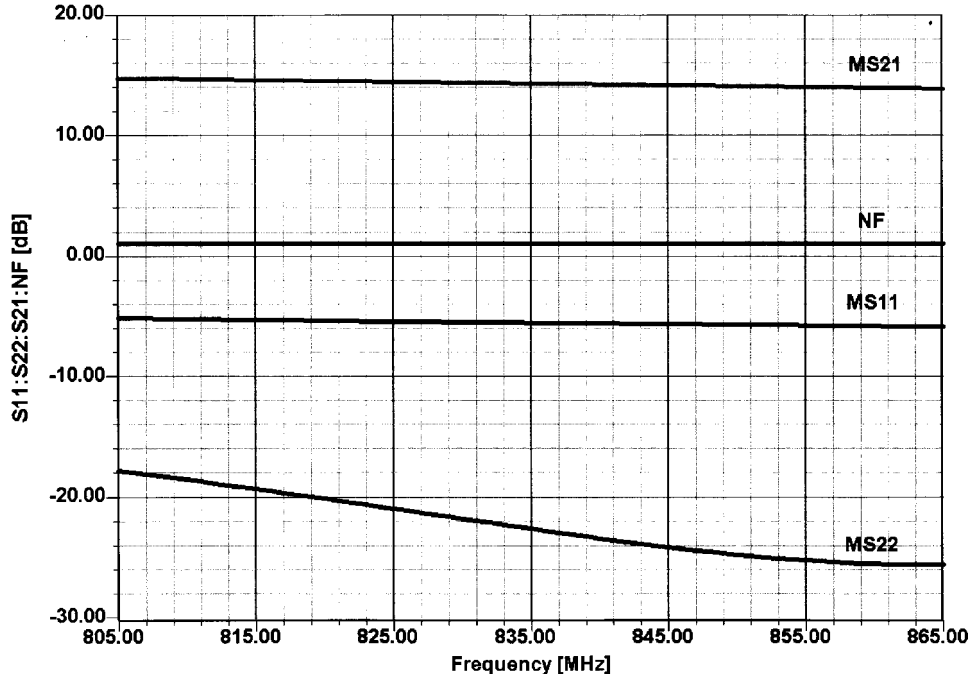


Figure 3-72 Optimized performance of input/output reflection, gain, and noise figure.

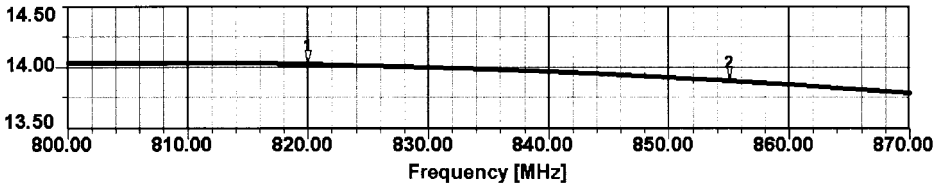


Figure 3-73 High-resolution display of gain as a function of frequency.

The next step is to look into the large-signal performance. The first order of business is to determine the third-order intercept point, which is based on a two-tone analysis. Given the fact that we run the device at low voltage and at a just-reasonable current, the actual intercept point of +13.5 dBm is quite a good number. Figure 3-74 shows the fundamental and third-order outputs. The crossover point between the two, as shown in Chapter 1, is defined as the third-order intercept point. If we had chosen the third harmonic output rather than the fundamental, we would have shown the fifth-order intercept point. The second-order IMD product ($f_1 \pm f_2$) typically can be minimized by input selectivity.

How an amplifier's linearity affects its ability to preserve the characteristics of digital signals is a major concern in wireless design. Figures 3-75 and 3-76 show constellation diagrams with the amplifier handling a PSK signal at drive levels of –40 dBm and +10 dBm, respectively. As expected, we see significant distortion at the higher drive levels due to circuit nonlinearities.

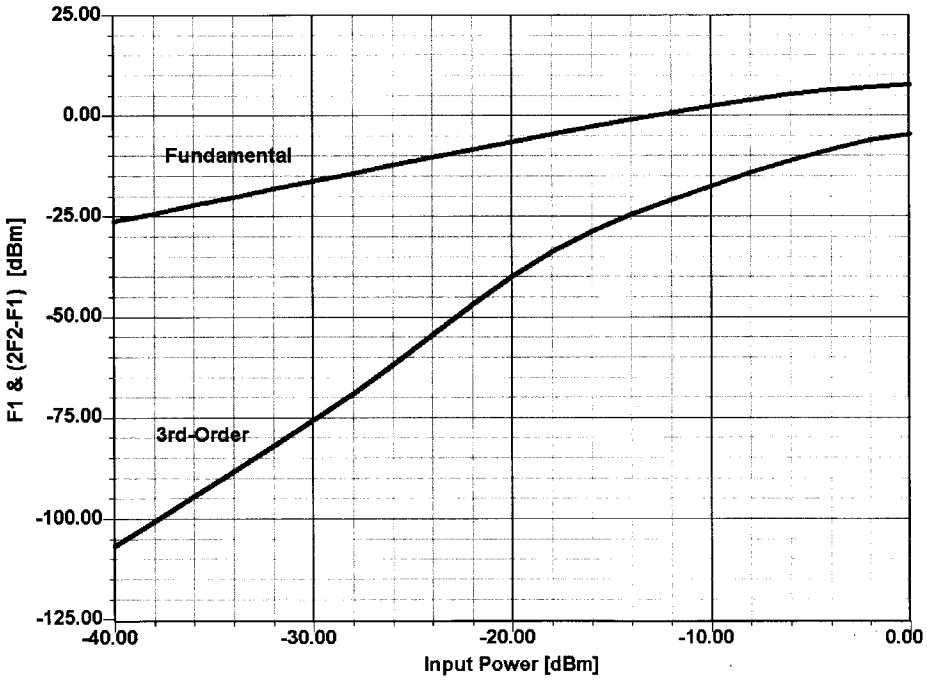


Figure 3-74 Simulated fundamental and third-order IMD outputs of the NE68133 amplifier.

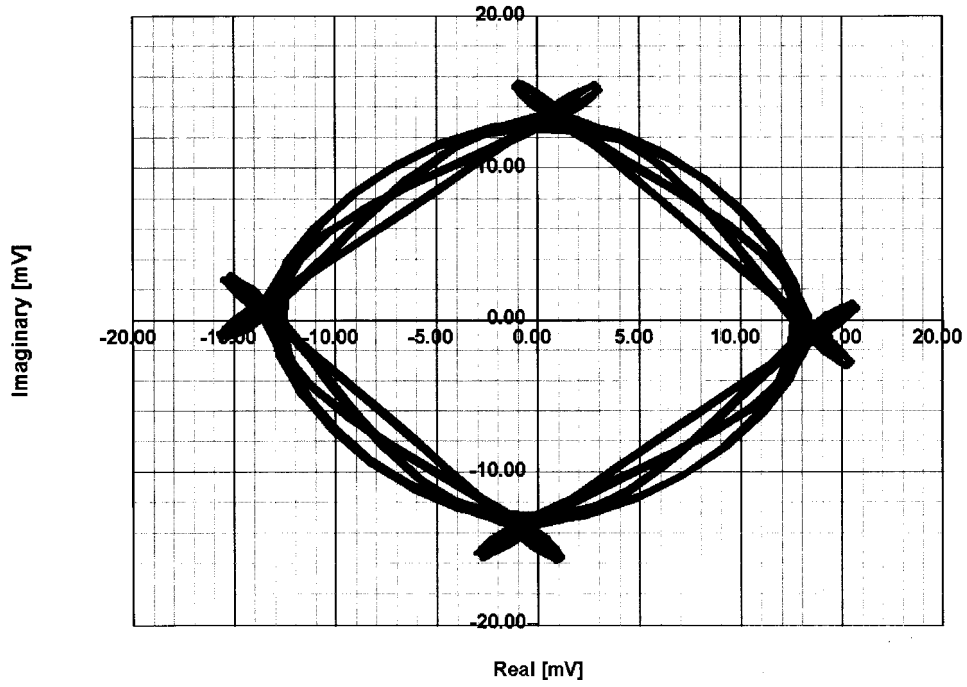


Figure 3-75 Output constellation at –40-dBm input.

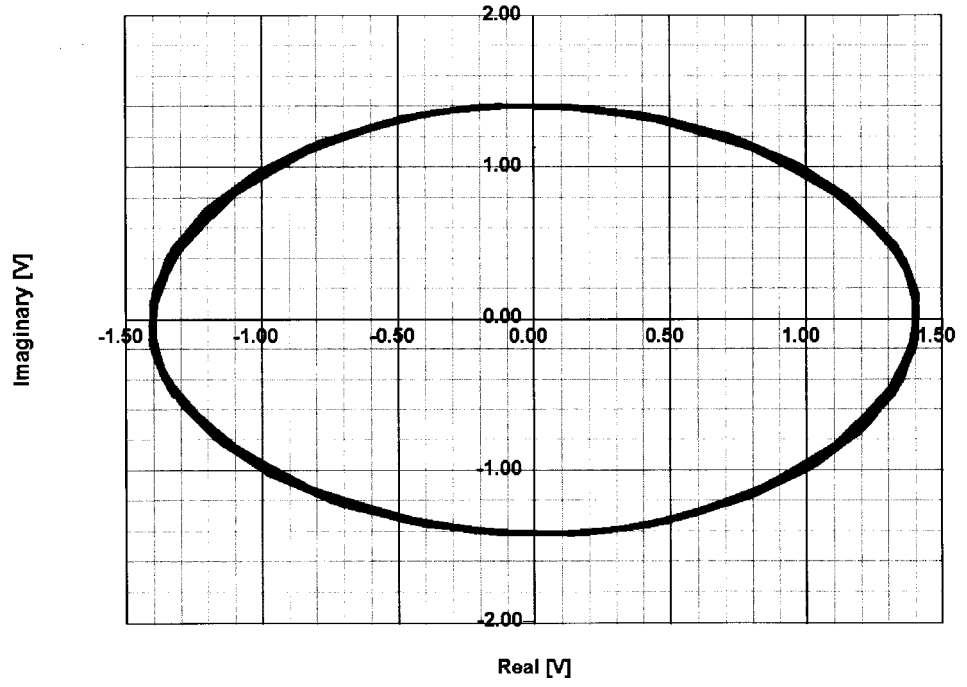


Figure 3-76 Output constellation at 10-dBm input.

BFP420 Matched Amplifier. The NE68133 has an attractive f_T of 8 GHz, but if more gain is needed, one may play with the concept of using a much "hotter" device, such as the BFP420. Recall that the NEC device was unconditionally stable, and our first worry is how the Siemens BFP420 will behave.

A study of Figure 3-77 shows that indeed the NEC device had to be stable all over and therefore was a good basis for a risk-free design. Things change dramatically with the BFP420, for which the K factor is essentially less than 1 up to a corner frequency of 2.75 GHz, where the transistor becomes unconditionally stable. Around 1.75 GHz, the K factor for the NEC transistor barely touches 1. We explained earlier that $G_{\max}$ deviating from the maximum stable gain is a more realistic number, since the conditions for maximum stable gain involve feedback and conjugate matching.

To improve the stability of the BFP420, we will now set out to use the same topology as the previous amplifier but with a resistive feedback between base and collector (Figure 3-78). A good value for this is always somewhere around 200–500 Ω. It must be remembered that this voltage feedback also stabilizes the input and output impedance and reduces the Miller effect because of the feedback being resistive and more dominant.

While the first example covered a fairly narrow frequency band, we are now attempting to cover a much wider frequency range; specifically, from 1.5 to 3 GHz reasonable gain and noise. This, by definition, requires a much higher gain-bandwidth product, and the simple matching network at the input and output provides a nonoptimal input and output matching (Figure 3-79). The feedback introduces some additional noise. This application shows the limits of single-stage UHF/SHF amplifiers with feedback.

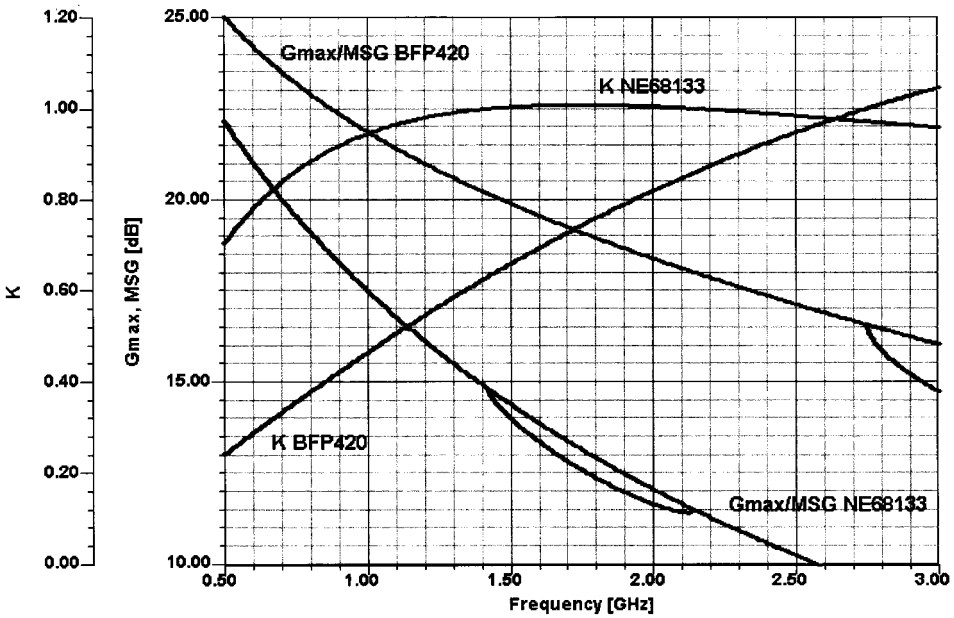


Figure 3-77 Comparison of K, $G_{\max}$, and maximum stable gain (MSG) for the NE68133 and BPF420 BJTs.

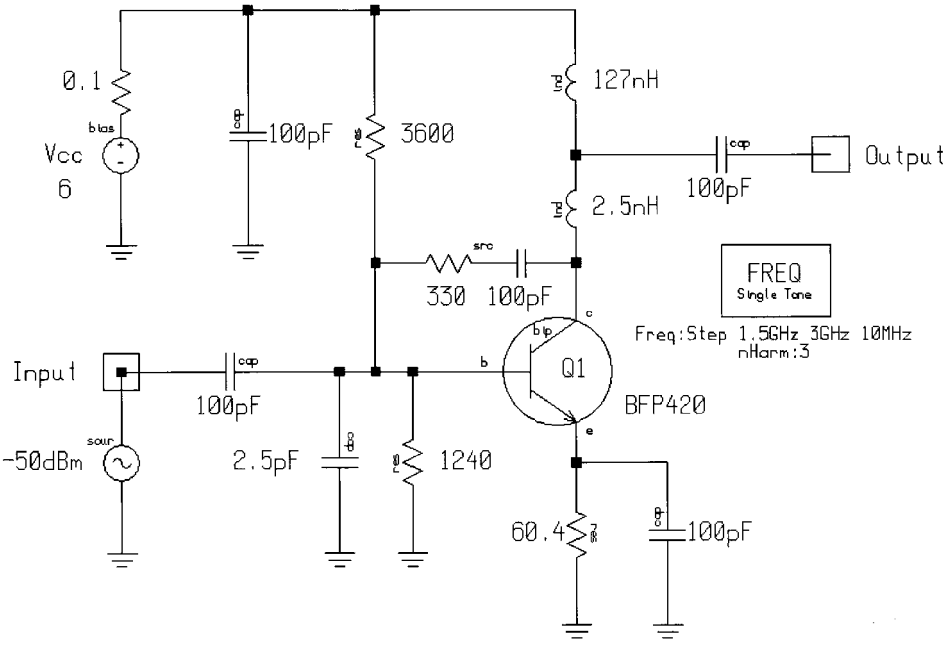


Figure 3-78 BFP420 amplifier with resistive voltage feedback from collector to base.

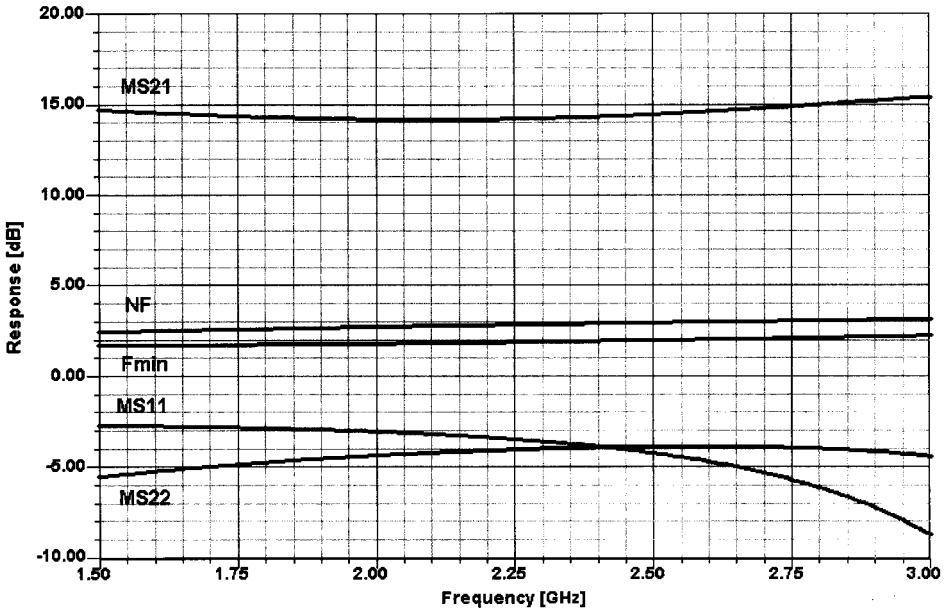


Figure 3-79 Frequency-dependent gain, matching, and noise performance of the BFP420 feedback amplifier.

Narrowband BFP420 Amplifier. Our next attempt will be to design a narrowband input stage that combines very high selectivity with good noise figure. This can only be achieved by using two tuned single-stage resonators with capacitive/magnetic coupling. The following amplifier is tailored to have a center frequency of 900 MHz. The transistor will operate at 12 V/10 mA to have a sufficient dynamic range. The input selectivity is achieved by the use of two top-coupled tuned circuits, followed by a BFP420 (Figure 3-80). The first order of business is to determine the input tuned circuit. The tuning capacitance for the two tuned circuits is arbitrarily set to a manageable 2 pF (small values have too much tolerances and larger values will require smaller inductances, resulting in lower Q). The reactance of the first tuned circuit is

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi \times 9\text{E}8 \times 2\text{E}{-}12} = 88.42\ \Omega$$

The necessary inductance for resonance will be

$$\omega L = 88.42$$

or

$$L = 88.42/\omega = 88.42/(2\pi \times 9\text{E}8) = 15.64\ \text{nH}$$

We will make both tuned circuits identical, but with different loaded Qs. By multiplying 88.42 times the loaded Q (20 for the input resonator and 50 for the output resonator), the resulting parallel-resonant resistance is $88.42 \times 20 = 1768\ \Omega$ at 900 MHz (for the input resonator) and $88.42 \times 50 = 4421\ \Omega$ (for the output resonator). These values are achieved by loading the input with the 50-Ω source and by loading the output with the input of the transistor. Therefore, the input transformation ratio $m_1 = 1768.4/50 = 35.37$. At the output, the equivalent number $m_2 = 4421/50 = 88.42$. For the purpose of simulation, and because it is extremely difficult to find the physical location of the appropriate tap on an air-wound coil for the necessary impedance, we do this "on paper" by incorporating an ideal transformer. We are also going to show later the actual transmission-line-based filter and its modeling. This filter will have a higher insertion loss based on the PC-board material used.

Using our nonlinear simulator, in small-signal ac mode, we obtain the frequency response and noise figure shown in Figure 3-81. In actually building this circuit, there is another problem besides these transformers—the top-coupling capacitor. We have used a grounded-T configuration to obtain a practically realizable value of 36 pF, whose lead inductance will not influence the circuit. (The actual coupling capacitance value for top coupling without the T configuration would have 0.1 pF. We would like to see this value repeatably realized.)

If the actual noise would be calculated relative to the output tuned circuit, the equivalent noise resistor R_n of 5 Ω would now be multiplied by m_2, resulting in 442 Ω. Fortunately, the Serenade simulator has a highly accurate BJT noise model that is both bias and temperature dependent, so the simulator will give us the answer directly.

We have not put any matching circuit at the output and leave this task to the reader. In practice, the slight tilt in the MS_{21} curve in Figure 3-81 would be tuned out. Also note that the minimum noise figure ($F_{\min}$) and the actual 50-Ω noise figure (NF) agree somewhat off center. This has to do with the Miller effect, which results in the well-known difference between noise matching and input-gain matching. By varying the coupling capacitor to move from the overcoupled response, which gives a low noise figure, to a critically coupled

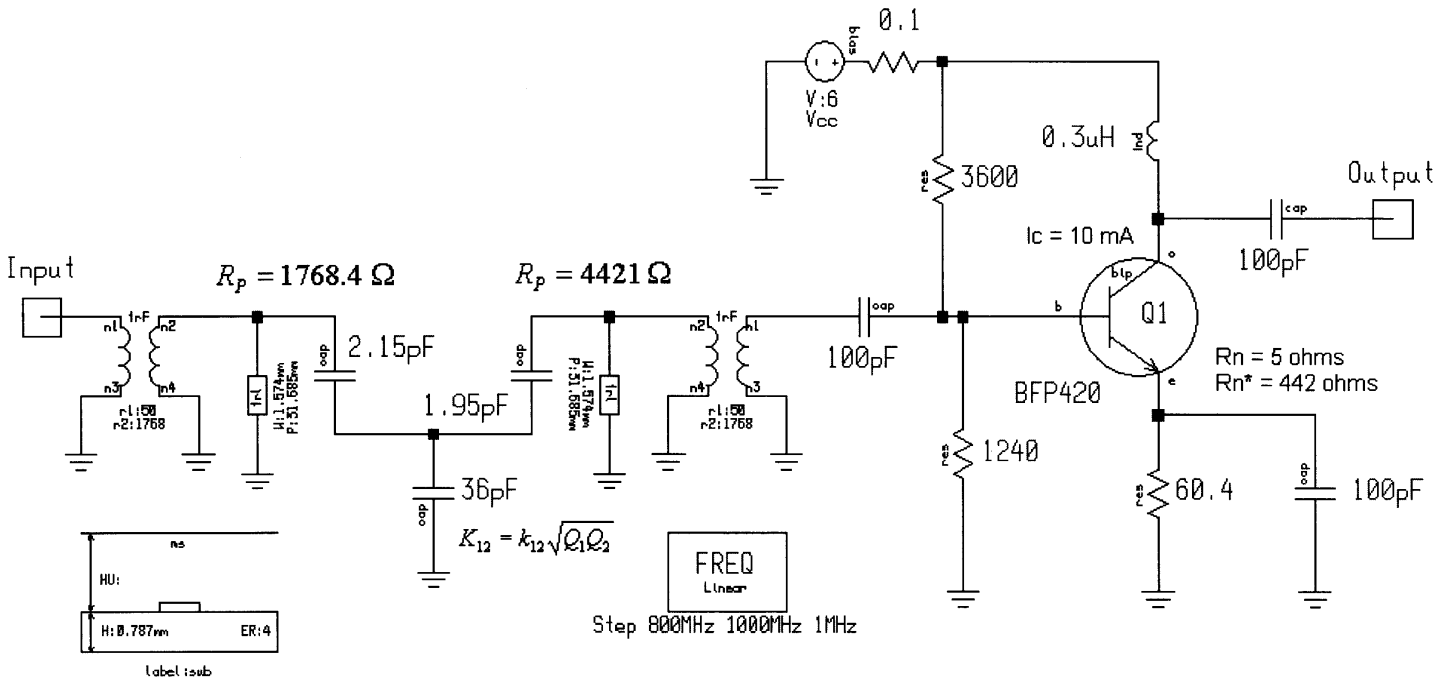


Figure 3-80 BFP420 amplifier with narrowband input filtering. Ideal transformers are used to match the filter's input and output to the input port and the BFP420 base, respectively.

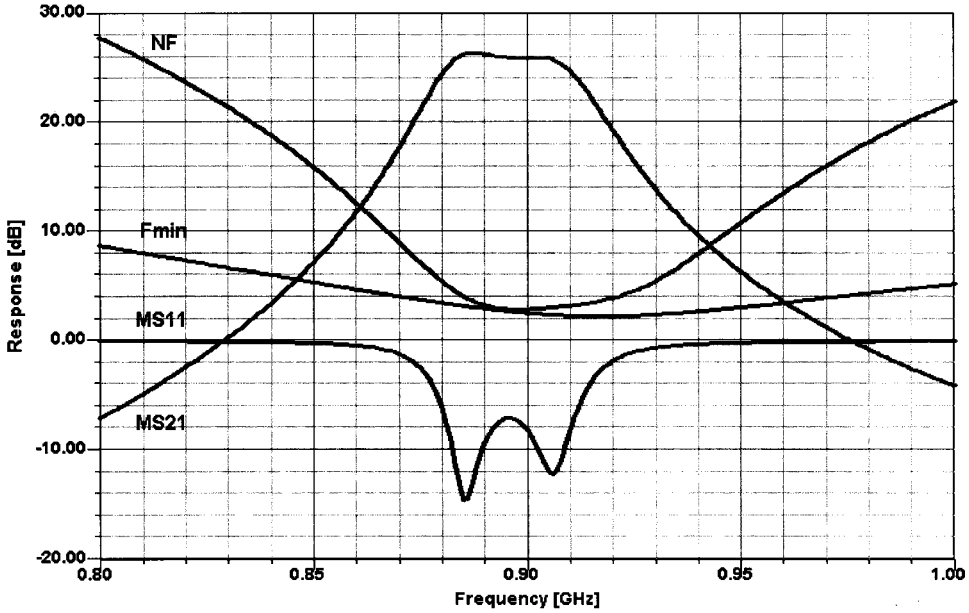


Figure 3-81 Frequency-dependent gain, matching, and noise performance of the narrowband BFP420 amplifier.

response ($K = 1$), we get significantly better selectivity and narrow bandwidth (at the obvious expense of the noise figure) as shown in Figure 3-82.

Since we somewhat compromised by not providing the reader with an actual input and output coupling that can be built, and furthermore, because the inductance in question (15.26 nH) is hard to realize with a high Q, we now show the printed circuit approach for the input filter. By using a Teflon-based material for the PC board, the losses would be less; however, the radiation would be more because ε would go down to 2.1 from the value we used (4). It is also convenient to introduce this filter to show that at higher frequencies, where lumped elements are really no longer available, one has to switch to distributed elements. A list of the commonly used distributed elements will be shown later; the transmission line and T junction are probably the most frequently used elements, followed by spiral and rectangular inductors, and bends.

The actual implementation of the input filter on a printed circuit board requires distributed elements (Figure 3-83). In practice, one has two parallel coupled lines with a tap somewhere toward the ground connection and, for modeling accuracy, we have to introduce a T junction with the appropriate impedance of 50 Ω for the input and a different impedance (if necessary) for the output. The coupling mechanism now is magnetic coupling, since no one can physically change the distance between the transmission lines by tuning some mechanical elements, one has to get it right the first time. This can only be achieved with a high-precision simulator. The reason why these simulators are so expensive is the investment of research time and personnel to produce valid models over a huge frequency and impedance range. Since we are now curious about the frequency response, we have plotted in Figure 3-84 the frequency response, which shows a higher (expected) insertion loss. The PC-board material is partially responsible for this.

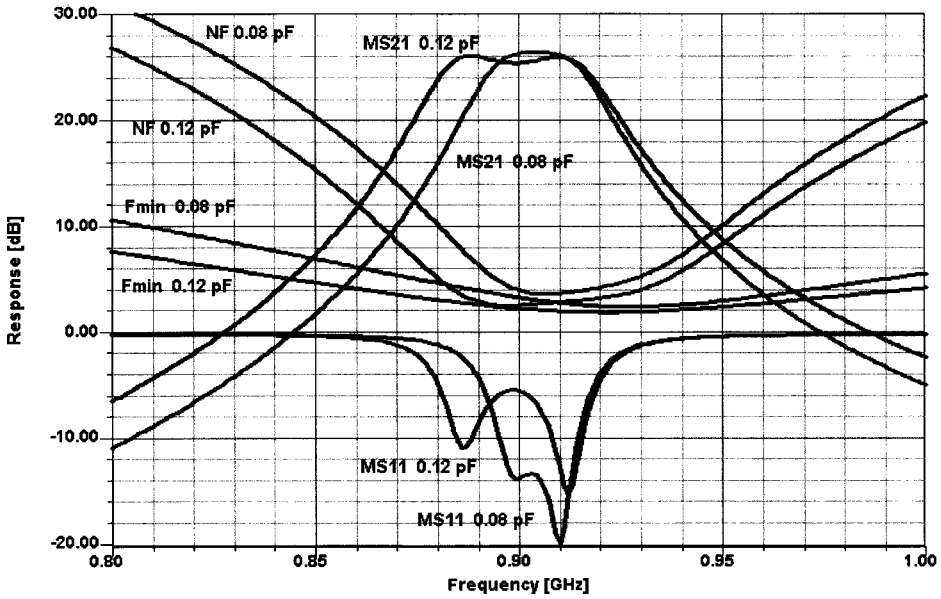


Figure 3-82 Comparison of filter responses with critical coupling (0.08-pF response) and overcoupling (0.12-pF response).

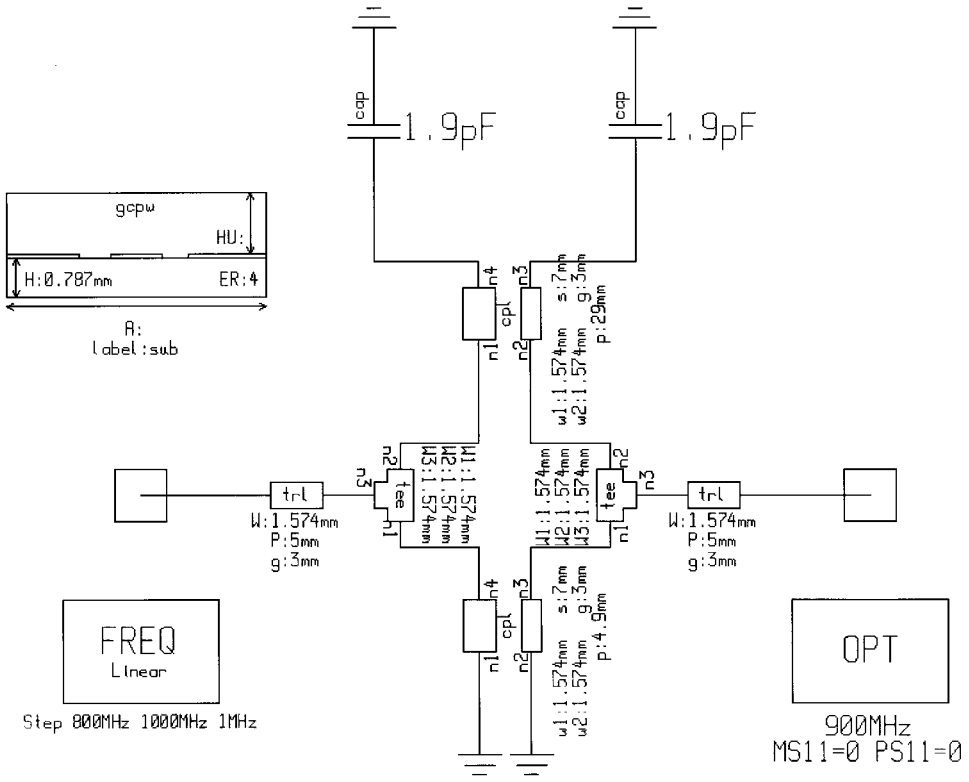


Figure 3-83 Input filter using distributed-element resonators and matching.

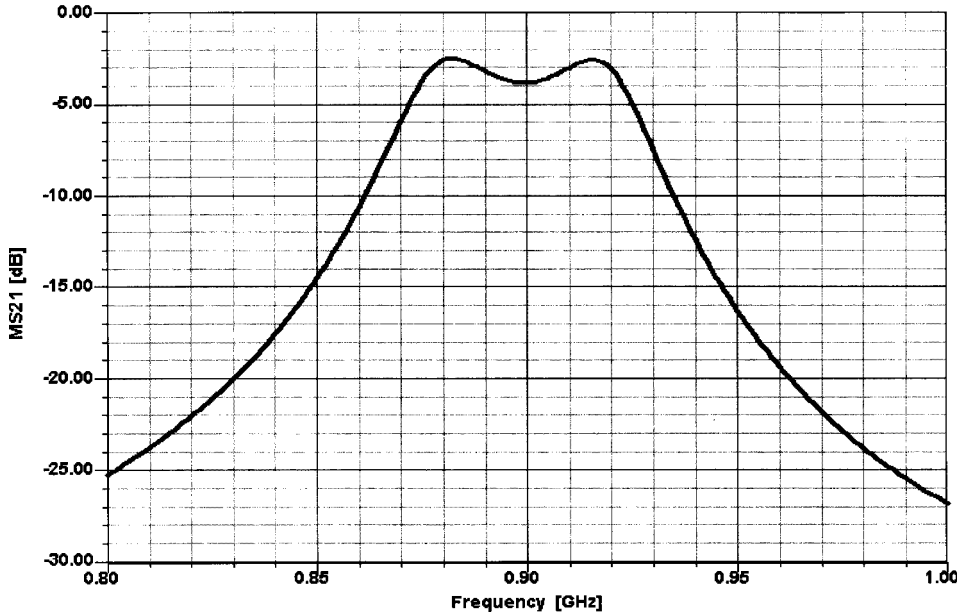


Figure 3-84 Gain versus frequency response for the distributed-element filter.

At the end of this chapter we will have a chance to discuss passive elements as needed for microwave frequencies, and one should consider frequencies above 1500 MHz as microwave frequencies.

GaAsFET Feedback Amplifier. The implementations we have shown so far have been based on bipolar transistors. Figure 3-85 shows a single-stage feedback amplifier using a GaAsFET. The feedback is based around an *LRC* network and some attempts at matching at the input and output are made. Figure 3-86 shows its frequency-dependent gain, matching, and noise performance.

3-2-3 High-Gain Amplifiers

In most cases, high-gain amplifiers receive an already amplified signal, which means we are less concerned about the signal-to-noise ratio but remain concerned about distortion. To reduce distortion, we have two options: (1) increase the dc current in the active device, which always improves the current-related distortion; and (2) add some voltage feedback, which reduces the voltage distortion. In simple terms again, only current feedback can reduce the current-related distortion of a square-law or exponential transfer characteristic and only voltage feedback can reduce the effect of overdrive as a function of the load, but only to some degree. Its main purpose remains to reduce the input impedance to useful values, specifically for FETs, since current feedback increases the input and output impedances. Some ultrasensitive circuits still require Class A operation, but now at higher current levels; these are typically CATV applications, and there are several transistors available that can be operated at 5–15 V but at 60 mA or higher. Because of the higher power consumption, this excludes handheld cordless or cellular telephones. The same circuit as used before (Figure 3-85) for

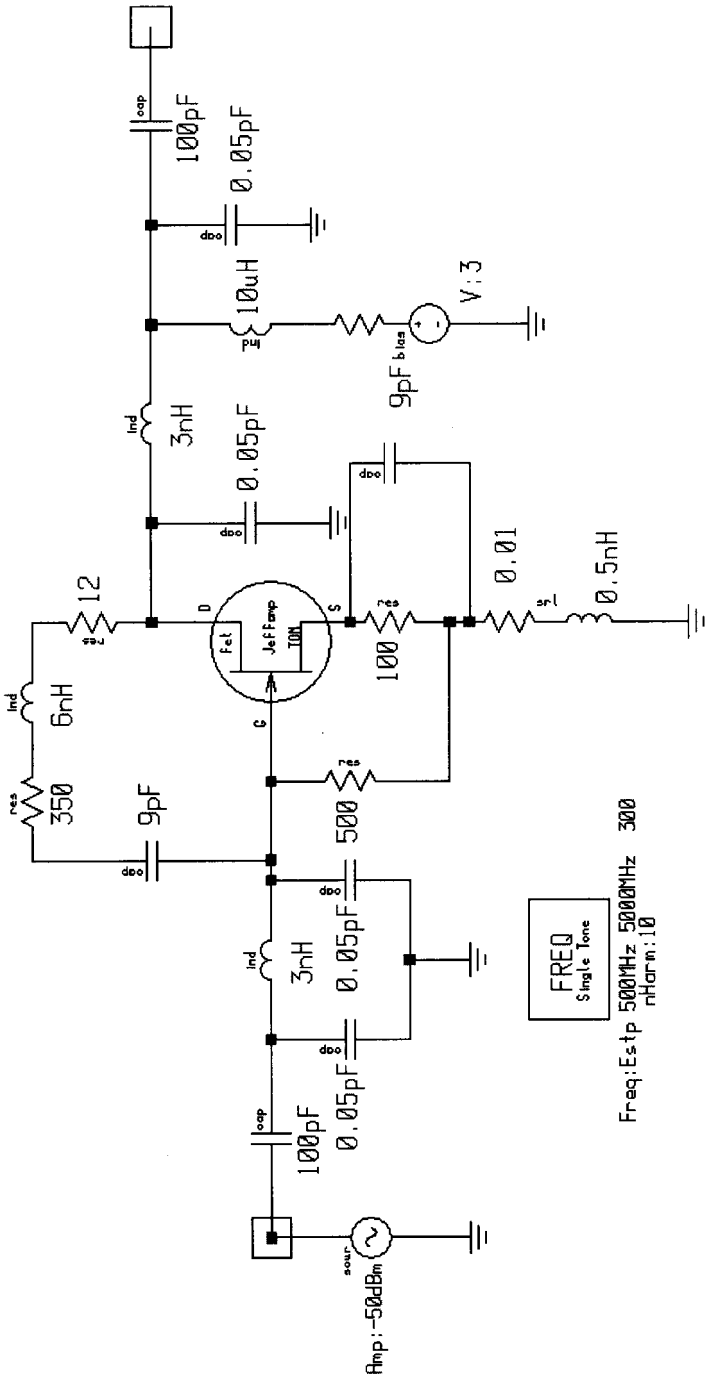


Figure 3-85 GaAsFET feedback amplifier.

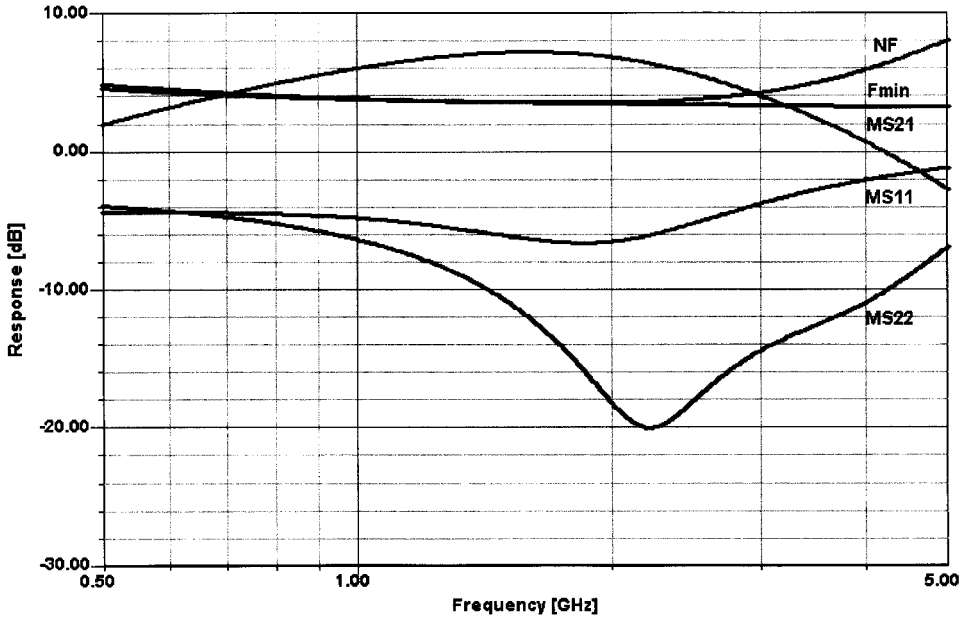


Figure 3-86 Frequency-dependent gain, matching, and noise responses for the GaAsFET amplifier.

Class A operation now can be operated at a higher current, which automatically increases the gain and the noise figure and reduces distortion. This is based on the fact that the transconductance (g_m) increases at higher currents. The second transistor in the TI circuit achieves this by being in reality a larger device, which could be modeled by several transistors, like the first stage, in parallel. This automatically increases the dc current by the scaling factor. The same is even more true for the output stage. Likewise, bipolar transistors can also be scaled. A visual inspection of some transistors in their packages will reveal that they are really not just one transistor, but several operating in parallel, each with its own ballast resistors (for bipolar transistors) or other means (for FETs).

Using tuned circuits rather than wideband matching, it is possible to go to Class AB or B operation. Class AB is really a hybrid between A and B, and Class B is probably better known. Figure 3-63 shows the Class B operating point. Most Class B stages are operated in push–pull because, together with the tuned circuit, one-half of the cycle is supplied by the other transistor. The definition of Class B operation is a conducting angle of 90° (see Figure 3-63). At this point it should be noted that some older texts define the conducting angle as being twice the value we have defined in this chapter.

Without going into further details, the highest efficiency in Class B operation is 78.5%. Based on the fact that modern wireless applications generally use a hybrid of amplitude and angle modulation, pure Class C operation finds little application today because of the severe nonlinear distortion it introduces. A Class B stage has less "power gain" than a Class A stage but provides much more output power. We have already mentioned the problem associated with characterizing a Class A stage in terms of power gain. That's the only time Kirchhoff's equations are valid, meaning that the sums of all currents and voltages have to be zero. Because Kirchhoff's equations do not consider nonlinearities and harmonic contributions,

the very moment we move into considering medium- and higher-power amplifiers, the nonlinearities will dominate and the performance of the stage is best analyzed with a high-quality circuit simulator with the appropriate dynamic range and good modeling capability. Without question, this applies to Class C operation also. As pointed out previously, the high-gain amplifier is optimized for gain rather than noise figure and typically requires more collector or drain current. This can easily be seen by rebiasing the FET amplifier as shown above to a significantly higher value of 33 mA. This is accomplished by changing the source resistor down to 3 Ω. As a result of this, we obtain the gain, matching, and noise figure responses shown in Figure 3-87.

It may be a surprise to the reader that the actual noise figure now is less than that shown in Figure 3-86. This device is now operating at 33 mA; the reason for this has to do with the fact that if the transconductance is too low, the noise figure also increases, and the NF curve for the FET has a similar-looking current-dependent minimum—a curve that is frequently not supplied by the manufacturer. This amplifier, now operating at 33 mA, should have quite good large-signal capabilities; we will examine this shortly.

Another interesting piece of information is the display of the dc I–V curves with the "load line." Since we have reactances in the circuit, we do not see a line but rather an ellipse, indicating the storage of energy. (The actual load line would be a line through the two foci of the ellipse.) This curve allows us to see immediately that at this operating point the transistor is neither going into current saturation nor becoming voltage limiting. The input power can be increased further until such a condition exists. Figure 3-88 shows the dc I–V curve for the amplifier shown in Figure 3-85.

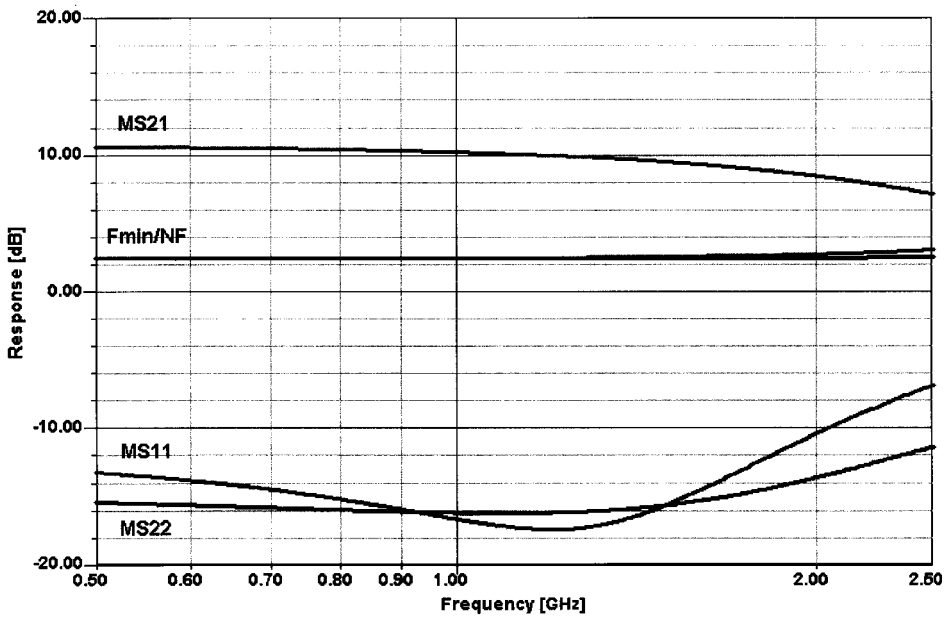


Figure 3-87 Frequency-dependent gain, matching, and noise performance for the GaAsFET amplifier of Figure 3-85 rebiased for a drain current of 33 mA.

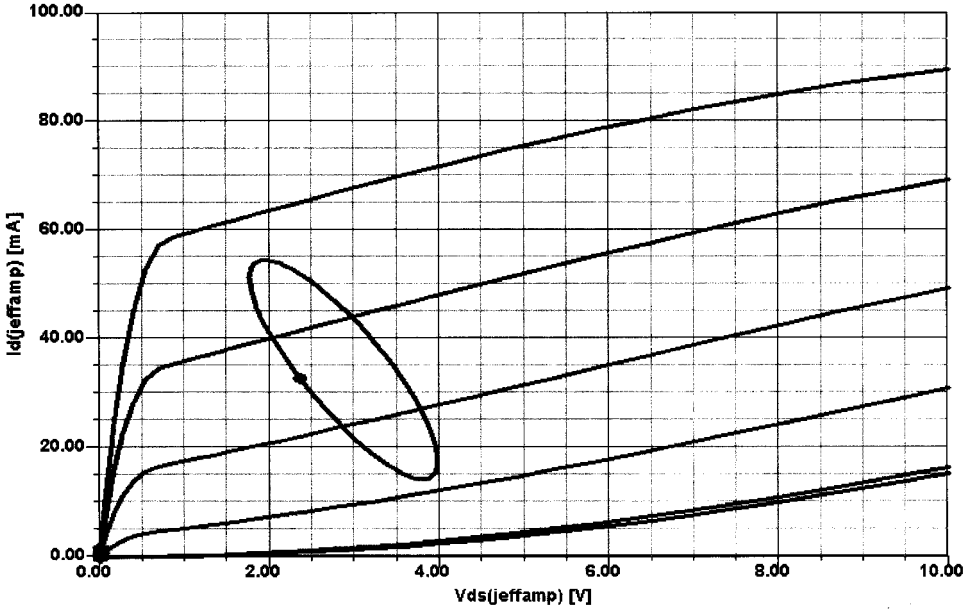


Figure 3-88 The dc *I–V* curves for the GaAsFET operating at I_D = 33 mA, including the ac load line. The load line is an ellipse rather than a straight line for the reason described in the text.

Besides looking for the adjacent-channel power ratio (ACPR), the other traditionally important points are third-order intercept point and the equivalent of a multitone simulation, as used for CATV application but relevant to the multisignal environment of wireless applications. Figure 3-89 shows the determination of the amplifier's third-order intercept point. It is most important for the reader to understand that very low values for input and output level have to be used to determine the crossover point. If this advice is not followed, totally erroneous numbers can occur.

Since an antenna supplies a large number of signals (this is true of handheld radios as well as base-station sites), a three-tone standard has been adopted as a good approximation of multisignal operation. This standard is in accordance with the German DIN 45004E,3.3 three-tone measurement. The requirement is that the intermodulation distortion products at the output are better than 60 dB suppressed relative to the largest of the three tones. One tone is used as the reference, and the other two tones are 6 dB less. This measurement has been derived from television, whereby the video transmitter, based on the vestigial sideband power, is 6 dB above the two sidebands produced by the FM sound transmitter using a modulation index of 5. The result of a three-tone test can be seen in Figure 3-90. This amplifier clearly meets the DIN specification for linearity at input levels up to –9/–15 dBm at the input.

Bipolar transistors are also quite popular for this purpose. Figure 3-91 shows a single-stage amplifier with an intrinsic transistor and all the other "hidden" intrinsics, normally not visible, added. The three diodes belong to those intrinsic elements, which are needed for accurate microwave modeling but which are not provided by the standard nonlinear models. The reader will see that we went into great detail to model all the necessary parts. This feedback amplifier, whose frequency response is shown in Figure 3-92, has been quite

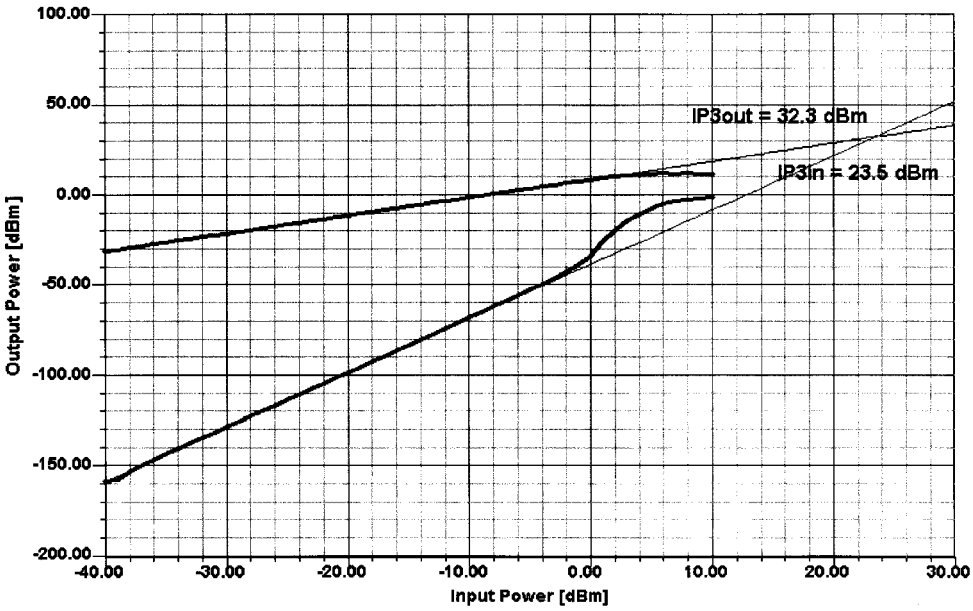


Figure 3-89 Determination of the third-order intercept point based on a two-tone measurement using tones spaced at 5% of the operating frequency (1.8 GHz).

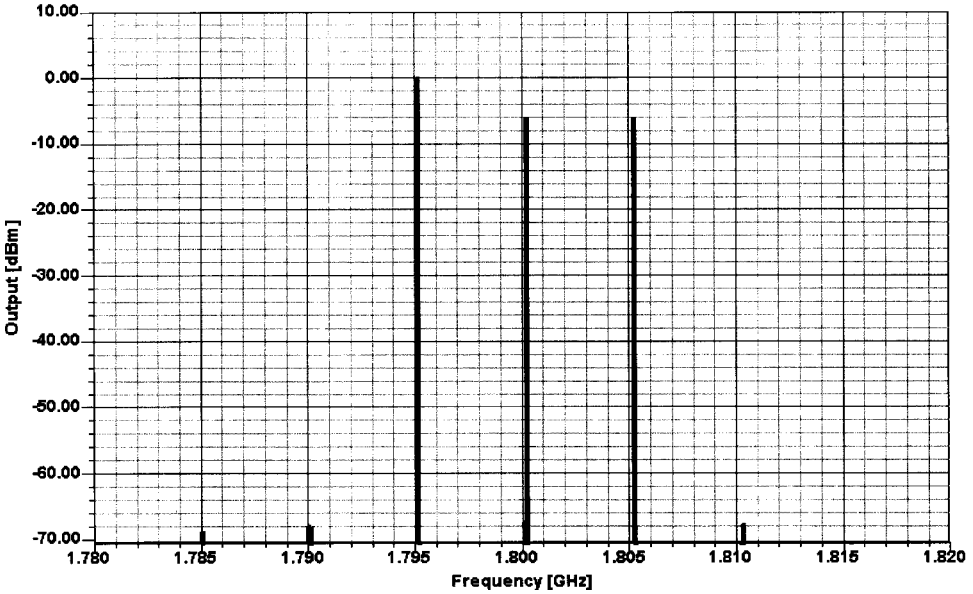


Figure 3-90 Simulated three-tone analysis of the GaAsFET amplifier with $I_D = 33$ mA.

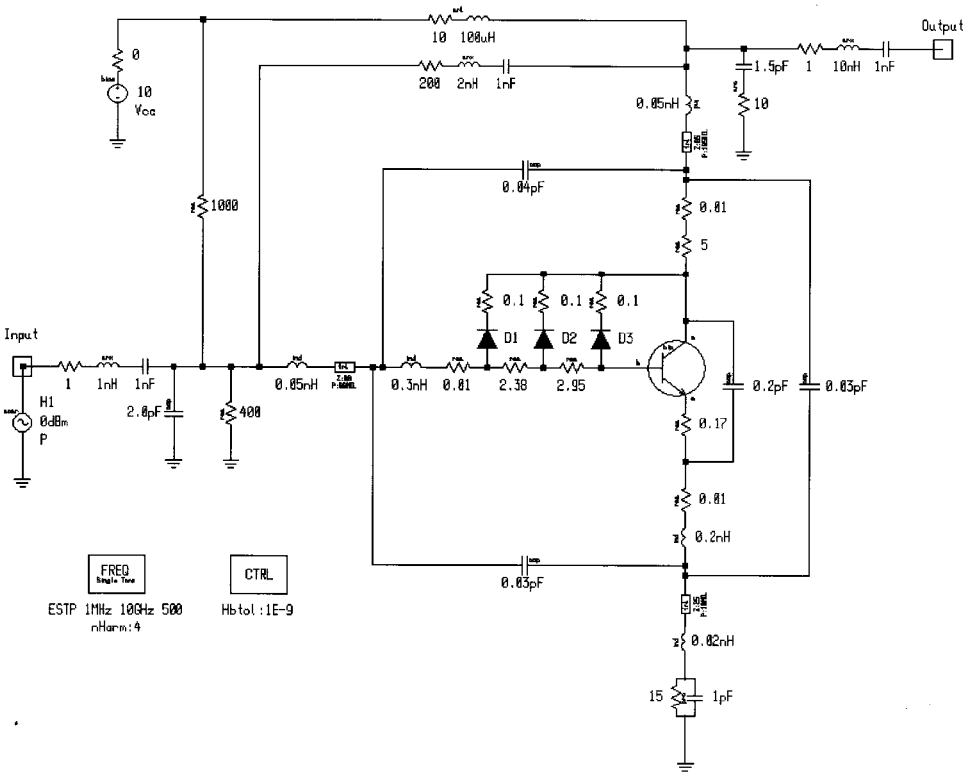


Figure 3-91 A high-linearity single-transistor amplifier modeled in great detail.

difficult to design, but there are monolithic circuits available where the higher gain–bandwidth product is being used to advantage.

Figure 3-92 shows the predicted frequency-dependent gain, matching, and noise figure at a bias point of 108 mA at 10 V. Using the same DIN three-tone test procedure, Figure 3-93 shows the expected good result.

Since this test requires two equal signals and one larger signal, we need to expect one IMD product larger than the rest. We have set our window for a dynamic range of 70 dB, and yes, in the lower left corner, we can see the product between the first two signals, while the other products are below the –50 dB level.

As a little task, Figure 3-94 shows a simple amplifier, based on a modern Siemens transistor, which is a feedback amplifier with fewer modeling details surrounding the transistor. Its associated frequency-dependent performance is shown in Figure 3-95. Interested readers are invited to improve its performance by using their skills and intuition.

Permitting higher currents and actually combining two transistors leads to our first MMIC. The following simulation is based on the MGA64135. Needless to say, there are some compensation networks inside the MMIC and distributed elements that affect the frequency response. However, with this arrangement, it is possible to obtain an appreciable gain at our wireless frequencies. The noise figure, based on the feedback, however, it not too impressive. This is a problem that occurs with most feedback amplifiers, which try to minimize the distortion and yet would like to keep the noise figure down. Figure 3-96 shows the schematic

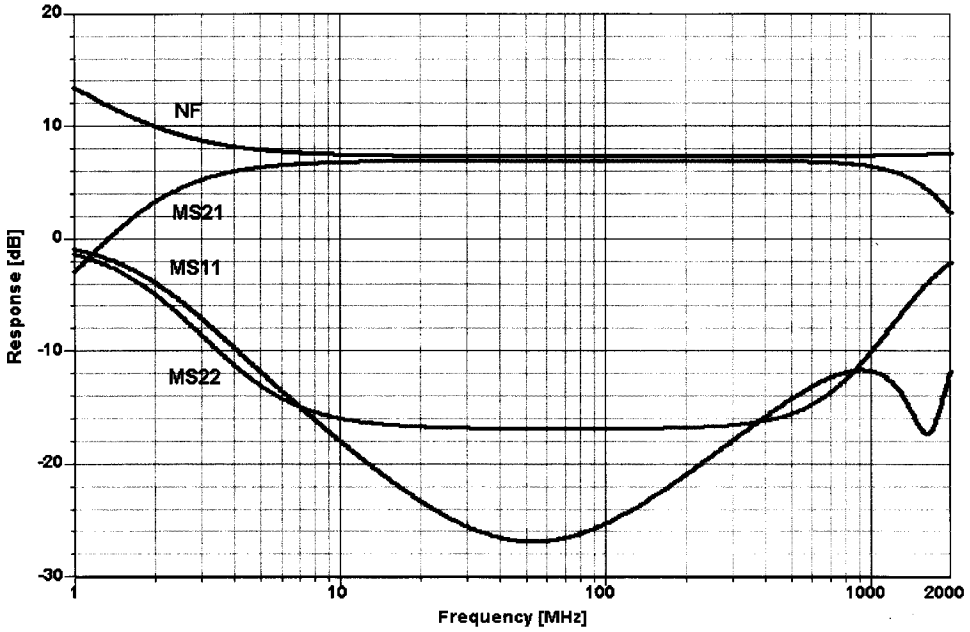


Figure 3-92 Simulated frequency-dependent gain, matching, and noise figure performance of the circuit in Figure 3-91.

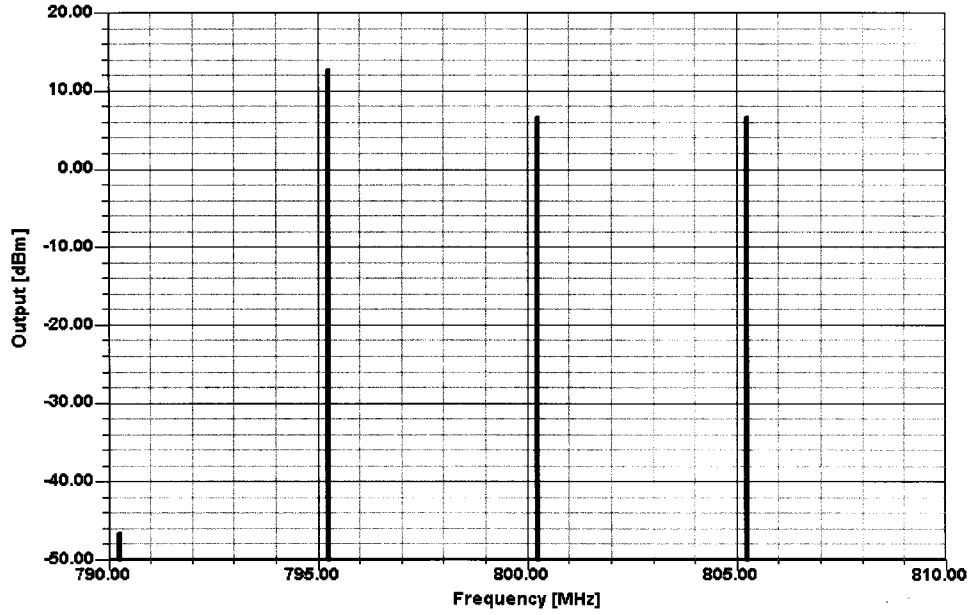


Figure 3-93 Simulated three-tone output spectrum of the amplifier in Figure 3-91.

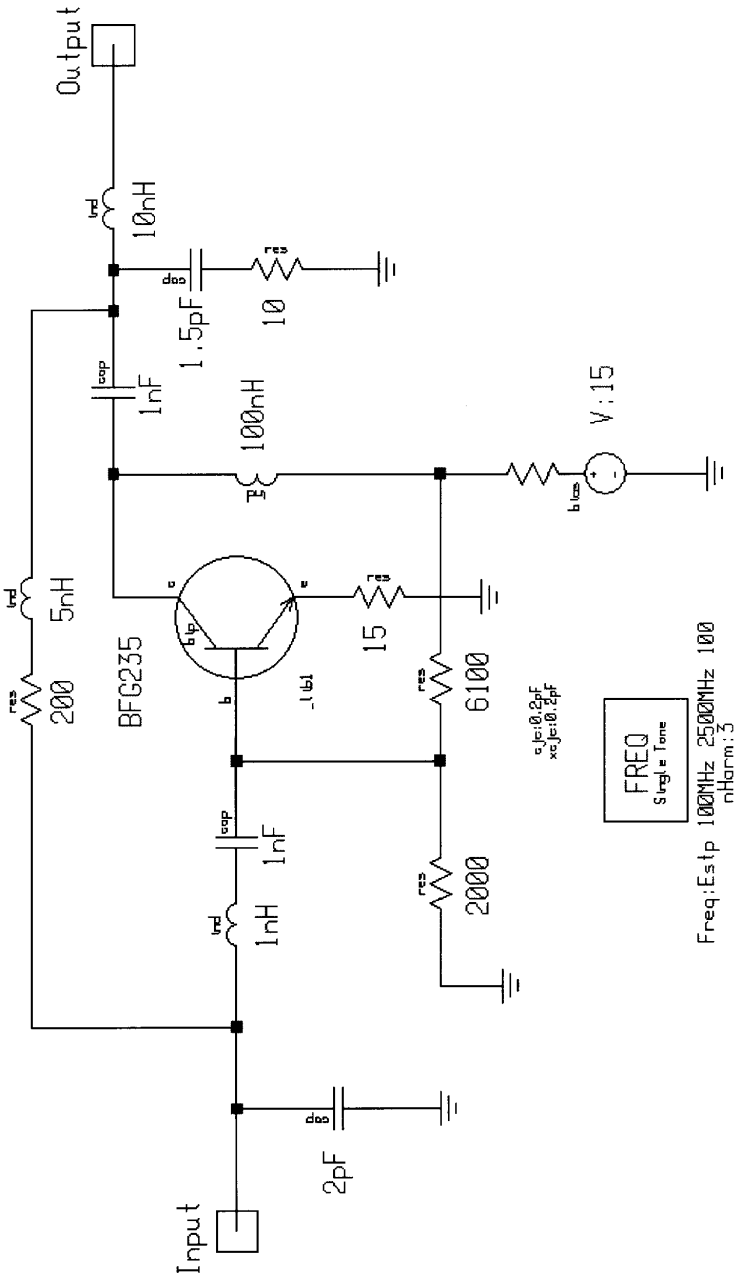


Figure 3-94 A high-linearity amplifier using the Siemens BFG235.

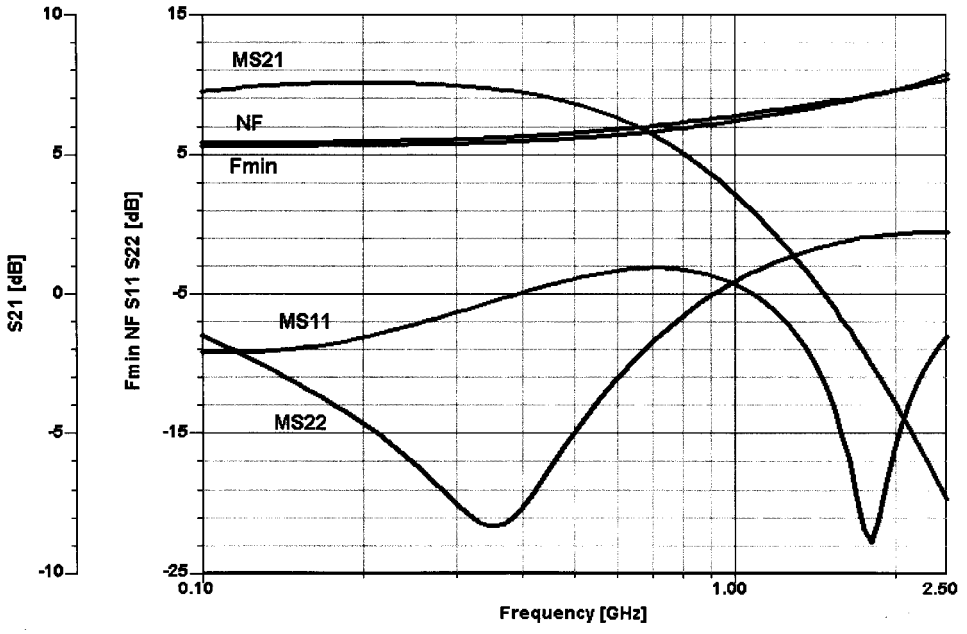


Figure 3-95 Frequency-dependent gain, matching, and noise performance of the amplifier in Figure 3-94.

of the GaAs MMIC; Figure 3-97 shows its frequency-dependent gain, matching, and noise performance; Figure 3-98 shows its two-tone response, which is used to obtain the intercept point ($IP_{3,in}$ = 8.5 dBm, $IP_{3,out}$ = 22 dBm); and, finally, Figure 3-99 shows the three-tone test results at an input RF level of –21/–27 dBm. While the gain–bandwidth product is considerably larger than the single-stage amplifier we examined before, the intercept point and three-tone performance do not reach the same values. As a general rule, the dynamic

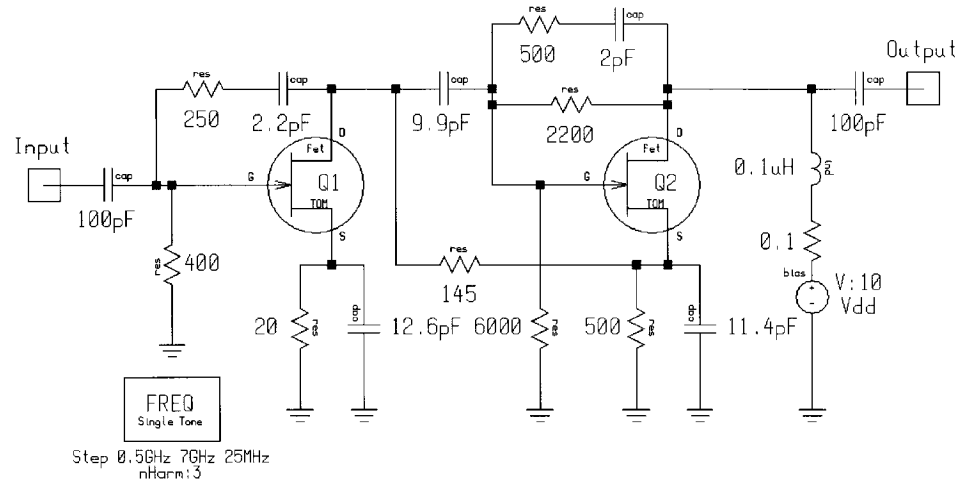


Figure 3-96 Simulation schematic for the MGA64135 MMIC amplifier. Q1 runs at 17.4 mA; Q2 at 32.2 mA.

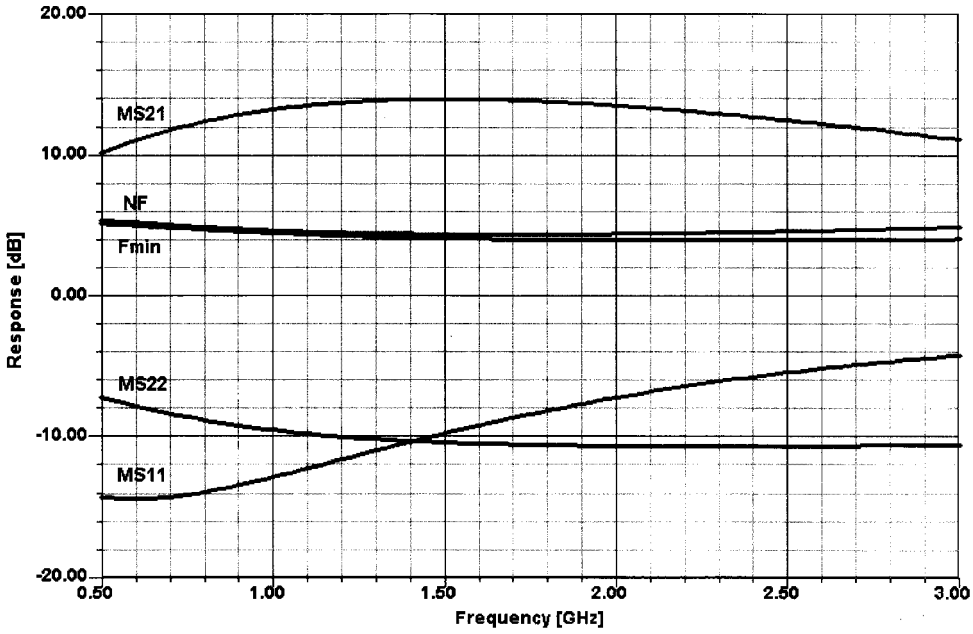


Figure 3-97 Frequency-dependent gain, matching, and noise figure performance for the MGA64135 MMIC amplifier.

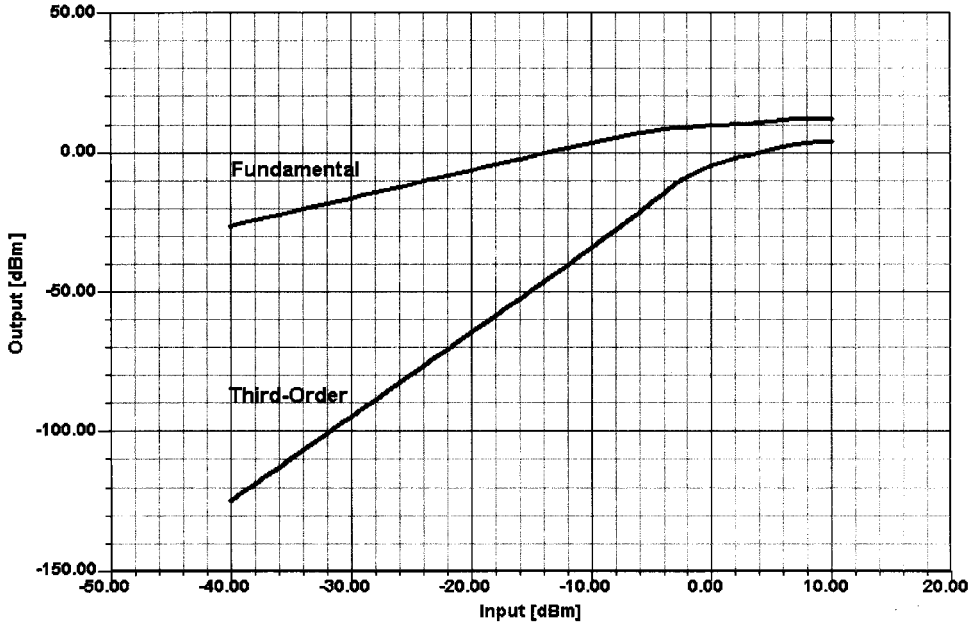


Figure 3-98 Simulated two-tone testing of the MGA64135 MMIC amplifier. $IP_{3,in}$ = 8.5 dBm and $IP_{3,out}$ = 22 dBm.

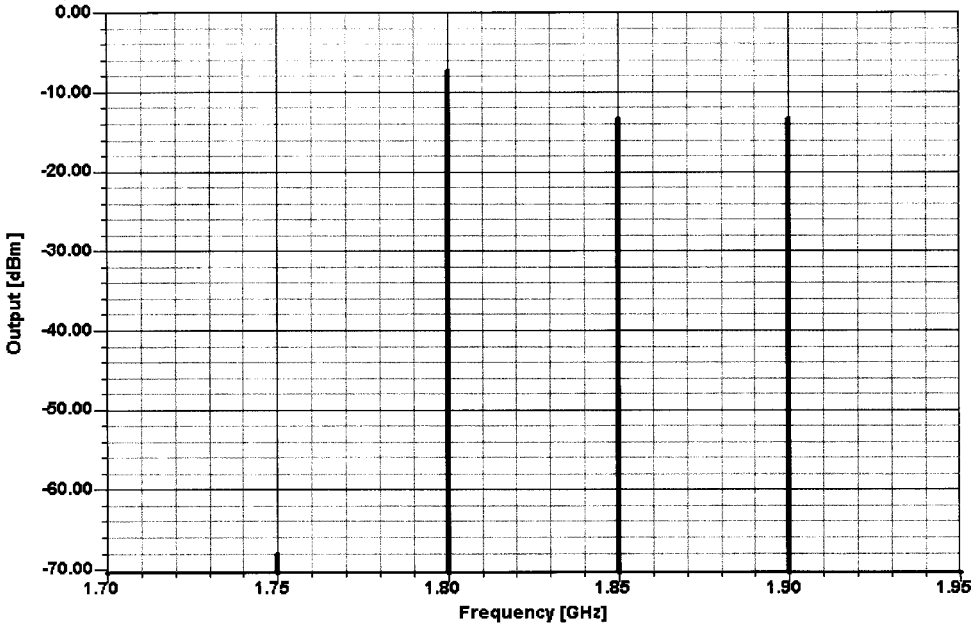


Figure 3-99 Simulated three-tone test of the MGA64135 MMIC. The test tones were 1.80 GHz (–21 dBm), 1.85 GHz (–27 dBm), and 1.90 GHz (–27 dBm). The 1.75-GHz IMD product is down 60.5 dB relative to the 1.80-GHz signal.

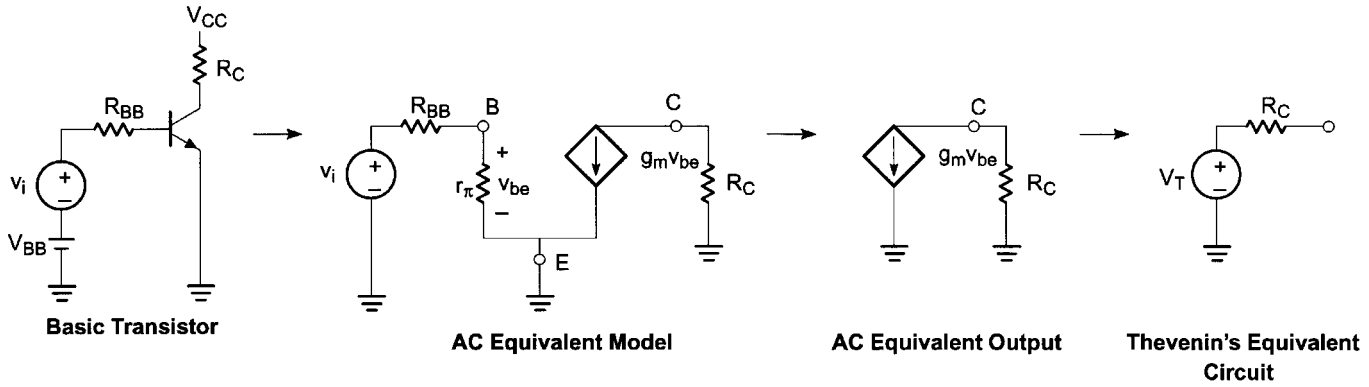
range is directly proportional to the power dissipated in the devices: specifically, the current. As explained earlier several times, the first-order nonlinearity comes from the voltage-to-current transfer characteristic, such as the transconductance itself, and can only be improved by a combination of resistive feedback in the emitter/source and by high currents. At these dc levels, the base–emitter (gate–source) junction gets linearized because the differential values are getting smaller compared to the parasitic loss resistances, including external dc/RF feedback. Once again, the voltage feedback reduces the voltage clipping but is initially more responsible for the input and output matching.

3-2-4 Low-Voltage Open-Collector Design [22–25]*

Why R_C Acts Like a Source Resistor. An open-collector output allows a designer the flexibility to choose the value of the R_C resistor. Choice of this resistor value not only sets the dc bias point of the device but also defines the source impedance value. Figure 3-100 shows the ac model of the transistor. Converting the output structure by applying a Norton and Thevenin transformation, one can conclude that R_C becomes the source resistance. Thus, by choosing R_C to be equal to the load, maximum power transfer will then occur.

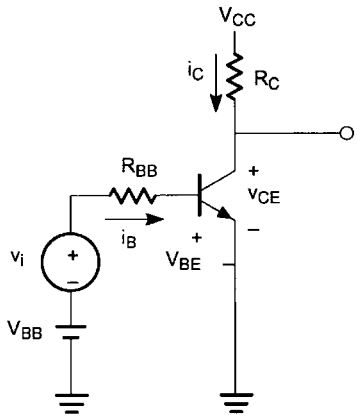
Figure 3-101 shows an active transistor with collector and base resistors. From basic transistor theory, Eq. A in Figure 3-101 is generated and has the same form as the general equation for a straight line ($y = mx + b$).

*Based on portions of the Philips Semiconductors/Signetics RF Communications Products Application Note AN1777, "Low-Voltage Front End Circuits: SA601, SA620," August 20, 1997. Used with permission.



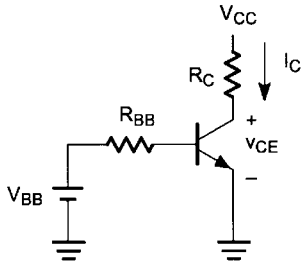
Symbol Convention:
DC: represented by uppercase symbols with uppercase subscripts, i.e., I_C
AC: represented by lowercase symbols with lowercase subscripts, i.e., i_C
Combined AC & DC: represented by lowercase symbol with uppercase subscript, i.e., i_C

Figure 3-100 *Rc* as source resistor.



DC Model

- Kill AC sources
- Open capacitors
- Short inductors

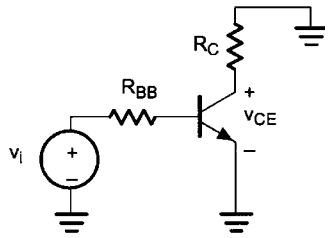


$$I_C = \left(\frac{-1}{R_C}\right) V_{CE} + \frac{V_{CC}}{R_C}$$

Eq. A

AC Model

- Kill DC sources
- Short capacitors
- Open inductors



$$i_C = \frac{-1}{R_C} (v_{CE})$$

Eq. B

Figure 3-101 Basic transistor analysis.

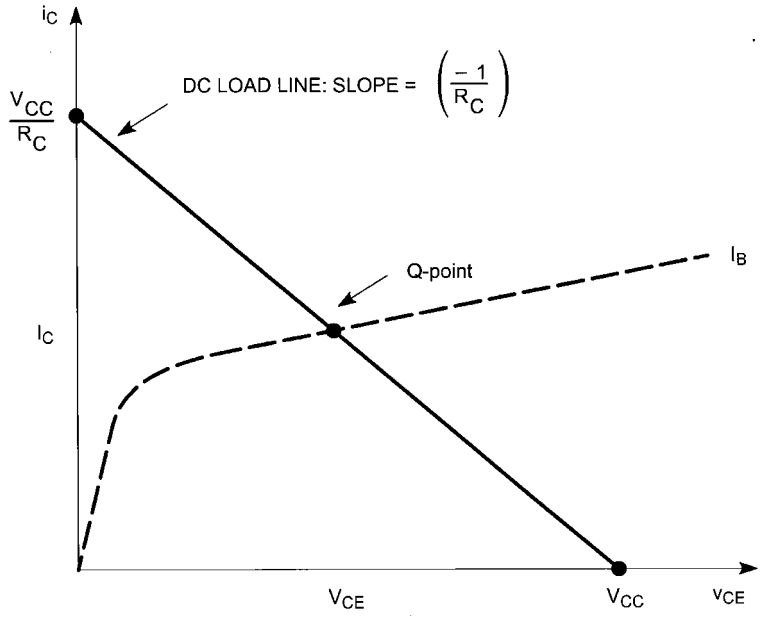


Figure 3-102 Load line and Q-point graph.

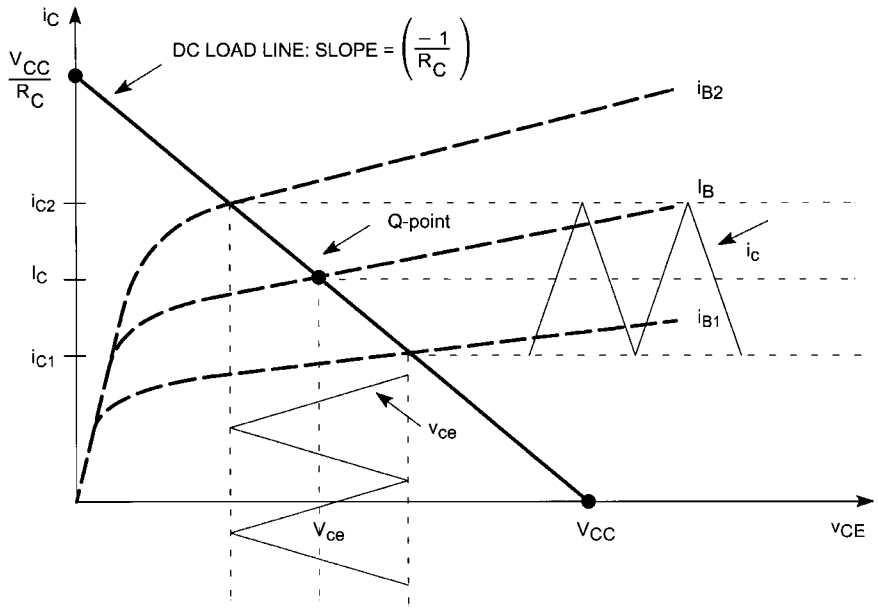


Figure 3-103 Graphical analysis for the circuitry in Figure 3-101.

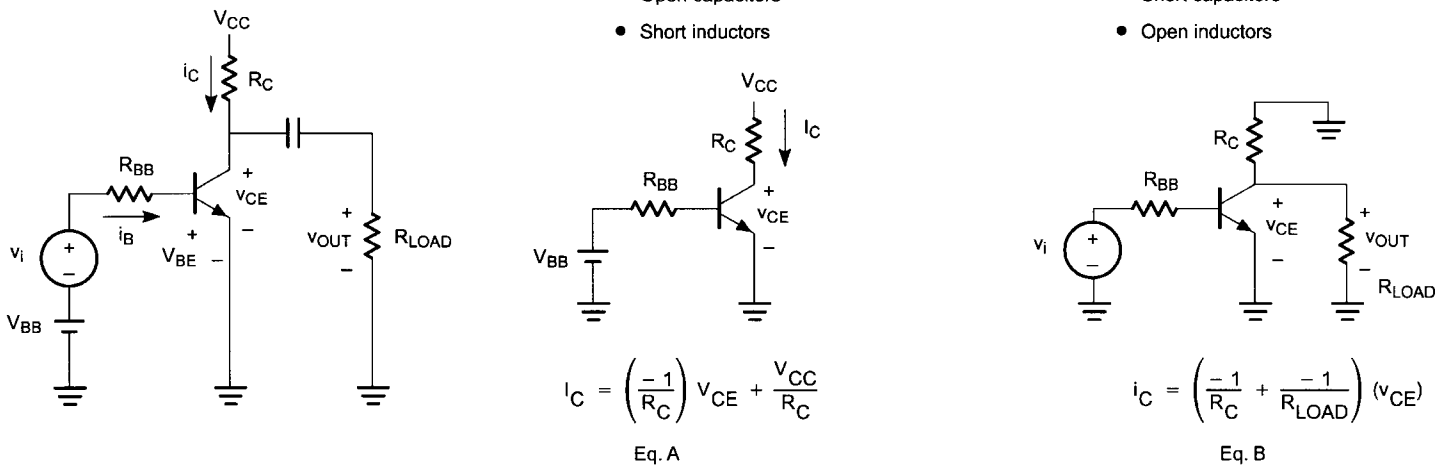


Figure 3-104 Basic transistor analysis with R_{LOAD}.

The slope of the dc load line is generated by the value of the collector resistor ($m = (1/R_C)$ and is shown in Figure 3-102. For a given small-signal base current, the collector current is shown by the dotted curve. The intersection of the dotted curve and the dc load line is called the *quiescent point* (Q-point) and dc bias determinant. The location of the Q-point is important because it determines whether the transistor is operated in the cutoff, active, or saturation region(s). In most cases, the Q-point should be in the active region because this is where the transistor acts like an amplifier.

Figure 3-103 shows the ac collector–emitter voltage (V_{CE}) output swing with respect to an ac collector current (I_C). Collector current is determined by the ac voltage presented to the input transistor's base (v_i) because it affects the base current (i_b), which then affects i_c. This is how v_i is amplified and seen at the output. Recall that this is with no external load (R_{LOAD}) present at the collector. Since no load is present, the slope of the ac load line is identical to that of the dc load line as seen in Eqs. A and B of Figure 3-101 ($m = -1/R_C$).

Open Collector with R_{LOAD}. A filter with some known input impedance is a typical load for the output of the transistor. For simplicity, we will assume a resistive load (R_{LOAD}) and neglect any reactance. Since a resistive load is used (see Figure 3-104), the ac output swing is measured at V_{OUT} or V_{CE}.

A dc blocking capacitor is used between the R_{LOAD} and the V_{CE} output to assure that the Q-point is not influenced by R_{LOAD}. It is also necessary to avoid passing dc to the load in applications where the load is a SAW filter. However, R_{LOAD} will affect the ac load line, which is seen in Eq. B in Figure 3-104. Note that the V_{CE} voltage swing is reduced; thus, the V_{OUT} signal is reduced (see Figure 3-105).

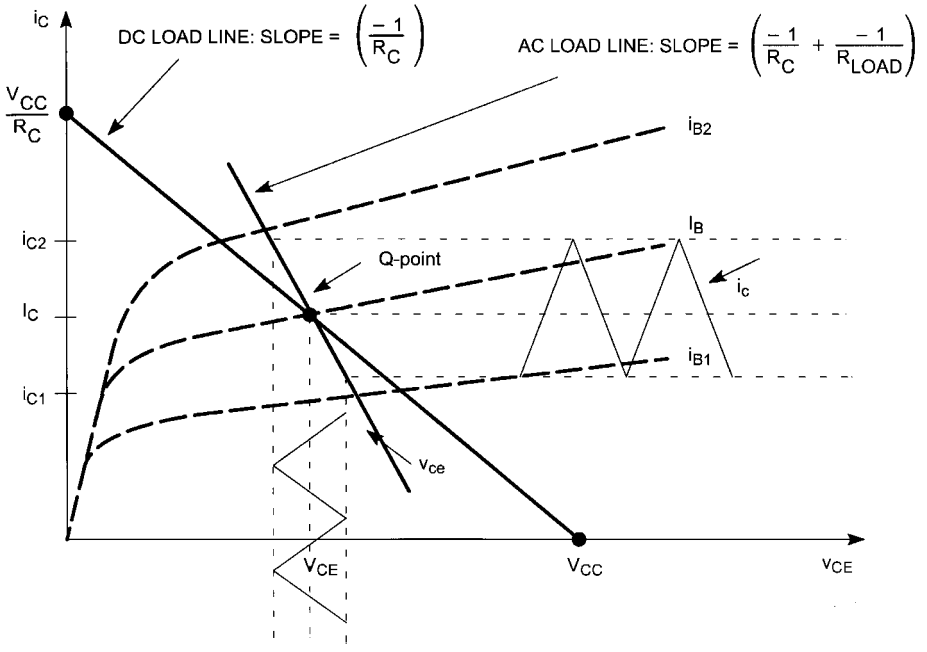


Figure 3-105 Graphical analysis for the circuitry in Figure 3-104. The ac load line intersects the Q point, and the slope is steeper than the dc load line due to R_{LOAD}, which makes V_{CE} smaller.

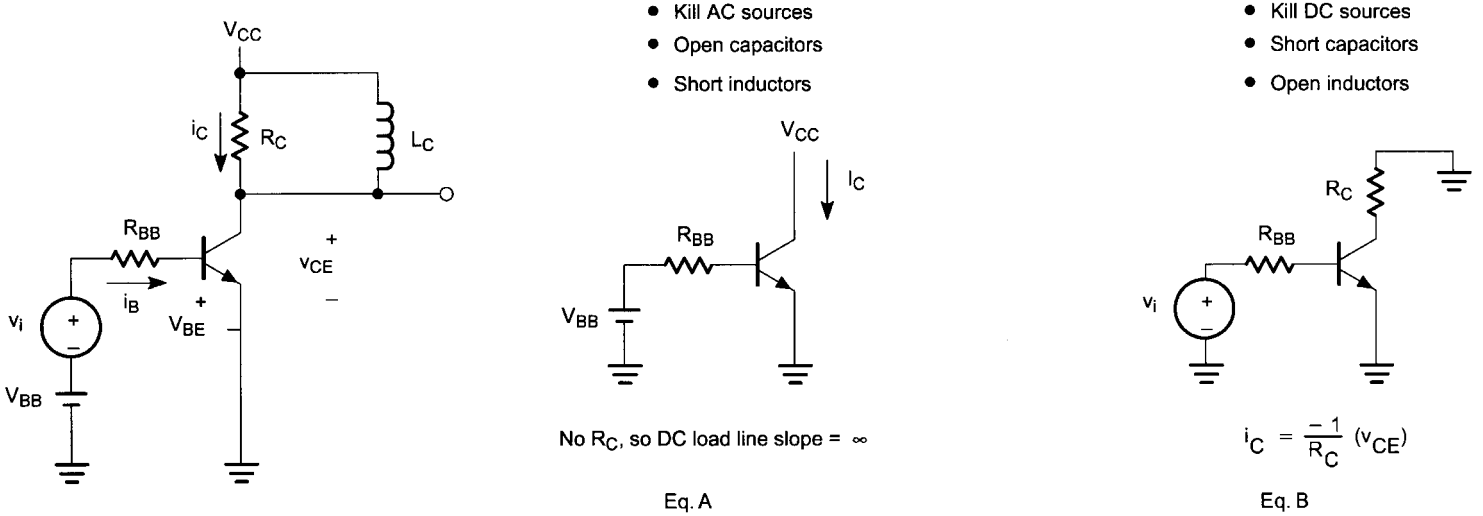


Figure 3-106 Basic transistor analysis with inductor added to collector.

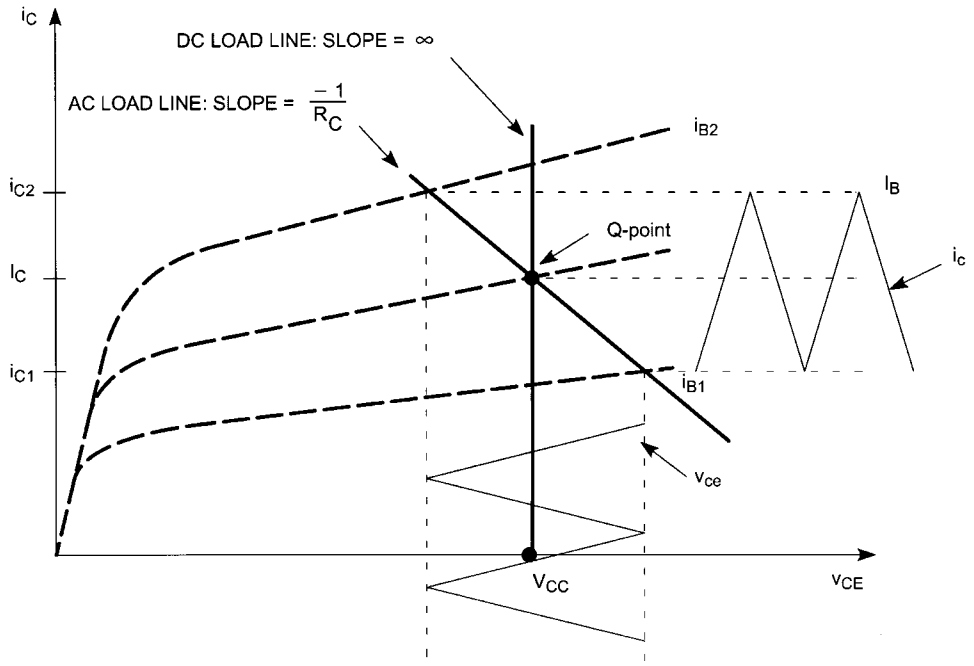


Figure 3-107 Graphical analysis for the circuitry in Figure 3-106.

Since the values of R_C and R_{LOAD} affect the ac load line slope, the value chosen is important. The higher the impedance of R_{LOAD} and R_C, the greater the ac output swing will be at the output, which means more conversion gain in a mixer. This is due to the slope getting flatter, thus allowing for more output swing.

Open Collector with Inductor (L_C). Adding an inductor in parallel with R_C can increase the ac output signal V_{CE}. Figure 3-106 shows the dc and ac analyses of this circuit configuration. In Eq. A in Figure 3-106, there is no R_C influence because the inductor acts like a short in the dc condition. This means the slope of the dc load is infinite and causes the Q-point to be centered around $V_{CC,}$ thus moving it to the right of the curve. The ac load line slope is set only by R_C because no load is present. Note that it has the same ac load line slope as the first condition in Eq. B of Figure 3-101.

Referring to Figure 3-107, one might notice that the base current (ac and dc) curves spread open as V_{CE} increases. This is caused by a noninfinite Early voltage (see Figure 3-108), which causes the collector current to be dependent on V_{CE}. Taking advantage of this nonideal condition, the peak-to-peak ac output swing V_{CE} can thereby be increased by moving the dc Q-point to the right due to the wider spreading between the curves corresponding to different base currents. Figure 3-109 combines Figures 3-103 and 3-107 to show the different ac output signals with different Q-points.

Looking at the ac output level, one might ask how the V_{CE} peak voltage can exceed the supply voltage V_{CC}. Recall that the inductor is an energy-storing device ($v = L\ di/dt$). Therefore, total instantaneous voltage is V_{CC} plus the voltage contribution of the inductor.

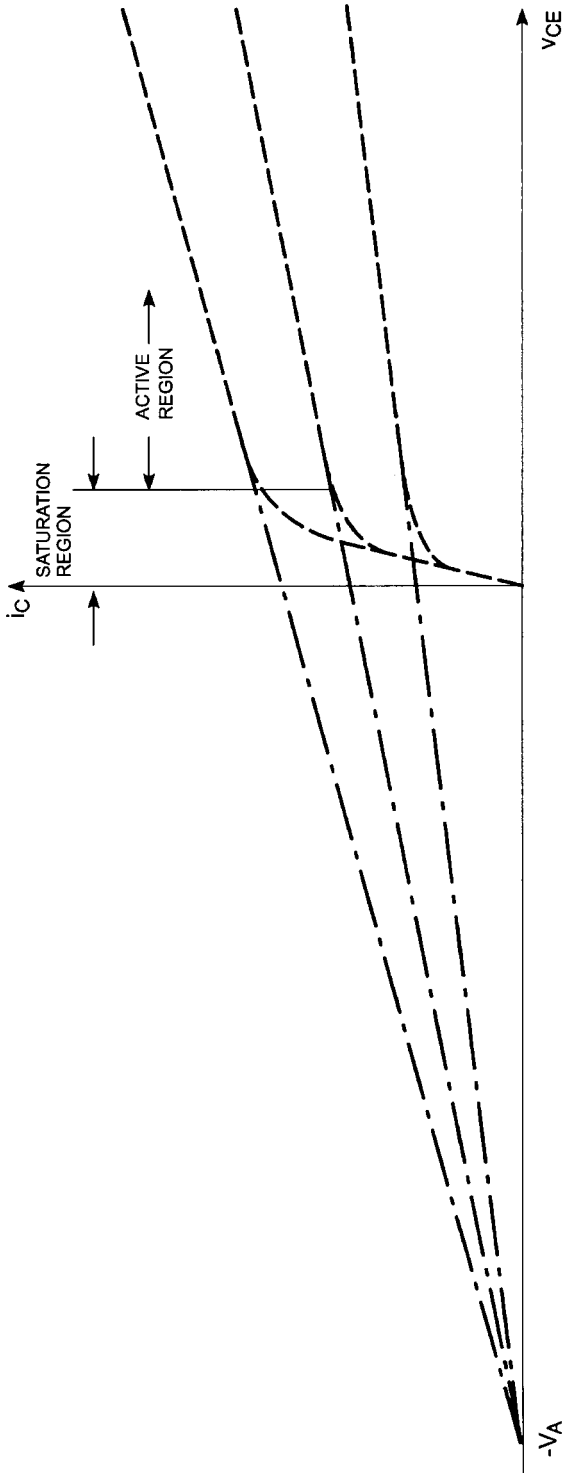


Figure 3-108 Graphical representation of Early voltage effect.

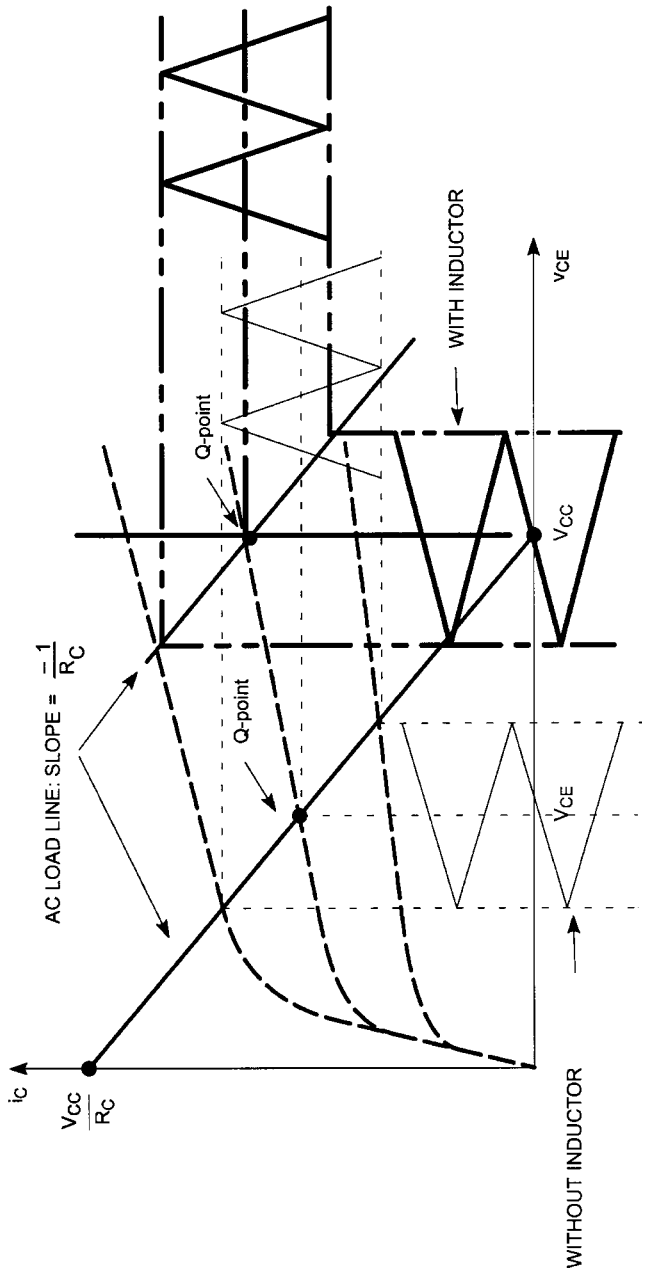


Figure 3-109 Comparison of open-collector circuit with inductor versus without inductor.

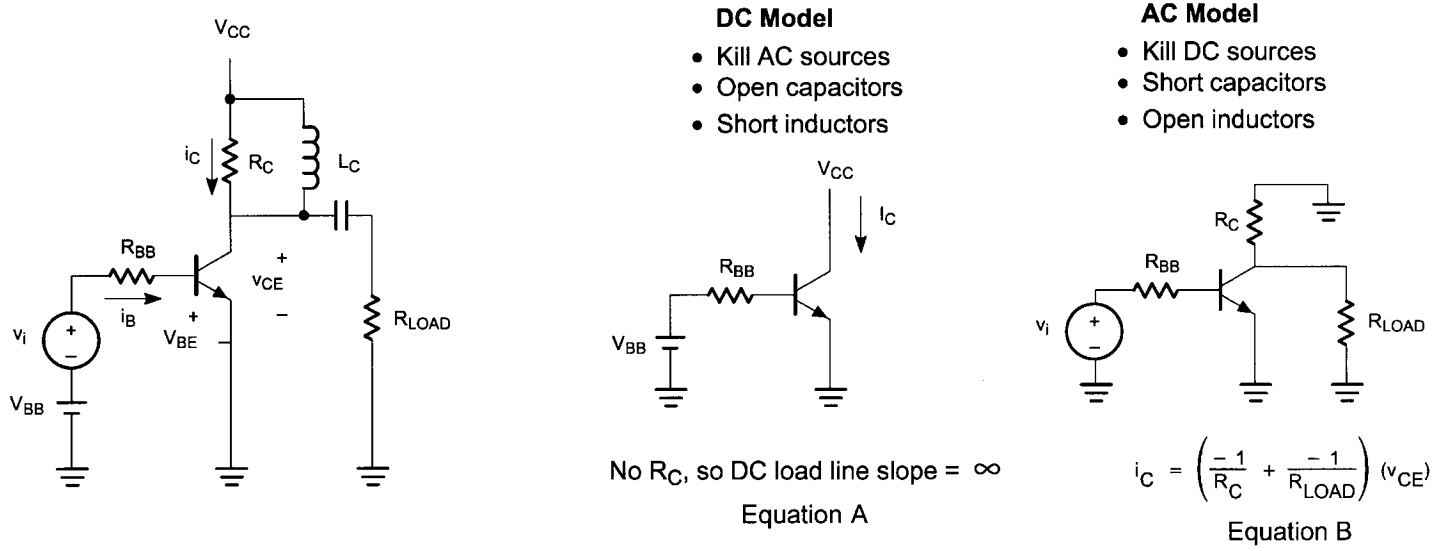


Figure 3-110 Basic transistor analysis with inductor and R_{LOAD}.

Open Collector with Inductor (L_C) and R_{LOAD}. Figure 3-110 shows the dc and ac analyses with the inductor and load resistor. Again, from the dc analysis, the inductor causes R_C to be nonexistent so the dc load line is vertical. In the ac analysis, the ac load line slope is influenced by both R_C and R_{LOAD} (see Eq. B in Figure 3-110). The ac load line slope is the same as the example of the open collector without the inductor. Figure 3-111 shows the response for the open collector with L_C and R_{LOAD}. Figure 3-112 combines Figures 3-105 and 3-111 showing the increase in ac output swing.

In conclusion, for the load line, R_C plays a role in setting up the bias-determined Q-point as well as the ac source impedance. However, when an inductor is placed in parallel with R_C, a different Q-point is set and the ac source impedance is altered. Moving the Q-point takes advantage of the transistor's nonideal i_c dependence on V_{CE} to get more signal output without having to change the base current. Since R_C is in parallel with R_{LOAD} in the ac condition, it influences the ac load line slope.

Flexible Matching Circuit [25]. A useful variation of the open-collector matching concepts previously outlined provided the capability of delivering equal power to two unequal resistive loads. This allows the power delivered to the load to be measured indirectly at another test point in the circuit where the impedance can be defined arbitrarily. If this impedance is defined to be 50 Ω, a spectrum analyzer can easily be placed directly into the circuit. This is an excellent troubleshooting technique and a valuable option to have available in high production environments.

Figure 3-113 shows the schematic for this flexible matching circuit. In this circuit, C_B functions only as a dc blocking capacitor and presents a negligible impedance at the frequency of interest. Recall from the previous open-collector matching discussions that,

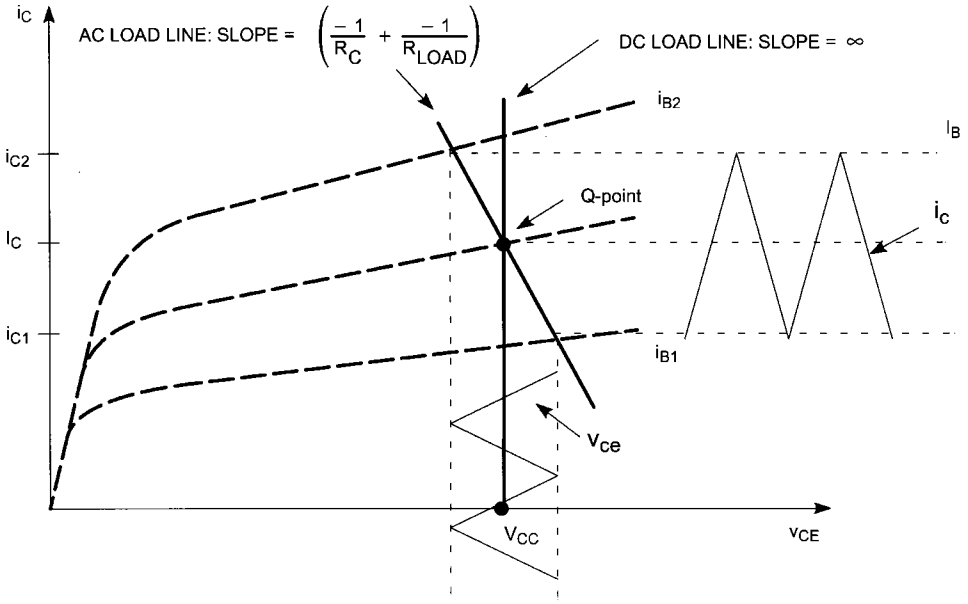


Figure 3-111 Graphical analysis for the circuitry in Figure 3-110.

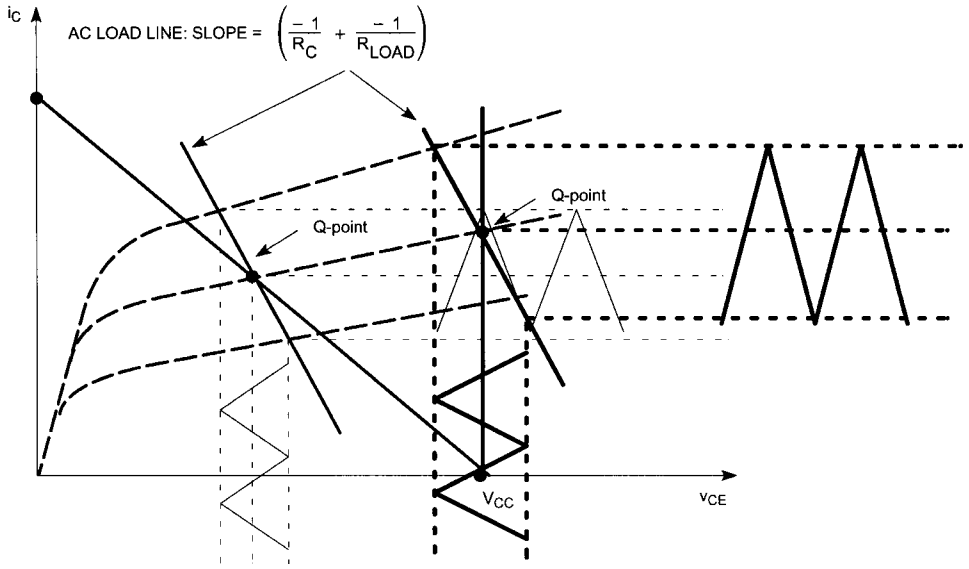


Figure 3-112 Comparison of open-collector R_{LOAD} circuit with inductor versus without inductor.

when R_C is placed in parallel with an inductor, it has no effect on the Q-point but does influence the slope of the ac load line as

$$\text{Slope} = \frac{-1}{R_C} + \frac{-1}{R_{\text{LOAD}}} \tag{3-166}$$

The capacitor C_S not only functions as a dc blocking capacitor but is also chosen such that the impedance presented by the combination of L, R_C, and C_S is equal to R_{LOAD} for optimum power transfer. The analysis is done in the following manner. First, note that inductor L is connected to V_{CC}, which is an affective ac ground. So L can be redrawn to ground. Next, R_C and C_S are converted to their parallel equivalent values as shown in Figure 3-114.

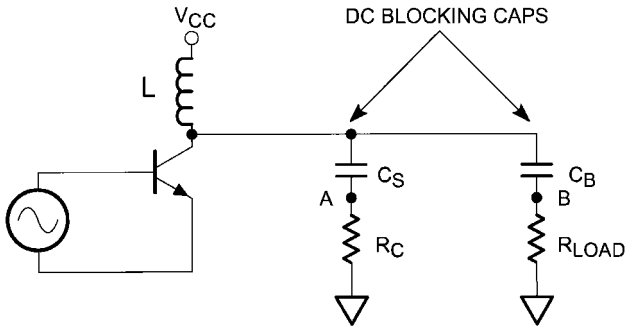


Figure 3-113 Flexible matching circuit.

Figure 3-114 Converting from series to parallel configuration.

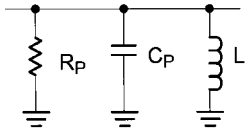


Figure 3-115 Converting from series to parallel configuration. Goal: Find L_C and C_p such that R_p looks like the R_{LOAD} value at the resonant frequency.

The resulting parallel *LCR* circuit is shown in Figure 3-115. At resonance, the parallel L–C_P combination will be an effective open circuit leaving only R_P. R_P is then simply chosen to be equal to R_{LOAD}.

Starting with the most simple transistor configuration and using *RC* types of dc-coupled amplifier circuits, we next present an overview of possible configurations for this application.

3-3 SINGLE-STAGE FEEDBACK AMPLIFIERS

This amplifier (Figure 3-116) is a standard minimal configuration amplifier with reduction of current distortion because of the emitter feedback. The voltage gain is

$$\text{VG} = \frac{R_L}{(26\text{mV})/I_C + R_e} \tag{3-167}$$

where R_L is the total load seen at the collector.

The feedback amplifier in Figure 3-117 uses voltage feedback to "correct" the input and output impedance to be close to 50 Ω but does not improve the current distortion. The voltage gain of this circuit is

$$\text{VG} = \frac{R_F}{R_g} \tag{3-168}$$

$$R_{\text{IN}} = R_{\text{OUT}}$$

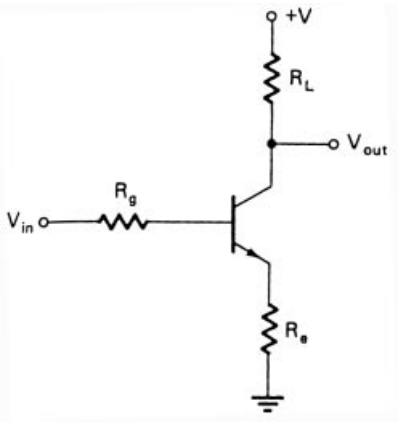


Figure 3-116 Single-stage amplifier with current feedback (provided by R_e).

$$= R_F R_D \quad \text{with} \quad R_D = \frac{26\text{mV}}{I_C} \tag{3-169}$$

Example. For $I_C = 10$ mA $\rightarrow R_D = 2.6\ \Omega$ and assuming $R_F = 200\ \Omega$, the input resistance $R_{\text{IN}} = 22.63\ \Omega$. The gain is calculated by getting the inverse of $R_D = g_m = 0.39$ A/V. Once this is multiplied with the output impedance of 22.63 Ω, the resulting gain becomes 10 log (8.6) or 9.3 dB. The gain can be increased by either using a larger collector current or a larger feedback resistor. It should be reminded that this circuit does not improve the distortion caused by the base–emitter junction.

To reduce the current distortion as well, we use the circuit shown in Figure 3-118, which contains both types of feedback. In reality, the collector–base resistor is frequently split into an unbypassed section responsible for the feedback and a larger resistor responsible for the appropriate biasing. The example circuit in Figure 3-119 shows this.

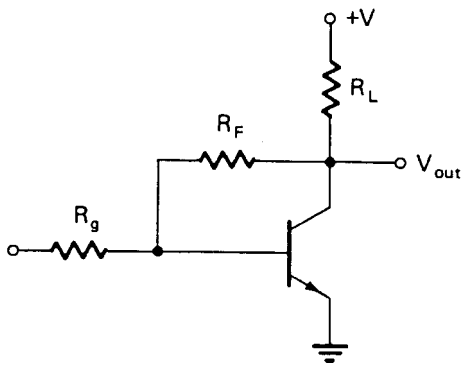


Figure 3-117 Single-stage amplifier with voltage feedback (provided by R_F).

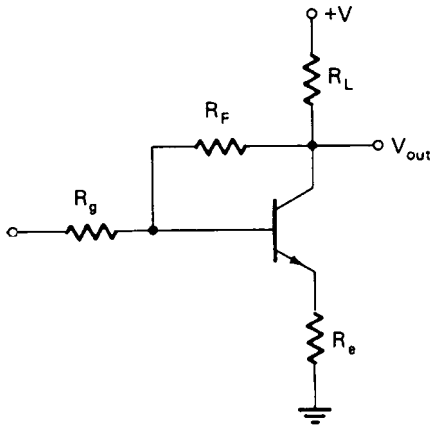


Figure 3-118 Single-stage amplifier with voltage and current feedback.

By using two of the fairly recent BFP420 transistors in parallel, and balancing the transistors by using the appropriate emitter resistors (which in total will result in 10 Ω) plus the bypassed value, we obtain the frequency-dependent gain, matching, and noise responses shown in Figure 3-120. To determine the appropriate feedback resistor R_F, we recommend the formula

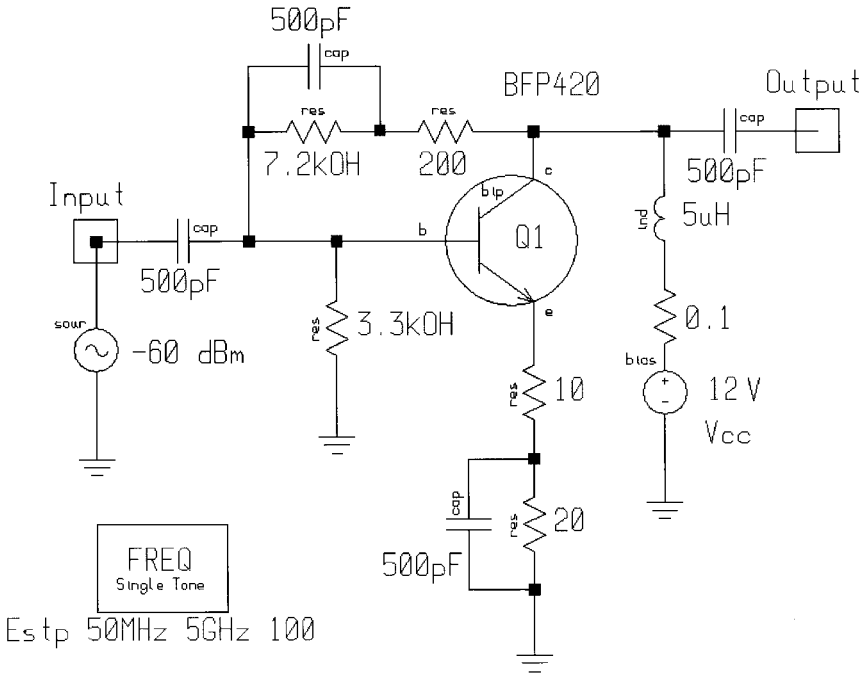


Figure 3-119 Practical amplifier circuit showing voltage and current feedback, and illustrating how reactances (in this case, 500-pF bypass capacitors) and multiple resistors can be used to provide different ac and dc resistances for the paths from collector to base and emitter to ground.

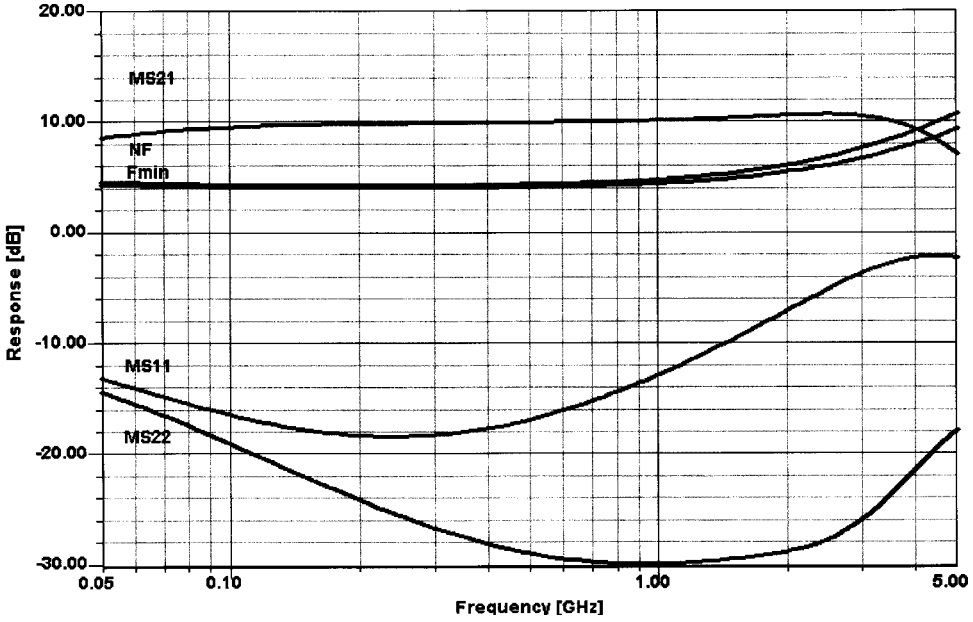


Figure 3-120 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-119.

$$R_F = \frac{(Z_{\text{in}}\, Z_{\text{out}})}{R_E} \tag{3-170}$$

This equation requires some additional comments. The combined shunt and series feedback only work as long as a sufficient transconductance is provided. In other words, there must be enough open-loop gain that will be reduced to obtain more bandwidth. A necessary condition for this approach, which really comes from the frequency range of less than 30 MHz, is that

$$Z_{\text{in}} = \frac{26\ \text{mV}}{I_C}\,\beta >> R_F \tag{3-171}$$

This requires a minimum transconductance of

$$g_m = \frac{1 + S_{21}}{Z_{\text{in}}} \tag{3-172}$$

where S_{21} must be a negative number that reflects the 180° phase shift. Of course, life becomes interesting when S_{21} goes through 1, but it (hopefully) never becomes negative and larger than 1 at the same time.

Example. The BFP420 has been used here frequently. A 10-dB gain would require an S_{21} of 3.16 ($\sqrt{10}$); therefore, we need to inspect the datasheet, which shows that the magnitude

of S_{21} is 3.16 around 4.5 GHz, yet the phase shift has changed from 180° to around 20°. In order to design an amplifier, we now set

$$R_F = Z_{\text{in}}(1 + |S_{21}|) \tag{3-173}$$

which is valid until S_{21} becomes too small. For a given 10 dB or 3.16 for S_{21}, the resulting RF is 4.16×50, or 208 Ω. The transconductance under these conditions becomes

$$g_m = \frac{1 + |S_{21}|}{Z_0} \tag{3-174}$$

where $Z_0 = 50$ Ω (the desired input impedance). g_m now becomes 83.2 mS. We now test the input and find

$$\frac{\beta}{g_m} = Z_{\text{in}} = \frac{50}{0.0832} = 600\ \Omega \tag{3-175}$$

Therefore, the emitter feedback resistor is not necessary to meet the previously mentioned condition whereby the input impedance needs to be higher than RF. Assuming the transconductance now is 0.00832, or 10% of the previous value, the input impedance would be 60 Ω, no longer meeting the prescribed requirement. Therefore, we need an unbypassed emitter resistor that fulfills the equation

$$R_E = \frac{Z_0^2}{R_F} - \frac{1}{g_{m_{\text{dc}}}} \tag{3-176}$$

where

$$g_{m_{\text{dc}}} = \frac{1 - S_{21_{\text{max}}}}{Z_0} \tag{3-177}$$

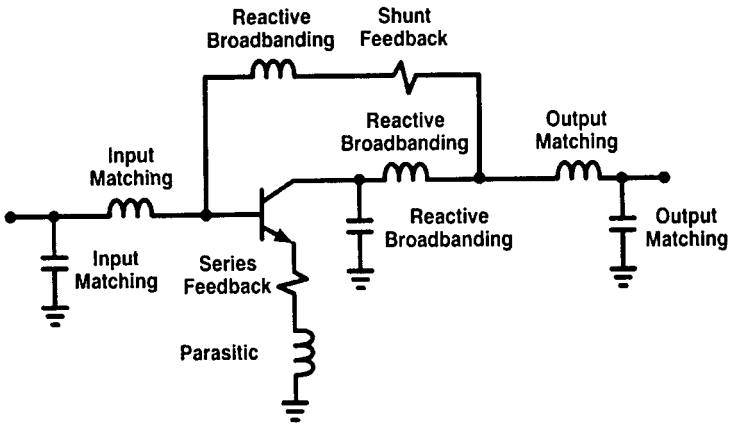


Figure 3-121 Elements of a feedback amplifier.

Again, the reader should be cautioned that in the common-emitter circuit, there is a 180° phase shift, and therefore $S_{21_{\max}}$ is a negative quantity, resulting in reality in

$$g_{m_{\text{dc}}} = \frac{1 + S_{21_{\max}}}{Z_0}$$

In our case,

$$g_{m_{\text{dc}}} = \frac{-S_{21}}{2Z_0} = \frac{38}{2 \times 50} = \frac{38}{100} = 0.38$$

so

$$R_E = \frac{50^2}{208} - \frac{1}{0.38} = 12 - 2.6 = 9.4\ \Omega \tag{3-178}$$

In practice, the designer would use a 10-Ω resistor.

Since the gain of the transistor continues to fall as a function of frequency, one can compensate to a degree for the cutoff frequency by putting an inductance in series that follows the equation

$$L_F = \frac{R_F}{2\pi f_{3\text{dB}}} \tag{3-179}$$

Assuming $f_{3\text{dB}}$ = 2 GHz, the resulting series inductance L_F = 16.6 nH. Basically, the same applies to circuits where the second transistor is used as an emitter follower to feed the signal back from the collector to the base of the first transistor.

In summary (see Figure 3-121):

- Resistive negative feedback at lower frequencies reduces open-loop gain and simultaneously matches input and output.
- Reactive elements set the gain and matching at high frequencies. Feedback is reduced at high frequencies.
- Emitter bypass capacitors are not used, as they would destroy the impedance match.

This is equally valid for FETs.

The next level of complexity can be reached by using two transistors with dc feedback.

3-3-1 Lossless or Noiseless Feedback

All the amplifiers shown above are based on a combination of resistive and reactive feedback. The very moment resistive feedback is used, the noise figure increases, as does the intercept point. As long as the intercept point increases overproportionally compared to the noise figure, this is an acceptable method. However, there are many cases in which an extremely low noise figure and a higher than normally achievable intercept point are required. To avoid this degradation of performance, Norton of Anzac developed a negative feedback structure based on the capabilities of the ferrite transformer. This patented technique (U.S. Patent Nos. 3,624,536 and 3,891,934, since expired), known as *lossless feedback*, provides lower noise figure and higher linear output than can be achieved using resistive feedback with the same transistor. Transformer feedback is used in which the transformer in the feedback network acts as a directional coupler. The coupling ratio between the input and output of the coupler

determines the magnitude of the feedback, and hence the gain. Across the frequency band over which the transformer operates as intended, the feedback is negative due to the in-phase coupling and the inversion of the active device.

This technique was used in amplifiers ranging from 5 to 500 MHz, 20 to 1000 MHz, and 300 to 1800 MHz. The difficulty associated with this type of amplifier is the fact that it is based on transformers that use twisted pairs of wires as transmission lines. In some implementations used at higher frequencies, actual directional couplers were used at the expense of bandwidth; in these circuits, the coupling was determined by the spacing of the coupler lines rather than transformer turns ratio. These microstrip couplers have much narrower bandwidths than transformers and are fairly large.

Another headache with the lossless feedback design is the fact that many of these transistor configurations are used in either grounded-base or grounded-emitter with transistors that have f_T values of 25 GHz and higher. Such devices are very hot, and achieving freedom from oscillation over a wide frequency range is a continuing problem, even for the circuits we have shown so far. A similar clever technique has been implemented in the termination-insensitive mixer, which we will examine in Chapter 4.

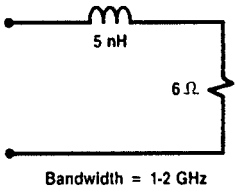
3-3-2 Broadband Matching

The circuitry above has enabled us to extend the frequency range of the amplifier and avoid the normal sharp cutoff. According to Fano [26], there is a limit to the bandwidth of matching. The concept is that impedance plus bandwidth determines the ability to match, since there is a Q of the circuit and there is a Q of the device. The following equations are valid:

$$|\rho_{\text{Bode}}| = e^{-(\pi Q_T / Q_{BP})} \tag{3-180}$$

$$Q_T = \frac{\sqrt{f_1 f_2}}{f_2 - f_1} \tag{3-181}$$

$$Q_{BP} = \frac{\text{Im}(Z)}{\text{Re}(Z)} = \frac{X_s}{R_s} \tag{3-182}$$



$$Q_{BP} = \frac{2\pi\sqrt{(2\text{ GHz})(1\text{ GHz})(5\text{ nH})}}{6\,\Omega} = 7.40$$

$$Q_T = \frac{\sqrt{(2\text{ GHz})(1\text{ GHz})}}{(2\text{ GHz} - 1\text{ GHz})} = 1.414$$

$$|\rho_{\text{Bode}}| = e^{-\left(\frac{\pi(1.414)}{7.40}\right)} = .549$$

$$\text{VSWR}_{\text{Bode}} = \frac{1 + .549}{1 - .549} = 3.43 : 1$$

Figure 3-122 Broadband matching example showing the Bode limit.

Figure 3-122 shows an example. This combination results in a VSWR of 3.43, which for practical purposes is already too high. The workaround is special circuitry (such as exponentially staggered impedance jumps in the matching network like 50 Ω > 20 Ω > 7 Ω > 3 Ω). This means we first transform the 50-Ω source of load impedance down to 3 Ω, and then look for a matching network that matches the transistor output/input, which is within a few ohms of 3 Ω. A detailed discussion of such matching can be found in the "Numerical Design" portion of Section 3-12-2 under the title "Broadband Matching Using Bandpass-Filter Networks—High-Q Case."

3-4 TWO-STAGE AMPLIFIERS

There are two popular configurations that have been used mostly for low-frequency applications. The first one (Figure 3-123, implemented in the circuit shown in Figure 3-124) shows a nice high-frequency performance assuming the emitter and collector currents are essentially the same. As for distortion products, this circuit behaves equivalently to the shunt impedance example from Figures 3-118 and 3-119, with the additional gain of the second transistor. Its frequency matching and noise performance are also shown, in Figure 3-125.

The next two-stage circuit (Figure 3-126, implemented in Figure 3-127) uses the collector current of the second stage as the feedback to the first. Proper biasing of this circuit is quite difficult and interactive. Figure 3-128 shows the frequency-dependent gain, matching, and noise figure. We believe that this circuit can be optimized further, but we leave such investigations to the reader.

A *cascode* amplifier is a special form of two-stage circuit. Although a cascaded pair acts like a single transistor, its merits are in its reduced feedback and resulting higher gain–bandwidth product. Figure 3-129 shows a simple cascode application that assumes the presence

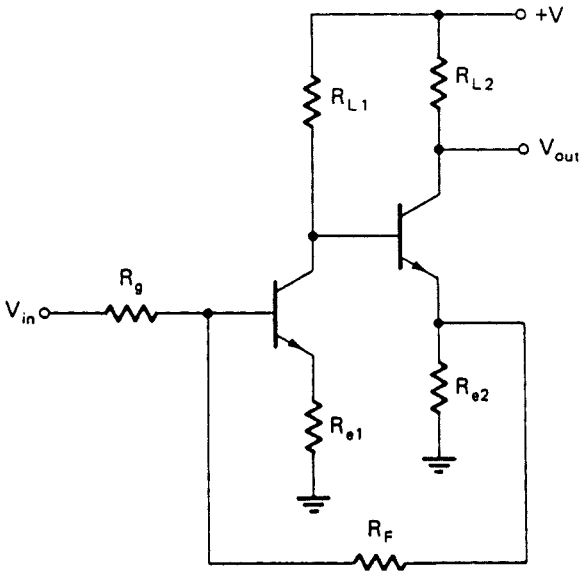


Figure 3-123 Two-stage amplifier with voltage feedback.

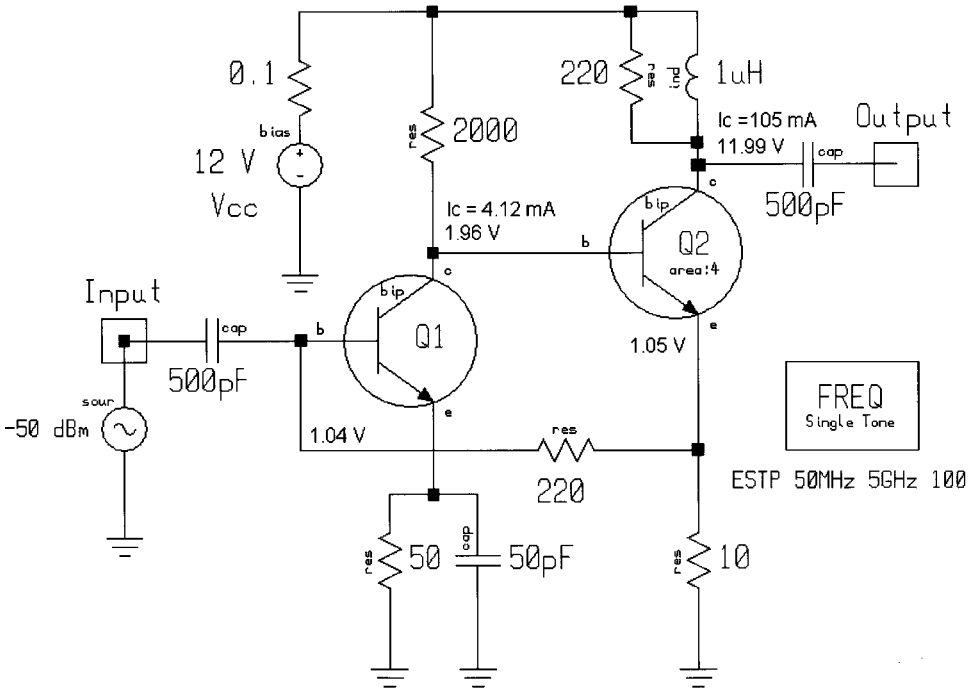


Figure 3-124 Implementation of the circuit in Figure 3-123.

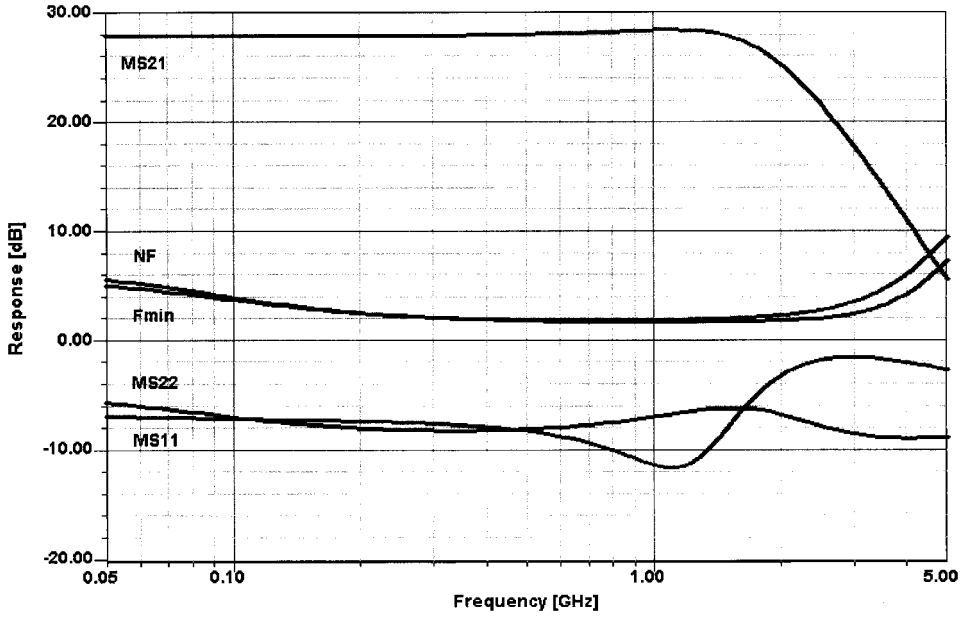


Figure 3-125 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-124.

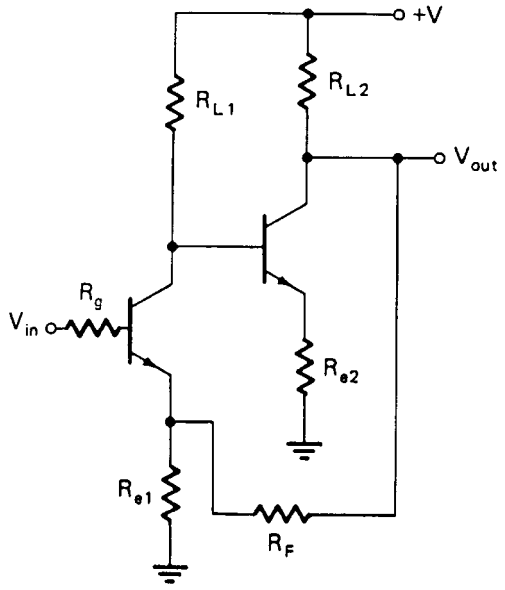


Figure 3-126 Two-stage amplifier with voltage feedback from Stage 2 to Stage 1, and current feedback in both stages.

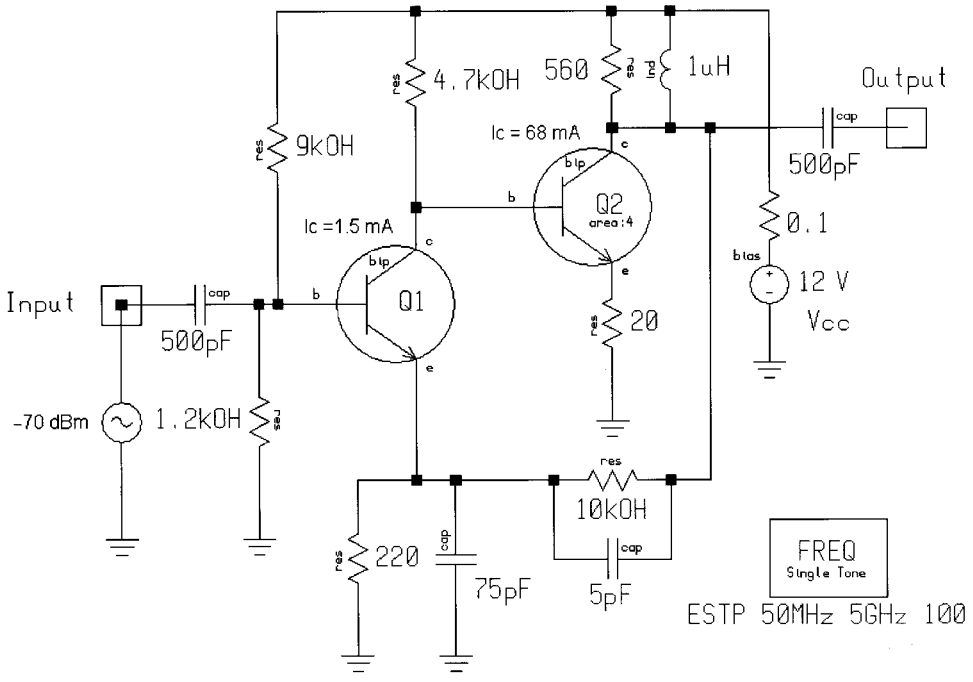


Figure 3-127 Simulated two-stage amplifier with feedback as shown in Figure 3-126.

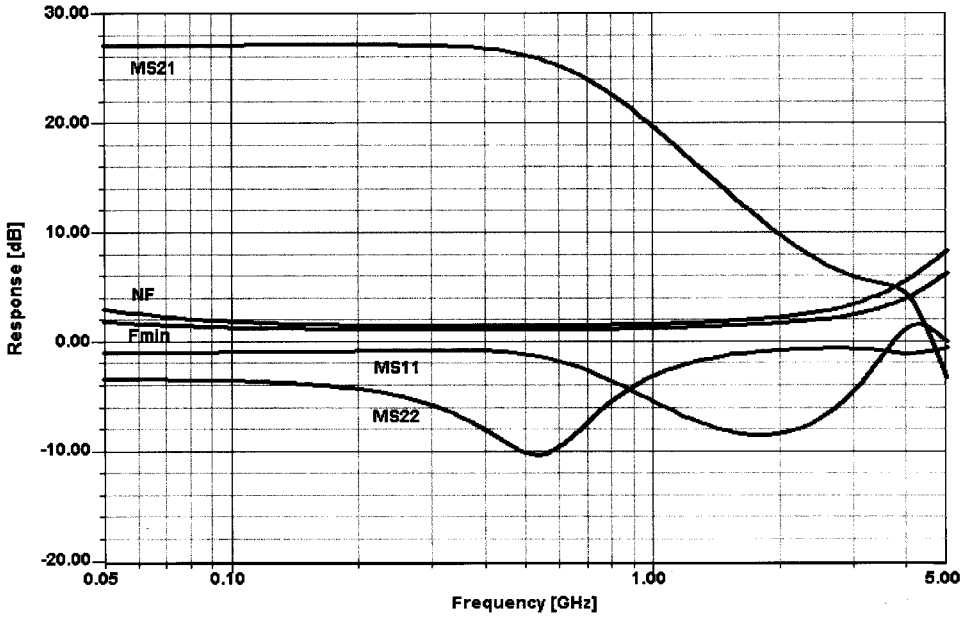


Figure 3-128 Gain, matching, and noise performance of the amplifier in Figure 3-127.

of an output transformer that absorbs the output capacitance. Figure 3-130 shows this circuit's frequency-dependent gain, matching, and noise performance.

It has been obvious to us that most published wireless application circuits are evaluated on a no-input-selectivity basis; that is, their reported performance does not take into account the filtering that is usually present in practical implementations. Evaluated in this way, a low-noise design may achieve a noise figure of less than 2 dB—1.2 dB or so is typical—but this is misleading because filters and connectors introduce insertion loss, resulting in a higher noise figure.

The cascode is a nice preamplifier combination. It combines (depending on the transistor) a low noise figure with essentially no feedback. Since the output of a grounded-base transistor provides a high-impedance source, this configuration is ideal to work into the SAW filters frequently used in wireless applications at the operating frequency. Special transistor pairs are available for use in cascode; as can be seen from this example, not every transistor combination is ideal. This one gives a higher noise figure than wanted and at frequencies above 2 GHz it tends to show possible instabilities. Both Siemens and Motorola are now offering transistors appropriate for these cascodes. Mostly, one would resort to transistors of less-exotic cutoff frequencies to guarantee stability. Typically, grounded-base/gate transistors tend to become oscillators, specifically because this configuration, like the emitter follower, allows a higher operating frequency while maintaining the same gain–bandwidth product; the actual gain for the last two circuit configurations is less. A grounded-base/gate stage has a high voltage gain and the emitter/source follower has power gain at low impedance levels. We invite our readers to experiment with the capacitor from the base of Q1 to ground in Figure 3-129 to get an interesting education about these transistors and their willingness to oscillate.

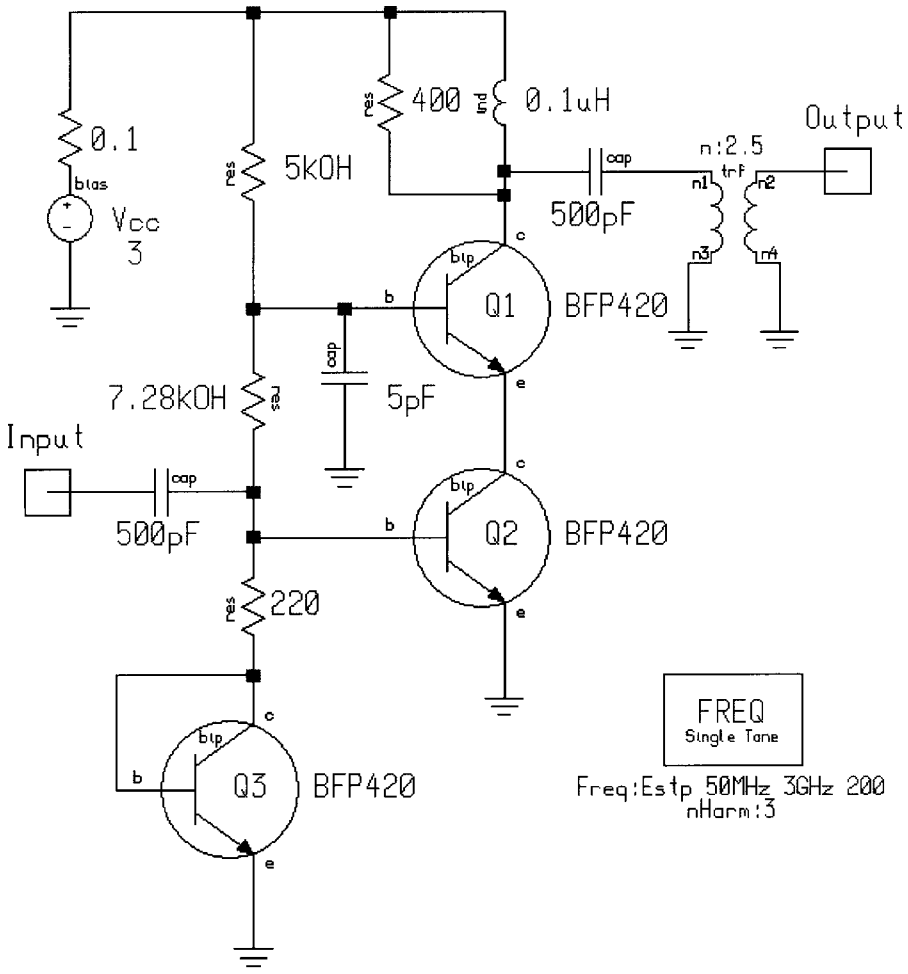


Figure 3-129 Schematic of the cascode amplifier.

The MC13144 is an example of a commercially available VHF to 2-GHz low-noise amplifier with programmable bias. As can be seen from Table 3-11, it is ideal for the 900-MHz range and its programmable gain/bias makes it very attractive. This design (Figure 3-131) is based on the cascode mentioned above and, because of high integration, uses many transistors for bias and level shifting.

The 900-MHz application circuit in Figure 3-132 shows the required external components for a successful design. Input and output are made tunable. The 5.6- and 39-nH inductances can be made from small air-wound coils. The output ceramic filter provides the necessary selectivity ahead of the mixer. With such a highly selective device in the input, the noise figure would deteriorate too much. The antenna is rarely in the vicinity of sufficiently strong transmitters to require such strong input filtering.

There are two more interesting applications for two-stage amplifiers: The one implemented in the Hewlett-Packard MSA-0735 MMIC has become a quite common one, and its usefulness up to higher frequencies depends mostly on the cutoff frequency of the transistors,

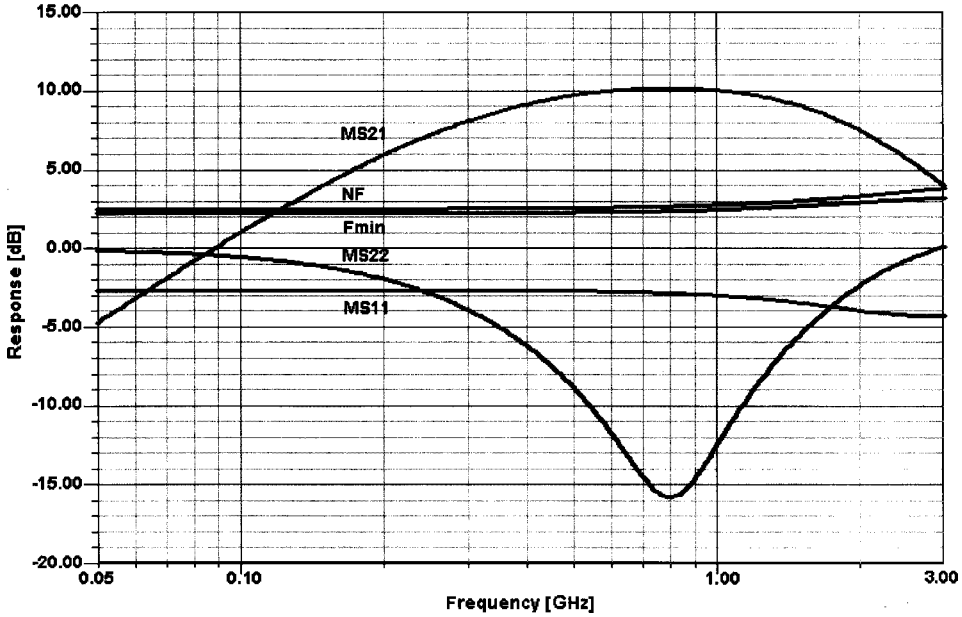


Figure 3-130 Frequency-dependent gain, matching, and noise performance of the cascode amplifier.

their parasitics, and the phase shift of the gain. The typical configuration is shown in Figure 3-133, and its frequency-dependent gain, match, and noise figure in Figure 3-134. There is an increase of S_{22} at higher frequencies; this indicates unwanted phase shift in the circuit. Some manufacturers hide the compensation circuits to overcome such peaking; most of the schematics shown by the manufacturers are really intended only to show operating principles, with proprietary details omitted.

Thanks to the properties of GaAs transistors, we can extend the frequency range significantly. The amplifier we are about to look at goes back several years and is described in great detail in Goyal [27]. Essentially what we have, seen in Figure 3-135, is an FET version of Figure 3-127. The main goal of this circuit was to have an all-monolithic circuit, avoiding all ac coupling and accomplishing all matching and interaction with dc coupling. The transistor Q1 is the main gain stage, which has the transistor Q2 as an active load. The output impedance of Q2 is $1/g_m$—typically in the area of 50–200 Ω, depending on the biasing,

Table 3-11 MC13144 features

Gain at 900 MHz, 17 dB
Noise figure at 900 MHz, 1.4 dB
P_{-1dB}, −7.0 dBm
$IP_{3,in}$, −5.0 dBm
Low operating supply voltage (1.8–6.0 Vdc)
Programmable bias with Enable 1 and Enable 2. [Enable 1 and Enable 2 programmed high for optimal noise figure and gain associated with NF. Enable can be overridden for external bias programming (up to 15 mA.)]

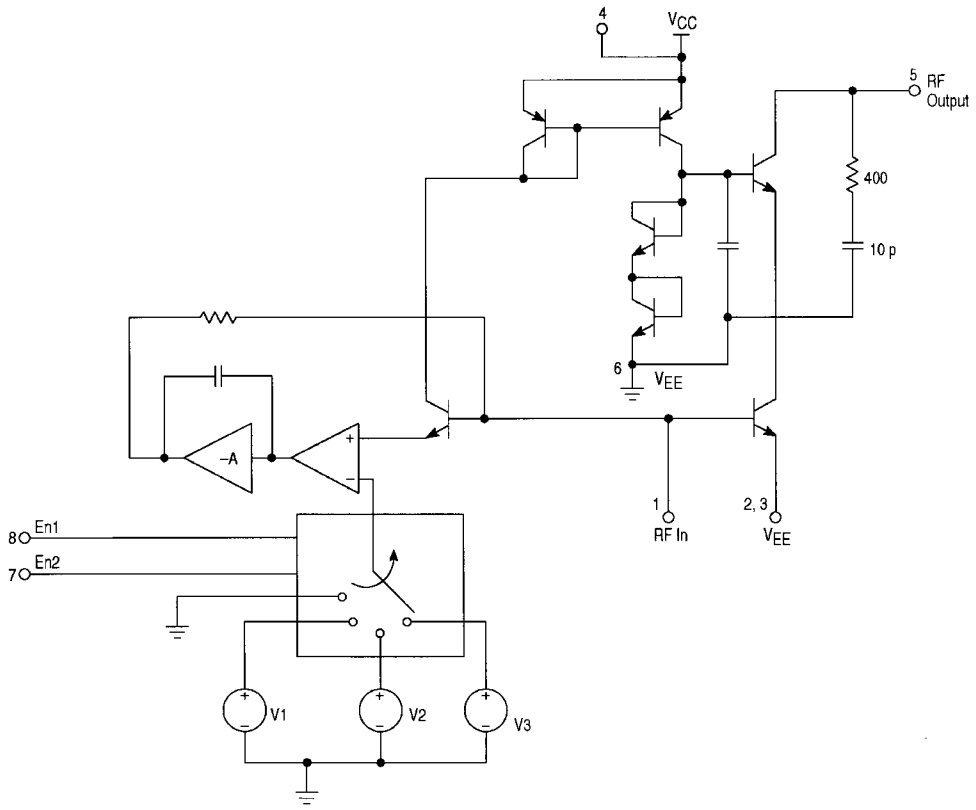


Figure 3-131 Schematic of the Motorola MC13144 cascode LNA.

which affects the transconductance. Transistor Q3 has the same function as an emitter/source follower; again, its dc bias determines its transconductance and therefore its output impedance. (Given this, a device that has a transconductance of 20 mS, or close to it, would establish a very good output match to 50 Ω.) The drains of Q1 and Q5 are at approximately 3.5 V. The diodes at the source of Q3 are level shifters that must shift the dc voltage at the gate of Q5 and at the drain of Q4 to approximately −1.5 V relative to ground—the necessary bias condition for Q1 and Q5. The standard shunt feedback circuit would show resistive feedback between the drain of Q1 (gate of Q2) and the gate of Q1; in the previous examples, we calculated the magic value of this part to be about 200 Ω. Adding in parallel to Q1 the transistor Q5 allows the designer to use a feedback scheme that isolates the feedback loop from the input. Dc-wise, the two transistors are tied together via a 10-kΩ resistor, which can be included in the actual packaged device. Insufficient information was published to allow us to duplicate the design, considering all the intricacies involved (we somewhat object to the fact that even in the *IEEE MTT-S Transactions*, so little information is given about the actual implementation of this design—using space constraints as an excuse—that its educational value is usually minimal); it shows a new principle without revealing the design steps necessary to obtain a practical circuit. Several variations of this type of circuit are available; the literature is full of discussions about them. The magnitude of the feedback is set by the ratio of the width of Q5 to that of Q1. Typical values for the Q5/Q1 ratio range

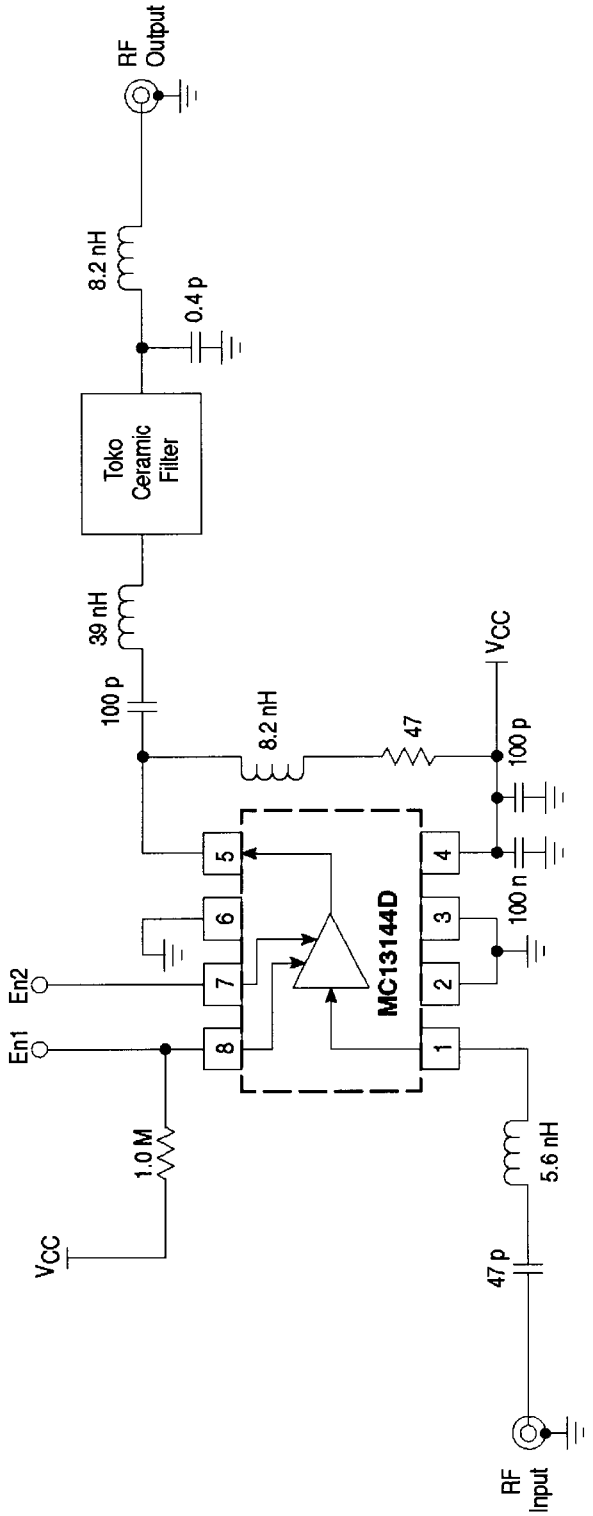


Figure 3-132 An MC13144 application circuit for 926.5 MHz.

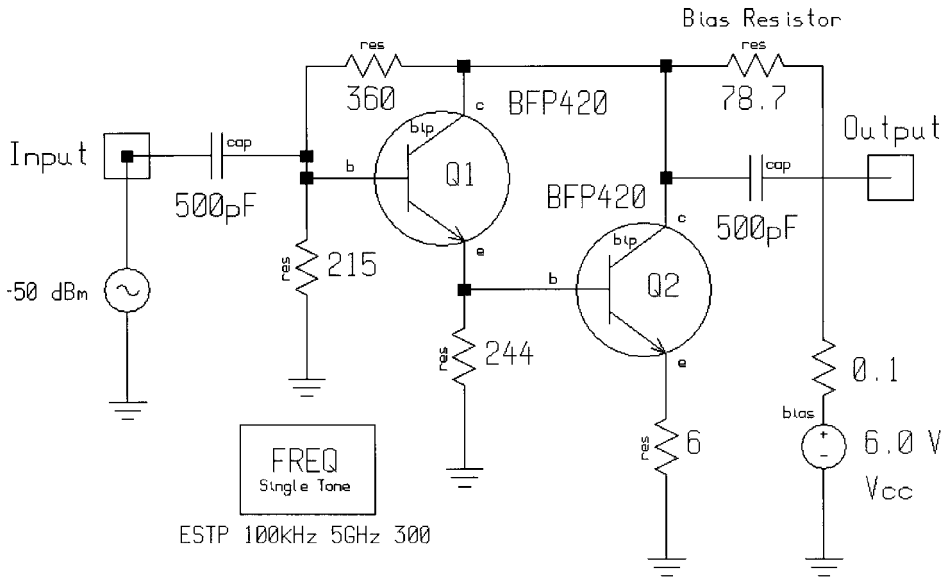


Figure 3-133 Schematic of the MSA-0375 MMIC amplifier.

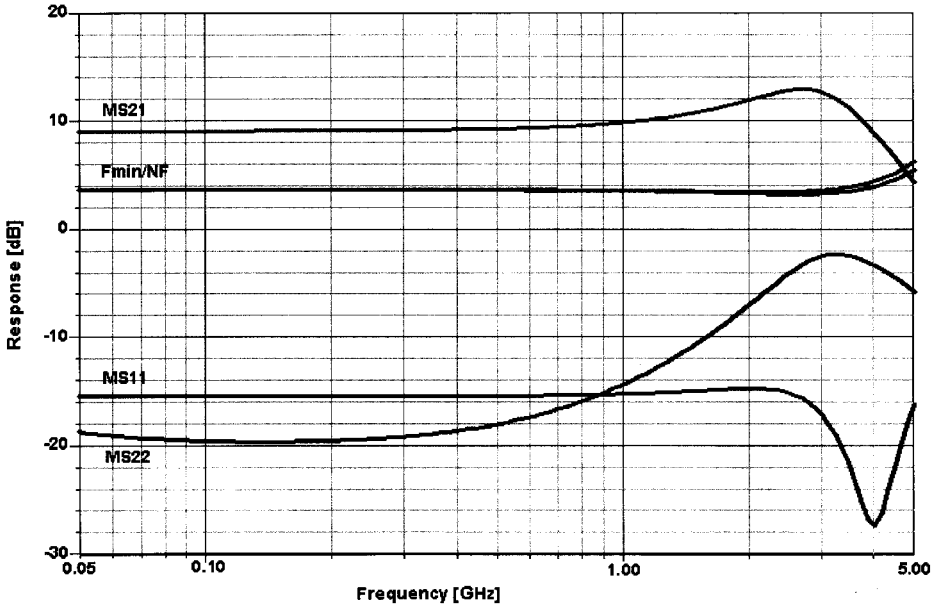


Figure 3-134 Frequency-dependent gain, matching, and noise performance of the MSA-0735 MMIC amplifier.

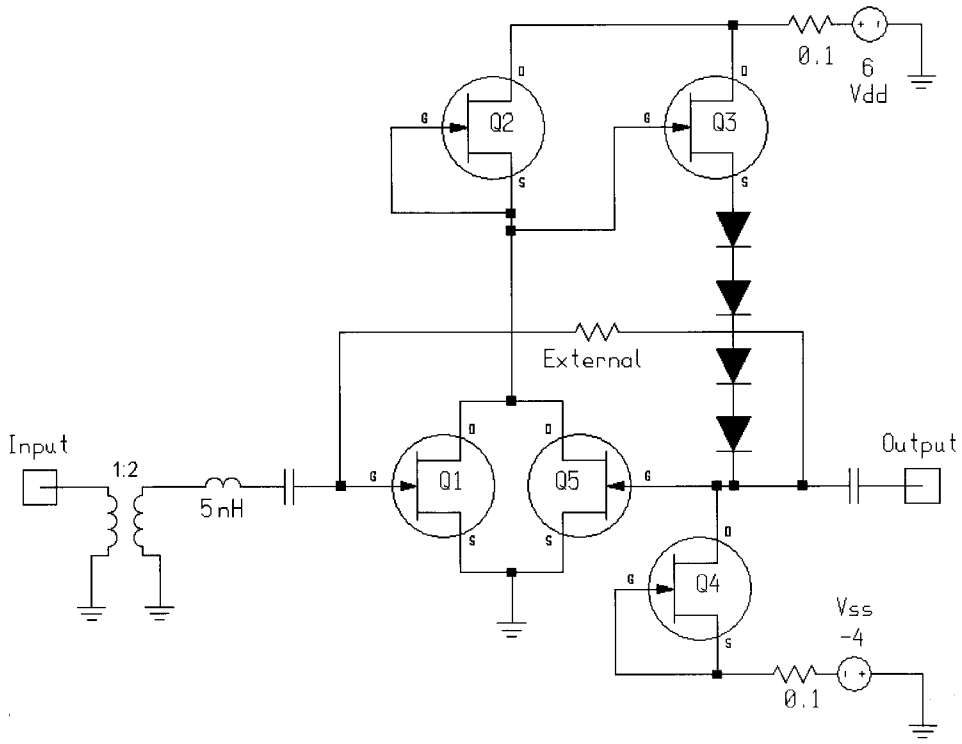


Figure 3-135 Schematic of the dc-coupled GaAsFET amplifier.

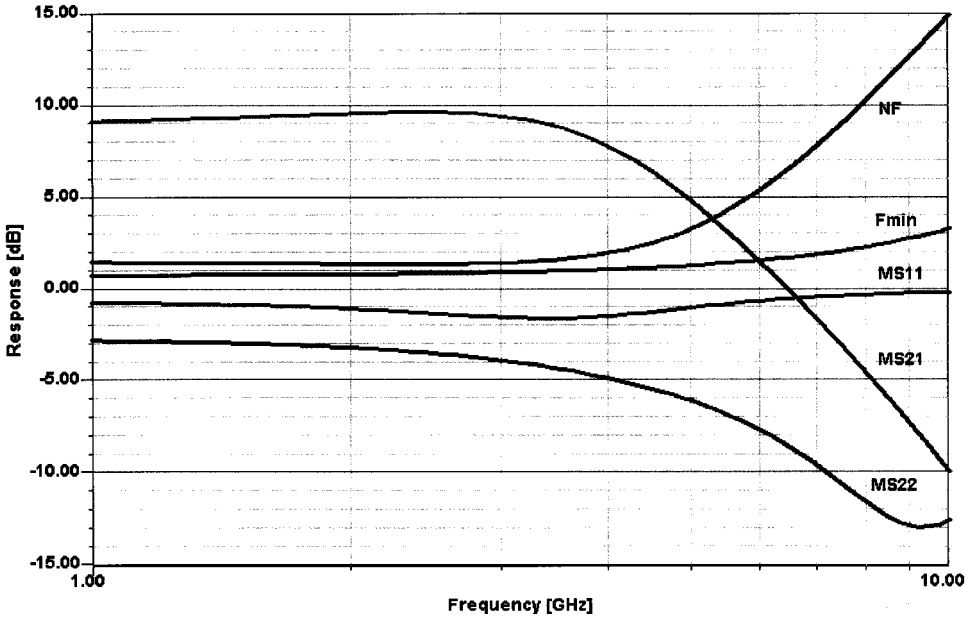


Figure 3-136 Frequency-dependent gain, matching, and noise performance of the dc-coupled GaAsFET amplifier.

from 0.15 to 0.30. In simulation, sizing the devices can best be accomplished by using the scaling factor for the FET model.

One of our questions was: "What would the perfect input termination for Q1 be?" By using an ideal transformer and playing with its turns ratio (needless to say, we did this with CAD), we determined experimentally that, relative to the gate, 200 Ω and a series inductor of 5–8 nH—the sort of details that designers often avoid telling their audiences—gave the ideal frequency response. Overall, we believe that the inventor of this circuit can be quite proud of the overall performance. The FETs we chose were taken from the Texas Instruments nonlinear foundry library; however, any similar device from another company would have worked equally well. The feedback loop used in this IC is an example of an approach frequently referred to as *active feedback*. The performance of this amplifier can be evaluated from Figure 3-136, which shows the frequency-dependent gain, matching, and noise figures (even including F_{min}).

3-5 AMPLIFIERS WITH THREE OR MORE STAGES

Most three-stage amplifiers are obtained by adding another buffer or power-gain stage to the previously described principal designs. As an example, we decided to evaluate the recent NEC UPC2749 IC, a silicon MMIC intended for application as a low-noise 1900-MHz amplifier operating at 3 V. Table 3-12 briefly summarizes its electrical specifications.

Table 3-12 NEC UPC2749 electrical characteristics

(T_A = 25 °C, $Z_L = Z_S$ = 50 Ω, V_{CC} = 3.0 V)			
	Part Number Package Outline		UPC2749T TO6
Symbols	Parameters and Conditions	Units	Typical
I_{CC}	Circuit current (no signal)	mA	6
G_S	Small-signal gain, f = 900 MHz dB	14.5	
	f = 1900 MHz	dB	16
f_{corner}	–3-dB Gain corner frequency	GHz	2.9
P_{1dB}	1 dB Compressed output power at 1900 MHz	dBm	–12.5
P_{SAT}	Saturated output power, f = 1900 MHz	dBm	–6
NF	Noise figure, f = 900 MHz	dB	3.2
	f = 1900 MHz	dB	4.0
RL_{IN}	Input return loss, f = 1900 MHz	dB	10
RL_{OUT}	Output return loss, f = 1900 MHz	dB	12.5
ISOL	Isolation, f = 1900 MHz	dB	30
OIP_3	SSB output third-order intercept, f_1 = 500 MHz, f_2 = 510 MHz	dBm	–3
	f_1 = 1000 MHz, f_2 = 1010 MHz	dBm	–3
	f_1 = 1900 MHz, f_2 = 1902 MHz	dBm	–3.5
	f_1 = 2000 MHz, f_2 = 2010 MHz	dBm	–4
$R_{TH(J-A)}$	Thermal resistance (junction to ambient)		
	Free air	°C/W	
	Mounted on a 50 × 50 × 1.6 mm epoxy glass PWB	°C/W	

In modeling the UPC2749's three transistors, we chose to use our favorite, the BFP420, which is not without its headaches at the higher frequencies. Figure 3-137 shows the circuit diagram, which closely resembles the one published by NEC; again, the only hard facts we had were the IC's basic circuit topology, gain, and dc current. Given those numbers, we staggered the current in the three devices by setting them at 0.64 mA for the first stage, 1.85 mA for the second stage, and 3 mA for the third stage. It should be noted that the two output transistors are not unlike the approach used for the Hewlett-Packard MSA-0735 (Figure 3-133). Taking (hopefully) all the right assumptions, we end up with a design that closely resembles the manufacturer's specifications for both power consumption, gain, and noise. This can be verified by looking at the results plotted in Figure 3-138.

The final example is a more complex derivative of the Philips NE/SA5204A amplifier, which was intended for standard 6-V operation and has a fairly "normal" circuit arrangement. This can be seen from Figure 3-139. A much more elaborate cousin of this is the low-voltage NEC UPC2710 amplifier. According to its specifications, it operates at 5 V but has significantly more gain than the NE/SA5204. Table 3-13 shows its electrical specifications and Figure 3-140 shows its measured gain as published by the manufacturer.

The actual interior of the circuit is shown in Figure 3-141. It consists of the standard preamplifier and then a Darlington configuration with the collectors tied together. The middle section of the circuit acts as a power supply responsible for the bias relative to the emitter of the first transistor influencing the rest of the circuit as well. Since NEC mentioned that they use their NESAT III process (f_T of 20 GHz at V_{ce} = 3) in producing the UPC2710T and we

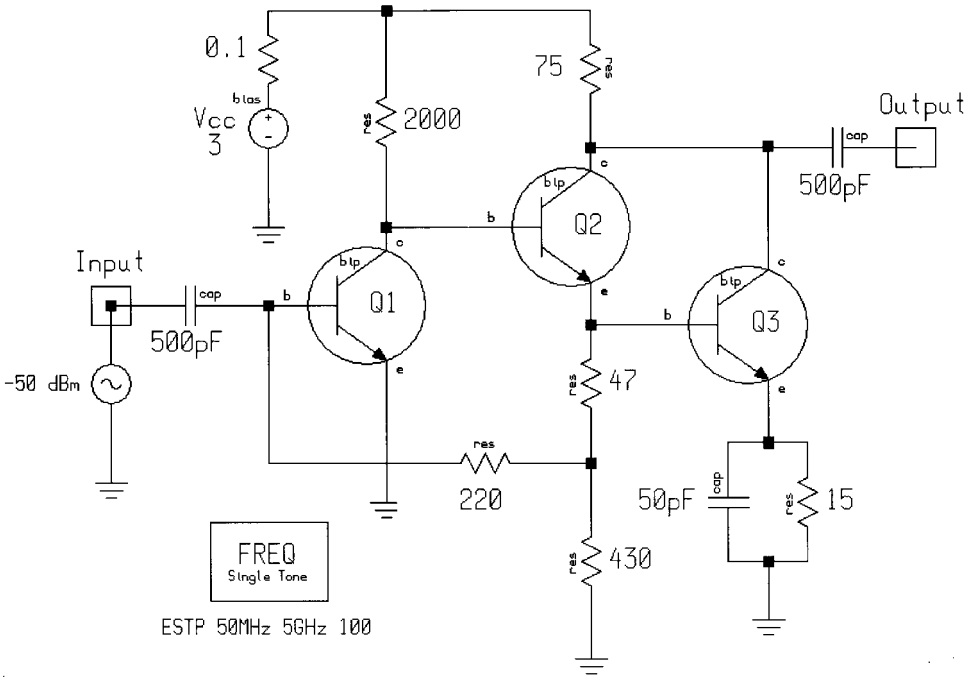


Figure 3-137 Circuit of the UPC2749 MMIC. The component values were not supplied by the manufacturer.

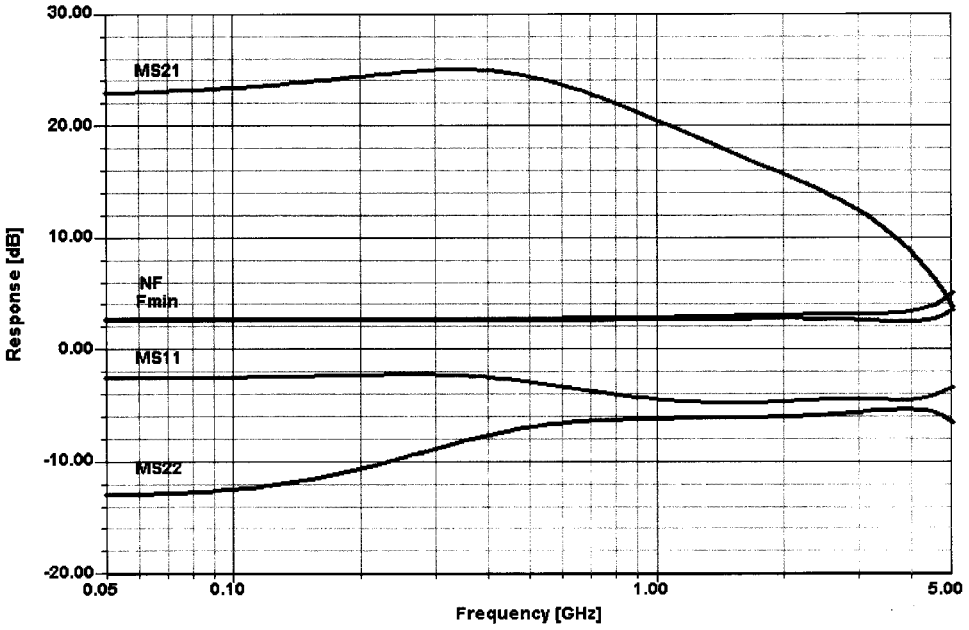


Figure 3-138 Simulated gain, matching, and noise performance of the NEC UP2749 MMIC.

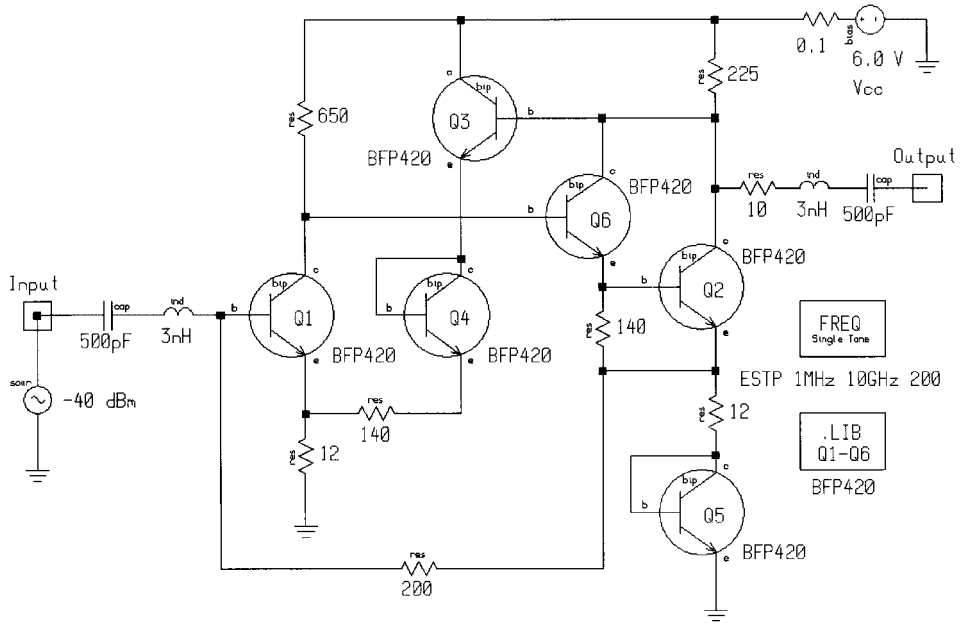


Figure 3-139 Schematic of the Philips NE/SA5204A amplifier IC entered for simulation.

Table 3-13 UPC2710 electrical specifications

	T_A = 25 °C, f = 0.5 GHz, V_{CC} = 5 V)		
	Part Number Package Outline		UPC2710T TO6
Symbols	Parameters and Conditions	Units	Typical
I_{CC}	Circuit current (no signal)	mA	22
G_S	Small-signal gain	dB	33
f_{corner}	–3-dB Gain corner frequency	GHz	1.0
ΔG_S	Gain flatness, f = 0.1–0.6 GHz f = 0.1–0.8 GHz	dB	±0.8
P_{SAT}	Saturated output power	dBm	13.5
P_{1dB}	Output power at 1dB compression point	dBm	7.5
NF	Noise figure	dB	3.5
RL_{IN}	Input return loss	dB	6
RL_{OUT}	Output return loss	dB	12
ISOL	Isolation	dB	39
ΔG_T	Gain–temperature coefficient	dB/°C	–0.006
R_{TH}	Thermal resistance (junction to ambient)	°C/W	

obtained some information regarding it, we decided to use our standard BFP420 (f_T of 25 GHz; see the datasheet in Chapter 2) for the purpose of simulation.

In our effort to simulate this circuit, we chose a combination of values that pretty much complies with the manufacturer's published specifications, but S_{22} is not close, as shown in Figure 3-142. We stopped at this point and invite our readers with CAD capabilities to provide the final touch in meeting all the published specifications for the circuit. We are particularly

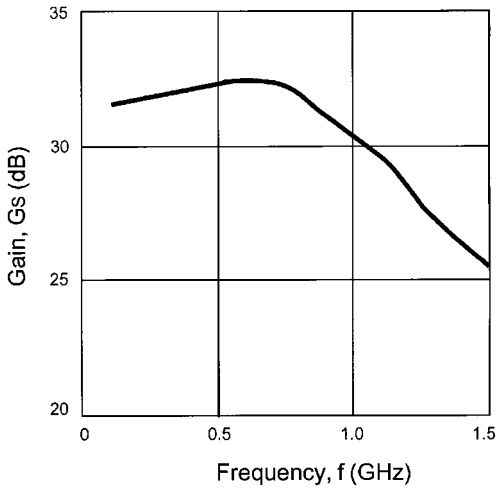


Figure 3-140 Manufacturer-supplied gain versus frequency for the NEC UPC2710T MMIC.

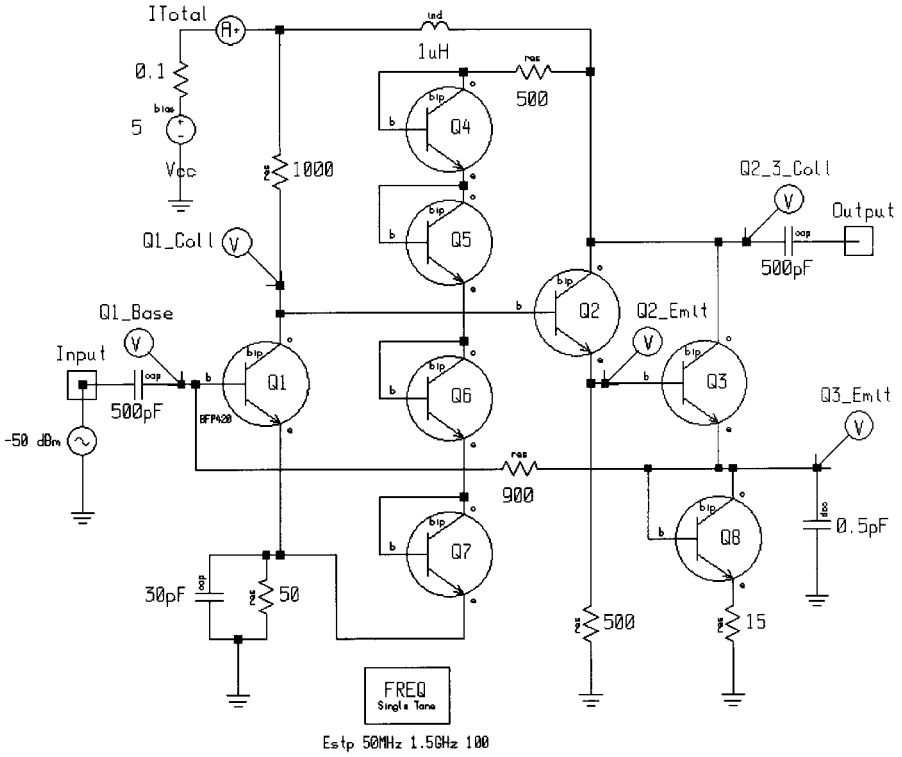


Figure 3-141 Schematic of the UPC2710T silicon MMIC entered for simulation.

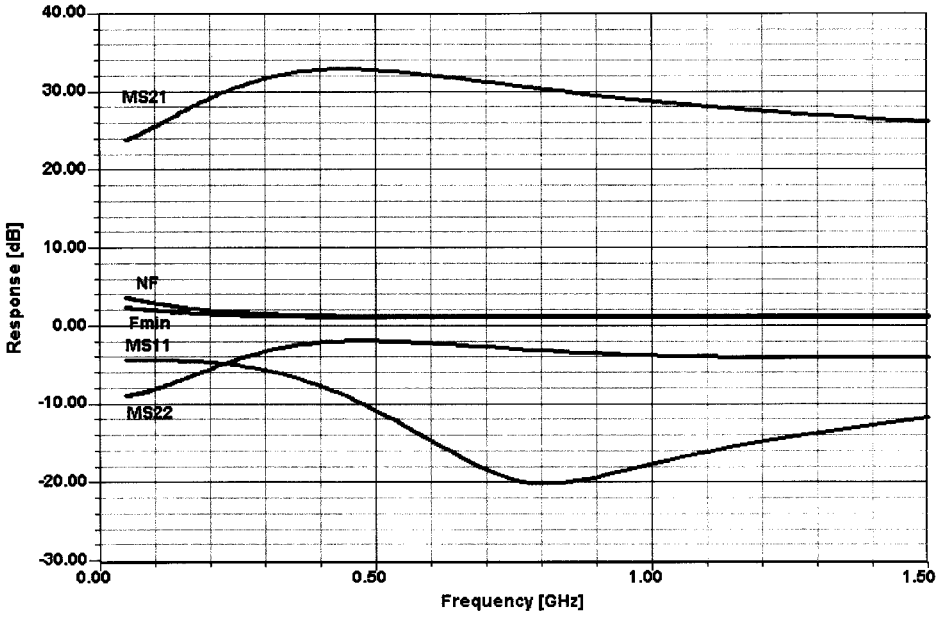


Figure 3-142 Simulated gain, match, and noise performance of the UPC2710T.

recommending such an effort because it shows very nicely the interaction of all the circuit elements in a fairly drastic way. Because of this, we admire the IC design even more, since the manufacturer has to sell a pretty consistent amplifier as a function of different production runs.

3-5-1 Stability of Multistage Amplifiers

Stability, needless to say, is a big issue, and the delay of the various stages adds to the phase shift, resulting in possible instabilities. The fact that S_{11} and/or S_{22} can become larger than 0 (in dB) means there is the potential of oscillating, depending on the reactances available. If a cable of inappropriate length is added, hell will break loose. Figure 3-143 shows the dangerous peaking of our simulation of the Philips NE5204A, in which the internal compensation was intentionally not assumed correctly. We will all agree that the high-end performance of this amplifier can generate high anxiety. As the production process of the aging device becomes more modern—meaning that transistors with much higher f_T will be the basis for the production—it will become quite difficult to maintain a previously achieved stability. This, by the way, applies also to discrete transistors, which for reasons of economics and multiple manufacturers will be improved over time. Such evolution results in generally smaller base-spreading resistances and capacitances, and other parameter changes, which, overall, cause headaches in manufacturing. We will leave the topic of three-stage amplifiers on this note.

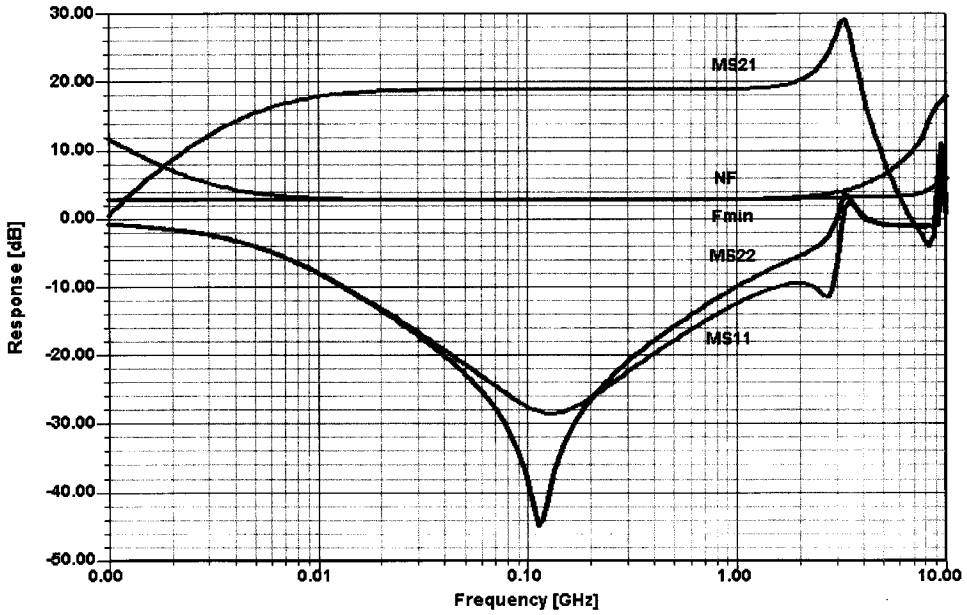


Figure 3-143 Simulated gain, match, and noise performance of the NE5204A IC.

3-6 A NOVEL APPROACH TO VOLTAGE-CONTROLLED TUNED FILTERS INCLUDING CAD VALIDATION [28]

Modern receivers control input stages as well as the oscillator band and frequency by electrical rather than mechanical means. Tuning is accomplished by voltage-sensitive capacitors (varactor diodes), and band switching by diodes with low forward conductance. Since the wireless band (essentially 400 MHz to 2.4 GHz) is so full of strong signals, the use of a tracking filter is desired as a solution to improve the performance and prevent second-order IMD products or other undesired overload effects. The dc control voltage needed for the filter can easily be derived from the VCO control voltage. There may be a small dc offset, depending on the IF used.

3-6-1 Diode Performance

The capacitance versus voltage curves of a varactor diode depend on the variation of the impurity density with the distance from the junction. When the distribution is constant, there is an "abrupt junction" and capacitance follows the law

$$C = \frac{K}{(V_d + V)^{1/2}} \tag{3-183}$$

where V_d is the contact potential of the diode and V is applied voltage.

Such a junction is well approximated by an alloyed junction diode. Other impurity distribution profiles give rise to other variations, and the above equation is usually modified to

$$C = \frac{K}{(V_d + V)^n} \tag{3-184}$$

where n depends on the diffusion profile and $C_0 = K/V_d^n$.

A so-called graded junction, having a linear decrease in impurity density with the distance from the junction, has a value of n. This is approximated in a diffused junction.

In all cases these are theoretical equations, and limitations on the control of the impurity pattern can result in a curve that does not have such a simple expression. In this case, the coefficient n is thought of as varying with voltage. If the impurity density increases away from the junction, a value of n higher than 0.5 can be obtained. Such junctions are called hyperabrupt. A typical n value for a hyperabrupt junction is about 0.75. Such capacitors are used primarily to achieve a large tuning range for a given voltage change. Figure 3-144 shows the capacitance–voltage variation for the abrupt and graded junctions as well as for a particular hyperabrupt-junction diode. Varactor diodes are available from a number of manufacturers, such as Motorola, Siemens, and Philips. Maximum values range from a few to several hundred picofarads, and useful capacitance ratios range from about 5 to 15.

Figure 3-145 shows three typical circuits that are used with varactor tuning diodes. In all cases, the voltage is applied through a large resistor R_e or, better, an RF choke in series with a small resistor. The resistance is shunted across the lower diode and may be converted to a shunt load resistor across the inductance to estimate Q. The diode also has losses that may result in lowering the circuit Q at high capacitance, when the frequency is sufficiently high. This must be considered in the circuit design.

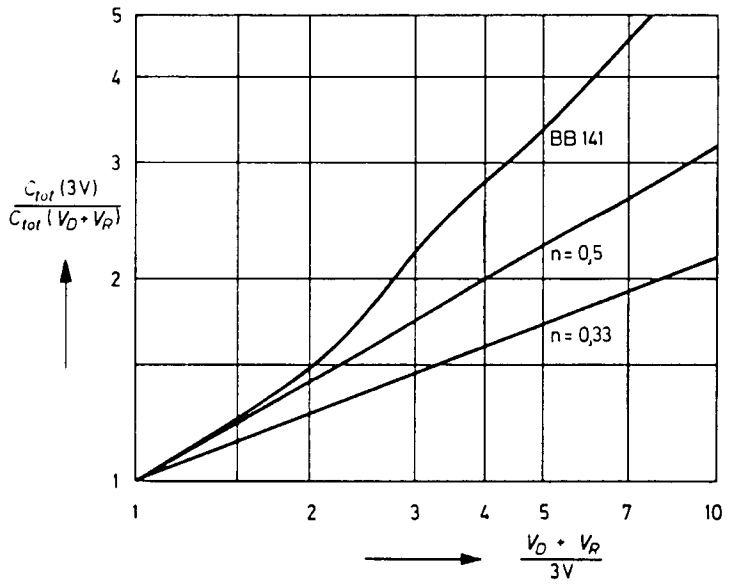


Figure 3-144 Voltage-dependent change of capacitance of different types of diodes. The BB141 is a hyperabrupt-junction diode with $n = 0.75$.

The frequency-dependent performance is not determined solely by applying the dc tuning voltage to Eq. (3-184). If the RF voltage is sufficient to drive the diode into conduction on peaks, an average current will flow in the circuits of Figure 3-145, which will increase the bias voltage. The current is impulsive, giving rise to various harmonics of the circuit. Even in the absence of conduction, Eq. (3-184) deals only with the small-signal capacitance. When the RF voltage varies over a relatively large range, the capacitance changes. In this case, Eq. (3-184) must be changed to

$$\frac{dQ}{dV} = \frac{K}{(V + V_d)^n} \tag{3-185}$$

Here Q is the charge on the capacitor. When this relation is substituted in the circuit differential equation, it produces a nonlinear differential equation, dependent on the parameter n. Thus, the varactor my generate direct current and harmonics of the fundamental frequency. Unless the diodes are driven into conduction at some point in the cycle, the direct current must remain zero.

The current of Figure 3-145c can be shown to eliminate the even harmonics, and it permits a substantially larger RF voltage without conduction than either circuit in Figure 3-145a or 3-145b. When $n = 0.5$, only the second harmonic is generated by the capacitor, and this can be eliminated by the back-to-back connection of the diode pair. Integrating Eq. (3-185), we find

$$Q + Q_A = \frac{K}{1 - n}(V + V_d)^{1-n}$$

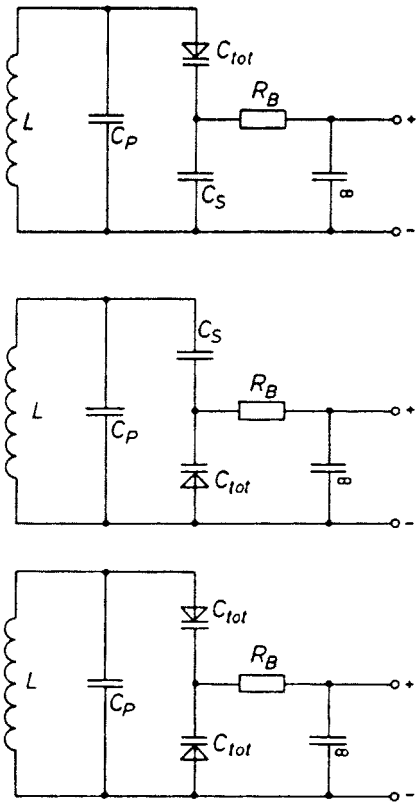


Figure 3-145 Various configurations to apply tuning diodes in a tuned circuit. The bottom version shows the lowest distortion.

$$= \frac{C_v}{1-n}(V + V_d) \tag{3-186}$$

C_v is the value of Eq. (3-184) for applied voltage V, and Q_A is a constant of integration. By letting $V = V_1 + v$ and $Q = Q_1 + q$, where the lowercase letters represent the varying RF and the uppercase letters indicate the values of bias when RF is absent, it follows that

$$q + Q_1 + Q_A = \frac{K}{1-n}[v + (V_1 + V_d)]^{1-n} \tag{3-187}$$

$$1 + \frac{v}{V'} = \left(1 + \frac{q}{Q'}\right)^{1/(1-n)} \tag{3-188}$$

where $V' = V_1 + V_d$ and $Q' = Q_1 + Q_A$. For the back-to-back connection of identical diodes, $K_{11} = K_{12} = K_1$, $V_1' = V_2' = V'$, $Q_1' = Q_2' = Q'$, $q = q_1 = -q_2$, and $v = v_1 - v_2$. Here the new subscripts 1 and 2 refer to the top and bottom diodes, respectively; v is the RF voltage across

the pair in series, and q is the charge transferred through the pair in series. This notation results in

$$\frac{v}{V'} \equiv \frac{v_1 - v_2}{V'} = \left(1 - \frac{q}{Q'}\right)^{1/(1-n)} - \left(1 - \frac{q}{Q'}\right)^{1/(1-n)} \tag{3-189}$$

For all n, this eliminates the even powers of q, and hence the even harmonics. This can be shown by expanding Eq. (3-189) in a series expansion and performing term by term combination of the equal powers of q. In the particular case $n = \frac{1}{2}$, $v/V' = 4q/Q'$, and the circuit becomes linear.

The equations hold as long as the absolute value of v_1/V' is less than unity, so that there is no conduction. At the point of conduction, the total value of v/V' may be calculated by noting that when $v_1/V' = 1$, $q/Q' = -1$, so $q_2/Q' = 1$, $v_2/V = 3$, and $v/V = -4$. The single-diode circuits conduct at $v/V' = -1$, so the peak RF voltage should not exceed this. The back-to-back configuration can provide a fourfold increase in RF voltage handling over the single diode. For all values of n, the back-to-back configuration allows an increase in the peak-to-peak voltage without conduction. For some hyperabrupt values of n, such that $1/(1 - n)$ is an integer, many of the higher-order odd harmonics are eliminated, although only $n = \frac{1}{2}$ provides elimination of the third harmonic. For example, $n = \frac{2}{3}$ results in $1/(1 - n) = 3$. The fifth harmonic and higher odd harmonics are eliminated, and the peak-to-peak RF without conduction is increased eightfold; for $n = \frac{3}{4}$ the harmonics 7 and above are eliminated, and the RF peak is increased 16 times. It must be noted in these cases that the RF peak at the fundamental may not increase so much, since the RF voltage includes the harmonic voltages.

Since the equations are only approximate, not all harmonics are eliminated, and the RF voltage at conduction, for the back-to-back circuit, may be different from that predicted. For example, abrupt-junction diodes tend to have n of about 0.46–0.48 rather than exactly 0.5. Hyperabrupt junctions tend to have substantial changes in n with voltage. The diode illustrated in Figure 3-144 shows a variation from about 0.6 at low bias to about 0.9 at higher voltages, with wiggles from 0.67 to 1.1 in the midrange. The value of V_d for varactor diodes tends to be in the vicinity of 0.7 V.

3-6-2 A VHF Example

The application of tuning diodes in double-tuned circuits in TV tuners has been common practice for many years. Figure 3-146 shows the circuit diagram. The input impedance of 50 Ω gets transformed up to 10 kΩ. The tuned circuits consist of the 0.3-μH inductor and two sets of antiparallel diodes. By dividing the RF current in the tuned circuit and using several diodes instead of just one pair, intermodulation distortion is reduced.

The coupling between the two tuned circuits is tuned via the 6-nH inductor that is common to both circuits. This type of inductance is usually printed on the circuit board. The diode parameters used for this application were equivalent to the Siemens BB515 diode. The frequency response of this circuit is shown in Figure 3-147.

The coupling is less than critical. This results in an insertion loss of about 2 dB and relatively steep passband sides. Once the circuit's large-signal performance (Figure 3-148) is seen, a third-order intercept point of about –2 dBm is not so unexpected. The reason for this poor performance is the high impedance (high L/C ratio), which provides a large RF voltage swing across the diodes. A better approach appears to be using even more diodes and at the same time changing the impedance ratio (L/C ratio).

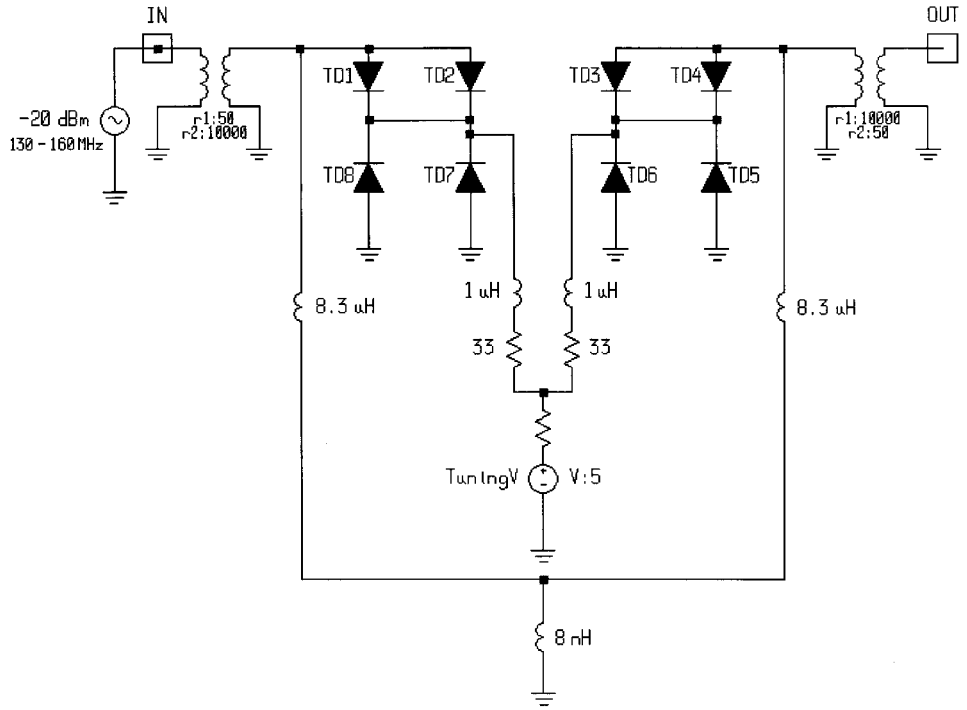


Figure 3-146 Double-tuned filter at 161 MHz using hyperabrupt-junction tuning diodes. By using several parallel diodes, the IMD performance improves.

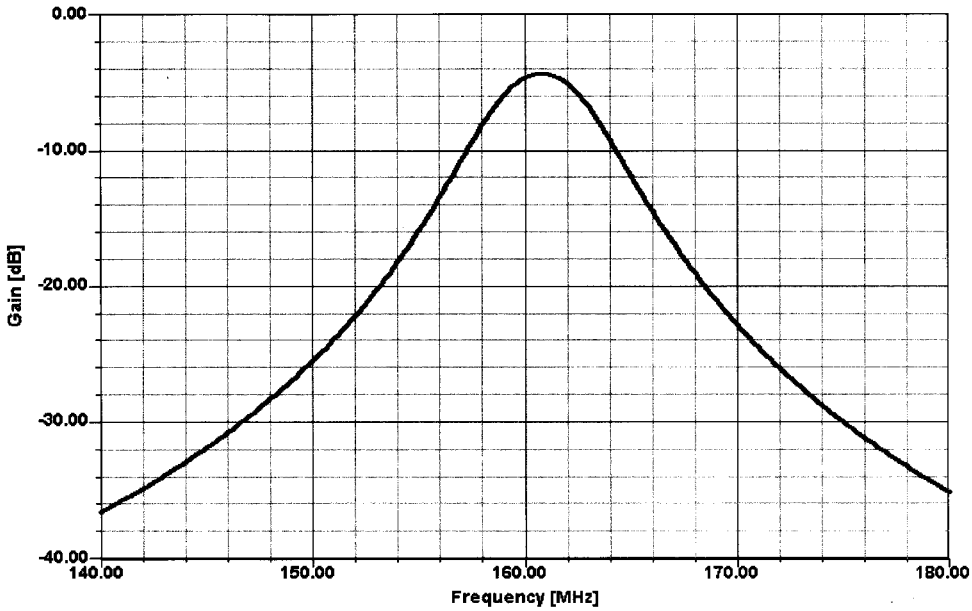


Figure 3-147 Frequency response of the tuned filter shown in Figure 3-146. The circuit is undercoupled (less than transitional coupling); $Q \times k < 1$.

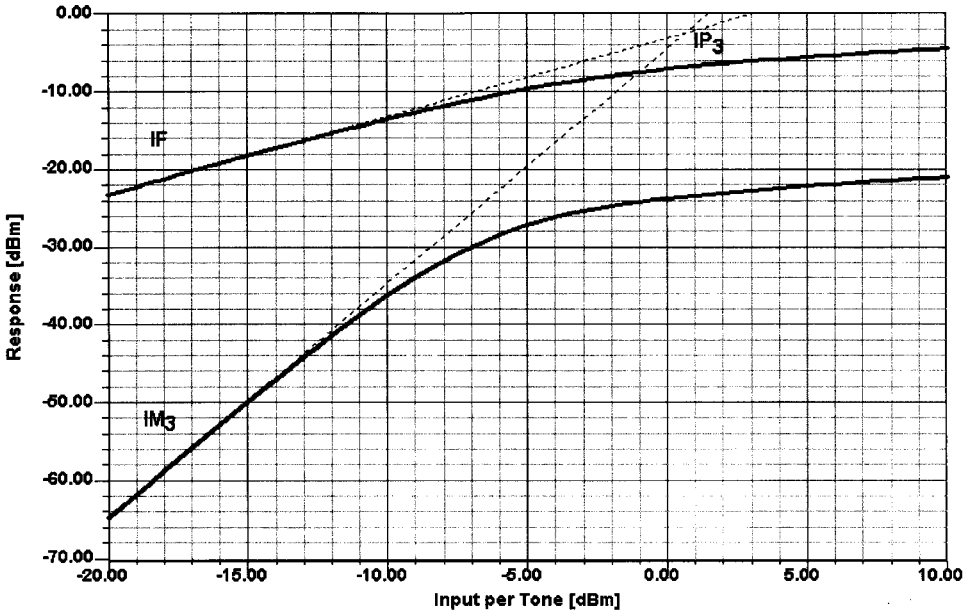


Figure 3-148 Prediction of intercept point of the double-tuned circuit shown in Figure 3-146. Note the compression of both the input signal and the IMD product.

3-6-3 An HF/VHF Voltage-Controlled Filter

The above example used a step-up procedure typically done by a tap at the input and output inductance. This method allows an impedance transformation; however, if one desires to change it into a series-tuned arrangement, it has to be done with a transformer. The large-signal conditions in the frequency range from 10 to 30 MHz on a medium- to large-sized antenna are equivalent to, if not worse than, the conditions for VHF operation. The only exception would be line of sight into the transmitter such as a tower in the middle of the city. Examples of such hostile conditions would be any large city, such as Munich, New York, Miami, Chicago, and San Francisco, where the authors had significant experience with intermodulation distortion problems.

By translating the circuit into a low-impedance-drive arrangement, building it symmetrically, and reducing the tuning range somewhat by capacitors and many series diode pairs, the filter's large-signal performance was significantly improved. The tuning diodes had about 125 pF at 1 Vdc per unit. Initial tests trying to use a high-capacitance diode, such as the BB112, resulted in much higher IMD products. The B112 has a 1-V capacitance of 470 pF down to about 20 pF at 8 V. Standard diodes show about 30 pF at 1 V, while the diodes used in the experiment had a 125-pF capacitance at 1 V. These experimental diodes were made available from a well-known manufacturer for this evaluation. Figure 3-149 shows the circuit diagram.

A step-down transformer drives the tuned circuit with a source impedance of 12.5 Ω. The circuit is symmetrical. There are two 0.8-μH inductors and 2×5 diodes in the loop. The output of the circuit is transformed back up to 50 Ω. The tuning voltage is supplied via a heavily filtered arrangement. This circuit has actually been implemented in Rohde & Schwarz field strength meter equipment. They guarantee an intercept point of about +20 dBm. The selectivity of these circuits is quite reasonable and mostly intended to reduce

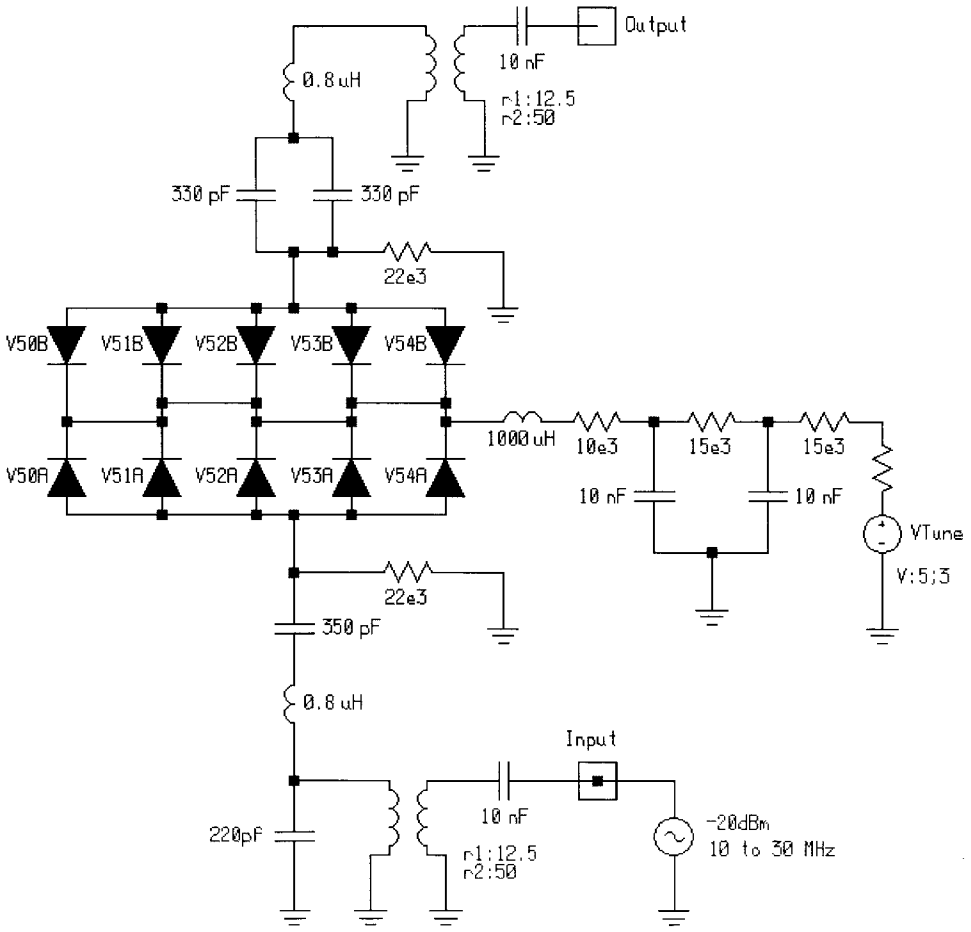


Figure 3-149 High dynamic range of a 10–20-MHz filter.

second-order intermodulation distortion products by about 10 dB, at the same time having a high third-order intercept point. Figure 3-150 shows the selectivity curves at two different tuning voltages.

The most interesting number, however, is the third-order intercept point (already determined as about 20 dBm), which still had to be simulated. This sometimes sounds like a contradiction: Once the measured values are available and are acceptable, why would one want—or, rather, need—to do a simulation besides the necessity to develop a high-input intercept filter? It was desirable to validate the nonlinear models and prove that the above-mentioned equations will hold true. This type of simulation is now more difficult because the number of nonlinear elements went from four to ten, and some numerical problems, such as convergence difficulties, can be expected, and the effect of harmonic frequency cancellation (compensation) can also be seen. By using diode combinations that result in $1(1 - n)$, n being an integer number, these IMD products can drastically be reduced. This implies that each of the five diode pairs has a selected value for n to meet this condition. Later, a final attempt will be made to improve the first VHF filter with the proper diode combinations.

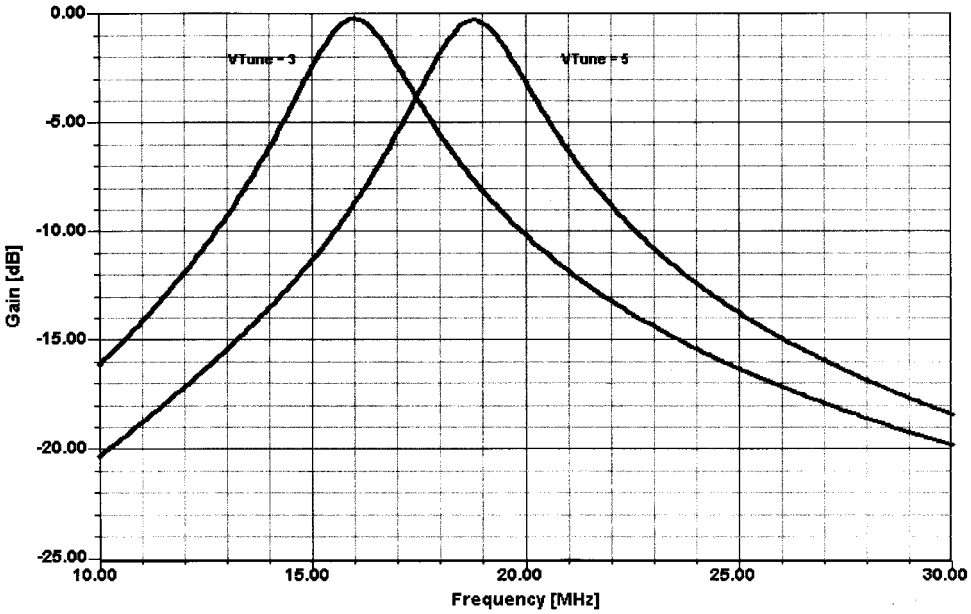


Figure 3-150 Frequency response of the filter shown in Figure 3-149. Filters of this type are intended to reduce second-order IMD by providing about 20-dB suppression at half the center frequency.

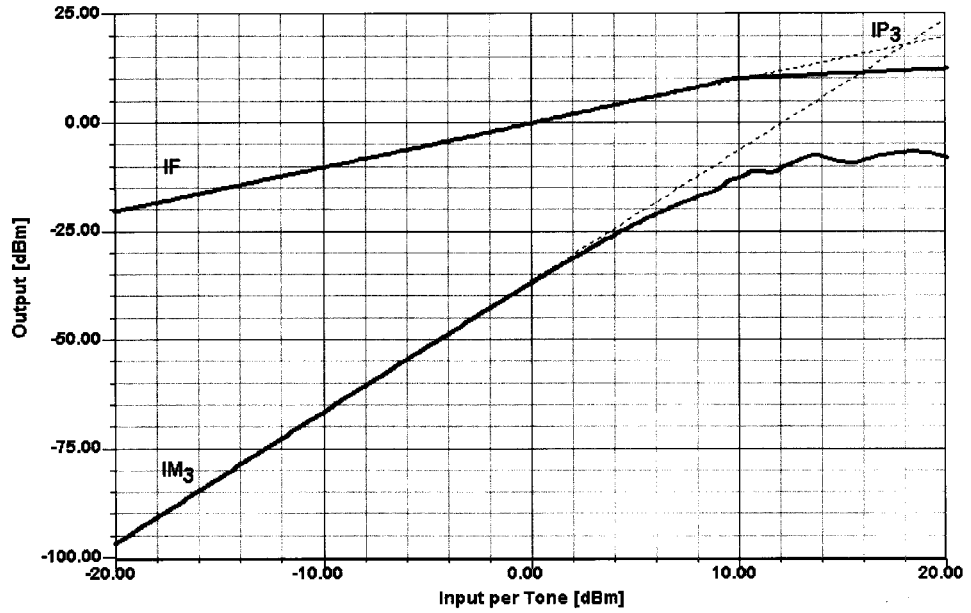


Figure 3-151 Prediction of third-order intercept point of the filter shown in Figure 3-149. The reason for the curves on the IMD plot lies in the interaction between the ten nonlinear devices.

Figure 3-151 shows the calculation of third-order intercept point for the arrangement shown in Figure 3-149. This number has now increased to +18 dBm. The difference between selection and measurement is approximately 2 dB. The reason for this more pessimistic value compared to the measured value probably has to do with slight variations of the exponent of the diodes' voltage-dependent capacitance. Therefore, the authors consider both the circuit performance and the simulation accuracy to be extremely good. This type of circuit, as mentioned earlier, has wide application in oscillator circuits. Danzeisen of Rohde & Schwarz was probably the first to have used this type of circuit by paralleling many diodes for improved performance.

3-6-4 Improving the VHF Filter

By selecting the appropriate diode combinations and circuit modifications as shown in Figure 3-152, a significant IMD improvement of the circuit shown in Figure 3-146 is obtained. A special shunt arrangement of several diodes with different exponents has been developed, which allows its value to be "adjusted." This circuit, for which a patent has been applied, showed an improvement from –2 dBm to 32 dBm. The same high operating Q was maintained. Figure 3-153 shows the frequency response after the modification. Note that a coupling slightly greater than critical ($Q \times k = 1.1$) has been selected.

3-6-5 Conclusion

After explaining some of the nonlinearities in mathematical terms, we have given some examples of voltage-tuned circuits and discussed their large-signal performance. This section has also shown that modern CAD tools can accurately predict the performance of such circuits.

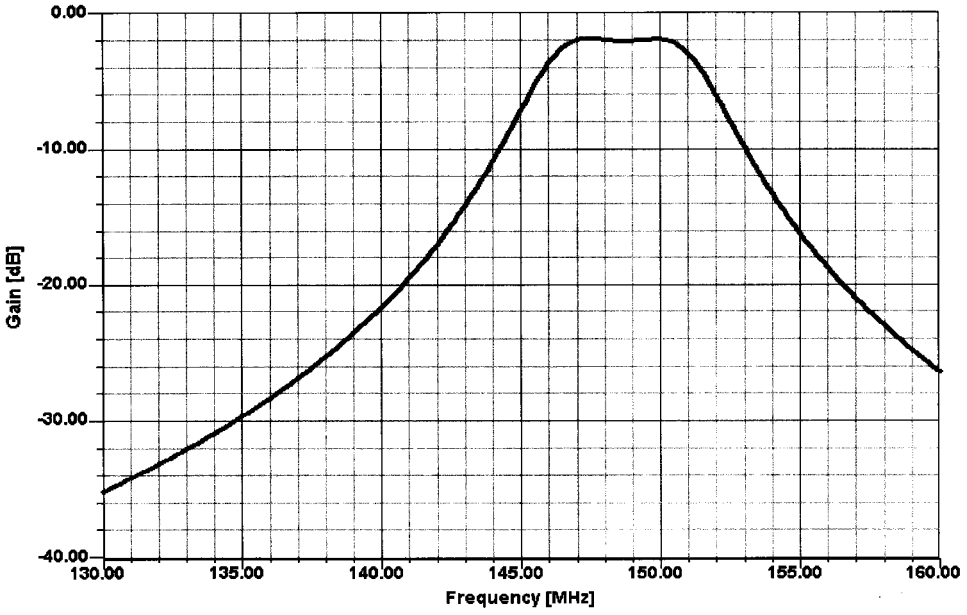


Figure 3-152 The frequency response of the improved circuit derived from Figure 3-146.

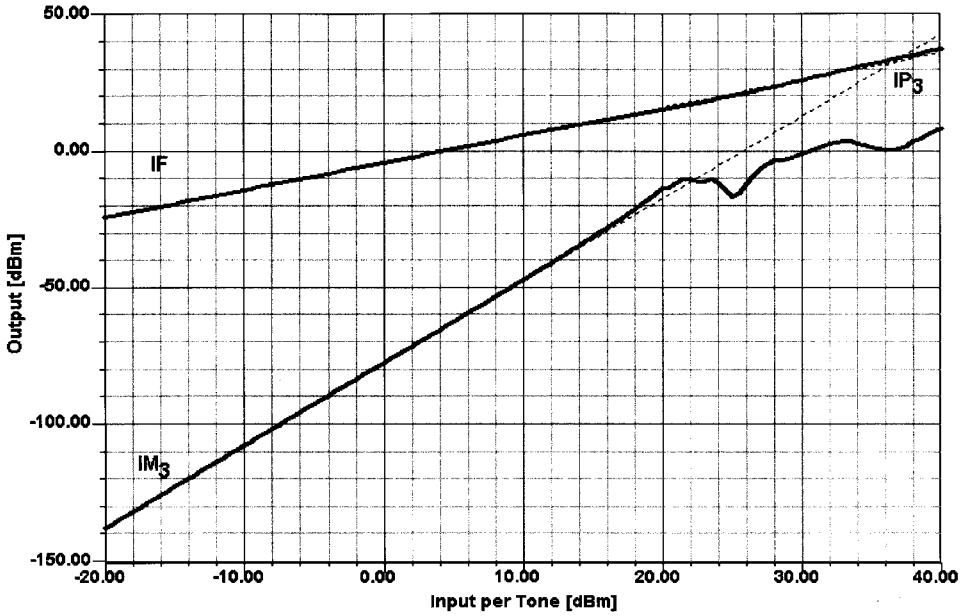


Figure 3-153 Predicted IMD performance of the improved version of the circuit shown in Figure 3-146. This type of improvement is significant for all applications. A patent is in the process of being obtained for this; therefore, the circuit cannot be disclosed at present.

3-7 DIFFERENTIAL AMPLIFIERS

The differential amplifier (Figure 3-154) goes back many years and also was the first step from TTL to ECL. The differential amplifier consists of an emitter follower (Q1) and a grounded-base stage (Q2). The constant-current generator (Q3) between ground and the emitters of Q1 and Q2 determines the current for these stages; the dc operating currents of Q1 and Q1 are the same. Most modern integrated circuits use this differential amplifier in one way or another.

If the circuit is operated about its point of symmetry, the dc component of the current through either Q1 or Q2 remains half of the emitter constant current through Q3 for all symmetrical driving signals. In addition, for input signals that are aperiodic, no even harmonics are generated. This assumes that the stage is driven symmetrically at both inputs. In our case, the differential amplifier has its second stage grounded, meaning that there is a base-to-ground capacitor, and the stage is asymmetrically fed. As the input voltage is increased, the differential stage becomes a limiter, which limits the current between plus and minus I_S, with I_S being the source current. The resulting output will become a square wave. This stage can also be used as a line receiver, translating a sine wave to an output logic suitable for driving emitter-coupled logic (ECL) stages.

One can assume that the differential amplifier has a transconductance of

$$g_m = \frac{\alpha g_{\text{in}}}{2} \tag{3-190}$$

where α is the ratio of collector to emitter current. Within the limits of the output being a sine wave, this stage has very low distortion. The output current now is determined by

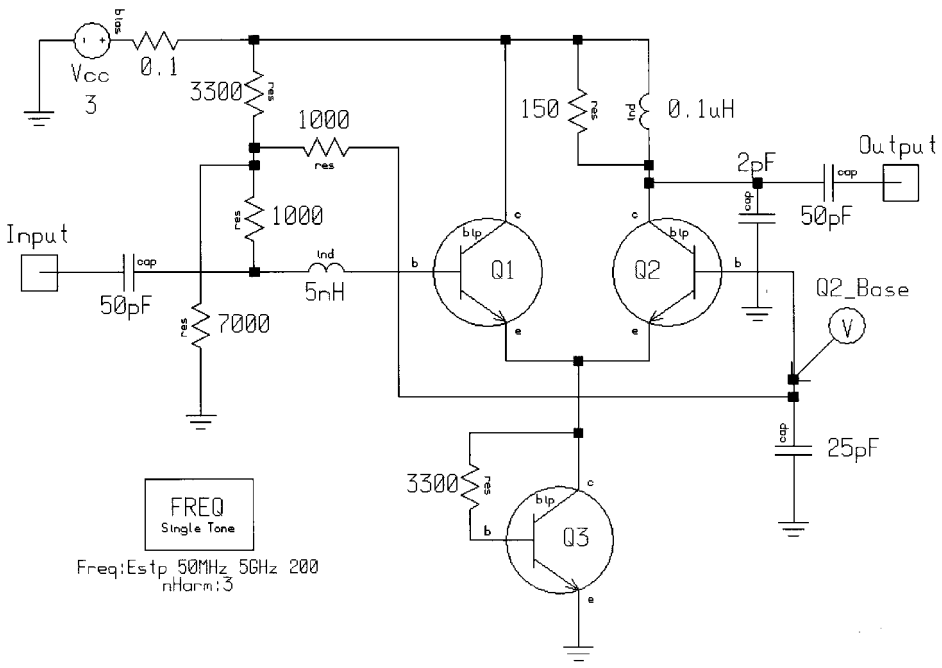


Figure 3-154 The differential amplifier schematic. The transistors are Siemens BFP420s. Figure 3-155 shows the circuit's frequency-dependent gain, matching, and noise performance.

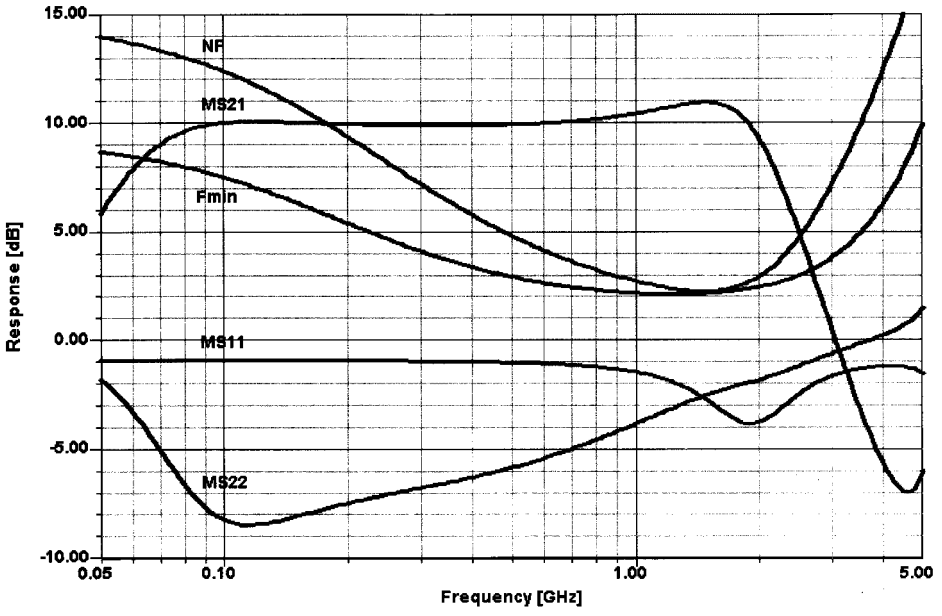


Figure 3-155 Frequency-dependent gain, matching, and noise performance of the differential amplifier.

$$i = \frac{I_{\text{in}}}{2} \tanh\left(\frac{x}{2} \cos \omega t\right) \tag{3-191}$$

where ωt is the input frequency and I_{in} is the current forced by Q3. The reason why the transconductance has to be divided by 2 comes from the fact that the current is split into two equal components for Q1 and Q2. Multiplying the new transconductance with the output load gives the voltage gain of this amplifier; its power gain can be obtained by multiplying this with $\sqrt{\text{output resistance} \div \text{load resistance}}$.

Having done the linear analysis and having learned that because of the hyperbolic tangent functions this amplifier can also be used as a line receiver, and converting sinusoidal waves to square waves, it may be useful to take a closer look at these issues. If we take a CAD oscilloscope, we see the resulting output waveform being a nonequal duty cycle output voltage, which results in high harmonic content. The amount of harmonic content can be controlled by the drive power and by the actual bias. Figure 3-156 shows the resulting output waveform.

If we now switch to a CAD spectrum analyzer, we can evaluate the harmonic content (Figure 3-157). Figure 3-158 shows the output transistor's dc I–V curves and ac load line under these conditions. Based on the previously defined conducting angle of the differential amplifier, one can optimize for harmonic content. Therefore, it is logical at this point to look at frequency multipliers, as they are sometimes needed. Finally, we are curious to see the so-called load line of the output transistor, which is, based on the interaction of all the stages, not looking the same way as the expected one, which we know from previous amplifier examples. Part of the reason for this is that for a certain voltage range, the collector voltage is in the negative region, and all kinds of saturation effects result in this surprising waveform. By reducing the input power, one would obtain a more "expected" load line. We invite readers with CAD capabilities to play with this example, as it gives a lot of insight into the operation of ECL stages.

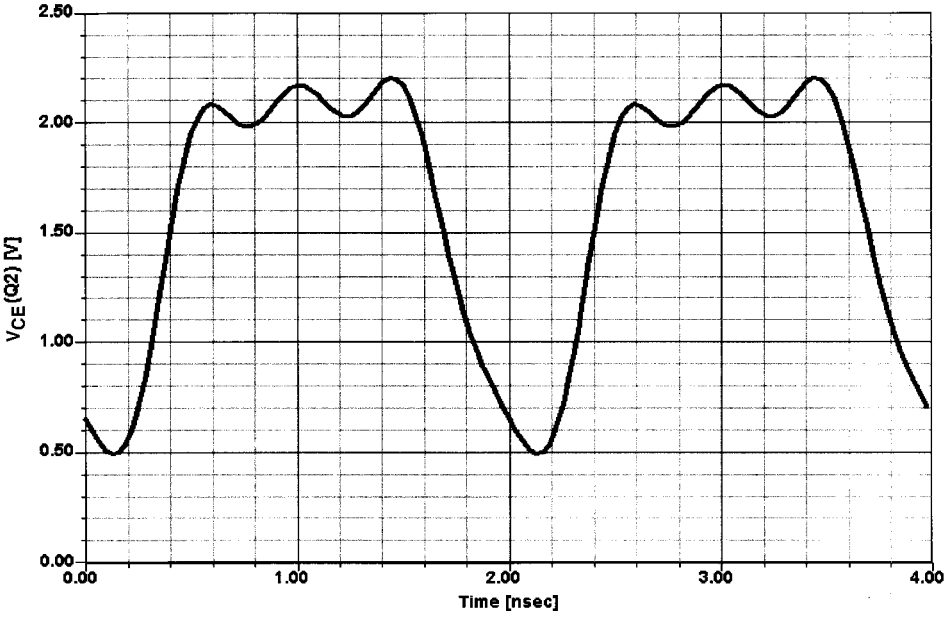


Figure 3-156 V_{CE} versus time for output transistor Q2 with a drive of 10 dBm.

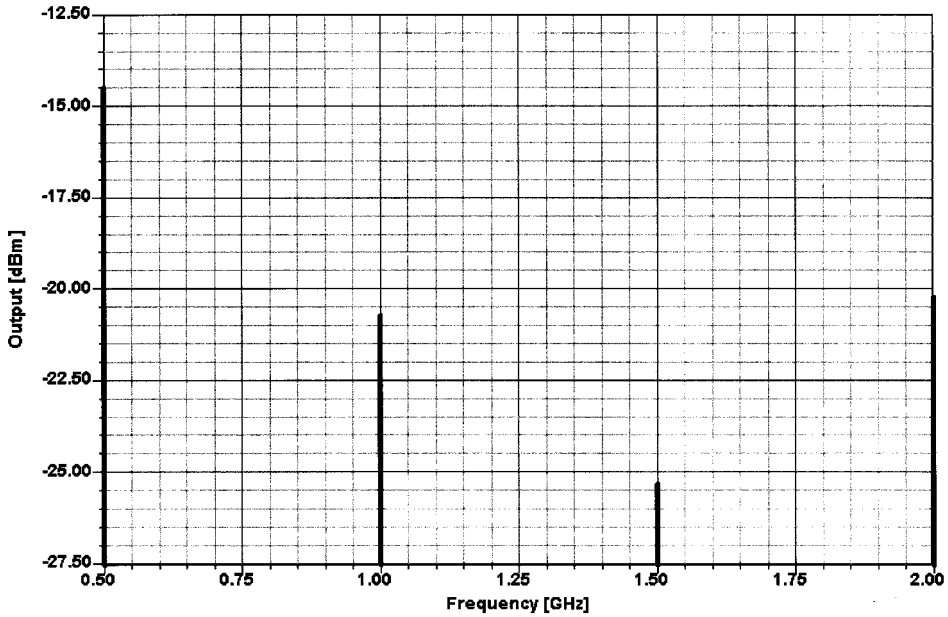


Figure 3-157 Output spectrum of the differential amplifier with a drive of 10 dBm. The biasing is not optimal for frequency multiplication.

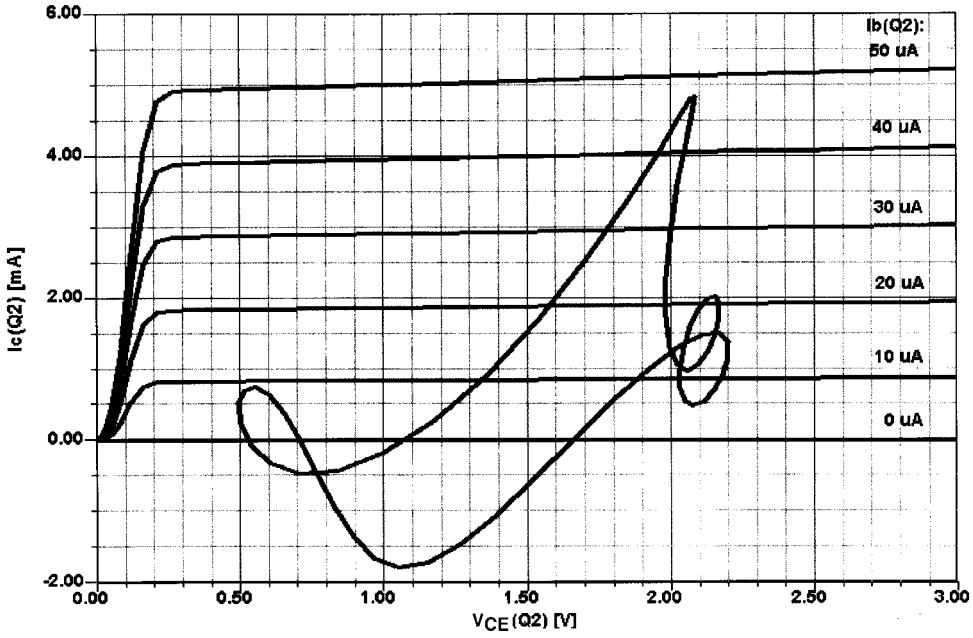


Figure 3-158 The dc *I–V* curves and ac load line for output transistor Q2 with a drive of 10 dBm.

3-8 FREQUENCY DOUBLERS

Having seen that the harmonic content invites us to bias a stage from Class A into Class B or C if our goal is a harmonically rich current, we will now evaluate a single stage that will be biased close to cutoff and driven by a relatively low frequency (500 MHz). To obtain the best results, we will use a parallel notch filter in the output, followed by a high-pass matching circuit to "catch" the desired harmonic. The input frequency is heavily suppressed by the notch filter. Based on its finite Q, there is a limited range over which this frequency can be varied with constant subharmonic attenuation. Figure 3-159 will be our test circuit.

This circuit consists of a bias transistor, Q2. This method has been discussed before and solves the temperature-dependent bias emitter thermal runaway problems through "temperature compensation." The purists among our readers may point out that the currents drawn by Q1 and Q2 are not quite identical (which would give 100% cancellation), but the approach shown is sufficient to do its job. Also, we have chosen a Class AB operation point, which is defined by Q1's 1-mA dc standing current. The conducting angle has not been optimized for the second harmonic—a task that the interested reader can do by using the conducting angle relationship shown at the beginning of this chapter. To give readers with access to nonlinear CAD some homework, we have not set this at the optimum point.

The second transistor, here called Q1 (because of its importance at RF), has a matching network at the input. This is recommended for optimum energy transfer. The matching condition can be validated by examining Figure 3-160. If we look at the higher frequency response, the input shows a possible trend of oscillation (negative loss); this could actually be eliminated by using a low-pass filter at the input instead of the high-pass filter used now. Findings of this type always make us nervous, and we hope that the nonlinear models are sufficiently accurate at higher frequencies to predict such behavior correctly. In the case of the FET doubler mentioned in Chapter 2, we found that the device modeling done by Compact (now Ansoft) gave the best results, the university answers were fairly close, and the device manufacturer data were worst. We have since become highly skeptical about device modeling unless we can validate it ourselves. Synergy Microwave has a complete set of test equipment in the laboratory, including network and spectrum analyzers with special software created by Rohde & Schwarz that can be used in addition to the Scout program to do in-house validation. This was achieved by the use of nonlinear optimization at a drive level of 0 dBm. One could say that this is similar to using a load-pull approach at the input, and then resorting to the large-signal S parameters obtained. Another note of caution: Capacitors at these low values tend to have a very large percentage error; for instance, a 1-pF capacitor, if not selected carefully, can vary between 0.7 and 1.3 pF—a 50% error. Needless to say, this is not acceptable for useful production.

The output consists of a high-pass filter and a notch filter with the purpose of reducing the subharmonic significantly. We have again done this on purpose, leaving room for the reader to come up with a more inventive output circuit like a Cauer-type (same as elliptical) low-pass filter with discrete notches at higher harmonics of the input frequency, such as 1.5 GHz, 2 GHz, and so on. In addition, we have shown three ways how to implement the inductance for the notch filter. This is equally applicable for other inductors in this value range. As frequency increases, this becomes more interesting but one needs to remember that many materials like silicon and gallium arsenide are extremely lossy. The same inductor approach has been used in Figure 5-61 of Chapter 5. For silicon, even the manufacturer becomes an issue. Figure 3-161 shows a top view of a "rectangular" inductor that actually approaches the spiral inductor; its implementation can be seen in Figure 3-162. Modeling this was a

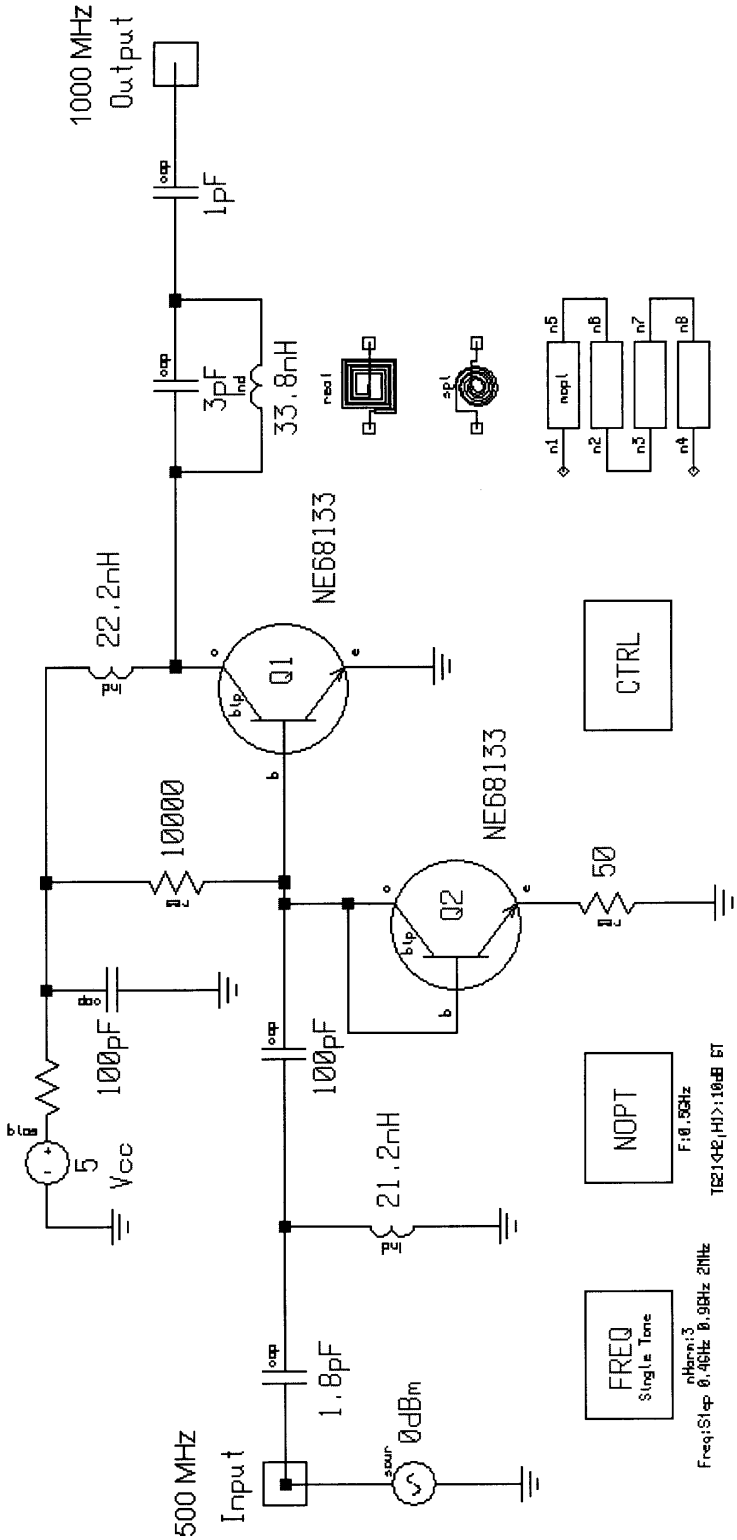


Figure 3-159 BJT frequency doubler schematic. A parallel-tuned trap in Q1's collector attenuates fundamental feedthrough.

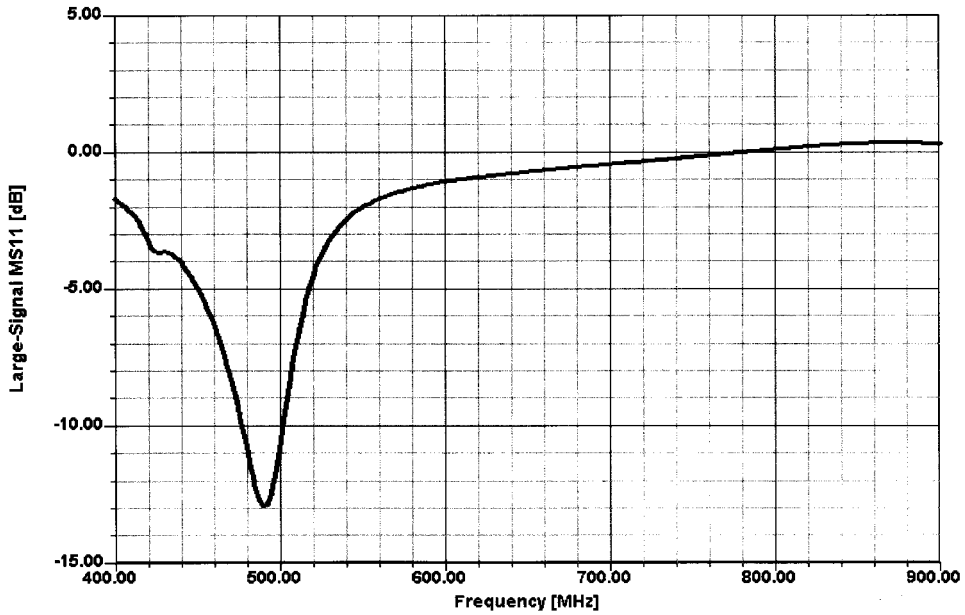


Figure 3-160 Large-signal MS_{11} for the BJT doubler. Note the positive region above 800 MHz.

nightmare, and during our Compact days, we actually had to modify the electromagnetic simulator to accommodate this and get results where measurements and simulation agreed. Figure 3-163 shows the measured and simulated Q as a function of frequency of a silicon-based inductor. In the case of other materials, such as PC board or GaAs, the models in the CAD tools are sufficient to predict the losses of the inductors. Following this initial exercise, Motorola then developed an improved silicon spiral inductor; its frequency-dependent Q is shown in Figure 3-164. Note that this Q was measured and not simulated.

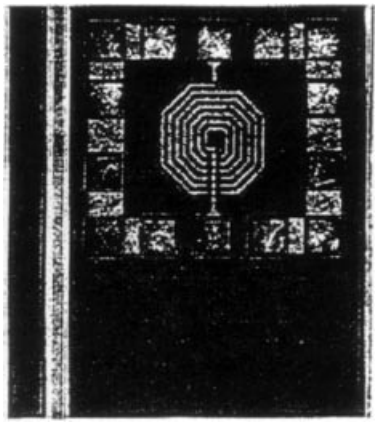


Figure 3-161 A 6.5-turn spiral inductor in silicon.

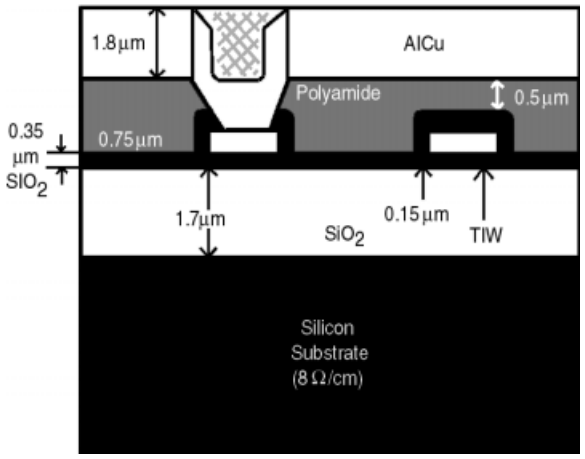


Figure 3-162 Cross section of the silicon inductor.

One of the reasons why we use a transistor multiplier is that we expect gain (because of the square-law characteristic, an FET would have been better tailored for a doubler); therefore, we have "measured" the gain of this circuit while driving it with 0 dBm. Another important factor is that, because of the way the CAD tool displays it, the gain is referenced to the 500-MHz input. Mentioning this is important; otherwise, the reader may look for the 1-GHz gain, which does not exist. Figure 3-165 shows the doubler gain.

By now we are curious to see the doubler in action, and by reviewing Figure 3-166 we can see the output spectrum. The output spectrum validates the statement that the output circuit, while doing a marvelous job of removing the 500-MHz input frequency, does not cure the problem of unwanted multiples of the input frequency. Again, a Cauer low-pass filter, while introducing some possible additional loss, should be considered for this circuit to have a bandpass response. On the other hand, many oscillators have very little harmonic suppression, and the same type of postfiltering needs to be considered. There are two-transistor solutions available also, such as push–push for frequency doubling (push–pull would

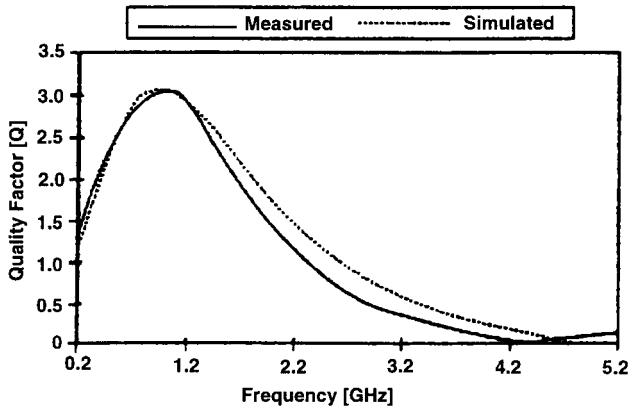


Figure 3-163 Measured and simulated *Q* versus frequency for the simulated inductor.

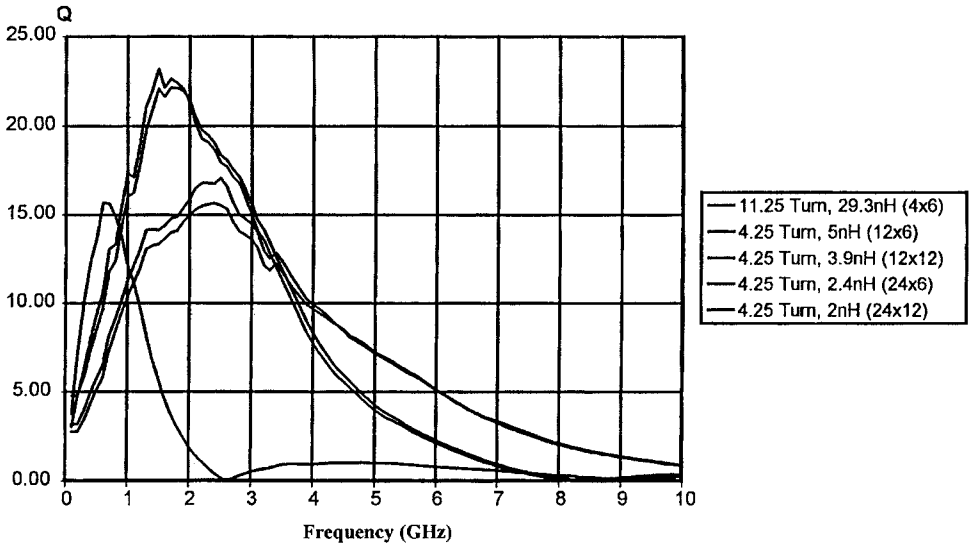


Figure 3-164 *Q* versus frequency for Motorola improved inductor.

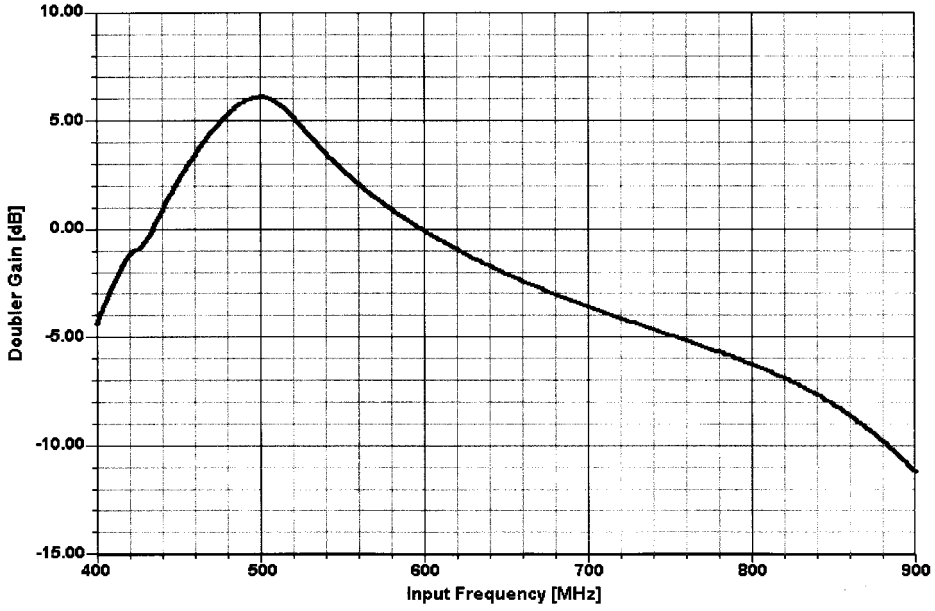


Figure 3-165 Frequency-dependent gain of the BJT doubler.

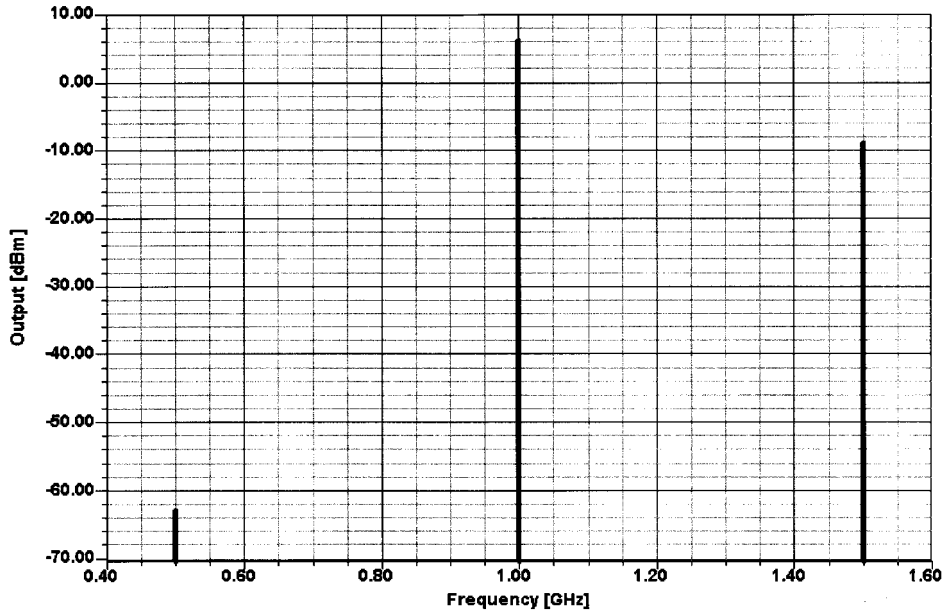


Figure 3-166 Output spectrum of the BJT doubler.

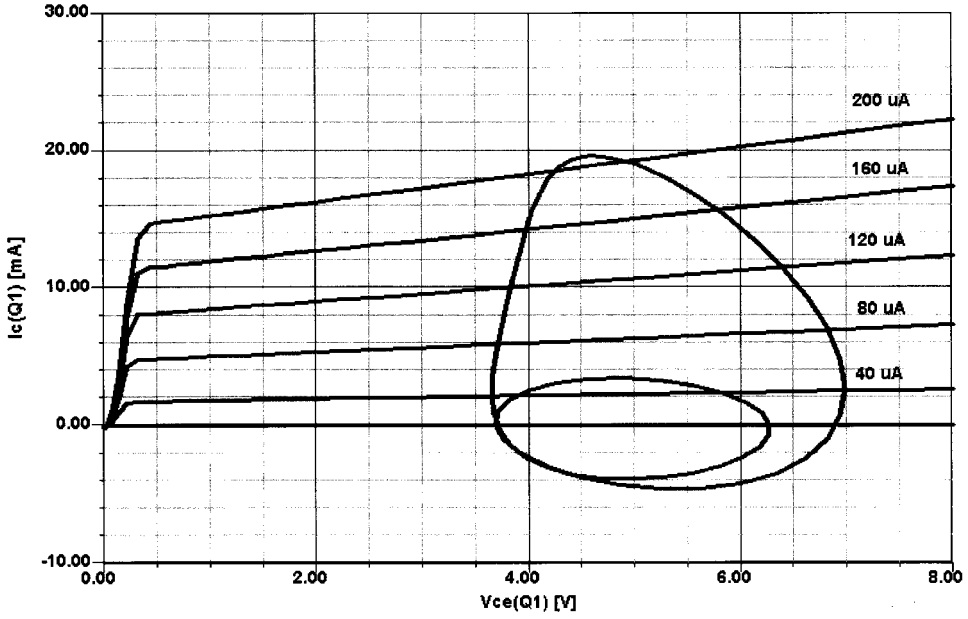


Figure 3-167 The dc *I–V* curves and ac load line for RF transistor Q1 in the BJT frequency doubler.

eliminate the desired second harmonic quite well). Actually, two diodes could also have been used for doubling, but since this requires transformers and a postamplifier, we have shown the preferred application. The only advantage of the diode application is that it is wideband, while this BJT circuit, as we have just learned, must be made narrowband for reasons of spectral purity.

Finally, we are curious to see the actual operating state of this amplifier, and this can be done by analyzing the load line (Figure 3-167). This is as complicated as the analysis for the differential amplifier. Because of the stored energy in the various tuned circuits, the load line looks quite impressive and would probably be difficult to predict without a nonlinear CAD tool.

3-9 MULTISTAGE AMPLIFIERS WITH AUTOMATIC GAIN CONTROL (AGC)

Having discussed the basic differential amplifier, we will now evaluate the performance of the MC1350/1490 video amplifier chip from Motorola. Most of the circuitry in the chip is self-explanatory. Figure 3-168 shows its inner circuit, and Figure 3-169 show a recommended application. By grounding one side of the input, we effectively get the same performance as the differential amplifier just discussed, with one difference being that the input is now fed to the constant-current generator. To minimize the distortion at the two lower transistors, they are decoupled by a 66-Ω resistor. Figure 3-170 shows the frequency-dependent gain, matching, and noise figure. Since this particular device has been specifically designed for AGC purposes, we also show in Figure 3-171 its gain reduction versus applied AGC voltage.

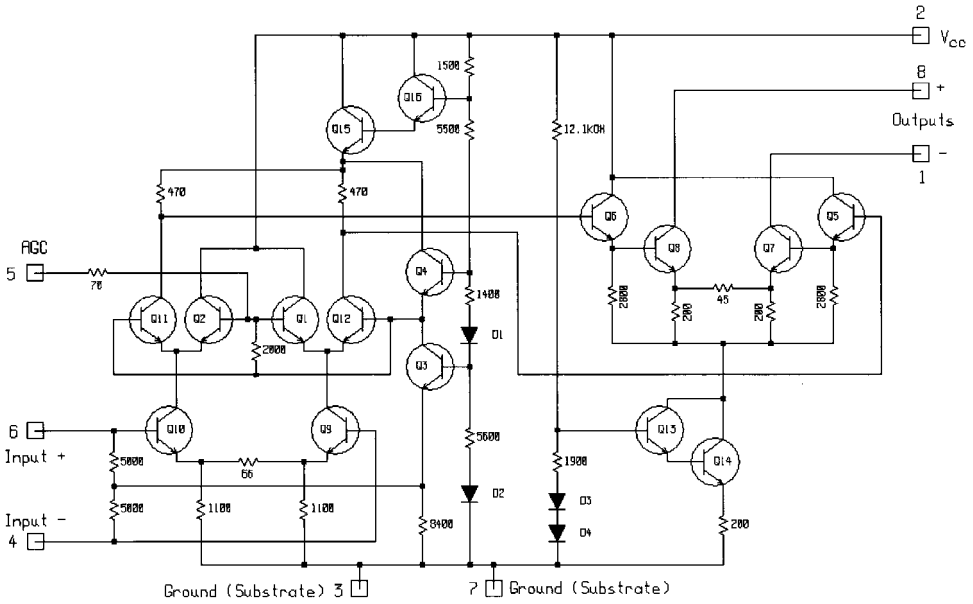


Figure 3-168 Schematic of the Motorola MC1350/1490 entered for simulation.

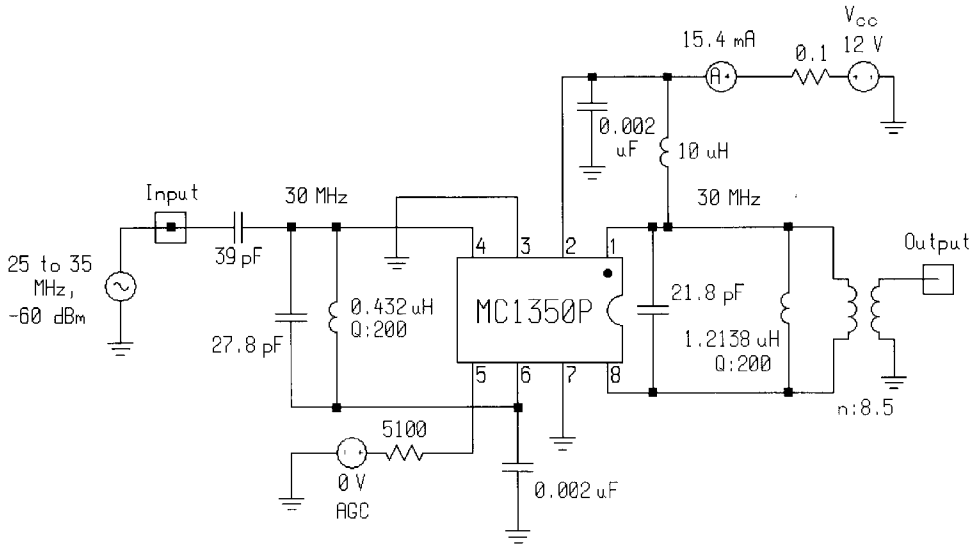


Figure 3-169 MC1350/1490 30-MHz application circuit.

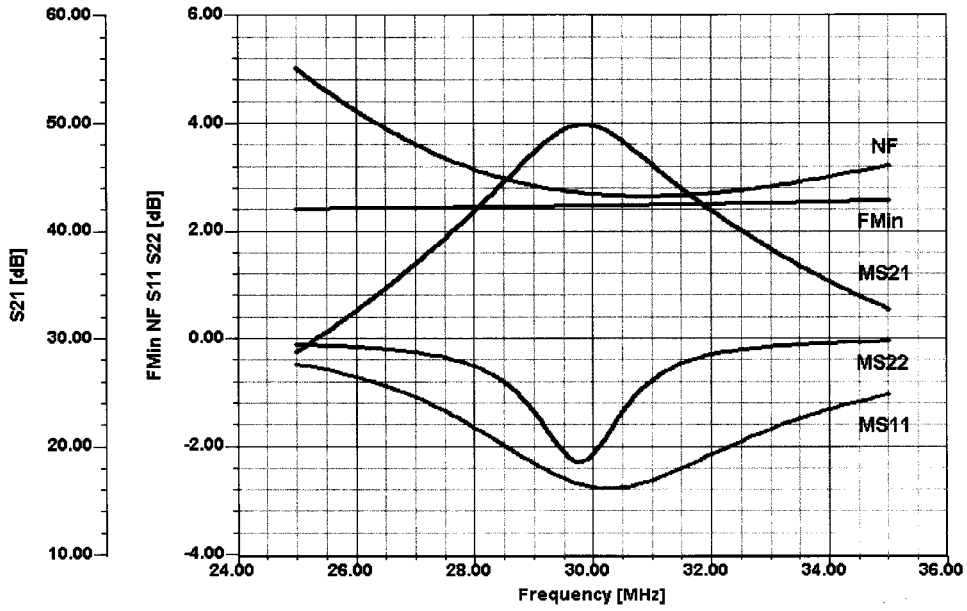


Figure 3-170 Frequency-dependent gain, matching, and noise performance of the MC1350/1490 in the circuit of Figure 3-169.

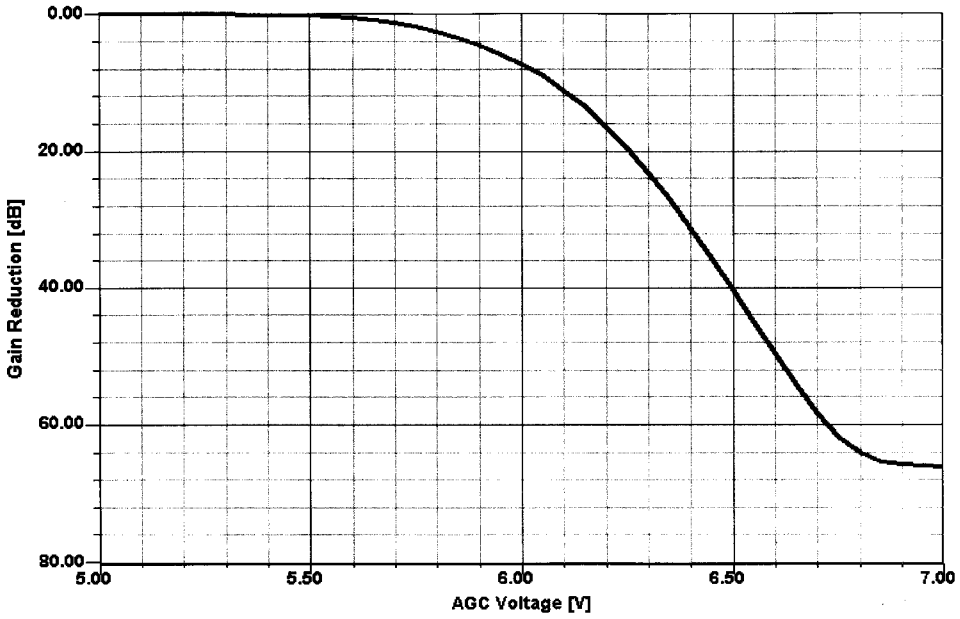


Figure 3-171 Gain reduction versus applied AGC voltage for the MC1350/1490 IC.

3-10 BIASING

A lot has been written about biasing, and since the purpose of this book is not to focus too much on dc, the following examples including RF consequences have been selected and considered sufficient. Early application reports by Motorola and others have provided a wealth of data. As a side effect of the introduction of the wireless era, we find that the engineers are so pressed for time that the number of complete application reports has approached zero. This is one of the reasons why we have tried to minimize the mathematical basis of this book (some is still needed), while still providing practical assistance. Figure 3-172 shows a textbook type of approach for RF.

In order to decouple the transistor from the biasing, we use RF chokes and dc decoupling (bypass) capacitors. This is the technique used in previous examples, so one might ask, "What's new?" As frequency increases, these inductors are either not manufacturable or have such a low Q that their use becomes questionable. This is the point where one may introduce the so-called distributed elements.

Figure 3-173 shows the very same circuit but resorting to distributed rather than lumped elements. The newly introduced elements (part of any good, up-to-date CAD tool; see element library for how to use their physical element descriptions) are described next.

- *Transmission Line.* Any printed connection between two points on a circuit board is a transmission line (Figure 3-174). Its characteristic electrical impedance is a function of the square root of the dielectric constant (ε_r), the width, metallization, thickness and height above substrate of the line, and the loss tangent of the substrate. Since lines frequently have to be laid out in the form of curved connections or have a bend in their direction, we have to add elements capable of describing the high-frequency conse-

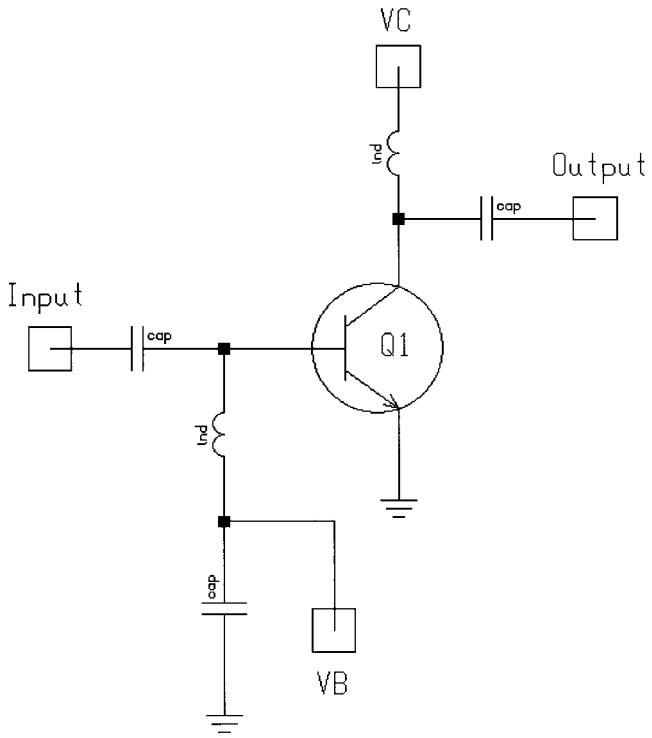


Figure 3-172 Simple BJT RF amplifier with lumped elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

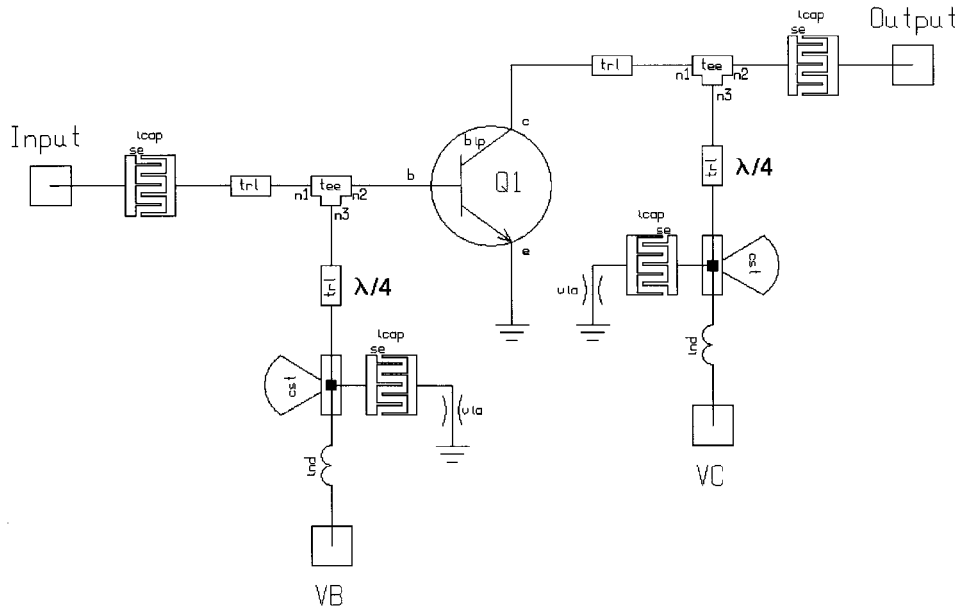


Figure 3-173 Simple BJT RF amplifier with distributed elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

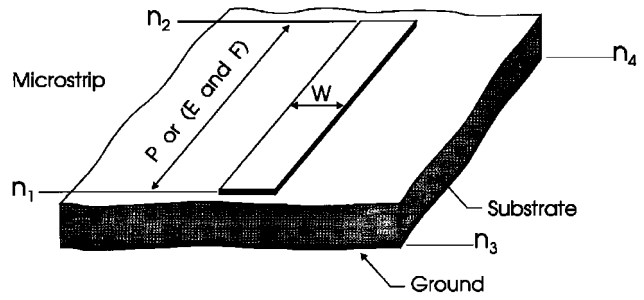


Figure 3-174 Transmission line in microstrip.

quences of such connections. Figures 3-175 and 3-176 show mitered and radial bend elements that perform this function.

- T *Junction, Cross, and* Y *Junction.* By the time a point like a collector or base, or its FET equivalent, spreads out into connections with other elements, we need additional modeling capability to describe ⊤ connections, crossings, and Y junctions. Figures 3-177, 3-178, and 3-179 show the way in which these connections must be modeled.

If the need exists, the standard inductances must be replaced with a transmission line whose length is λ/8 at the operating frequency. At higher frequencies, these transmission lines, however, then go into λ/4 resonant mode and later become capacitive. This type of design makes it fairly narrowband. A way around this is the use of printed inductors, as shown in Figures 3-180 and 3-181.

These inductors have a self-resonant frequency similar to the transmission line mentioned above, but the safety margin is significantly higher.

Talking about printed inductors, a logical extension of this is the printed transmission-line-based transformer, as shown in Figure 3-182. One can consider this as two interlaced rectangular inductors and, based on the substrate material, they are useful over a wide frequency range. Besides being used as a transformer, they can also be used to transit from

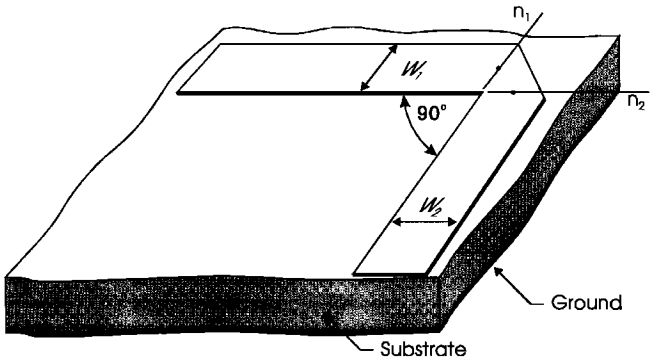


Figure 3-175 Mitered bend.

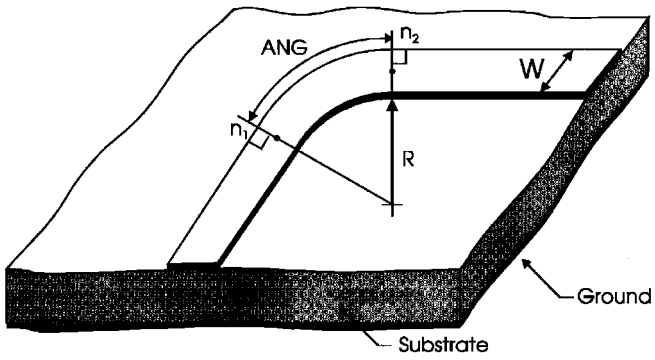


Figure 3-176 Radial bend.

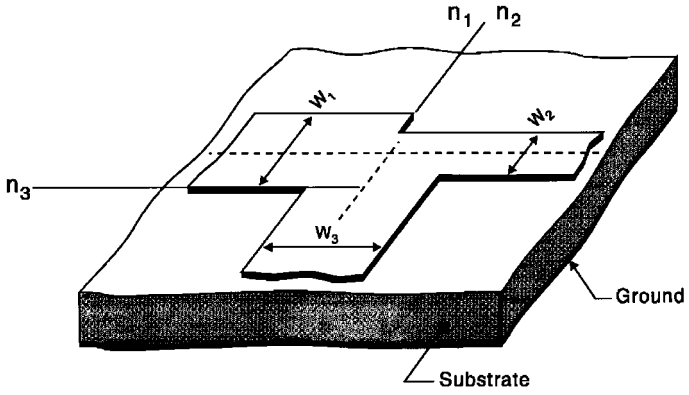


Figure 3-177 T junction.

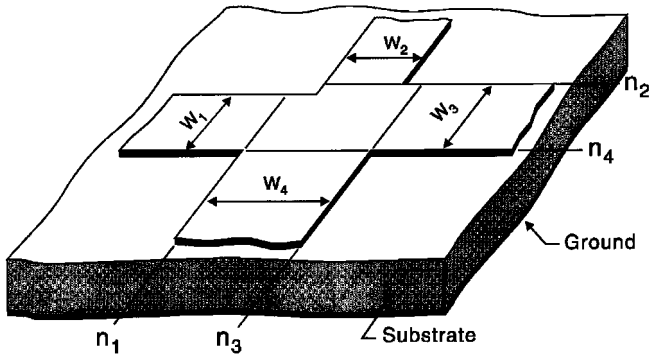


Figure 3-178 Cross.

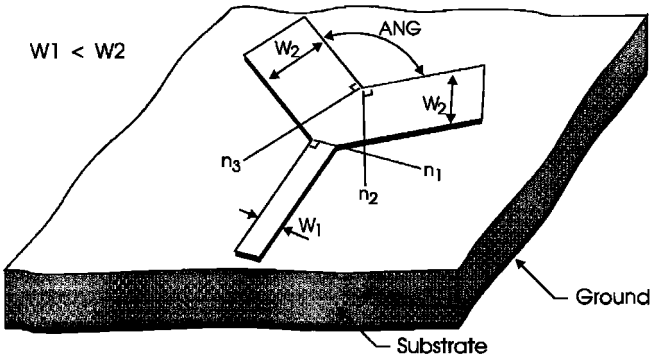


Figure 3-179 Y junction.

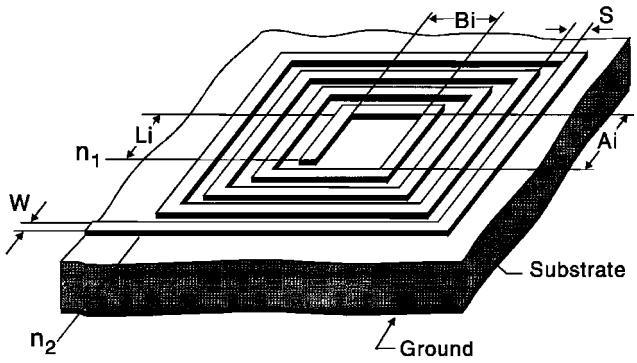


Figure 3-180 Rectangular inductor.

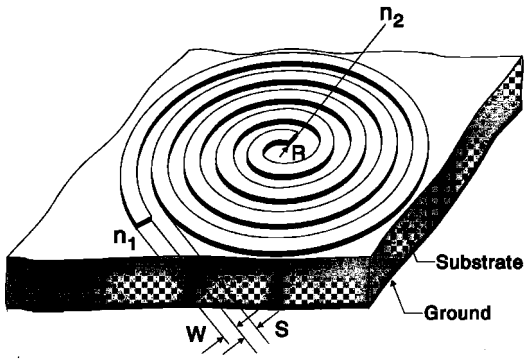


Figure 3-181 Spiral inductor.

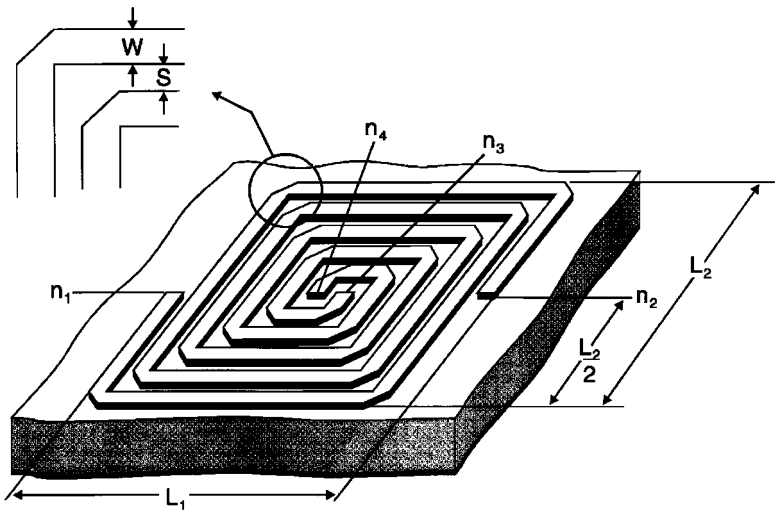


Figure 3-182 Transformer in microstrip.

unbalanced to balanced transmission provided that the difference in length from a connection point of view does not cause any problems (this is a layout issue).

A popular form of combining stages is the so-called Lange coupler (Figure 3-183) invented by the German Julius Lange. It is one of the major contributions in wideband applications.

Useful application of the Lange coupler probably starts at 4 GHz. It consists of parallel transmission lines with the appropriate connections as shown. Lange couplers are typically built with four, six, and eight fingers. The Ansoft Serenade product has a Lange coupler synthesis program that can be used to gain more insight into this coupler's application. We assume that other modern software has similar capabilities.

Where meander-type inductors are necessary, a neat way to implement and simulate them is to use the multiple coupled lines element (Figure 3-184) of the Serenade product, which both quickly and accurately calculates the behavior of the meander, including self-resonances and losses. We made use of this arrangement in our previous examples.

- *Interdigital Capacitors.* The issue of tolerances of small capacitors already has been brought up. The interdigital capacitor can be made on printed circuit board material as

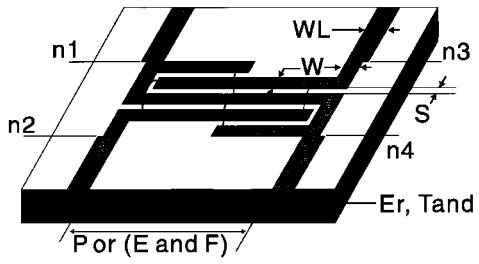


Figure 3-183 Four-strip Lange coupler.

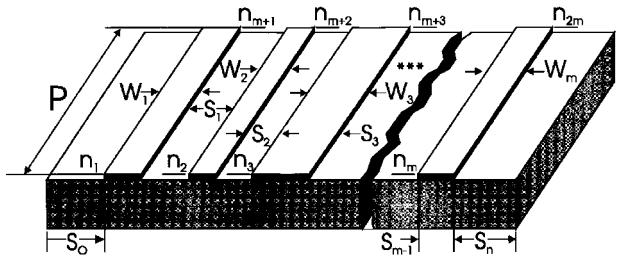


Figure 3-184 Multiple coupled lines element in microstrip.

well as gallium arsenide, and if its dimensions are continuous with the transmission-line width it does not cause any abrupt changes in impedance. This type of capacitor permits one to obtain very small values. An alternative to this is the use of transmission lines with a length of $\frac{3}{8}\lambda$. We learned earlier that a transmission line is inductive at frequencies below its fundamental resonance, is resistive at below resonance, and becomes capacitive above resonance. Again, bandwidth is also an issue. An interdigital capacitor consists of a number of parallel fingers as shown in Figure 3-185, and its capacitance can be varied by adjusting the number of fingers and their spacing. The advantage of the interdigital capacitor compared to discrete components is its low variation in value.

- *Radial Stubs.* The radial stub (Figure 3-186) is not much different from a $\lambda/4$ resonator, but its bandwidth is much greater than a simple transmission line. This is another way to ground the "cold" side of a transmission line or part of a circuit that needs to be grounded for RF. Of course, the interdigital capacitor comes in a version that is a combination of a capacitor and a via hole.
- *Via Holes.* The "cold" end of the transmission line, being considered either an inductor or a capacitor, needs to be connected to a sufficiently large copper backplane. One very efficient way to do this, especially if there is not enough copper left on the top of the board, is the use of via holes (Figure 3-187). One could theoretically generate

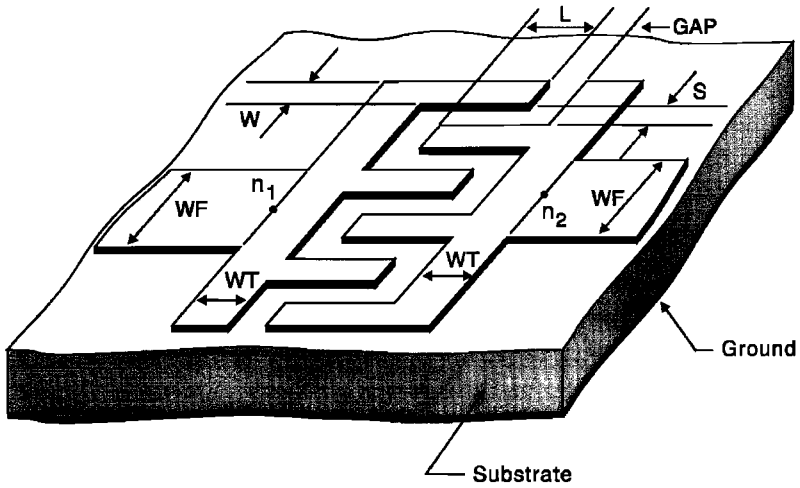


Figure 3-185 Interdigital capacitor.

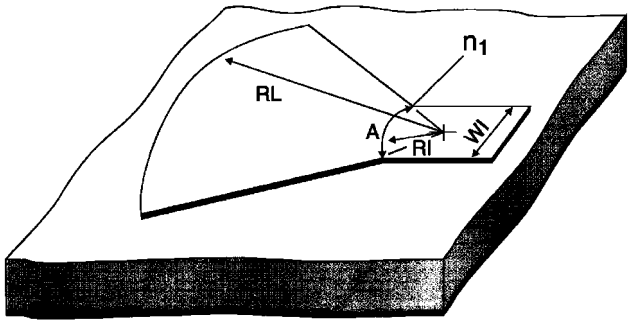


Figure 3-186 Radial stub.

a via hole with a rivet, but most manufacturing processes don't allow this; the normal solution is to use plated-through holes left open. In PC boards, via holes are typically cylindrical; on substrates like GaAs, they may be conical.

- *Correction Elements.* Although the behavior of an actual circuit proceeds regardless of our ability to measure and describe it, we do not enjoy this luxury in simulating circuit behavior in software. In a simulator, effects that are insufficiently described will be inaccurately simulated. For instance, segments of high-impedance transmission line (e.g., 120 Ω) are frequently used for dc feeds and RF chokes. In predicting the effect of a transition from 120-Ω line to 50-Ω line, a simulator must be alerted to the discontinuity so it can do the necessary mathematical corrections to account for the impedance jump. To do this, a specific circuit element, the STEP (Figure 3-188), must be inserted between the 120-Ω and 50-Ω line elements in the simulation circuit file. In addition to substrate data, we characterize a STEP by providing the widths of its input and output lines. The element itself has no physical length.

A similar correction is necessary if a transmission line is used as a resonator or just "left open" at one end. Such a transmission line tends to radiate and, because of its high-impedance properties, it reacts differently as far as its electrical length is concerned. A zero-length one-port element, the OPEN (Figure 3-189), must be added to such a line for mathematically correct calculation.

We end this excursion into distributed elements here. Interested readers may want to obtain a CD containing the element library of their favorite CAD tools that combine capabilities in the microwave area with a full set of distributed circuit elements. It is most important to keep

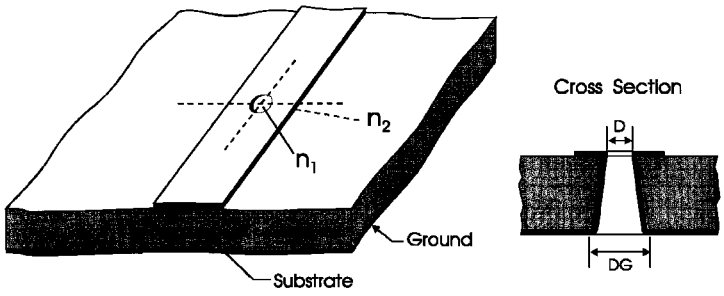


Figure 3-187 Via hole.

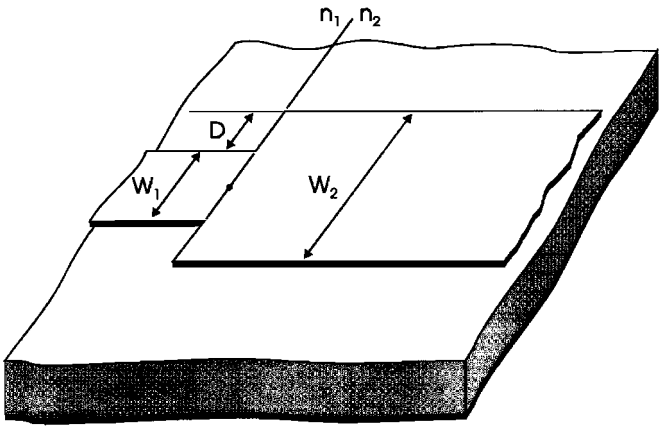


Figure 3-188 The STEP element tells the simulator to calculate the effects of joining transmission lines of differing characteristics.

in mind that, as frequency increases, we rapidly move into the area where we must consider all these distributed elements to achieve accurate simulations—even if doing so makes simulation a painful and time-consuming effort.

Finally, anyone who adventures into this area *must* obtain a foundry manual from the company that will build the integrated circuit or hybrid under design. There are basically two foundries, one applicable for MOS technology and one for GaAs. In the case of the bipolar transistor, the foundry service is not yet well-established. While we were able to interest one German company and one U.S. company in generating a custom bipolar IC within the activities of this book, the only "open" foundry we know of for bipolar (actually HBT) technology is the one operated by TRW. As of 1999, this seems to be the leading company for BJT-related products, and many currently available ICs have been developed with this foundry service.

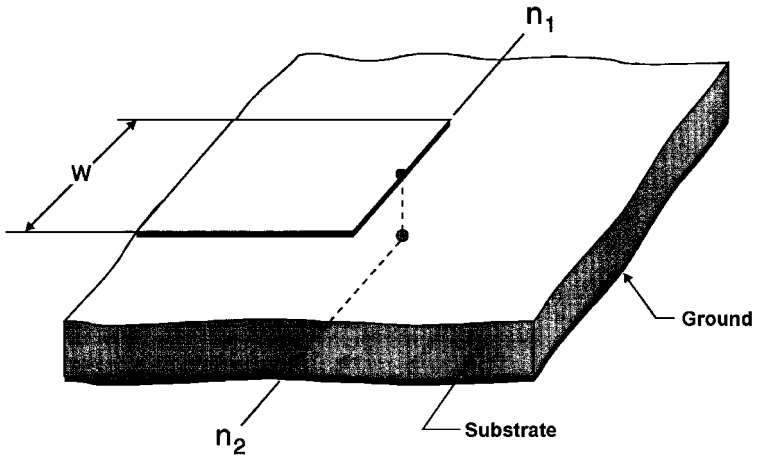


Figure 3-189 The OPEN element tells the simulator to calculate the effects of leaving the end of a transmission line unconnected.

We mention the issue of foundries here again because each foundry has its own proprietary approach to modeling discontinuities. The availability of a foundry service somewhat eases the requirement that a designer be fully up to speed on the nuances of discontinuities, because a foundry's designer service will help customers account for all relevant parasitics or discontinuities in their designs. In addition, there are tables of *S* parameters for standard cells of either capacitors, resistors, or inductors. The designer may then be forced to adjust the circuit so that it will work with a particular inductance value or value of another component within the resolution of the table that describes these elements. Information on the active parts, such as diodes and transistors, was given in Chapter 2.

3-10-1 RF Biasing

Applying the knowledge we have just acquired about distributed elements, Figure 3-190 shows a somewhat exaggerated case of using the various elements. While it is consistent with the simple RF case with which we started, modeling of this type is certainly necessary at higher frequencies. Many but not all of the elements we have just described are used, but we have not solved the question of the actual dc biasing.

3-10-2 dc Biasing

The following discussion is the recommended procedure for biasing a transistor. We will start with the collector current (in our example, Figure 3-191, 10 mA). To separate RF and dc, we have a combination of a load resistor (R_L) and an inductance (L_C). Information about this operation was given already in Section 3-2-6. The emitter resistor should have a voltage

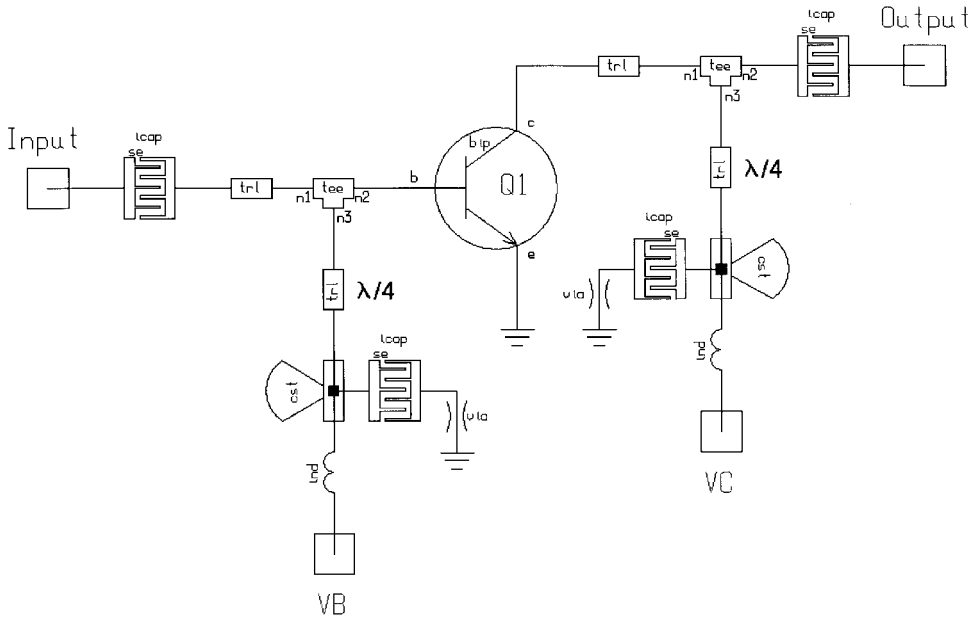


Figure 3-190 Simple BJT RF amplifier with distributed elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

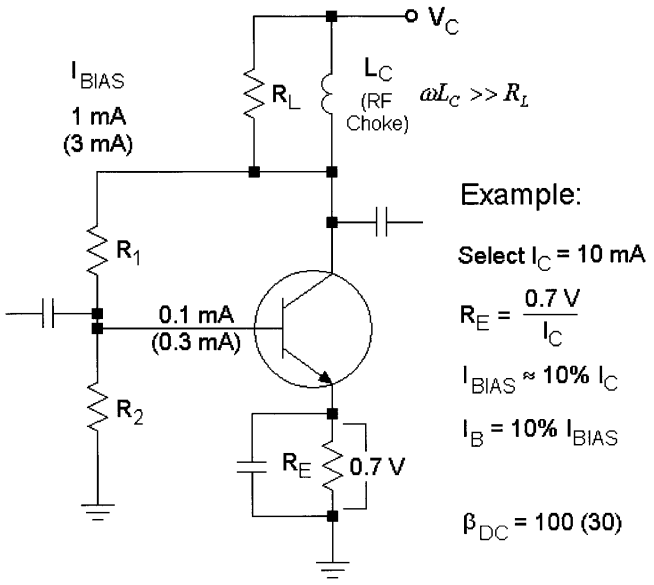


Figure 3-191 A dc biasing example.

drop of about 0.7 V, and therefore the emitter resistor is $0.7/I_C$. The next assumption is that the transistor has a dc current gain of 100, and therefore we decide to have the bias resistor chain (R_1 and R_2) draw 10% of I_C, or 1 mA. Assuming a collector voltage of 5 V, the value of $R_1 + R_2$ has to be 5 kΩ to result in the required 1 mA. Because the base current now is 10% of the divider current, the base will draw 0.1 mA. The voltage across R_2 has to be the voltage drop across R_E (0.7 V) + V_{BE}, which is also 0.7 V, or 1.4 V. Given the 1 mA, R_2 = 1.4/1E–3 = 1.4 kΩ. R_1 therefore is 5 kΩ – 1.4 kΩ = 3.6 kΩ. This concludes the calculation.

In the case of a dc beta of only 30, we need to raise I_{BIAS} from 1 mA to 3 mA since the base current now will be 0.3 mA. This changes the resistor values. $R_1 + R_2$ = 5 V/3 mA = (rounded) 1.6 kΩ. Following the same approach, we need 1.4/3E–3 = (rounded) 470 Ω. This makes R_1 = 1600 Ω – 470 Ω = (rounded) 1100 Ω. These roundings were necessary because resistors can only be bought in certain resistance steps.

However, for low battery voltages, the 0.7 V between the emitter and ground cannot be wasted. Therefore, a dc stabilization circuit, to which we have already referred, will be considered next. Figure 3-192 shows a combination of an RF transistor, Q1, and a dc stabilization transistor, Q2. The most important thing about this circuit is the fact that the difference between the supply voltage and the collector of Q1 is 0.3 V, which is approximately half of the voltage we had to "waste" across R_E in the above-mentioned circuit. Plus, the voltage drop is now in the collector, and given the fact that we feed Q1's collector through the inductor of a quasi-tuned circuit rather than through only a resistor, we get a much higher voltage swing. The base–collector diode temperature dependency of Q2 is compensated for with the silicon diode feeding the base. It is also fascinating to see that, because of the low currents, both the voltage drop of the diode and V_{BE} of Q2 are about the same, and yet Q1, operating at a reasonable current, shows the expected 0.7–0.8 V V_{BE}. We have supplied all the necessary dc voltages to give the reader an incentive to calculate the mechanism of this stabilization. The key equation to this is to determine the current in the voltage divider feeding

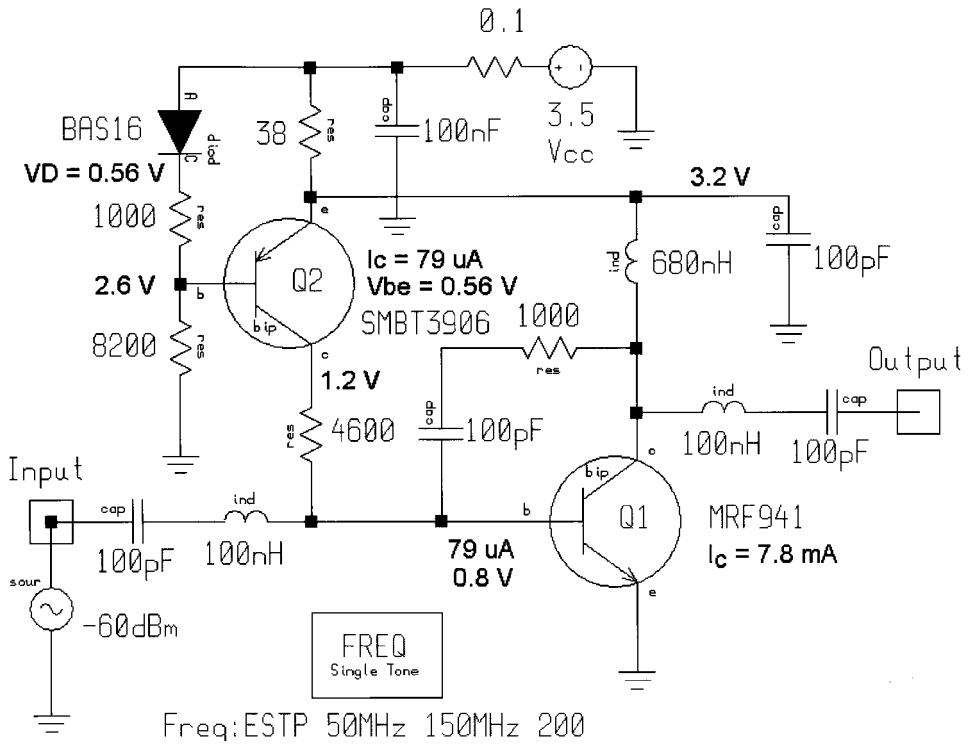


Figure 3-192 An RF amplifier with active biasing. (After the low-noise amplifier circuit in [29].)

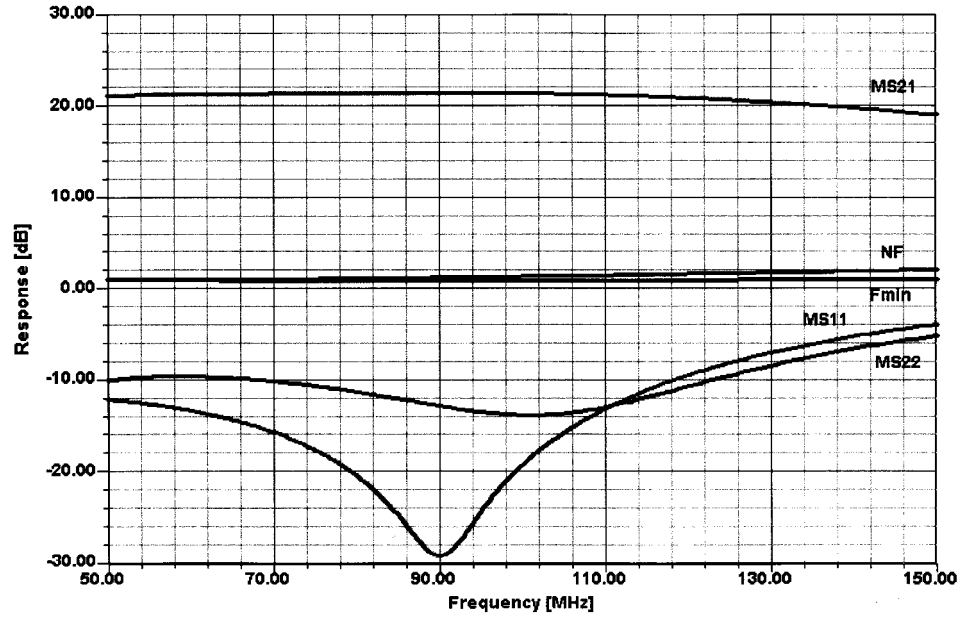


Figure 3-193 Frequency-dependent gain, matching, and noise responses for the amplifier in Figure 3-192.

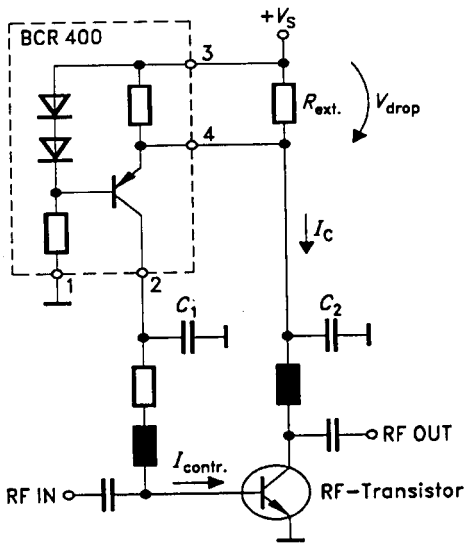


Figure 3-194 The Siemens BCR400 active bias controller can be applied to FETs as well as BJTs.

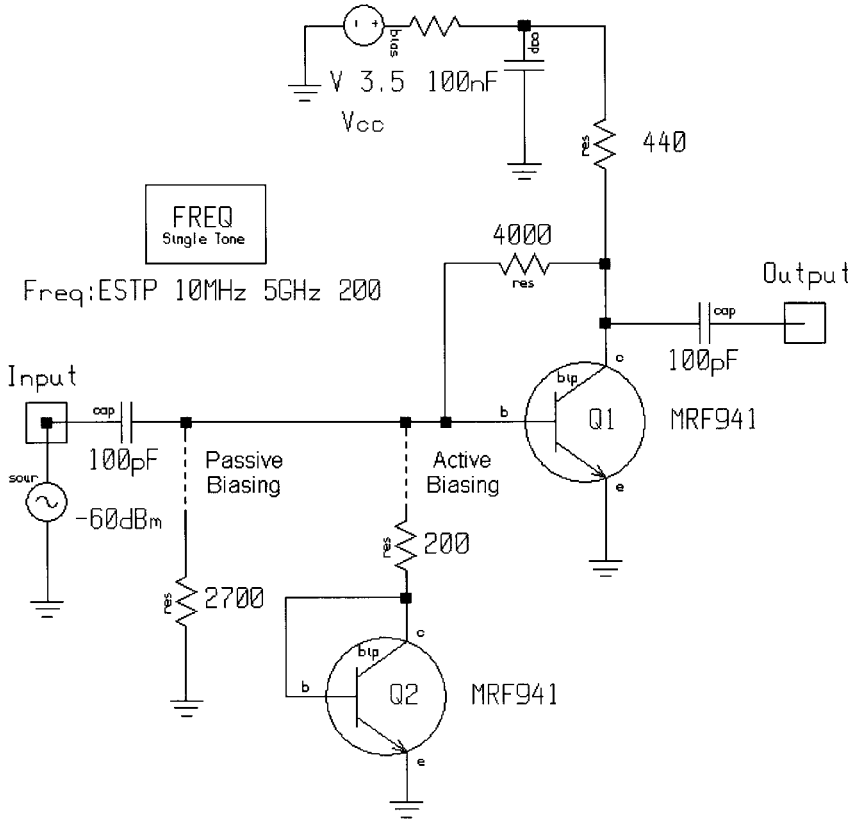


Figure 3-195 Amplifier stage (I_C = 3 mA) showing active and passive bias alternatives.

Q2, and proceeding from there. On the other hand, having a nice CAD tool that provides insight into the dc voltages and allows the addition of dc voltage probes makes life much easier. To prove that this circuit actually works, Figure 3-193 shows its frequency-dependent gain, matching, and noise figure performance. It is based on a Motorola application, which had excessive voltage drop for Q1 and which we have improved to be suitable for low-voltage operation [29].

Integrated active bias solutions are also available. Figure 3-194 shows the Siemens BCR400 active bias controller applied to a BJT; it can also be applied to FETs and TR switching diodes, as we saw in Figures 3-57 and 3-59, respectively.

3-10-3 dc Biasing of IC-Type Amplifiers

In integrated circuits, we do not have the unlimited flexibility described earlier. This results frequently in constant-current sources or transistors being used as diodes for biasing purposes. Our first suspicion is that this is going to cause a lot of noise in the IC; Figure 3-195 shows a test example.

As suspected, choosing the same bias point of 3 mA, the noise figure using active bias is quite a bit higher than the passive one, as shown in Figure 3-196.

3-11 PUSH–PULL/PARALLEL AMPLIFIERS

Since it is practically impossible to obtain $S_{11} = S_{22} = 0$, using a combiner at the input and output reduces the problem. Figure 3-197 shows the 3-dB hybrid approach. It has a practical

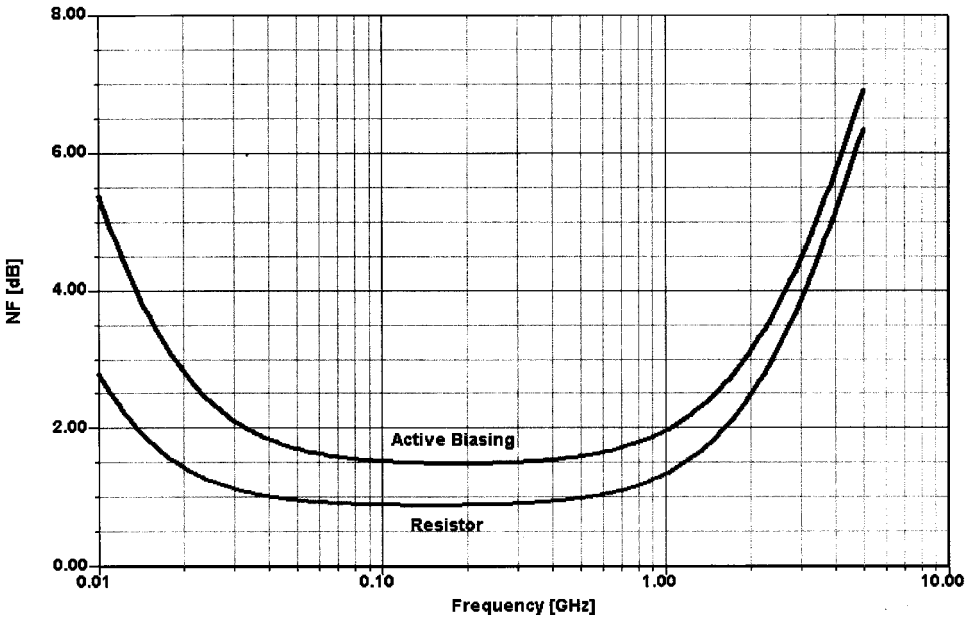


Figure 3-196 Frequency-dependent noise figure of the amplifier circuit shown in Figure 3-195. The active biasing curve is considerably noisier.

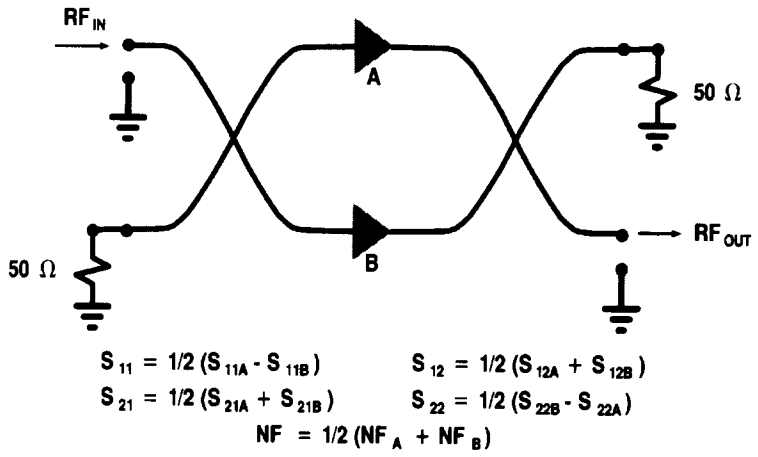


Figure 3-197 Combining power with 3-dB hybrid couplers.

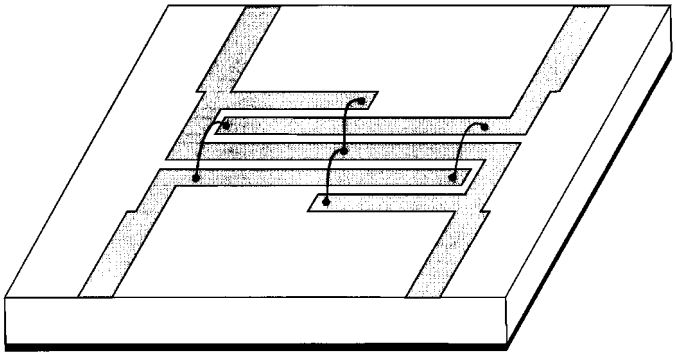


Figure 3-198 Four-strip version of the Lange coupler, a broadband 90° hybrid.

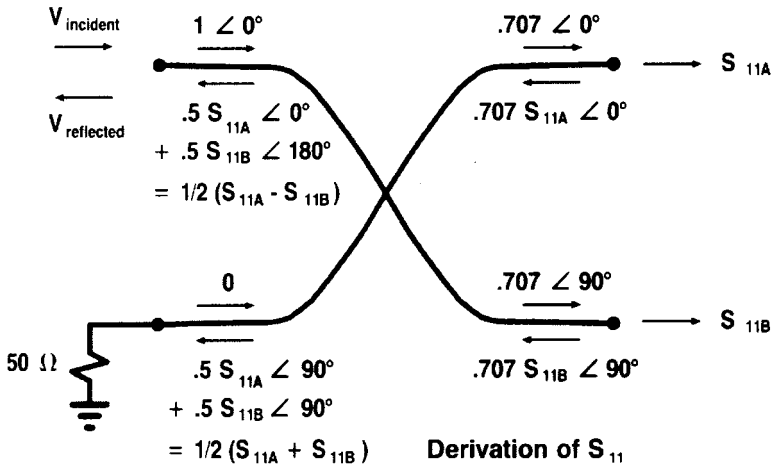


Figure 3-199 3-dB hybrids. This analysis method may be applied to any coupler.

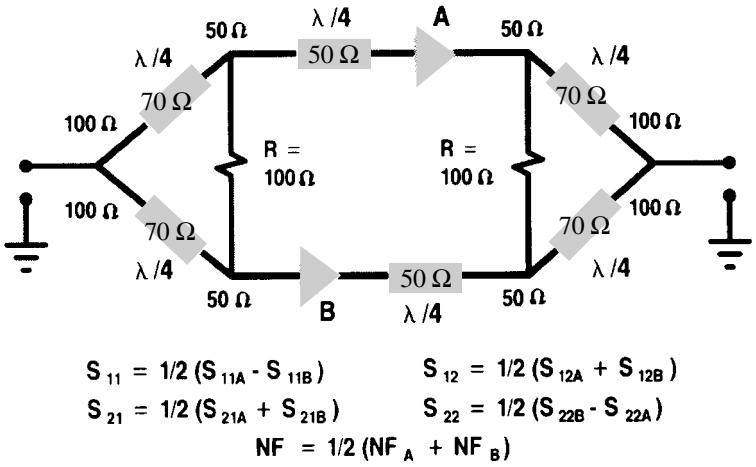


Figure 3-200 Wilkinson divider/combiners.

bandwidth limit of about 4:1; its advantage is that it is compact and ideal for cascading. If one stage opens, P_{OUT} and gain drop only 6 dB; however, it does not protect against load mismatch. A recommended distributed form is the Lange coupler (Figure 3-198). To analyze this approach, Figure 3-199 shows the resulting impedances and necessary phase shifts. If a Wilkinson coupler [30] is used instead of the 3-dB hybrid, the bandwidth is reduced. The Wilkinson coupler is larger in size than the 3-dB hybrid but more easily realized, leading to lower cost. Figure 3-200 shows the actual impedances necessary for matching, with A1 and A2 being the actual amplifiers combined.

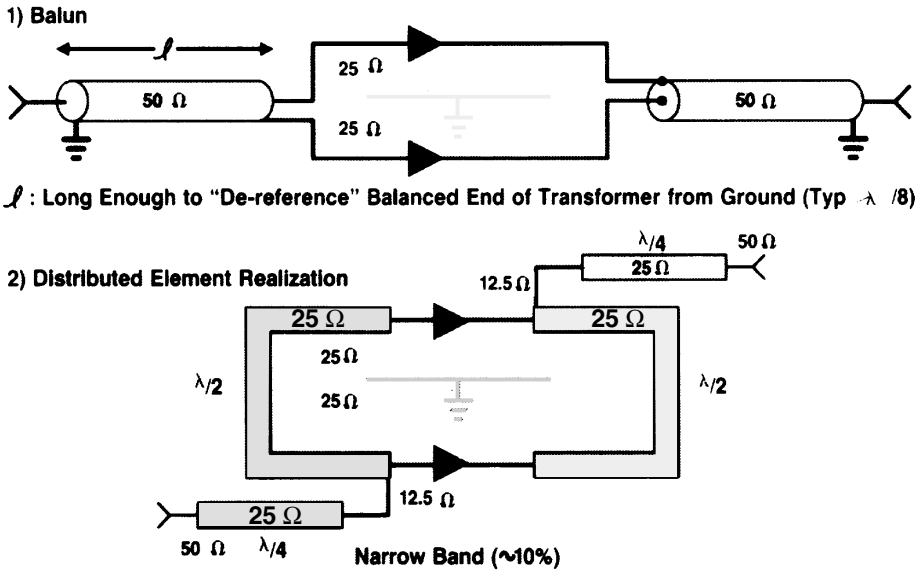


Figure 3-201 Push–pull amplifiers.

The solution for push–pull amplifiers resorts to a balun with the following results obtained: Bandwidths of greater than one decade are achievable; the balun structure gives 4:1 impedance advantage; there is no isolation; $S_{11} \neq 0$, $S_{22} \neq 0$; and cancellation of even-order products is provided.

Since most amplifiers have 50-Ω rather than 25-Ω impedances, Figure 3-201 shows a circuit that matches 25 Ω to 50 Ω on both sides. Since the standard amplifier operates at 50 Ω rather than 25 Ω, an additional network may be necessary to transform the impedance up to 50 Ω again. A similar problem occurs at the amplifier output.

3-12 POWER AMPLIFIERS

Power amplifiers are devices that are tasked to transform as much dc power into RF/microwave power as possible. Depending on the technology, devices are available that can produce at least 100 W up to 2 GHz. Since the transconductance of FETs at conducting angles of less than 180° becomes very small, the actual power gain sharply decreases compared to Class A operation while the power-added efficiency is optimized to values of 55% (Class B) or so. The dissipation of these devices is, of course, quite high, while in Class B operation this is even more of an issue. Power devices in Class C are dying to have big heat sinks to survive. One of the nice things about FETs is a much reduced thermal runaway effect compared to their bipolar brothers.

Power amplifier classes are defined in terms of the device conduction angle and/or the type of device operation:

Class A	"Linear" operation; 180° conduction angle
Class AB	"Linear" operation; conduction angle between 90° and 180°
Class B	"Linear" operation; 90° conduction angle
Class C	Fixed drive; less than 90° conduction angle
Class D	Switched operation, conduction range may vary with time from 0° to 180° or may be fixed.

There are also Classes E and F, which are currently under evaluation and showing some promising results.

Figure 3-202 shows Class C operation. The output of the amplifier is a current pulse train, which frequently leads to misunderstanding as to the calculation of its load. Given the fact that the amplifier is a current generator, it is silly to assume one has to do a conjugate matching looking backward into the amplifier. This is the condition for Class A operation. Unfortunately, this is still not understood by some engineers, but we hope we can clear this up today. The theoretical efficiency is 89.7% at a conducting angle of less than 180°; in practice, efficiencies more than 85% are not obtainable. (All of these efficiencies are those of the device itself; losses in matching networks are not taken into account in these efficiency values.) Considering optimal gain and optimal output power, the conducting angle actually settles somewhere between 100° and 140°. Again, the power gain in Class C operation is less than in Class A and Class B operation.

3-12-1 Example 1: 7-W Class C BJT Amplifier for 1.6 GHz

The following demonstrates how to design a 1.6-GHz, 7-W amplifier operating in Class C. We need to point out again that all modulation forms that contain the information in both

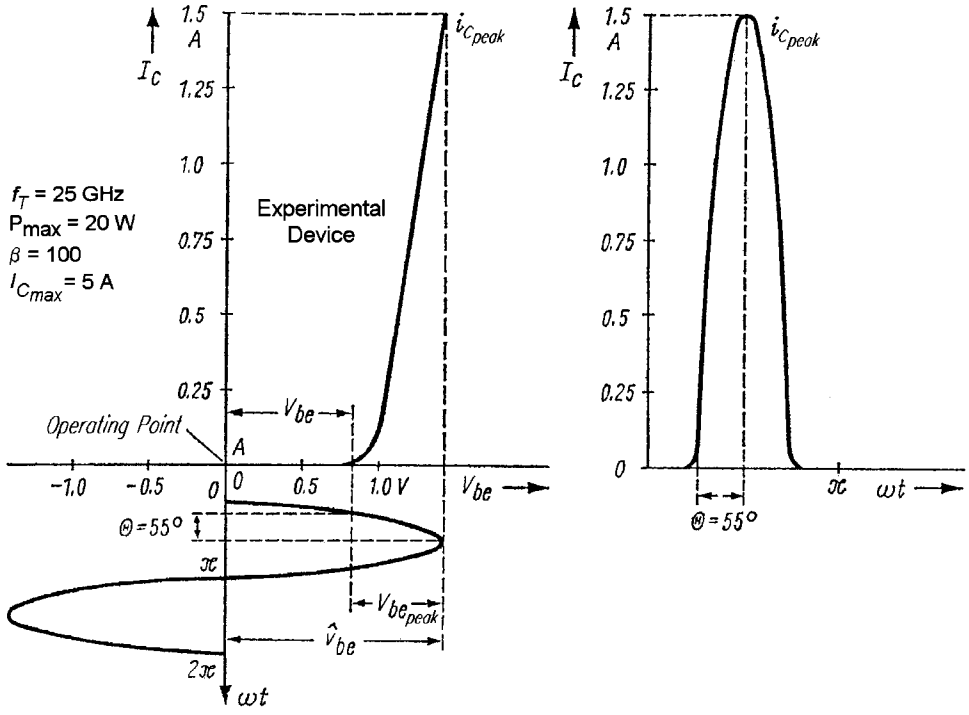


Figure 3-202 Output current as a function of drive voltage, taking into consideration the knee voltage of the base–emitter *pn* junction. The output is a series of current pulses with a duty cycle determined by the conduction angle.

amplitude and phase require more linearity than Class C affords. Needless to say, this affects the efficiency; therefore, at a later point we will also show a Class AB power amplifier. This is a tutorial, so we will present a UHF example and walk the reader through all the problems involved in the design. The selection of the transistor is done by first looking at the f_T, which in our case should be at least 4–5 GHz. The first order of business is to define the point of operation for the output amplifier. To do this, we need to determine the maximum voltage and current values supplied by the manufacturer. We have chosen a device that operates at 28 V and can sustain a peak collector current of 1.5 A. Figure 3-202 lets us determine the conducting angle Θ from the I_C/V_{BE} plane. This figure shows on the left side the I_C/V_{BE} plane and the necessary voltage to drive the amplifier.

Initially, we assume that the base–emitter junction is operated without any bias ($V_{BE} = 0$). The dc I–V curve shows that the collector current starts at about 650 mV. The amplifier itself, therefore, is already in Class C operation. For full output power, the drive voltage $V_{BE,\text{peak}}$ has to be 1.14 V. Therefore, the conducting angle can be determined by

$$\cos \Theta = \frac{V_{BE}}{\hat{\boldsymbol{v}}_{BE}} = \frac{0.65}{1.14} = 0.57$$

$$\cos 0.57 = 55^\circ$$

V_{BE} is the dc voltage at which the transistor starts conducting. From Figure 3-65 and Eqs. (3-160), (3-161), and (3-162), the fundamental current i_{c1} can be determined from

$$i_{c1} = i_{\text{peak}} \cdot f_1(\Theta)$$

or

$$i_{c1} = 1.5 \text{ A} \times 0.36 = 0.54 \text{ A}$$

The average collector dc current can be determined from

$$i_{c2} = i_{\text{peak}} \cdot f_2(\Theta)$$

or

$$i_{c2} = 1.5 \text{ A} \times 0.21 = 0.315 \text{ A}$$

The collector saturation voltage $V_{CE(\text{sat})}$, unfortunately, is around 2–3 V; for the purpose of the calculation, we assume that the saturation voltage is 3 V, and, therefore, we can assume that the maximum collector ac voltage $\hat{v}_{CE} \approx (V_B - 3) = 25$.

The appropriate collector load has to be

$$R_C = \frac{\hat{v}_{CE}}{i_{C1}} = \frac{25}{0.54} = 46.3 \ \Omega$$

The maximum output power is

$$P_{\text{out}} = \frac{(V_{CE} - V_{CE(\text{sat})})^2}{2R_C} = \frac{625}{92.6} = 6.75 \text{ W}$$

The only way to increase the power would be to have a smaller collector load at a higher current. In this family of Motorola transistors, there was no 10-W transistor; that's why we have chosen a 20-W experimental device, which for this purpose is overkill. The theoretical efficiency of the amplifier is defined as

$$\eta = \frac{P_{\text{out}}}{P_{\text{in}}} = \frac{6.75 \text{ W}}{0.315 \times 28} = 85\%$$

If the saturation voltage would be less, our efficiency would automatically increase, since for the given load, the actual output power would increase (6.75 W $\rightarrow$ 7.5 W). Assuming that this stage would operate at 1900 MHz with a gain of roughly 22 dB, the required drive would be significantly less than the output power, or roughly 40 mW. For the purpose of the following calculations, we will assume that the drive power is 0.7 W. The total dissipation of the device now adds up to be

$$P_T = P_{\text{in(dc)}} - P_{\text{out}} + P_{\text{drive}}$$

$$P_T = (28 \times 0.315) - 6.75 + 0.7 = 2.77$$

Power-added efficiency is derived from this. Knowing the power dissipation, one can calculate a thermal resistance R_{therm} of the heat sink and the transistor:

$$R_{\text{therm}} = \frac{T_J - T_A}{P_T}$$

Where T_J is the device junction temperature and T_A is the ambient temperature. For a maximum ambient temperature of 60 °C and a maximum junction temperature of 150 °C, the thermal resistance is

$$R_{\text{therm}} = \frac{120 - 60}{2.77} = 21.7\ ^\circ\text{C/W}$$

R_{therm} is the sum of thermal resistance of the transistor and its package, and the thermal resistance between the transistor package and the heat sink, and the thermal resistance between the heat sink and the ambient air. Therefore,

$$R_{\text{therm}} = R_{\text{th(case)}} + R_{\text{th(case/heat sink)}} + R_{\text{th(heat sink)}}$$

The required thermal resistance of the heat sink is calculated from

$$R_{\text{th(heat sink)}} = 21.7 - 10 - 1.4 = 15.6\ ^\circ\text{C/W}$$

Assuming an aluminum heat sink (a 2-mm-thick plate), the required heat-sink surface area A (cm^2) is

$$A = \frac{1}{\alpha R_{\text{th(heat sink)}}}; \quad \alpha \approx 1.5\ \frac{\text{mW}}{^\circ\text{Ccm}^2}$$

Therefore, the required area is

$$A = \frac{1}{1.5 \times 0.0156} = 42.7\ \text{cm}^2$$

The next task is to match both the input and output to the (most likely) 50-Ω source and termination. The following section will guide us through the procedure. The input impedance Z_{11} (after some lengthy search) was found to be $1 + j10\ \Omega$. This means that the input consists of the base-spreading resistance (approximately 1 Ω) and an input capacitance, which because of phase shift now turns out to be inductive. Since the popular method for input matching these days is a pi configuration, we are going to use a simple pi filter both for input and output matching. There are several equations in the literature that give the values for the two capacitors and inductors; most of them result in an incorrect resonance frequency. Since we have access to software tools, we decided to take three approximate values and use the optimizer to give us the appropriate matching. The only remaining variable in such an approach is the bandwidth. If we specify $S_{21} = 1$ and $S_{11} = 0$ (100% energy transfer and perfect return loss), we will realize that this is possible at only one frequency. We therefore need to identify a frequency band or bandwidth over which an acceptable match is needed. As the bandwidth increases, the match for such a simple circuit will get poorer; there are not enough elements available to do better. In this case, we took the center frequency ±100 MHz.

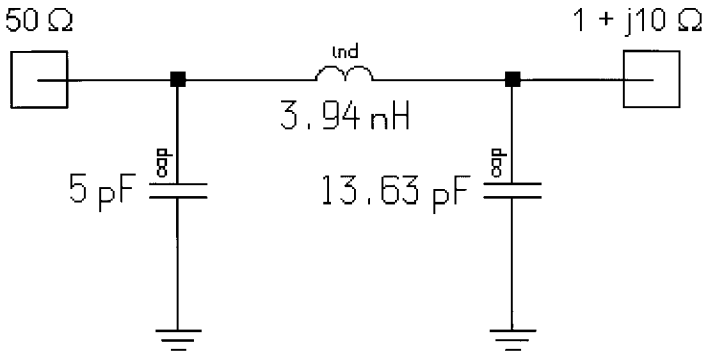


Figure 3-203 Input matching network for the amplifier used in this tutorial.

Figures 3-203 and 3-204 show the actual circuit with the element values and the resulting input match. Likewise, we did the same thing at the output, but the actual resulting dynamic output capacitance is not known *a priori*. For low frequencies, the integral equation gives a value of $2C_{CE}$; however, practical tests in frequencies above 500 MHz show that the actual increase is only 1.3 times the dc output capacitance. Therefore, we decided to design the filter with no output capacitance assumed, and the input capacitance of about 5 pF must be reduced for proper reactive matching. Mathematically, this could result in a negative number, specifically if the resulting output capacitance exceeds the 5-pF value. To compensate for this, one needs to use an RF choke that tunes out the resulting dynamic output capacitance and then one can revert to the original 5 pF as calculated.

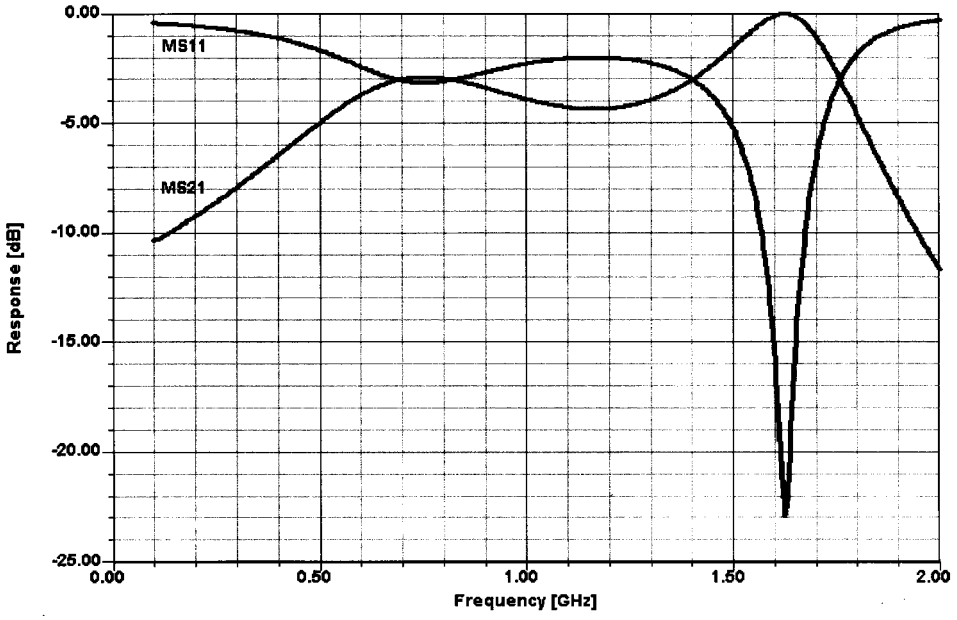


Figure 3-204 Frequency response of the input matching network.

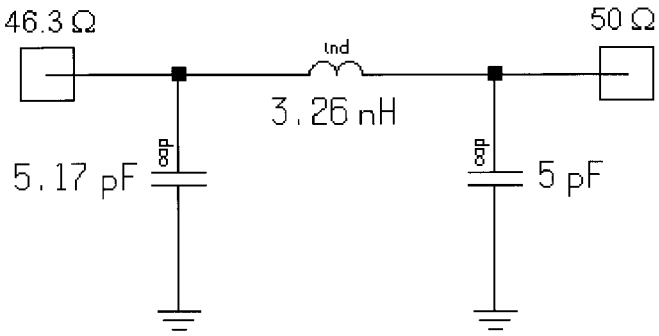


Figure 3-205 Output matching network for the amplifier.

The selection of the actual component is another issue, because these capacitors need to handle several amperes of RF current. It is not always easy to find capacitors with the appropriate value that can handle the current. The inductors can be printed. The paragraph following this section therefore will deal with the issue of how to translate the lumped values into distributed values, which will make the current handling much easier (we hope).

For this, we did the same optimizing approach by using a 50-Ω source and a 46.3-Ω load. Figures 3-205 and 3-206 show the circuit with the appropriate values and its simulated frequency-dependent responses.

The next step is to connect the input and output matching network to an "actual" transistor. This is a point where difficulties arise, since very few, if any, companies provide SPICE-type

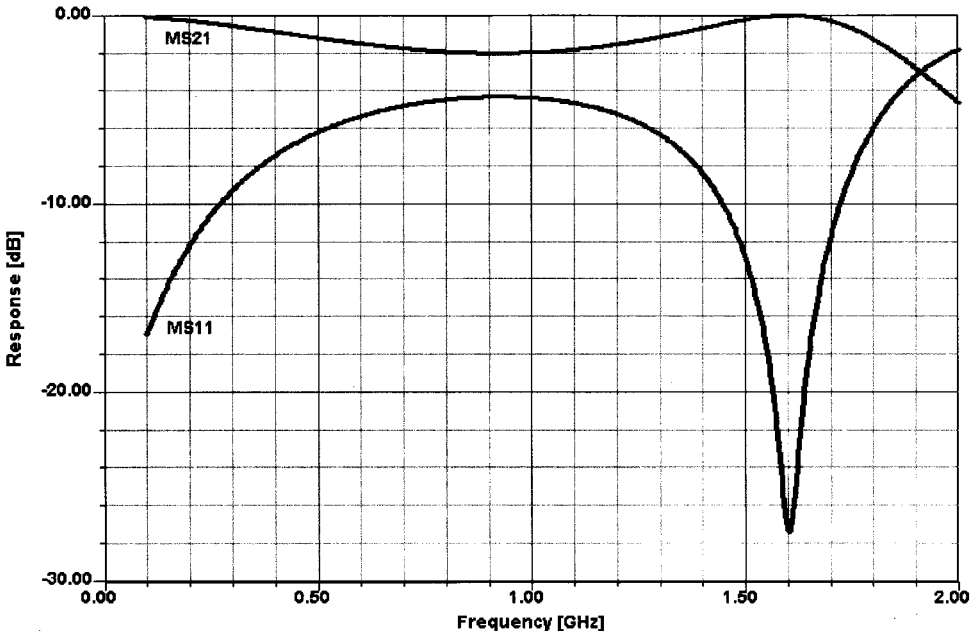


Figure 3-206 Frequency response of the output matching network.

parameters for transistors operating at currents of more than 200 mA. Figures 3-207 and 3-208 show the overall schematic and frequency response of the actual amplifier with all filters connected. Figure 3-209 details the amplifier's peak gain response. If more bandwidth is needed, the matching circuit must be increased in complexity, as will be shown in the following section.

Now is a good time to look at the final results of our simulation, which is based on the various assumptions as outlined above. Figure 3-210 shows the dc *I–V* curves with a straight load line and no reactances. This was accomplished by using a wideband transformer instead of an output matching circuit. Figure 3-211, however, shows the "load line" that results when the transformer is replaced with a real matching network. Based on the harmonics content and its improper termination, one can see how the output load line opens and shows ringing in the low-current area. If we expect to see the saturation voltage, we will be disappointed because the saturation voltage is frequency dependent (increases with frequency) and this figure gives an erroneous impression. Hence, we show dc *I–V* curves. To display the conduction angle, which we showed previously, we examine the output current as a function of time at a drive level of +16.5 dBm (Figure 3-212). During the negative half, the transistor is not conducting as predicted, so this validates our assumptions. Next, we are interested in determining the power gain. Therefore, we look at the output spectrum for a drive level of +16.5 dBm (Figure 3-213). It shows an output level of 38.6 dBm. If we subtract the drive level (16.5 dBm) from this, we see that the power gain is 22.1 dB at this drive level. Since we are used to thinking in watts, Figure 3-214 shows the absolute output power and the harmonics prior to filtering. Figure 3-215 shows the output spectrum after replacing the transformer with a simple pi output network.

During the conduction period, there will be a voltage across the transistor 180° out of phase, as shown in Figure 3-216. Finally, to highlight the nonlinearity of a Class C amplifier, Figure 3-217 shows the amplifier's output power versus drive, and Figure 3-218 shows the amplifier's gain versus drive. When the transistor turns on is clearly evident.

A particular difficulty arises in the case of FETs, where the charge energy makes for a difference between dc and pulsed measurements. Figure 3-219 shows this. Some of the more modern transistors, such as SiGe types, may need some pulsing also; as can be seen in Figure 3-219, even depending on the pulsing, one may get different dc *I–V* curves. This area is still under investigation, and even the "experts" do not agree on all aspects of this. In real life, it turns out that a lot of experimentation is the answer. Anyone who thinks that a pure CAD approach will give the right design answers immediately will encounter many surprises! This is not a fault of the CAD tools—assuming that they are properly engineered—but rather the availability of appropriately "tweaked" models for high-power applications, and the willingness of device suppliers to standardize on certain models and support them.

Because of the difficulties involved in cooling transistors under test and simulating RF pulse and CW conditions, depending on the application, manufacturers have so far stayed away from investing money in this area. The designer must therefore resort to application notes published by device manufacturers. Because even slight layer changes can have a large impact on the frequency response of a given device, application-note-based designs can run into difficulty if a device's fabrication has significantly evolved since the publication of its application notes.

Given the multitude of models outlined in Chapter 2, the CAD user is sometimes left hanging as to which is the best model to use. The latest indications are that the MEXTRAM (*m*ost *e*xquisite *t*ransistor *m*odel), developed by the University of Delft, Holland, with Philips of Holland, shows the most promising result, specifically in the area of IMD products. While

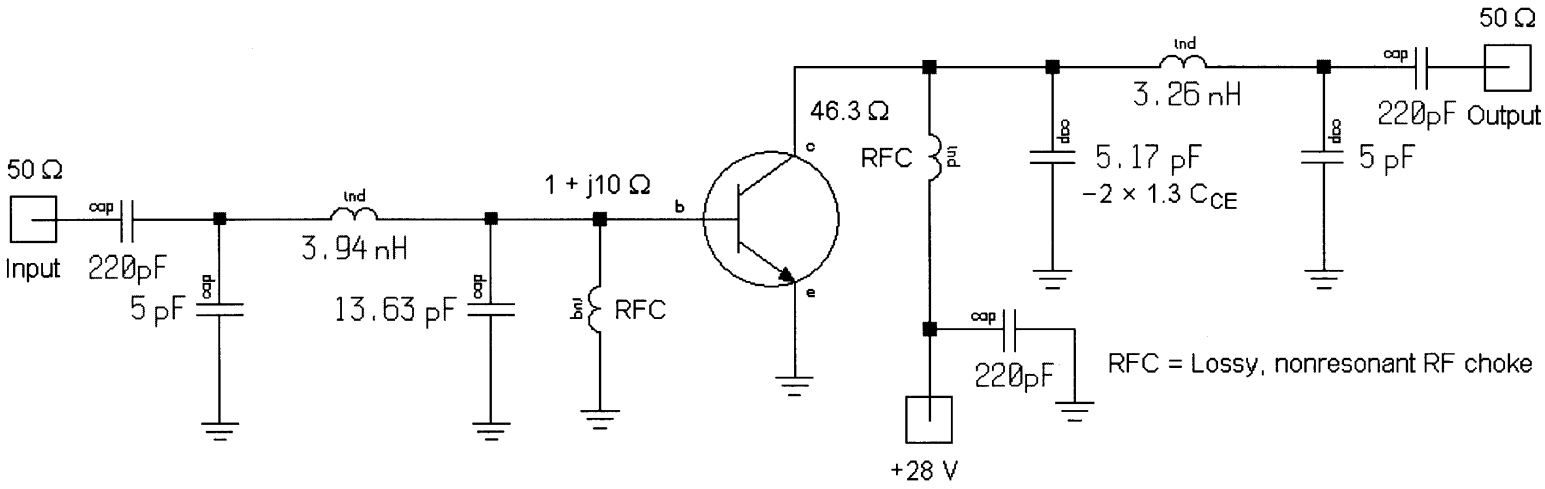


Figure 3-207 Actual schematic of our experimental 1.6-GHz amplifier. The RF chokes have enough RF losses not to be resonant and yet don't load the circuit. This is frequently accomplished with either ferrite beads or resistors in parallel with the chokes.

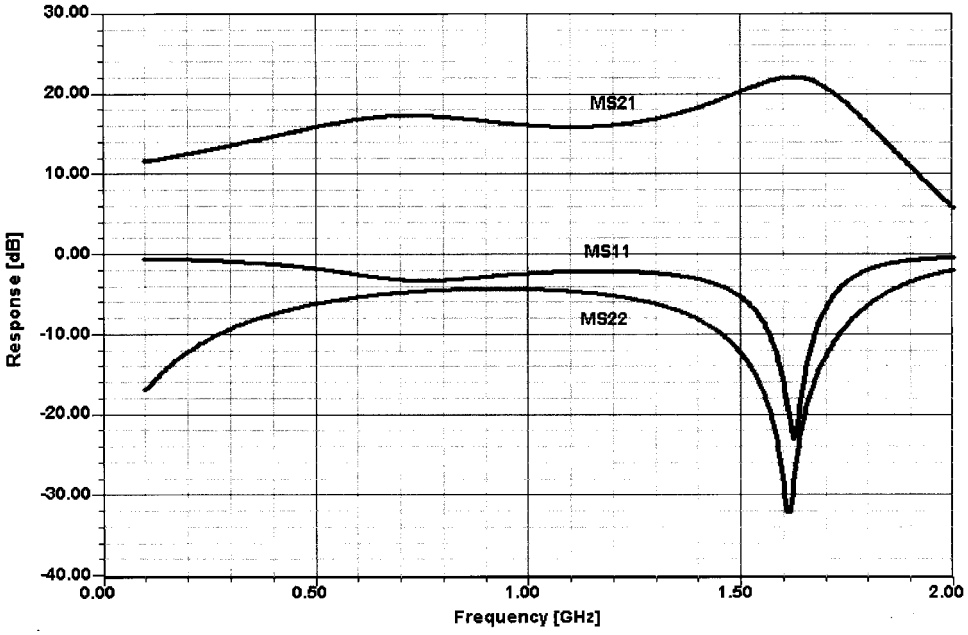


Figure 3-208 Overall frequency response of our experimental amplifier. Input and output matching, and gain, are shown as a function of frequency.

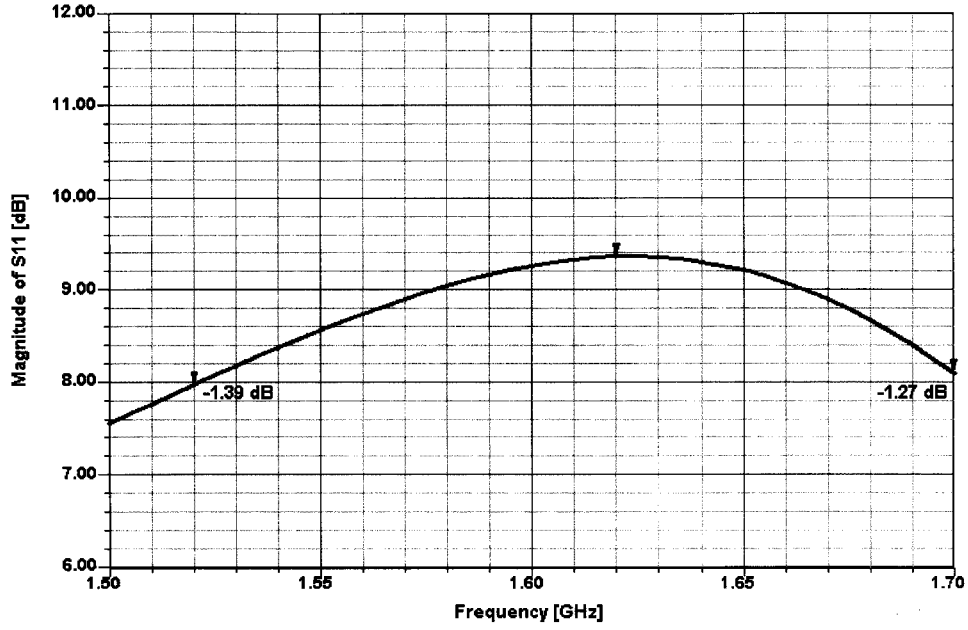


Figure 3-209 Close-in display of the frequency-dependent gain.

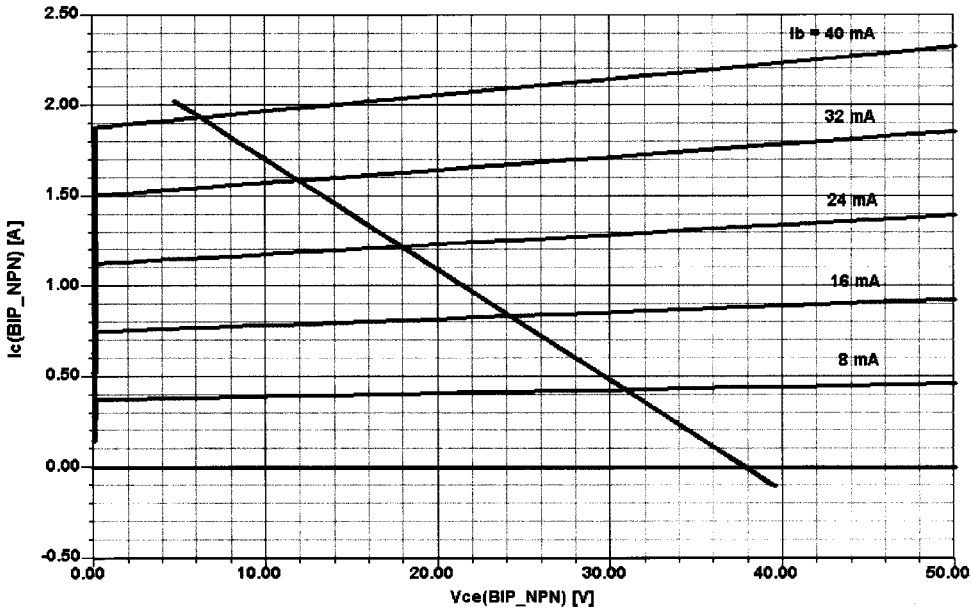


Figure 3-210 The dc *I–V* curves assuming a perfect resistive load. No RF saturation voltage or other RF saturation effects are evident.

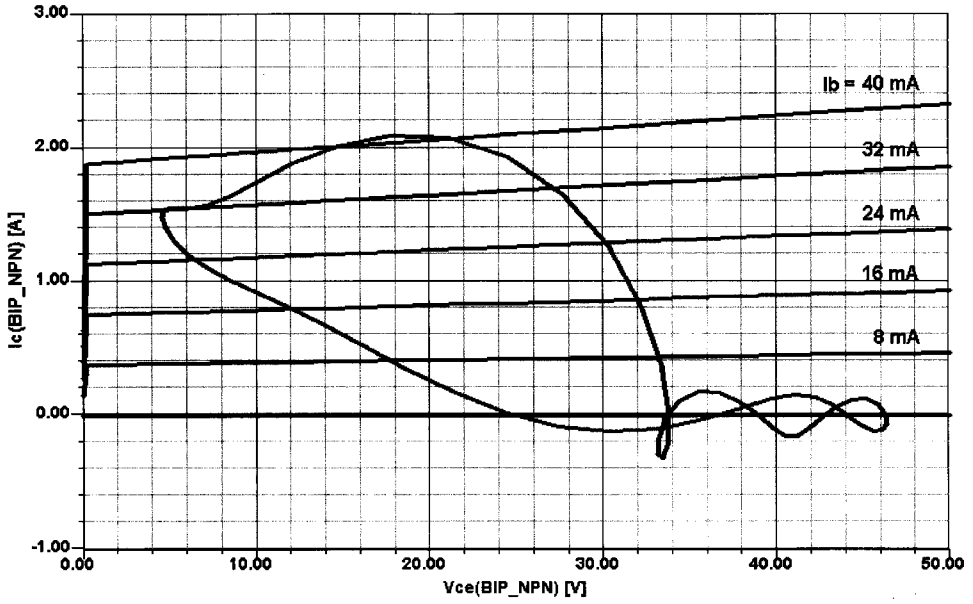


Figure 3-211 Load line with the output transformer replaced with a real matching circuit.

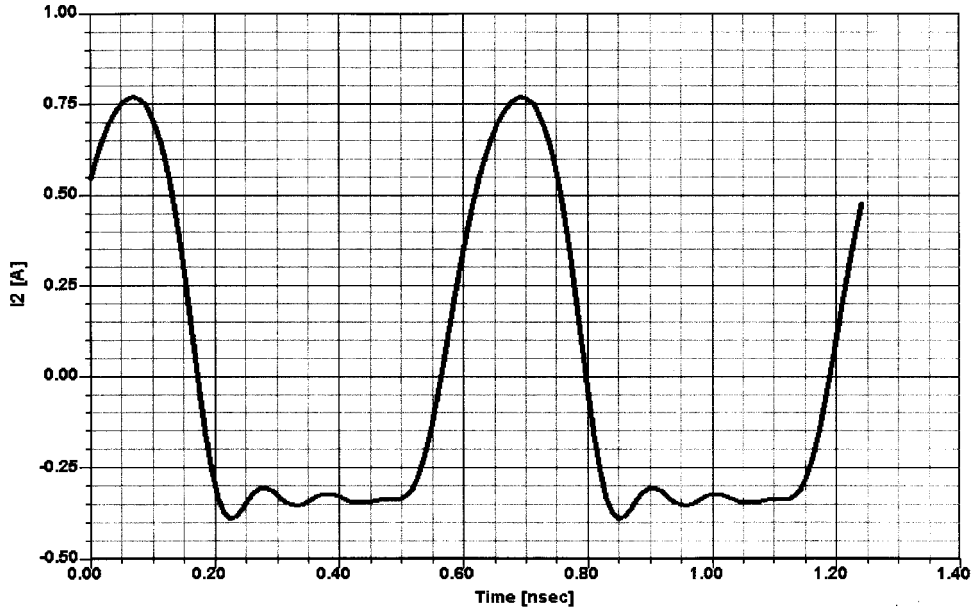


Figure 3-212 Output current as a function of time. The duty cycle is determined by the device conduction angle.

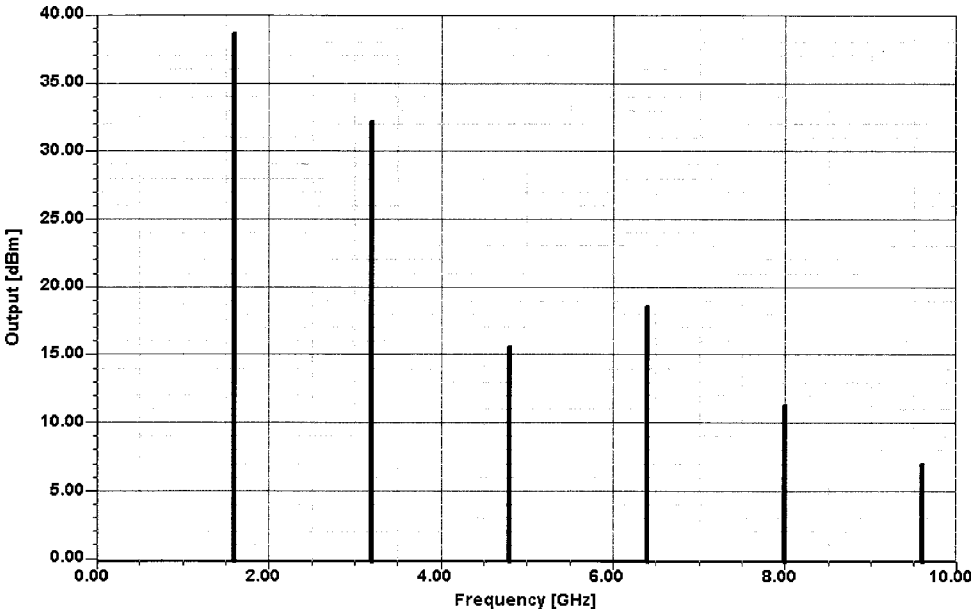


Figure 3-213 Power output at the collector prior to any filtering, assuming a real output termination.

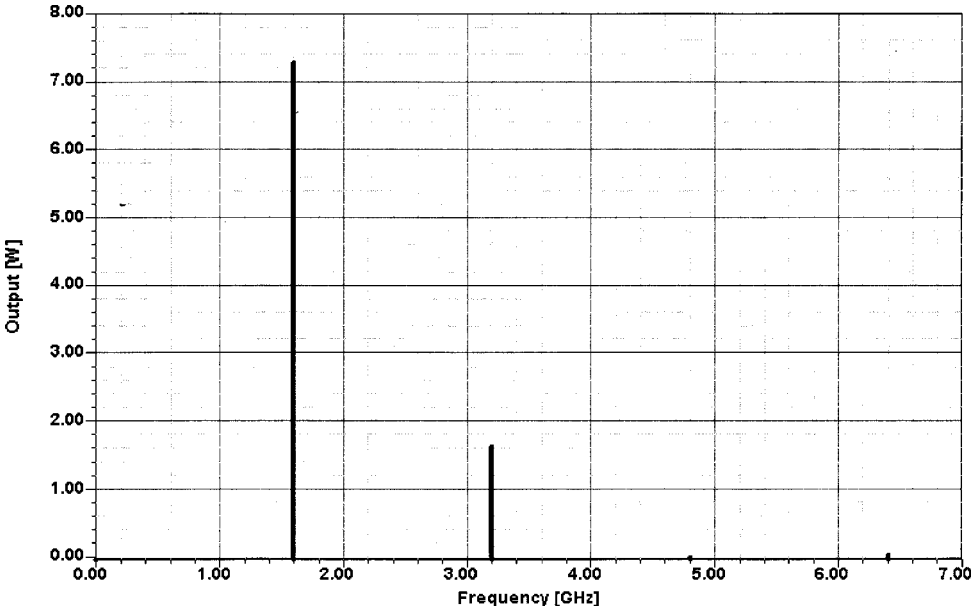


Figure 3-214 Available output power in watts, prior to any filtering, assuming a real output termination.

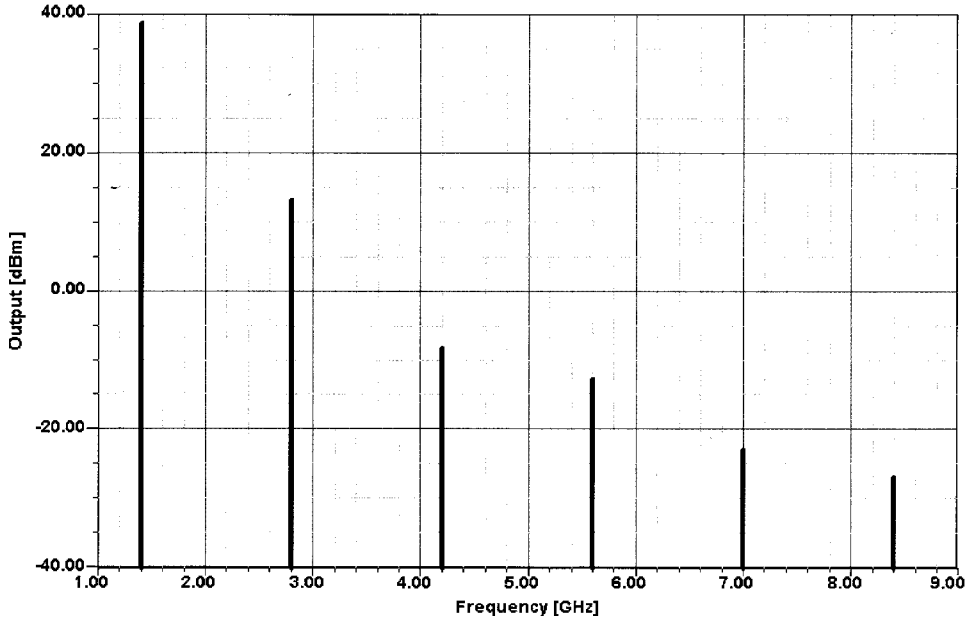


Figure 3-215 The output spectrum after replacing the transformer with a simple pi output network.

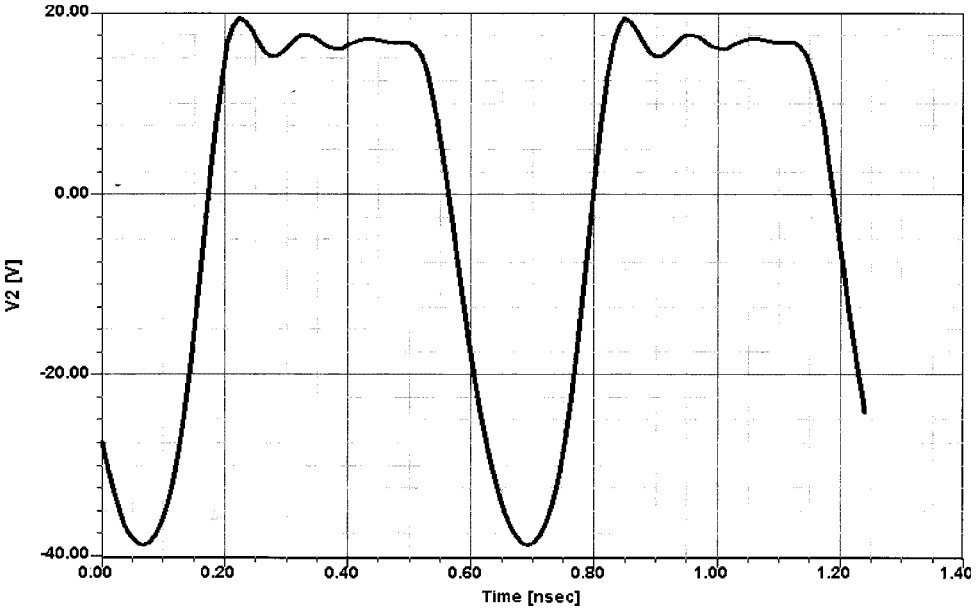


Figure 3-216 Output voltage at the collector. The voltage is asymmetrical compared to 0; this is important in the selection of a device capable of handling these voltages.

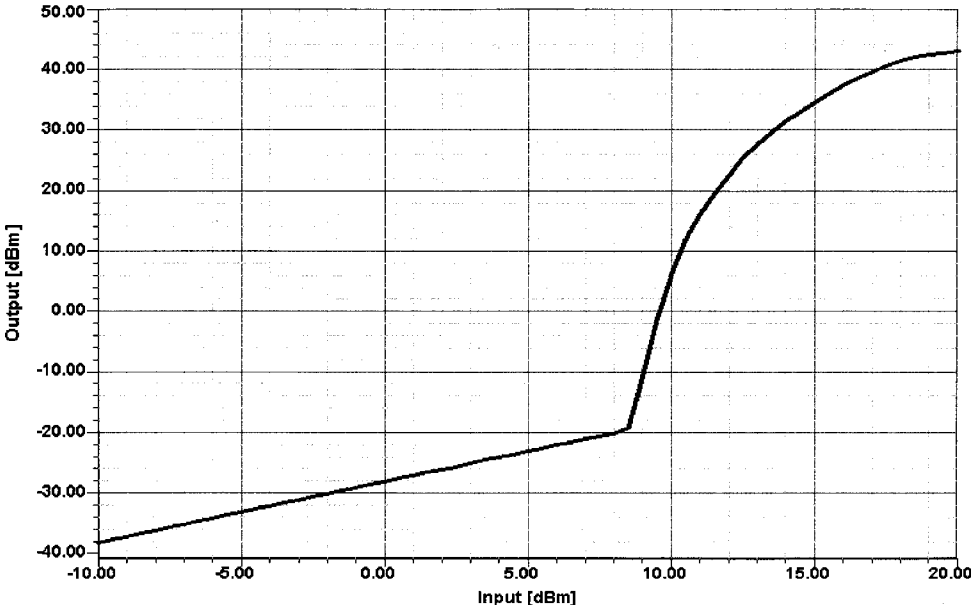


Figure 3-217 Output power versus input power for the Class C amplifier. Until the transistor turns on (at a drive level of about 8.5 dBm), only fedthrough power appears at the output.

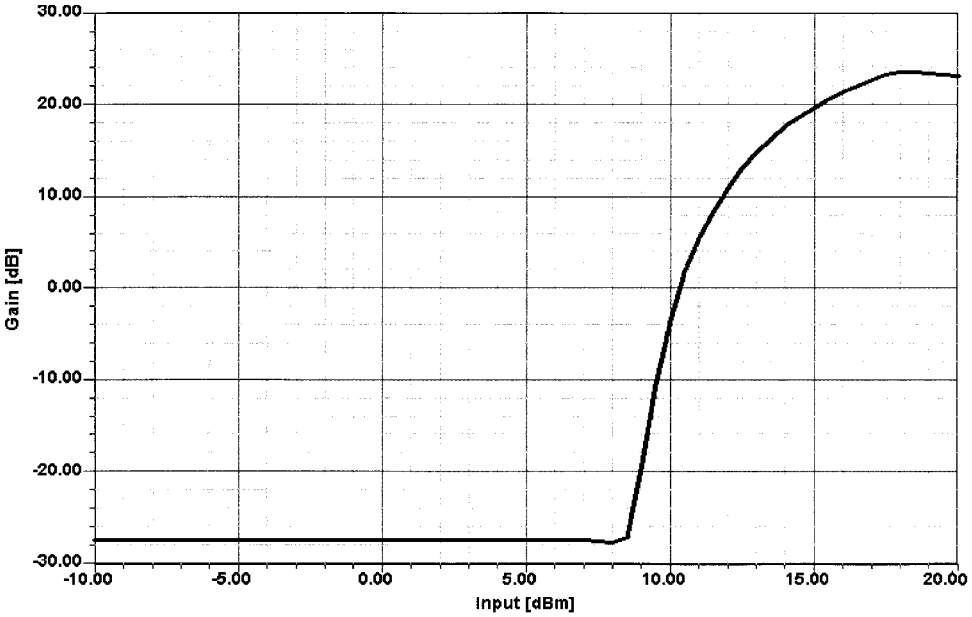


Figure 3-218 Amplifier gain as a function of drive. It is evident that one needs to cross the 0.65-V point for the transistor to become conductive and show gain. The initial amplification as a function of drive shows a very steep curve that then levels off and drops as the amplifier moves into saturation.

other later models have up to seven internal nodes, MEXTRAM fortunately uses only five. (The MEXTRAM model actually consists of three transistors, which makes modeling fun.) Power-added efficiency simulation is not available, since the MEXTRAM model was developed for small-signal applications, meaning circuits that operate at less than 100 mW of RF output power.

As if the absence of trustworthy high-power device models is not challenge enough, in developing this 1.6-GHz Class C amplifier example, we ran into another complication: Using a CAD optimizer to simultaneously match the circuit's input and output networks failed because of the device's input-impedance variation with drive. Every matching adjustment that improved the drive to the device resulted in an input-impedance change that reduced the drive to the device! In the end, the optimizer couldn't keep up. In actually building such a circuit, one would make the matching-network values adjustable on the initial breadboard and then, with the necessary input drive level applied, iteratively adjust the input and output matching for the desired output level. It appears that the CAD program could not handle such a degree of nonlinearity.

We then went to the approach of using large-signal *S* parameters at a drive level slightly less than 16 dBm, but this detuned the input filter, resulting in a poorer input match than the one initially predicted. Likewise, the output matching, which in a power stage must result in maximum output power rather than a conjugate match, was less impressive. The large-signal *S* parameters at the output for a power amplifier are somewhat dubious because Class C operation generates so many harmonics that the output termination will reflect a lot of energy and, depending on the output filter topology, different results can be expected. We firmly believe that there will be further mathematical improvement in the CAD tools, and we very much invite

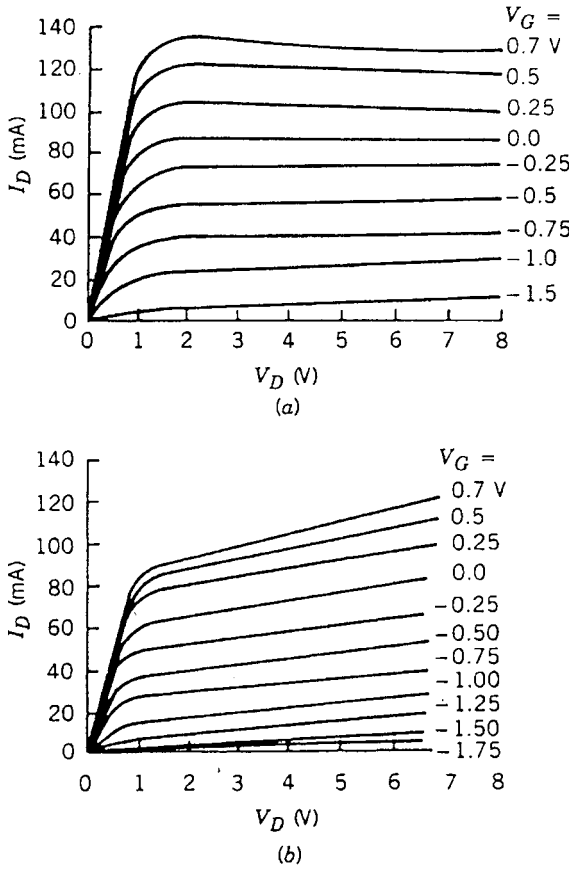


Figure 3-219 Comparison of FET drain current characteristics obtained from (a) pulsed *I–V* and (b) rectified sine-wave evaluation. The rectified sine-wave curves can be considered equivalent to curves obtained at dc.

our readers to use such a nonlinear example to test their tools, provided that a high-power model becomes available in this power range.

It has been rumored that the load-pull technique, rather than the application of large-signal *S* parameters, is a vehicle to solve such matching problems. However, the load-pull technique is really applied only to the fundamental frequency and does not deal with the issue of harmonics of the output current. More details about the load-pull technique and its capabilities are found in Appendix B.

Actual measurements of amplifiers of this type show that our simulated efficiency and drive level come quite close to reality, but at a frequency somewhat offset from that chosen for the design. As an example, the input and output filters tuned at 1.4 GHz instead of 1.6 GHz using the large-signal *S* parameters. Again, an iterative process, similar to manually trimming a prototype, will give a similar result. This is consistent with our earlier statement that most final designs are really hand-tweaked, and that CAD can only bring us close, but not all the way, to a good solution. For passive problems, the CAD does a significantly better job; the nonlinear cases are the ones that still cause headaches.

3-12-2 Impedance Matching Networks Applied to RF Power Transistors*

Introduction. Some graphical and numerical methods of impedance matching will be reviewed here. The examples given refer to high-frequency power amplifiers.

Although matching networks normally take the form of filters and therefore are also useful to provide frequency discrimination, this aspect will be considered only as a corollary of the matching circuit.

Matching is necessary for the best possible energy transfer from stage to stage. In RF power transistors, the input impedance is of low value, decreasing as the power increases or as the chip size becomes larger. This impedance must be matched either to a generator—of generally 50-Ω internal impedance—or to a preceding stage. Impedance matching has to be made between two complex impedances, which makes the design still more difficult, especially if matching must be accomplished over a wide frequency band.

Device Parameters

Input Impedance. The general shape of the input impedance of RF power transistors is as shown in Figure 3-220. It is a large-signal parameter, expressed here by the parallel combination of a resistance R_p and a reactance X_p [31].

The equivalent circuit shown in Figure 3-221 accounts for the behavior illustrated in Figure 3-220. With the presently used stripline or flange packaging, most of the power devices for VHF low band will have their R_p and X_p values below the series resonant point f_s. The input impedance will be essentially capacitive.

Most of the VHF high-band transistors will have the series-resonant frequency within their operating range; that is, they will be purely resistive at one single frequency f_s, which the parallel-resonant frequency f_p will be outside.

Parameters for 1- or 2-GHz transistors will be beyond f_s and approach f_p. They show a high value of R_p and X_p with inductive character.

A parameter that is very often used to judge on the broadband capabilities of a device is the input Q or Q_{IN}, defined simply as the ratio R_p/X_p. Practically, Q_{IN} ranges around 1 or less for VHF devices and around 5 or more for microwave transistors.

Q_{IN} is an important parameter to consider for broadband matching. Matching networks normally are low-pass or pseudo-low-pass filters. If Q_{IN} is high, it can be necessary to use bandpass filter-type matching networks and to allow insertion losses. But broadband matching is still possible. This will be discussed later.

Output Impedance. The output impedance of RF power transistors, as given by all manufacturers' datasheets, generally consists of only a capacitance C_{OUT}. The internal resistance of the transistor is supposed to be much higher than the load and is normally neglected. In the case of a relatively low internal resistance, the efficiency of the device would decrease by the factor

$$1 + R_L/R_T \tag{3-192}$$

*Based on Motorola application note AN-721 ("Impedance Matching Networks Applied to RF Power Transistors"). Copyright of Motorola; used by permission.

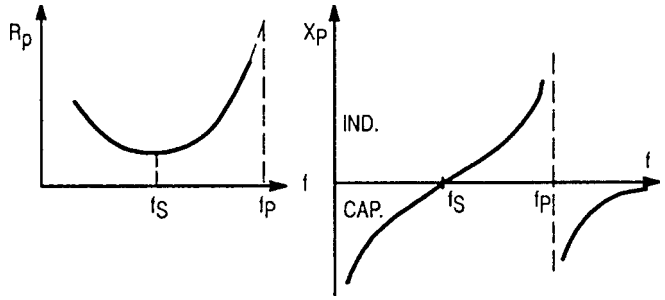


Figure 3-220 Input impedance of RF power transistors as a function of frequency.

where R_L is the load resistance, seen at the collector–emitter terminals, and R_T is the internal transistor resistance equal to

$$\frac{1}{\omega_T (C_{TC} + C_{DC})} \tag{3-193}$$

defined as a single parameter, where ω_T is the transit angular frequency and C_{TC} and C_{DC} are the transition and diffusion capacitances, respectively, at the collector junction.

The output capacitance, C_{OUT}, which is a large-signal parameter, is related to the small-signal parameter C_{CB}, the collector–base transition capacitance.

Since a junction capacitance varies with the applied voltage, C_{OUT} differs from C_{CB} in that it has to be averaged over the total voltage swing. For an abrupt junction and assuming certain simplifications, $C_{\text{OUT}} = 2C_{CB}$.

Figure 3-222 shows the variation of C_{OUT} with frequency. C_{OUT} decreases partly due to the presence of the collector lead inductance, but mainly because of the fact that the base–emitter diode does not shut off anymore when the operating frequency approaches the transit frequency, f_T.

Output Load. In the absence of a more precise indication, the output load R_L is taken as

$$R_L = \frac{[V_{CC} - V_{CE(\text{sat})}]^2}{2P_{\text{OUT}}} \tag{3-194}$$

with $V_{CE(\text{sat})}$ equal to 2 or 3 V, increasing with frequency.

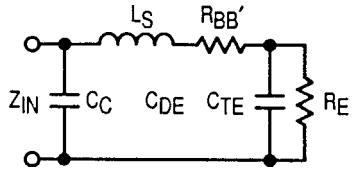


Figure 3-221 Equivalent circuit for the input impedance. R_E is the emitter diffusion resistance; C_{DE} and C_{TE} are the diffusion and transition capacitances, respectively, of the emitter junction; $R_{BB'}$ is the base-spreading resistance; C_C is the package capacitance; and L_S is the base lead inductance.

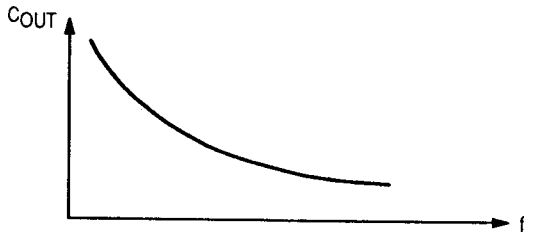


Figure 3-222 Output capacitance C_{OUT} as a function of frequency.

The above equation just expresses a well-known relation, but it also shows that the load, in first approximation, is not related to the device, except for $V_{CE(sat)}$. The load value is primarily dictated by the required output power and the peak voltage; it is not matched to the output impedance of the device.

At high frequencies this approximation becomes less exact, and for microwave devices the load that must be presented to the device is indicated on the datasheet.

Strictly speaking, impedance matching is accomplished only at the input. Interstage and load matching are more impedance transformations of the device input impedance and of the load into a value R_L (sometimes with additional reactive component) that depends essentially on the power demanded and the supply voltage.

Matching Networks. In the following, matching networks will be described by the order of complexity. These are ladder-type reactance networks. The different reactance values will be calculated and determined graphically. Increasing the number of reactances broadens the bandwidth. However, matching networks consisting of more than four reactances are rare. Above four reactances, the improvement is small.

Numerical Design

TWO-RESISTANCE NETWORKS. Resistance terminations will first be considered. Figure 3-223 shows the reactive L section and the terminations to be matched.

Matching or exact transformation from R_2 into R_1 occurs at a single frequency f_0. At f_0, X_1 and X_2 are equal to

$$X_1 = \pm R_1 \sqrt{\frac{R_2}{R_1 - R_2}} = R_1 \frac{1}{\sqrt{n-1}} \tag{3-195}$$

$$X_2 = \mp \sqrt{R_2(R_1 - R_{2)}} = R_1 \frac{\sqrt{n-1}}{1} \tag{3-196}$$

At f_0, $X_1X_2 = R_1R_2$. X_1 and X_2 must be of opposite sign. The shunt reactance is in parallel with the larger resistance.

The frequency response of the L section is shown in Figure 3-224, where the normalized current is plotted as a function of the normalized frequency.

If X_1 is capacitive and consequently X_2 is inductive, then

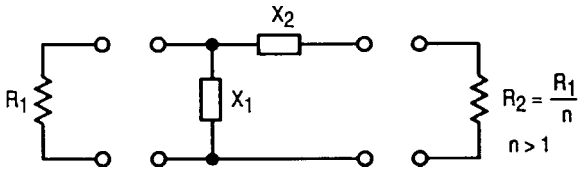


Figure 3-223 Two-reactance matching network.

$$X_1 = -\frac{f_0}{f} R_1 \sqrt{\frac{R_2}{R_1 - R_2}} = -\frac{f_0}{f} R_1 \frac{1}{\sqrt{n-1}} \tag{3-197}$$

and

$$X_2 = \frac{f}{f_0} \sqrt{R_2(R_1 - R_2)} = \frac{f}{f_0} R_1 \frac{\sqrt{n-1}}{1} \tag{3-198}$$

The normalized current absolute value is equal to

$$\left|\frac{I_2}{I_0}\right| = \frac{2\sqrt{n}}{\sqrt{(n-1)^2 \cdot (f/f_0)^4 - 2\,(f/f_0)^2 + (n+1)^2}} \tag{3-199}$$

where $I_0 = \sqrt{n}E/2R_1$ and is plotted in Figure 3-224 [32].

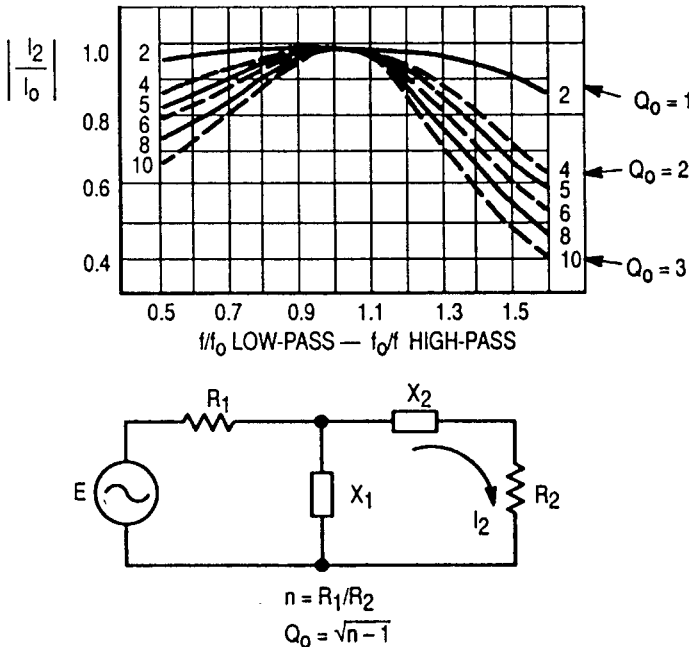


Figure 3-224 Normalized frequency response for the L section in low-pass or high-pass form.

If X_1 is inductive and consequently X_2 is capacitive, the only change required is a replacement of f by f_0 and vice versa. The L section has low-pass form in the first case and high-pass form in the second case.

The Q of the circuit at f_0 is equal to

$$Q_0 = \frac{X_2}{R_2} = \frac{R_1}{X_1} = \sqrt{n-1} \tag{3-200}$$

For a given transformation ratio n, there is only one possible value of Q. On the other hand, there are two symmetrical solutions for the network that can be either a low-pass filter or a high-pass filter.

The frequency f_0 does not need to be the center frequency, $(f_1 + f_2) / 2$, of the desired band limited by f_1 and f_2. In fact, as can be seen from the low-pass configuration of Figure 3-224, it may be interesting to shift f_0 toward the high-band edge frequency f_2 to obtain a larger bandwidth w, where

$$w = \frac{2(f_1 + f_2)}{f_2 - f_1} \tag{3-201}$$

This will, however, be at the expense of poor harmonic rejection.

Example. For a transformation ratio of $n = 4$, we can determine the following values from the above relations:

Bandwidth w	0.1	0.3
Maximum insertion losses	0.025	0.2
X_1/R_1	1.730	1.712

If the terminations R_1 and R_2 have a reactive component X, the latter may be taken as part of the external resistance as shown in Figure 3-225. This compensation is applicable as long as

$$Q_{int} = \frac{X_{\text{int}}}{R_2} \quad \text{or} \quad \frac{R_1}{X_{\text{int}}} < n - 1 \tag{3-202}$$

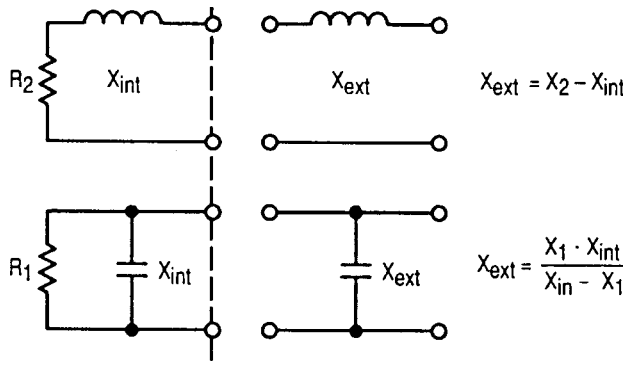


Figure 3-225 Termination reactance compensation.

Tables giving reactance values can be found in Matthaei et al. [33, 34].

USE OF TRANSMISSION LINES AND INDUCTORS. In the preceding section, the inductance was expected to be realized by a lumped element. A transmission line can be used instead (Figure 3-226).

As can be seen from the computed selectivity curves (Figure 3-227) for the two configurations, transmission lines result in a larger bandwidth. The gain is important for a transmission line having a length $L = \lambda/4$ ($\Theta = 90°$) and a characteristic impedance $Z_0 = \sqrt{R_1 R_2}$. It is not significant for lines short with respect to $\lambda/4$. One will notice that there is an infinity of solutions, one for each value of C, when using transmission lines.

THREE-REACTANCE MATCHING NETWORKS. The networks that will be investigated are shown in Figure 3-228. They are made of three reactances alternately connected in series and shunt.

A three-reactance configuration allows the designer to make the quality factor (Q) of the circuit and the transformation ratio $n = R_2/R_1$ independent of each other and consequently to choose the selectivity between certain limits.

For narrowband designs, one can use the following formulas (see [35], where tables are given):

Network (a)

$$X_{C1} = R_1/Q \quad Q \text{ must be selected first} \tag{3-203}$$

$$X_{C2} = R_2 \sqrt{\frac{R_1 R_2}{(Q^2 + 1) - R_1/R_2}} \tag{3-204}$$

$$X_L = \frac{QR_1 + (R_1 R_2/X_{C2})}{Q^2 + 1} \tag{3-205}$$

Network (b)

$$X_{L1} = R_1 Q \quad Q \text{ must be selected first} \tag{3-206}$$

$$X_{L2} = R_2 B \tag{3-207}$$

$$A = R_1(1 + Q^2) \tag{3-208}$$

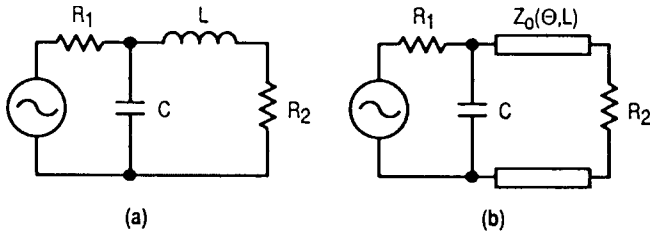


Figure 3-226 Use of a transmission line in the L section.

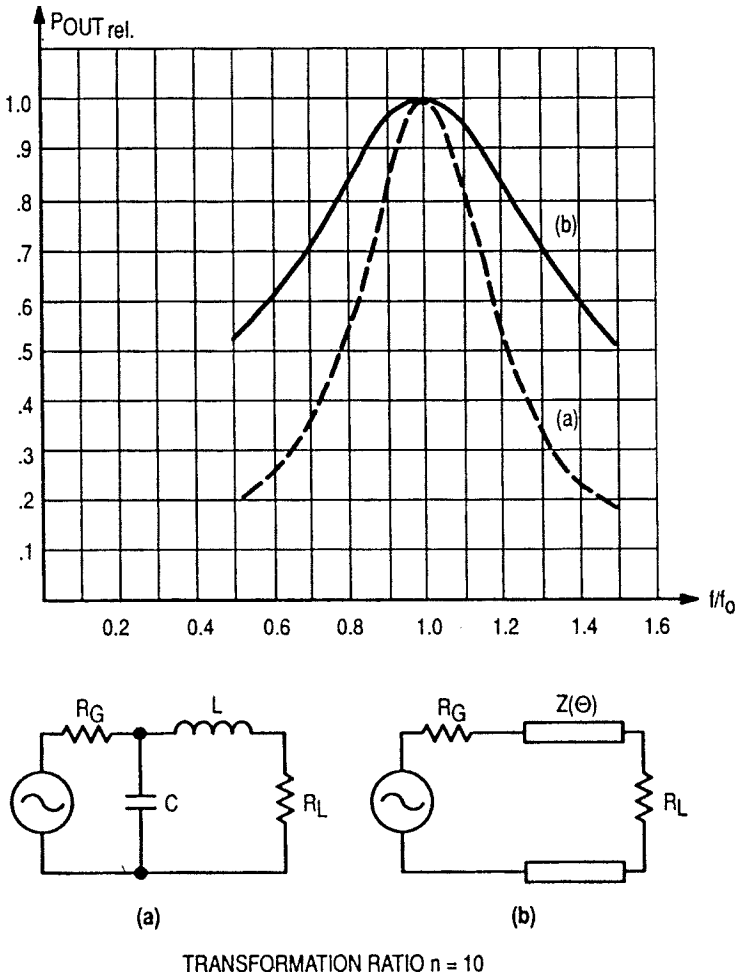


Figure 3-227 Bandwidth of the L section for $n = 10$ (a) with lumped constants and (b) with transmission line (λ/4).

$$X_{C1} = \frac{A}{Q + B} \tag{3-209}$$

$$B = \sqrt{\frac{A}{R_2} - 1} \tag{3-210}$$

Network (c)

$$X_{L1} = QR_1 \quad Q \text{ must be selected first} \tag{3-211}$$

$$X_{C2} = AR_2 \tag{3-212}$$

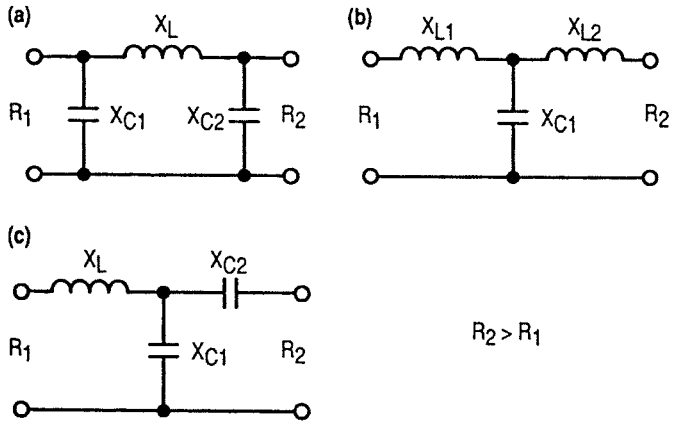


Figure 3-228 Three-reactance matching networks.

$$A = \sqrt{\frac{R_1(1+Q^2)}{R_2} - 1} \tag{3-213}$$

$$X_{C1} = \frac{B}{Q - A} \tag{3-214}$$

$$B = R_1(1+Q^2) \tag{3-215}$$

The network that yields the most-practical component values should be selected for a given application.

The three-reactance networks can be thought of as being formed of an L section (two reactances) and of a compensation reactance. The L section essentially performs the impedance transformation, while the additional reactance compensates for the reactive part of the transformed impedance over a certain frequency band.

Figure 3-229 shows a representation in the Z plane of the circuit of Figure 3-228a split into two parts: $R_1 - C_1 - L_1$ and $C_2 - R_2$. Exact transformation from R_1 into R_2 occurs at the points of intersection M and N. Impedances are then conjugate, or $Z' = R' + jX'$ and $Z'' = R'' + jX''$ with $R' = R''$ and $X' = -X''$.

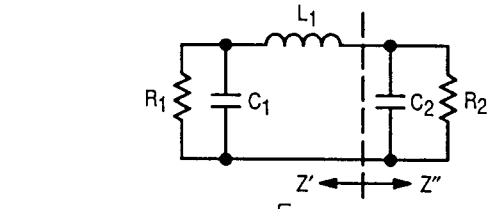
The only possible solution is obtained when X' and $-X''$ are tangential to each other. For the dashed curve, representing another value of L_1 or C_1, a wider frequency band could be expected at the expense of some ripple inside the band. However, this can only be reached with four reactances, as will be shown below.

With a three-reactance configuration, there are not enough degrees of freedom to permit $X' = -X''$ and simultaneously obtain the same variation of frequency on both curves from point M' to point N'. Exact transformation can, therefore, be obtained at only one frequency.

The values of the three reactances can be calculated by making

$$X' = -X'', \quad R' = R'', \quad \text{and} \frac{dX'}{dR'} = -\frac{dX''}{dR''} \tag{3-216}$$

The general solution of these equations leads to complicated calculations. Therefore, computed tables should be used.



$$Z' = \frac{R_1}{1 + \omega^2 R_1^2 C_1^2} + j\omega \left[L1 - \frac{R_1^2 C_1}{1 + \omega^2 R_1^2 C_1^2} \right] = R' + jX'$$

$$Z'' = \frac{R_2}{1 + \omega^2 C_2^2 R_2^2} - \frac{j\omega^2 C_2 R_2^2}{1 + \omega^2 C_2^2 R_2^2} = R'' + jX''$$

$$|x'| = \left(\frac{L_1}{C_1 R_1^2} - \frac{R'}{R_1} \right) \sqrt{\frac{R_1}{R'} - 1}$$

$$|x''| = \sqrt{R' R_2 - R'^2}$$

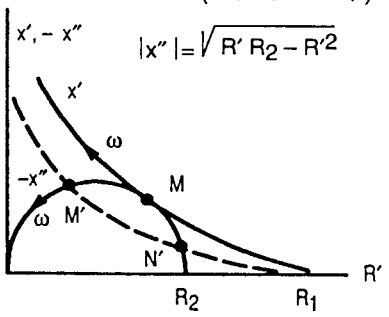


Figure 3-229 *Z*-plane representation of the circuit of Figure 3-228a.

One will note from Figure 3-229 that the compensation reactance contributes somewhat to the impedance transformation; that is, R' varies when going from M to R_2.

The circuit of Figure 3-228b is dual with respect to the first one and gives exactly the same results in a *Y*-plane representation. The circuit of Figure 3-228c is somewhat different since only one intersection M exists, as shown in Figure 3-230. Narrower frequency bands must be expected from this configuration. The widest band is obtained for $C_1 = \infty$.

Again, if one of the terminations has a reactive component, the latter can be taken as a part of the matching network, provided that it is not too large (see Figure 3-227).

FOUR-REACTANCE NETWORKS. Four-reactance networks are used essentially for broadband matching. The networks that will be considered in the following consist of two two-reactance sections in cascade. Some networks have pseudo-low-pass filter character; others, bandpass filter character. In principle, the former show narrower bandwidth since they extend the impedance transformation to very low frequencies unnecessarily, while the latter ensure good matching over a wide frequency band around the center frequency only (see Figure 3-233).

The two-reactance sections used in the networks in Figure 3-231 have either transformation properties or compensation properties. Impedance transformation is obtained with one series reactance and one shunt reactance. Compensation is made with both reactances in series or in shunt. If two cascaded transformation networks are used, transformation is accomplished partly by each one.

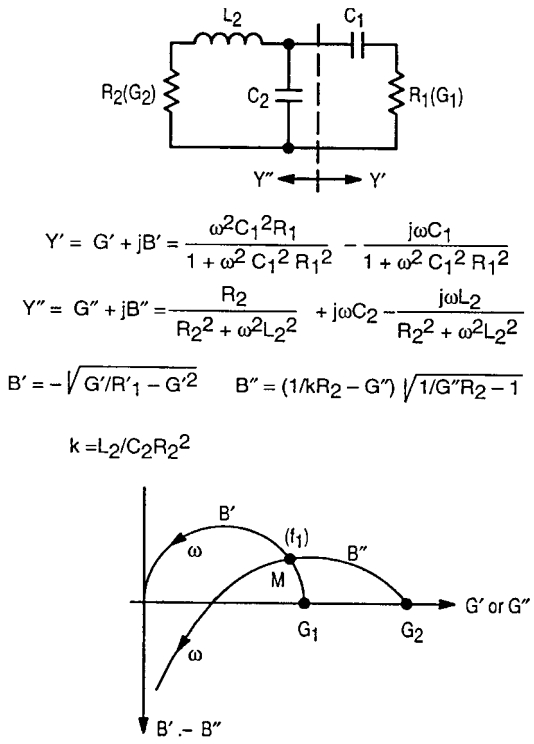


Figure 3-230 *Y*-plane representation of the circuit of Figure 3-228c.

With four-reactance networks there are two frequencies, f_1 and f_2, at which the transformation from R_1 into R_2 is exact. These frequencies may also coincide. For the network in Figure 3-231b, for instance, at point M, R_1 or R_2 is transformed into $\sqrt{R_1R_2}$ when both frequencies fall together. At all points (M), Z_1 and Z_2 are conjugate if the transformation is exact.

In the case of Figure 3-231b, the reactances are easily calculated for equal frequencies:

$$X_1 = \frac{R_1}{\sqrt{(n-1)^{1/2}}} \tag{3-217}$$

$$X_2 = R_1\sqrt{\frac{n^{1/2}-1}{n}} \tag{3-218}$$

$$X_1X_4 = R_1R_2 = X_2X_3 \tag{3-219}$$

$$X_3 = \frac{R_1}{\sqrt{n(n^{1/2}-1)}} \tag{3-220}$$

$$X_4 = \frac{R_1}{n}\sqrt{n^{1/2}-1} \tag{3-221}$$

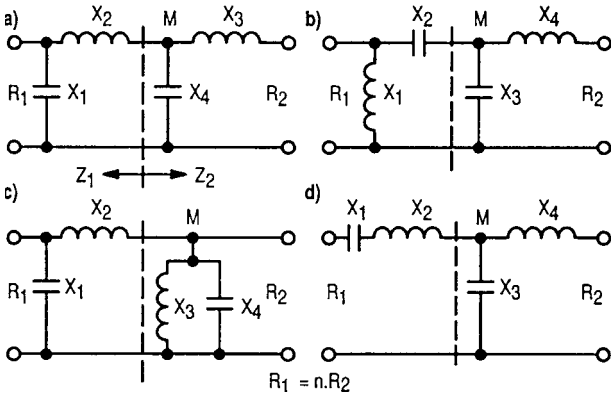


Figure 3-231 Four-reactance networks.

For the network in Figure 3-231a, normally at point M, Z_1 and Z_2 are complex. This pseudo-low-pass filter has been computed elsewhere [33]. Many tables can be found in the literature for networks of four and more reactances having Chebyshev character or maximally flat response [33, 34, 36].

Figure 3-232 shows the transformation path from R_1 to R_2 for networks (a) and (b) in Figure 3-231 on a Smith Chart.

Case (a) of Figure 3-232 has been calculated using tables in Matthaei [34]. Case (b) has been obtained from the relationship given above for $X_1, \ldots, X_4$. Both apply for a transformation ratio equal to 10 and for $R_1 = 1$. There is no simple relationship for $X_1', \ldots, X_4'$ of network (b) if f_1 is made different from f_2 for larger bandwidth. Figure 3-233 shows the respective bandwidths of networks (a) and (b) for the circuits shown in Figure 3-232.

If the terminations include a reactive component, the computed values for X_1 or X_4 may be adjusted to compensate for this.

For configuration (a), it can be seen from Figure 3-232 that, in the considered case, the Q values are equal to 1.6. For configuration (b), Q_1', which is equal to Q_2', is fixed for each transformation ratio

n	2	4	8	10	16
$Q'_1 = Q'_2$	0.65	1	1.35	1.46	1.73

$$Q' = \sqrt{n^{1/2} - 1} \qquad (3\text{-}222)$$

The maximum value of reactance that the terminations may have for use in this configuration can be determined from the above values of Q'.

If R_1 is the load resistance of a transistor, the internal transistor resistance may not be equal to R_1. In this case, the selectivity curve will be different from the curves given in Figure 3-233. Figure 3-234 shows the selectivity for networks (a) and (b) when the source resistance R_1 is infinite. From Figure 3-234 it can be seen that network (a) is more sensitive to changes in R_1 than network (b).

As mentioned earlier, the four-reactance network can also be thought of as two cascaded two-reactance sections: one used for transformation, the other for compensation. Figure 3-235 shows commonly used compensation networks, together with the associated L section.

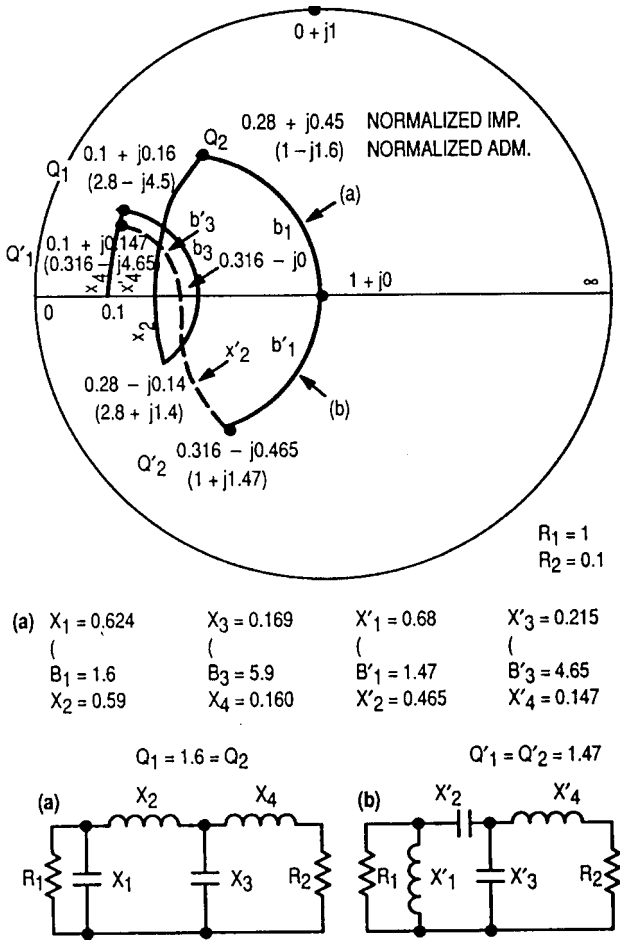


Figure 3-232 Transformation paths for networks (a) and (b).

The circuit of Figure 3-235a can be compared to the three-reactance network shown in Figure 3-228c. The difference is that capacitor C_2 of that circuit has been replaced by an *LC* circuit. The resulting improvement can be seen by comparing Figure 3-236 with Figure 3-230.

By adding one reactance, exact impedance transformation is achieved at two frequencies. It is now possible to choose component values such that the point of intersection M' occurs at the same frequency f_1 on both curves and simultaneously that N' occurs at the same frequency f_2 on both curves. Among the infinite number of possible intersections, only one allows the achievement of this.

When M' and N' coincide in M, the new $dX'/df = dX''/df$ condition can be added to the condition $X' = -X''$ (for three-reactance networks) and similarly $R' = R''$ and $dR'/df = dR''/df$.

Again, a general solution of the above equations leads to still more complicated calculations than in the case of three-reactance networks. Therefore, tables are preferable [33, 34, 36].

The circuit of Figure 3-235b is the dual of the circuit of Figure 3-233a and does not need to be treated separately. It gives exactly the same results in the *Z* plane. Figure 3-235c shows a higher-order compensation requiring six reactive elements.

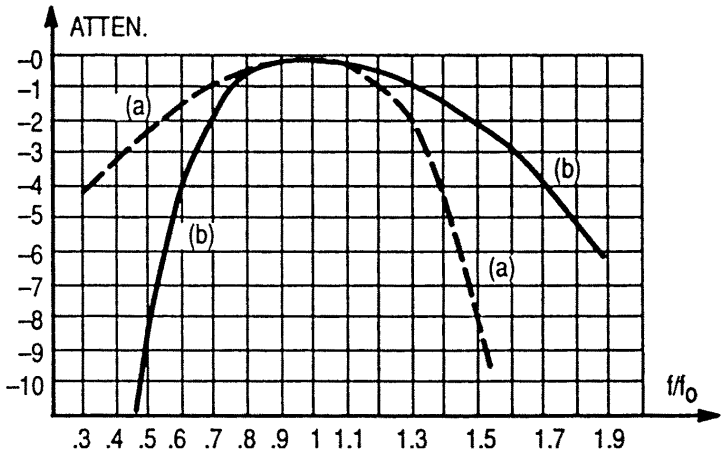


Figure 3-233 Selectivity curves for networks (a) and (b) of Figure 3-232.

The above-discussed matching networks employing compensation circuits result in narrower bandwidths than the former solutions (see p. 573) using two transformation sections. A matching arrangement with higher-order compensation, such as in Figure 3-235c, is not recommended. Better use can be made of the large number of reactive elements by using them all for transformation.

When the above configurations are realized using short portions of transmission lines, the equations or the usual tables no longer apply. The calculations must be carried out on a computer because of their complexity. However, a graphical method can be used (see the next section) that will consist essentially in tracing a transformation path on the Z–Y chart using the computed lumped element values and replacing it by the closest path obtained with distributed constants. The bandwidth change is not significant as long as short portions of lines are used.

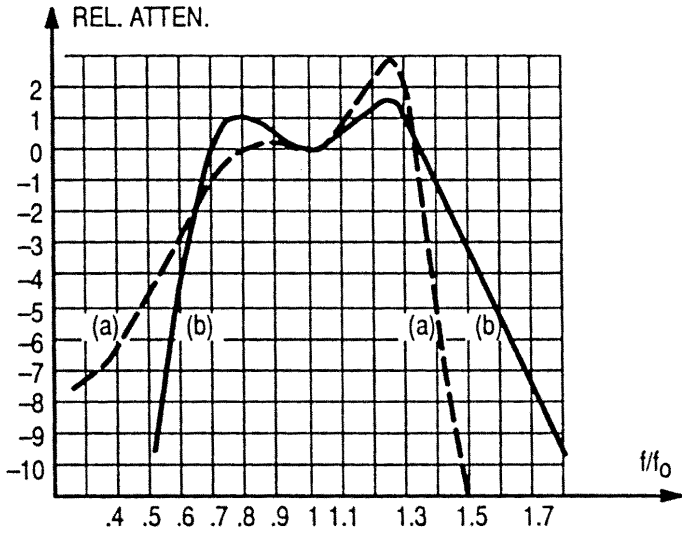


Figure 3-234 Selectivity curves for networks (a) and (b) of Figure 3-232 with infinite R_1.

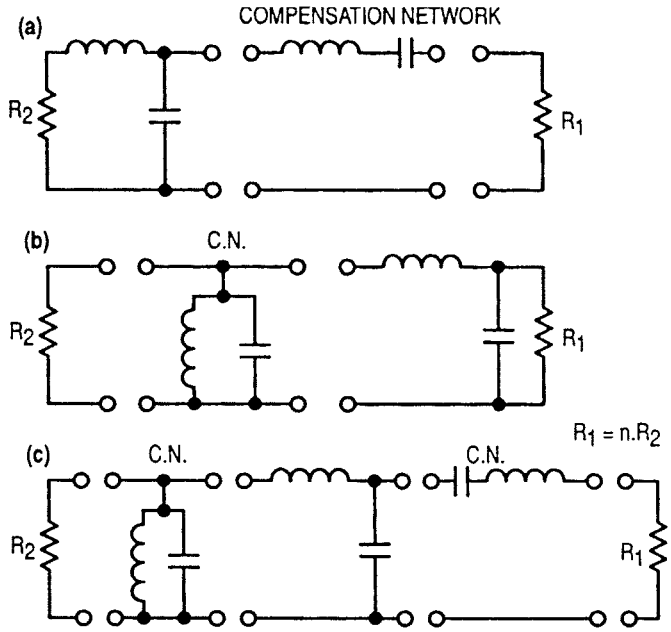


Figure 3-235 Compensation networks used with an L section.

MATCHING NETWORKS USING QUARTER-WAVE TRANSFORMERS. At sufficiently high frequencies, where λ/4-long lines of practical size can be realized, broadband transformations can easily be accomplished by the use of one or more λ/4 sections. Figure 3-237 summarizes the main relations for (a) one-section and (b) two-section transformations. A compensation network can be realized using a λ/2-long transmission line. Figures 3-238 and 3-239 show the selectivity curves for different transformation ratios and section numbers.

EXPONENTIAL LINES. Exponential lines have largely frequency-independent transformation properties. The characteristic impedance of such lines varies exponentially with their length l:

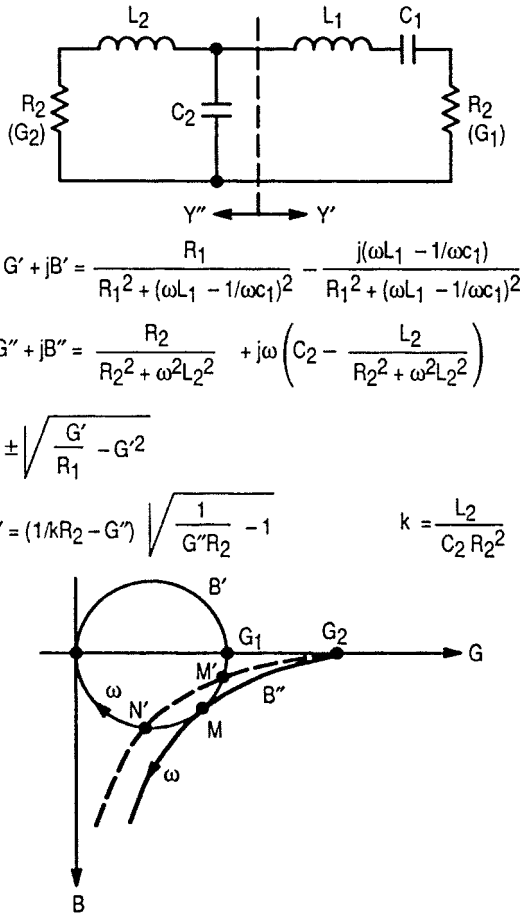
$$Z = Z_0 e^{kl} \tag{3-223}$$

where k is a constant, but these properties are preserved only if k is small.

BROADBAND MATCHING USING BANDPASS FILTER NETWORKS—HIGHQ CASE. The above circuits are applicable to devices having low input or output Q, if broadband matching is required. Generally, if the impedances to be matched can be represented, for instance, by a resistor R in series with an inductor L (sometimes a capacitor C) within the band of interest and if L is sufficiently low, the latter can be incorporated into the first inductor in the matching network. This is also valid if the representation consists of a shunt combination of a resistor and a reactance.

Practically, this is feasible for Q values around 1 or 2. For higher Q values or for input impedances consisting of a series or parallel resonant circuit (see Figure 3-221), as it appears to be for large bandwidths, a different treatment must be followed.

Let us first recall that, as shown by Bode [37] and Fano [38], limitations exist on the impedance matching of a complex load. In the example of Figure 3-240, the load to be matched consists of a capacitor C and a resistor R in shunt.



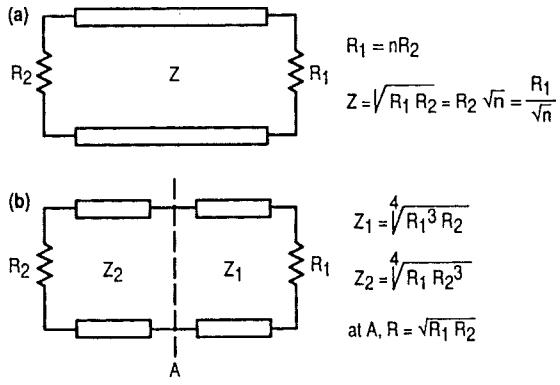
$$Y' = G' + jB' = \frac{R_1}{R_1^2 + (\omega L_1 - 1/\omega c_1)^2} - \frac{j(\omega L_1 - 1/\omega c_1)}{R_1^2 + (\omega L_1 - 1/\omega c_1)^2}$$

$$Y'' = G'' + jB'' = \frac{R_2}{R_2^2 + \omega^2 L_2^2} + j\omega\left(C_2 - \frac{L_2}{R_2^2 + \omega^2 L_2^2}\right)$$

$$B' = \pm\sqrt{\frac{G'}{R_1} - G'^2}$$

$$B'' = (1/kR_2 - G'')\sqrt{\frac{1}{G''R_2} - 1} \qquad k = \frac{L_2}{C_2 R_2^2}$$

Figure 3-236 *Y*-plane representation of the circuit of Figure 3-235a.



$$R_1 = nR_2$$

$$Z = \sqrt{R_1 R_2} = R_2\sqrt{n} = \frac{R_1}{\sqrt{n}}$$

$$Z_1 = \sqrt[4]{R_1^3 R_2}$$

$$Z_2 = \sqrt[4]{R_1 R_2^3}$$

$$\text{at A, } R = \sqrt{R_1 R_2}$$

Figure 3-237 Transformation networks using λ/4-long transmission lines.

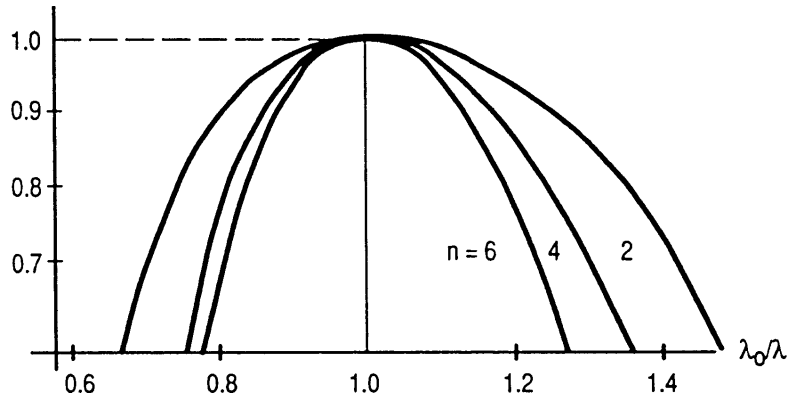


Figure 3-238 Selectivity curves for two λ/4-section networks at different transformation ratios.

The reflection coefficient between the transformed load and generator is equal to

$$\Gamma = \frac{Z_{\text{IN}} - R_G}{Z_{\text{IN}} + R_G} \tag{3-224}$$

When $\Gamma = 0$, there is perfect matching; when $\Gamma = 1$, there is total reflection. The ratio of reflected to incident power is

$$\frac{P_r}{P_i} = |\Gamma|^2 \tag{3-225}$$

The fundamental limitation on the matching takes the form

$$\int_{\omega=0}^{\infty} \ln\left(\frac{1}{|\Gamma|}\right) d\omega \le \frac{\pi}{RC} \qquad \text{Bode equation} \tag{3-226}$$

and is represented in Figure 3-241.

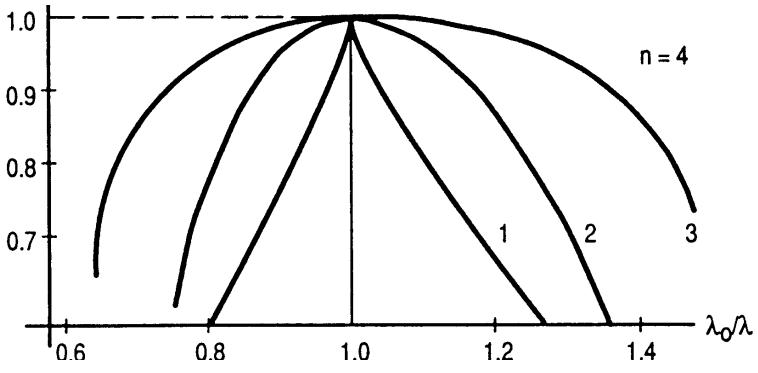


Figure 3-239 Selectivity curves for one, two, and three λ/4 sections.

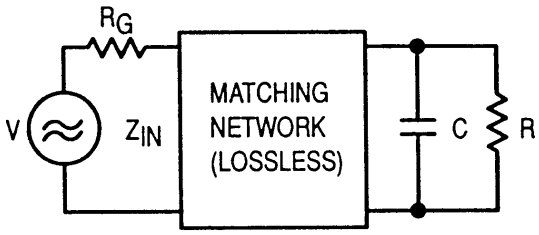


Figure 3-240 General matching conditions.

The meaning of the Bode equation is that the area S under the curve cannot be greater than π/RC and, therefore, if matching is required over a certain bandwidth, this can only be done at the expense of less power transfer within the band. Thus, power transfer and bandwidth appear as interchangeable quantities.

It is evident that the best utilization of the area S is obtained when $|\Gamma|$ is kept constant over the desired band ω_c and made equal to 1 over the rest of the spectrum. Then

$$|\Gamma| = e^{-\pi/\omega_c RC} \tag{3-227}$$

within the band and no power transfer happens outside. A network fulfilling this requirement cannot be obtained in practice because an infinite number of reactive elements would be necessary.

If the attenuation is plotted versus frequency for practical cases, one may expect to have curves like the ones shown in Figure 3-242 for a low-pass filter having Chebyshev character.

For a given complex load, an extension of the bandwidth from ω_1 to ω_2 is possible only with a simultaneous increase of the attenuation a. This is especially noticeable for Q values exceeding 1 or 2 (see Figure 3-243). Thus, devices having relatively high input Q values are usable for broadband operation, provided the consequent higher attenuation or reflection introduced is acceptable.

For average insertion losses or attenuation a (neglecting the ripple), the attenuation decreases if the number n of the network element increases. But above $n = 4$, the improvement is small. For a given attenuation a and bandwidth, the larger the n the smaller the ripple; for a given attenuation, the larger the n the larger the bandwidth.

Computations show that for $Q < 1$ and $n \leq 3$, the attenuation is below 0.1 dB (approximately). The impedance transformation ratio is not free here. The network is a

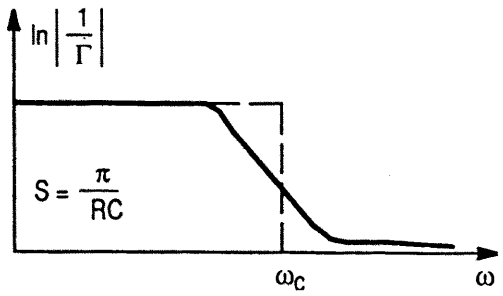


Figure 3-241 Representation of the Bode equation.

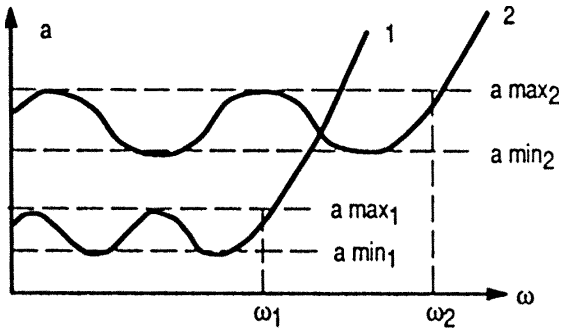


Figure 3-242 Attenuation versus angular frequency for different bandwidths with the same load.

true low-pass filter. For a given load, the optimum generator impedance will result from the computation.

Before impedance transformation is introduced, a conversion of the low-pass prototype into a bandpass filter network must be made. Figure 3-244 summarizes the main relations for this conversion. r is the conversion factor. For the bandpass filter, $Q_{\text{IN max}}$, or the maximum possible input Q of the device to be matched, has been increased by the factor r (from Figure 3-244, $Q'_{\text{IN max}} = rQ_{\text{IN max}}$).

Impedance inverters will be used for impedance transformation. These networks are suitable for insertion into a bandpass filter without affecting the transmission characteristics.

Figure 3-245 shows four impedance inverters. Note that one of the reactances is negative and must be combined in the bandpass network with a reactance of at least equal positive value. Insertion of the inverter can be made at any convenient place [32, 39].

When using the bandpass filter for matching the input impedance of a transistor, reactances $L'_1C'_1$ should be made to resonate at ω_0 by addition of a convenient series reactance.

As stated earlier, the series combination of R_0, L'_1 and C'_1 normally constitutes the equivalent input network of a transistor when considered over a large bandwidth. This is a good approximation up to about 500 MHz.

In practice, the normal procedure for using a bandpass filter matching network will be the following:

1. For a given bandwidth center frequency and input impedance of a device to be matched—for example, 50 Ω—first determine Q'_{IN} from the datasheet as

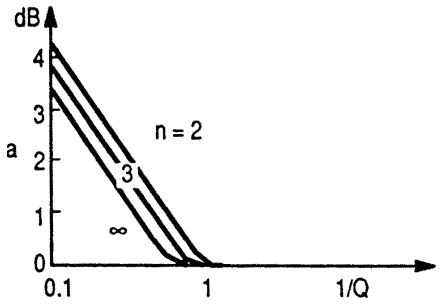


Figure 3-243 Insertion losses as a function of 1/*Q*.

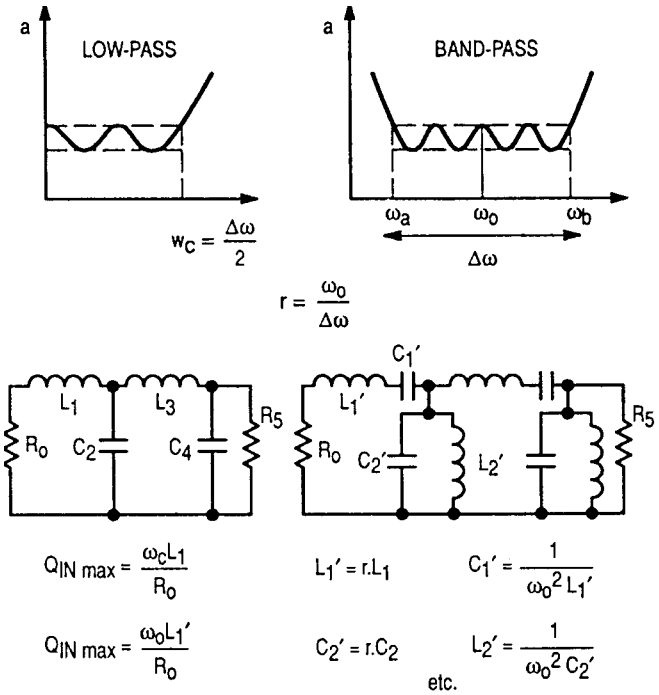


Figure 3-244 Conversion from low-pass filter into bandpass filter.

$$\frac{\omega_0 L_1'}{R_0} \tag{3-228}$$

after having eventually added a series reactor for centering.

2. Convert the equivalent circuit $R_0 L_1' C_1'$ into a low-pass prototype $R_0 L_1$ and calculate Q_{IN} using the formulas of Figure 3-244.
3. Determine the other reactance values from tables [33] for the desired bandwidth.
4. Convert the element values found by Step 3 into series- or parallel-resonant circuit parameters.
5. Insert the impedance inverter in any convenient place.

In the above discussions, the gain roll-off has not been taken into account. This is of normal use for moderate bandwidths (e.g., 30%). However, several methods can be employed to obtain a constant gain within the band despite the intrinsic gain decrease of a transistor with frequency. Tables have been computed elsewhere [40] for matching networks approximating 6 dB/octave attenuation versus frequency.

Another method consists of using the above-mentioned network and then adding a compensation circuit, as shown in Figure 3-246.

Resonance ω_b is placed at the high edge of the frequency band. Choosing Q correctly, roll-off can be made 6 dB/octave. The response of the circuit of Figure 3-246 is expressed by

$$\frac{1}{1 + Q^2(\omega/\omega_b - \omega_b/\omega)^2} \tag{3-229}$$

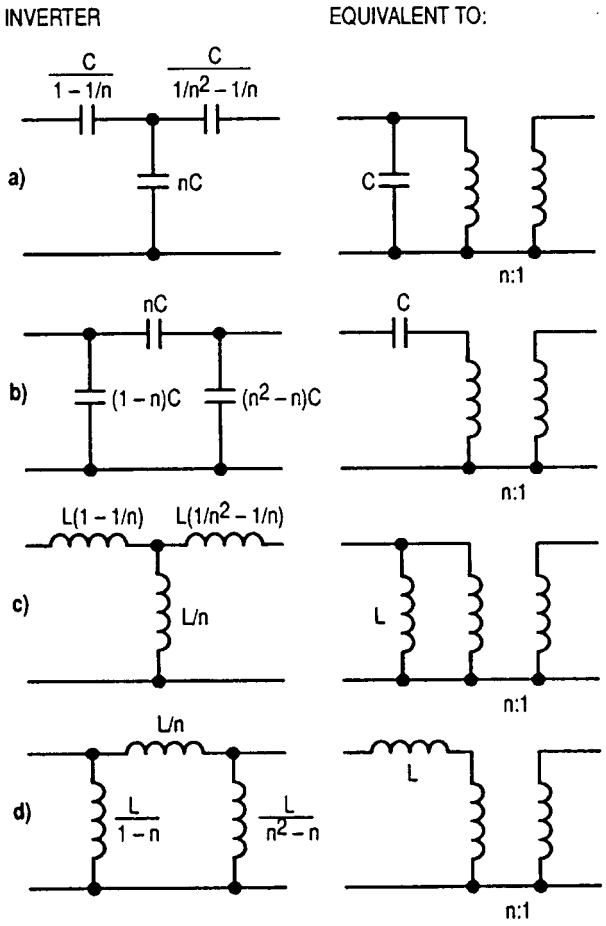


Figure 3-245 Impedance inverters.

where $\omega < \omega_b$. This must be equal to ω/ω_b for 6 dB/octave attenuation.

At the other band edge a, exact compensation can be obtained if

$$Q = \frac{(\omega_b/\omega_a)^2 - 1}{\omega_a/\omega_b - (\omega_b/\omega_a)^2} \tag{3-230}$$

There are several methods available for synthesis of matching networks. They should not be confused with filters because they can be either high-pass, low-pass, or bandpass, and their main purpose is to match one complex impedance to another. In many cases, one of the terminations happens to be 50 Ω and real, but this is just a "lucky" case of the general approach. An excellent paper on this [41] refers to the "real frequency" technique. Another equally important publication is the one by Cuthbert [42], since followed by his book [43]. Two more good references are Vendelin et al. [44] and Potter [45].

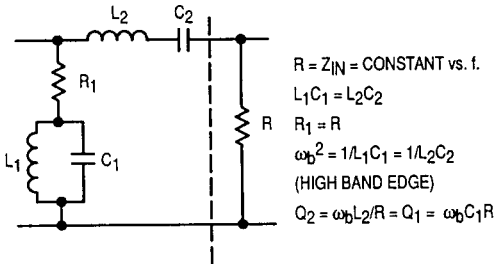


Figure 3-246 Roll-off compensation network.

3-12-3 Example 2: Low-Noise Amplifier Using Distributed Elements

In our previous examples, the amplifiers were built around lumped elements. Since a transmission line length of less than $\lambda/4$ is an inductor and $\frac{3}{4}\lambda$ is a capacitor, the whole amplifier can be built around microstrip technology.

If anyone wonders why we are going to show a low-noise amplifier in the middle of the power section, the reason is that we will select a low-power Siemens transistor, which, being operated in Class A, has an intercept point of +30 dBm and can easily achieve a noise figure of less than 2 dB. Such an amplifier is attractive when feeding an input stage from an antenna that sees many signals and curtails the bandwidth with an input filter of good selectivity that needs a good output termination, which in many cases works against the achievement of a good noise figure because when designing for best noise figure the input impedance of the transistor stage is far from 50 Ω.

First, we will start with the textbook approach and use the Smith diagram to obtain the best noise figure of which the device is capable. Figure 3-247 shows the actual approach to input and output matching.

We start our design by querying the Smith Tool about the minimum noise figure and the gain that comes with it. We mark the point of the lowest noise figure (labeled Γ_{opt}). The small

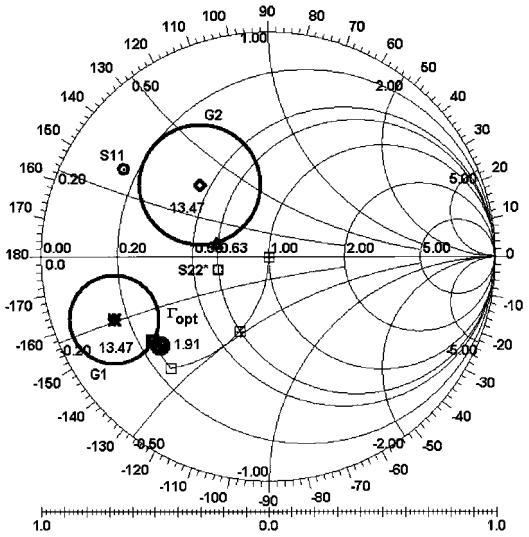


Figure 3-247 Input and output matches for 13.47 dB of gain and NF = 1.91 dB as displayed by the interactive Smith Tool.

noise circle yields the location for Γ_{opt} and tells us the minimum noise figure is going to be 1.91 dB. (By the way, this assumes no further losses, even in soldering the transistor to the PC board.) Then we construct the input gain circle (G1), which touches Γ_{opt} and turns out to represent 13.47 dB of available gain.

The next step is to find a matching network between Γ_{opt} and the 50-Ω source impedance. Immediately there is a conflict: If we look at S_{11}, which is the input reflection coefficient shown in the Smith diagram, it is on the other side of the complex plane and sufficiently far away, indicating that there is going to be quite a difference between noise matching and power-gain matching. We have two options for matching the input: Either we start from the origin and move toward the position of Γ_{opt}, or we start with Γ_{opt} and go to 50 Ω. Depending on the direction in which we do this, the result can be different topologies, such as a high-pass or low-pass filter. The high-pass filter is obtained by starting with Γ_{opt}; the low-pass topology is obtained by starting at the 50-Ω point. The Smith Tool then provides us with the corresponding values of capacitors and inductors after the transformation has been accomplished.

The final step is to do the same type of matching at the output. We start with S_{22}, and either select S_{22}^* (the complex conjugate for highest gain) or S_{22load}. As a rule, the second one is always a lower impedance (in terms of the real part) than the conjugate match. This graphic procedure is somewhat trial and error. As to the topology, a high-pass match should be preferred at the input and a low-pass match at the output. The reason for this is that the high-pass filter will tend to protect the stage from the enormous number of signals present below 1 GHz; using a low-pass filter at the output suppresses unwanted harmonics that the amplifier may generate. Figure 3-248 shows one of the approaches we like best simply because it gives a good input termination and we were trying for a simultaneous match for best noise figure and power gain.

The way we were able to achieve this was to take advantage of the existing feedback in the transistor and select a less-than-optimal output termination but gaining at the input. Taking advantage of the Miller effect, we rotate the phase angle of the feedback with the load to move the optimum noise and gain matching points closer. The danger in this is that the circuit is now highly sensitive to both input and output termination. The output *must* see 50 Ω to maintain this good input match, but since the second transistor stage can now be optimized for best input match, a two-stage combination can be configured to meet this

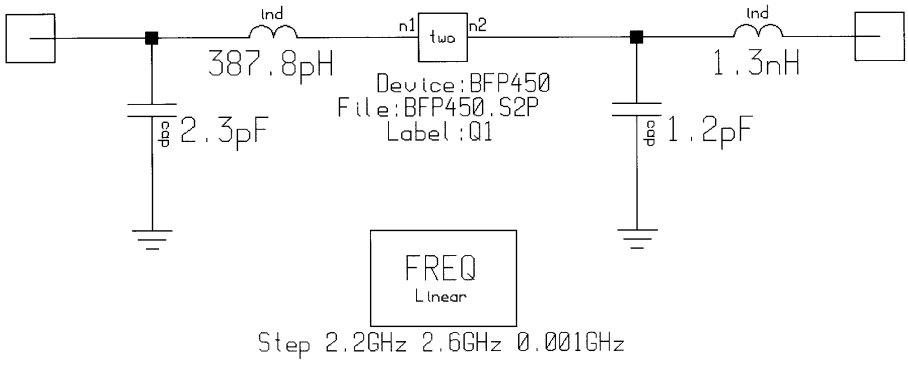


Figure 3-248 A 2.4-GHz amplifier using a BFP450 transistor and lumped-element input and output matching.

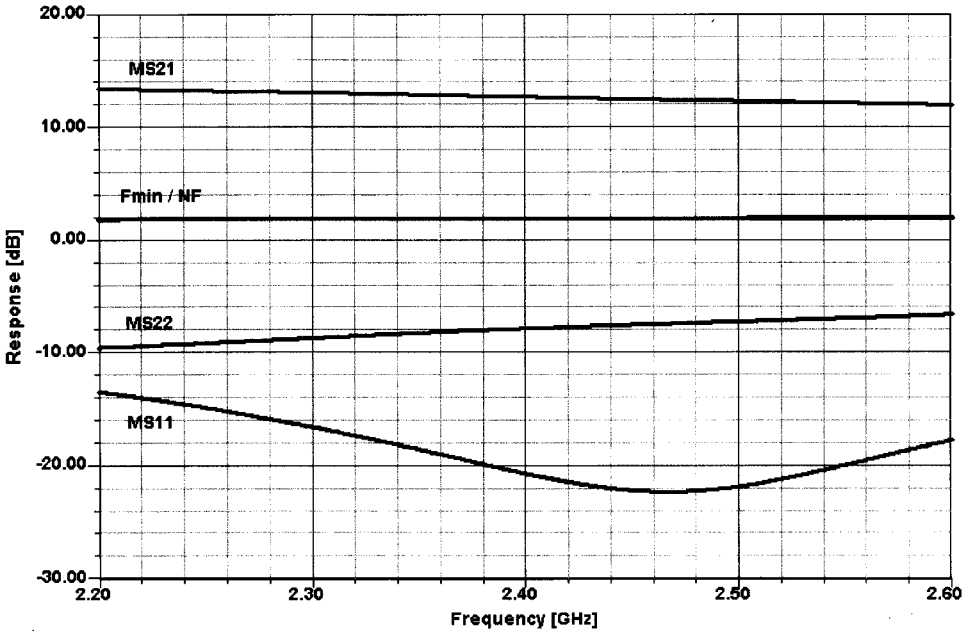


Figure 3-249 Frequency-dependent gain, matching, and noise characteristics of the matched high-dynamic-range BFP450 2.4-GHz amplifier.

requirement. Also, let us not forget that the device is running at 50 mA and as far as bipolar transistors are concerned, this is probably the best amplifier combination we have seen for getting the highest dynamic range and lowest noise figure. The resulting gain, matching, and noise figure characteristics are shown in Figure 3-249.

Now we turn to using distributed elements and at the same time use a shunt feedback at the output; to be specific, we will load it with 150 Ω. The reason for this is to be able to obtain a good compromise at the input that will depend less on the output and at the same

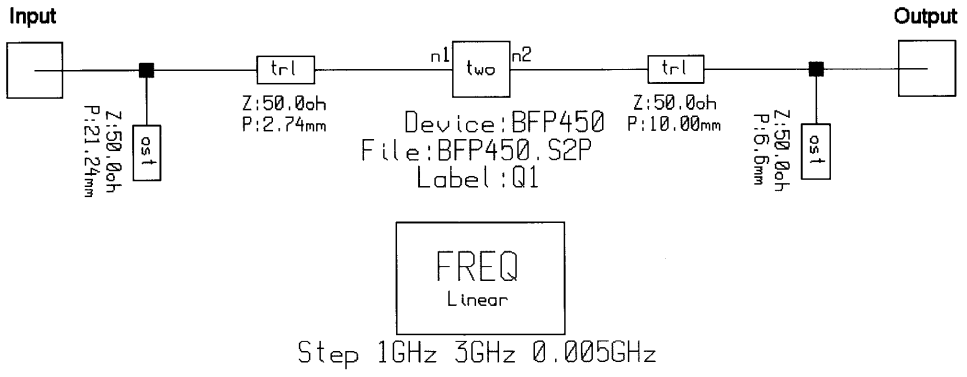


Figure 3-250 Distributed-element input and output matching for the 2.4-GHz BFP450 amplifier.

time improve the S_{22} matching at the output. The procedure for doing the matching with a transmission line offers only certain degrees of freedom, while doing the matching with lumped elements allows the choice of series or parallel inductors or capacitors. In the case of a transmission line, we have open stubs, shorted stubs, and transmission lines. This may result in somewhat more complex matching schemes. On the positive side, they can easily be trimmed, are reproducible, and certainly cost less than exotic capacitors, which need to have small tolerances and to be able to handle high RF currents.

The resulting input and output matching is shown in Figure 3-250, and its frequency response in Figure 3-251. We were able to maintain a good input and output match with an insignificant deterioration of the noise figure.

Since this is a high-dynamic-range amplifier, we also show its predicted two-tone response in Figure 3-252. The manufacturer specifies an $IP_{3,in}$ of 29 dB for a bias point of $V_{CE} = 3$ V and $I_C = 50$ mA with $Z_S = Z_{S,opt}$ and $Z_L = Z_{L,opt}$ at 1.8 GHz. Significantly, however, the manufacturer-supplied nonlinear BFP450 library we used exhibits considerably different characteristics at $V_{CE} = 4$ V and $I_C = 50$ mA than those reflected in the corresponding S parameters supplied by the same manufacturer. In our circuit, the S parameters resulted in about 13.5 dB of gain (based on a compromise between matching for optimum noise and maximum gain) versus about 9.5 dB of gain for the nonlinear library (match optimized for maximum gain). The difference between the 2.4-GHz F_{min} values returned for the two modeling approaches is similarly striking: 1.90 dB for the bare S parameter set and 7.76 dB for the bare nonlinear library biased to $V_{CE} = 4$ V and $I_C = 50$ mA. Clearly, the nonlinear modeling of the BFP450 needs work!

Having fully evaluated the amplifier at 2.4 GHz as a working example, we note that this frequency is used by radio amateurs for some applications. We will therefore supply

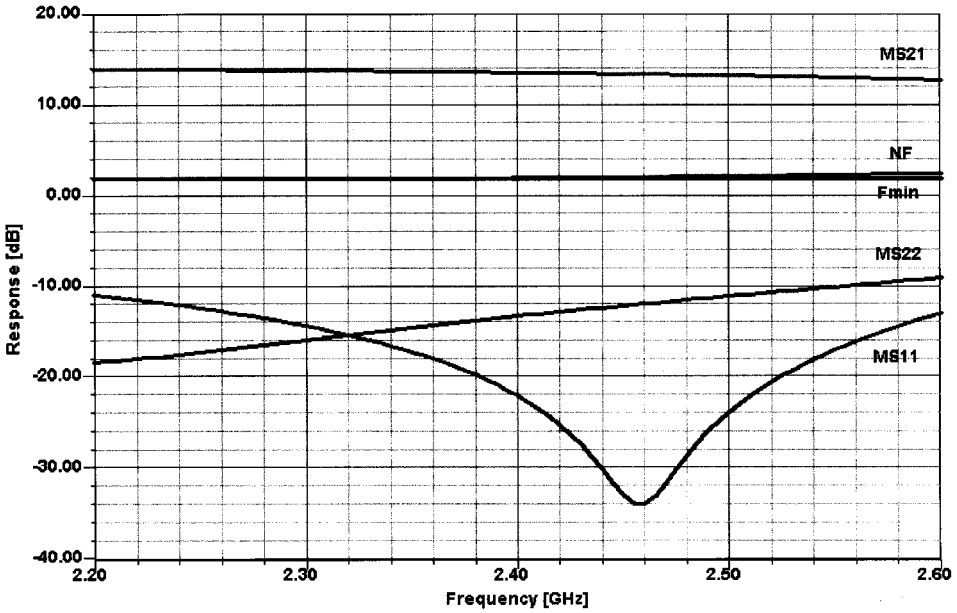


Figure 3-251 Predicted frequency-dependent gain, matching, and noise performance for the BFP450 amplifier with distributed-element matching.

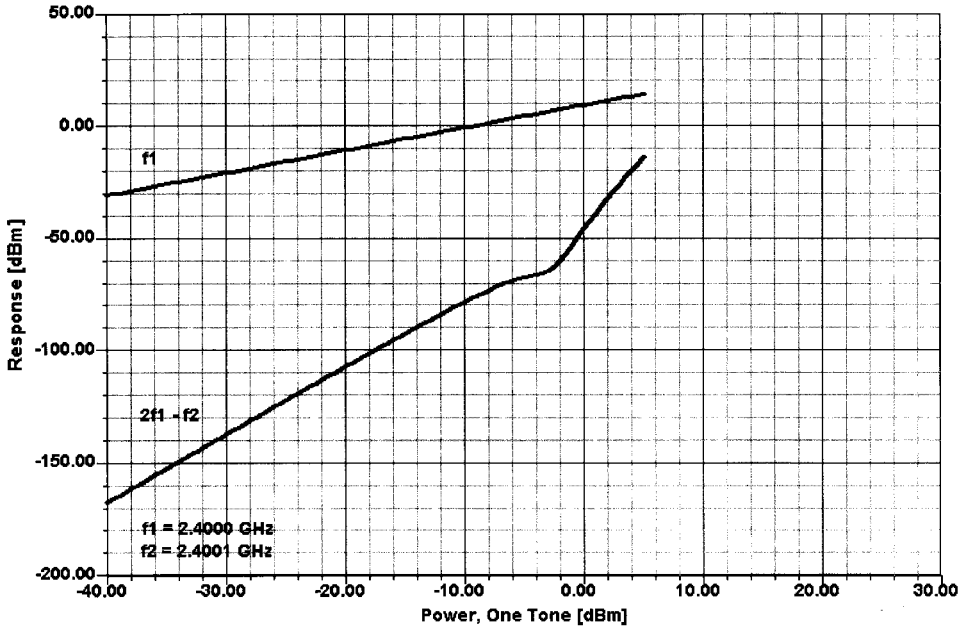


Figure 3-252 Predicted two-tone response of the BFP450 amplifier with the circuit adjusted for maximum gain (9.5 dB): $IP_{3,in} \approx 28.5$ dBm and $IP_{3,out} \approx 38$ dBm. That this version's gain is significantly lower than that of the *S*-parameter-based equivalent illustrates the need for better accuracy in parameter extraction.

examples for 432 and 1296 MHz, popular spot frequencies in the amateur 70- and 23-cm bands, respectively. Figures 3-253 and 3-254 show the 432-MHz version's schematic and frequency-dependent responses. Figures 3-255 and 3-256 show the schematic and frequency-dependent responses for the 1296-MHz version.

Using a bipolar transistor that operates at much less current, the noise figure can be improved at the expense of the dynamic range/intercept point. It is possible to reduce the noise figure by about 1 dB or so, but the reduction in intercept point will be much more dramatic. Another choice is the use of GaAsFETs for this application, but their matching in comparable circuits has been described as a nightmare simply because the devices have too much gain and very high input impedances (essentially 0.2 pF in series with 3 Ω) combined with reasonable output impedances in the vicinity of 240 Ω. Here is a wide field in which engineers can experiment and examine price/performance and performance/real estate issues. GaAsFETs also tend to require more than 30–40 mA to be alive, and even the low operating voltage of 4–5 V makes the device consume a lot of dc power.

3-12-4 Example 3: 1-W Amplifier Using the CLY15

The following is an application showing a fairly wideband amplifier with +30 dBm (1 W) output. We need to point out that in accordance with the various digital modulation modes used with this amplifier, this is not continuous operation, but a much less than 1:1 duty cycle operation. The device would probably not survive operating at full output under CW conditions. This is subject to a discussion with the manufacturer, Siemens, who has not

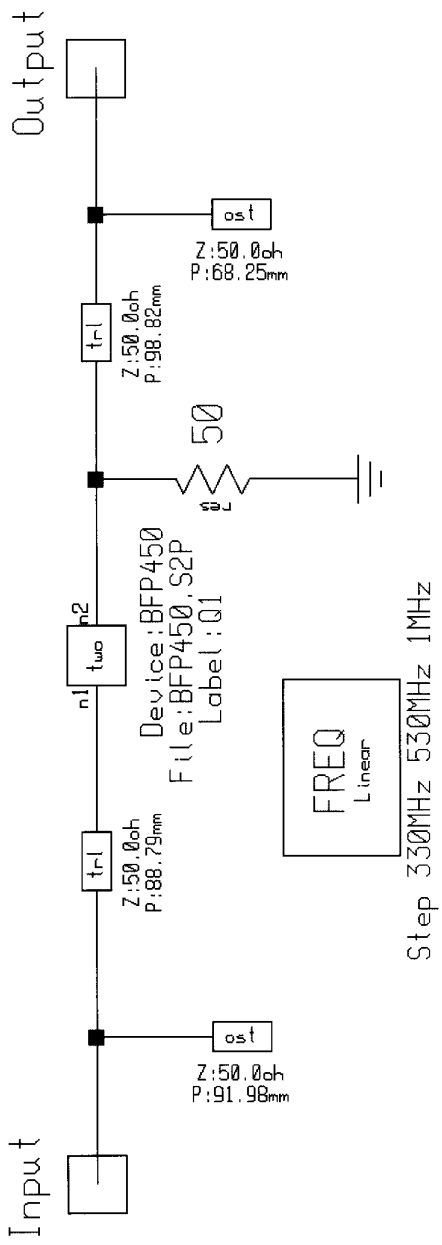


Figure 3-253 Schematic of BPF450 amplifier for 432 MHz. Because f_T = 25 GHz for the BFP450, at this lower frequency we must shunt the device output with a 50-Ω resistor to maintain stability.

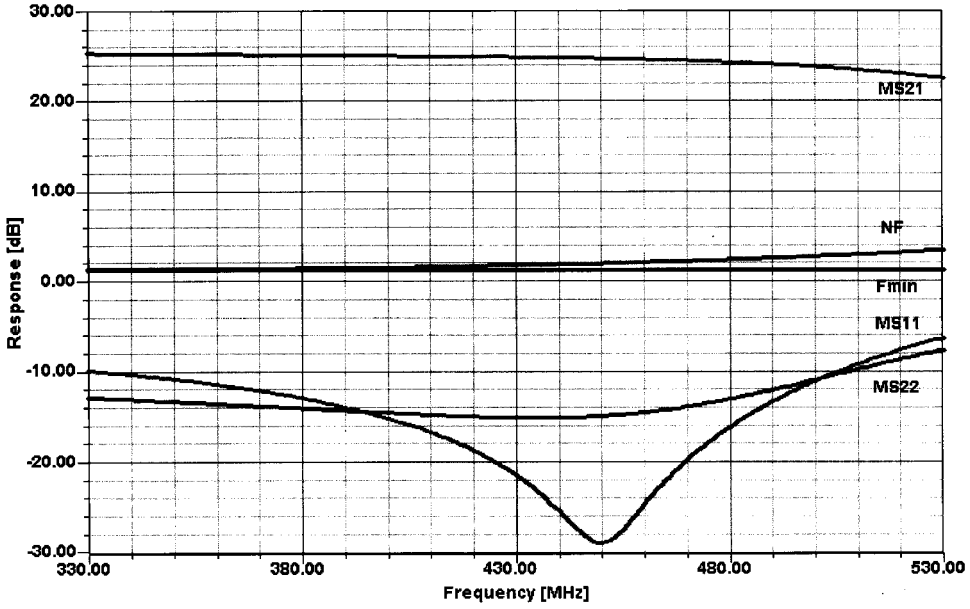


Figure 3-254 Frequency-dependent gain, matching, and noise performance for the BPF450 amplifier at 432 MHz. As with the circuit in Figure 3-253, the input and output network values have been chosen to achieve a low NF and good input matching at the expense of the output match.

released sufficient data for us to make a judgment as to the CLY15's continuous power capability. Since there is no useful large-signal model available at the time of this writing, we will design this Class A amplifier using large dc bias *S* parameters. In the actual design there will be some cut-and-try, but we find that in this type of application this is still the best way of handling the design.

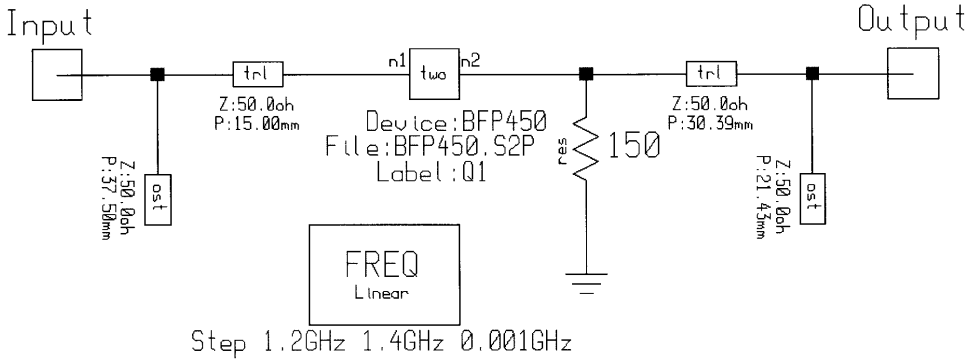


Figure 3-255 Schematic of the BPF450 amplifier for 1296 MHz. A loading resistor is still necessary to ensure stability, but because the transistor exhibits less gain at this frequency than at 432 MHz, we can increase the loading resistor to 150 Ω.

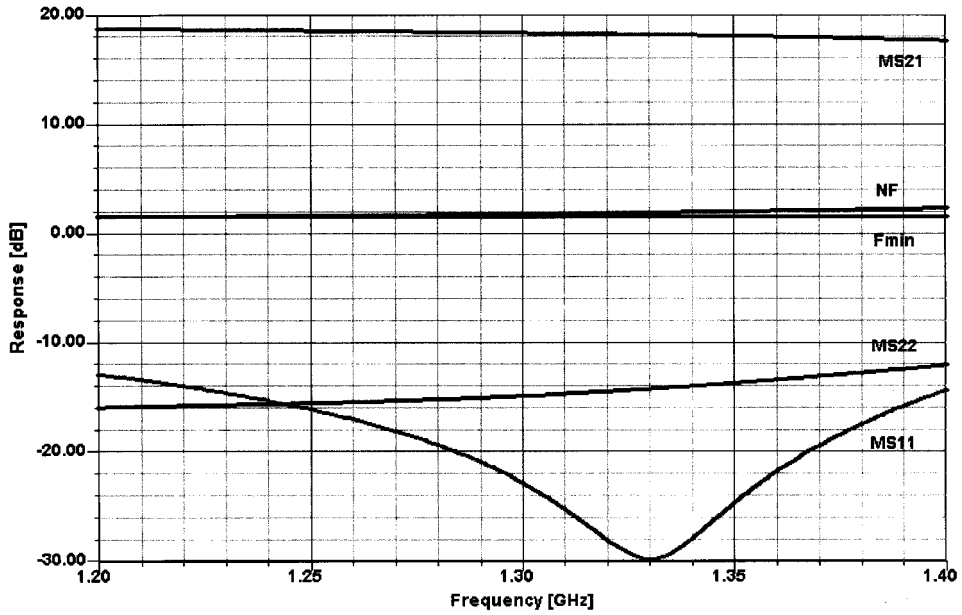


Figure 3-256 Frequency-dependent gain, matching, and noise responses of the BFP450 amplifier at 1296 MHz. As with the two previous versions, the input and output network values have been chosen to achieve a low NF and good input matching at the expense of the output match.

The CLY15 (gate width 16 mm) is the largest of a series of Siemens GaAsFETs; it was already mentioned in Chapter 2, where its datasheet is reproduced. Its device family also includes the CLY2 (gate width 2 mm), CLY5 (4 mm), and CLY10 (8 mm). It is possible that bigger devices will be offered in the future. One of the most critical parameters will be the saturation voltage; from the datasheet of the CLY15, we can see that at 1.5 A the saturation voltage is approximately 1 V. This drop of 1 V includes losses that result from other parasitic elements; the dc *I–V* curve (Figure 3-257a) indicates a saturation voltage of 0.75. Since we aim for 1-W output power, the resistive portion of the load will be $(5 - 1)^2 / (1 \times 2) = 8\ \Omega$. This assumes a supply voltage of 5 V.

The next step is to design the input and output matching network. Previewing the *S* parameters, it will be noticed that they are essentially inductive, which means that the input match, using lumped elements, will result in a combination of capacitors at both the input and the output. Figures 3-258 and 3-259 show the ways in which the input and output matches were done on the Smith diagram resulting in lumped matching elements. Inspecting the lumped input matching, we find the following. We start with S_{11}, which is in the lower part of the Smith diagram, and determine its conjugate match, S_{11}^*. We already pointed out that this device is inductive, which means that matching it from the conjugate point will require two capacitors. The first one is a series capacitor, as identified in the Smith diagram, and followed by a shunt capacitor to transform the impedance to 50 Ω. A close inspection of the Smith diagram shows that the resulting point is not quite on S_{11}; it is an iterative process to exactly find the crossover point, yet the input match of better than 20 dB is more than acceptable.

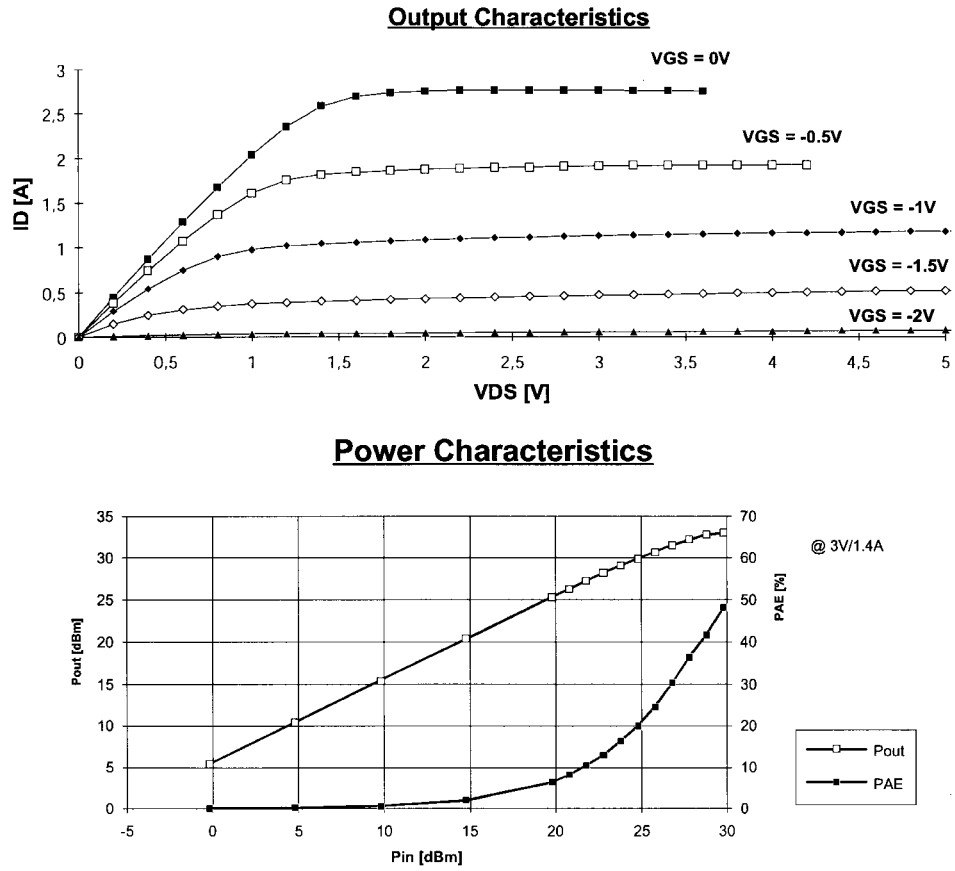


Figure 3-257 Output and power characteristics for the CLY15 medium-power GaAsFET.

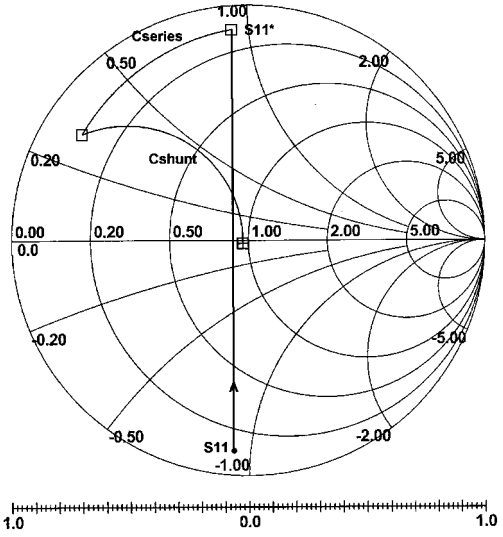


Figure 3-258 Input match for S_{11}^{*} using lumped elements.

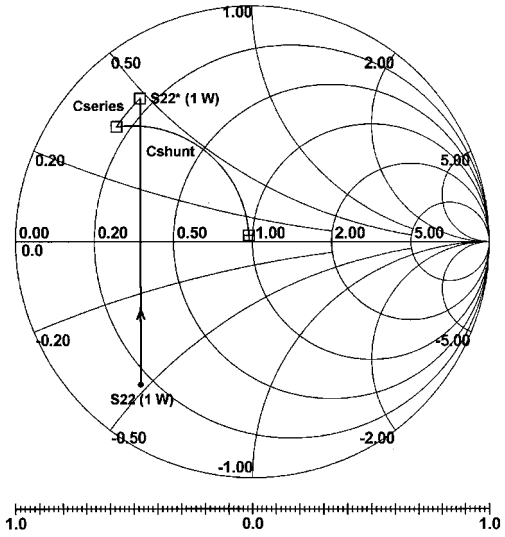


Figure 3-259 Output match for S_{22}^* (1 W) using ideal lumped elements.

As far as the output is concerned, we use the same technique by moving from S_{22} to S_{22}^*. S_{22} has been determined from the required output power; as we showed above, the real portion is 8 Ω. Using the lumped elements, the matching technique is identical; we first need to use a series capacitor and then a shunt capacitor for the transformation. The arc for the series capacitor this time is much longer, resulting in a series capacitance of 11.2 pF instead of the 2.1-pF capacitance used at the input. The actual schematic is shown Figure 3-260. The drawback of lumped elements is that it is not practical to buy components that can handle the heavy RF currents that flow in the input and output networks. Taking the values obtained

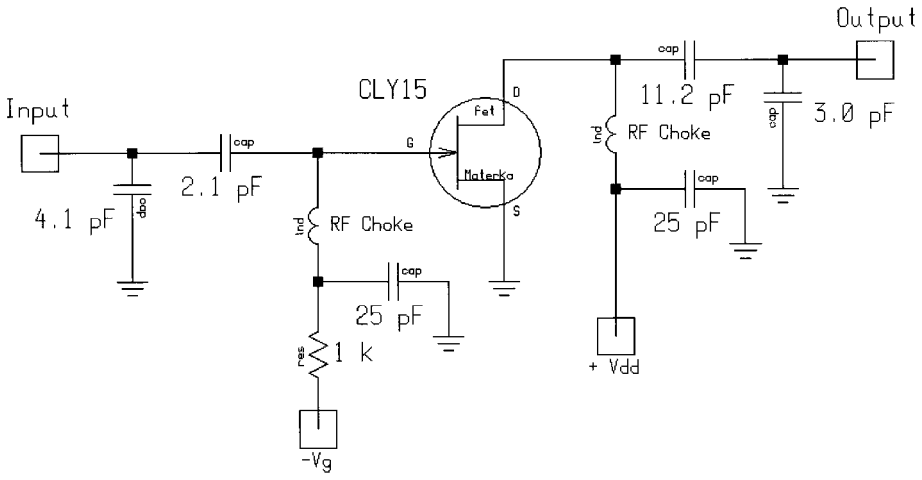


Figure 3-260 The complete amplifier using the CLY15 medium-power GaAsFET, with lumped-element input and output matching.

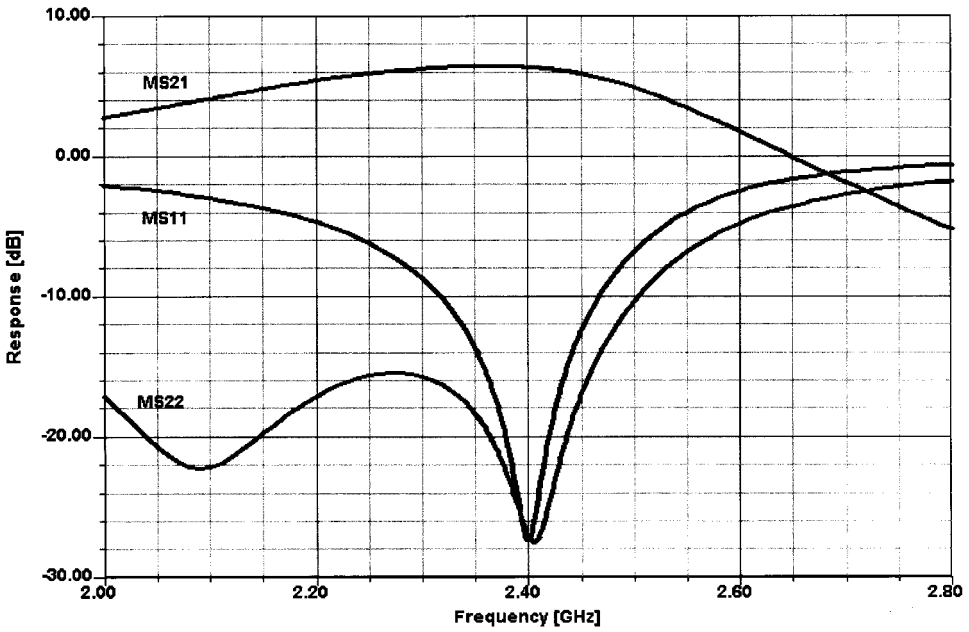


Figure 3-261 Frequency-dependent gain and matching responses of the CLY15 amplifier using ideal lumped elements.

from the lumped input and output matching approach, Figure 3-261 shows the resulting frequency response.

The next step is to provide the same matching condition using distributed elements, which of course have the advantage that they can be printed on the PC board. Instead of using ideal elements, we used a loss tangent of 2E–3, or a Q of 500. The manufacturer of the R4000 material (Rogers) provided this specification. Now we start with the same approach by moving from S_{11} to S_{11}^*. From there, we use an open stub with a transmission-line impedance of 50 Ω, followed by a transmission line to the point in the left lower corner, and, finally, a shorted transmission-line stub allows the 50-Ω matching. For this amplifier, we are not looking for noise matching but for input and output power matching. Figure 3-262 shows this match on the Smith Chart.

At the output, the procedure is again similar: We move from the S_{11} for 1 W to a conjugate value, then use an open stub that acts like a capacitor until we cross the real axis; finally, a λ/4 transmission line is used to match from approximately 75 Ω to 50 Ω. Figure 3-263 shows this match on the Smith Chart. This concludes the distributed matching design for the CLY15 amplifier.

Finally, the complete circuit using distributed elements is shown in Figure 3-264. From a dc point of view, we took advantage of the incoming shorted stub by using a transmission line of the appropriate length and a bypass capacitor to ground. This allows us to feed the gate with the required negative voltage without interfering with the frequency response of the amplifier. When evaluating the total frequency response (Figure 3-265), we noticed that this circuit seems to have a higher Q than the lumped equivalent. Again, to obtain a wider bandwidth more transformation stages will be needed. We leave this task to the interested reader.

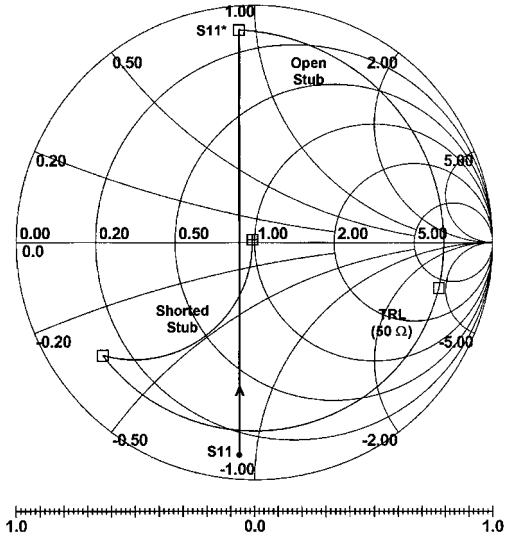


Figure 3-262 Input match for S_{11}^* (1 W) using distributed elements.

The 1-dB bandwidth of the amplifier using distributed, lossy elements is 120 MHz. If a wider bandwidth is required, matching networks with more elements are required. Because of the gain–bandwidth product, the peak gain for such a wider-bandwidth arrangement will be less than this narrowband example.

Having used a lumped to distributed transformation, we next present a power amplifier that uses coaxial lines, baluns, and other transformation steps capable of handling the increased power requirement.

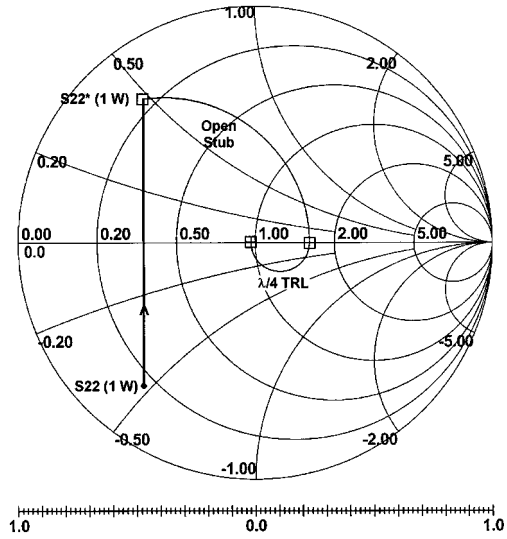


Figure 3-263 Output match for S_{22}^* (1 W) using distributed elements.

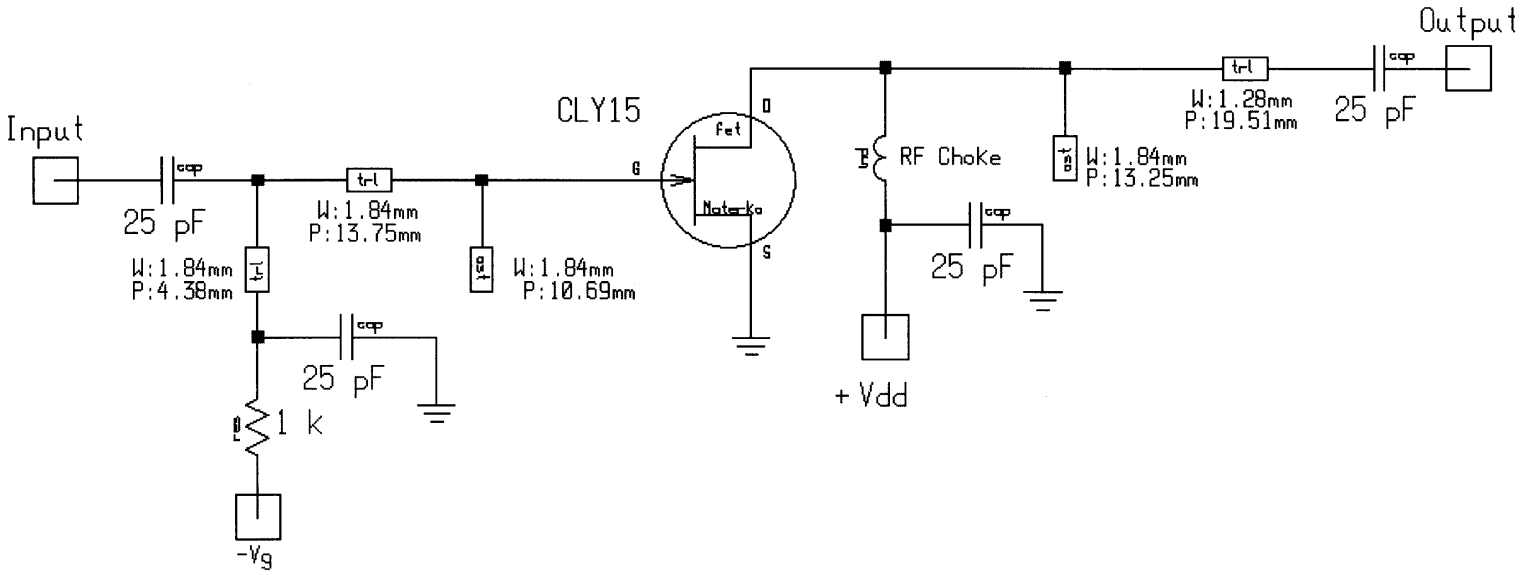


Figure 3-264 The complete amplifier using the CLY15 medium-power GaAsFET, with distributed-element input and output matching.

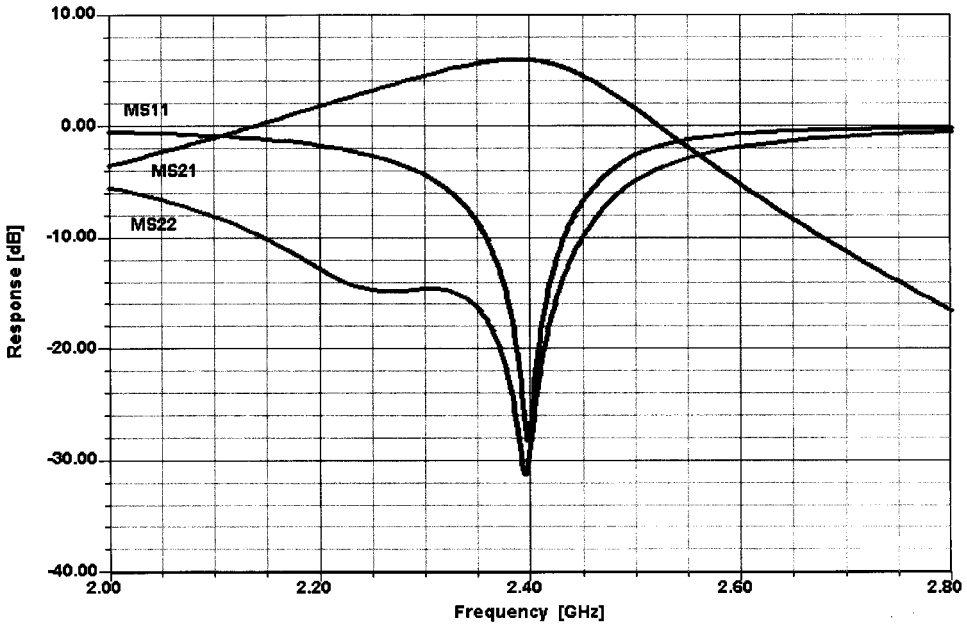


Figure 3-265 Frequency-dependent gain and matching using distributed, lossy elements.

3-12-5 Example 4: 90-W Push–Pull BJT Amplifier at 430 MHz

The last complete application we want to examine is a 90-W output, push–pull amplifier using two TRW transistors (that are no longer made—an all-too-frequent occurrence after a design is complete, resulting in a scramble for replacements). Figure 3-266 shows the circuit diagram of this UHF transistor amplifier, and Figure 3-267 shows its mechanical design. The amplifier requires 28 Vdc and will draw approximately 6 A, requiring 15 W of drive at 438 MHz. This amplifier operates in Class AB, which means it is not limited to amplifying FM and other angle-modulated signals, and can accommodate even single-sideband (SSB) operation.

The drive power coming from the source connected to the input N connector (15 W) is fed through a short piece of coaxial cable to the PC board sitting on top of the heat sink. To change from single ended to push–pull, a 50-Ω coaxial cable is coiled to provide sufficient electrical length to act as a balun that provides two symmetrical paths through the amplifier. The RC section of 10 Ω in parallel with a 220-pF capacitor stabilizes the amplifier at lower frequencies, while at higher frequencies the 10-Ω resistor has no contribution. There are two independent 4:1 transformers, one for each section; they are built using 25-Ω coaxial cable with a set of two binocular ferrite cores to force the transformation. The dc supply is also fed to the transistors at this point. The stabilization circuit consists of a temperature monitor using the *pn* junction of a transistor, side-mounted on the heat sink, and a combination of operational amplifiers and adjustments. This guarantees proper, temperature-independent biasing. The 4:1 input baluns result in 12.5 Ω impedance, or 3.125 Ω per transistor. Since the input impedance at this drive level is somewhere in the vicinity of $[(1 \ldots 2) + j(3 \ldots 5)]$ Ω; the final matching is accomplished by an *RC* section in which the inductance is actually printed

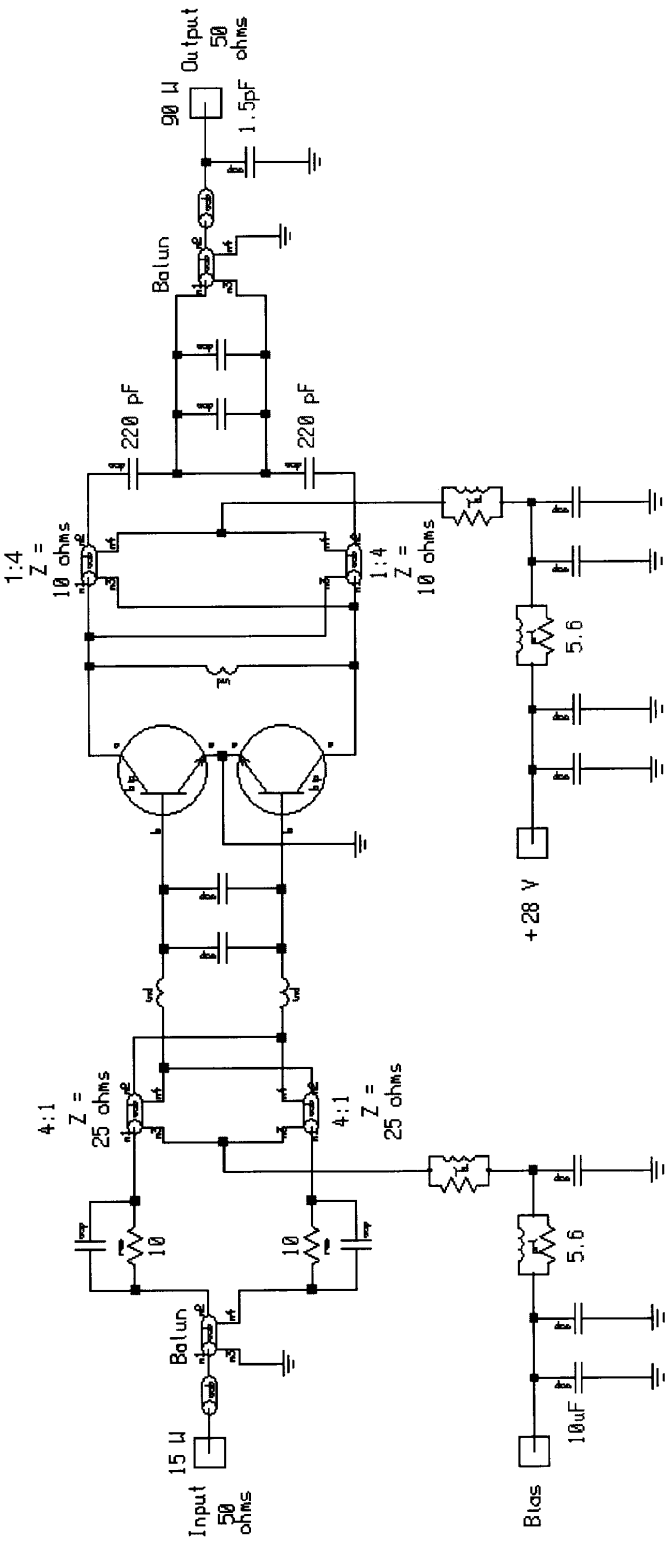


Figure 3-266 Circuit diagram of the 90-W, push–pull, 438-MHz linear amplifier.

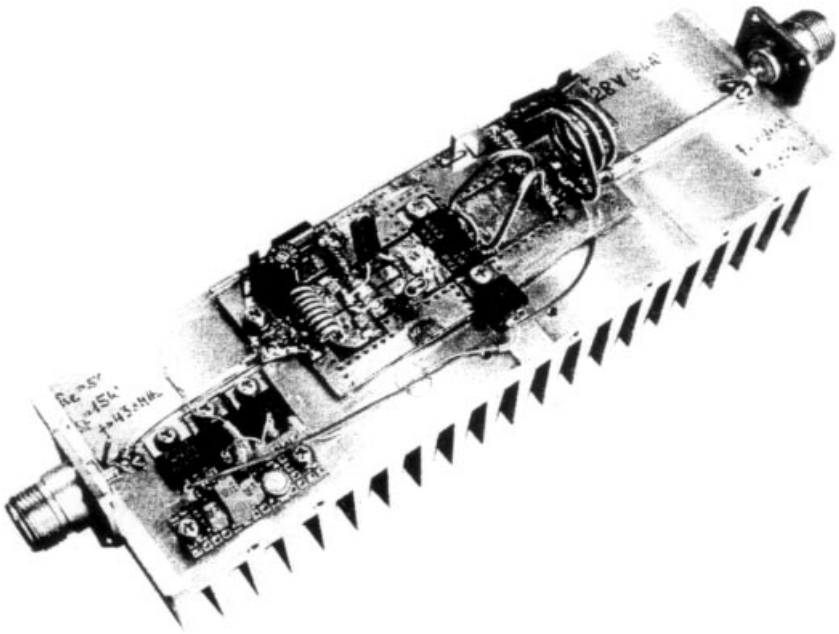


Figure 3-267 Mechanical assembly of the 90-W amplifier shown in Figure 3-266.

and the shunt capacitor across is soldered from base to base as close to the transistor bodies as possible. A ceramic-dielectric variable capacitor in parallel with this allows fine tuning.

The output consists of an inductor from collector to collector that tunes out the capacitive reactance and makes the transistor output impedance real. We then follow the same pattern by taking a 1:4 transformer, but this time built using 10-Ω rigid line instead of 25-Ω rigid line. Together with the output capacitance, prior to the output balun, this is a medium-bandwidth matching arrangement that transforms up to 50 Ω. The collector impedance per collector had to be $(28 - 2)^2/2 \times 90$, resulting in roughly 3.8 Ω—not that different from the input match at the balun point. This way, the matching networks at the input and output can be identical. At the output of the matching network, another coiled 50-Ω coaxial cable serves as balun; because it must operate at the 90-W level, larger-diameter line was used than in the input balun. Finally, another section of rigid line is used to connect the output balun on the PC board to the output N connector. The output capacitor again helps in the transformation.

Both the base and collector voltage supplies are heavily filtered using ferrite-bead-based RF chokes and heavy wire inductors wound on low-impedance resistors. A fairly large number of these amplifiers were actually built and are still in use.

3-12-6 Quasiparallel Transistors for Improved Linearity

Some applications require really low intermodulation distortion. One of the authors (Rohde), during his tenure at RCA Government Systems Division, and his team developed a parallel-device topology that is somewhat similar to the conventional parallel approach but shows vastly increased linearity. In creating higher-power devices, manufacturers parallel several transistors in one package, connecting them in a clever way so the overall drive, dissipation, and output power are equally divided between all the devices. Special nickel–chromium

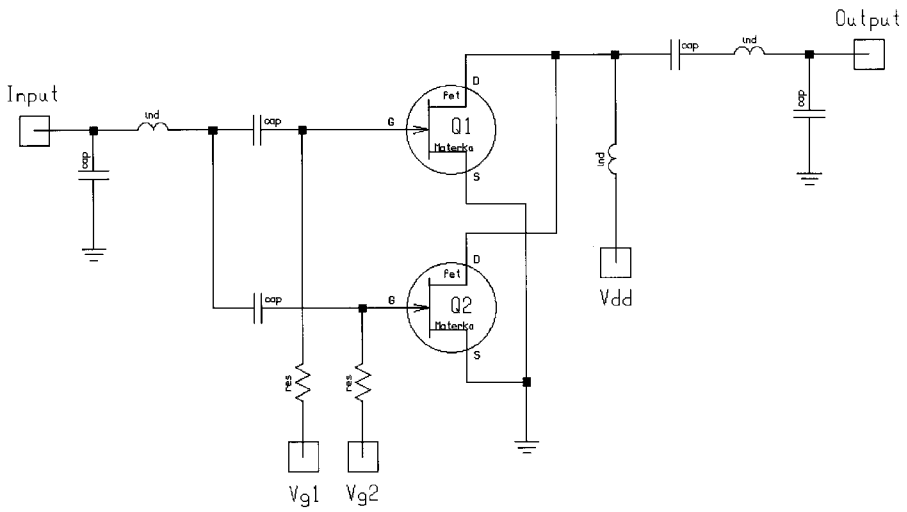


Figure 3-268 The quasiparallel amplifier topology looks like a standard parallel amplifier. Biasing the devices to different operating points nets an improvement in IMD performance.

ballasting resistors in the many fingers of the emitter connections of the multiple internal devices equalize any differences among the transistors while increasing the saturation voltage by only a small amount. In our high-linearity design, we used two transistors in parallel, one being biased somewhere between Class A and B, and the second one somewhere between Class B and C (Figure 3-268). The exact bias voltages and operating points had to be found through experiment. In practice, Q1, the "initially on" transistor, is driven to a point close to that at which its IMD products would increase. Q2 is biased such that at this same drive level it adds to the output power.

We frequently referred to this configuration as "turbocharged." The IMD products at full output power were down 40–45 dB, compared to 26–30 dB for a single transistor. While such an arrangement shows superior IMD performance, because of the somewhat abrupt kick-in of the second transistor, the transconductance response is semismooth, a characteristic that might make alert readers think that IMD will occur at this point. On the other hand, the transconductance for Q1, in effect, saturates or stays constant, so when the transconductance responses of both transistors are overlaid, the result is a smooth response overall.

This technique has been validated up to 100 W and up to at least 50 MHz. Its drawbacks include the need for two transistors with their associated costs, additional real estate, and higher input and output capacitances. Nonetheless, considering that MOS transistors capable of operation to 1 GHz were unavailable when this technique was developed, we believe that it should be revisited.

The Class D and higher operating modes require a large number of additional components and will make the efficiency higher (selective in frequency), but not necessarily reduce the IMD products. A further point of consideration is that it is incorrect to assume that reducing the input level to a given Class A will necessarily reduce its IMD products. As with tubes, the closer one gets to the knee, even in Class A operation, the worse the IMD. (This phenomenon was particularly troublesome in systems using 1-kW amplifiers up to 150 MHz;

underdrive with the intent of getting better IMD performance produced just the opposite result.) If a wide power range is considered, we recommend using a bias scheme that includes sufficient intelligence to find the best operating point.

Another hot amplifier issue is that of IMD as a function of load impedance. If an amplifier is used at output power considerably below its design value, its load impedance will be too low and, depending on the circuit configuration, the result may be additional IMD.

3-12-7 Distribution Amplifiers

It is not uncommon that a signal from a particular source like an antenna has to be distributed to several places. A condominium is a typical situation in which energy from a main signal source must be split for distribution to a larger number of users—there must be enough voltage or power to produce a noise-free picture. An amplifer capable of doing this is a *distribution* amplifier, not to be confused with a *distributed* (or traveling-wave) amplifier, such as that shown earlier in this chapter in Figure 3-6. An interesting means of achieving this is to use multiple Wilkinson couplers following an amplifier. Figure 3-269 shows an arrangement that makes heavy use of multiple power dividers to provide multiple outputs. Theoretically, the reverse is also possible: If we want to collect energy from a number of sources and combine them, as on a cable, we can use Wilkinson couplers at the *input* of a gain block. Note, however, that depending on the level and number of the combined signals, this may put enormous strain on the linearity of the amplifier. What we want to avoid is a lot of distortion, which, in the case of television, will show up as color confetti in the noise and signal.

3-12-8 Stability Analysis of a Power Amplifier

One of the hidden pitfalls in narrowband circuit design is that a circuit may be unstable outside the frequency range of interest. Circuits containing structures built from distributed elements, which exhibit multiple resonances, can be particularly susceptible to such instabilities. In this example, we analyze the stability of a simple FET power amplifier designed using small-signal concepts and the interactive Smith Tool utility in Ansoft's Serenade

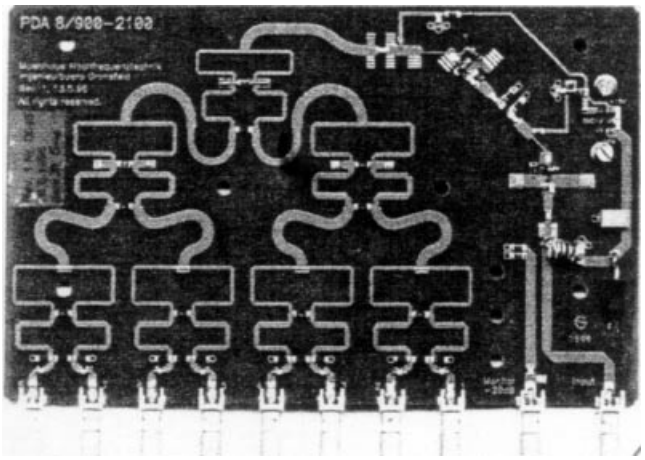


Figure 3-269 Distribution amplifier showing the use of multiple Wilkinson power dividers.

Design Environment. The evaluation shows how K-factor analysis alone is not sufficient to check for stability of a design.

Small-Signal ac Analysis. We begin by taking the Figure 3-270 amplifier, which was designed for about 16 dB of gain at 2.2 GHz. Due to the very low, capacitive input impedance of the FET and the proximity of the source-plane gain circles and stability circles to the edge of the Smith Chart, the matching network attempts to transform 50 Ω to an impedance of approximately $(2.5 + j15)$ Ω. The impedance lies close to the stability circles at 2 GHz but remains in a stable region. The bias point selected (–3.5 V) is chosen for Class AB operation. Higher linear gain can be achieved by reducing the gate voltage to –2.0 V; at this bias point, the amplifier gain will increase to about 19 dB.

First, we will do a small-signal ac analysis and examine the S parameters and K and B_1 factors [see Eqs. (3-154) and (3-157), respectively] from 1 GHz to 6 GHz—well beyond the bandwidth of the design. Figure 3-271 shows the circuit's frequency-dependent gain and matching characteristics.

If we examine B_1 and the stability factor K, we see that the amplifier is not unconditionally stable in the 2.2-GHz passband as shown in Figure 3-272. K dips slightly below 1, although B_1 remains above zero. In practice, a slight change in design would be required to bring K above 1 and remove this potential instability. However, for our example, this is of no consequence, as we shall see below.

Nyquist Stability Analysis. Now we will utilize a Nyquist stability analysis to check the stability of the amplifier. To fully realize the advantages of a Nyquist stability analysis, we must sweep frequency over a very wide range. We want to be able to pick up instabilities at *any* frequency, not just near the design frequency. In this example we sweep from 1 kHz to 20 GHz. A start frequency of 1 kHz is chosen so we can determine the starting point of the Nyquist plot and we ensure that any low-frequency resonances are picked up. The stop frequency is chosen past the f_{max} of the transistor (where we know it cannot oscillate).

The analysis results are shown in polar plot form in Figure 3-273 and in magnitude–angle form in Figure 3-274. The polar graph shows the real and imaginary parts of the system determinant as frequency is swept. The magnitude–angle graph shows the magnitude and cumulative angle (no cut at ±180°) of the system determinant as frequency is swept. The key to interpreting these graphs lies in the following statement drawn from the Nyquist criterion:

The circuit will be unstable if the Nyquist plot encircles the origin in a clockwise direction.

The term "unstable" refers to the existence of natural frequencies lying in the right-half of the complex frequency plane (RHP). These natural frequencies on the RHP will cause oscillations. To encircle the origin in a clockwise direction means that the Nyquist plot will travel in a clockwise path, cross the negative real axis, and continue to completely surround the origin. If the path appears to encircle the origin but then "unravels" itself in a counter-clockwise direction, the origin is not actually encircled. Figure 3-273 shows that the origin is not encircled and therefore this circuit is stable.

The magnitude–angle plot can assist in determining if a clockwise encirclement has actually been made with the help of the following:

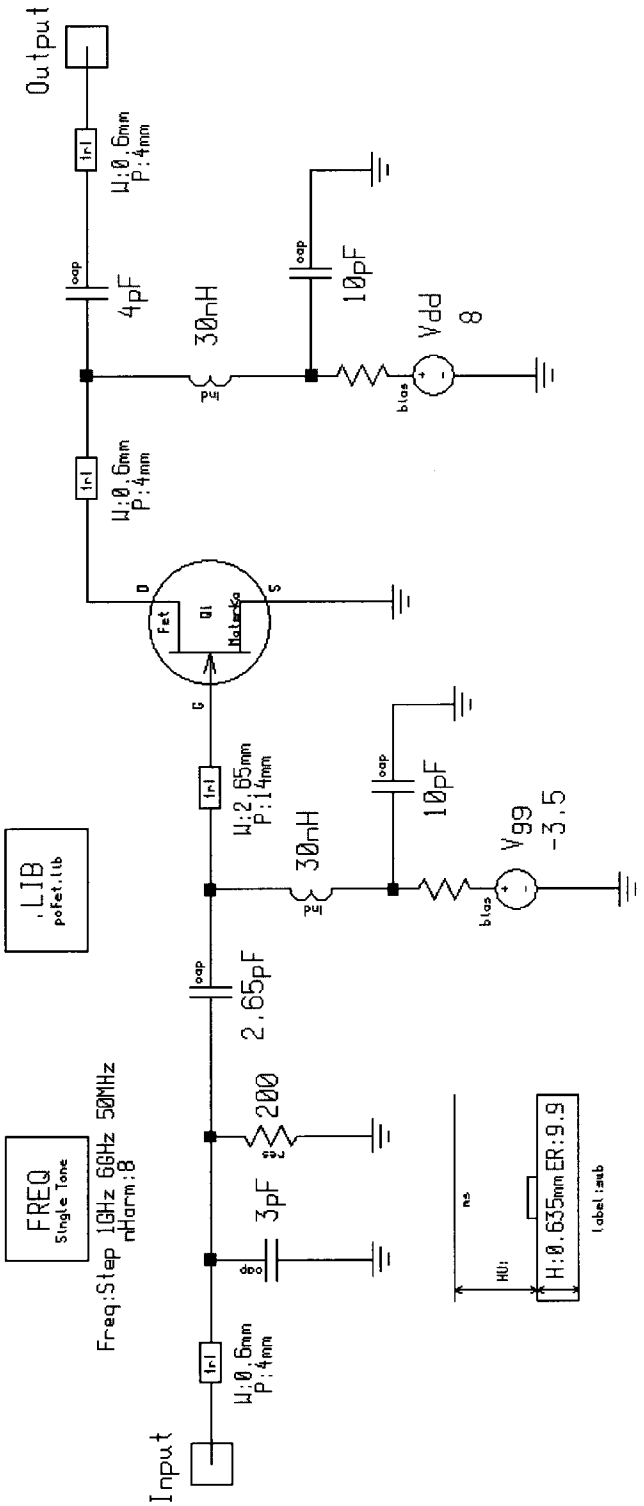


Figure 3-270 A 2.2-GHz amplifier design for stability analysis.

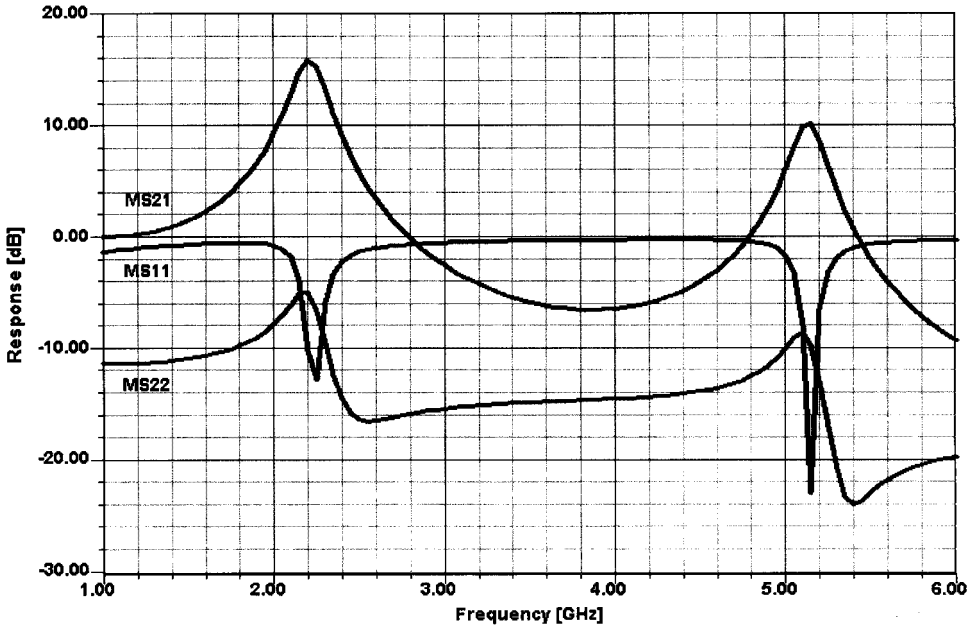


Figure 3-271 Frequency-dependent gain and matching characteristics of the amplifier. Although the circuit is designed for 2.2-GHz operation, the periodic responses of its distributed-element-based input and output matching networks result in a second gain peak near 5.1 GHz. As we shall see, however, instabilities may result at frequencies that cannot be intuited from this graph.

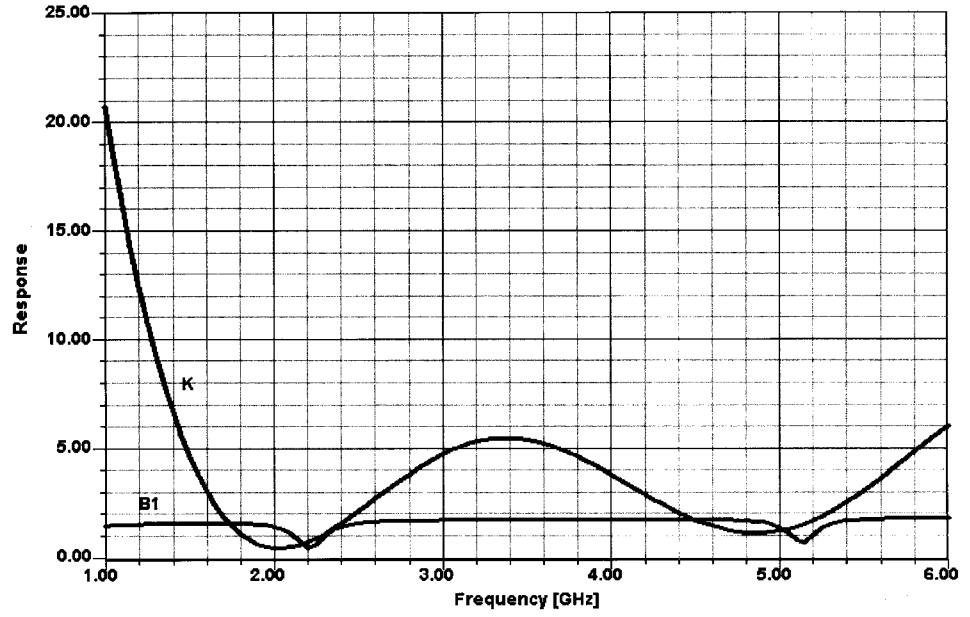


Figure 3-272 K and B_1 for the 2.2-GHz amplifier.

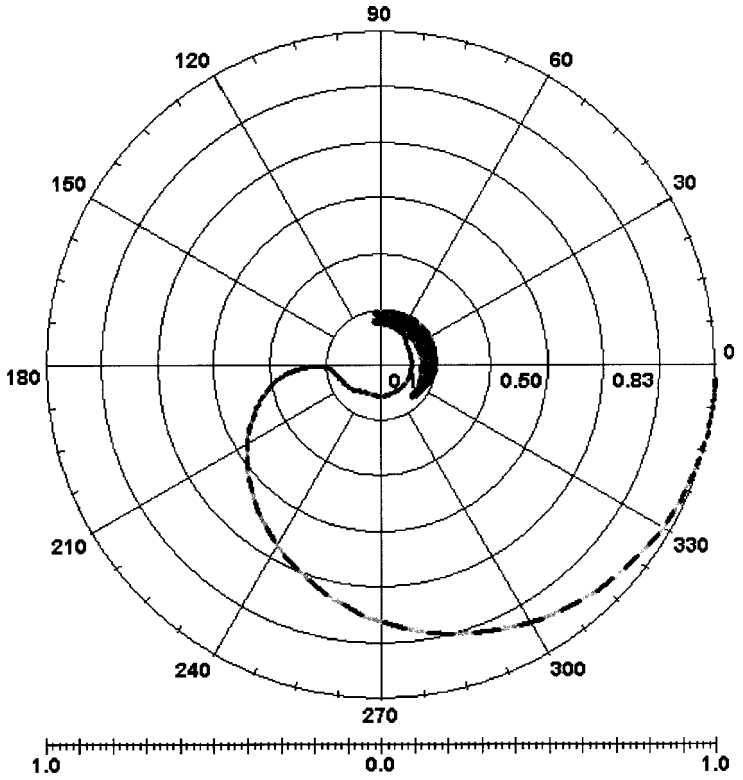


Figure 3-273 Polar Nyquist plot for the 2.2-GHz amplifier covering 1 kHz to 20 GHz.

- To travel in a clockwise direction on a polar chart means to decrease phase angle. On a magnitude–angle plot, the phase will become more negative as frequency is increased (although it may not be monotonic).
- If the path crosses the negative real axis on the polar chart, it will cross –180° on the magnitude–angle plot.
- Encirclement will occur if the phase angle continues to decrease to, say, –360°. If "unraveling" occurs, the phase angle will not stay below –360°. Figure 3-274 clearly shows that the phase angle decreases to about –180° but does not decrease any more. Rather, the angle returns to a positive value and oscillates between about –40° and +100°. Therefore an encirclement did not occur, and the circuit is stable.

An Unstable Case. Let us now change the gate bias to –2 V—a value that increases the amplifier gain enough to cause instability. Figures 3-275 and 3-276 show the results. From Figure 3-275 we can see that the origin is encircled in a clockwise direction. Figure 3-276 confirms this since the phase angle goes through –180° and continues without unraveling.

Approximate Frequency of Oscillation. The frequency where the Nyquist plot crosses –180° is often the approximate frequency of the unstable right-half-plane zeros and provides an estimate of the potential oscillation. For this example, there is a sharp crossing at 200 MHz that is characteristic of a circuit resonance. Note that the frequency is not related

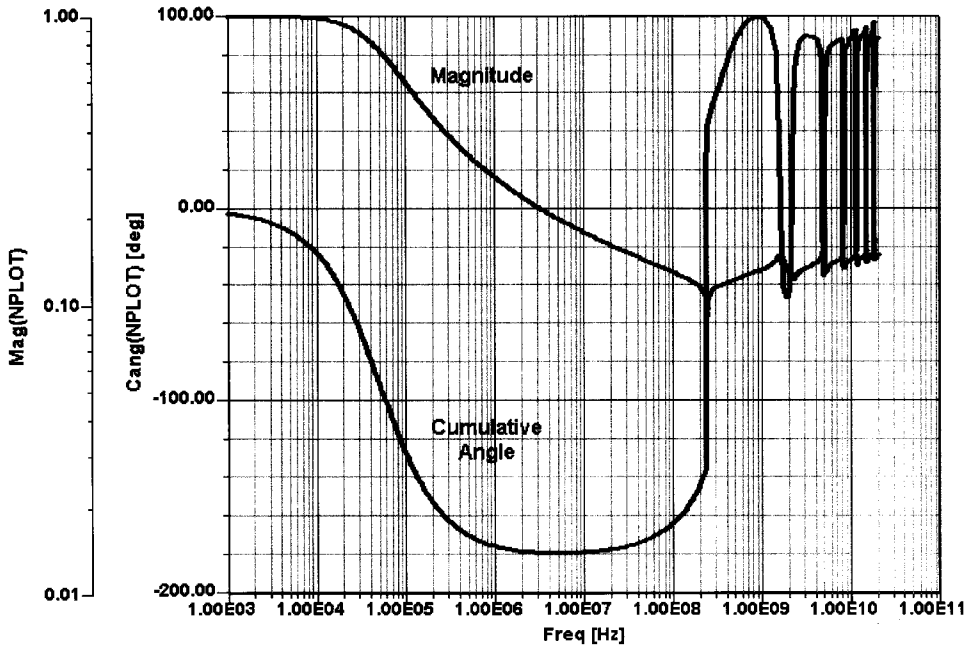


Figure 3-274 Magnitude and cumulative angle Nyquist plot for the 2.2-GHz amplifier covering 1 kHz to 20 GHz.

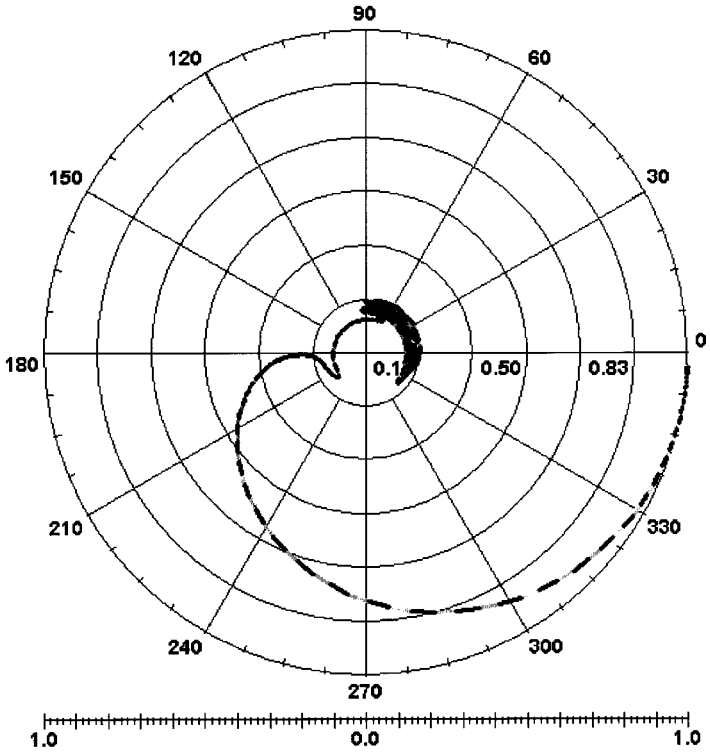


Figure 3-275 Polar Nyquist plot of the unstable amplifier.

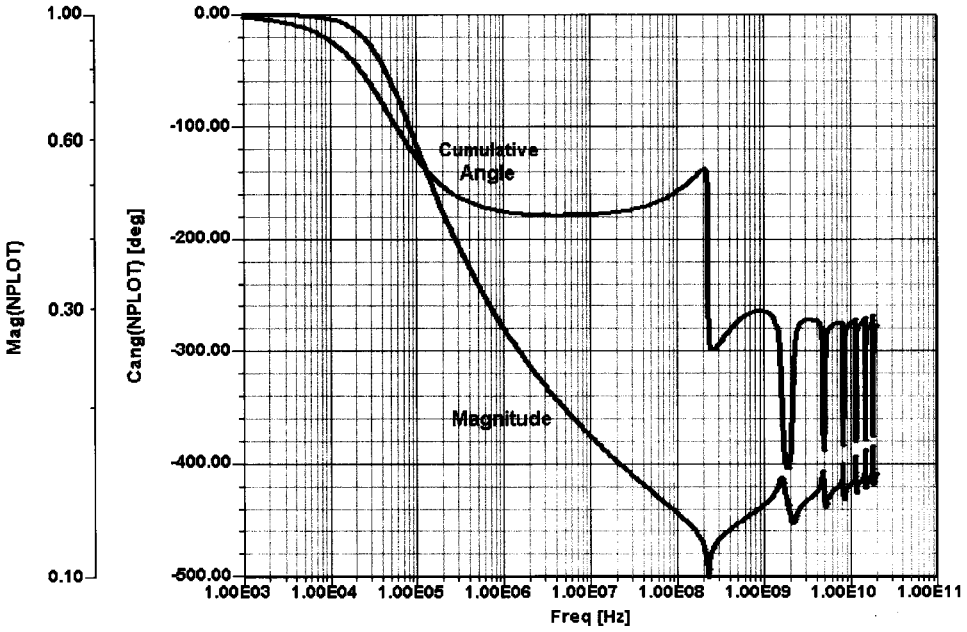


Figure 3-276 Magnitude–cumulative angle Nyquist plot of the unstable amplifier.

to a resonant frequency in a rigorous mathematical fashion, but it has been our experience that this frequency approximates the resonant frequency. To confirm the circuit's readiness for oscillation in the vicinity of this frequency, we evaluate it with Serenade's Oscillator Design Aid from 10 MHz to 1 GHz. The Design Aid finds an initial resonant frequency near 230.6 MHz—not too far from the –180° crossing frequency of 200 MHz of the Nyquist plot. Figure 3-277 shows this point.

Oscillator analysis finds a final oscillation frequency of 228.48 MHz. Figure 3-278 shows the resulting output spectrum at Port 2. At this point we decided to do a quick jump into Chapter 5 and take advantage of the CAD capability to predict the resulting phase noise, which is shown in Figure 3-279. As we compare this with other measured results, the approximately 228-MHz oscillator has a poor phase noise since the same phase noise is achievable at 1 GHz and higher. We remind the reader that this is an unwanted effect, but all of us who have built amplifiers will have experienced either low-frequency oscillations (motorboating) or very-high-frequency oscillation, sometimes caused by the components in the immediate vicinity of the transistor itself and found somewhere above 3 GHz. This case is really a demonstration case; a closer look reveals that the transconductance required for this condition is unrealistically high; but nonetheless it is a good example to demonstrate the stability issue.

Where Does Oscillation Begin? An interesting question to solve now is, "What gate bias voltage is needed for oscillation to begin?" Ansoft's Serenade Design Environment includes oscillator synchronous stability analysis, which uses a modification of the harmonic-balance technique to determine critical points, such as the voltage needed to "turn on" an oscillator. Mathematically, this is called a *Hopf bifurcation* because the behavior of

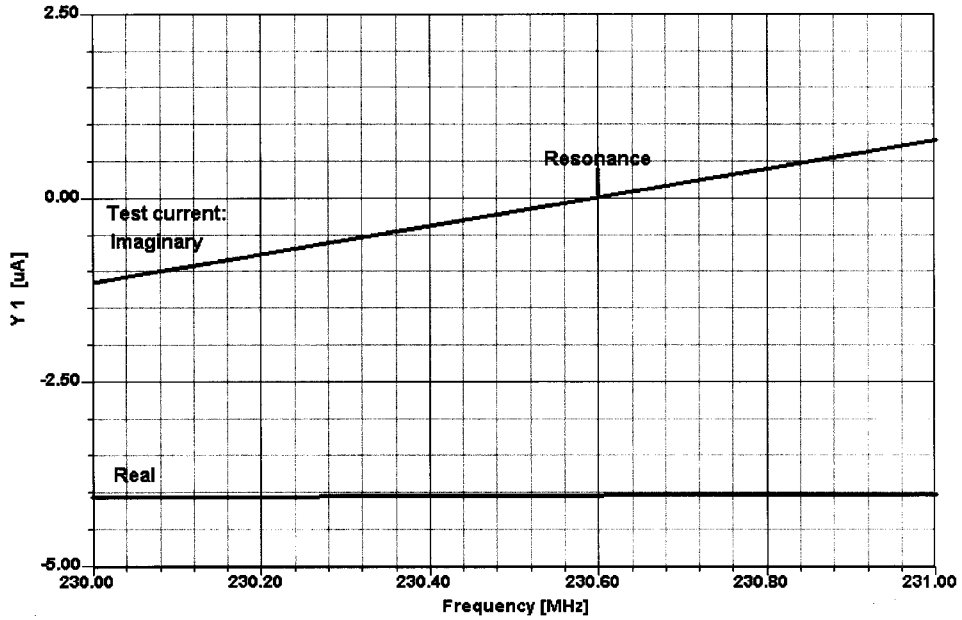


Figure 3-277 Graph from Serenade Oscillator Design Aid analysis of the amplifier. A resonant frequency is indicated where the imaginary part of the test current equals zero and the real part is negative.

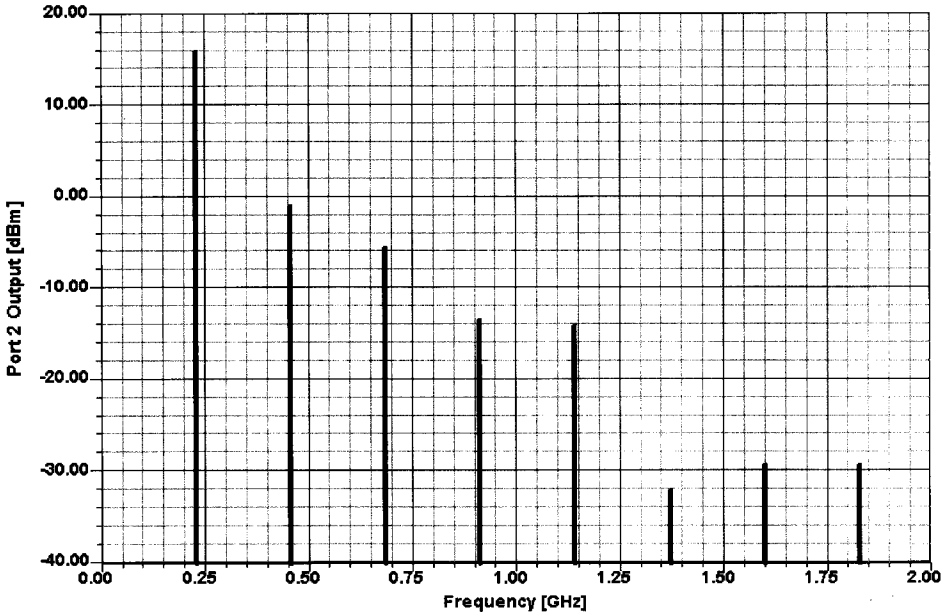


Figure 3-278 Predicted output spectrum of the unintended oscillator.

the circuit changes dramatically from a dc steady state to an oscillatory response. Figure 3-280 shows this effect by plotting output power versus gate bias. The turning point at –3.2 V (22 mW) is clearly visible, as is the bifurcation at –3.1 V (0 mW output).

Figure 3-280 can be interpreted as follows:

1. If the bias voltage is increased from, say, –4.0 V, the circuit will begin to oscillate at the bifurcation voltage and the output power will jump to 26 mW.
2. If the circuit is oscillating and the voltage is decreased from –2.0 V, the circuit will cease oscillating at the turning voltage of –3.2 V.
3. The branch between the turning point and the bifurcation point is unstable and cannot be realized in practice. This forms a hysteresis loop around points (–3.1 V, 0 mW), (–3.1 V, 26 mW), (–3.2 V, 22 mW), and (–3.2 V, 0 mW).

This amplifier stability analysis section is based on the stability analysis examples provided by Ansoft for the Serenade Design Environment.

An interesting example that combines both CAD and the feed-forward amplifier technique can be found in the Serenade Design Environment system simulator examples manual. This technique has been under evaluation for the last two years, but our feeling is that not enough measured data at frequencies above 500 MHz with reliable performance information have yet been made available in the literature. Along these lines, work done at Ansoft under a U.S. Air Force contract has provided very good insight into high-efficiency amplifiers at microwave frequencies [46].

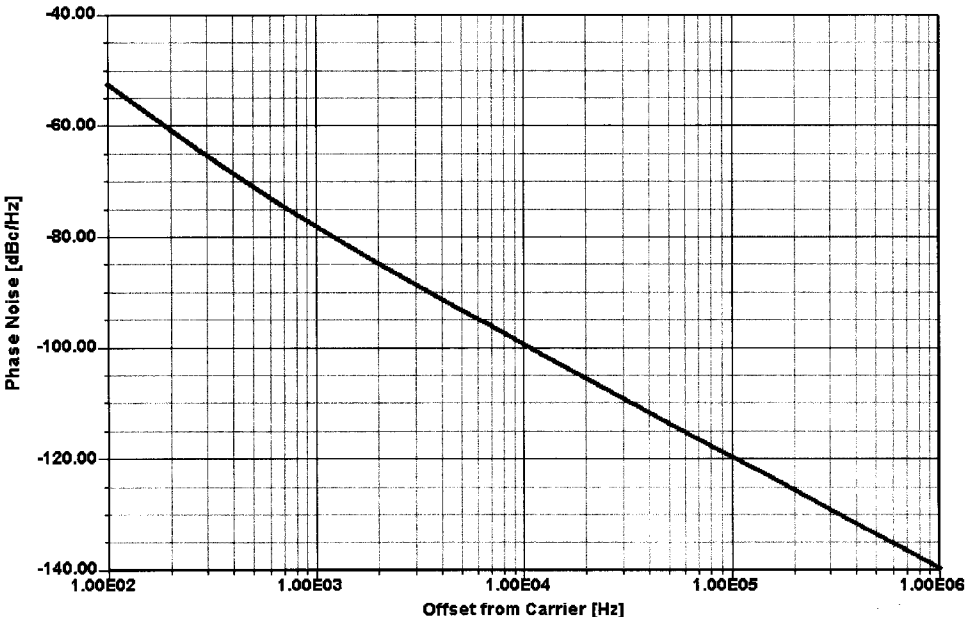


Figure 3-279 Predicted phase noise of the oscillating amplifier.

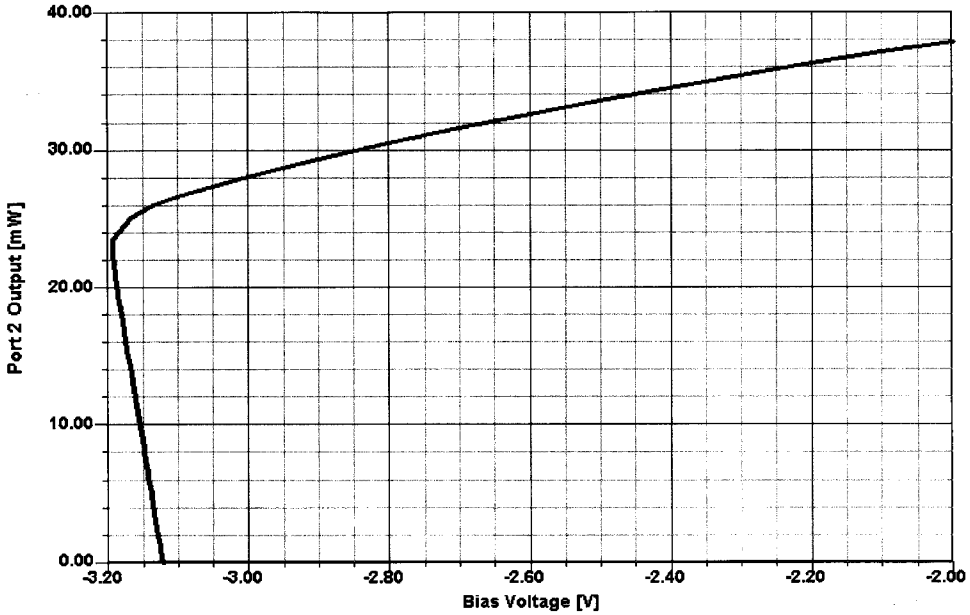


Figure 3-280 Output power versus gate bias for the unstable amplifier.

3-13 POWER AMPLIFIER DATASHEETS AND MANUFACTURER-RECOMMENDED APPLICATIONS

Having discussed the various aspects of BJTs and FETs, we have decided to use MOS examples from Ericsson and Motorola, and a bipolar example from Motorola, as good examples to show the technology required to build high-power transistor amplifiers in this frequency range. In the following pages, we reproduce, with permission, datasheets for Ericsson's PTF 10009 FET and Motorola's MRF186 FET, followed by Motorola's MRF899 BJT.

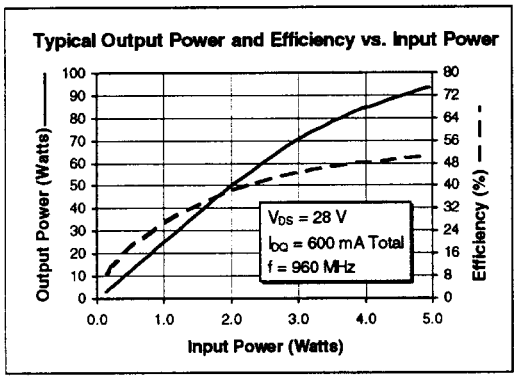
ERICSSON

PTF 10009
85 Watts, 1.0 GHz
LDMOS Field Effect Transistor

Description

The 10009 is a common source N-channel enhancement-mode lateral MOSFET intended for large signal amplifier applications to 1.0 GHz. It is rated at 85 watts minimum output power. Nitride surface passivation and gold metallization are used to ensure excellent device lifetime and reliability. 100% lot traceability is standard.

- Performance at 960 MHz, 28 Volts
  - Output Power = 85 Watts
  - Power Gain = 13.0 dB Typ
  - Efficiency = 50% Typ
- Gold Metallized
- Silicon Nitride Passivated
- Excellent Thermal Stability



Package 20230

Maximum Ratings

Parameter	Symbol	Value	Unit
Drain-Source Voltage	V_{DSS}	65	Vdc
Gate-Source Voltage	V_{GS}	±20	Vdc
Operating Junction Temperature	T_J	200	°C
Total Device Dissipation at T_{flange} = 25°C Above 25°C derate by	P_D	270 1.54	Watts W/°C
Storage Temperature Range	T_{STG}	-65 to 150	°C
Thermal Resistance (T_{flange} = 70°C)	$R_{\theta JC}$	0.65	°C/W

PTF 10009

ERICSSON

Electrical Characteristics (100% Tested)

Characteristic (per side)	Conditions	Symbol	Min	Typ	Max	Units
Drain-Source Breakdown Voltage	V_{GS} = 0 V, I_D = 25 mA	$V_{(BR)DSS}$	65	—	—	Volts
Drain-Source Leakage Current	V_{DS} = 28 V, V_{GS} = 0 V	I_{DSS}	—	—	1.0	mA
Gate Threshold Voltage	V_{DS} = 10 V, I_D = 75 mA	$V_{GS(th)}$	—	2.0	—	Volts
Forward Transconductance	V_{DS} = 10 V, I_D = 3 A	g_{fs}	—	2.8	—	Siemens

Dynamic Characteristics

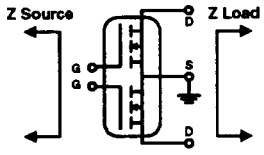
Characteristic (per side)	Symbol	Min	Typ	Max	Units
Input Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{iss}	—	90	—	pF
Output Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{oss}	—	36	—	pF
Reverse Transfer Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{rss}	—	1.9	—	pF

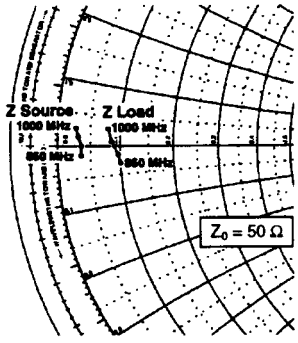
RF Specifications (100% Tested)

Characteristic	Symbol	Min	Typ	Max	Units
Gain (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz)	G_{ps}	12.0	13.0	—	dB
Drain Efficiency (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz)	η	47	50	—	%
Load Mismatch Tolerance (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz— all phase angles at frequency of test)	Ψ	—	—	5:1	—

Impedance Data (data shown for fixed-tuned broadband circuit)

(V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA)



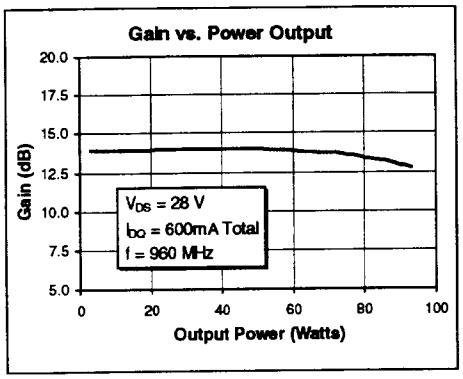


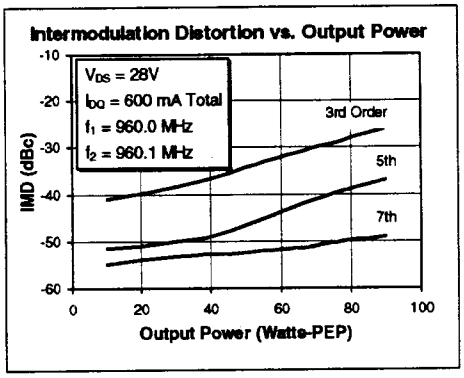
Frequency	Z Source		Z Load	
MHz	R	jX	R	jX
860	1.76	-0.78	5.00	-1.50
900	1.80	-0.05	4.80	-0.78
960	1.58	0.69	4.24	0.36
1000	1.39	1.35	3.95	1.41

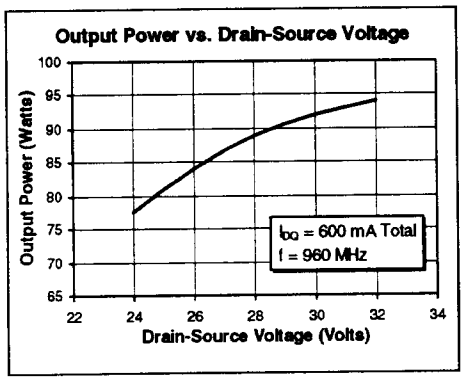
ERICSSON

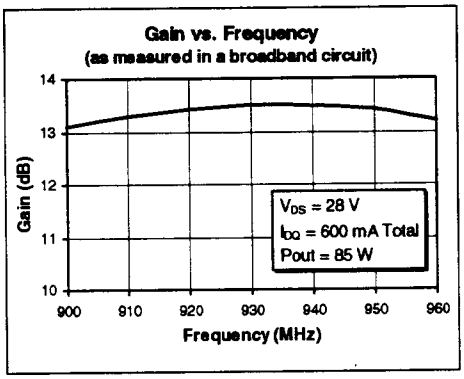
PTF 10009

Typical Performance





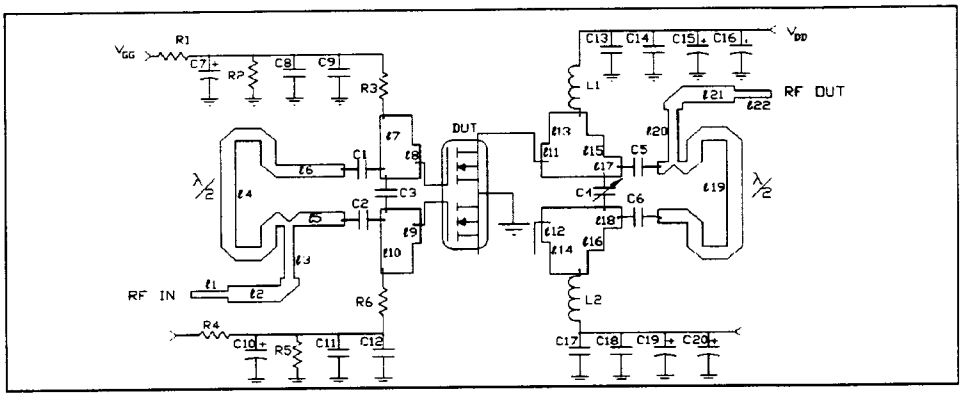




PTF 10009

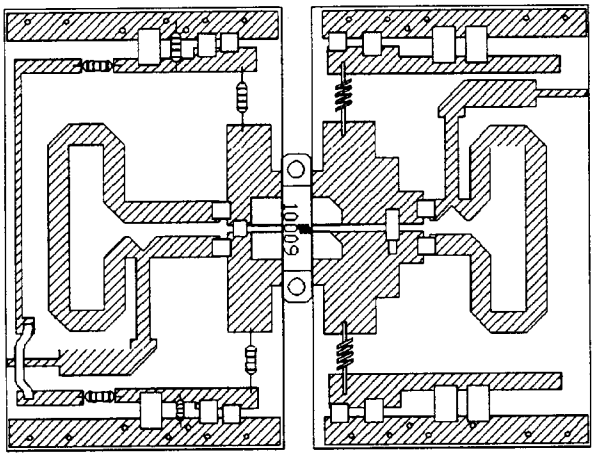
ERICSSON

Test Circuit



Schematic for f = 960 MHz

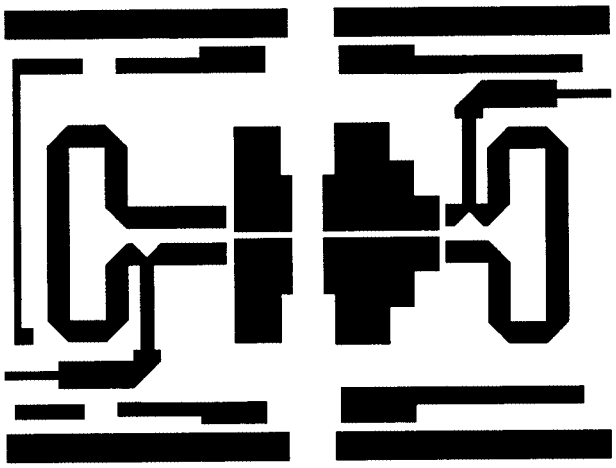
DUT	10009
C1-2, C5-6, C9, C12-13, C17	33 pF, Capacitor ATC 100 B
C3	11 pF, Capacitor ATC 100 B
C4	6.0 pF, Variable Capacitor, JMC 5701
C7, C10	10 µF, +10 V Electrolytic Capacitor
C8, C11, C14, C18	0.01 µF, Capacitor ATC 100 B
C15, C16, C19, C20	10 µF, +30 V Electrolytic Capacitor
L1. L2	4 Turn, #20 AWG, .120" I.D.
R1, R2, R4, R5	1.0 K, Ω Resistor
R3, R6	5.1 K, 1/4 Ω Resistor
ℓ1, ℓ22	50 Ω, .030 λ
ℓ2, ℓ21	20 Ω, .080 λ
ℓ3, ℓ20	32 Ω, .191 λ
ℓ4, ℓ19	25 Ω, .500 λ
ℓ5, ℓ6	25 Ω, .091 λ
ℓ7, ℓ10	7 Ω, .056 λ
ℓ8, ℓ9	13.0 Ω, .017 λ
ℓ11, ℓ12	13.0 Ω, .017 λ
ℓ13, ℓ14	7.0 Ω, .064 λ
ℓ15, ℓ16	10.0 Ω, .029 λ
ℓ17, ℓ18	19.0 Ω, .028 λ
Circuit Board	.031" Thick, ε_r = 4.0, AlliedSignal, G200



Parts Layout (not to scale)

ERICSSON

PTF 10009



Artwork (1 inch ——— *)*

Ericsson Components
RF Power Products
675 Jarvis Drive
Morgan Hill, CA 95037 USA
Telephone: 408-778-9434

1-877-GOLDMOS
(1-877-465-3667)
e-mail: rfpower@ericsson.com
www.ericsson.com/rfpower

Specifications subject to change
without notice.
L1
© 1997 Ericsson Inc.
EUS/KR 1301-PTF 10009 Uen Rev. B 09-29-98

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF186/D

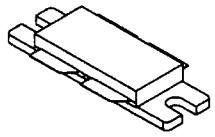
The RF MOSFET Line
RF Power Field-Effect Transistor
N–Channel Enhancement–Mode Lateral MOSFET

MRF186

120 W, 1.0 GHz
28 VOLTS
LATERAL N–CHANNEL
BROADBAND
RF POWER MOSFET

Designed for broadband commercial and industrial applications at frequencies from 800 MHz to 1.0 GHz. The high gain and broadband performance of this device makes it ideal for large–signal, common source amplifier applications in 28 volt base station equipment.

- Guaranteed Performance @ 960 MHz, 28 Volts
  Output Power — 120 Watts (PEP)
  Power Gain — 11 dB
  Efficiency — 30%
  Intermodulation Distortion — –28 dBc
- Excellent Thermal Stability
- 100% Tested for Load Mismatch Stress at all Phase Angles with 5:1 VSWR @ 28 Vdc, 960 MHz, 120 Watts CW



CASE 375B–02, STYLE 2

MAXIMUM RATINGS (2)

Rating	Symbol	Value	Unit
Drain–Source Voltage	V_{DSS}	65	Vdc
Drain–Gate Voltage (R_{GS} = 1 MΩ)	V_{DGR}	65	Vdc
Gate–Source Voltage	V_{GS}	±20	Vdc
Drain Current — Continuous	I_D	14	Adc
Total Device Dissipation @ T_C = 70°C Derate above 70°C	P_D	162.5 1.25	Watts W/°C
Storage Temperature Range	T_{stg}	– 65 to +150	°C
Operating Junction Temperature	T_J	200	°C

THERMAL CHARACTERISTICS (2)

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	0.8	°C/W

NOTE – **CAUTION** – MOS devices are susceptible to damage from electrostatic charge. Reasonable precautions in handling and packaging MOS devices should be observed.

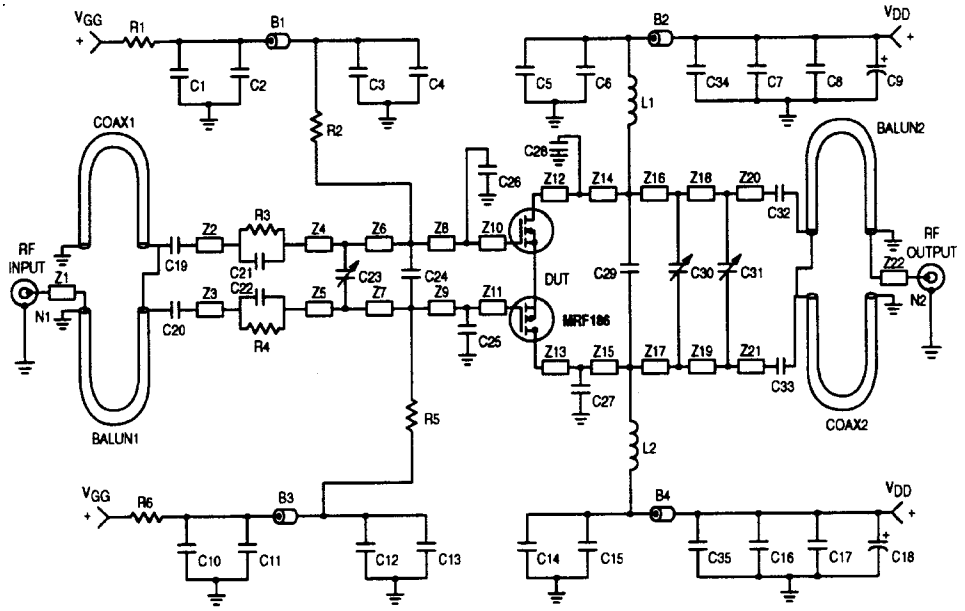
© Motorola, Inc. 1999

MOTOROLA

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted)

Characteristic	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS (1)					
Drain–Source Breakdown Voltage (V_{GS} = 0 Vdc, I_D = 50 μAdc)	$V_{(BR)DSS}$	65	—	—	Vdc
Zero Gate Voltage Drain Current (V_{DS} = 28 Vdc, V_{GS} = 0)	I_{DSS}	—	—	1	μAdc
Gate–Source Leakage Current (V_{GS} = 20 Vdc, V_{DS} = 0)	I_{GSS}	—	—	1	μAdc
ON CHARACTERISTICS (1)					
Gate Quiescent Voltage (V_{DS} = 26 Vdc, I_D = 300 μAdc Per Side)	$V_{GS(th)}$	2.5	3	4	Vdc
Gate Quiescent Voltage (V_{DS} = 26 Vdc, I_D = 300 mAdc Per Side)	$V_{GS(Q)}$	3.3	4.2	5	Vdc
Delta Gate Threshold Voltage (Side to Side) (V_{DS} = 28 V, I_D = 300 mA Per Side)	$\Delta V_{GS(Q)}$	—	—	0.3	Vdc
Drain–Source On–Voltage (V_{GS} = 10 Vdc, I_D = 3 Adc Per Side)	$V_{DS(on)}$	—	0.58	0.7	Vdc
Forward Transconductance (V_{DS} = 10 Vdc, I_D = 3 Adc Per Side)	g_{fs}	2.4	2.8	—	S
DYNAMIC CHARACTERISTICS (1)					
Input Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{iss}	—	177	—	pF
Output Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{oss}	—	45	—	pF
Reverse Transfer Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{rss}	—	3.4	—	pF
FUNCTIONAL CHARACTERISTICS (In Motorola Test Fixture) (2)					
Two–Tone Common Source Amplifier Power Gain (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	G_{ps}	11	12.2	—	dB
Two–Tone Drain Efficiency (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	η	30	35	—	%
3rd Order Intermodulation Distortion (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	IMD	—	–32	–28	dBc
Input Return Loss (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	IRL	9	16	—	dB
Two–Tone Common Source Amplifier Power Gain (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	G_{ps}	—	12	—	dB
Two–Tone Drain Efficiency (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	η	—	33	—	%
3rd Order Intermodulation Distortion (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	IMD	—	–32	—	dBc
Input Return Loss (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	IRL	—	16	—	dB
Output Mismatch Stress (V_{DD} = 28 Vdc, P_{out} = 120 W CW, I_{DQ} = 2 x 400 mA, f = 960 MHz, VSWR = 5:1, All Phase Angles at Frequency of Tests)	Ψ	No Degradation In Output Power Before and After Test			

(1) Each side of device measured separately.
(2) Device measured in push–pull configuration.



B1 – B4	Fair Rite Products Short Ferrit Bead, 2743021446	C31	0.8 – 8.0 pF, Variable Capacitor, Johanson Gigatrim
C1, C7, C8, C10, C16, C17	10 µF, 50 V, Tantalum	L1, L2	3 Turns, #20 AWG, IDIA 0.126", 24.7 nH
C2, C11, C34, C35	0.1 µF, Chip Capacitor	N1, N2	Type N Connectors
C3, C6, C12, C15	330 pF, Chip Capacitor	R1, R6	1 kΩ, 1/4 W, Carbon Resistor
C4, C5, C13, C14, C19, C20, C32, C33	47 pF, Chip Capacitor	R2, R5	1.2 kΩ, 0.1 W, Chip Resistor
C9, C18	250 µF, 50 V, Electrolytic Capacitor	R3, R4	75 Ω, 0.1 W, Chip Resistor
C21, C22	12 pF, Chip Capacitor	Z1 – Z22	Microstrip (See Component Placement)
C23, C30	0.6 – 4.5 pF, Variable Capacitor, Johanson Gigatrim	Balun1, Balun2, Coax1, Coax2	2.20" 50 Ω, 0.086" OD Semi–Rigid Coax
C24, C25, C26	5.1 pF, Chip Capacitor	Board	1/32" Glass Teflon®, ε_r = 2.55
C27, C28	3.9 pF, Chip Capacitor		

Figure 1. 930 – 960 MHz Test Circuit Schematic

TYPICAL CHARACTERISTICS

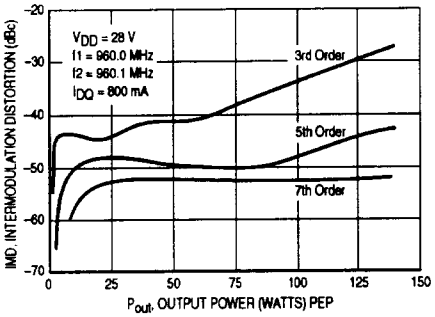


Figure 2. Intermodulation Distortion versus Output Power

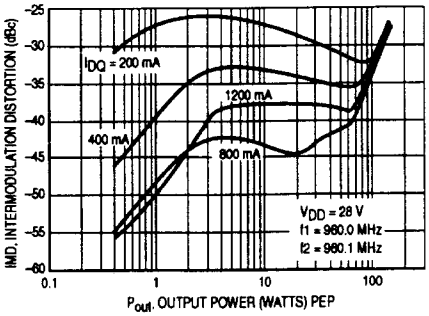


Figure 3. Intermodulation Distortion versus Output Power

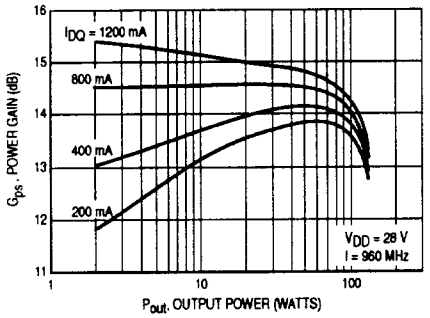


Figure 4. Power Gain versus Output Power

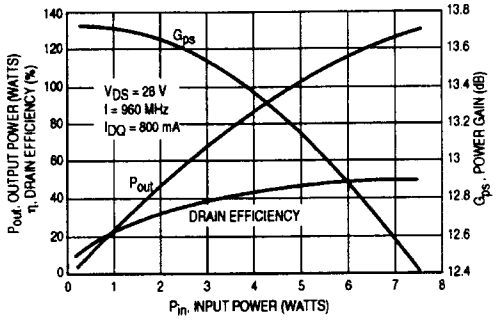


Figure 5. Output Power versus Input Power

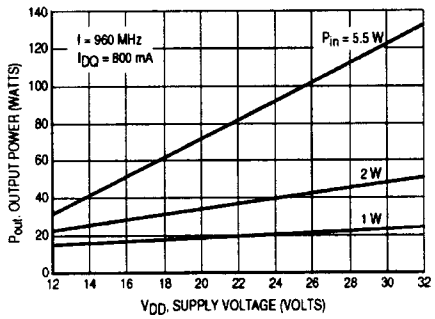


Figure 6. Output Power versus Supply Voltage

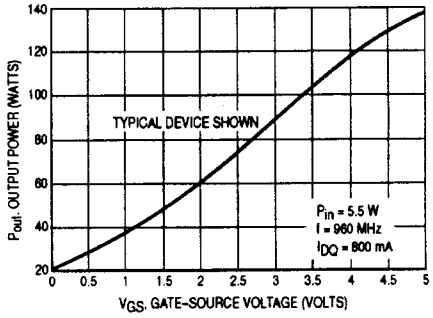


Figure 7. Output Power versus Gate Voltage

TYPICAL CHARACTERISTICS

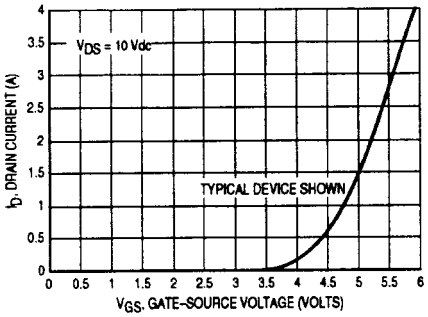


Figure 8. Drain Current versus Gate Voltage

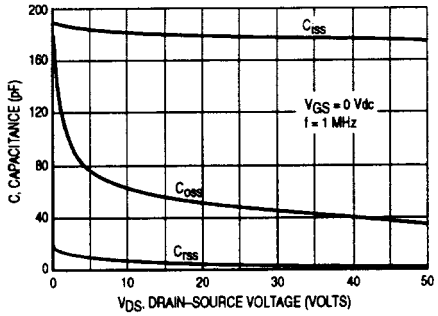


Figure 9. Capacitance versus Voltage

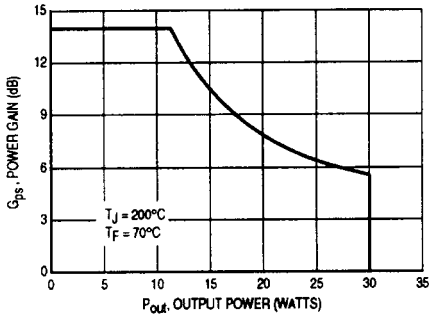


Figure 10. DC Safe Operating Area

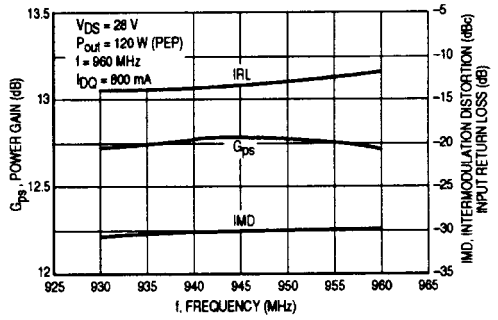
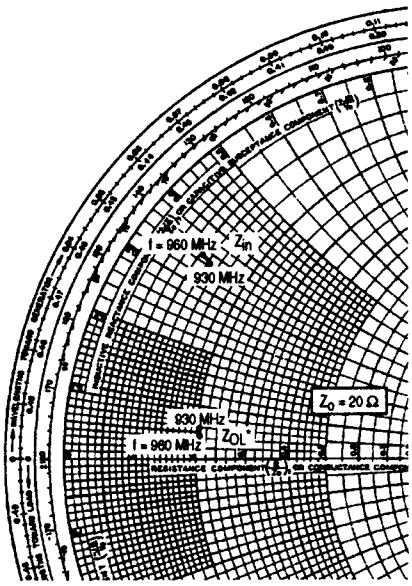


Figure 11. Broadband Circuit Performance



V_{CC} = 28 V, I_{DQ} = 2 × 400 mA, P_{out} = 120 Watts (PEP)

f MHz	Z_{in} Ω	Z_{OL}^* Ω
930	2.5 + j6.9	4.3 + j1.2
945	2.5 + j7.0	4.3 + j1.0
960	2.2 + j7.1	4.3 + j0.9

Z_{in} = Conjugate of impedance.

Z_{OL}^* = Conjugate of the optimum load impedance at a given output power, voltage, IMD, bias current, efficiency and frequency.

Note: Z_{OL}^* was chosen based on tradeoffs between gain, output power, drain efficiency and intermodulation performance. Impedances shown represent a single channel (1/2 of MRF186) impedance measurement.

Figure 12. Series Equivalent Input and Output Impedance

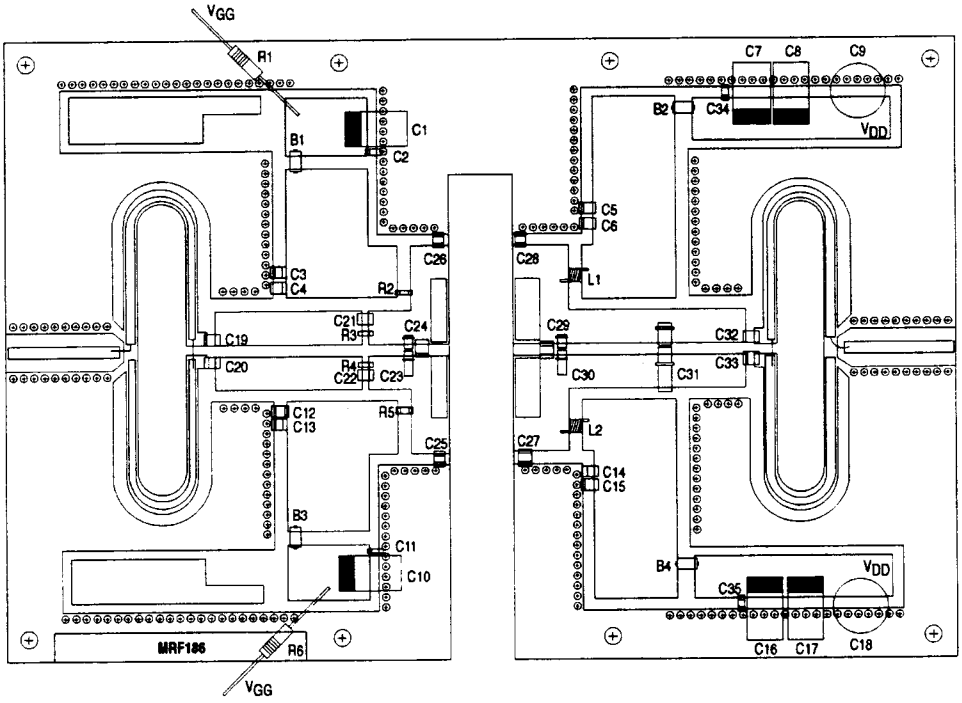
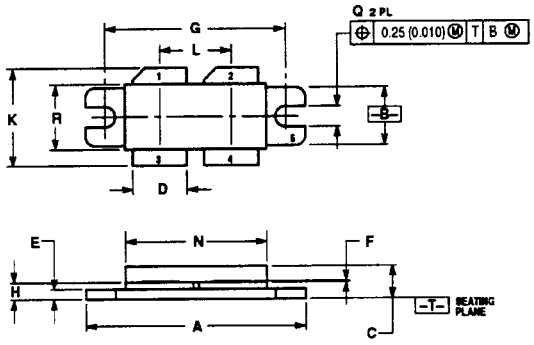


Figure 13. Component Placement Diagram of 930 – 960 MHz Broadband Test Fixture

PACKAGE DIMENSIONS



NOTES:
1. DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2. CONTROLLING DIMENSION: INCH.

DIM	INCHES MIN	INCHES MAX	MILLIMETERS MIN	MILLIMETERS MAX
A	1.330	1.350	33.78	34.29
B	0.375	0.395	9.52	10.03
C	0.180	0.205	4.57	5.21
D	0.320	0.340	8.13	8.64
E	0.060	0.070	1.52	1.77
F	0.004	0.006	0.11	0.15
G	1.100 BSC		27.94 BSC	
H	0.082	0.097	2.09	2.46
K	0.580	0.620	14.73	15.75
L	0.435 BSC		11.05 BSC	
N	0.845	0.875	21.46	22.23
Q	0.118	0.130	3.00	3.30
R	0.390	0.410	9.91	10.41

STYLE 1:
PIN 1. COLLECTOR
2. COLLECTOR
3. BASE
4. BASE
5. EMITTER

CASE 375A–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Motorola data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and Ⓜ are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

Mfax is a trademark of Motorola, Inc.

How to reach us:
USA/EUROPE/Locations Not Listed: Motorola Literature Distribution; P.O. Box 5405, Denver, Colorado 80217. 303–675–2140 or 1–800–441–2447

Mfax™: RMFAX0@email.sps.mot.com – TOUCHTONE 602–244–6609 – US & Canada ONLY 1–800–774–1848

INTERNET: http://motorola.com/sps

JAPAN: Nippon Motorola Ltd.: SPD, Strategic Planning Office, 4–32–1, Nishi–Gotanda, Shinagawa–ku, Tokyo 141, Japan. 81–3–5487–8488

ASIA/PACIFIC: Motorola Semiconductors H.K. Ltd.; 8B Tai Ping Industrial Park, 51 Ting Kok Road, Tai Po, N.T., Hong Kong. 852–26629298

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF899/D

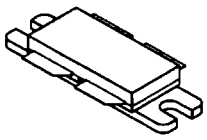
The RF Line
NPN Silicon RF Power Transistor

MRF899

150 W, 900 MHz
RF POWER
TRANSISTOR
NPN SILICON

Designed for 26 Volt UHF large–signal, common emitter, Class AB linear amplifier applications in industrial and commercial FM/AM equipment operating in the range 800–960 MHz.

- Specified 26 Volt, 900 MHz Characteristics
  Output Power = 150 Watts (PEP)
  Minimum Gain = 8.0 dB @ 900 MHz, Class AB
  Minimum Efficiency = 35% @ 900 MHz, 150 Watts (PEP)
  Maximum Intermodulation Distortion –28 dBc @ 150 Watts (PEP)
- Characterized with Series Equivalent Large–Signal Parameters from 800 to 960 MHz
- Silicon Nitride Passivated
- 100% Tested for Load Mismatch Stress at all Phase Angles with 5:1 VSWR @ 26 Vdc, and Rated Output Power
- Gold Metallized, Emitter Ballasted for Long Life and Resistance to Metal Migration
- Circuit board photomaster available upon request by contacting RF Tactical Marketing in Phoenix, AZ.



CASE 375A–01, STYLE 1

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Collector–Emitter Voltage	V_{CEO}	28	Vdc
Collector–Emitter Voltage	V_{CES}	60	Vdc
Emitter–Base Voltage	V_{EBO}	4.0	Vdc
Collector–Current — Continuous	I_C	25	Adc
Total Device Dissipation @ T_C = 25°C Derate above 25°C	P_D	230 1.33	Watts W/°C
Storage Temperature Range	T_{stg}	–65 to +150	°C

THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	0.75	°C/W

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS					
Collector–Emitter Breakdown Voltage (I_C = 100 mAdc, I_B = 0)	$V_{(BR)CEO}$	28	37	—	Vdc
Collector–Emitter Breakdown Voltage (I_C = 50 mAdc, V_{BE} = 0)	$V_{(BR)CES}$	60	85	—	Vdc
Emitter–Base Breakdown Voltage (I_E = 10 mAdc, I_C = 0)	$V_{(BR)EBO}$	4.0	4.9	—	Vdc
Collector Cutoff Current (V_{CE} = 30 Vdc, V_{BE} = 0)	I_{CES}	—	—	10	mAdc
ON CHARACTERISTICS					
DC Current Gain (I_{CE} = 1.0 Adc, V_{CE} = 5.0 Vdc)	h_{FE}	30	75	120	—
DYNAMIC CHARACTERISTICS					
Output Capacitance (V_{CB} = 26 Vdc, I_E = 0, f = 1.0 MHz) (1)	C_{ob}	—	75	—	pF

(1) For information only. This part is collector matched.

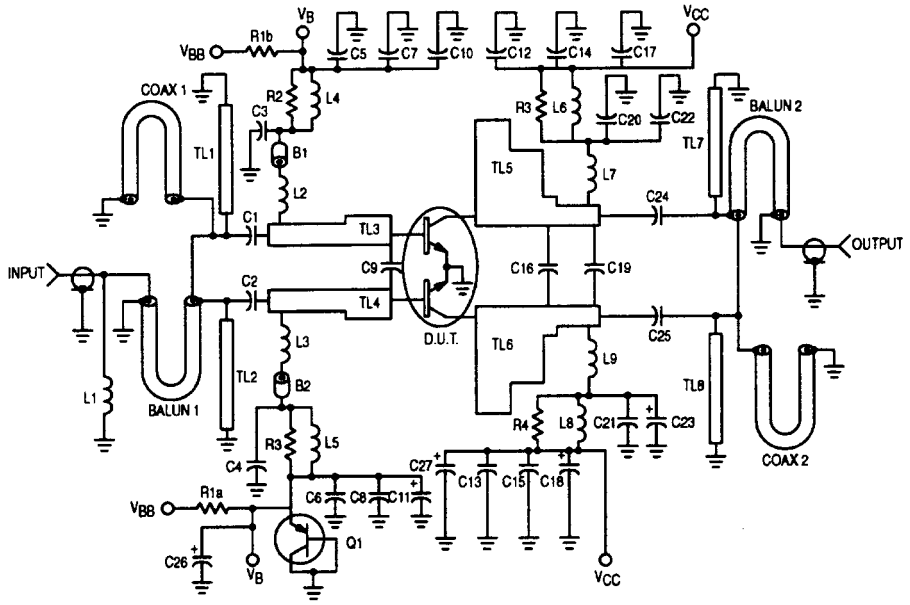
(continued)

© Motorola, Inc. 1997

MOTOROLA

ELECTRICAL CHARACTERISTICS — continued (T_C = 25°C unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
FUNCTIONAL CHARACTERISTICS					
Common–Emitter Amplifier Power Gain V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	G_{pe}	8.0	9.0	—	dB
Collector Efficiency V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	η	30	40	—	%
3rd Order Intermodulation Distortion V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	IMD	—	–32	–28	dBc
Output Mismatch Stress V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz, VSWR = 5:1 (all phase angles)	Ψ	No Degradation in Output Power Before and After Test			



B1, B2 — Ferrite Bead, Ferroxcube #56–590–65–3B
C1, C2, C24, C25 — 43 pF, B Case, ATC Chip Capacitor
C3, C4, C20, C21 — 100 pF, B Case, ATC Chip Capacitor
C5, C6, C12, C13 — 1000 pF, B Case, ATC Chip Capacitor
C7, C8, C14, C15 — 1800 pF, AVX Chip Capacitor
C9 — 9.1 pF, A Case, ATC Chip Capacitor
C10, C11, C17, C18, C22, C23 — 10 µF, Electrolytic Capacitor Panasonic
C16 — 3.9 pF, B Case, ATC Chip Capacitor
C19 — 0.8 pF, B Case, ATC Chip Capacitor
C26 — 200 µF, Electrolytic Capacitor Mallory Sprague
C27 — 500 µF Electrolytic Capacitor

L1 — 5 Turns 24 AWG IDIA 0.059″ Choke, 19.8 nH
L2, L3, L7, L9 — 4 Turns 20 AWG IDIA 0.163″ Choke
L4, L5, L6, L8 — 12 Turns 22 AWG IDIA 0.140″ Choke
N1, N2 — Type N Flange Mount, Omni Spectra
Q1 — Bias Transistor BD136 PNP
R2, R3, R4, R5 — 4.0 x 39 Ohm 1/8 W Chips in Parallel
R1a, R1b — 56 Ohm 1.0 W
TL1–TL8 — See Photomaster
Balun1, Balun2, Coax 1, Coax 2 — 2.20″ 50 Ohm 0.086″ o.d. Semi–rigid Coax, Micro Coax
Board — 1/32″ Glass Teflon, ε_r = 2.55″ Arlon (GX–0300–55–22)

Figure 1. 900 MHz Power Gain Test Circuit

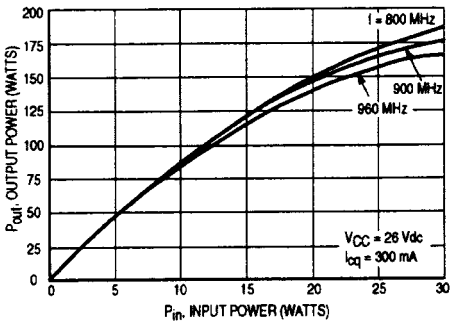


Figure 2. Output Power versus Input Power

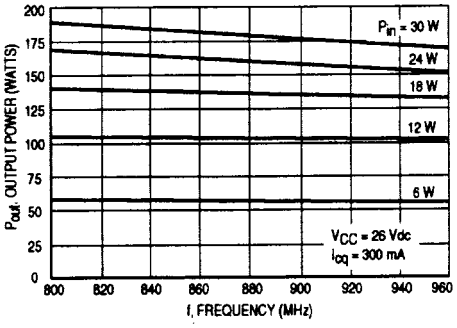


Figure 3. Output Power versus Frequency

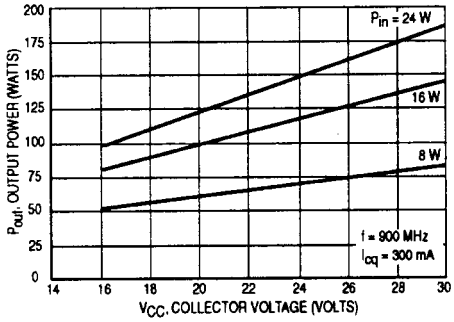


Figure 4. Output Power versus Supply Voltage

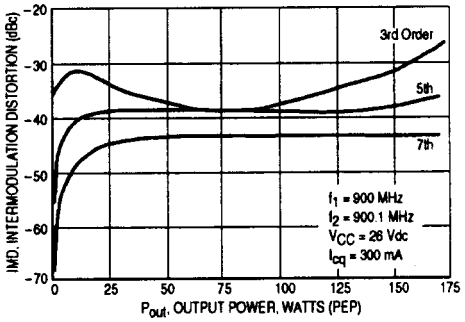


Figure 5. Intermodulation versus Output Power

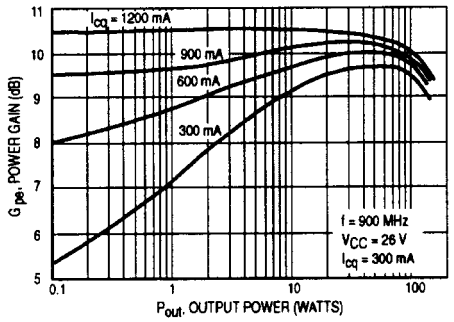


Figure 6. Power Gain versus Output Power

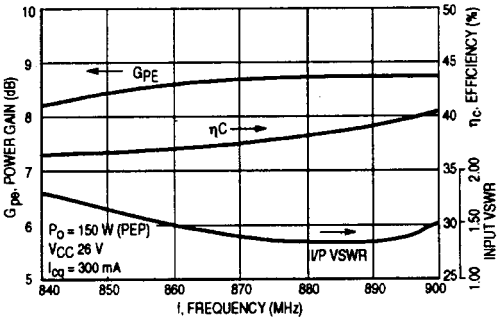
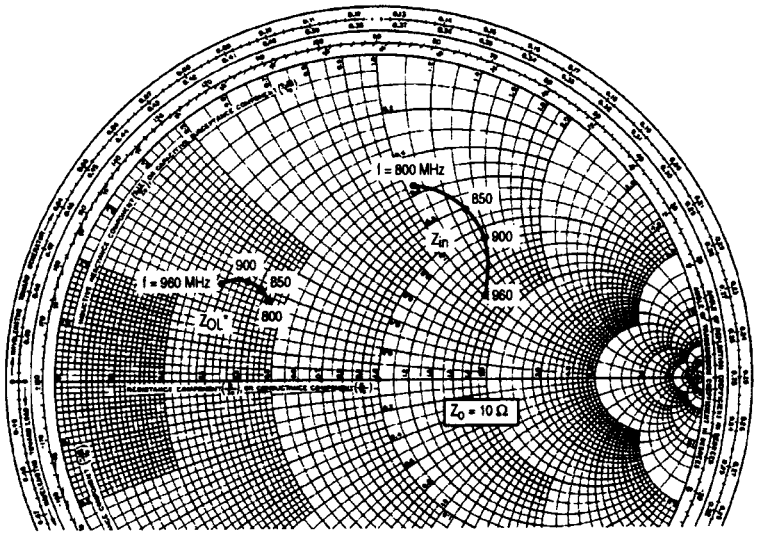


Figure 7. Broadband Test Fixture Performance



f MHz	Z_{in} Ohms	Z_{OL}^* Ohms
800	5.51 + j10.6	4.52 + j2.64
850	8.17 + j13.2	4.21 + j2.98
900	11.2 + j13.8	3.68 + j2.97
960	16.8 + j10.1	2.98 + j2.71

Z_{OL}^* = Conjugate of optimum load impedance into which the device operates at a given output power, voltage and frequency.

NOTE: Z_{in} & Z_{OL}^* are given from base–to–base and collector–to–collector respectively

Figure 8. Input and Output Impedances with Circuit Tuned for Maximum Gain @ P_O = 150 W (PEP), V_{CC} = 26 V

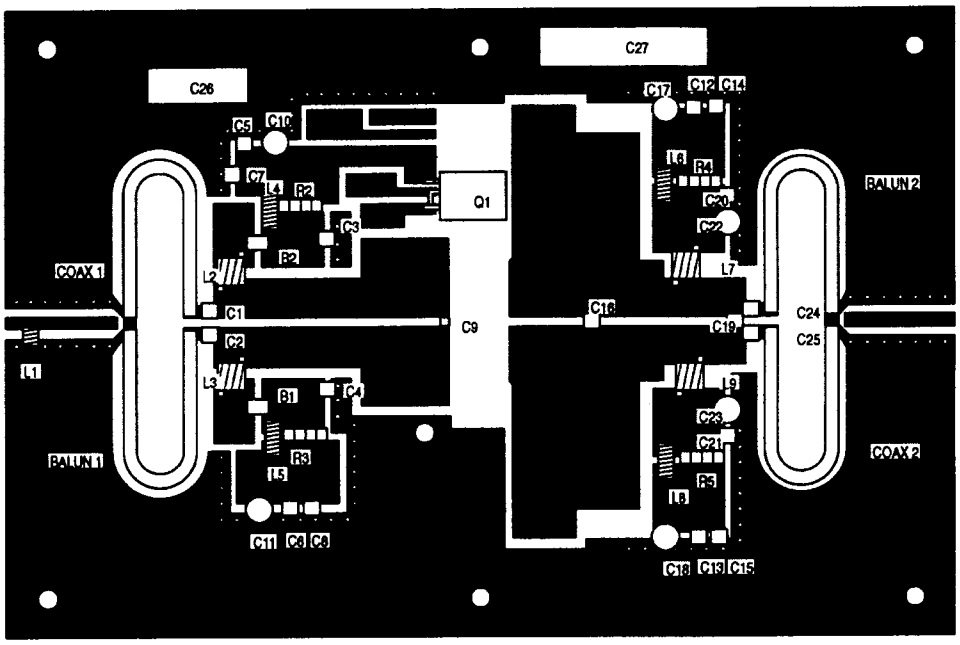
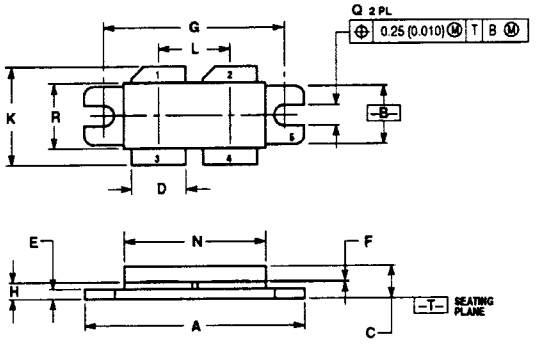


Figure 9. MRF899 Test Fixture Component Layout

PACKAGE DIMENSIONS



NOTES:
1. DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2. CONTROLLING DIMENSION: INCH.

DIM	INCHES MIN	INCHES MAX	MILLIMETERS MIN	MILLIMETERS MAX
A	1.330	1.350	33.79	34.29
B	0.375	0.395	9.52	10.03
C	0.160	0.205	4.57	5.21
D	0.320	0.340	8.13	8.64
E	0.060	0.070	1.52	1.77
F	0.004	0.006	0.11	0.15
G	1.100 BSC		27.94 BSC	
H	0.062	0.097	2.08	2.46
K	0.580	0.620	14.73	15.75
L	0.435 BSC		11.05 BSC	
N	0.845	0.875	21.46	22.23
Q	0.116	0.130	3.00	3.30
R	0.390	0.410	9.91	10.41

STYLE 1:
PIN 1. COLLECTOR
2. COLLECTOR
3. BASE
4. BASE
5. EMITTER

CASE 375A–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Motorola data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and M are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

Mfax is a trademark of Motorola, Inc.

How to reach us:
USA/EUROPE/Locations Not Listed: Motorola Literature Distribution;
P.O. Box 5405, Denver, Colorado 80217. 303–675–2140 or 1–800–441–2447

Mfax™: RMFAX0@email.sps.mot.com – TOUCHTONE 602–244–6609
– US & Canada ONLY 1–800–774–1848

INTERNET: http://motorola.com/sps

JAPAN: Nippon Motorola Ltd.: SPD, Strategic Planning Office, 4–32–1, Nishi–Gotanda, Shinagawa–ku, Tokyo 141, Japan. 81–3–5487–8488

ASIA/PACIFIC: Motorola Semiconductors H.K. Ltd.; 8B Tai Ping Industrial Park, 51 Ting Kok Road, Tai Po, N.T., Hong Kong. 852–26629298

REFERENCES

1. Ulrich L. Rohde, Anthony M. Pavio, and Robert A. Pucel, "Accurate Noise Simulation of Microwave Amplifiers Using CAD," *Microwave J.*, **31**(12), 130–141, December 1988.
2. H. T. Friis, "Noise Figures of Radio Receivers," *Proc. IRE*, **32**, 419–422, July 1944.
3. "IRE Standards on Methods of Measuring Noise in Linear Twoports," *Proc. IRE*, **48**, 60–68, January 1960.
4. H. Rothe and W. Dhalke, "Theory of Noisy Fourpoles," *Proc. IRE*, **44**, 811–818, June 1956.
5. "Representation of Noise in Linear Twoports," *Proc. IRE*, **48**, 69–74, January 1960.
6. H. A. Haus and R. B. Adler, *Circuit Theory of Linear Noisy Networks*, Cambridge, MA: MIT Press, 1959/New York: Wiley, 1959.
7. H. Fukui, "Noise Performance of Microwave Transistors," *IEEE Trans. Electron Devices*, **ED-13**, 329–341, March 1966.
8. S. Iversen, "The Effect of Feedback on Noise Figure," *Proc. IEEE*, **63**, 540–542, March 1975.
9. L. Besser, "Stability Considerations of Low-Noise Transistor Amplifiers with Simultaneous Noise and Power Match," *IEEE MTT-S International Microwave Symposium Digest*, 1975, pp. 327–329.
10. G. D. Vendelin, "Feedback Effects on the Noise Performance of GaAs Mesfets," *IEEE MTT-S International Microwave Symposium Digest*, 1975, pp. 324–326.
11. W. A. Suter, "Feedback and Parasitic Effects on Noise," *Microwave J.* **26**(2), 123–129, February 1983.
12. K. Hartmann and M. J. O. Strutt, "Changes in Four Noise Parameters Due to General Changes in Linear Two-Port Circuits," *IEEE Trans. Electron Devices*, **ED-20**(10), 874–877, October 1973.
13. H. Fukui, "Available Power Gain, Noise Figure, and Noise Measure of Two-Ports and Their Graphical Representation," *IEEE Trans. Circuit Theory*, **CT-13**, 137–142, June 1966.
14. C. R. Poole and D. K. Paul, "Optimum Noise Measure Terminations for Microwave Transistor Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**(11), 1254–1257, November 1985.
15. H. Hillbrand and P. H. Russer, "An Efficient Method for Computer Aided Noise Analysis of Linear Amplifier Networks," *IEEE Trans. Circuits Syst.*, **CAS-23**(4), 235–238, April 1976.
16. W. E. Pastori, *Topics on Noise*, Eaton Electronics Instrumentation Division (EID) Seminar Notes, Los Angeles, 1983.
17. W. E. Pastori, "Image and Second Stage Corrections Resolve Noise Figure Measurement Confusion," *Microwave Syst. News*, **13**(5), 67–86, May 1983.
18. R. Q. Lane, "The Determination of Device Noise Parameters," *Proc. IEEE*, **57**, 1461–1462, August 1969.
19. R. Q. Lane, "A Microwave Noise and Gain Parameter Test Set," *IEEE International Solid-State Circuits Conference*, 1978, pp. 172–173.
20. N. Kuhn, "Accurate and Automatic Noise Figure Measurements with Standard Equipment," *IEEE MTT-S International Microwave Symposium Digest*, 1980, pp. 425–427.
21. E. W. Strid, "Measurement of Losses in Noise-Matching Networks," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**, 247–252, March 1981.
22. Chris Bowick, *RF Circuit Design*. Indianapolis: Howard W. Sams & Co., 1982.
23. Adel S. Sedra and Kenneth C. Smith, *Microelectronic Circuits*. Chicago: Saunders College Publishing, 1991.
24. Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed. New York: Wiley, 1984.

25. S. H. Lee, A. K. Wong, and M. G. Wong, *A Current-Combiner Circuit for Better Mixer Conversion Gain*, Philips Semiconductors applications document.
26. F. M. Fano, "Theoretical Limitations in the Broadband Matching Arbitrary Impedances," *J. Franklin Inst.*, **249**, 57–85, 139–155, 1950.
27. Ravender Goyal, ed., *High-Frequency Analog Integrated-Circuit Design*. New York: Wiley, 1995, Chap. 6.
28. Stefan Georgi and Ulrich L. Rohde, "A Novel Approach to Voltage Controlled Tuned Filters Using CAD Validation," *Microwave Eng. Eur.*, 35–42, October 1997.
29. *Analog/Interface ICs Device Data*, Vol. II, Q4/96, DL128, Rev 6, Motorola, Inc., MC13158 Wideband FM IF Subsystem specifications, pp. 8-308 to 8-329.
30. E. J. Wilkinson, "An *N*-Way Power Divider," *IRE Trans. Microwave Theory Tech.*, **MTT-8**, 116, January 1960.
31. Motorola Application Note AN-282A, *Systemizing RF Power Amplifier Design*.
32. W. E. Everitt and G. E. Anner, *Communication Engineering*. New York: McGraw-Hill, 1956.
33. G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks and Coupling Structures*. New York: McGraw-Hill, 1980.
34. G. L. Matthaei, "Tables of Chebyshev Impedance-Transforming Networks of Low-Pass Filter Form," *Proc. IEEE*, August 1964.
35. Motorola Application Note AN-267, *Matching Network Designs with Computer Solutions*.
36. E. G. Cristal, "Tables of Maximally Flat Impedance-Transforming Networks of Low-Pass Filter Form," *IEEE Trans. Microwave Theory Tech.*, **MTT 3**(5), September 1965.
37. H. W. Bode, *Network Analysis and Feedback Amplifier Design*. New York: Van Nostrand, 1943.
38. R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," *J. Franklin Inst.*, 139–154, January-February 1950.
39. J. H. Horwitz, "Design Wideband UHF-Power Amplifiers," *Electron. Des.* **11**, x–x, May 24, 1969.
40. O. Pitzalis and R. A. Gilson, "Tables of Impedance Matching Networks Which Approximate Prescribed Attenuation Versus Frequency Slopes," *IEEE Trans. Microwave Theory Tech.*, **MTT-9**(4), April 1971.
41. Anthony N. Gerkis, "Broadband Impedance Matching Using the 'Real Frequency' Network Synthesis Technique," *Appl. Microwave & Wireless*, **10**(6), 26–36, July/August 1998.
42. T. Cuthbert, "Broadband Impedance Matching Methods," *RF Des.*, **17**(8), 64–71, August 1994.
43. T. Cuthbert, *Broadband Direct-Coupled and Matching RF Networks*. Greenwood, AR, T. R. Cuthbert, 1999.
44. George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990, pp. 279–283.
45. Andy Potter, "HP RF Compiler Automates Schematic Capture and Extends Capabilities of Circuit Synthesis." *Appl. Microwave & Wireless*, **11**(6), 106–118, June 1999.
46. Aron Kain, *Final Report for an Integrated CAD Tool for High-Power, High-Efficiency RF and Microwave Amplifiers*, STTR Phase II, Contract no. F33615-95-C-174, Dept. of the Air Force, Small Business Technology Transfer (STTR) Program.
47. Kenneth K. Clarke and Donald T. Hess, *Communication Circuits: Analysis and Design.* Reading, MA: Addison-Wesley, 1971.

FURTHER READING

P. L. D. Abrie, *The Design of Impedance-Matching Networks for Radio-Frequency and Microwave Amplifiers*. Norwood, MA: Artech House, 1985.

P. L. D. Abrie and P. Rademeyer, "A Method for Evaluating and the Evaluation of the Influence of the Reverse Transfer Gain on the Transducer Power Gain of Some Microwave Transistors," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**(8), 711–713, August 1985.

A. Ambrózy, *Electronic Noise*. New York: McGraw-Hill, 1982.

W. Baechtold and M. J. O. Strutt, "Noise in Microwave Transistors," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**(9), 578–585, September 1968.

A. Cappy, "Noise Modeling and Measurement Techniques," *IEEE Trans. Microwave Theory Tech.*, **MTT-36**, 1–10, January 1988.

R. S. Carson, *High Frequency Amplifiers*. New York: Wiley, 1975.

Cezar A. A. Carioca and Paulo H. P. de Carvalho, "Computer-Aided Design of Diode Frequency Multipliers," *Appl. Microwave & Wireless*, **11**(4), 38–46, April 1999.

H. J. Carlin, "A New Approach to Gain-Bandwidth Problems," *IEEE Trans. Circuits Syst.*, **CAS-23**, 170–175, April 1977.

H. J. Carlin and P. Amstutz, "On Optimum Broad-Band Matching," *IEEE Trans. Circuits Syst.*, **CAS-23**, 401–405, April 1977.

H. J. Carlin and P. Amstutz, "On Optimum Broadband Matching," *IEEE Trans. Circuits Syst.*, **CAS-28**(5), May 1981.

H. J. Carlin and J. J. Komiak, "A New Method of Broad-Band Equalization Applied to Microwave Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**, 93–99, February 1979.

H. J. Carlin and B. Yarman, "A Simplified Real Frequency Technique Applied to Broad-Band Multistage Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-30**(12), December 1982.

H. J. Carlin and B. S. Yarman, "The Double Matching Problem: Analytic and Real-Frequency Solutions," *IEEE Trans. Circuits Syst.*, **CAS-30**(1), 15–28, January 1983.

G. Caruso and M. Sannino, "Computer-Aided Determination of Microwave Two-Port Noise Parameters," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 639–642, September 1978.

Complex Math User's Guide, Compact Software, Paterson, NJ, 1988.

Steve C. Cripps, *RF Power Amplifiers for Wireless Communications*. Norwood, MA: Artech House, 1999.

T. Cuthbert, *Circuit Design Using Personal Computers*. New York: Wiley, 1983.

J. A. Eisenberg, "Designing Amplifiers for Optimum Noise Figure," *Microwaves*, **13**, 36–41, April 1974.

R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," *J. Franklin Inst.*, **249**, 57–83, January 1960/139–155, February 1960.

H. Fukui, *Low-Noise Microwave Transistors and Amplifiers*. New York: IEEE Press, 1981.

C. Gentili, *Microwave Amplifiers and Oscillators*. New York: McGraw-Hill, 1987.

Stefan Georgi and Ulrich L. Rohde, "A Novel Approach to Voltage-Controlled Tuned Filters Using CAD Validation," *Microwave Eng. Eur.*, 35–42, October 1997.

G. Gonzales, *Microwave Transistor Amplifier Analysis and Design*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

H. A. Haus and R. B. Adler, "Optimum Noise Performance of Linear Amplifiers," *Proc. IRE*, **46**, 1517–1533, August 1958.

Aron Kain, *Final Report for Bias Dependence Noise Modeling of Heterojunction Bipolar Transistors*, USAF SBIR Phase II (PIIN), F33615-95-C-1707, November 1997. Issued by USAF/AFMC/ASC, Wright Laboratory WL/AAKE BLD 7, 2530 C ST, Wright-Patterson AFB, OH 45433-7607.

M. J. Köppen, "A Single-Stage Power Amplifier for 17 watts Output in the 470-MHz Band Using the 266BLY, Operating from 12.5 volt Supply Voltage," C.A.B. Communications, Eindhoven, Holland, Electronic Application Laboratory Report, August 8, 1970.

William B. Kuhn and Naveen K. Yanduru, "Spiral Inductor Substrate Loss Modeling in SiliconRFICs," *Microwave J.*, **11**(3), 66–81, March 1999.

J. Lange, "Noise Characterization of Linear Two-Ports of Invariant Parameters," *IEEE J. Solid-State Circuits*, **SC-2**, 37–40, June 1967.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

David Lovelace, Natalino Camilleri, and George Kannell, "Silicon MMIC Inductor Modeling for High Volume, Low Cost Applications," *Microwave J.*, **39**(8), 60–71, August 1994.

Jarek Lucek and Robbin Damen, "Designing an LNA for a CDMA Front End," *RF Des.*, **22**(2), 20–30, February 1999.

Stephen A. Maas, "FET Models for Volterra-Series Analysis," *Microwave J.*, **42**(5), 260–270, May 1999.

Donald A. McClure, KB2Z, "Transmission-Line Transformers," *Commun. Q.*, 45–58, Summer 1997.

D. J. Mellor and J. G. Linvill, "Synthesis of Interstage Networks and Prescribed Gain Versus Frequency Slopes," *IEEE Trans. Microwave Theory Tech.*, **MTT-23**(12), 1013–1020, December 1975.

Sean Mercer, "Exact Simulation of LNAs Reduces Design Cycle Time," *Appl. Microwave & Wireless*, 74–82, January 1999.

R. P. Meys, "A Wave Approach to the Noise Properties of Linear Microwave Devices," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 34–37, January 1978.

C. K. S. Miller, W. C. Daywitt, and M. G. Arthur, "Noise Standards, Measurements and ReceiverNoise Definitions," *Proc. IEEE*, **55**, June 1967.

Brian J. Minnis, *Designing Microwave Circuits by Exact Synthesis* . Boston: Artech House, 1996.

R. Minton, *Design of Large Signal VHF Transistor Power Amplifiers* , RCA Application Note SMA-36, July 1964.

R. Minton and H. C. Lee, *Frequency Multiplication Using Overlay Transistors* , RCA Application Note SMA-40, September 1965.

M. Mitama and H. Katoh, "An Improved Computational Method for Noise Parameter Measurement," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**, 612–615, June 1979.

Motorola tutorial, "The Cascode Amplifier: A Useful RF Building Block," *Microwave Prod. Digest*, 39, 59, March 1999.

T. Mukaihata, "Applications and Analysis of Noise Generation in *N*-Cascaded Mismatched Two-Port Networks," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 699–708, September 1968.

W. W. Mumford and E. H. Scheibe, *Noise Performance Factors in Communications Systems* . Dedham, MA: Horizon House, 1968.

D. E. Norton, "High Dynamic Range Transistor Amplifier Using Lossless Feedback," *Microwave J.*, **19**(5), 53–57, May 1976.

D. E. Norton, "Strong Signal Survivability Design for Lossless Feedback Amplifiers," *Proceedings of the Eleventh Annual European Microwave Conference* , 1981.

H. S. Okean and P. O. Lombardo, "Noise Performance of M/W and MM-Wave Receivers," *Microwave J.*, **16**(1), January 1973.

T. Y. Otoshi, "The Effect of Mismatched Components on Microwave Noise-Temperature Calibrations," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 675–686, September 1968.

P. Penfeld, "Wave Representations of Amplifier Noise," *IRE Trans. Circuit Theory*, 84–86, March 1962.

Ulrich L. Rohde, Jerry C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*, 2nd ed. New York: McGraw-Hill, 1997, Chap. 3.

C. Rorabaugh, *Circuit Design and Analysis—Featuring C Routines* . New York: McGraw-Hill, 1992.

M. Sannino, "On the Determination of Device Noise and Gain Parameters," *Proc. IEEE*, **67**, 1364–1366, September 1979.

SYNMAT, matching synthesis program by Radom, Inc., developed by Douglas Mellor.

A. van der Ziel, *Noise in Solid State Devices and Circuits* . New York: Wiley, 1986.

Dave Vondran, "Noise Figure Measurement: Corrections Related to Match and Gain," *Microwave J.*, **42**(3), 22–38, March 1999.

D. F. Wait, "Thermal Noise from a Passive Linear Multiport," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 687–691, September 1968.

B. S. Yarman and H. J. Carlin, "A Simplified 'Real Frequency' Technique Applied to Broad-Band Multistage Microwave Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-30**(12), 2216–2222, December 1982.

G. P. Young and S. O. Scanlan, "Matching Network Design Studies for Microwave Transistor Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**(10), 1027–1035, October 1981.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

3

AMPLIFIER DESIGN WITH BJTs AND FETs

3-1 PROPERTIES OF AMPLIFIERS

3-1-1 Introduction

The goal of this chapter is to move from the details on the semiconductor devices themselves to our first practical application. One of the more interesting features of amplifiers is that they can easily become what will be covered in Chapter 5: *oscillators*. This is due to the high gain of the devices, unaccounted for parasitic elements, and other design flaws. The amplifiers we need will fall into three categories:

- Low-noise amplifiers
- High-gain amplifiers
- Medium- to high-power amplifiers

See Figure 3-1. The low-noise amplifier always operates in Class A, typically at 15–20% of its maximum useful current. The high-gain amplifier can operate in Class A, as well as Class B (mostly push–pull). The higher dc current for the same device results in a higher noise figure and more gain, and ultimately more output power. Class C operating mode is really reserved for either FM transmissions or constant-carrier modes like CW. Some of the modern digital modulation types are sensitive to phase distortion rather than amplitude changes and because of the resulting output spectrum designers have stayed away from Class C operation.

Some of these amplifiers, such as bipolar versions, can be dc-coupled with very few difficulties; others, like those in the FET families, will cause more headaches. An interesting example, although much too high in frequency for the purpose of this book, is shown in Figures 3-2, 3-3, 3-4, and 3-5. It shows a three-stage amplifier that has noise feedback for the first two stages and resistive feedback for the output stage. These are the three types of amplifiers, regardless of technology, we will be evaluating. The achievement of accurate noise analysis of this circuit, reported in Rohde et al. [1], marked the first time a complete

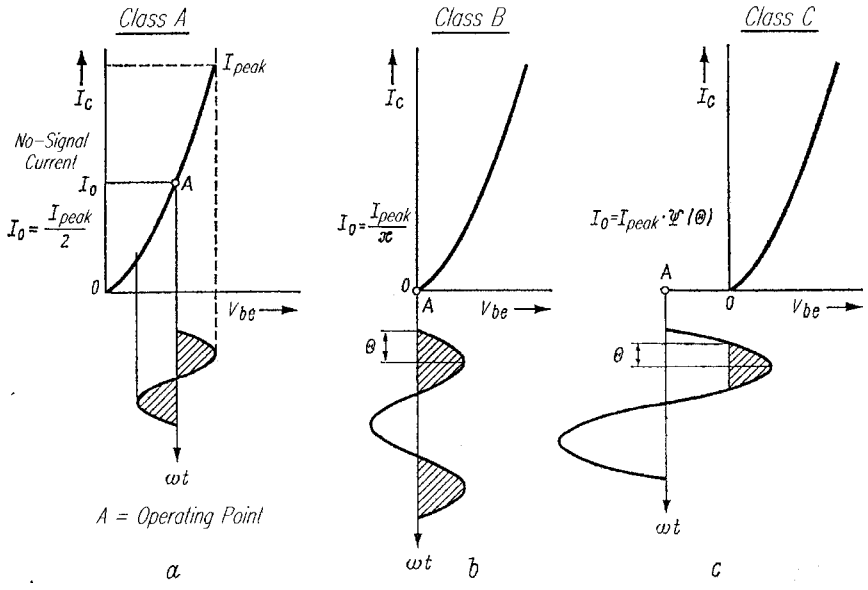


Figure 3-1 Definition of Class A, B, and C operation and resulting bias, including conduction angle (Θ, to be discussed in Section 3-2-2). The transfer characteristic can be either quadratic or exponential. This results in different distortion but does not change the basic operating mode.

linear noise model with essentially no frequency limitation had been developed. Its accuracy depends solely on the accuracy of the element values of the linear equivalent circuit, which can be obtained from measured *S*-parameters by an optimization process. Needless to say, this technique is also applicable at lower frequencies, such as 1.5–3 GHz, and its deviation from measured values has never been worse than 0.2 dB or, relative to the noise figure, 10% expressed in dB.

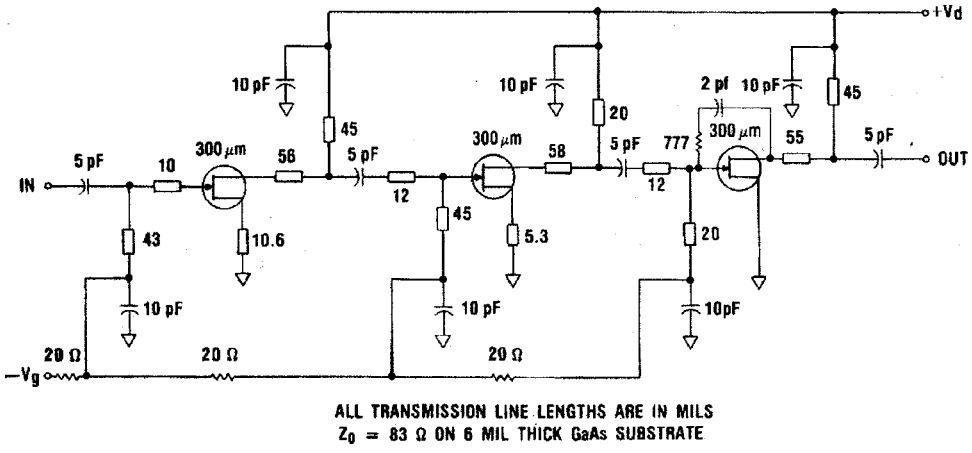


Figure 3-2 Schematic of the X-band GaAs monolithic low-noise amplifier (Texas Instruments EG8021). For reasons of linearity, all three stages operate in Class A.

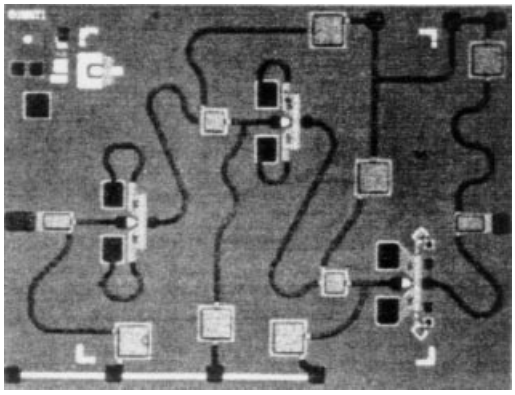


Figure 3-3 Photograph of the EG8021 monolithic amplifier chip. The area pictured is 0.09 inch by 0.12 inch in size.

The results of Figure 3-4 were obtained by using the linear FET model as supplied by Texas Instruments. It becomes somewhat obvious that a match for S_{11} of –60 dB is not likely. We then replaced the linear transistor with the SPICE-type nonlinear model, resulting in the responses shown in Figure 3-5. S_{11} now looks much more realistic in the center of the range, but at the same time exhibits a tendency toward instability at frequencies around 8 GHz. The actual circuit, when measured, did not exhibit this potential instability; on the other hand,

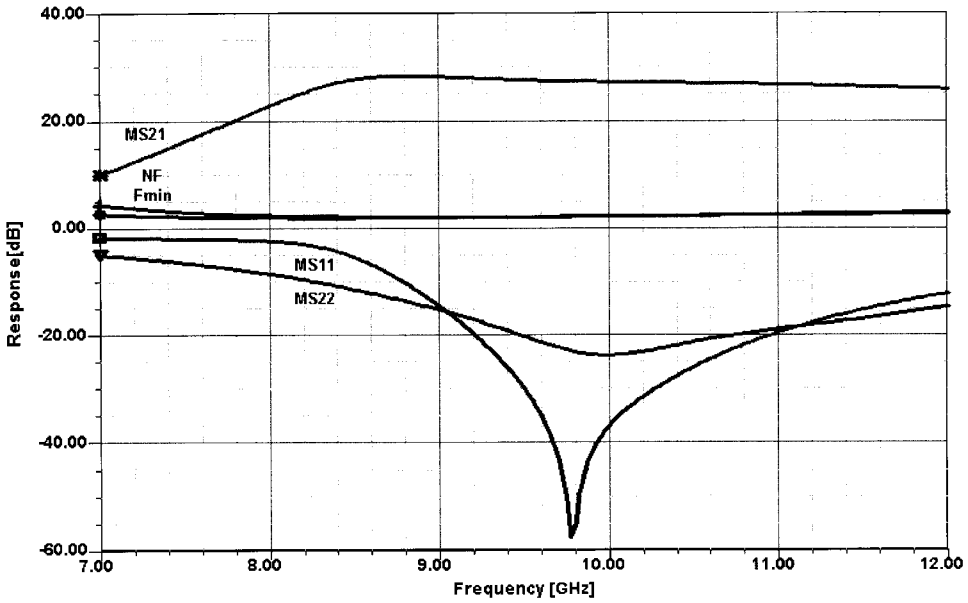


Figure 3-4 Simulated F_{min}, noise figure, S_{11}, S_{21}, and S_{22} responses for the three-stage GaAsFET amplifier using the TI linear FET model. The values at 10 GHz are: F_{min}, 2.20 dB; NF, 2.21 dB; S_{11}, –36.8 dB; S_{21}, 27.2 dB; and S_{22}, –23.7 dB.

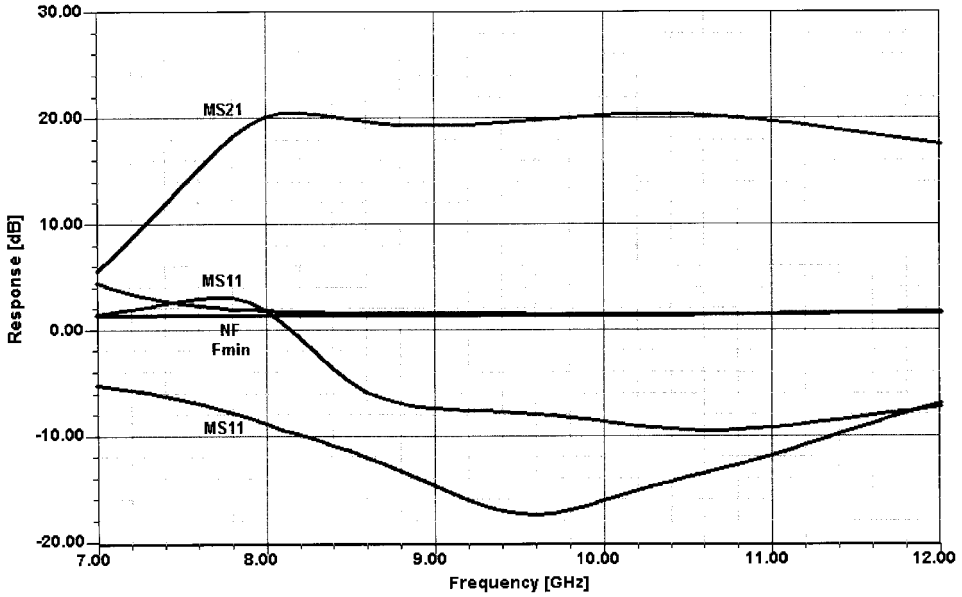


Figure 3-5 Simulated F_{min}, noise figure, S_{11}, S_{21}, and S_{22} responses for the three-stage GaAsFET amplifier using the nonlinear FET model. The values at 10 GHz are: F_{min}, 1.53 dB; NF, 1.65 dB; S_{11}, –8.5 dB; S_{21}, 20.3 dB; and S_{22}, –15.9 dB.

the noise figure and gain agreed quite well. The discrepancies between the results with the linear and nonlinear models emphasize the importance of accurate device modeling.

Figure 3-6 shows a wideband amplifier, specifically a distributed amplifier, that covers 1–20 GHz. Figure 3-7 shows its simulated frequency-dependent gain, matching, and noise performance. The example given here as the upper frequency for wireless systems is rapidly expanding and the technology to accommodate this is already in place. Modern CAD tools can handle such high frequencies accurately, which means one can trust them at the significantly lower wireless frequencies of 500 MHz to 3 GHz. Above 1500 MHz, one really should model using distributed elements. We will come to this later in this chapter.

Input/output and interstage matching will become another challenging task, specifically if one requires sufficient bandwidth. As circuits become more complex and as the operating frequency increases, passive components play an increasingly important role; the use of

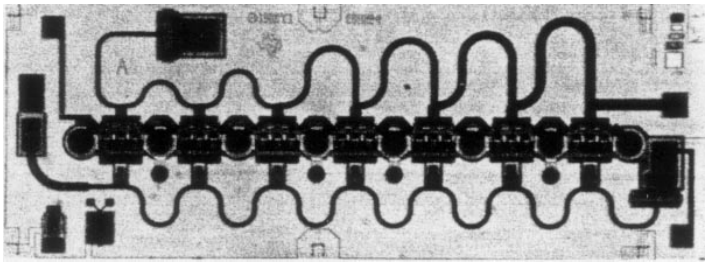


Figure 3-6 Photo of the 1–20-GHz distributed amplifier.

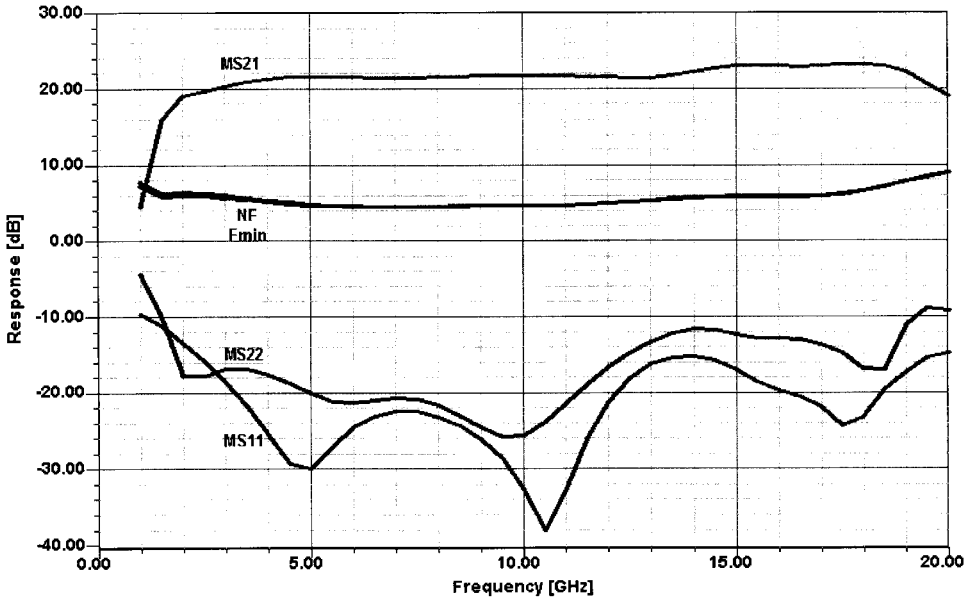


Figure 3-7 Simulated F_{min}, NF, S_{11}, S_{21}, and S_{22} responses. At the center of the range, the following numbers were found: F_{min}, 4.80 dB; NF, 4.80 dB; S_{11}, –32.8 dB; S_{21}, 21.8 dB; and S_{11}, –25.2 dB.

distributed elements is practically a necessity above 1500 MHz. For this reason, we begin this chapter by covering the application of transistors as amplifiers, following this with a discussion of the surrounding external circuitry, such as transmission lines, inductors, capacitors, and resistors. Monolithic ICs on silicon are a totally different ball game, and neither of the authors has sufficient expertise in this area. For this reason, we refer readers who will concentrate on MMIC design to the literature mentioned at the end of this chapter.

The goal of the applications presented here is to address, where possible, high-performance wireless stages that are less subject to constraints of space, power consumption, and cost, than portable/mobile applications. This is consistent with our preface, in which we have clearly stated our goal. The designers of handsets typically cannot consider the full variety of circuits we will present here. These types of stages (low noise, high gain and power) frequently will be used with filtering, and we will give some examples of input filters, including a tracking high-performance design.

A key topic besides the frequently mentioned gain is the whole noise issue. We will give the reader a complete introduction to this important area, followed by a similar introduction to the gain and matching of transistors. *Caution*: The majority of applications published go through the exercise of using *S*-parameters manually to design low-noise and small-signal amplifiers. While we have been tempted to repeat the information found in several textbooks, we feel that today's engineers will have access to CAD tools one way or another. It is highly unlikely that readers will undertake the inefficient process of going through all the design steps manually (a procedure that can take days); hopefully, engineers will resort to a linear/nonlinear CAD tool to do so. Only for the purpose of summarizing the necessary steps will we provide the relevant information regarding gain, stability, and related topics, demonstrating their usefulness or possible misuse with a modern CAD tool—in our case,

the Serenade Design Environment product from Ansoft. At this time, several additional companies beside Ansoft and Hewlett-Packard have emerged, but it is dangerous to list all of them as they come and go.

3-1-2 Gain

In practical terms, when we talk about gain we mean the output power relative to the input power. The industry standard now uses 50 Ω; however, at the point at which we have to match high-impedance devices such as monolithic crystal filters where the impedance jumps, things are not so obvious. The transformation equation that allows us to transform from one impedance to another is

$$m = \sqrt{\frac{R_2}{R_1}} \tag{3-1}$$

This correction factor, m, is required to do the transformation from R_1 (50 Ω) to R_2 (e.g., 1.2 kΩ). It needs to be understood that while the power gain remains the same, the voltage gain requires this correction factor. Since the power gain is expressed as 10 log A, A being the loop gain, 20 log A is the voltage gain, both expressed in dB. The reason for 10 versus 20 is that the power gain is always the voltage squared, which accounts for the 2. As a reminder of these relationships we present the following:

Distinction: Power Ratios–Voltage Ratios

$$n \text{ dB} = 10 \log_{10}(P_2/P_1) \quad \xrightarrow{PP = V^2/Z} \quad n = 10 \log_{10}\left(\frac{V_2^2/Z_2}{V_1^2/Z_1}\right)\text{dB}$$

$$Z_2 = Z_1 \rightarrow n \text{ dB} = 10 \log_{10}(V_2^2/V_1^2) = 10 \log_{10}(V_2/V_1)^2$$

$$\Longrightarrow \boldsymbol{n \text{ dB} = 20 \log_{10}(V_2/V_1) \text{ dB}}$$

The gain is typically defined under the condition of real resistive source and load terminations, which means that all the available power is being supplied to the amplifier, and likewise through proper matching at the output all the available power is fed to the termination. This gain figure is typically measured at a single spot frequency, but since most amplifiers are not totally flat, one needs to specify the gain tolerances or gain flatness. A good number is x dB ± 0.5 dB gain variation. In the case of a narrowband amplifier, the gain typically is shaped like a tuned circuit or like a critically coupled bandpass filter frequency response, while in the case of a real wideband application, the desired gain has some kind of ripple, which is due to the various compensation components in the circuit. An extreme case of a wideband amplifier is a 6–18-GHz single-stage amplifier that uses some very clever matching. This is shown in the following example. Figure 3-8 shows the circuit diagram, and Figure 3-9 shows its frequency-dependent performance.

While this is a cute approach for generating a single-stage amplifier over such a wide bandwidth, in real life it is not acceptable. The reason for this is that neither S_{11} nor S_{22} meets any useful specifications over the entire range. Typically, S_{11}/S_{22} values are only a few dB, while in reality more than 10 dB is required. We have chosen this example to show that gain

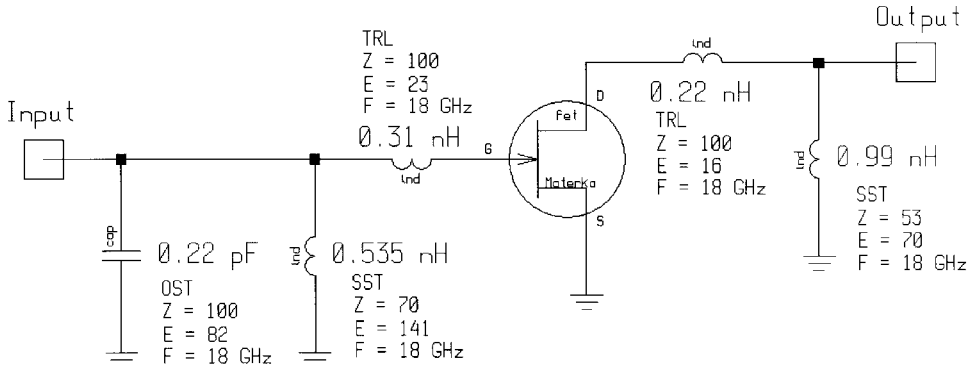


Figure 3-8 A 6–18-GHz FET amplifier.

and matching, depending on the circuit topology, are connected. We get similar results for the noise figure, whereby the best achievable noise figure ($F_{\min}$) and the actual 50-Ω noise figure are only a few dB apart; however, its frequency dependence in a wideband stage is certainly not acceptable.

Another parameter associated with gain is the stability factor K of the circuit. A value of $K > 1$ is mandatory to guarantee stability together with the S_{11} and S_{22} magnitudes being less than 1. Figure 3-10 plots the K stability factor, which for absolute stability must be >1, and maximum stable gain (MSG) as well as $G_{\max}$. It turns out that MSG is a highly artificial

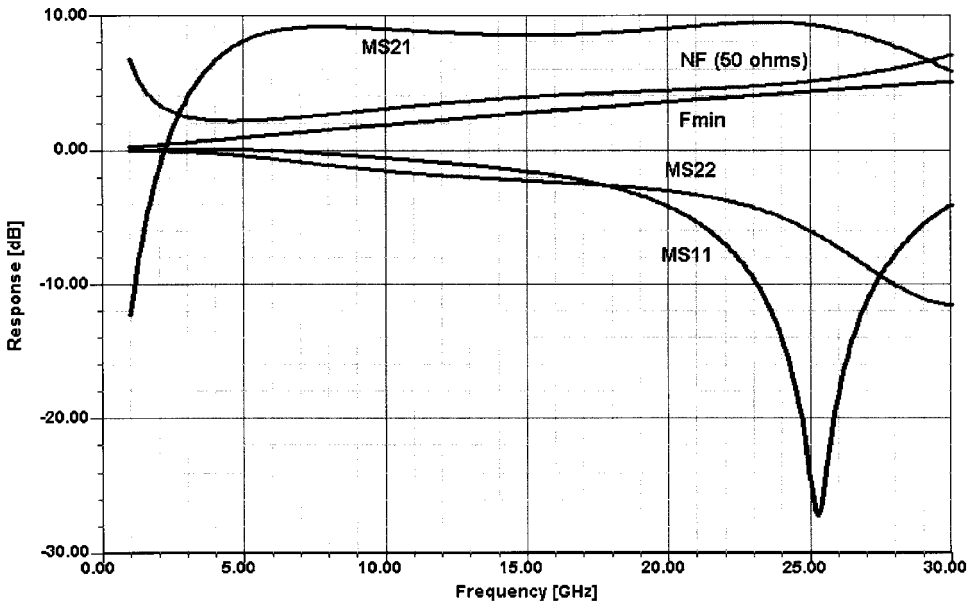


Figure 3-9 Frequency-dependent gain, matching, and noise performance of the amplifier shown in Figure 3-8.

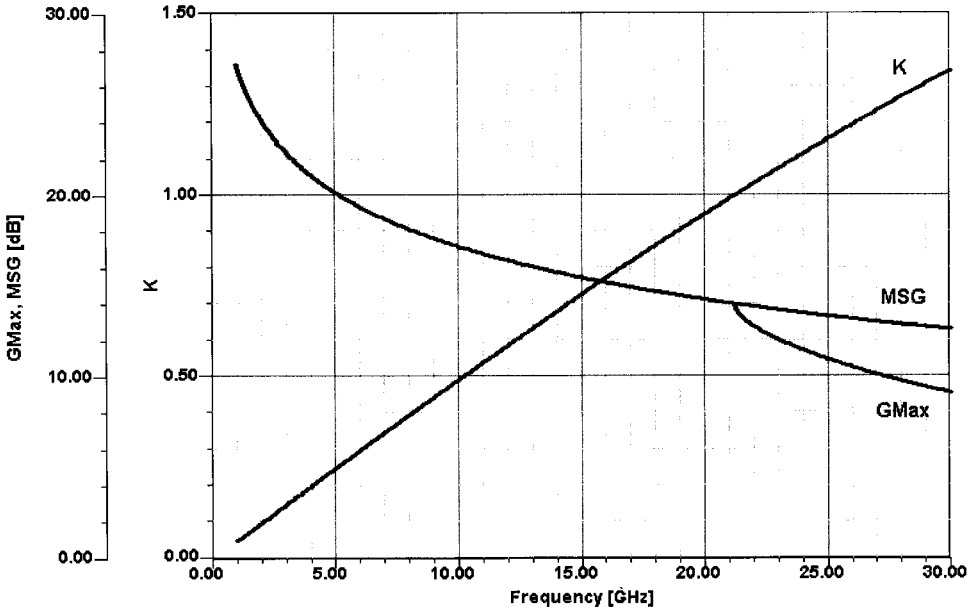


Figure 3-10 G_{max}, maximum stable gain (MSG), and K performance of the amplifier in Figure 3-8.

definition of stability that implies that input and output are terminated with pure resistances and then conjugately matched. There are four ways of achieving this resistive loading, either in the emitter as a shunt resistor between gate/drain or base/collector, or in parallel to the input and output to ground. Needless to say, this is not a real operating mode because it deteriorates both noise and gain. In our case, the entity G_{max} is a function of S_{12} of the circuit, not just the transistor, and depends highly on S_{12}. If we bring the system close to oscillation, then G_{max} will increase at the expense of bandwidth. Unfortunately, some of these microwave definitions of gain (see Table 3-1) have become widely used in industry without always being useful. Evaluating these equations for S_{12} is not legitimate because other quantities like K become infinite, and the entire system of equations falls apart. The power gain with input and output conjugately matched, or the transducer power gain for arbitrary source and load matching, still give more insight into the system.

In the above example, we specifically stated that the input and output VSWR was totally unacceptable, and this was due to a one-stage design. Figure 3-11 shows an amplifier chip built around two transistors with standard feedback techniques applied; Figures 3-12 and 3-13 show its frequency-dependent gain, matching, stability, and noise performance. This amplifier was one of the early examples of GaAsFET wideband amplifiers, but it is interesting because while it has wild (and wide) gain swings, it exhibits a fairly constant input and output impedance or reflection coefficient. A further evaluation of the circuit shows that not only is the circuit unconditionally stable, but that the maximum stable gain has an extremely high value based on essentially minimal S_{12}, the K factor is always significantly higher than 1, and G_{max} follows, to some degree, the actual gain response of the amplifier. (Ultimately, designers gave up on simple *RCL* feedback amplifiers for such a wide bandwidth, replacing them with distributed amplifiers previously called traveling wave amplifiers.) The problem of gain and stability starts at audiofrequency amplifiers, where the

Table 3-1 Nine power gains

Transducer power gain in 50-Ω system	$G_T = \lvert S_{21}\rvert^2$
Transducer power gain for arbitrary Γ_G and Γ_L	$G_T = \dfrac{(1-\lvert\Gamma_G\rvert^2)\,\lvert S_{21}\rvert^2\,(1-\lvert\Gamma_L\rvert^2)}{\lvert(1-S_{11}\Gamma_G)\,(1-S_{22}\Gamma_L)-S_{12}S_{21}\Gamma_G\Gamma_L\rvert^2}$
Unilateral transducer power gain	$G_{\mathrm{TU}} = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_G\rvert^2)\,(1-\lvert\Gamma_L\rvert^2)}{\lvert 1-S_{11}\Gamma_G\rvert^2\lvert 1-S_{22}\Gamma_L\rvert^2}$
Power gain with input conjugate matched	$G = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_L\rvert^2)}{\lvert 1-S_{22}\Gamma_L\rvert^2(1-\lvert S'_{11}\rvert^2)} = \dfrac{\lvert S_{21}\rvert^2}{1-\lvert S_{11}\rvert^2}$ (for $\Gamma_L = 0$)
Available power gain with output conjugate matched	$G_A = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_G\rvert^2)}{\lvert 1-S_{11}\Gamma_G\rvert^2(1-\lvert S'_{22}\rvert^2)} = \dfrac{\lvert S_{21}\rvert^2}{1-\lvert S_{22}\rvert^2}$ (for $\Gamma_G = 0$)
Maximum available power gain[a]	$G_{ma} = \left\lvert\dfrac{S_{21}}{S_{12}}\right\rvert\,(k-\sqrt{k^2-1})$
Maximum unilateral transducer power gain	$G_{\mathrm{TUmax}} = \dfrac{\lvert S_{21}\rvert^2}{(1-\lvert S_{11}\rvert^2)(1-\lvert S_{22}\rvert^2)}$
Maximum stable power gain	$G_{ms} = \dfrac{\lvert S_{21}\rvert}{\lvert S_{12}\rvert}$
Unilateral power gain	$U = \dfrac{1/2\lvert S_{21}/S_{12}-1\rvert^2}{k\lvert S_{21}/S_{12}\rvert-\mathrm{Re}(S_{21}/S_{12})}$

[a]Definition for $G_{\max}$: If $k > 1$, it is called G_{ma}, and if $k < 1$, it is called G_{ms}.

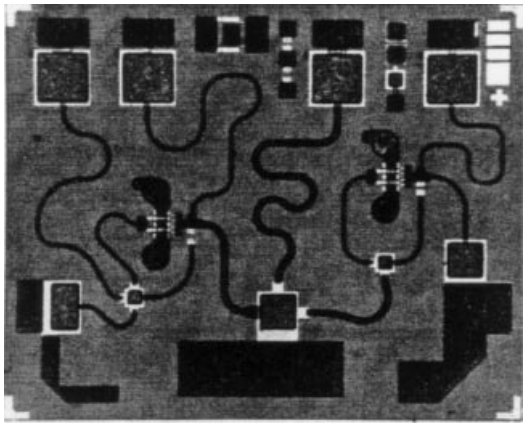


Figure 3-11 A two-stage, wideband feedback amplifier.

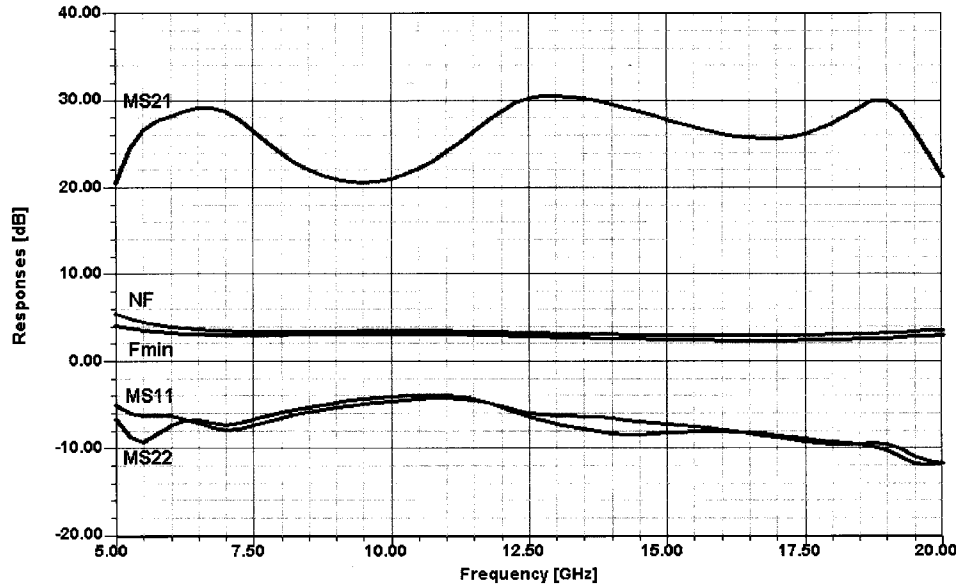


Figure 3-12 Frequency-dependent gain, matching, and noise performance for the amplifier in Figure 3-11.

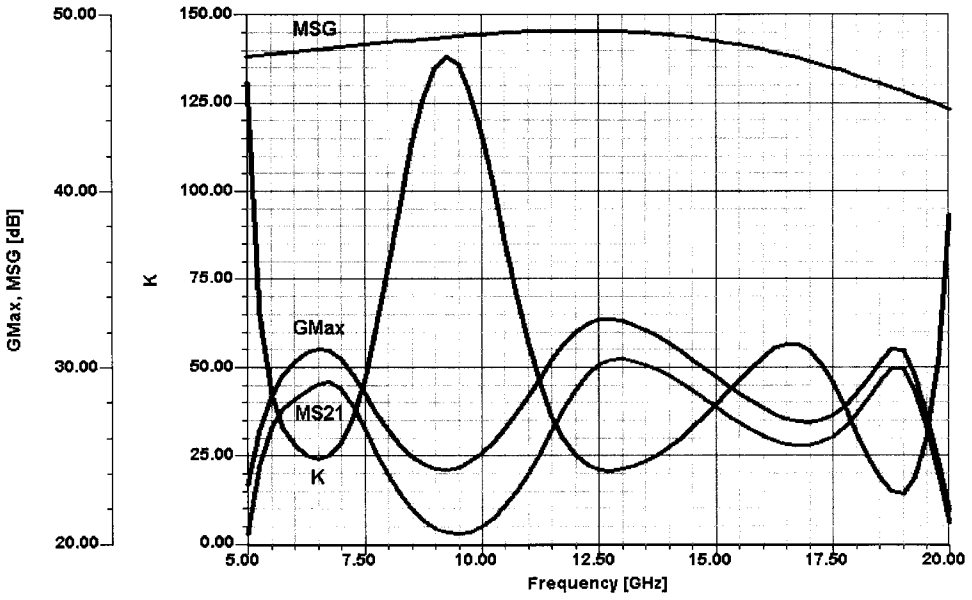


Figure 3-13 Frequency-dependent G_{max}, K, maximum stable gain, and gain performance for the amplifier in Figure 3-11.

motorboating phenomenon has been known for years, and the problem moves up in frequency with more modern technology and is largely dependent not only on the device and the circuit, but also on the layout. The motorboating effect has to do with ground loops and the fact that if the printed circuit traces from the power supply connection to the output power stage present too high a resistance, then at high currents the voltage will drop to a point below useful operation and then, as the volume is lowered, become stable again. In this case the amplitude of the signal through the amplifier controls the stability of the output stage. Needless to say, this most unpleasant effect applies to high-power stages for all output transistors, where we use Class B or C operation, meaning that the dc operating current as a function of output does not remain constant.

In terms of gain, there is also an important gain characteristic called *differential gain* and *differential phase*. These expressions were first used in television circuits, where linearity in both areas is an absolute must. The amplitude of the signal determines the black and white contrast, and the phase determines the color. As an example, transmitting an image of one of our presidents playing golf in California to the New York area is done by many microwave hops. Any multipath reception, such as that caused by reflection from airplanes, may have little effect on the signal amplitude, but the resulting phase shift can transform a green lawn to a red lawn if a total shift of 180° occurs. Given a particular operating point, if the 1-dB compression point of multiple amplifiers in the chain is approached or exceeded, their differential gain, which is defined by the IEEE as "the difference between (a) the ratio of the output of a small high-frequency sine-wave signal at two different levels of a low-frequency signal on which it is superimposed, and (b) unity," may increase. This mode of saturation charges all the capacitances of the amplifier transistors, causing huge phase shifts that result in the known problem of changing colors. The same applies, of course, in test equipment amplifiers, in which such changes cannot be tolerated. A somewhat similar, but not quite the same, problem is spectral regrowth [evaluated in terms of adjacent-channel power ratio (ACPR)], which is also due to nonlinearities. The question of linearity has already been covered in detail in Chapter 1. Needless to say, any overdrive condition also changes an amplifier's input and output SWR.

3-1-3 Noise Figure (NF)*

Introduction. In this section we look at the best way to describe noise in active devices (black boxes). Even when a two-port is linear, the output waveform will differ from the input, because of the failure to transmit all spectral components with equal gain (or attenuation) and delay. By careful design of the two-port, or by limitation of the bandwidth of the input waveform, such distortions can largely be avoided. However, noise generated within the two-port can still change the waveform of the output signal. In a linear passive two-port, noise arises only from the losses in the two-port; thermodynamic considerations indicate that such losses result in the random changes that we call noise. When the two-port contains active devices, such as transistors, there are other noise mechanisms present. A very important consideration in a system is the amount of noise that it adds to the transmitted signal. This is often judged by the ratio of the output signal power to the output noise power

*Section 3-1-3 is reprinted from Chapter 2 of George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*, copyright 1990 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

(S/N). The ratio of signal plus noise power to noise power [$(S + N)/N$] is generally easier to measure and approaches S/N when the signal is large.

In the evaluation of a two-port it is important to know the amount of noise added to a signal passing through it. An important parameter for expressing this characteristic is the noise factor. The signal energy coming from a generator or antenna is amplified or attenuated in passing from the input to the output of a two-port, as is the noise that accompanies the input signal energy. A system generally includes a cascade of two-port networks that constitute one overall two-port that amplifies the signal to a high-enough power level for its intended use. The *noise factor* of a system is defined as the ratio of signal-to-noise ratios available at input and output:

$$F = \frac{(S/N)_{\text{input}}}{(S/N)_{\text{output}}} \geq 1 \tag{3-2}$$

The noise figure (or factor) of a receiver is an easily measured quantity that describes the signal-to-noise ratio reduction of that receiver.

When this ratio of powers is converted to decibels, it is generally referred to as the *noise figure* rather than noise factor. Various conventions are used to distinguish the symbols used for noise factor and noise figure. Here we use F to represent the noise factor and NF to represent the noise figure, although the terms are usually used interchangeably.

For an amplifier with the power gain G, the noise factor can be rearranged as

$$F = \frac{S_i/N_i}{GS_i/G(N_i + N_a)} \tag{3-3}$$

when N_a is the additional noise power added by the amplifier referred to the input. This can be computed to be

$$F = 1 + N_a/N_i \tag{3-4}$$

The noise factor is often replaced by the noise figure (NF), which is defined in decibels as

$$\text{NF} = 10 \log_{10} F \tag{3-5}$$

In applications like satellite receivers, the noise factor becomes such a small number that it is inconvenient to work with. Many people have adopted the use of an equivalent noise temperature for a circuit to remedy this situation. Since the thermal noise power available from a resistor at temperature T_e is

$$N = kT_eB \tag{3-6}$$

where k is Boltzmann's constant (1.38×10^{-23} J/K), T_e is the effective temperature in kelvins, and B is the bandwidth in hertz. The equation above may be used to associate an effective noise temperature with circuits containing more than just thermal noise sources. This allows Eq. (3-4) to be written as

$$F = 1 + \frac{kT_eB}{kT_0B} = 1 + \frac{T_e}{T_0} \tag{3-7}$$

where T_e is the effective noise temperature of the circuit and T_0 is the temperature of the generator resistor in kelvins. The noise temperature T_e now characterizes our circuit noise contribution and can directly be related to the noise factor.

Assuming a reference noise temperature of 290 K (–273 + 290 = 17 °C), let us determine the noise temperature of the system with a noise factor of 2.6 (4.15 dB):

$$T_e = (2.6 - 1)\ (290) = 464 \text{ K}$$

This temperature T_e should not be confused with the environmental operating temperature T_0. It is quite common to operate low-noise amplifiers with T_e below 100 K at an ambient temperature of 290 K.

Signal-to-Noise Ratio. Let us consider the signal-to-noise ratio of power delivered from a generator to a load as shown in Figure 3-14. The signal power delivered to the input is given by

$$S_{\text{in}} = P_{\text{in}} = \frac{E_g^2 \text{Re}(Z_{\text{in}})}{|Z_g + Z_{\text{in}}|^2} \tag{3-8}$$

where E_g is the rms voltage of the input signal supplied to the system, and the noise power supplied to the input is expressed by

$$N_{\text{in}} = \frac{\overline{v_n^2} \text{Re}(Z_{\text{in}})}{|Z_g + Z_{\text{in}}|^2} \tag{3-9}$$

where the noise power at the input is provided by the noise energy of the real part of Z_g. The input impedance Z of the system in the form $Z = R_{\text{in}} + jX_{\text{in}}$ is assumed to be complex.

The Johnson noise of a resistor [here $\text{Re}(Z_g)$], given by the mean-square voltage

$$\overline{v_n^2} = 4kTRB \tag{3-10}$$

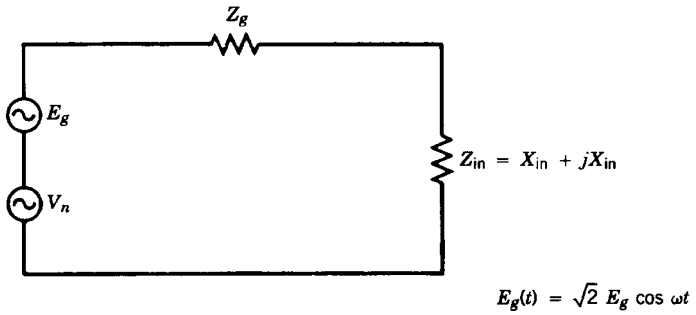


Figure 3-14 Combination of signal and noise voltages supplied to a complex termination.

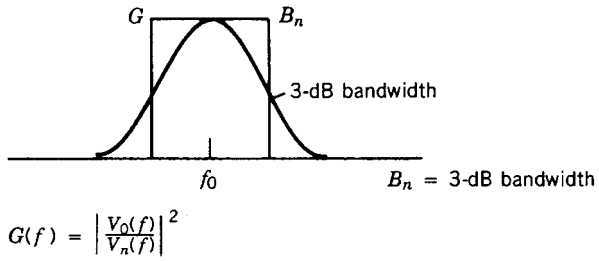
where k (Boltzmann's constant) = 1.38×10^{-23} J/K, T is the absolute temperature of the resistor, and B is the bandwidth, is sufficiently small that the resistive component of impedance does not change. The available signal power from the generator has a lower limit, even if the signal is attenuated by the highest possible attenuation. The generator resistor acts as a Johnson noise generator, its power being

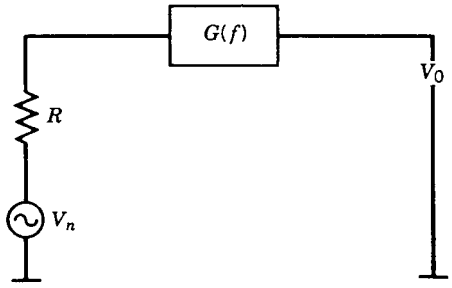
$$P_A = \frac{4kTRB}{4R} = kTB \tag{3-11}$$

where k is Boltzmann's constant, T is the absolute temperature, and B is the bandwidth. This power is the maximum available output power.

For an ambient temperature of 290 K, $kT = 4 \times 10^{-21}$ W/Hz. This expression is also given as kT = –204 dBW/Hz = –174 dBm/Hz = –114 dBm/MHz. We can combine Eqs. (3-8) to (3-10) to obtain

$$\left(\frac{S}{N}\right)_{\text{in}} = \frac{E_g^2}{4kT\text{Re}(Z_g)B} \tag{3-12}$$





$$V_0^2 = \int_0^\infty 4kT_0RG(f)\,dF$$
$$= 4kT_0R \int_0^\infty G(f)\,dF$$
$$B_n = \frac{1}{G} \int_0^\infty G(f)\,dF$$
B_n = noise bandwidth

Figure 3-15 Graphical and mathematical explanations of the noise bandwidth from a comparison of the Gaussian-shaped bandwidth to the rectangular filter response.

This is the value of *S/N* contributed by the generator, which does not include the noise generated by the load, in this case $\mathrm{Re}(Z_{\mathrm{in}})$, which would need to be included in the measurement of the total *S/N* across the input impedance.

A critical parameter is the noise bandwidth, B_n, which is defined as the equivalent bandwidth, as shown in Figure 3-15. For reasons of group delay correction, most practical filters have round rather than sharp corners. The noise figure measurements shown later can be used to determine the "integrated" bandwidth, which is B_n.

An active system such as a combination of amplifiers and mixers will add noise to the input signals, and the noise factor that describes this is defined as the *S/N* ratio at the input to the *S/N* ratio at the output, which is always greater than unity [2]. In practice, a certain minimum signal-to-noise ratio is required for operation. For example, in a communication system such a minimum is required for intelligible transmission, either voice or data. For high-performance TV reception, to provide a noise-free picture to the eye, a typical requirement is for a 60-dB *S/N*. In the case of a TV system, a large dynamic range is required, as well as a very large bandwidth to reproduce all colors truthfully and all shades from high-intensity white to black. Good systems will have a bandwidth of 8 MHz or more.

Noise Figure Measurements. Some of the noise equations are based on mathematical models and physics. To understand some of these expressions, it is useful to look at a practical system, with amplifiers, that has to be evaluated.

Let us look at Figure 3-16, which consists of a signal generator, the system or device under test (DUT), and a selective receiver with a built-in root-mean-square (rms) voltmeter to determine the signal and noise voltages. It is necessary that the system have enough gain so that the noise voltage supplied by the generator will be indicated [3].

If we assume that our selective receiver is a video noise meter calibrated in rms voltage levels, we can perform two measurements. With an input termination connected to the TV system (typically, 75 Ω for cable TV, 50 Ω for satellite TV), the noise receiver/meter will read a value for proper termination that can easily be calculated. Since one-half of the mean-square noise voltage appears across the input,

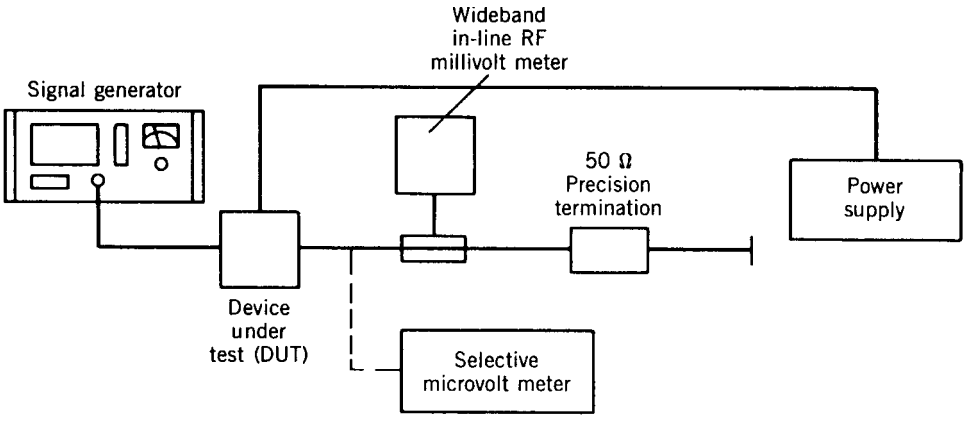


Figure 3-16 Test setup to measure signal-to-noise ratio.

$$v_{\text{in}} = \frac{v_n}{2} = \frac{\sqrt{4kTRB}}{2} \tag{3-13}$$

With $B = 10$ MHz, $T = 290$ K (T is always expressed in absolute temperature; $T_0 = -273$ °C), and $k = 1.38 \times 10^{-23}$ J/K, then for $R = 75\ \Omega$,

$$v_{\text{in}} = \frac{v_n}{2} = 1.73\ \mu\text{V} \tag{3-14}$$

where the rms noise voltage has been referred to the input port. We can verify this with our first measurement.

Now we increase the input voltage of the signal generator to a value that indicates a 60-dB *S/N* ratio at the output port. This should be about

$$E_g = \frac{v_n}{2}\sqrt{F} \times 1000 = 1.73\sqrt{F} \quad \text{mV}$$

where F is the noise factor of the receiver. For a receiver noise factor of 10, we would obtain $E_g = 5.48$ mV (rms value). If the noise energy equivalent to a noise factor of F is assumed, we need $\sqrt{F}$ times more voltage. For a 60-dB ratio, this means that $E_g = 1000 \times (v_n/2) \times \sqrt{F}$.

As they are done here, over a power range of 60 dB, the measurements can be performed over such a wide range only if special equipment is available. In cases where the internal detector of a piece of communication equipment is used, the signal-to-noise ratio measurements are performed over much smaller power ranges.

Let us assume that for the above-mentioned case ($F = 10$) we find a *S/N* ratio of 10 dB at the output for an input signal of 5.47 μV. By rewriting Eq. (3-7) as

$$E_g = \frac{v_n}{2}\sqrt{F} = \sqrt{KTRBF} \tag{3-15}$$

with F being the noise factor, we can solve for F with

$$F = \frac{P_s}{P_n} = \frac{E_g^2/R}{kTB} \tag{3-16}$$

While the input power from the thermal energy of the input termination resistor was $kTB = 4 \times 10^{-4}$ W, the input power required for the 10-dB *S/N* ratio was

$$P_s = \frac{(5.47 \times 10^{-6})^2}{75} = 3.98 \times 10^{-13}\ \text{W} \tag{3-17}$$

The noise factor is defined as the ratio P_s/P_n:

$$F = \frac{3.98 \times 10^{-13}}{4 \times 10^{-14}} = 10 \tag{3-18}$$

which is the proof.

This method is used more frequently at the 3-dB point, or double the input power if the dynamic range of the detector is small or only a linear indicator is available. Because of hum and other pickup, this is not an easy measurement. Using a signal generator is very expensive because in a laboratory or production environment a wide frequency range requires several generators.

Another method is the use of a wideband noise generator. Modern gas discharge diodes or avalanche diodes are available that provide essentially white noise energy over a large frequency range. These microwave diodes typically have an output of 30 dB above kT when switched on and kT when switched off. To provide good matching at microwave frequencies, a 15-dB attenuator is cascaded. This means that the noise power of the source in the ON condition is about 15 dB above kT.

In the early 1960s, low-cost noise figure test equipment was built around vacuum diodes whose operating range was limited to 1200 MHz due to the resonance effects of the structure. Today, automatic noise gain analyzers use calibrated solid-state noise sources up to 26.5 GHz. It appears that the upper frequency limit has to do with matching, and the lower-frequency limit with $1/f$ noise.

Noisy Two-Port Description. Based on the convention by Rothe and Dalke [4], any linear two-port can be in the form shown in Figure 3-17. This general case of a noisy two-port can be redrawn showing noise sources at the input and at the output. Figure 3-17b shows this in admittance form and Figure 3-17c in impedance form. The internal noise sources are assumed to produce very small currents and voltages, and we assume that linear two-port equations are valid. The internal noise contributions have been expressed by using external noise sources:

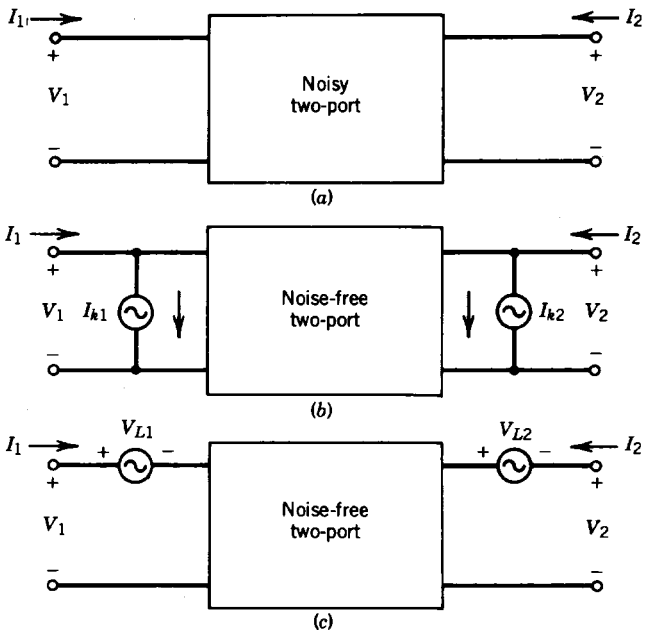


Figure 3-17 Noisy linear two-ports: (a) general form, (b) admittance form, and (c) impedance form.

$$I_1 = y_{11}V_1 + y_{12}V_2 + I_{K1}$$
$$I_2 = y_{21}V_1 + y_{22}V_2 + I_{K2} \tag{3-19}$$

$$V_1 = z_{11}I_1 + z_{12}I_2 + V_{L1}$$
$$V_2 = z_{21}I_1 + z_{22}I_2 + V_{L2} \tag{3-20}$$

where the external noise sources are I_{K1}, I_{K2}, V_{L1}, and V_{L2}.

Since we want to describe our noisy circuit in terms of the noise figure, the *ABCD*-matrix description will be more convenient since it refers both noise sources to the input of the two-port [5]. This representation is given below (note the change in direction of I_2):

$$V_1 = AV_2 + BI_2 + V_A$$
$$I_1 = CV_2 + DI_2 + I_A \tag{3-21}$$

where V_A and I_A are the external noise sources.

It is important to remember that all of these matrix representations are interrelated. For example, the noise sources for the *ABCD*-matrix description can be obtained from the *z*-matrix representation shown in Eq. (3-20). This transformation is

$$V_A = -\frac{I_{K2}}{y_{21}} = V_{L1} - \frac{V_{L2}z_{11}}{z_{21}} \tag{3-22}$$

$$I_A = I_{K1} - \frac{I_{K2}y_{11}}{y_{21}} = -\frac{V_{L2}}{z_{21}} \tag{3-23}$$

The *ABCD* representation is particularly useful because it allows us to define a noise temperature for the two-port referenced to the input. The two-port itself (shown in Figure 3-18) is assumed to be noise-free.

In the past, *z* and *y* parameters have been used, but in microwave applications it has become common to use *S*-parameter definitions. This is shown in Figure 3-19. The previous equations can be rewritten in their new form using *S*-parameters:

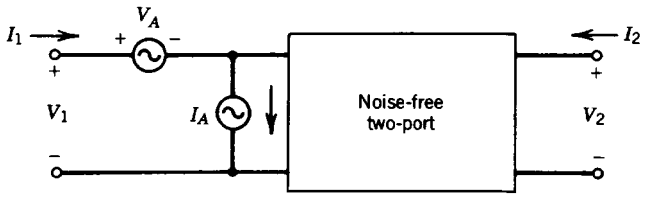


Figure 3-18 Chain-matrix form of linear noisy two-ports.

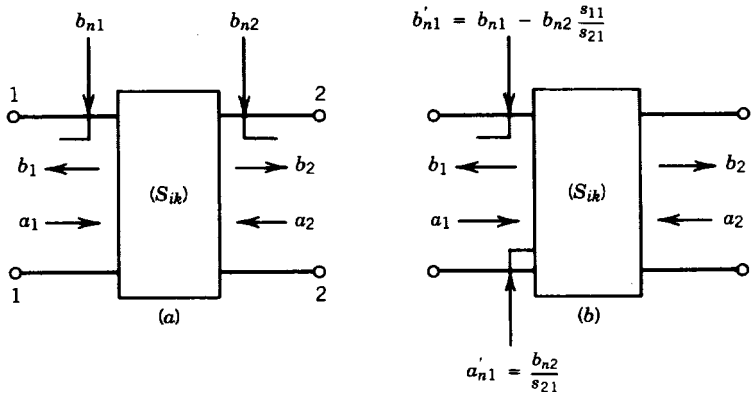


Figure 3-19 *S*-parameter form of linear noisy two-ports.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_{n1} \\ b_{n2} \end{bmatrix} \tag{3-24}$$

There are different physical origins for the various sources of noise. Typically, thermal noise is generated by resistances and loss in the circuit or transistor, whereas shot noise is generated by current flowing through semiconductor junctions and vacuum tubes. Since these many sources of noise are represented by only two noise sources at the device input, the two equivalent noise sources are often a complicated combination of the circuit internal noise sources. Often, some fraction of V_A and I_A is related to the same noise source. This means that V_A and I_A are not independent in general. Before we can use V_A and I_A to calculate the noise figure of the two-port, we must calculate the correlation between the V_A and I_A shown in Figure 3-18.

The noise source V_A represents all the device noise referred to the input when the generator impedance is zero; that is, when the input is short-circuited. The noise source I_A represents all the device noise referred to the input when the generator admittance is zero; that is, the input is open-circuited.

The correlation of these two noise sources considerably complicates analysis. By defining a correlation admittance, we can simplify the mathematics and get some physical intuition for the relationship between noise figure and generator admittance. Since some fraction of I_A will be correlated with V_A, we split I_A into correlated and uncorrelated parts as follows:

$$I_A = I_n + I_u \tag{3-25}$$

I_u is the part of I_A uncorrelated with V_A. Since I_n is correlated with V_A, we can say that I_n is proportional to V_A and the constant of proportionality is the correlation admittance:

$$I_n = Y_{\text{cor}} V_A \tag{3-26}$$

This leads us to

$$I_A = Y_{\text{cor}} V_A + I_u \tag{3-27}$$

The following derivation of noise figure will use the correlation admittance. Y_{cor} is not a physical component located somewhere in the circuit. Y_{cor} is a complex number derived by correlating the random variables I_A and V_A. To calculate Y_{cor}, we multiply each side of Eq. (3-27) by V_A^* and average the result. This gives

$$\overline{V_A^* I_A} = Y_{\text{cor}} \overline{V_A^2} \tag{3-28}$$

where the I_u term averaged to zero since it was uncorrelated with V_A. The correlation admittance is thus given by

$$Y_{\text{cor}} = \frac{\overline{V_A^* I_A}}{V_A^2} \tag{3-29}$$

Often, people use the term "correlation coefficient." This normalized quantity is defined as

$$c = \frac{\overline{V_A^* I_A}}{\sqrt{\overline{V_A^2} I_A^2}} = Y_{\text{cor}} \sqrt{\frac{\overline{V_A^2}}{I_A^2}} \tag{3-30}$$

Note that the dual of this admittance description is the impedance description. Thus, the impedance representation has the same equations as above with Y replaced by Z, I replaced by V, and V replaced by I.

V_A and I_A represent internal noise sources in the form of a voltage source acting in series with the input voltage and a source of current flowing in parallel with the input current. This representation conveniently leads to the four noise parameters needed to describe the noise performance of the two-port. Again using the Nyquist formula, the open-circuit voltage of a resistor at the temperature T is

$$\overline{V_A^2} = 4kTRB \tag{3-31}$$

This voltage is a mean-square fluctuation (or spectral density). It is the method used to calculate the noise identity. We could also define a noise equivalent resistance for a noise voltage as

$$R_n = \frac{\overline{V_A^2}}{4kTB} \tag{3-32}$$

The resistor R_n is not a physical resistor but can be used to simulate different portions of the noise equivalent circuit.

In a similar manner a mean-square current fluctuation can be represented in terms of an equivalent noise conductance G_n, which is defined by

$$G_n = \frac{\overline{I_A^2}}{4kTB} \tag{3-33}$$

and

$$G_u = \frac{\overline{I_u^2}}{4kTB} \tag{3-34}$$

for the case of the uncorrelated noise component. The input generator to the two-port has a similar contribution:

$$G_G = \frac{I_G^2}{4kTB} \tag{3-35}$$

with Y_G being the generator admittance and G_G being the real part. With the definition of F above, we can write

$$F = 1 + \left| \frac{I_A + Y_G V_A}{I_G} \right|^2 \tag{3-36}$$

The use of the voltage V_A and the current I_A has allowed us to combine all the effects of the internal noise sources.

We can use the previously defined [Eq. (3-29)] correlation admittance, $Y_{\text{cor}} = G_{\text{cor}} + jB_{\text{cor}}$, to simplify Eq. (3-36). First, we determine the total noise current:

$$\overline{I_A^2} = 4kT(|Y_{\text{cor}}|^2 R_n + G_u)B \tag{3-37}$$

where R_n and G_u are defined in Eqs. (3-32) and (3-34). The noise factor can now be determined:

$$F = 1 + \frac{G_u}{G_g} + \frac{R_n}{G_g}[(G_G + G_{\text{cor}})^2 + (B_G + B_{\text{cor}})^2] \tag{3-38}$$

$$F = 1 + \frac{R_u}{R_g} + \frac{G_n}{R_g}[(R_G + R_{\text{cor}})^2 + (X_G + X_{\text{cor}})^2] \tag{3-39}$$

The noise factor is a function of various elements, and the optimum impedance for the best noise figure can be determined by minimizing F with respect to generator reactance and resistance. This gives

$$R_{0n} = \sqrt{\frac{R_u}{G_n} + R_{\text{cor}}^2} \tag{3-40}$$

$$X_{0n} = -X_{\text{cor}} \tag{3-41}$$

and

$$F_{\min} = 1 + 2G_n R_{\text{cor}} + 2\sqrt{R_u G_n + (G_n R_{\text{cor}})^2} \tag{3-42}$$

(To distinguish between optimum noise and optimum power, we have introduced the convention 0*n* instead of the more familiar abbreviation *opt*.) At this point we see that the optimum condition for minimum noise figure is not a conjugate power match at the input port. We can explain this by recognizing that the noise source V_A and I_A represent all the

two-port noise, not just the thermal noise of the input port. We should observe that the optimum generator susceptance, $-X_{corr}$, will minimize the noise contribution of the two noise generators.

In rearranging for conversion to S parameters, we write

$$F = F_{\min} + \frac{g_n}{R_G}|Z_G - Z_{0n}|^2 \tag{3-43}$$

$$F = F_{\min} + \frac{R_n}{G_G}|Y_G - Y_{0n}|^2 \tag{3-44}$$

From the definition of the reflection coefficient,

$$\Gamma_G = \frac{Y_0 - Y_G}{Y_0 + Y_G} \tag{3-45}$$

and with

$$g_G = \frac{G_G}{Y_0} \tag{3-46}$$

$$r_n = \frac{R_n}{Z_0} \tag{3-47}$$

the normalized equivalent noise resistance,

$$F = F_{\min} + \frac{4r_n|\Gamma_G - \Gamma_{0n}|^2}{g_G(1 - |\Gamma|^2)|1 + \Gamma_{0n}|^2} \tag{3-48}$$

$$r_n = (F_{50} - F_{\min})\frac{|1 + \Gamma_{0n}|^2}{4|\Gamma_{0n}|^2} \tag{3-49}$$

$$\Gamma_{0n} = \frac{Z_{0n} - Z_0}{Z_{0n} + Z_0} \tag{3-50}$$

The noise performance of any linear two-port can now be determined if the values of the four noise parameters, $F_{\min}$, r_n (= $R_n/50$), and Γ_{0n} (magnitude and phase), are known.

Figure 3-20 shows the noise factor of a high-frequency transistor as a function of B_g for G_g = constant and as a function of G_g for $B_g = B_{opt}$.

Noise Figure of Cascaded Networks. In a system with many circuits connected in cascade (Figure 3-21), we must consider the contributions of the various circuits. In considering the equivalent noise resistor R_n in series with the input circuit,

$$F = \frac{R_G + R_n}{R_G} \tag{3-51}$$

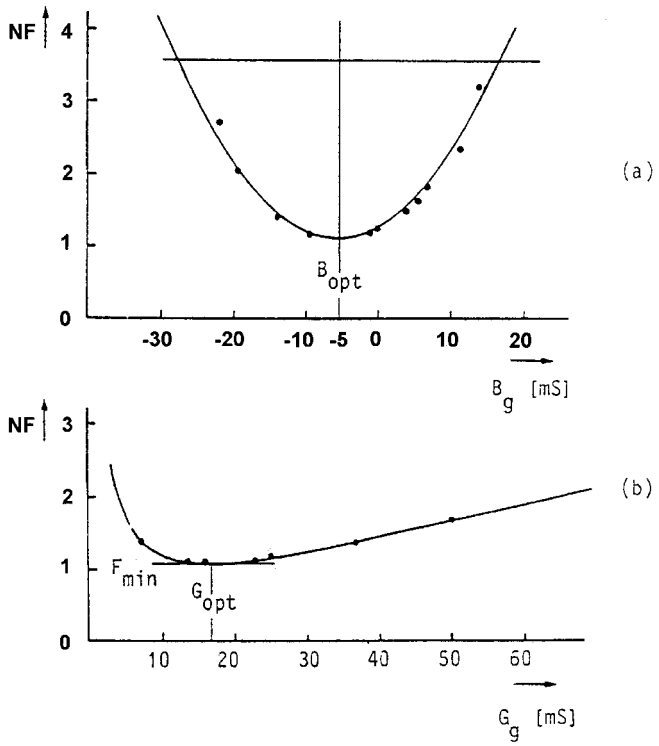


Figure 3-20 Noise factor in high-frequency BJTs for $f = 600$ MHz: (a) as a function of B_g for G_g = constant and (b) as a function of G_g for $B_g = B_{opt}$.

$$F = 1 + \frac{R_n}{R_G} \tag{3-52}$$

The excess noise added by the circuit is R_n/R_G.

In considering two cascaded circuits a and b, by definition the available noise at the output of b is

$$N_{ab} = F_{ab} G_{ab} kTB \tag{3-53}$$

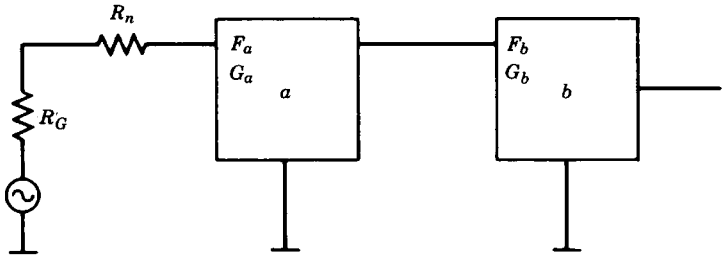


Figure 3-21 Cascaded noisy two-ports with the noise factors F_a and F_b and the gain figures G_a and G_b.

with B the equivalent noise bandwidth in which the noise is measured. The total available gain G is the product of the individual available gains, so

$$N_{ab} = F_{ab} G_a G_b kTB \tag{3-54}$$

The available noise from network a at the output of network b is

$$N_{a/b} = N_a G_b = F_a G_a G_b kTB \tag{3-55}$$

The available noise added by network b (its excess noise) is

$$N_{b/b} = (F_b - 1) G_b kTB \tag{3-56}$$

The total available noise N_{ab} is the sum of the available noise contributed by the two networks:

$$\begin{aligned} N_{ab} &= N_{a/b} + N_{b/b} = F_a G_a G_b kTB + (F_b - 1) G_b kTB \\ &= \left(F_a + \frac{F_b - 1}{G_a} \right) G_a G_b kTB \end{aligned} \tag{3-57}$$

$$F_{ab} = F_a + \frac{F_b - 1}{G_a} \tag{3-58}$$

For any number of circuits, this can be extended to

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \cdots \tag{3-59}$$

When considering a long chain of cascaded amplifiers, there will be a minimum noise figure achievable for this chain. This is a figure of merit and was proposed by Haus and Adler [6]. It is calculated by rearranging Eq. (3-59):

$$(F_{\text{tot}})_{\min} = (F_{\min} - 1) + \frac{F_{\min} - 1}{G_A} + \frac{F_{\min} - 1}{G_A^2} + \cdots + 1 \tag{3-60}$$

where $F_{\min}$ is the minimum noise figure for each stage and G_A is the available power gain of the identical stages. Using

$$\frac{1}{1 - X} = 1 + X + X^2 + \cdots \tag{3-61}$$

we find a quantity $(F_{\text{tot}} - 1)$, which is defined as the noise measure M. The minimum noise measure

$$(F_{\text{tot}})_{\min} - 1 = \frac{F_{\min} - 1}{1 - 1/G_A} = M_{\min} \tag{3-62}$$

refers to the noise of an infinite chain of optimally tuned, low-noise stages, so it represents a lower limit on the noise of an amplifier.

The minimum noise measure $M_{\min}$ is an invariant parameter and is not affected by feedback. It is somewhat similar to a gain–bandwidth product, in its use as a system invariant. The minimum noise measure is achieved when the amplifier is tuned for the available power gain and $\Gamma_G = \Gamma_{0n}$, given by Eq. (3-50).

Influence of External Parasitic Elements. Mounting an active two-port such as a transistor usually adds stray capacitance and lead inductance to the device, as shown in Figure 3-22. These external components consisting of transmission lines and parasitic reactances modify the noise parameters and the gain. Some researchers have published the results of these parasitic effects and have made manual computations or used some limited computer programs.

In a paper by Fukui [7], an attempt was made to determine the necessary equations, but the formulas are too involved even for pocket calculators. A more generic study by Iversen [8] is also very involved because of the various matrix manipulations and is more suitable for a computer. Besser's paper [9] in the *IEEE MTT-S* in 1975 and Vendelin's paper [10] in the same issue have shown for the first time some practical results using computers and even optimization methods, using an early version of COMPACT. The intention of these investigations was to find feedback that modifies the device noise and scattering parameters such that a noise match could also provide a low-input VSWR. It can be seen from these discussions that some feedback, besides resulting in some gain reduction, may improve the noise matching at the input for a limited frequency range.

A more recent paper by Suter [11] based on a report by Hartmann and Strutt [12] has given a simple transformation starting from the *S*-parameters and the noise parameters from common-source (or common-emitter) measurements. The noise parameters for the "packaged" device are calculated. This means that the parameters for the "new" device, including

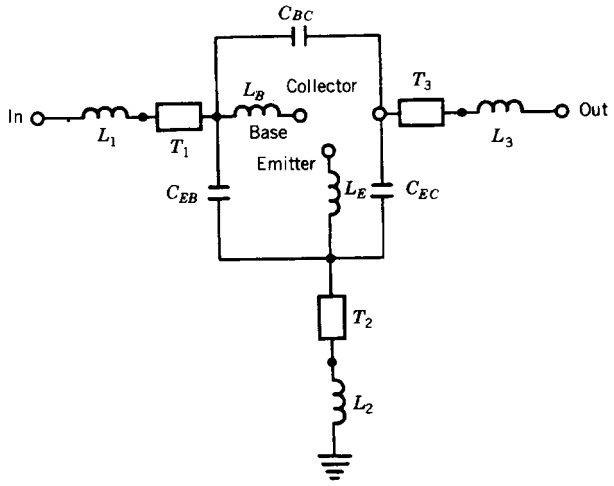
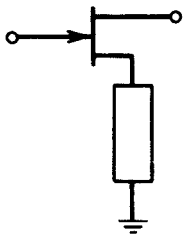


Figure 3-22 Equivalent circuit of the transistor package.

Table 3-2 Transformation matrix parameters

Series Feedback



$$Z_S = R_S + jX_S = Z_S$$

$$[n] = \begin{bmatrix} n_{11} = 1 & n_{12} = Z_0 \dfrac{S_{21}M - S'_{21}N}{S_{21}C'_1 + S'_{21}C_1}\ \Omega \\ n_{21} = 0 & n_{22} = \dfrac{S_{21}C'_1}{S_{21}C'_1 + S'_{21}C_1} \end{bmatrix}$$

where

$$S'_{11} = S'_{22} = \frac{-1}{1 + 2Z_S}$$

$$S'_{12} = S'_{21} = \frac{2Z_S}{1 + 2Z_S}$$

$$M = (1 + S'_{11})(1 - S'_{22}) + S'_{12}S'_{21}$$

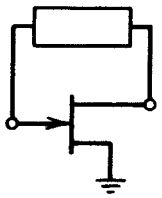
$$N = (1 + S_{11})(1 - S_{22}) + S_{12}S_{21}$$

$$C_1 = (1 - S_{11})(1 - S_{22}) - S_{12}S_{21}$$

$$C'_1 = (1 + S'_{11})(1 - S'_{22}) - S'_{12}S'_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

Shunt Feedback



$$Z_P = R_P + jX_P = Z_P$$

$$[n] = \begin{bmatrix} n_{11} = \dfrac{S_{21}C'_2}{S_{21}C'_2 + S'_{21}C_2} & n_{12} = 0\ \Omega \\ n_{21} = \dfrac{1}{Z_0}\dfrac{S_{21}P - S'_{21}Q}{S_{21}C'_2 + S'_{21}C_2}\,S & n_{22} = 1 \end{bmatrix}$$

where

$$S'_{11} = S'_{22} = \frac{Z_P}{2 + Z_P}$$

(*continued*)

Table 3-2 (*Continued*)

$$S'_{12} = S'_{21} = \frac{2}{2 + Z_P}$$

$$P = (1 - S'_{11})(1 + S'_{22}) + S'_{12}S'_{21}$$

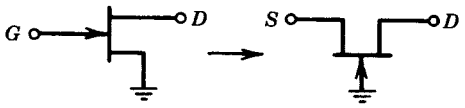
$$Q = (1 - S_{11})(1 + S_{22}) + S_{12}S_{21}$$

$$C_2 = (1 + S_{11})(1 + S_{22}) - S_{12}S_{21}$$

$$C'_2 = (1 + S'_{11})(1 + S'_{22}) - S'_{12}S'_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

Common Gate



$$[n] = \begin{bmatrix} n_{11} = \dfrac{2S_{21}}{-2S_{21} + C_4} & n_{12} = 0\ \Omega \\ n_{21} = \dfrac{1}{Z_0}\dfrac{C_3C_4 - 4S_{12}S_{21}}{V(-2S_{21} + C_4)}S & n_{22} = -1 \end{bmatrix}$$

where

$$V = (1 + S_{11})(1 + S_{22}) - S_{12}S_{21}$$

$$C_3 = (1 - S_{11})(1 + S_{22}) + S_{12}S_{21}$$

$$C_4 = (1 + S_{11})(1 - S_{22}) + S_{12}S_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \text{common-source } S \text{ parameters}$$

the common-gate (or common-base) case, are calculated. The equations are device dependent. They are valid for any active two-port.

A transformation matrix, n, may be used to combine the noise sources of the various circuit configurations. The transformation matrix parameters are given in Table 3-2 for series feedback, shunt feedback, and the common-gate (base) case, which will be important for oscillator analysis. The transformation matrix gives the new noise parameters as follows:

$$R'_n = R_n|n_{11} + n_{12}Y_{\text{cor}}|^2 + G_n|n_{12}|^2 \tag{3-63}$$

$$G'_n = \frac{G_nR_n}{R'_n}|n_{11}n_{22} - n_{12}n_{21}|^2 \tag{3-64}$$

$$Y'_{\text{cor}} = \frac{R_n}{R'_n}(n_{21} + n_{22}Y_{\text{cor}})(n^*_{11} + n^*_{12}Y^*_{\text{cor}}) + \frac{G_n}{R'_n}n_{22}n^*_{12} \tag{3-65}$$

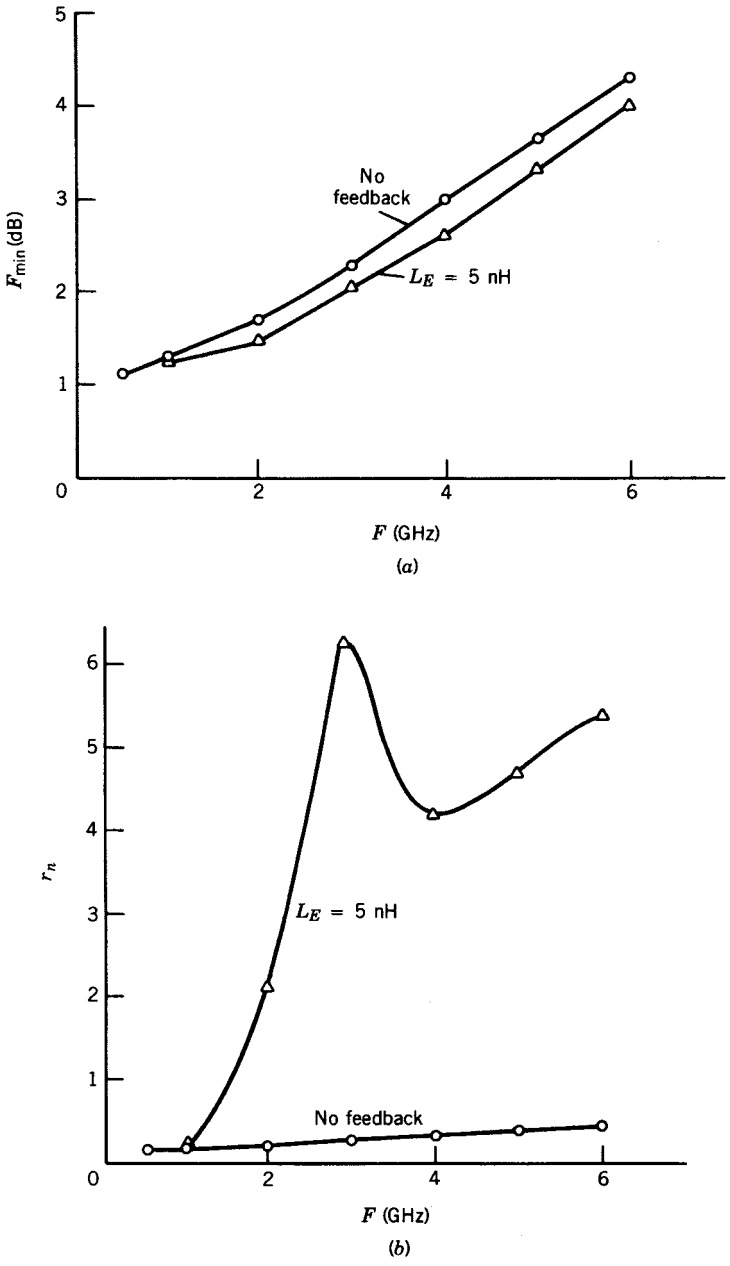


Figure 3-23 Noise parameters versus feedback for the AT-41435 silicon bipolar transistor: (a) F_{min} versus frequency and feedback; (b) r_n versus frequency and feedback; (c) Γ_{0n} versus frequency and feedback.

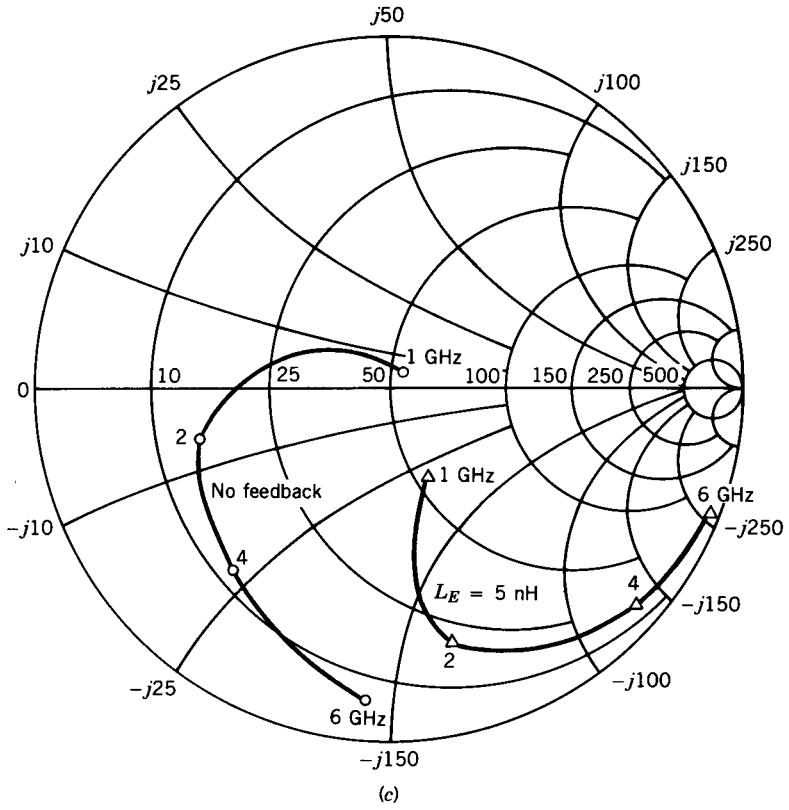


Figure 3-23 (*Continued*)

A final transformation to the more common noise-parameter format given by Eq. (3-44) is still needed [10]:

$$F_{\min} = 1 + 2R_n'(G_{\text{cor}}' + G_{0n}') \tag{3-66}$$

$$R_n = R_n' \tag{3-67}$$

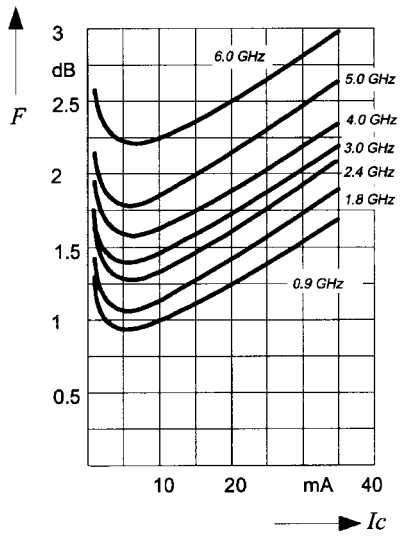
$$G_{0n} = \sqrt{\frac{G_n'}{R_n'} + G_{\text{cor}}'^2} \tag{3-68}$$

$$B_{0n} = -B_{\text{cor}}' \tag{3-69}$$

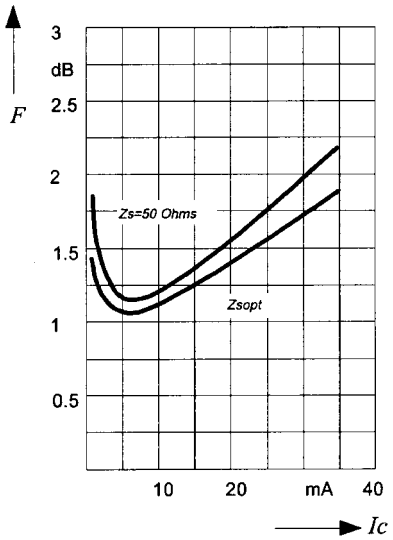
Figure 3-23 shows the noise figure as a function of external feedback for a low-noise microwave bipolar transistor, the AT-41435.

Bias-Dependent Noise Parameters. Since we have elaborated on the effect of parasitics, it is now useful to point out that noise is also a strong function of bias point and frequency. Figure 3-24 shows the noise figure expressed in dB as a function of different dc currents and frequencies for a BFP420 BJT by Siemens. A change in the collector voltage will also affect the noise parameters, not quite as strongly as shown here, but still significantly. The reason

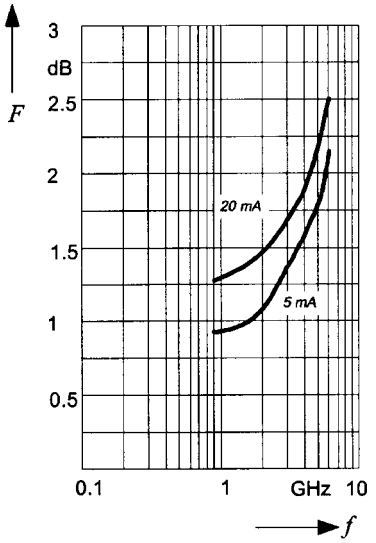
Noise Figure
versus Collector Current
V_{CE} = 2 V, $Z_S = Z_{Sopt}$



Noise Figure
versus Collector Current
VCE = 2 V, f = 1.8 GHz



Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA,
$Z_S = Z_{Sopt}$



Source Impedance for min.
Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA

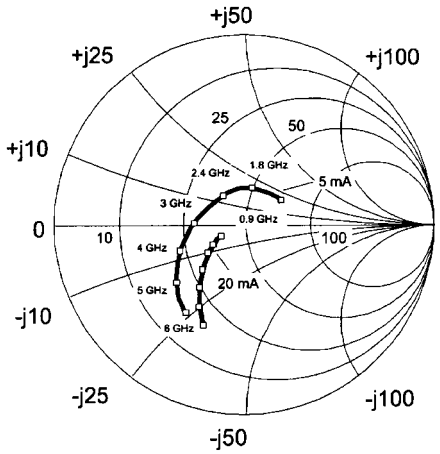


Figure 3-24 Noise figure versus collector current, noise figure versus frequency, and source impedance for minimum noise figure versus frequency for the Siemens BFP420 transistor.

Table 3-3 BFP420 common-emitter noise parameters

f	$F_{\min}{}^a$	$G_a{}^a$	Γ_{opt}		R_N	r_n	$F_{50\Omega}{}^b$	$\lvert S_{21} \rvert^{2b}$
(GHz)	dB	(dB)	(Magnitude)	(Angle)	(Ω)		(dB)	(dB)
V_{CE} = 2 V, I_C = 5 mA								
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	–175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	–157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	–142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	–123.0	15.0	0.30	3.30	5.3

[a]Input matched for minimum noise figure, output for maximum gain.
[b]$Z_S = Z_L = 50\ \Omega$.

for the turning point of the noise figure has to do with the fact that for lower currents the f_T cutoff frequency has not reached its peak yet, and therefore the noise increases. The increase of the noise at higher currents is due to the Schottky noise calculated from the equation

$$P_n = 2Iq \tag{3-70}$$

where P_n is the noise power, I is the saturation, collector, or drain current, and q is the charge of an electron (1.60×10^{-19} coulomb).

Table 3-3 shows the common-emitter noise parameters for the BFP420.

Noise Circles. From the section entitled "Influence of External Parasitic Elements," we see that the noise factor is a strong function of the generator admittance (or impedance) presented to the input terminals of the noisy two-port. Noise tuning is used to change the values of the input admittance to obtain the best noise performance. There is a range of values of input reflection coefficient over which the noise figure is constant. In plotting these points of constant noise figure, we obtain the so-called noise circles, which can be drawn on the Smith chart Γ_G plane [13]. Starting with the noise equation [see Eq. (3-44)] for a 50-Ω generator impedance, we find that

$$F_{50} = F_{\min} + 4r_n \frac{|\Gamma_{0n}|^2}{|1 + \Gamma_{0n}|^2} \tag{3-71}$$

We want to find the position of the reflection coefficient on the Smith chart, as in the case of the gain circles, for which F is constant. First we rearrange Eq. (3-67) to read

$$r_n = (F_{50} - F_{\min}) \frac{|1 + \Gamma_{0n}|^2}{4|\Gamma_{0n}|^2} \tag{3-72}$$

By introducing

$$N_i = \frac{F_i - F_{\min}}{4r_n} |1 + \Gamma_{0n}|^2 \tag{3-73}$$

we can find an expression for a circle of constant noise figure as introduced by Rothe and Dahlke [4] (see also Fukui [13]). The center for the noise circle is

$$C_i = \frac{|\Gamma_{0n}|}{1 + N_i} \tag{3-74}$$

and the radius is

$$r_i = \frac{\sqrt{N_i^2 + N_i(1 - |\Gamma_{0n}|^2)}}{1 + N_i} \tag{3-75}$$

with the definition of N used previously. However, if we consider only the minimum noise figure for a given device, we will not obtain the minimum noise figure for the multistage amplifier system. This was explained when the noise measure was introduced. (See Figure 3-25.) Therefore, a better way to design the amplifier would be to use circles at constant noise measure instead of circles of constant noise figure. This was recently done by Poole

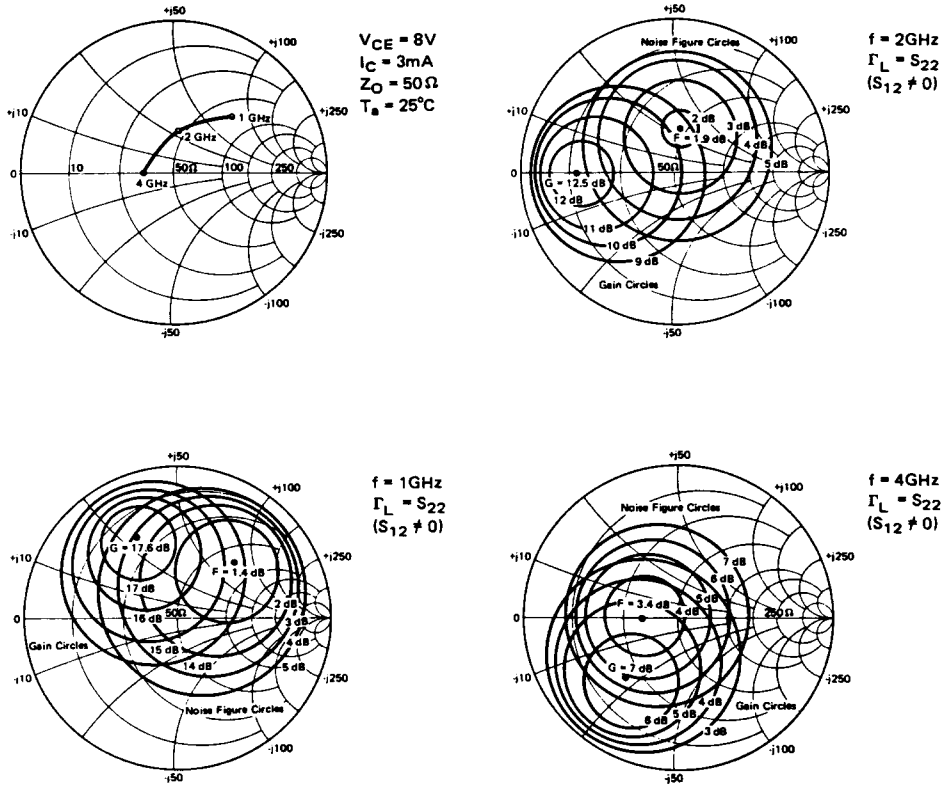


Figure 3-25 Typical noise figure circles and gain circles.

and Paul [14]. They derived the expressions for the noise measure circles as a function of S parameters, noise parameters, and Γ_G, using

$$G_A = \frac{|S_{21}|^2(1 - |\Gamma_G|^2)}{(1 - |S_{22}|^2) + |\Gamma_G|^2 (|S_{11}|^2 - |\Delta|^2) - 2\,\mathrm{Re}(\Gamma_G C_1)} \tag{3-76}$$

where

$$C_1 = S_{11} - S_{22}^* \Delta$$

$$\Delta = S_{11}S_{22} - S_{12}S_{11}$$

In terms of the generator reflection coefficient, the noise measure can be expressed as

$$M = [(F_{\min} - 1)(1 - |\Gamma_G|^2)|S_{21}|^2|1 + \Gamma_{0n}|^2 + 4r_n|S_{21}|^2|\Gamma_G - \Gamma_{0n}|^2]$$

$$\times \left\{|1 + \Gamma_{0n}|^2[|S_{21}|^2(1 - |\Gamma_G|^2) - (1 - |S_{22}|^2) - |\Gamma_G|^2(|S_{11}|^2 - |\Delta|^2) + 2\,\mathrm{Re}(\Gamma_G C_1)]\right\}^{-1} \tag{3-77}$$

Equation (3-77) can be shown to represent circles in the source reflection coefficient plane described by the following:

$$|\Gamma_G|^2 + |\Gamma_m|^2 - \Gamma_G\Gamma_m^* - \Gamma_G^*\Gamma_m = r_m^2 \tag{3-78}$$

The centers and radii of the constant-noise-measure circles are given by

$$C_m = \frac{M|1 + \Gamma_{0n}|^2 C_1^* + 4r_n|S_{21}|^2\Gamma_{0n}}{M|1 + \Gamma_{0n}|^2 P + |S_{21}|^2(4r_n - W)} \tag{3-79}$$

$$r_m = \frac{\sqrt{M^2M_a + MM_b + M_c}}{M|1 + \Gamma_{0n}|^2 P + |S_{21}|^2(4r_n - W)} \tag{3-80}$$

where

$$P = |S_{21}|^2 + |S_{11}|^2 - |\Delta|^2$$

$$Q = |S_{21}|^2 + |S_{22}|^2 - 1$$

$$W = |1 + \Gamma_{0n}|^2(F_{\min} - 1)$$

$$M_a = |1 + \Gamma_{0n}|^4(PQ + |C_1|^2)$$

$$M_b = |1 + \Gamma_{0n}|^2|S_{21}|^2[8r_n\,\mathrm{Re}(\Gamma_{0n}C_1)$$

$$- (4r_n|\Gamma_{0n}|^2 + W)P - (W - 4r_n)Q]$$

$$M_c = |S_{21}|^4[W - 4r_n(1 - |\Gamma_{0n}|^2)]$$

The value of the minimum noise measure can be found by considering the noise measure circle of zero radius; that is, set r_m equal to zero in Eq. (3-80). This results in

$$M_{\min} = \frac{-M_b \pm \sqrt{M_b^2 - 4M_a M_c}}{2M_a} \tag{3-81}$$

Equation (3-81) yields the same value of $M_{\min}$ as would have been obtained by using the immittance parameter equation given by Fukui [13] and can, therefore, be considered as the reflection coefficient plane analog of Fukui's expression. The elimination of the need for the parameter R_{eg}, however, results in a considerable simplification as compared with the earlier approach [13].

The value of the minimum noise measure is taken as the smallest nonnegative value of $M_{\min}$ given by Eq. (3-81). The source reflection coefficient that results in the minimum noise measure can now be obtained by employing Eq. (3-79):

$$\Gamma_{0m} = \frac{M_{\min}|1 + \Gamma_{0n}|^2 C_1^* + 4r_n|S_{21}|^2\Gamma_{0n}}{M_{\min}|1 + \Gamma_{0n}|^2 P + |S_{21}|^2\,(4r_n - W)} \tag{3-82}$$

The output reflection coefficient of the device, when Γ_{0m} is presented to the input port, is given by

$$S'_{22} = \frac{S_{22} - \Delta\Gamma_{0m}}{1 - S_{11}\Gamma_{0m}} \tag{3-83}$$

Noise Correlation in Linear Two-Ports Using Correlation Matrices. In the introduction to two-port noise theory, it was indicated that noise correlation matrices form a general technique for calculating noise in n-port networks. Haus and Adler [6] have described the theory behind this technique. In 1976, Hillbrand and Russer [15] published equations and transformations that aid in supplying this method to two-port CAD.

This method is useful because it forms a base from which we can rigorously calculate the noise of linear two-ports combined in arbitrary ways. For many representations, the method of combining the noise parameters is as simple as that for combining the circuit element matrices. In addition, noise correlation matrices can be used to calculate the noise in linear frequency conversion circuits. The following is an introduction to this subject.

A linear, noisy two-port can be modeled as a noise-free two-port with two additional noise sources. These noise sources must be chosen so that they add directly to the resulting vector of the representation, as shown in Eqs. (3-84) and (3-85) and Figure 3-17:

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \tag{3-84}$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{3-85}$$

where the i and v vectors indicate noise sources for the y and z representations, respectively. This two-port example can be extended to n-ports in a straightforward, obvious way.

Since the noise vector for any representation is a random variable, it is much more convenient to work with the noise correlation matrix. The correlation matrix gives us deterministic numbers with which to calculate. The correlation matrix is formed by taking the mean value of the outer product of the noise vector. This is equivalent to multiplying the noise vector by its adjoint (complex conjugate transpose) and averaging the result:

$$\langle \bar{i}\ \bar{i}^{+} \rangle = \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} [i_1^* \quad i_2^*] = \begin{bmatrix} \langle i_1 i_1^* \rangle & \langle i_1 i_2^* \rangle \\ \langle i_1^* i_2 \rangle & \langle i_2 i_2^* \rangle \end{bmatrix} + [C_y] \tag{3-86}$$

where the angular brackets denote the average value.

Note that the diagonal terms are the "power" spectrum of each noise source and the off-diagonal terms are complex conjugates of each other and represent the "cross-power" spectra of the noise sources. "Power" is used because these magnitude-squared quantities are proportional to power.

To use these correlation matrices in circuit analysis, we must know how to combine them and how to convert them between various representations. An example using y matrices will illustrate the method for combining two-ports and their correlation matrices. Given two matrices y and y', when we parallel them we have the same port voltages, and the terminal currents add (Figure 3-26):

$$\begin{aligned} I_1 &= y_{11}V_1 + y_{12}V_2 + y'_{11}V_1 + y'_{12}V_2 + i_1 + i'_1 \\ I_2 &= y_{21}V_1 + y_{22}V_2 + y'_{21}V_1 + y'_{22}V_2 + i_2 + i'_2 \end{aligned} \tag{3-87}$$

or

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} + y'_{11} & y_{12} + y'_{12} \\ y_{21} + y'_{21} & y_{22} + y'_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 + i'_1 \\ i_2 + i'_2 \end{bmatrix} \tag{3-88}$$

Here we can see that the noise current vectors add just as the y parameters add. Converting the new noise vector to a correlation matrix yields

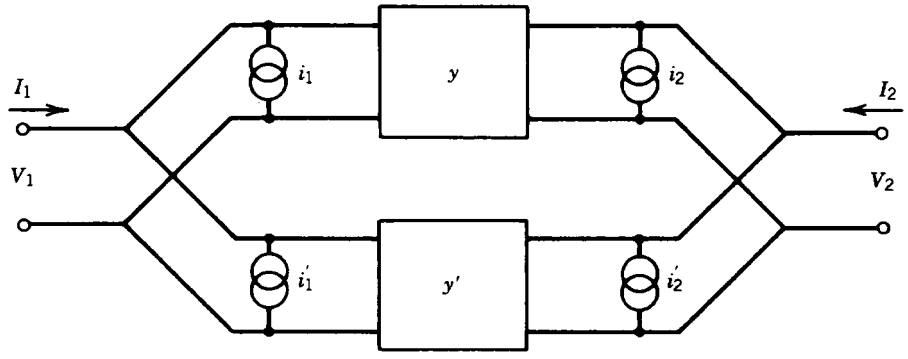


Figure 3-26 Parallel combination of two-ports using y parameters.

$$\langle \bar{i}_{\text{new}} \bar{i}^{+}_{\text{new}} \rangle = \left\langle \begin{bmatrix} i_1 + i_1' \\ i_2 + i_2' \end{bmatrix} [i_1^* + i_1'^* \quad i_2 i_2'^*] \right\rangle \tag{3-89}$$

$$= \begin{bmatrix} \langle i_1 i_1^* \rangle + \langle i_1' i_1'^* \rangle & \langle i_1 i_2^* \rangle + \langle i_1' i_2'^* \rangle \\ \langle i_2 i_1^* \rangle + \langle i_2' i_1'^* \rangle & \langle i_2 i_2^* \rangle + \langle i_2' i_2'^* \rangle \end{bmatrix} \tag{3-90}$$

The noise sources from different two-ports must be uncorrelated, so there are no cross products of different two-ports. By inspection, Eq. (3-90) is just the addition of the correlation matrices for the individual two-ports, so

$$[C_{y\text{new}}] = [C_y] + [C_y'] \tag{3-91}$$

The same holds true for g, h, and z parameters, but $ABCD$ parameters have the more complicated form shown below. If

$$[A_{\text{new}}] = [A][A'] \tag{3-92}$$

then

$$[C_{A\text{new}}] = \{C_A] + [A]\,[C_{A'}]\,[A]^{+} \tag{3-93}$$

The transformation of one representation to another is best illustrated by an example. Let us transform the correlation matrix for a Y representation to a Z representation. Starting with

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = [Y] \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \tag{3-94}$$

we can move the noise vector to the left side and invert y:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = [Y^{-1}] \begin{bmatrix} I_1 - i_1 \\ I_2 - i_2 \end{bmatrix} = [Y^{-1}] \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + [Y^{-1}] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-95}$$

Table 3-4 π Noise matrix $T_{\alpha\beta}$ transformation

		Original Form (α Form)					
		Y		Z		A	
Resulting Form (β Form)	Y	1	0	y_{11}	y_{12}	$-y_{11}$	1
		0	1	y_{21}	y_{22}	$-y_{21}$	0
	Z	z_{11}	z_{12}	1	0	1	$-z_{11}$
		z_{21}	z_{22}	0	1	0	$-z_{21}$
	A	0	A_{12}	1	$-A_{11}$	1	0
		1	A_{22}	0	$-A_{21}$	0	1

Since $(Y)^{-1} = (Z)$, we have

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = [Z] \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + [Z] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-96}$$

so

$$= [Z] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} = [T_{yz}] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-97}$$

where the signs of i_1 and i_2 are superfluous since they will cancel when the correlation matrix is formed. Here the transformation of the Y noise current vector to the Z noise voltage vector is done simply by multiplying by (Z). Other transformations are shown in Table 3-4.

To obtain the noise correlation matrix, we again form the mean of the outer product:

$$\langle vv^+ \rangle = \begin{bmatrix} \langle v_1 v_1^* \rangle & \langle v_1 v_2^* \rangle \\ \rangle v_1^* v_2 \rangle & \langle v_2 v_2^* \rangle \end{bmatrix} = [Z] \left\langle \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} [i_1^* \quad i_2^*] \right\rangle [Z]^+ \tag{3-98}$$

or

$$[C_z] = [Z]\,[C_y]\,[Z]^+ \tag{3-99}$$

where

$$v^+ = [i_1^* \quad i_2^*]\,[Z]^+ \tag{3-100}$$

This is called a congruence transformation. The key to all of these derivations is the construction of a correlation matrix from the noise vector, as shown in Eq. (3-90).

These correlation matrices may easily be derived from the circuit matrices of passive circuits with only thermal noise sources. For example,

$$[C_z] = 2kT\Delta f\,\mathrm{Re}([Z]) \tag{3-101}$$

and

$$[C_y] = 2kT\Delta f\,\mathrm{Re}([Y]) \tag{3-102}$$

The $2kT$ factor comes from the double-sided spectrum of thermal noise. The correlation matrix from the $ABCD$ matrix may be related to the noise figure, as shown by Hillbrand and Russer [15]. We have

$$F = 1 + \frac{\overline{Y}[C_a]\overline{Y}^+}{2kT\,\mathrm{Re}(Y_G)} \tag{3-103}$$

where

$$\overline{Y} = \begin{bmatrix} Y_G \\ 1 \end{bmatrix}$$

Expressing the noise factor in terms of the correlation matrix, we obtain a complete formula:

$$F = 1 + \frac{C_{22}^{A}(f) + 2\,\mathrm{Re}\{Y_g(f)\,C_{12}^{A}(f)\} + |Y_g(f)|^2\,C_{11}^{A}(f)}{2kT_0\,\mathrm{Re}\{Y_g(f)\}} \tag{3-104}$$

Once we transform this in the Y parameter form, we obtain the following equation:

$$F(f) = F_{\min}(f) + \frac{R_n(f)|Y_{\mathrm{opt}}(f) - Y_g(f)|^2}{\mathrm{Re}\{Y_g(f)\}} \tag{3-105}$$

It should be noted that all these values are frequency dependent, as expressed in this equation.

The *ABCD* correlation matrix can be written in terms of the noise figure parameters as (double-sided spectrum)

$$[C_a] = 2kT\begin{bmatrix} R_n & \dfrac{F_0 - 1}{2} - R_n Y_{0n}^{*} \\ \dfrac{F_0 - 1}{2} - R_n Y_{0n} & R_n|Y_{0n}|^2 \end{bmatrix} \tag{3-106}$$

The noise correlation matrix method forms an easy and rigorous technique for handling noise in networks. This technique allows us to calculate the total noise for complicated networks by combining the noise matrices of subcircuits. It should be remembered that although noise correlation matrices apply to n-port networks, noise figure calculations apply only to pairs of ports. The parameters of the C_a matrix can be used to give the noise parameters:

$$Y_{0n} = \sqrt{\frac{C_{ii*}}{C_{uu*}} - \left[\mathrm{Im}\left(\frac{C_{ui*}}{C_{uu*}}\right)\right]^2} + j\,\mathrm{Im}\left(\frac{C_{ui*}}{C_{uu*}}\right) \tag{3-107}$$

$$F_0 = 1 + \frac{C_{ui*} + C_{uu*}Y_{0n}^{*}}{kT} \tag{3-108}$$

$$R_u = C_{uu*} \tag{3-109}$$

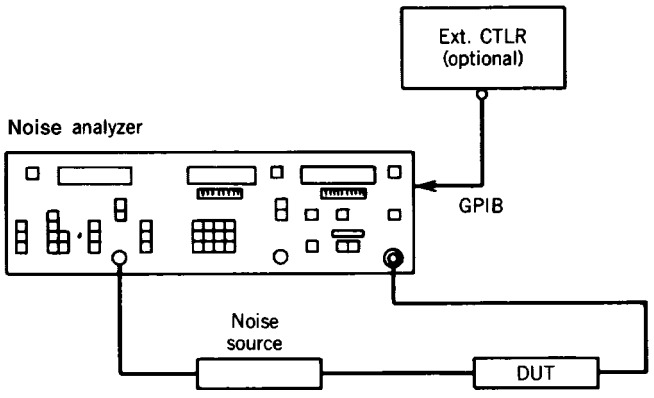


Figure 3-27 Noise figure measurement.

Noise Figure Test Equipment. Figure 3-27 shows the block diagram of a noise test setup. It includes the noise source and the other components. The metering unit has a special detector that is linear and over a certain dynamic range measures linear power. The tunable receiver covers a wide frequency range (e.g., 10–1800 MHz) and controls the noise source. The receiver is a double-conversion superheterodyne configuration with sufficient image rejection to avoid double-sideband noise measurements that would give the wrong results.

These receivers are microprocessor controlled and the measurement is a two-step procedure. The first is a calibration step that measures the noise figure of the receiver system and a reference power level. Then the device under test (DUT) is inserted and the system noise figure and total output power are measured. The noise factor is calculated by

$$F_1 = F_{\text{system}} - \frac{F_2 - 1}{G_1} \tag{3-110}$$

and the gain is given by the change in output power from the reference level [16]. The noise of the system is calculated by measuring the total noise power with the noise source on and off. With the excess noise ratio (ENR) known [16],

$$F_{\text{system}} = \frac{\text{ENR}}{Y - 1} \tag{3-111}$$

The noise bandwidth is usually set by the bandwidth of the receiver, which is assumed to be constant over the linear range. The ENR of the noise source is given by

$$\text{ENR} = \frac{T_{\text{hot}}}{T_{\text{cold}}} - 1 \tag{3-112}$$

where T_{cold} is usually room temperature (290 K). This ENR number is about 15 dB for noise sources with a 15-dB pad and 5 dB for noise sources with a 25-dB pad. Since both gain and noise were stored in the initial calibration, a noise/gain sweep can be performed.

For frequencies above 1800 MHz, we can extend the range with the help of the external signal generators, as shown in Figure 3-28. As shown, a filter ahead of the external mixer reduces the noise energy in the image band. If the DUT has a very broad frequency range

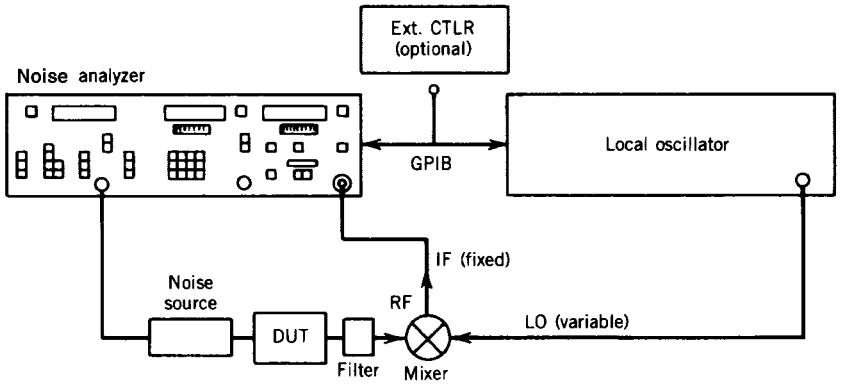


Figure 3-28 Single-sideband (SSB) noise figure measurements using an external mixer.

and has flat gain and noise over that range, a double-sideband (DSB) measurement is possible, with the image-rejection filter removed. However, a single-sideband (SSB) measurement is always more accurate [17].

How To Determine the Noise Parameters. The noise figure of a linear two-port network as a function of source admittance may be represented by

$$F = F_{\min} + \frac{R_n}{G_G}\left[(G_{0n} - G_G)^2 + (B_{0n} - B_G)^2\right] \tag{3-113}$$

where $G_G + jB_G$ = generator admittance presented to the input of the two-port
$G_{0n} + jB_{0n}$ = generator admittance at which optimum noise figure occurs
R_n = empirical constant relating sensitivity of the noise figure to generator admittance, with dimensions of resistance

It may be noted that for an arbitrary noise figure measurement with a known generator admittance, Eq. (3-113) has four unknowns—$F_{\min}$, R_n, G_{0n}, and B_{0n}. By choosing four known values of generator admittance, a set of four linear equations is formed and solutions of the four unknowns can be found [18, 19]. Eq. (3-113) may be transformed to

$$F = F_{\min} + \frac{R_n|Y_{0n}|^2}{G_G} - 2R_nG_{0n} + \frac{R_n|Y_G|^2}{G_G} - 2R_nB_{0n}\frac{B_G}{G_G} \tag{3-114}$$

or

$$F = F_{\min} + \frac{R_n}{G_G}|Y_G - Y_{0n}|^2 \tag{3-115}$$

Let

$$X_1 = F_{\min} - 2R_nG_{0n}$$

$$X_2 = R_n|Y_{0n}|^2$$

$$X_3 = R_n$$

$$X_4 = R_nB_{0n}$$

Then the generalized equation may be written

$$F_i = X_1 + \frac{1}{G_{si}}X_2 + \frac{|Y_{si}|^2}{G_{si}}X_3 - 2\frac{G_{si}}{B_{si}}X_4 \tag{3-116}$$

or, in matrix form,

$$[F] = [A]\,[X] \tag{3-117}$$

and the solution becomes

$$[X] = [A]^{-1}\,[F] \tag{3-118}$$

These parameters completely characterize the noise behavior of the linear two-port network. Direct measurement of these noise parameters by this method would be possible only if the receiver on the output of the two-port were noiseless and insensitive to its input admittance. In actual practice, the receiver itself behaves as a noisy two-port network and can be characterized in the same manner. What is actually being measured is the system noise figure of the two-port and the receiver. Thus, it becomes apparent that to do a complete two-port noise characterization, the gain of the two-port must be measured [20]. In addition, any losses in the input-matching networks must be carefully accounted for, because they add directly to the measured noise figure reading [21].

3-1-4 Linearity

Consistent with the elaborate discussion of linearity questions above, here is a quick refresher on what must be considered in amplifiers.

Dynamic Range, Compression, and IMD. An amplifier's dynamic range is related to the minimum discernible signal and the 1-dB compression point according to the equation

$$DR_n = \frac{(n-1)(IP_{n(\text{in})} - MDS)}{n} \tag{3-119}$$

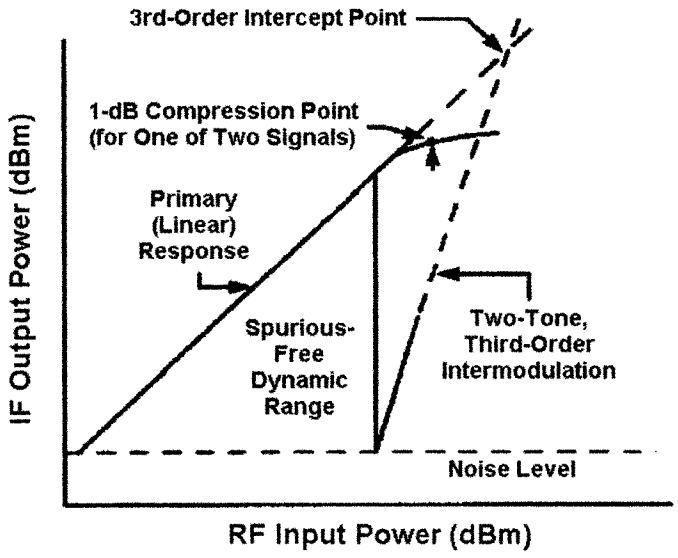


Figure 3-29 Amplifier linearity evaluation, including compression and two-tone IMD dynamic range. $P_{-1\text{dB}}$ for a single-tone cannot be read directly from this graph because the values shown are the result of two equal-power tones.

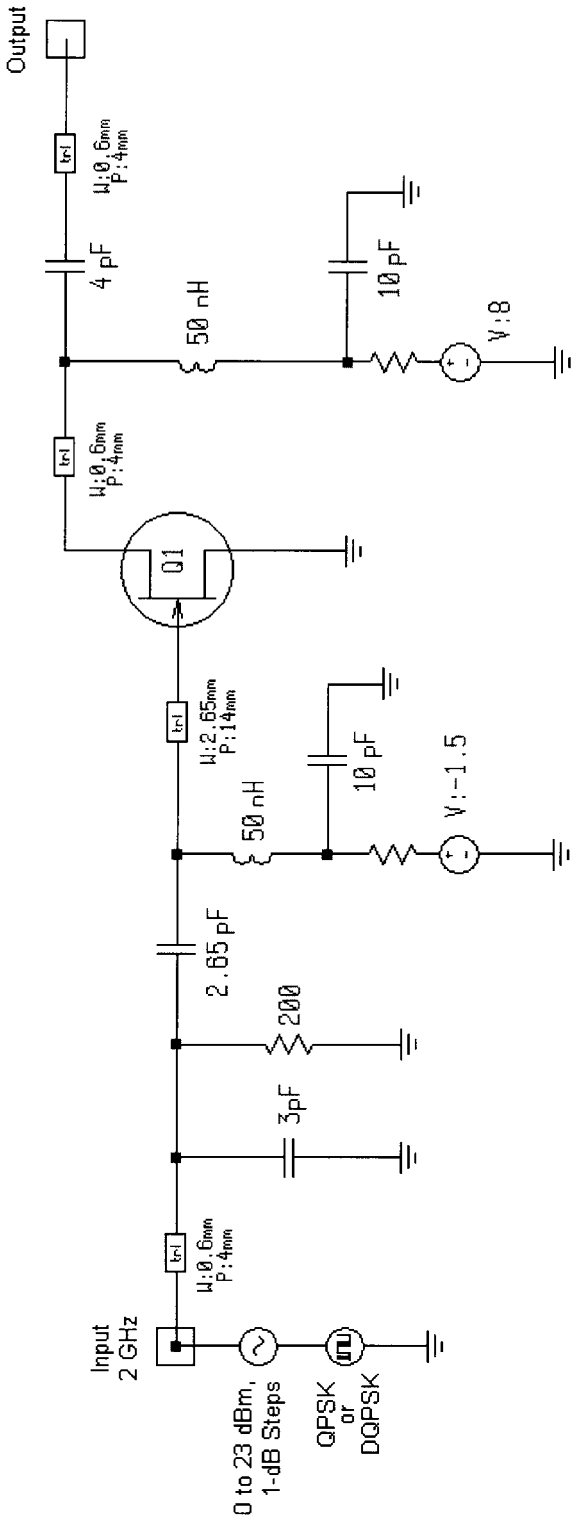


Figure 3-30 The power amplifier schematic.

where DR is the dynamic range in decibels, n is the order, $IP_{(in)}$ is the input intercept power in dBm, and MDS is the minimum detectable signal power in dBm. Once the 1-dB compression point is known, it is a fair assumption that the third-order intercept point, expressed in dBm, is 10 dB above this.

The intermodulation distortion gives the quality of an amplifier to withstand multiple signals without generating large intermodulation products or when using amplitude-modulated signals (cross-modulation). Figure 3-29 provides a quick reminder of IMD issues.

In some amplifiers, we would like to have control over the gain. While the better way to control gain while obtaining high linearity is a PIN-diode attenuator, special AGC circuits have become part of the amplifiers. Figure 3-48 shows a circuit of the SL610, made by Plessey, which has an AGC input provision. The noise figure of the PIN-diode attenuator is essentially equal to its attenuation. The AGC in multistage amplifiers is distributed over several stages and there is a lesser correlation between noise and AGC than in purely passive attenuators.

Special Case of Linearity Requirements for Digital Modulation. For cellular telephone systems we use digital modulation formats, such as QPSK and π/4-DQPSK, that combine phase and amplitude modulation. Amplifiers handling such signals must be carefully characterized and designed if adequate amplitude and phase linearity are to be maintained.

Examples of Power Amplifiers: Looking into the Effects of Distortion. This example presents modulation analysis using QPSK in a 2-GHz power amplifier (Figure 3-30). The simple amplifier was designed using the electronic Smith Tool of Serenade to determine the matching network for a narrowband design. The example demonstrates setup of modulation analysis and available results.

Siemens among others provides integrated amplifiers based on the concept depicted in Figure 3-30, the most popular ones being the CGY96 and CGY94. For the purpose of measurements, these amplifiers were actually put on a breadboard. Figure 3-31 shows the big brother, the CGY96. It is a GaAs MMIC intended as a power amplifier for GSM Class 4 phones with 3.2 W (35 dBm) at 3.5 V at an overall power added efficiency of 50%. Table 3-5 shows its electrical characteristics.

For the purpose of our experiments, we considered the CGY94 because it requires less external peripheral circuits. Table 3-6 shows its electrical characteristics. For those of us interested in seeing part of the internal circuit, Figure 3-32 shows the chip. This two-stage amplifier uses an interesting internal feedback scheme. Its schematic is shown in Figure 3-33.

For the purpose of looking at the waveforms, we will continue with the single-stage amplifier (Figure 3-30) and then compare the measured versus predicted performance of this CGY94-based amplifier.

A modulation source is connected to the RF source at the input port of the amplifier. The RF source specifies the carrier power (or voltage or current) and the modulation source specifies the modulation format and properties of the modulated signal.

A brief review of the modulation source used in this example indicates the following properties:

Property	Value	Description
NB	128	Number of bits
BR	1.2288E6	Bit rate
N	8	Number of samples per bit
BW1	590.E3	One-sided bandwidth of the main channel in the PIB or P1IB calculation
FS1	740.E3	Start baseband frequency of the fist adjacent channel
BW2	590.E3	One-sided bandwidth of the main channel in the P2IB calculation
FS2	1990.E3	Start baseband frequency of the second adjacent channel
BW3	590.E3	One-sided bandwidth of the main channel in the P3IB calculation
FS3	3240.E3	Start baseband frequency of the third adjacent channel
M	4	Order of the signal space; M = 4 is QPSK
DLY	0.0	Fractional bit delay
IASC	1.0	I channel amplitude scale; multiplier for the I waveform to model amplitude imbalance
QASC	1.0	Q channel amplitude scale; multiplier for the Q waveform to model amplitude imbalance
LPF	3	Baseband low-pass filter; 3 = root-raised cosine filter
LPFC	665.E3	–3 dB cutoff frequency of LPF
LPFN	3	Number of resonators if LPF = 1
LPFR	0.35	Roll-off factor if LPF = 3 or LPF = 4
FILE	N/A	File name of user-defined modulated signal file

The modulation source is a QPSK stream with a bit rate of 1.2288 Mbps. A total of 128 bits will be analyzed, and each bit is sampled eight times to construct an accurate analog waveform. The BW and FS parameters are defined here to specify the main (in-band) channel and adjacent channel for ACPR calculations. Imbalance of the I and Q amplitudes and phases can be described by the IASC, QASC, and DLY parameters, if desired. The modulation signal

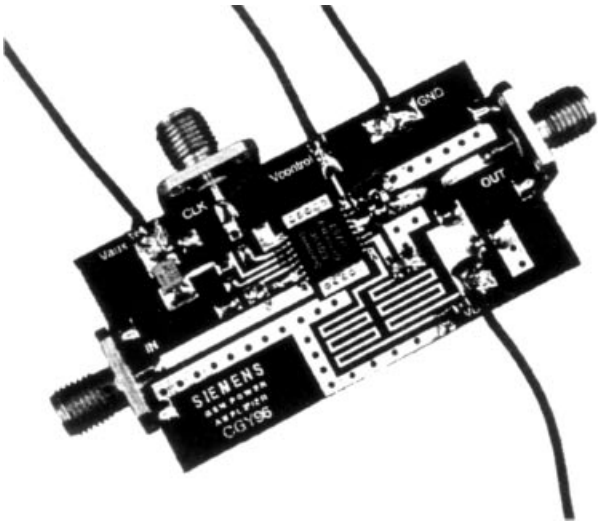


Figure 3-31 CGY96 GaAs MMIC power amplifier.

Table 3-5 CGY96 electrical characteristics

(T_A = 25 °C, V_{neg} = 5 V, $V_{control}$ = 2.2 V; duty cycle 12.5%, t_{on} = 577 µs)

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
Frequency range	f	880	—	915	MHz
Supply current P_{in} = 0 dBm	I_D	—	1.8	—	A
Supply current negative voltage generator V_{aux} = 3.5 V	I_{AUX}	—	10	—	mA
Gain (small signal)	G	—	40	—	dB
Power gain P_{in} = 0 dBm	G_P	—	35	—	dB
Output power P_{in} = 0 dBm, $V_{control}$ = 2.0–2.5 V	P_{OUT}	—	35	—	dBm
Overall power added efficiency P_{in} = 0 dBm	η	—	50	—	%
Dynamic range output power $V_{control}$ = 0.2–2.2 V		—	80	—	dB
Harmonics P_{in} = 0 dBm	$H(2f_0)$	—	–40	—	dBc
	$H(3f_0)$	—	–43	—	dBc
	$H(4f_0)$	—	–44	—	dBc
Noise power in RX (935–960 MHz) P_{in} = 0 dBm, P_{out} = 35 dBm, 100 kHz RBW	N_{RX}	—	–81	—	dBm
Stability *all spurious outputs* < –60 dBc, *VSWR load, all phase angles*		—	10:1	—	—
Input VSWR		—	1.7:1	—	—

Table 3-6 CGY94 electrical characteristics

T_A = 25 °C, f = 0.9 GHz, $Z_S = Z_L$ = 50 Ω, V_D = 3.6 V, V_G = 4 V, VTR pin connected to ground, unless otherwise specified; pulsed with a duty cycle of 10%, ton = 0.33 ms)

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
Supply current V_D = 3.0 V P_{in} = 10 dBm	I_{DD}	—	1.18	—	A
Negative supply current *(normal operation)*	I_G	—	2	—	mA
Shut-off current *VTR n.c.*	I_D	—	400	—	µA
Negative supply current (*shut off mode, VTR pin n.c.*)	I_G	—	10	—	µA
Gain P_{in} = *5* dBm	G	27.0	29.0	—	dB
Power gain V_D = 3.6 V; P_{in} = 10 dBm	G	22.8	23.6	—	dB
Output power V_D = 3.0 V*;* P_{in} = 10 dBm	P_O	31.5	32.3	—	dBm
Output power V_D = 3.6 V*;* P_{in} = 10 dBm	P_O	32.8	33.6	—	dBm
Output power V_D = 5 *V;* P_{in} = 10 dBm	P_O	34.5	35.5	—	dBm
Overall power added efficiency V_D = 3.0 V; P_{in} = 10 dBm	η	43	48	—	%
Overall power added efficiency V_D = 3.6 V; P_{in} = 10 dBm	η	42	47	—	%
Overall power added efficiency V_D = 5 V; P_{in} = 10 dBm	η	41	46	—	%
Harmonics (P_{in} = 10 dBm, *CW*) $2f_0$	—	—	–49	—	dBc
V_D = 3.6 V; (P_{out} = 33.1 dBm) $3f_0$	—	—	–45	—	dBc
Input VSWR V_D = 3.6 V	—	—	1.5:1	2.0:1	—

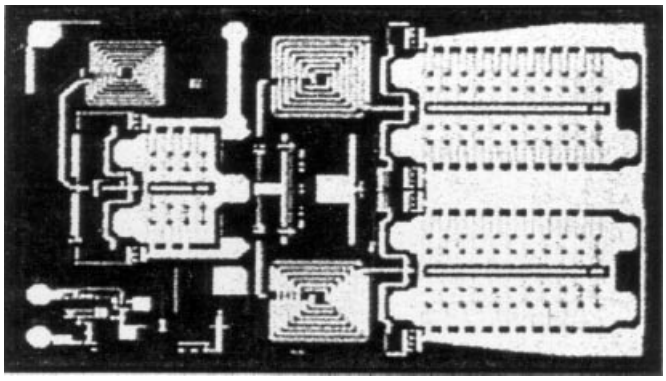


Figure 3-32 The CGY94 amplifier chip.

is often filtered, and several types of filters are available. Here a root-raised cosine filter is used with a cutoff frequency of 665 kHz and a roll-off factor of 0.35.

Analysis. Figure 3-34 shows the compression characteristics of the amplifier as the available RF source power (a sinusoid) is swept. Note that P_{-1dB} is 11.5 dBm referred to the input and 28.5 dBm referred to the output. We will look at modulation characteristics in the linear region of operation at a source power of 0 dBm, and the nonlinear region at a source power of 20 dBm.

Next, we will view the voltage across the transistor to examine its behavior under compression. Figure 3-35 shows the drain–source voltage of the FET as the source power is swept. It is clear from this graph that clipping due to pinchoff and forward-gate conduction

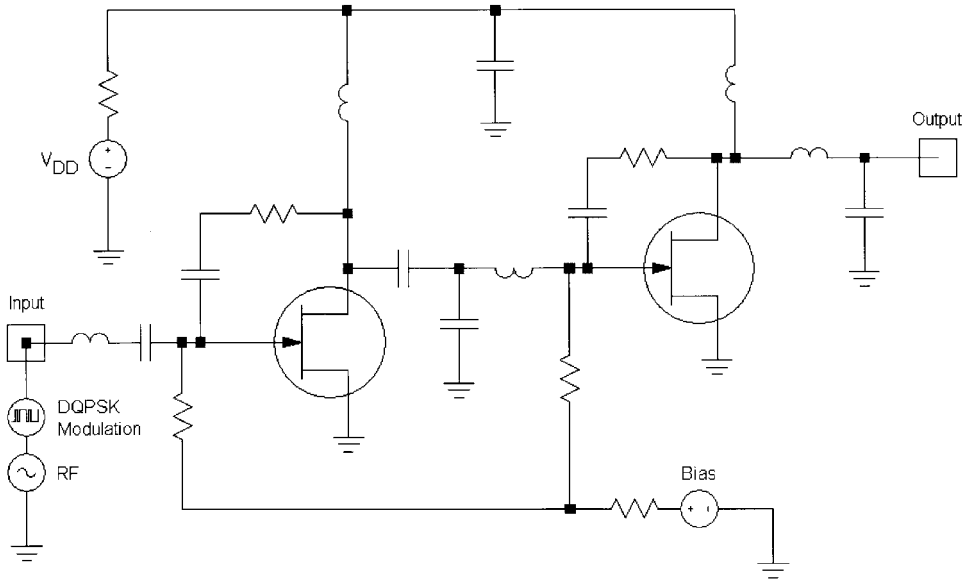


Figure 3-33 Schematic of the CGY94 GaAs MMIC power amplifier.

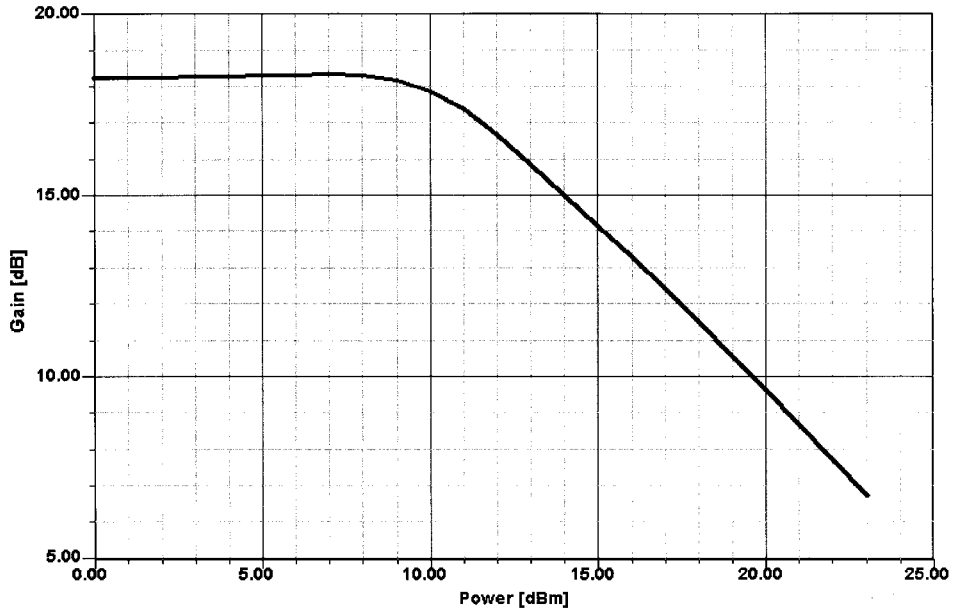


Figure 3-34 Single-tone RF power sweep analysis of the FET amplifier.

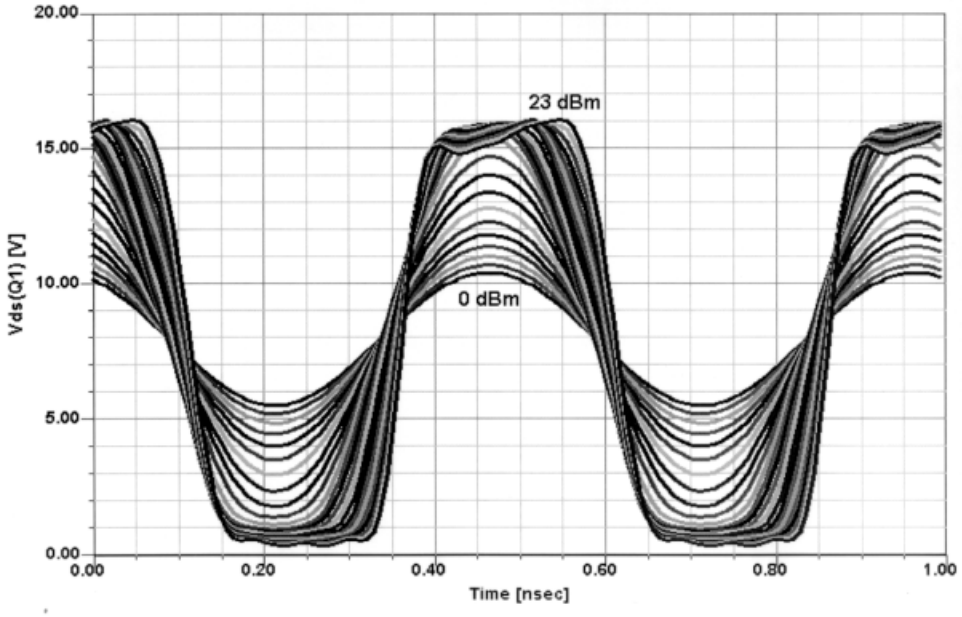


Figure 3-35 V_{DS} of the FET versus RF source power.

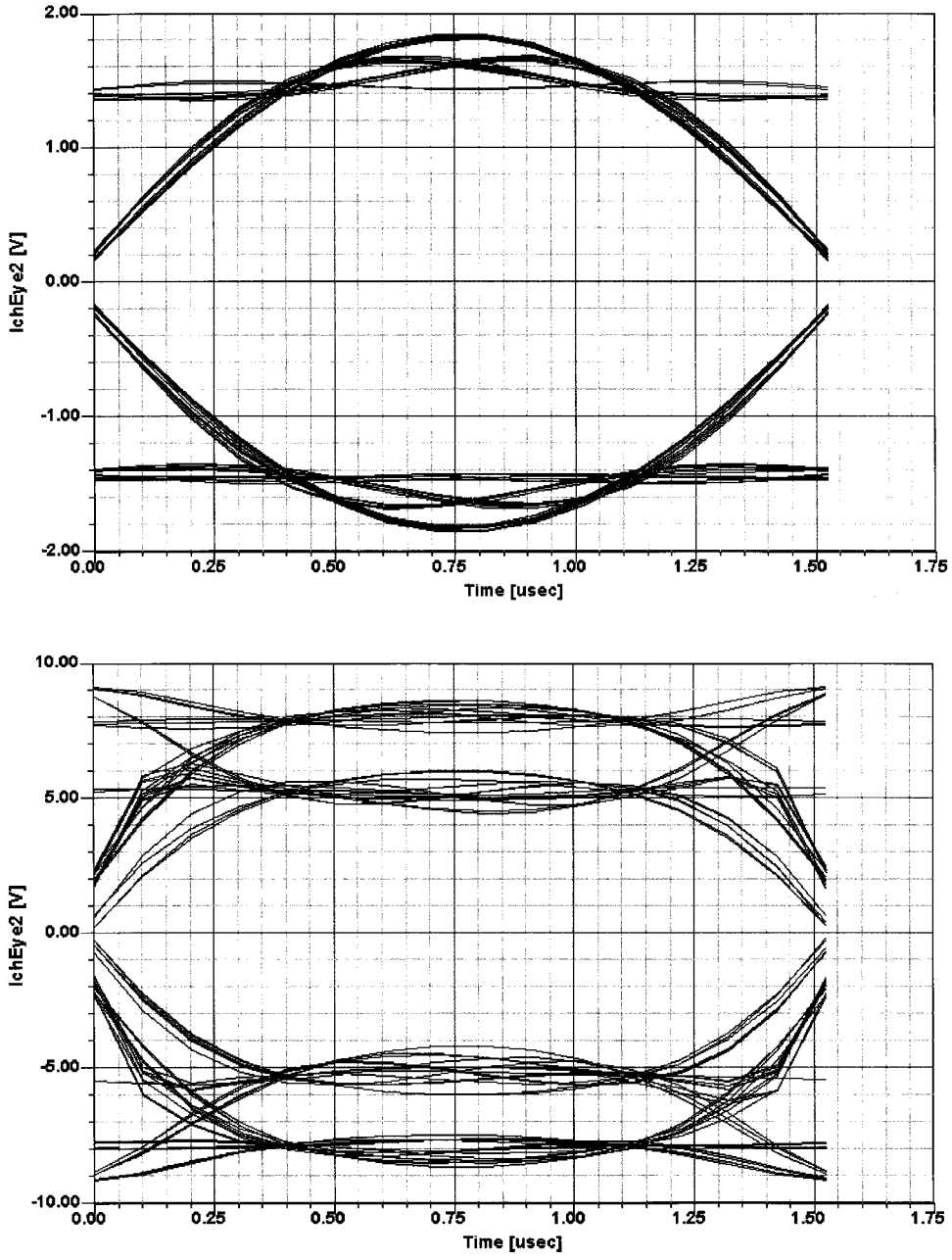


Figure 3-36 Eye diagrams of the I channel for source powers of (a) 0 dBm and (b) 20 dBm.

is limiting the performance of the device. A phase shift of the voltage waveform as power is increased is also apparent. The combination of the power compression and phase shift, or AM to PM conversion, will be used in the modulation analysis to determine the overall distortion of the modulated signal.

The output available from modulation analysis includes:

- I and Q channel waveforms
- Eye and constellation diagrams
- Spectral plots
- ACPR, in-band, and adjacent power

Figure 3-36 shows the eye diagrams for the 128-bit QPSK signal at two RF source powers, 0 and 20 dBm. Operating linearly at 0 dBm input, the amplifier doesn't distort the eye and it remains wide open. Note that in this example the filter bandwidth doesn't produce any intersymbol interference, as evidenced by the open eye. At an input power of 20 dBm, the amplifier compresses the signal, and distortion in the eye is evident.

We can also view the modulation spectrum at these two power levels to investigate the intermodulation distortion and spectral regrowth that takes place. Figure 3-37 shows the spectral plots. The lower trace corresponds to 0 dBm source power and shows almost no regrowth, while the upper trace at 20 dBm source power shows considerable regrowth.

The uncorrupted modulation source can also be shown. This corresponds to the source as a pure voltage without any circuit influence.

Now we evaluate the CGY94 board in a similar fashion. Figure 3-38 shows the waveform used for CAD purposes. Needless to say, such a signal cannot be produced by any signal

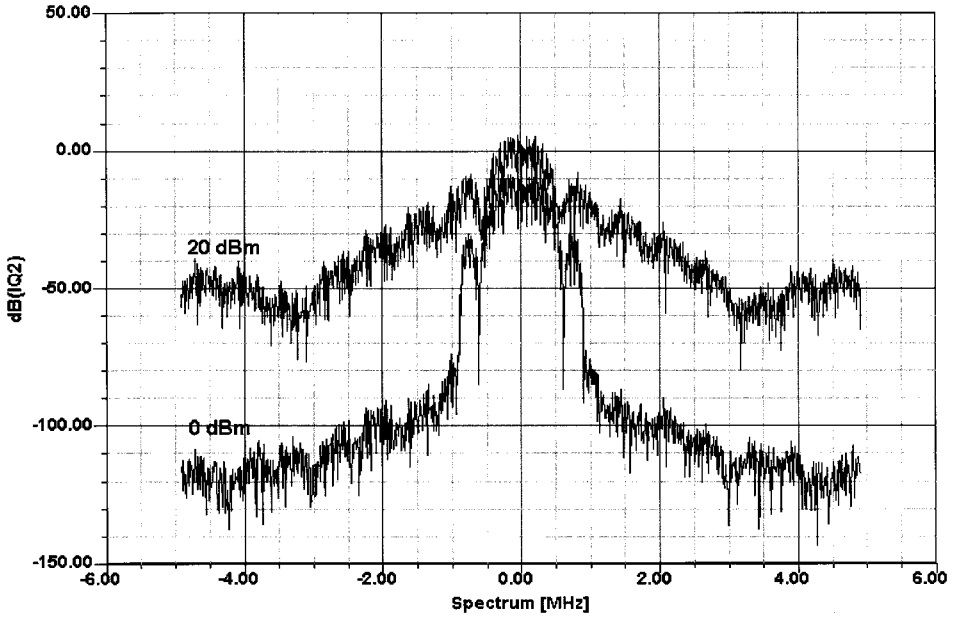


Figure 3-37 Spectrum of the modulation signal at the amplifier output. Source powers are 0 and 20 dBm.

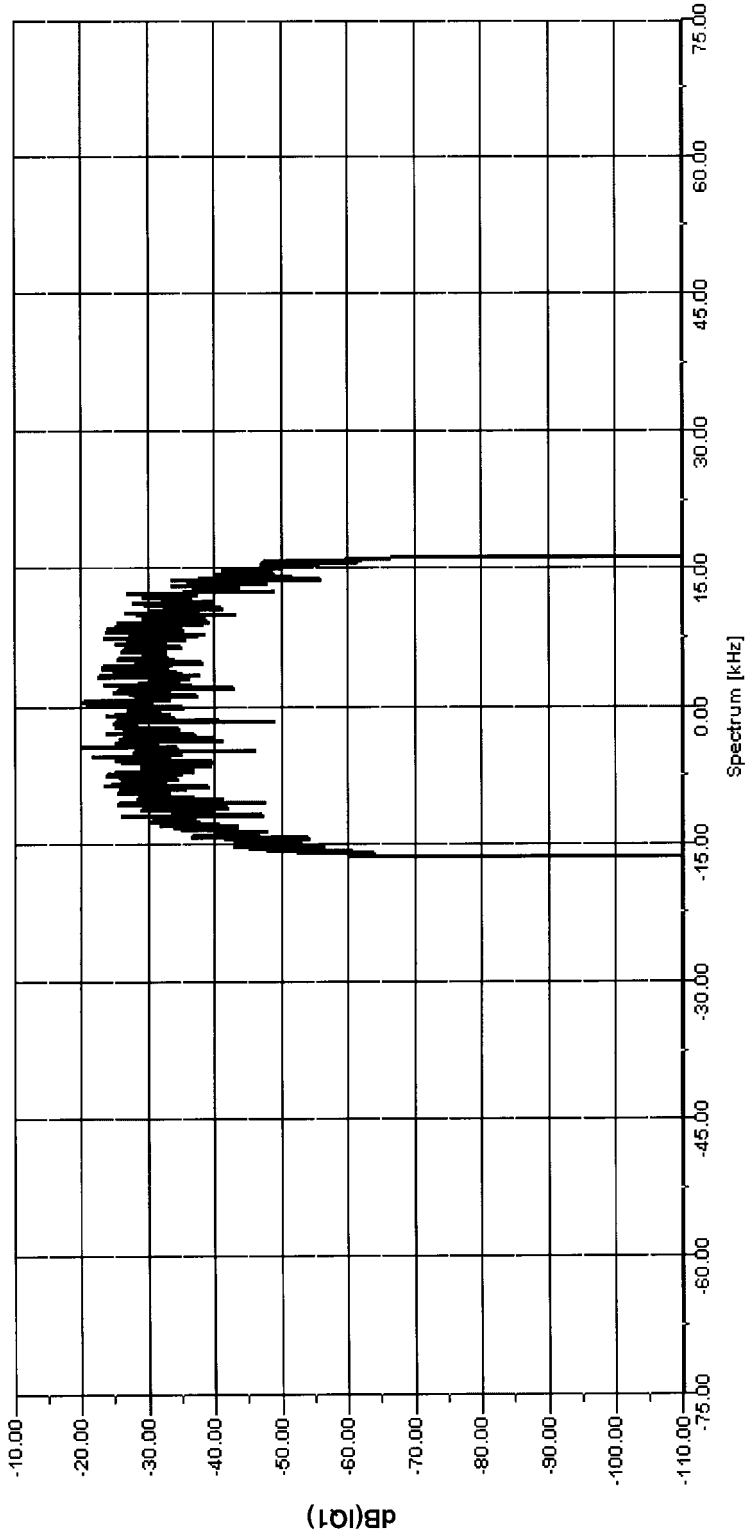


Figure 3-38 Simulated signal used for CAD analysis of the CGY94 amplifier.

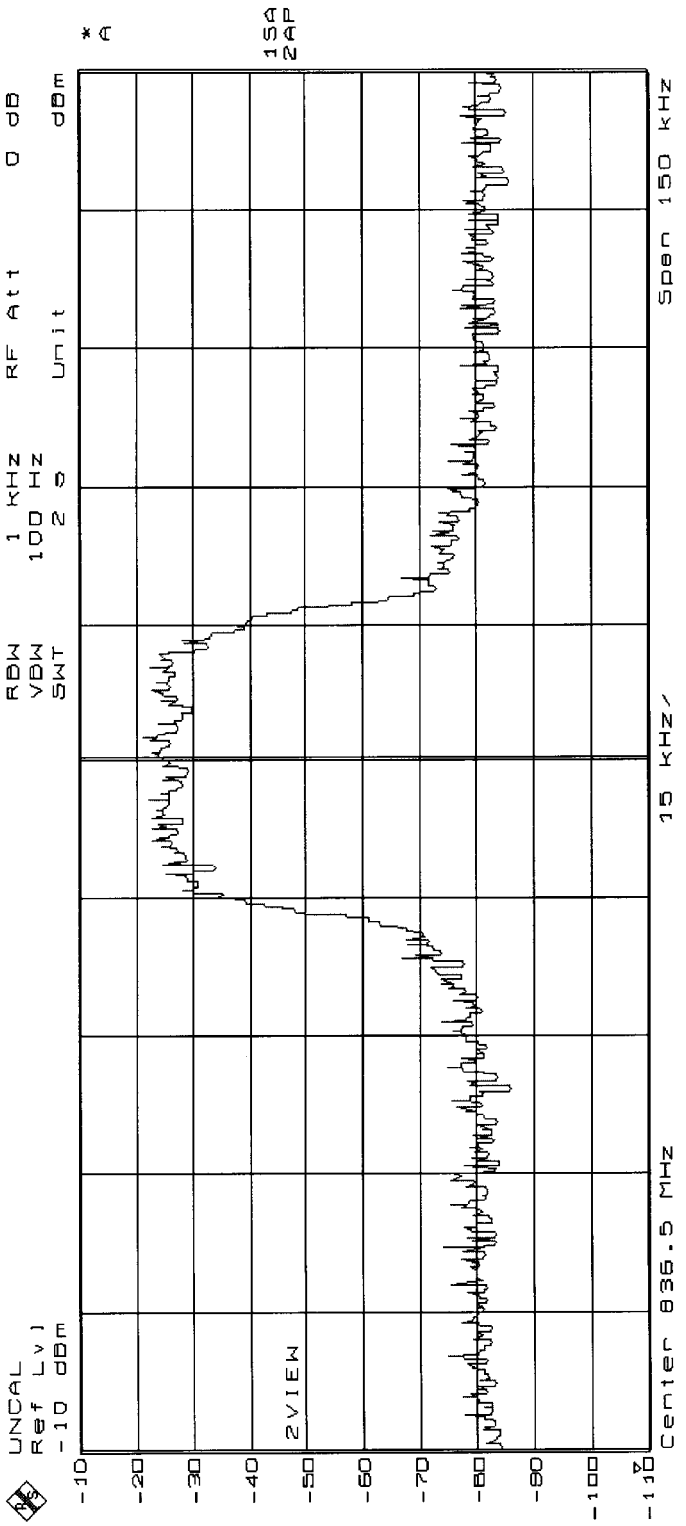


Figure 3-39 Real signal used for testing the actual amplifier.

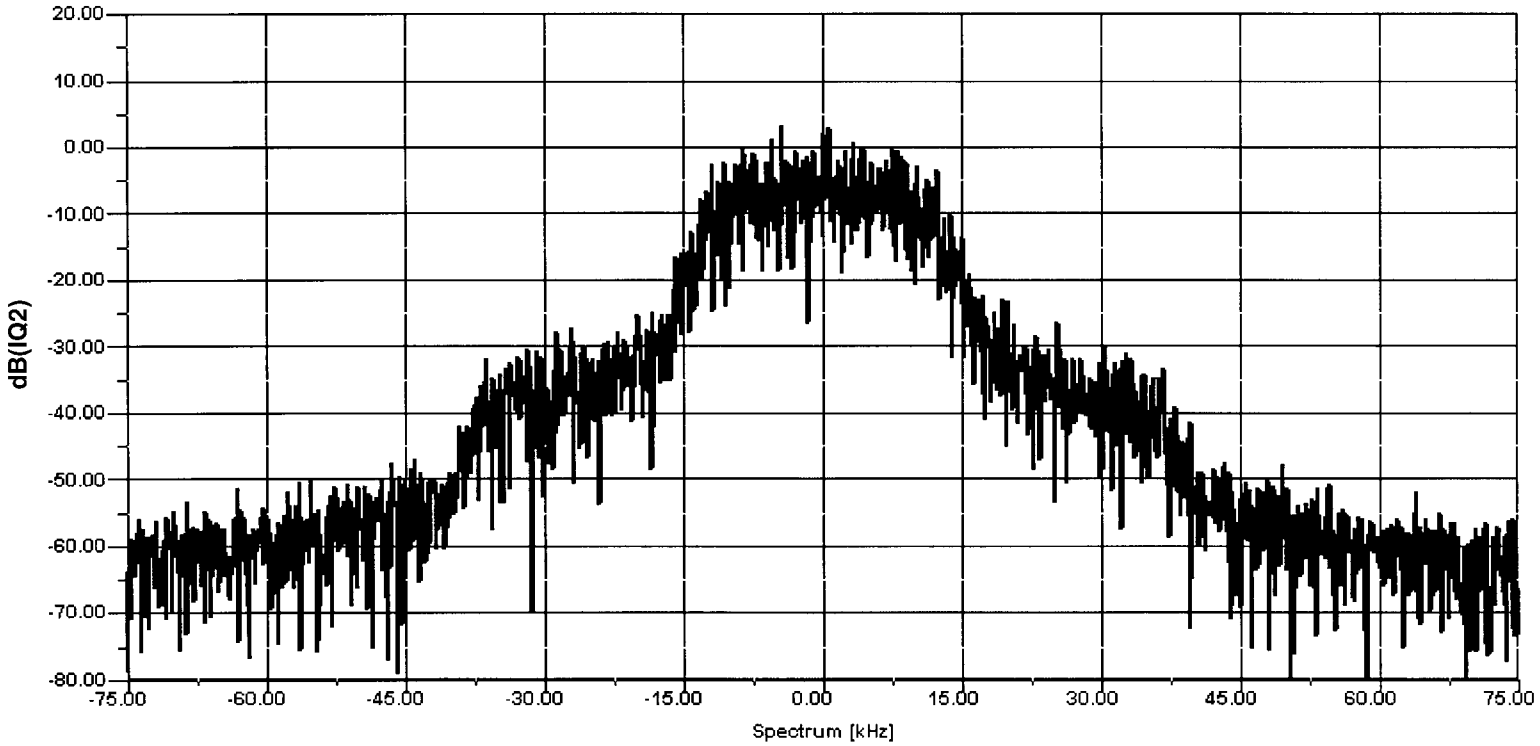


Figure 3-40 Spectral regrowth predicted by the simulation.

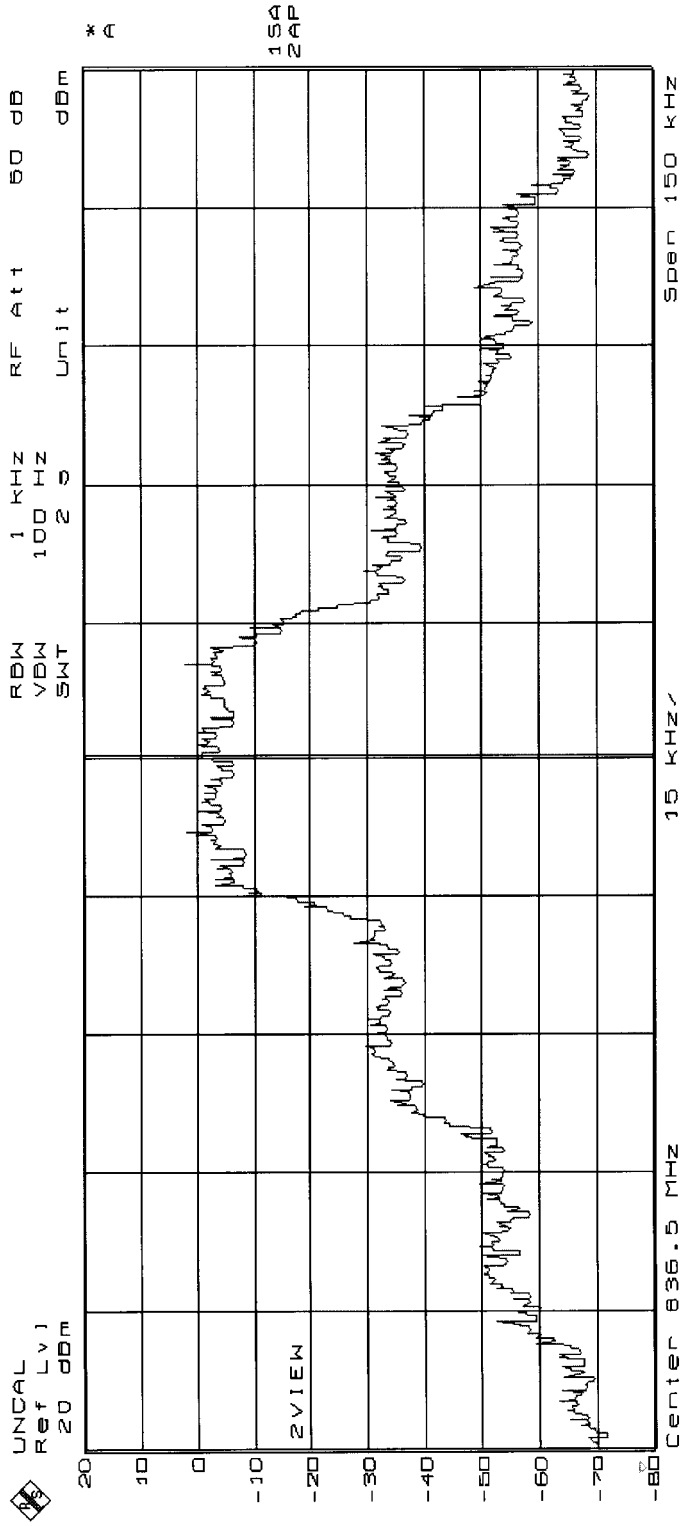


Figure 3-41 Measured spectral regrowth of the CGY94 amplifier.

generator. The one generated by an actual signal generator is shown in Figure 3-39. The wideband noise is around –80 dBm and the input level is about –25 dBm, consistent with the simulation signal. Looking at the output, Figure 3-40 shows the simulated signal with the expected increase in bandwidth. This is due to the nonlinearity of the amplifier and is consistent with Figure 3-37 for the single-stage amplifier. The actual measured output for the CGY94 shown in Figure 3-41 is quite close to its simulated output, indicating that the mathematical approach used for simulating is correct. This is only one example; in reality, many more waveforms can be analyzed and predicted using this method. Figure 3-41 shows essentially three distinct steps, which are sufficiently close to the response shown in the simulation. A dynamic range of 70 dB, while displayed in Figures 3-40 and 3-41, is not really necessary for good-quality transmission, and the discrepancy between the two pictures is in part due to the limited modeling quality of the active device as provided by the company that did the parameter extraction for this transistor.

Another useful view to aid in the understanding of the compression and distortion within the amplifier is to look at the magnitude of the complex waveform. Since the signal is composed of the in-phase and quadrature-phase components it is represented as

$$s(t) = i(t) + j \cdot q(t) \tag{3-120}$$

The time-domain magnitude is written

$$|s(t)| = \sqrt{i(t)^2 + q(t)^2} \tag{3-121}$$

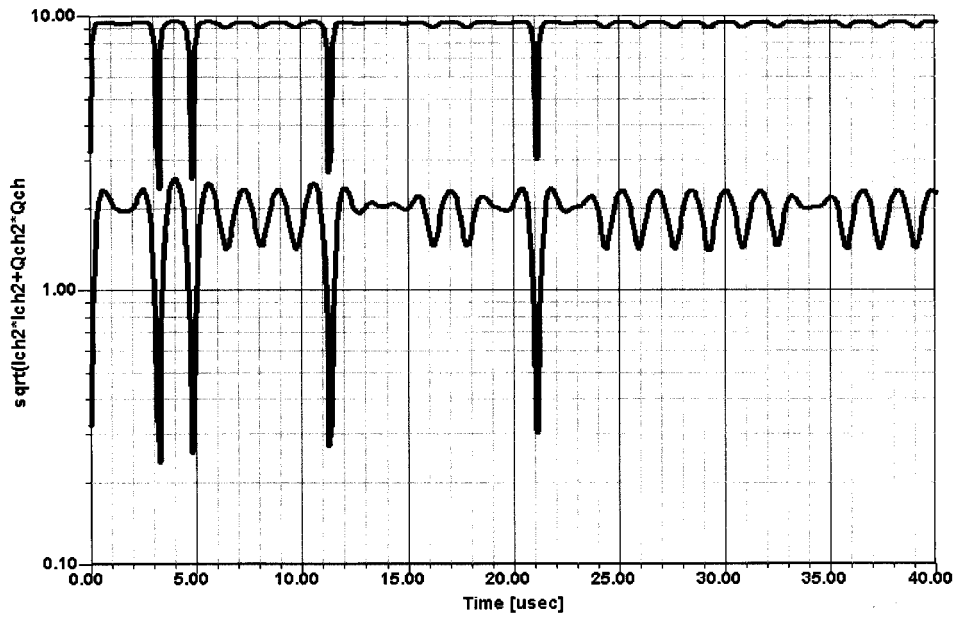


Figure 3-42 Time-domain magnitude of the complex modulation signal at source powers of 0 and 20 dBm.

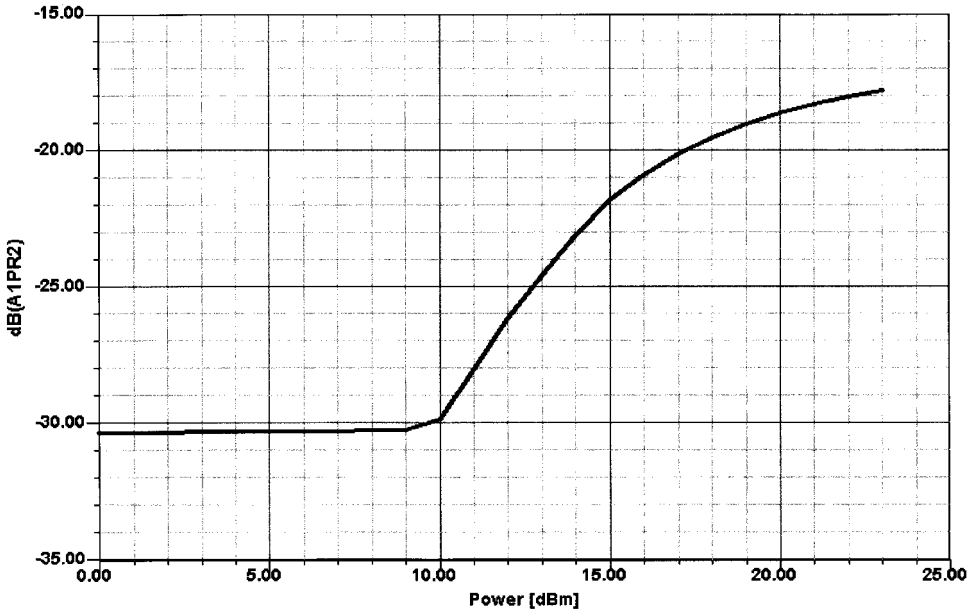


Figure 3-43 Adjacent-channel power ratio as function of RF source power.

Figure 3-42 shows the magnitude at source powers of 0 and 20 dBm. The *X* axis has been rescaled to 40 μs for improved viewing. Note the significant compression of the 20-dBm waveform at the higher signal levels. It is clear from this view that the signal is severely distorted.

We can also compute the adjacent-channel power ratio (ACPR) against the RF power sweep. ACPR is the ratio of the adjacent-channel power to the in-band channel power. The bandwidth (BW) and adjacent-channel start frequency (FS) are used for this calculation. For accurate ACPR calculation, a large number of bits are needed. This example uses 128 bits so the computation time is short, but 512 or more bits should be used for a more accurate computation.

Figure 3-43 shows the ACPR as a function of RF source power. The ACPR is nearly constant (and nonzero due to the gradual skirt of the baseband filter and inherent spillover to the adjacent channel) up to the P_{-1dB} compression point . As the intermodulation products spill into the adjacent channel and spectral regrowth occurs, the ACPR degrades.

π/4-DQPSK Circuit Analysis. Next, we will examine the same amplifier using a π/4-DQPSK modulation source. The task is identical to the previous one except for the source and the number of bits. The number of bits for this project has been increased from 128 to 256. Figure 3-44 shows the eye diagrams from this modulation format, at source powers of 0 and 20 dBm. The distortion and related intersymbol interference are clearly evident.

The constellation plot at 0 dBm is shown in Figure 3-45.

It may be interesting to take a look at the different waveforms currently in use. Figure 3-46 shows a spectrum analyzer picture of the most popular ones.

For reasons of time-division management, the European standard, GSM, which is now offered by Omnipoint at the 1.8-GHz band in the United States, uses gating to reduce the actual bandwidth of the spectrum. Figure 3-47 compares the GSM signal with and without gating.

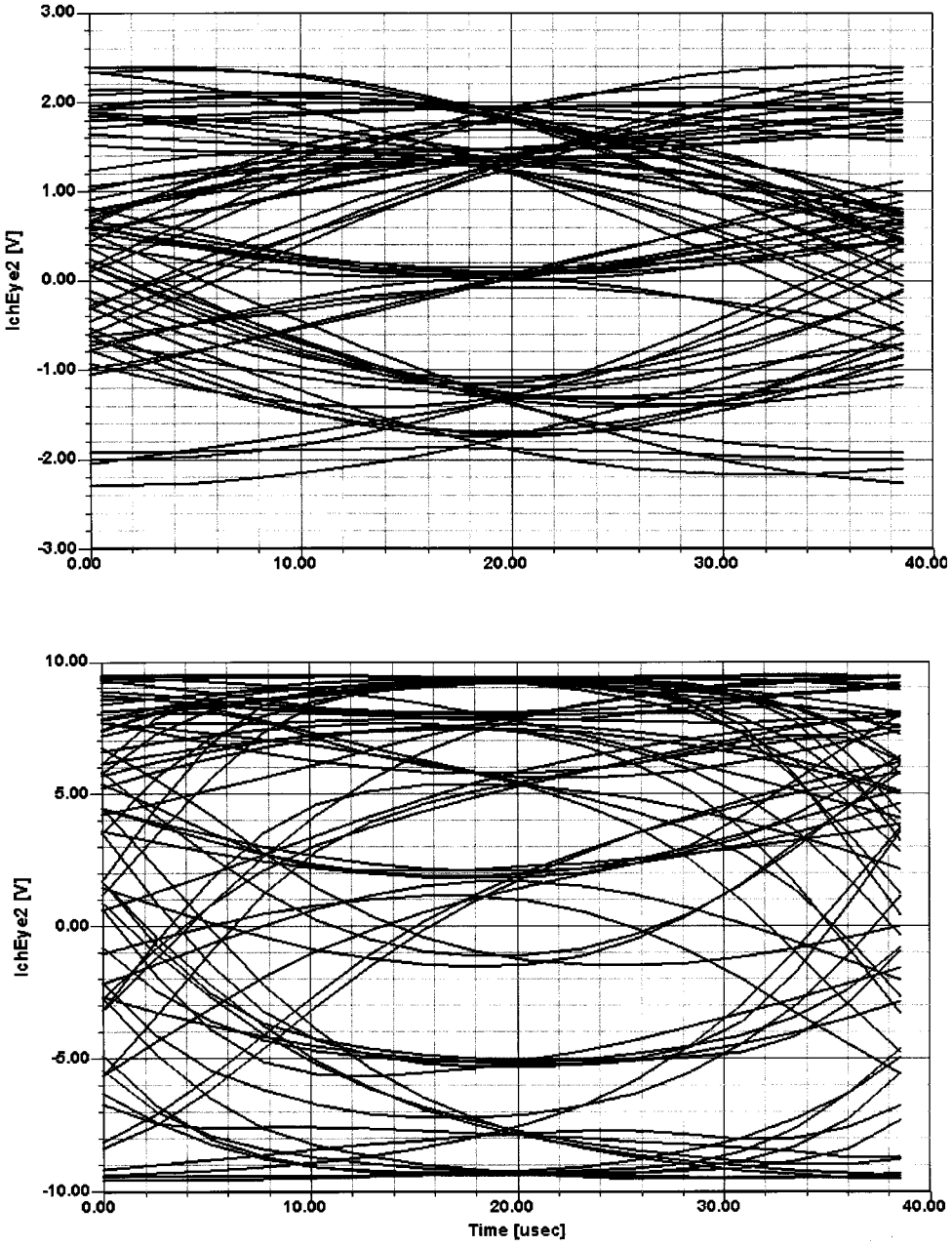


Figure 3-44 Eye diagrams for $\pi/4$-DQPSK at source powers of (a) 0 dBm and (b) 20 dBm. The distortion and related intersymbol interference are clearly visible.

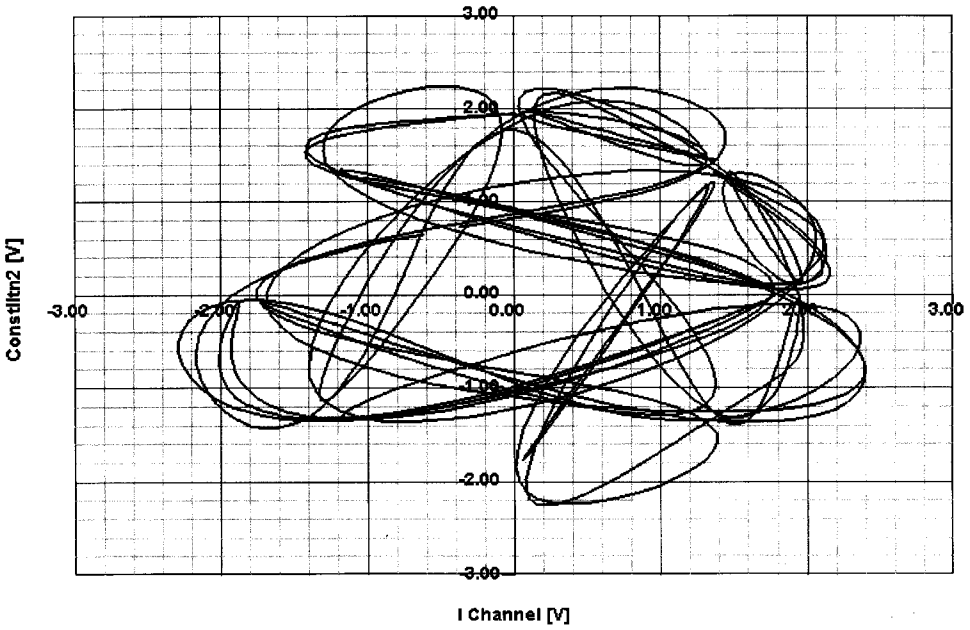


Figure 3-45 Constellation diagram for π/4-DQPSK at Port 2 and 0-dBm source power.

3-1-5 AGC

Many wireless systems consist of an up conversion or down conversion using an intermediate frequency, typically 10% or less of the receiving frequency (the FM broadcast IF has been set at 10.7 MHz for the same reason).

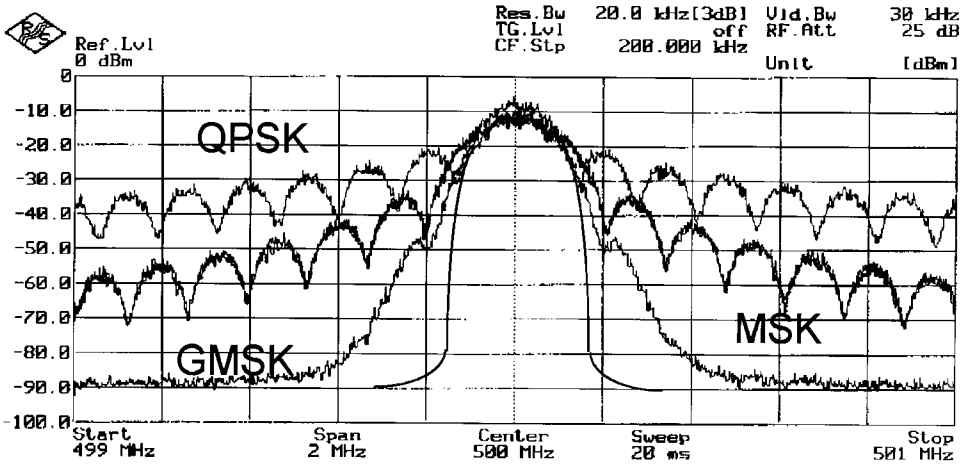


Figure 3-46 Spectrogram showing the characteristics of MSK, GMSK, and QPSK.

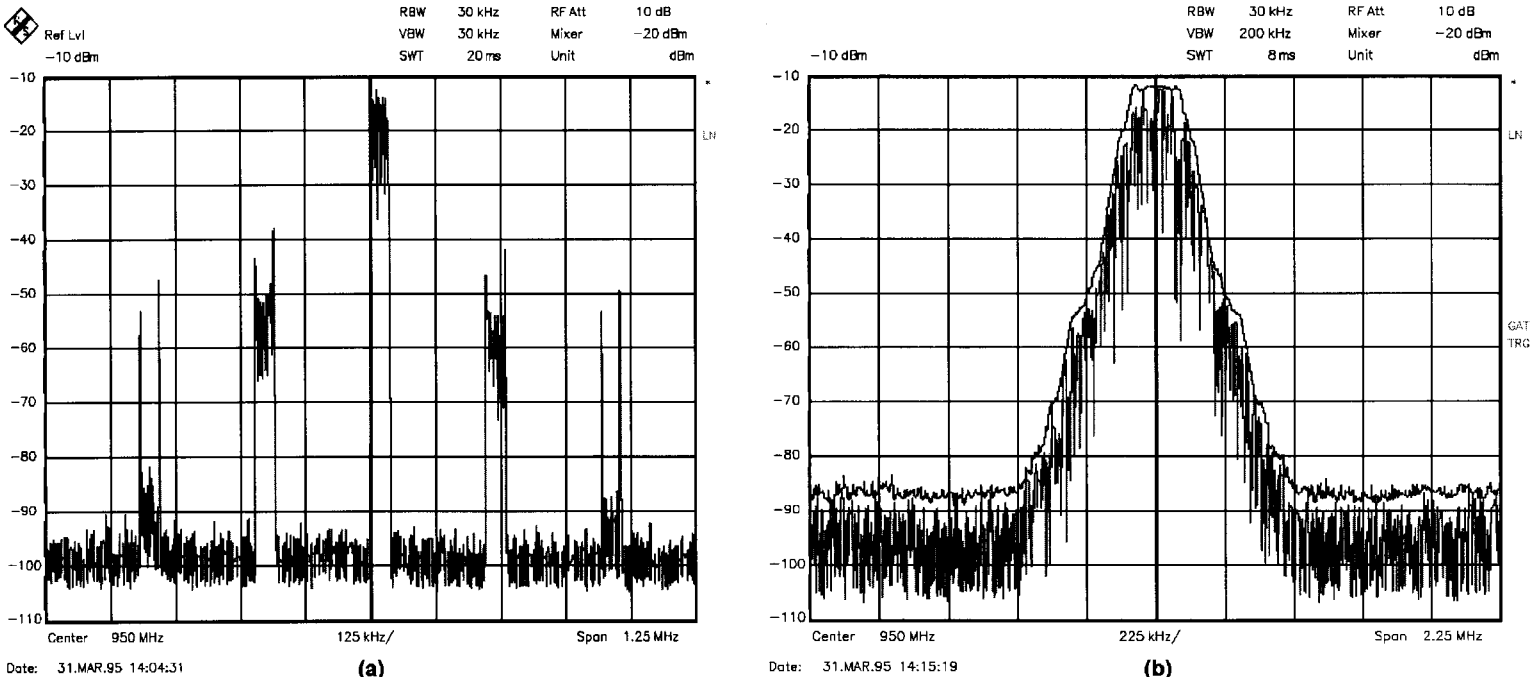


Figure 3-47 Pulsed GSM signal (a) without gating and (b) with gating.

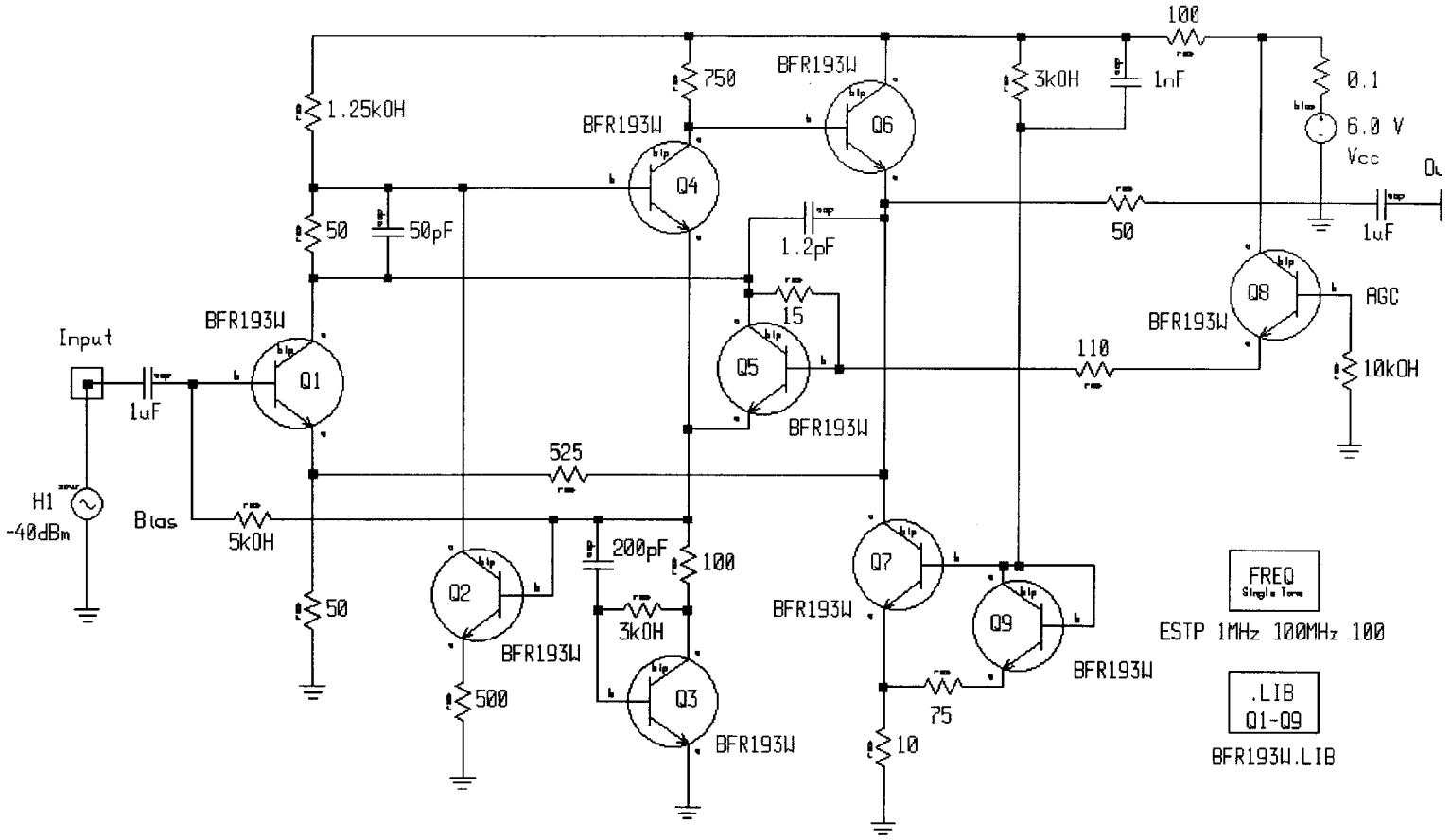


Figure 3-48 Schematic of the Plessey SL610 wideband amplifier with AGC. The 75-Ω resistor in the emitter of Q9 had to be added to obtain the appropriate dc current.

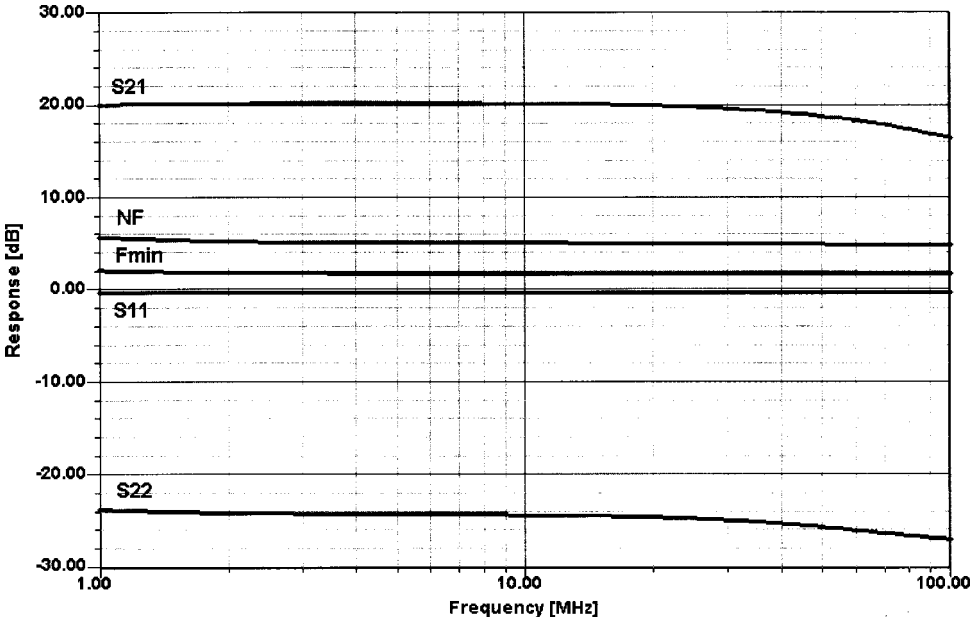


Figure 3-49 Simulated frequency response of the Plessey SL610 wideband amplifier IC with AGC = 0 V.

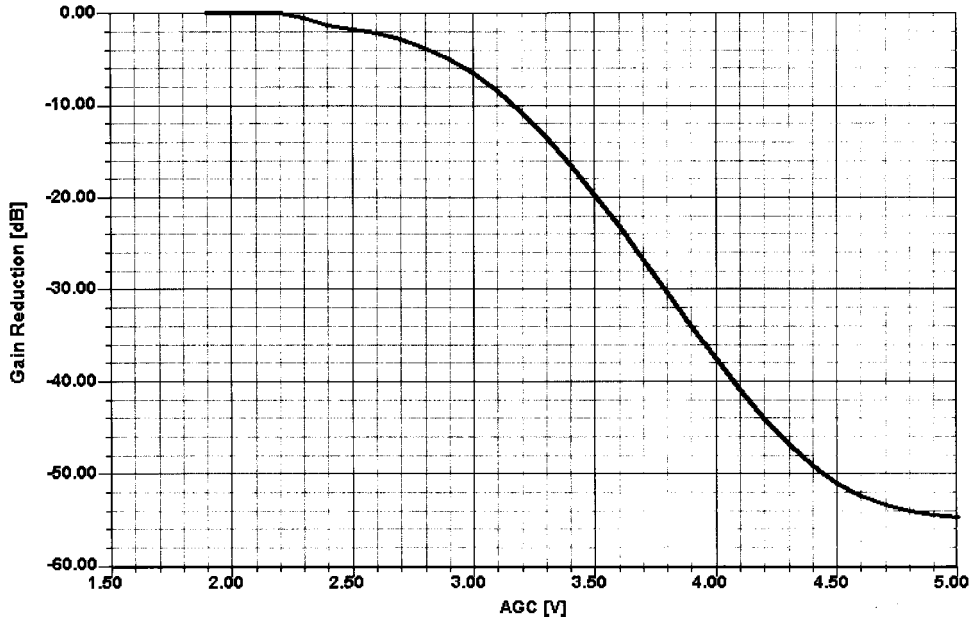


Figure 3-50 Simulated gain reduction versus AGC of the Plessey SL610 wideband amplifier IC.

Table 3-7 CGY121A features

- Variable gain amplifier (MMIC amplifier) for mobile communication
- Typical gain control range over 50 dB
- Positive control voltage
- 50-Ω input and output matched
- Low power consumption
- Operating voltage range: 2.7–6 V
- Frequency range 800 MHz to 2.5 GHz

We therefore will show two amplifiers with AGC and available data. The first one is a wideband amplifier type SL610, which has been manufactured by Plessey; its current fate is unknown. However, since we put the circuit into the simulator it provides significant insight into the operation of such AGC stages. Figure 3-48 shows the actual schematic of the SL610 entered in Ansoft's Serenade Design Environment for simulation.

The AGC action is derived from Q8 through Q5, which, together with Q4, forms a differential amplifier. Each consecutive stage (Q1, Q4, Q6) has increasing current, totaling about 20 mA; the last transistor takes about 13 mA. Based on the feedback surrounding the transistors, the actual transistor characteristics are less important. We have shown this by putting much higher f_T devices into the circuit than were available for the original design (20 years ago). The circuit is still interesting today because of its good large-signal-handling capability.

Table 3-8 CGY121A electrical characteristics at 900 MHz and 1.8 GHz

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
(T_A 25 °C, f = 900 MHz, V_g = –4 V, $RS = RL$ = 50 Ω unless otherwise specified)					
Power gain V_d = 3 V; I = 45 mA; V_{con} = 3 V	G	17	19	—	dB
Input return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{in}	—	11	—	dB
Output return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{out}	—	10	—	dB
Gain control range V_{con} = 3 V ... 0 V; V_d = 3 V; I = 45 mA	dG	48	53	—	dB
1-dB gain compression V_d = 3 V; I = 45 mA; V_{con} = 3 V	P_{1dB}	—	14	—	dBm
(T_A 25 °C, f = 1800 MHz, V_g = –4 V, $RS = RL$ = 50 Ω unless otherwise specified)					
Power gain V_d = 3 V; I = 45 mA; V_{con} = 3 V	G	15.5	17.5	—	dB
Input return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{in}	—	10	—	dB
Output return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{out}	—	8.5	—	dB
Gain control range V_{con} = 3 V ... 0 V; V_d = 3 V; I = 45 mA	dG	48	53	—	dB
1-dB gain compression V_d = 3 V; I = 45 mA; V_{con} = 3 V	P_{1dB}	—	14	—	dBm

Figure 3-49 shows the amplifier's frequency response which is extremely close to the measured data. The good simulation also can be seen from comparing the published AGC curve with the one we found in the simulation. The bias of the first transistor remains essentially constant, resulting in a noise figure more independent of the AGC than the PIN-diode attenuator, and also stabilizes the input impedance as a function of AGC. Figure 3-50 shows the SL610's predicted AGC response.

The same requirement is applicable for the wireless frequency range. Tables 3-7 and 3-8 and Figures 3-51 through 3-54 present specifications for the recently introduced CGY121A GaAs MMIC by Infineon Technologies AG, formerly the Semiconductor Group part of Siemens AG.

3-1-6 Bias and Power Voltage and Current (Power Consumption)

Biasing of transistors operating in Class A or B follows some standard rules. The only deviation from this tends to be when the operating voltage is very low, below 3 V, because then the luxury of using voltage drops for good dc stability as a function of temperature disappears. (We will discuss low-voltage design in detail in Section 3-2-6.) In our simple example (Figure 3-55), we are operating from a 12-V power supply, and the first rule of thumb is to make sure that the voltage between the emitter and ground is at least 0.7 V. This is necessary to compensate the base–emitter junction threshold voltage, which decreases by about 2 mV/°C as the temperature increases. A larger voltage drop in the emitter–ground connection greatly reduces the influence. In this amplifier the emitter–ground path consists of an emitter unbypassed resistor of 10 Ω in series with a bypassed 501-Ω resistor, resulting in an overall voltage drop of 2.6. We could reduce the supply voltage by about 1.9 V at the collector side and still maintain a safety margin of 0.7 V at the emitter–ground connection. The current 2.6 V, so to speak, is overkill.

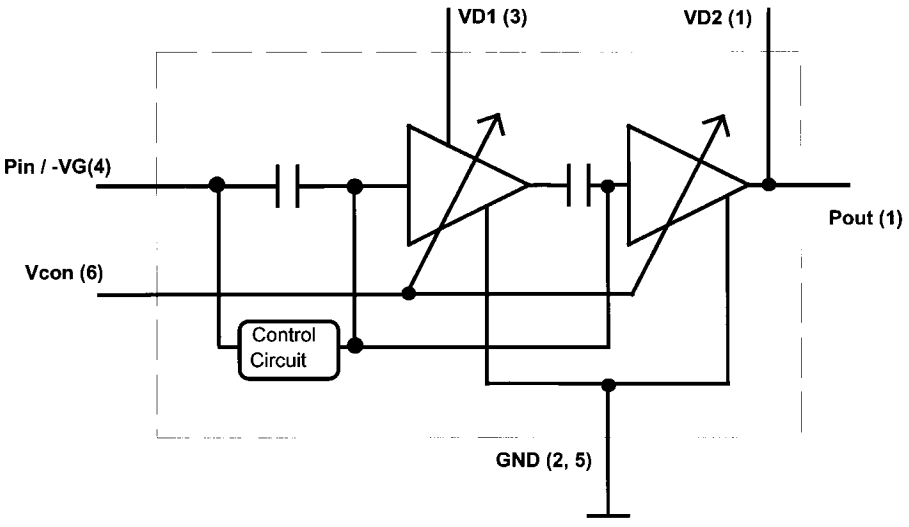
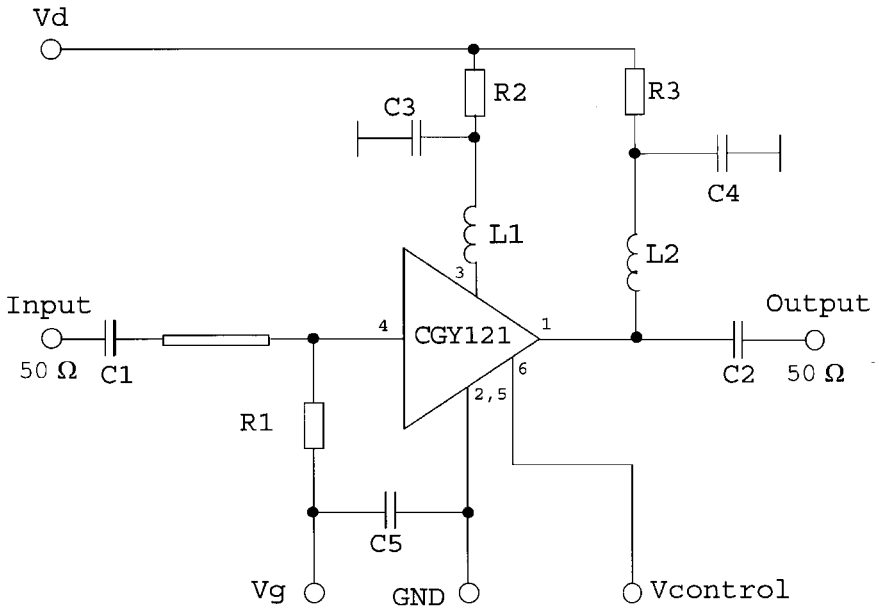


Figure 3-51 Functional block diagram of the gain-controllable CGY121A GaAs MMIC. Gain control is achieved by applying 0–3 V dc to the Vcon pin.



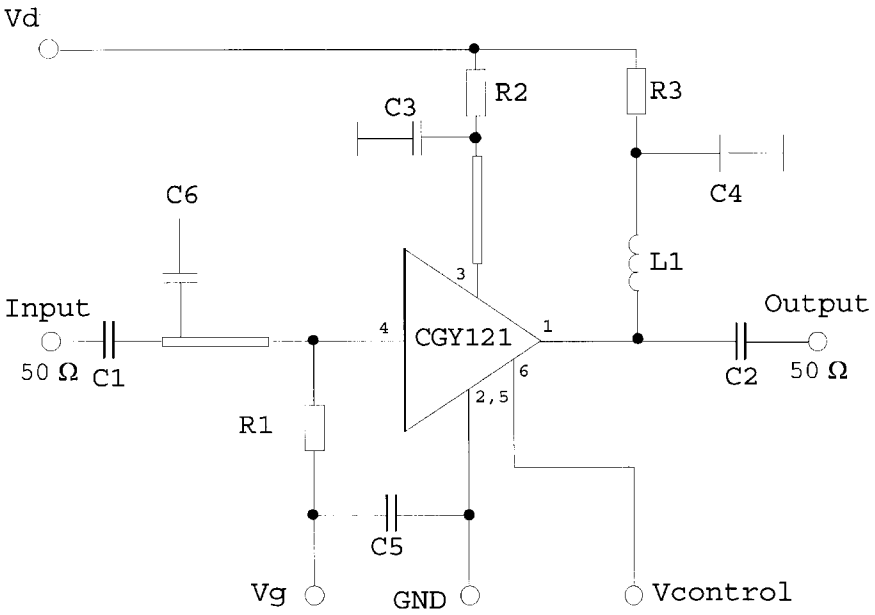
Parts List

Frequency	900 MHz
C1, C2 (Siemens Size 0603)	22 pF
C3, C4 (Siemens Size 0603)	100 nF
C5 (Siemens Size 0603)	47 nF
L1 (Coilcraft 0805CS-150XKBC)	15 nH
L2 (Coilcraft 0805CS-270XMBC)	27 nH
R1 (Siemens B 54102-A1271-J60)	270 Ω
R2 (Siemens B 54102-A1120-J60)	12 Ω
R3	6.8 Ω

Figure 3-52 CGY121A 900-MHz application circuit and parts list.

On the collector side, for an *RC* type of amplifier, it is not a bad idea to set the collector voltage at about half the supply, which allows the RF swing to be half the supply voltage. In this case, we have violated the rule because of our 2.6-V drop. Assuming we had chosen a 0.7-V drop, then the collector–emitter voltage would be 6.3 V available for the swing. The choice of 5 mA for the transistor has come from a combination of good linearity and low noise figure. While the 10-Ω unbypassed resistor will grade the noise figure somewhat, it reduces the emitter-current-dependent distortion. The base–emitter diffusion resistance equals 26 mV/5 mA, or roughly 5 Ω. The 10-Ω resistor is twofold larger and therefore adds to the linearity and to the noise.

To determine the base bias resistor it is a good assumption that the current going through the two resistors (in our case, 8.2 and 3.3 kΩ) is about 10% (or a slightly higher fraction) of



Parts List

Frequency	1.9 GHz
C1, C2 (Siemens size 0603)	12 pF
C3, C4 (Siemens size 0603)	100 nF
C5 (Siemens size 0603)	47 nF
C6 (Siemens size 0603)	1.2 pF
L1 (Coilcraft 0805CS-150XKBC)	15 nH
R1 (Siemens B 54102-A1271-J60)	270 Ω
R2 (Siemens B 54102-A1120-J60)	12 Ω
R3	2.7 Ω

Figure 3-53 CGY121A 1.9-GHz application circuit and parts list.

the collector current. Since the dc current gain of the transistor is typically between 50 and 100, it is sufficient if the current through the resistive network is between 5 and 10 times higher than the dc bias current taken off the network. This type of bias scheme is only valid for Class A to AB_1. Later, we will see that for monolithic circuits, additional transistors will be used to generate the required voltage drops. Lately, combinations of *npn* and *pnp* transistors have become popular to avoid using any resistance in the emitter for dc biasing purposes. Figure 3-56 shows such a circuit. This, of course, invites thermal runaway. To minimize the circuitry, there are bias ICs, such as the Siemens BCR400R, now available. Figures 3-57 and 3-58 show its application in biasing BJTs and FETs, respectively; it can also be used to control PIN-diode bias current for TR antenna switching, as shown in Figure 3-59.

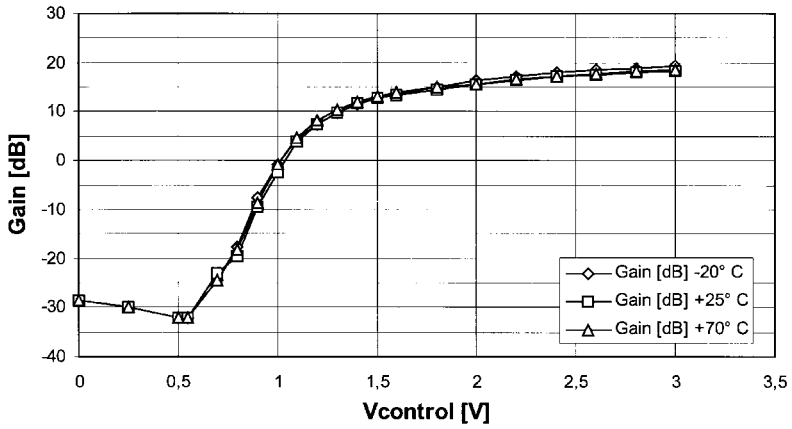


Figure 3-54 CGY121A gain versus $V_{control}$ at 1.9 GHz.

To validate this, we have used the Serenade 8.0 CAD tool, which has a built-in bias- and temperature-dependent noise model for BJTs and FETs. The SPICE-type dc analysis determines the actual currents, which were mentioned above. By maintaining the same bias and just bypassing both emitter resistors at the same time, the set of curves shown in Figure 3-60 will be obtained. As an exercise, we recommend the user to determine the optimum generator impedance for F_{min} with a 10-Ω resistor. Again, both the emitter current and collector voltage affect the noise figure, the intermodulation, and f_T. If a transistor is used

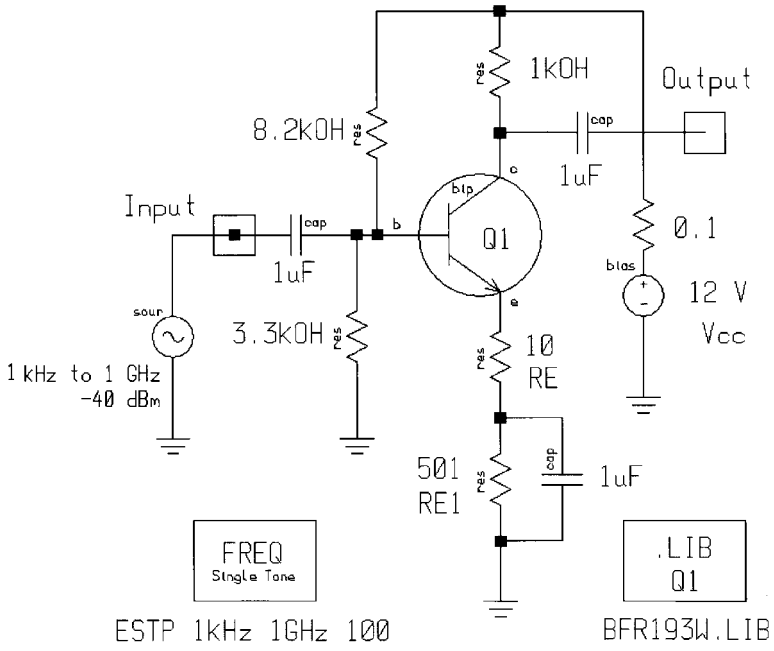


Figure 3-55 BJT amplifier example.

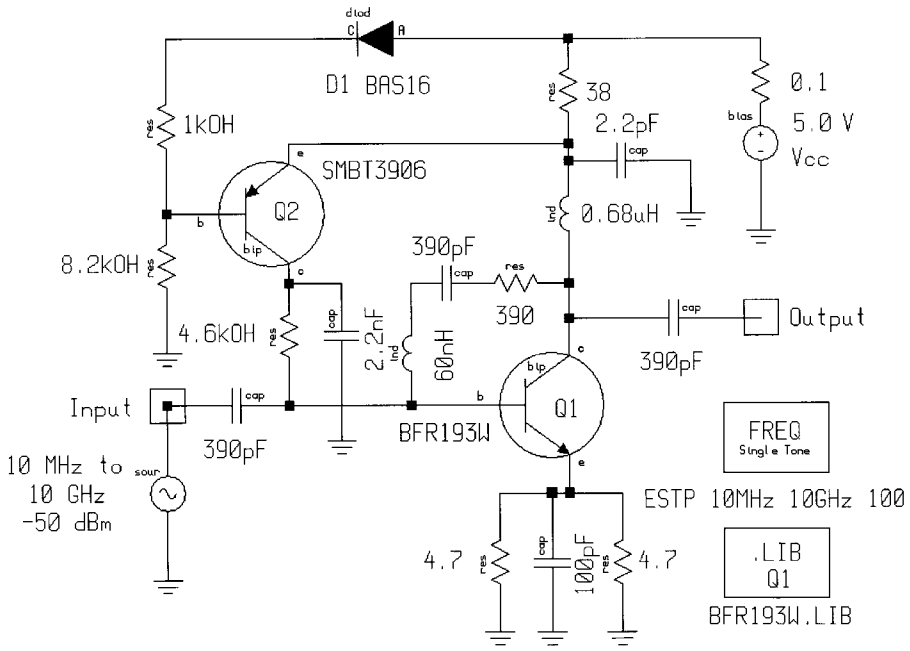


Figure 3-56 Wideband amplifier with active biasing. The 4.7-Ω resistors in the BFR193W emitter compensate the circuit's gain at low frequencies.

that is intended to be operated at significantly higher currents, then f_T will suffer while the other two parameters are less affected.

The dc input power translates into thermal dissipation, and Figure 3-60 is a plot of the noise figure as a function of temperature with and without emitter feedback. The heavy dc stabilization prevents a significant effect on the other parameters. The actual dc shift is from 5.0 to 5.2 mA for a temperature change of 70 K.

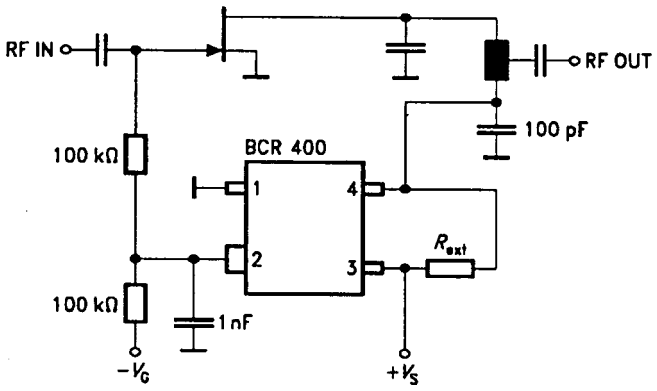


Figure 3-57 The Siemens BCR400 bias controller applied to a GaAsFET.

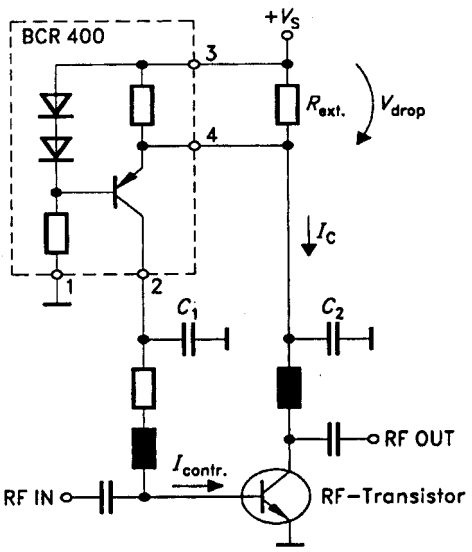


Figure 3-58 The Siemens BCR400 bias controller applied to a BJT.

As far as FETs are concerned, the enhancement types, typically PMOS, can use a similar type of bias, while the NMOS types complicate the designer's life as they require a negative gate voltage. For low-noise operation, these devices are operated somewhat close to the pinchoff voltage. Putting a source resistor in place causes mostly headaches simply because the gain is so high that the transistors are always ready to find a frequency at which they love to oscillate. The transistor will look for every possible parasitic to turn the amplifier into an oscillator.

3-2 AMPLIFIER GAIN, STABILITY, AND MATCHING

Most semiconductor manufacturers describe their devices these days by providing either their *S* parameters (small-signal) or SPICE-type parameters (large-signal). One of our headaches in writing this book has been that many of the published models don't translate

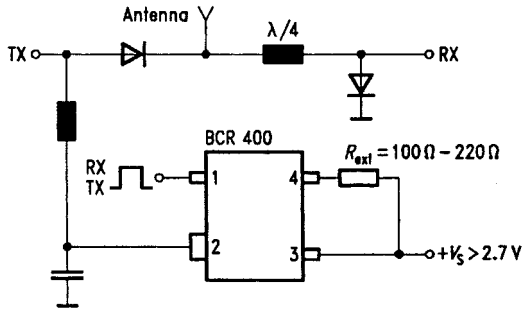


Figure 3-59 The Siemens BCR400 bias controller applied to TR switching.

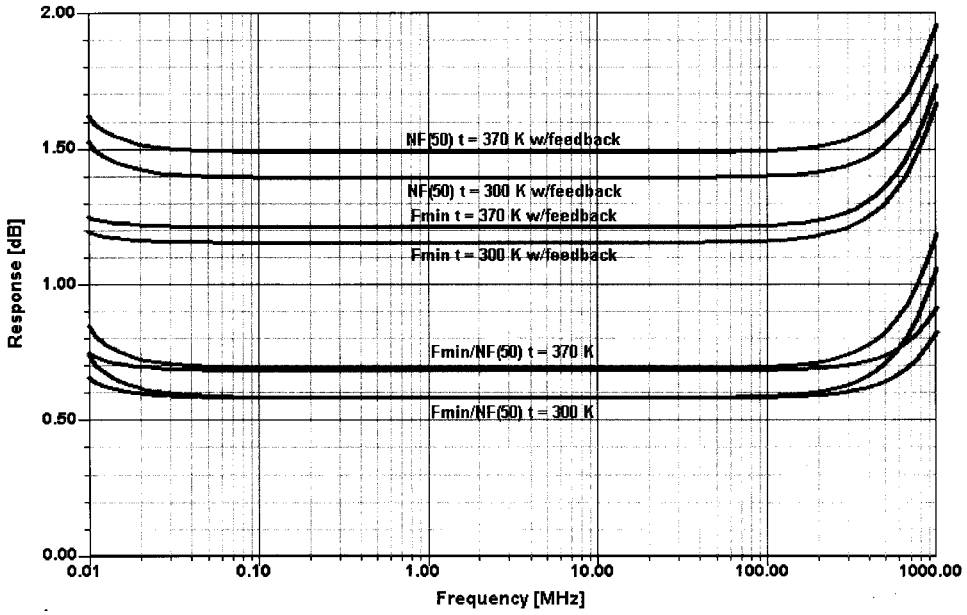


Figure 3-60 F_{min} and NF for the amplifier in Figure 3-55, with and without emitter feedback, at 300 and 370 K.

from large signal to small signal without giving unexpected or false results. Despite the availability of tools to do it right, many companies still can only provide data that are practically inadequate, and unfortunately the CAD manufacturers then get the blame for results for which they are not really responsible.

Tables 3-9 and 3-10 show typical data for a bipolar transistor (the BFP420 by Siemens). The next section presents a summary of the derivation and use of *S* parameters as published by Hewlett-Packard and others.

3-2-1 Scattering Parameter Relationships

Scattering (*S*) parameters characterize a network in terms of incident and reflected waves (Figures 3-61 and 3-62).

$$b_1 = s_{11}a_1 + s_{12}a_2 \tag{3-122}$$

$$b_2 = s_{21}a_1 + s_{22}a_2 \tag{3-123}$$

Notice that

$$s_{11} = \frac{b_1}{a_1} = \frac{V_1/I_1 - Z_0}{V_1/I_1 + Z_0} = \frac{Z_1 - Z_0}{Z_1 + Z_0} \tag{3-124}$$

and

Table 3-9 BFP420 common-emitter *S* parameters[a]

f	S_{11}		S_{21}		S_{12}		S_{22}	
(GHz)	(Magnitude)	(Angle)	(Magnitude)	(Angle)	(Magnitude)	(Angle)	(Magnitude)	(Angle)
				V_{CE} = 2 V, I_C = 20 mA				
0.01	0.452	–2.3	37.62	178.3	0.0011	94.4	0.956	–0.6
0.1	0.447	–52.1	36.30	164.7	0.0068	82.5	0.941	–12.4
0.5	0.386	–101.1	23.41	121.0	0.0262	61.7	0.632	–47.2
1.0	0.378	–146.2	13.99	96.0	0.0395	57.8	0.395	–63.9
2.0	0.405	173.5	7.18	70.8	0.0664	54.0	0.222	–87.3
3.0	0.446	149.4	4.77	52.6	0.0949	47.1	0.133	–111.3
4.0	0.501	130.0	3.52	36.8	0.1206	38.5	0.133	–158.5
6.0	0.599	104.8	2.27	8.2	0.1646	18.9	0.196	142.0
8.0	0.700	78.5	1.51	–20.8	0.1800	–2.4	0.289	99.3
9.0	0.758	67.6	1.25	–34.4	0.1820	–13.0	0.379	84.1
10.0	0.800	62.0	1.04	–43.5	0.1800	–19.3	0.465	76.6
				V_{CE} = 2 V, I_C = 5 mA				
0.01	0.790	–1.0	15.14	179.2	0.0012	83.4	0.988	–0.7
0.1	0.786	–11.6	14.98	171.8	0.0092	84.1	0.982	–6.5
0.5	0.702	–55.7	12.86	140.1	0.0398	62.8	0.857	–29.8
1.0	0.589	–99.1	9.63	112.6	0.0603	46.5	0.647	–48.6
2.0	0.507	–156.0	5.60	79.4	0.0798	34.6	0.401	–70.3
3.0	0.511	168.5	3.84	57.1	0.0957	29.8	0.267	–84.2
4.0	0.549	142.0	2.87	38.5	0.1121	25.1	0.207	–110.5
5.0	0.604	123.9	2.26	22.1	0.1285	19.4	0.150	–137.3
6.0	0.633	110.0	1.86	6.7	0.1442	13.1	0.173	–169.8

Table 3-10 BFP420 noise parameters

f	$F_{\min}$[a]	G_a[a]	Γ_{opt}		R_N	r_n	$F_{50\Omega}$[b]	$\lvert S_{21} \rvert^2$ [b]
(GHZ)	dB	(dB)	(Magnitude)	(Angle)	(Ω)		(dB)	(dB)
				V_{CE} = 2 V, I_C = 5 mA				
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	–175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	–157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	–142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	–123.0	15.0	0.30	3.30	5.3

[a]Input matched for minimum noise figure, output for maximum gain.
[b]$Z_S = Z_L = 50\ \Omega$.

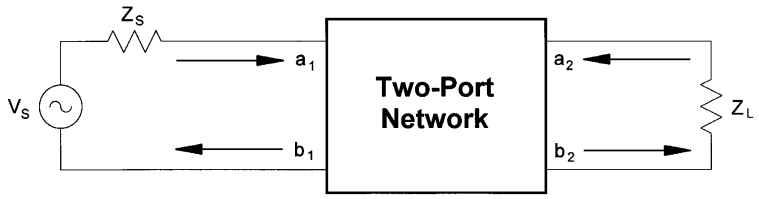


Figure 3-61 *S* parameters characterize a network (a two-port is shown) in terms of incident waves (a_1, a_2) and reflected waves (b_1, b_2). Figure 3-62 shows the flow graph for this network.

$$Z_1 = Z_0 \frac{(1 + s_{11})}{(1 - s_{11})} \tag{3-125}$$

where $Z_1 = V_1/I_1$ is the input impedance at Port 1.

This relationship between reflection coefficient and impedance is the basis of the Smith Chart transmission-line calculator. Consequently, the reflection coefficients S_{11} and S_{22} can be plotted on Smith charts, converted directly to impedance, and easily manipulated to determine matching networks for optimizing a circuit design.

The above equations show one of the important advantages of *S* parameters: They are simply gains and reflection coefficients, both familiar quantities to engineers. By comparison, some of the *Y* parameters described earlier in the article are not so familiar. For example, the *Y* parameter corresponding to insertion gain S_{21} is the "forward transadmittance," Y_{21}. Clearly, insertion gain gives by far the greater insight into the operation of the network.

S parameters are simply related to power gain and mismatch loss, quantities that are often of more interest than the corresponding voltage functions:

$$|s_{11}|^2 = \frac{\text{Power reflected from the network input}}{\text{Power incident on the network input}} \tag{3-126}$$

$$|s_{21}|^2 = \frac{\text{Power delivered to a } Z_0 \text{ load}}{\text{Power available from } Z_0 \text{ source}} \tag{3-127}$$

$$|s_{12}|^2 = \text{Reverse transducer power gain with } Z_0 \text{ load and source.} \tag{3-128}$$

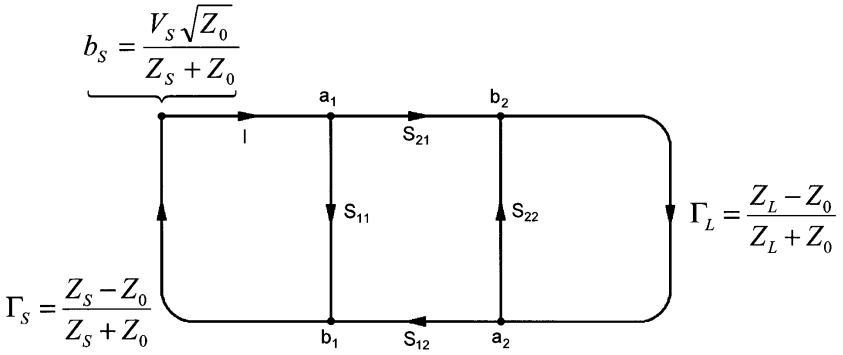


Figure 3-62 Flow graph for the network shown in Figure 3-61.

$$|s_{22}|^2 = \frac{\text{Power reflected from the network output}}{\text{Power incident on the network output}} \tag{3-129}$$

From this we obtain:
Input reflection coefficient with arbitrary Z_L:

$$s'_{11} = s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1 - s_{22}\Gamma_L} \tag{3-130}$$

Output reflection coefficient with arbitrary Z_S:

$$s'_{22} = s_{22} + \frac{s_{12}s_{21}\Gamma_S}{1 - s_{11}\Gamma_S} \tag{3-131}$$

Voltage gain with arbitrary Z_L and Z_S:

$$A_V = \frac{V_2}{V_1} = \frac{s_{21}(1 + \Gamma_L)}{(1 - s_{22}\Gamma_L)(1 + s'_{11})} \tag{3-132}$$

$$\text{Power gain} = \frac{\text{Power delivered to load}}{\text{Power input to network}}:$$

$$G = \frac{|s_{21}|^2 (1 - |\Gamma_L|^2)}{(1 - |s_{11}|^2) + |\Gamma_L|^2 (|s_{22}|^2 - |D|^2) - 2\,\text{Re}\,(\Gamma_L N)} \tag{3-133}$$

$$\text{Available power gain} = \frac{\text{Power available from network}}{\text{Power available from source}}:$$

$$G_A = \frac{|s_{21}|^2 (1 - |\Gamma_S|^2)}{(1 - |s_{22}|^2) + |\Gamma_S|^2 (|s_{11}|^2 - |D|^2) - 2\,\text{Re}\,(\Gamma_S M)} \tag{3-134}$$

$$\text{Transducer power gain} = \frac{\text{Power delivered to load}}{\text{Power available from source}}:$$

$$G_T = \frac{|s_{21}|^2 (1 - |\Gamma_S|^2)(1 - |\Gamma_L|^2)}{|(1 - s_{11}\Gamma_S)(1 - s_{22}\Gamma_L) - s_{12}s_{21}\Gamma_L\Gamma_S|^2} \tag{3-135}$$

Unilateral transducer power gain ($S_{12} = 0$):

$$G_{Tu} = \frac{|s_{21}|^2(1 - |\Gamma_S|^2)(1 - |\Gamma_L|^2)}{|1 - s_{11}\Gamma_S|^2|1 - s_{22}\Gamma_L|^2}$$

$$= G_0 G_1 G_2 \tag{3-136}$$

$$G_0 = |s_{21}|^2 \tag{3-137}$$

$$G_1 = \frac{1 - |\Gamma_S|^2}{|1 - s_{11}\Gamma_S|^2} \tag{3-138}$$

$$G_2 = \frac{1 - |\Gamma_L|^2}{|1 - s_{22}\Gamma_L|^2} \tag{3-139}$$

Maximum unilateral transducer power gain when $|S_{11}| < 1$ and $|S_{22}| < 1$:

$$G_u = \frac{|s_{21}|^2}{|(1 - |s_{11}|^2)(1 - |s_{22}|)^2|}$$

$$= G_0\, G_{1\max}\, G_{2\max} \tag{3-140}$$

$$G_{i\max} = \frac{1}{1 - |s_{ii}|^2} \quad i = 1, 2 \tag{3-141}$$

Constant-gain circles (unilateral case: $S_{12} = 0$):

- Center of constant-gain circle is on line between center of Smith Chart and point representing S_{ii}^*
- Distance of center of circle from center of Smith Chart:

$$r_i = \frac{g_i |s_{ii}|}{1 - |s_{ii}|^2(1 - g_i)} \tag{3-142}$$

- Radius of circle:

$$\rho_i = \frac{\sqrt{1 - g_i}(1 - |s_{ii}|^2)}{1 - |s_{ii}|^2(1 - g_i)} \tag{3-143}$$

where $i = 1, 2$ and

$$g_i = \frac{G_i}{G_{i\max}} = G_i(1 - |s_{ii}|^2) \tag{3-144}$$

Unilateral figure of merit:

$$u = \frac{|s_{11}s_{22}s_{12}s_{21}|}{|(1 - |s_{11}|^2)(1 - |s_{22}|^2)|} \tag{3-145}$$

Error limits of unilateral gain calculation:

$$\frac{1}{(1 + u^2)} < \frac{G_T}{G_{Tu}} < \frac{1}{(1 - u^2)} \tag{3-146}$$

Conditions for absolute stability. No passive source or load will cause a network to oscillate if a, b, and c are all satisfied:

$$\text{a.}\quad |s_{11}| < 1, |s_{22}| < 1 \tag{3-147}$$

$$\text{b.} \quad \left| \frac{|s_{12}s_{21}| - |M^*|}{|s_{11}|^2 - |D|^2} \right| > 1 \tag{3-148}$$

$$\text{c.} \quad \left| \frac{|s_{12}s_{21}| - |N^*|}{|s_{22}|^2 - |D|^2} \right| > 1 \tag{3-149}$$

where

$$D = s_{11}s_{22} - s_{12}s_{21} \tag{3-150}$$

$$M = s_{11} - Ds_{22}^* \tag{3-151}$$

$$N = s_{22} - Ds_{11}^* \tag{3-152}$$

Condition that a two-port network can be simultaneously matched with a positive real source and load:

$$K > 1 \tag{3-153}$$

$$K = \frac{1 + |D|^2 - |s_{11}|^2 - |s_{22}|^2}{2|s_{12}s_{21}|} \tag{3-154}$$

Source and load for simultaneous match

$$\Gamma_{mS} = M^* \left[\frac{B_1 \pm \sqrt{B_1^2 - 4|M|^2}}{2|M|^2} \right] \tag{3-155}$$

$$\Gamma_{mL} = N^* \left[\frac{B_2 \pm \sqrt{B_2^2 - 4|N|^2}}{2|N|^2} \right] \tag{3-156}$$

where

$$B_1 = 1 + |s_{11}|^2 - |s_{22}|^2 - |D|^2 \tag{3-157}$$

$$B_2 = 1 + |s_{22}|^2 - |s_{11}|^2 - |D|^2 \tag{3-158}$$

Maximum available power gain. If $K > 1$,

$$G_{A\max} = \left| \frac{s_{21}}{s_{12}} (K \pm \sqrt{K^2 - 1}) \right| \tag{3-159}$$

In conclusion, *a two-port network (active or passive) is unconditionally stable* when

$$K > 1$$

$$|S_{12}S_{21}| < 1 - |S_{11}|^2$$

and

$$|S_{12}S_{21}| < 1 - |S_{22}|^2$$

3-2-2 Low-Noise Amplifiers

A low-noise amplifier combines a low noise figure, reasonable gain, and stability without oscillation over its entire useful frequency range. These amplifiers are typically operated in Class A, which is characterized by a bias point more or less at the center of maximum current and voltage capability of the device used, and by RF current and voltages that are sufficiently small relative to the bias point that the bias point does not shift. This means that an amplifier operating at, say, 3 V and 1 mA for low-noise, high-gain operation will not be Class A if the RF current is larger than 0.1% of the current or the voltage swing is more than 0.1% of the current swing. These amplifiers can be designed using standard linear *S* parameters, which are (hopefully) available from the device manufacturer for the bias point of interest. If suitable *S* parameters are not available, we must resort to a CAD tool that allows us to determine a device's bias point as a function of its bias network and to generate small-signal *S* parameters for the device. With suitable *S* parameters in hand, and assuming the amplifier will be frequency selective, we then connect tuned circuits to the input and output to provide matching.

In designing amplifiers, the "power gain" of a circuit is mentioned frequently. This is actually deceptive because power gain is somewhat associated with dc power dissipation, and in considering it, we are on the verge of discussing large-signal performance. Therefore it would be better to define the power gain as a function of operating mode, such as "Class A power gain." We have also outlined that there is a difference between noise matching and gain matching. If we can be allowed to step back into the vacuum-tube era for a moment, the reader should be reminded that the difference between power gain and noise matching had to do with the grid-plate feedback capacitance which, via the Miller effect, showed up in parallel with the input. This actually detunes the input slightly off-center relative to the best gain. The reason for this is that as the feedback increases, Y_{12} or S_{12} starts playing a role, and the frequently assumed unilateral case no longer exists. The difference between noise matching and gain matching increases as a function of frequency. It can be brought closer together if appropriate reactive feedback is used. The two choices are: (1) increase the emitter or source inductance or (2) slightly increase the base-to-collector/gate-to-drain capacitance. Needless to say, the penalty for this is potential instability, and a sweep must be done from a few megahertz to several gigahertz to make sure the stage doesn't "take off."

The most efficiency this circuit provides is 50%; even this is a theoretical limit if one tries to keep distortion under control.

While the equation $P_{\text{OUT}} = (V_{\text{bat}} - V_{\text{sat}})^2/2R_L$ is valid, it really applies only to power amplifiers in the sense that for Class A, high-gain operation R_L can be made significantly higher as the absolute power output is not the important item. While the supply voltage V_{bat} is always known, the saturation voltage V_{sat} has to be obtained by the load line dc analysis. Note: The RF saturation voltage is always higher than the dc one. When looking at Figure 3-63, it becomes obvious that the conducting angle is 180°, which implies that the entire sine wave at the input is reproduced as an essentially undistorted sine wave at the output.

We can now calculate the components of the distorted current in separating them into the dc component, the fundamental frequency, and its first and second harmonics:

$$\text{Fundamental} \quad i_{c1} = i_{\text{peak}} \cdot f_1(\Theta) \tag{3-160}$$

$$\text{Second harmonic} \quad i_{c2} = i_{\text{peak}} \cdot f_2(\Theta) \tag{3-161}$$

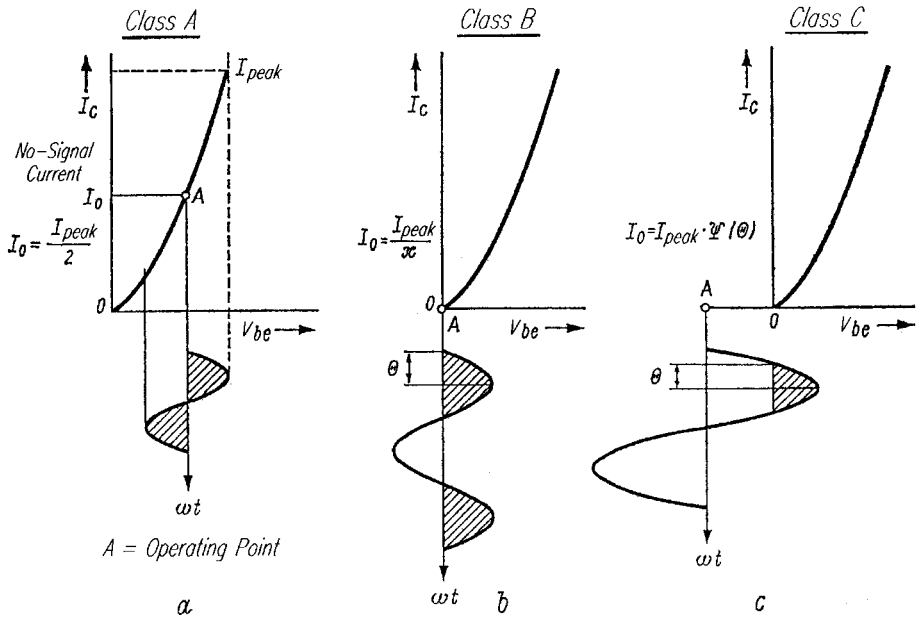


Figure 3-63 Definitions of Class A, B, and C operation, including conduction angle (Θ). Figure 3-64 shows the determination of Θ.

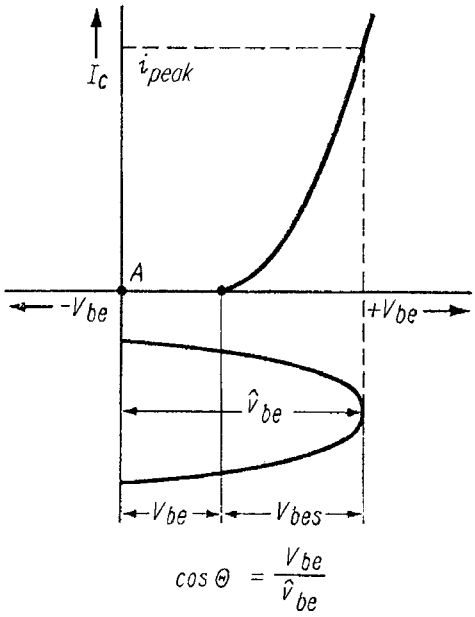


Figure 3-64 Determination of the conduction angle Θ from the base threshold voltage V_{be} and the base RF voltage $\hat{v}_{be}$.

$$\text{Third harmonic} \quad i_{c3} = i_{\text{peak}} \cdot f_3(\Theta) \tag{3-162}$$

For the collector dc current we obtain

$$I_c = i_{\text{peak}} \cdot \Psi(\Theta) \tag{3-163}$$

The frequency-dependent currents are a function of the conduction angle Θ. Figure 3-65 shows these transfer characteristics as a function of the conduction angle for a square-law dynamic transfer function; for higher-order transfer functions, the values don't change much. Tables for this can be found in various sources, including Clark and Hess [47]. Since the transistors exhibit a mixture of an exponential and a linear transfer characteristic (due to the unwanted parasitic emitter/source and collector/drain resistances when building the device), the plot in figure 3-65 is sufficiently valid for practical purposes. This function plot allows us to determine the amplitude of the current at various frequencies once the conducting angle has been set. For example, for Θ = 70°, the maximum collector current is $i_{\text{peak}} = I_c/0.253$, meaning that the peak current of the transistor is approximately four times larger than the dc value one would measure. Once the peak collector current is known, the fundamental and harmonic currents can be calculated:

$$i_{c1} = \frac{I_c}{0.253} f_1(\Theta)$$

$$i_{c1} = \frac{I_c}{0.253} 0.436$$

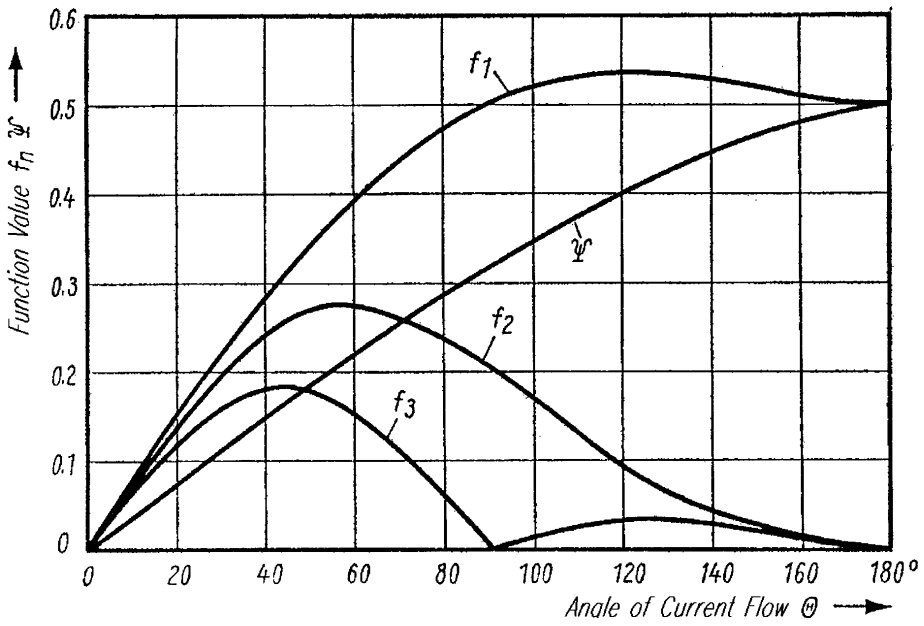


Figure 3-65 Function values for the fundamental, second harmonic, third harmonic, and dc components (f_1, f_2, f_3, and Ψ, respectively) versus device conduction angle.

In order to obtain i_{c2} and i_{c3} one needs to do the same calculations as above with the appropriate functions for f_2 and f_3.

It becomes obvious that the distorted collector current generates a lot of harmonics, which definitely requires sufficient low-pass filtering at the output, even for Class A operation. We already stated that for Class A operation the conduction angle is 180° while it is 90° for Class B operation. The conduction angle for Class C operation can be solved from the ratio of the knee voltage V_{be} and the peak value of the ac voltage at the base. We find

$$\cos \Theta = \frac{V_{be}}{\hat{V}_{be}} \tag{3-164}$$

where $\Theta = 90°$ and $V_{be} = 0$. Given the voltage V_{bes}, which determines the peak collector current (up to i_c peak), we can calculate the effective RF voltage at the base as

$$\hat{v}_{be} = \frac{v_{bes}}{1 - \cos \Theta} \tag{3-165}$$

Therefore, the peak value of v_b becomes v_{bes} for $\Theta = 90°$.

The necessary base voltage V_{be} for maximum current is obtained by the transistor model. Then one finds the threshold voltage at which the transistor becomes conductive and by using Eqs. (3-164) and (3-165) it can be checked whether this results in the appropriate conducting angle. An example for this will be shown in the high-power stage section.

Design Guidelines. The design of a one-stage amplifier consists of finding

- An input lossless matching network M_1
- An input lossless matching network M_2

so that the maximum or desired transistor gain is achieved over the operating bandwidth of the amplifier. Usually, the common-emitter or common-source configuration is chosen for highest gain per stage. If the stability factor K is greater than unity, these two networks can be found to give the maximum available gain $G_{\max}$. If the stability factor is less than or equal to unity, the amplifier could be terminated in a matching structure that causes oscillation; that is, $G_{\max}$ is infinite. This should be avoided by locating the regions of instability in the Γ_G and Γ_L planes. The input and output terminations (Γ_G and Γ_L) must be designed to avoid the instability regions. Usually, these unstable regions are near the conjugate match for S_{11} and S_{22}. Thus, a stable amplifier will require some input and/or output mismatch if K is less than or equal to unity.

There are at least two alternative approaches for potentially unstable amplifiers:

1. Add resistive matching elements to make $K \geq 1$ and $G_{\max} \approx G_{ms}$.
2. Add feedback to make $K \geq 1$ and $G_{\max \approx Gms}$.

For narrowband amplifiers, it is usually recommended that one accept a transistor with $K < 1$, design the amplifier for a gain approaching G_{ms}, and ensure that the Γ_G and Γ_L terminations provide stability at all frequencies, both inside and outside the amplifier passband.

The design of an amplifier would usually have the following specifications:

Gain and gain flatness
Bandwidth and center frequency ($f_2 - f_1, f_0$)
Noise figure
Linear output power
Input reflection coefficient ($VSWR_{in}$)
Output reflection coefficient ($VSWR_{out}$)
Bias voltage and current

For small-signal amplifiers, the small-signal S parameters are sufficient to complete the design. After selecting an appropriate transistor based on these specifications, a one-stage amplifier design should be considered if sufficient gain can be achieved; otherwise, a two-stage amplifier should be designed.

The circuit topology should be chosen to allow dc bias for the transistor. Usually, RF short-circuited stubs are placed near the transistor to allow dc biasing. If the topology does not allow dc bias, a broadband bias choke or high-resistance bias circuit that does not affect the amplifier performance must be used.

The following steps can be tabulated for the design of a one-stage amplifier (see Figure 3.66):

1. Select a transistor based on a datasheet description of the S parameters, noise figure, and linear output power.
2. Calculate K and G_{max} or G_{ms} versus frequency.
3. For $K > 1$, select the topologies that match the input and output (and allow dc biasing) at the upper band edge f_2. Ideally, this will give G_{max} and $S'_{11} = S'_{22} = 0$. Usually, $S_{12} = 0$ is assumed for the initial design. Next, the topology may be varied to flatten gain versus frequency at the expense of S'_{11} and S'_{22}.
4. For $K < 1$, plot the regions of instability on the Γ_G and Γ_L planes and select topologies that partially match the input and output at the upper band edge and avoid the unstable regions. The gain will approach G_{ms} as an upper limit. Next, the topology may be varied to flatten gain versus frequency.
5. After finding initial M_1 and M_2, plot the amplifier S parameters versus frequency; make adjustments in topology until the specifications for gain, input reflection coefficient, and output reflection are satisfied. Also plot Γ_G and Γ_L versus frequency to verify amplifier stability.
6. Design the dc bias circuit. Lay out the elements of the complete amplifier and check realizability.

As a beginning point in an amplifier design, the circuit topologies may be designed with the assumption that S_{12} is zero. Later, an exact design procedure will be described that includes the topologies with terminations in the stable region.

Design Examples. Throughout this book we always address the issue that we have to separate the battery-operated handheld applications and the more stationary designs like those for base stations. In the case of the handheld or low-voltage applications, the dynamic range requirements are significantly less than the front-end of a base station requires.

NE68133 Matched Amplifier. We will first consider the design of a low-voltage, low-current, low-noise amplifier based on the NEC BJT NE68133. This device was chosen

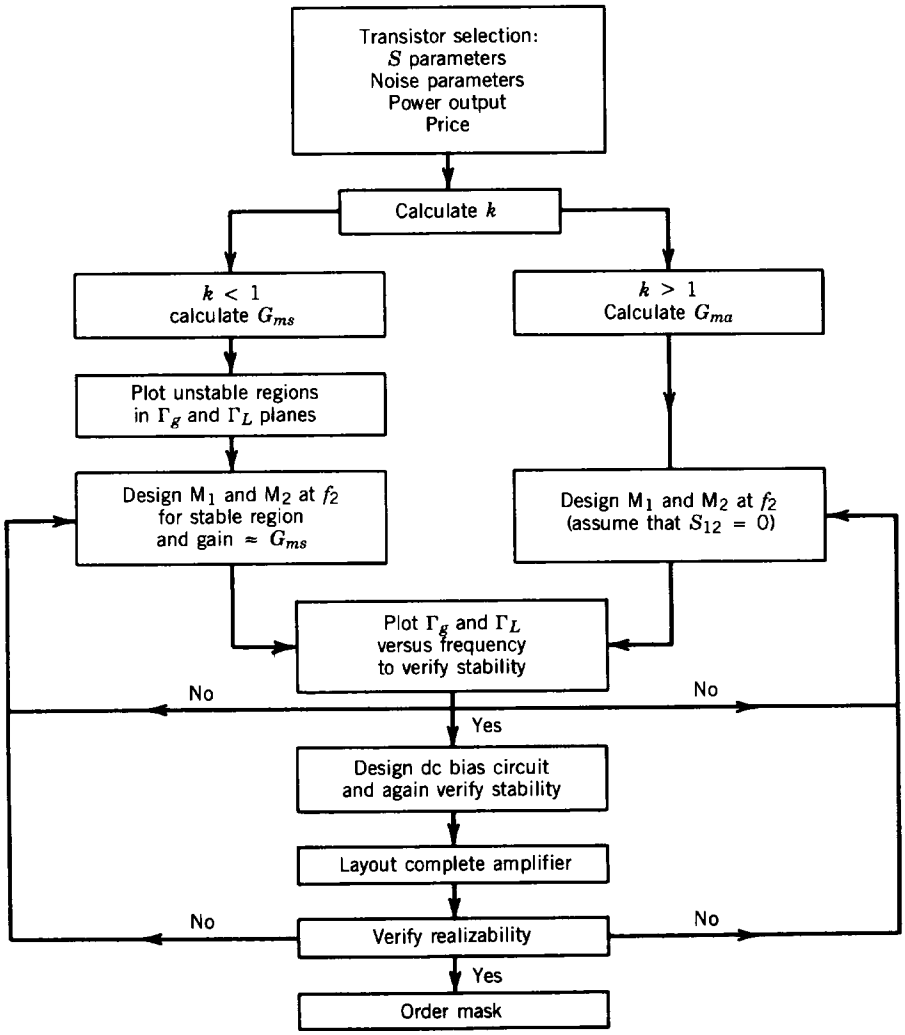


Figure 3-66 Simplified amplifier design procedure.

because it is unconditionally stable, has a minimum noise figure of about 1.2 dB at 7 mA, has 15.5 dB gain, and has an f_T of 8 GHz. While it is possible to start the design based on the published S and noise parameters, we will choose also to involve the nonlinear model, as this allows us more freedom in selecting the bias point including the collector voltage. The actual topology of the amplifier is shown in Figure 3-67.

Determining appropriate values for the input and output matching networks is essential to finalizing the design. The next several figures illustrate matching-network syntheses undertaken with the interactive Smith Tool in Ansoft's Serenade Design Environment. We evaluate the point of tangency as well as stability circles in the source and load planes, indicating that our impedance match will result in a stable amplifier.

In Figure 3-68, the center circle (with 1.1 marker) refers to the noise circle for which the noise figure of 1.1 dB can be obtained. The left circle (with the 14.3 marker) corresponds to

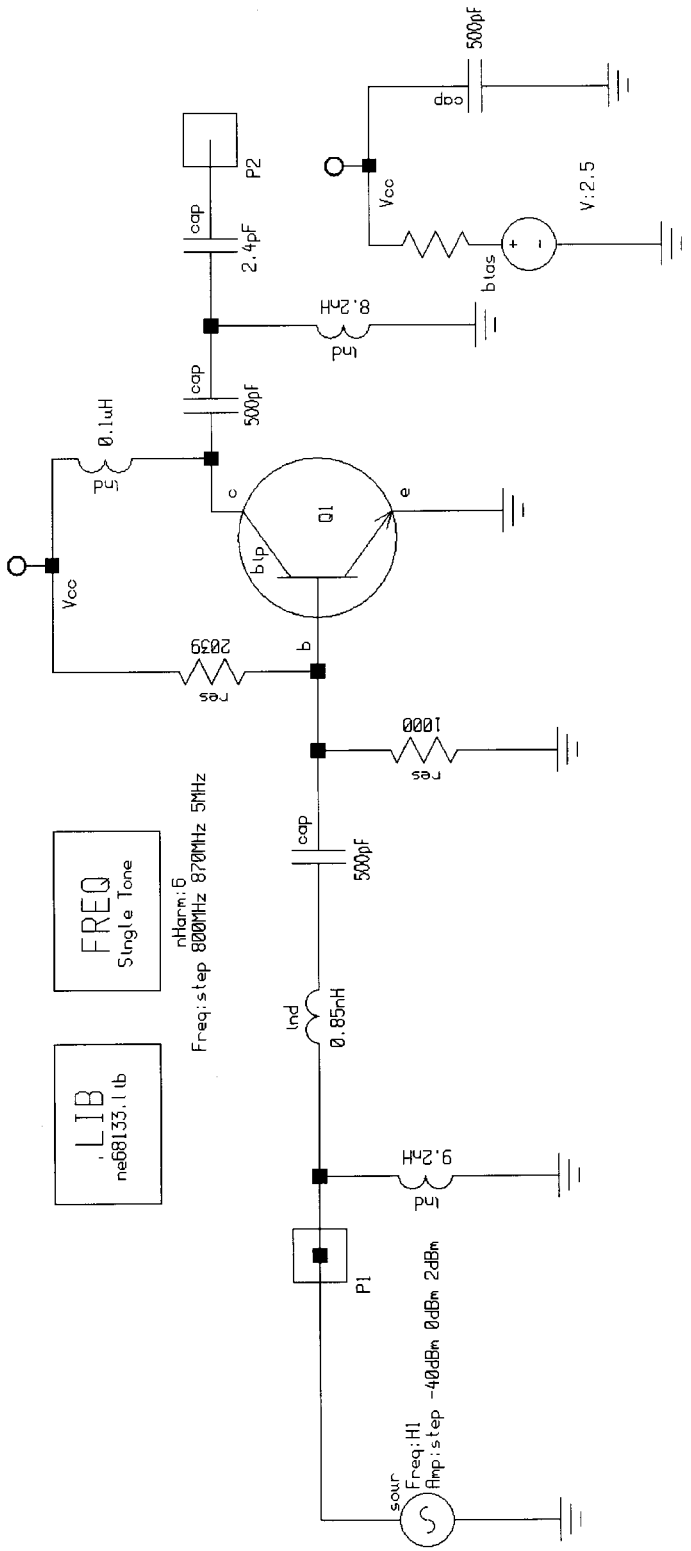


Figure 3-67 The NE68133 low-noise amplifier.

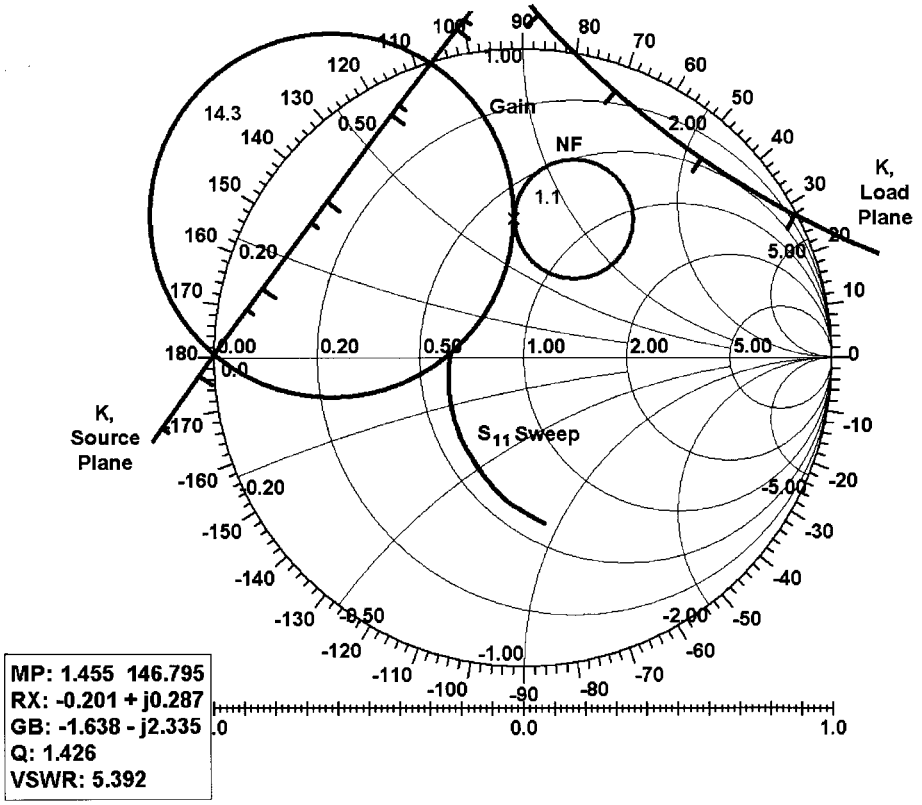


Figure 3-68 Display from the interactive Smith Tool, showing circles for gain, noise figure, and source- and load-plane stability, as well as a sweep of S_{11} versus frequency. In synthesizing an input matching circuit, we choose a single working frequency and work from the chart's origin to the point of tangency between the gain and noise figure circles. See Figure 3-69.

source-plane terminations that correspond to 14.3 dB of available gain. In order to get the required noise figure *and* gain, the input termination must be at the point of tangency of both circles. The arc touching the X axis close to 0.50 is the amplifier's S_{11} response from 800 to 875 MHz. The shallow arc that nearly bisects the 14.3-dB gain circle is actually a portion of the source-plane stability circle; the arc at the top right is a portion of the load-plane stability circle. Because we will be terminating the amplifier input with a value inside the source-plane stability circle, and terminating the amplifier output with a value outside the load-plane stability circle, the amplifier will be unconditionally stable at the frequency at which the match is obtained.

The next task is to obtain the values for the network circuit that matches the input for the 1.1-dB noise figure. A noise matching will not result in the best possible S_{11} matching (Figure 3-69). This translates into an inductor of 9.2 nH to ground and a series inductor of 0.85 nH based on the large-signal model of the NE68133 at the chosen bias point.

After we have determined our input matching network, it is time to find the output matching (Figure 3-70). First, find the termination point on the output plane that corresponds to 14.3-dB gain and 1.1-dB NF at the input plane. The output gain circle goes through the

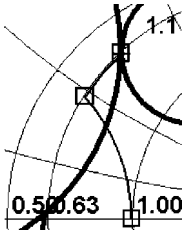


Figure 3-69 Input matching-network extraction, assuming lossless matching components and no parasitic effects.

1.1-dB NF circle at two points, and for reasons of easy matching we selected the lower point. The next step is to find the conjugate reflection coefficient $\Gamma_L = S_{22}^*$. This is the lower square marker in Figure 3-70. The remaining task is to find the matching network that brings this value to 50 Ω. By doing so, we first obtain a shunt inductance $L_{\text{shunt}} = 8.2$ nH and a series capacitor $C_{\text{series}} = 2.4$ pF. Again, these are the values applicable to the large-signal equivalent circuit.

The careful reader will have noticed that we have been avoiding the issue of small-signal exact matching using the manufacturer-supplied parameters. If this is done, the narrow band response shown in Figure 3-71 will be obtained. Therefore, we must optimize the circuit for a better frequency response to obtain the one shown in Figure 3-72.

After this optimization, we were well able to meet the specification of a center band gain of 14 dB and a corresponding NF of 1.1 dB. To show the flatness of the gain, Figure 3-73 shows a gain variation in the range of 0.25 dB.

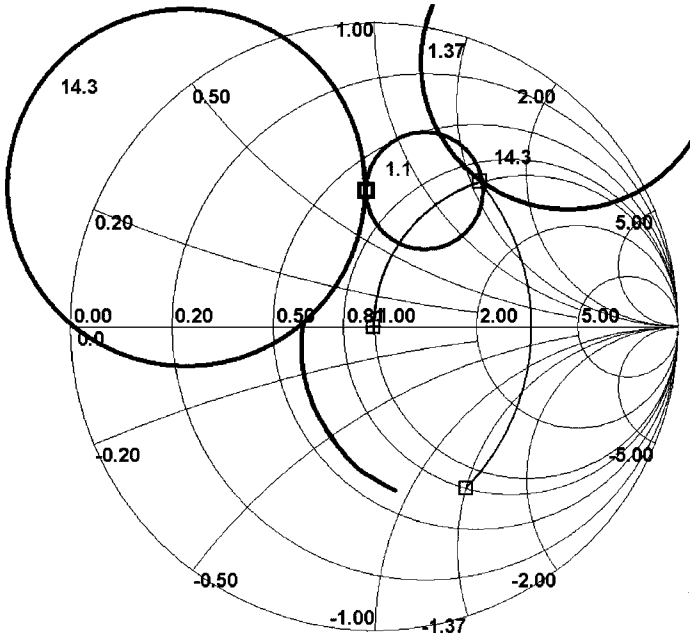


Figure 3-70 Output matching-network extraction.

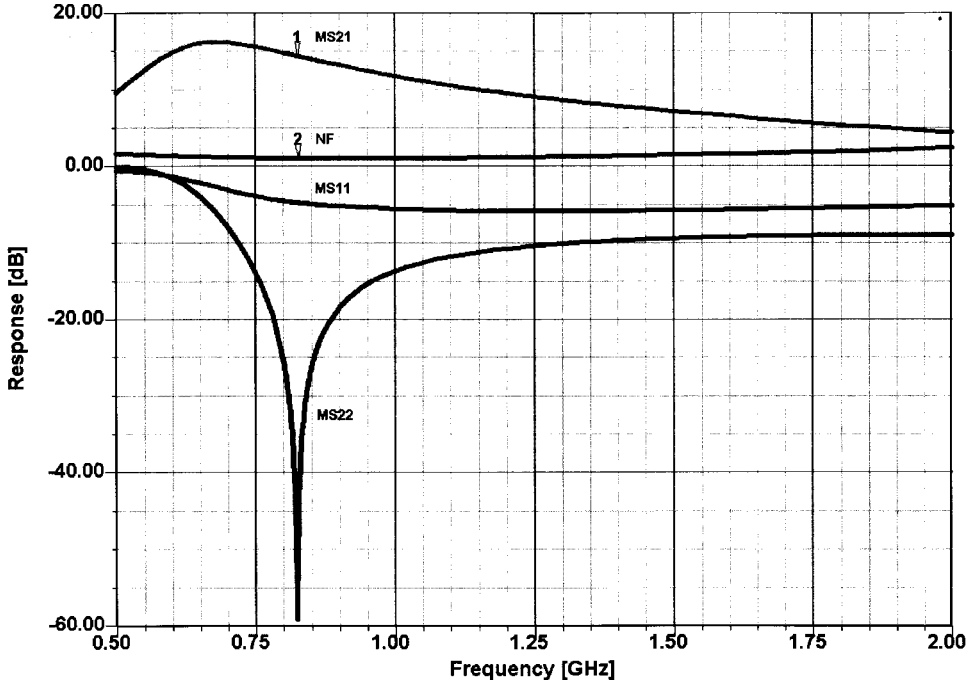


Figure 3-71 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-67.

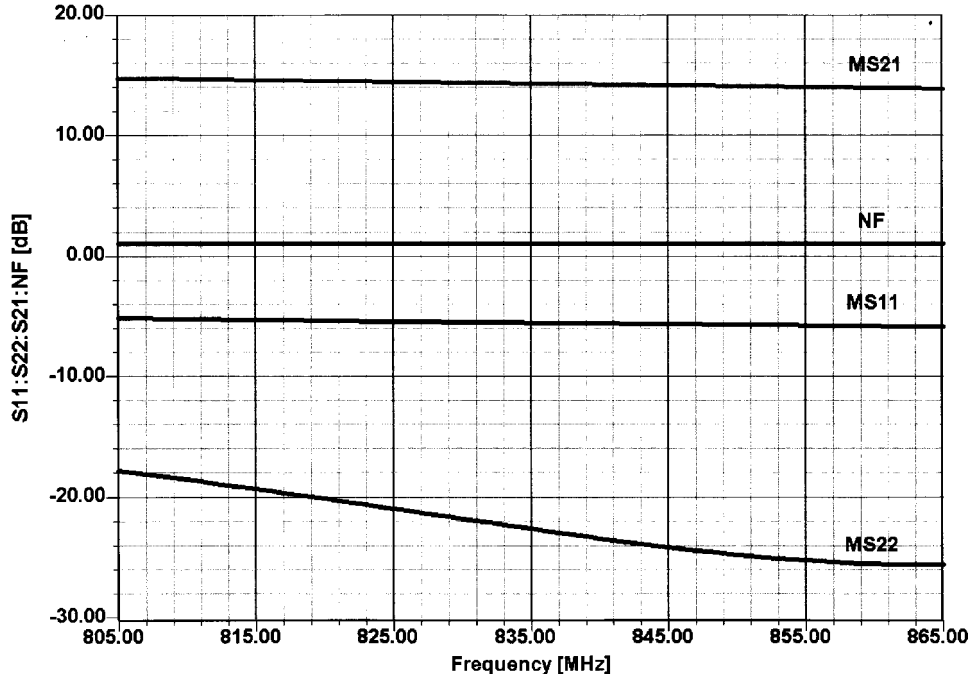


Figure 3-72 Optimized performance of input/output reflection, gain, and noise figure.

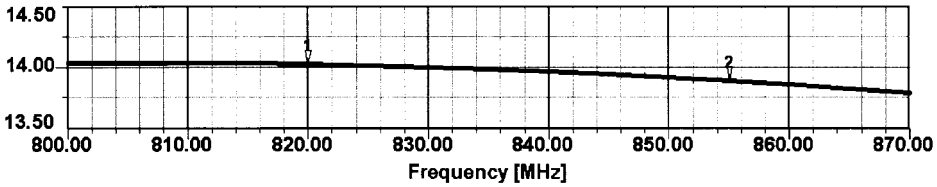


Figure 3-73 High-resolution display of gain as a function of frequency.

The next step is to look into the large-signal performance. The first order of business is to determine the third-order intercept point, which is based on a two-tone analysis. Given the fact that we run the device at low voltage and at a just-reasonable current, the actual intercept point of +13.5 dBm is quite a good number. Figure 3-74 shows the fundamental and third-order outputs. The crossover point between the two, as shown in Chapter 1, is defined as the third-order intercept point. If we had chosen the third harmonic output rather than the fundamental, we would have shown the fifth-order intercept point. The second-order IMD product ($f_1 \pm f_2$) typically can be minimized by input selectivity.

How an amplifier's linearity affects its ability to preserve the characteristics of digital signals is a major concern in wireless design. Figures 3-75 and 3-76 show constellation diagrams with the amplifier handling a PSK signal at drive levels of –40 dBm and +10 dBm, respectively. As expected, we see significant distortion at the higher drive levels due to circuit nonlinearities.

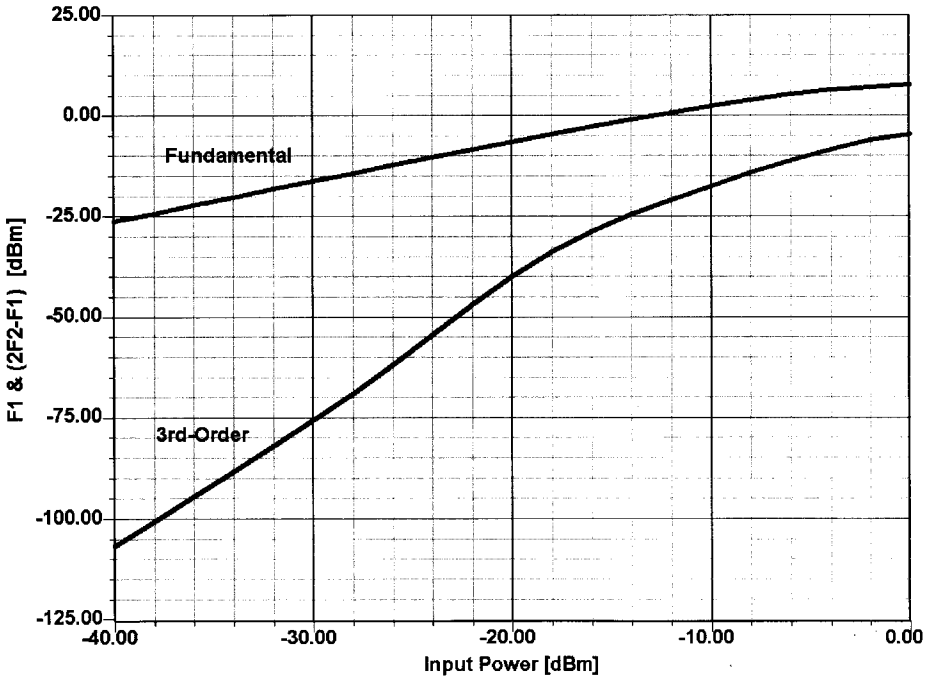


Figure 3-74 Simulated fundamental and third-order IMD outputs of the NE68133 amplifier.

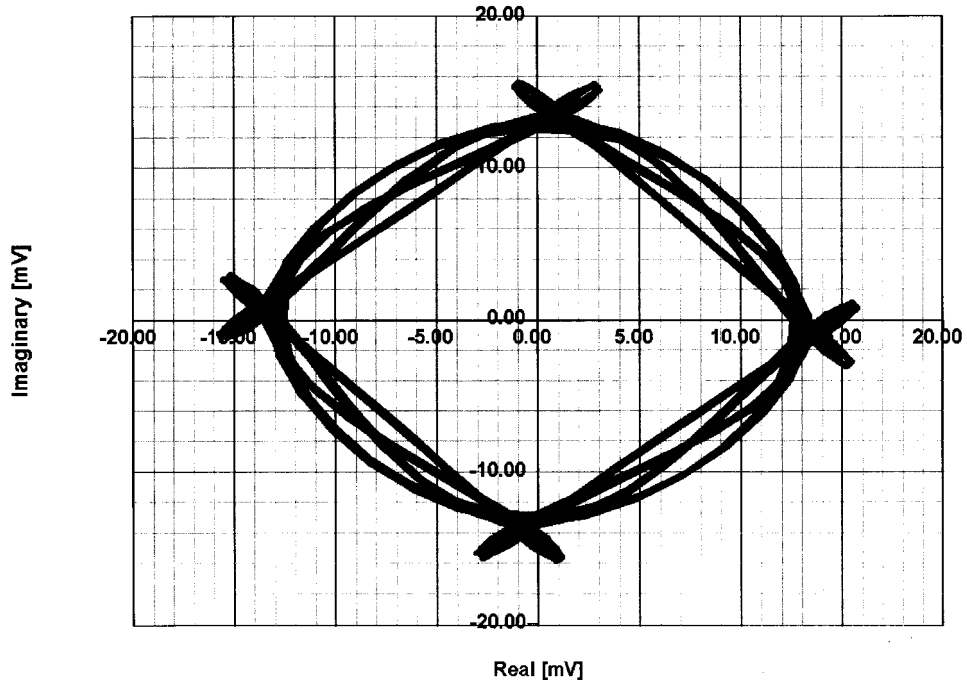


Figure 3-75 Output constellation at –40-dBm input.

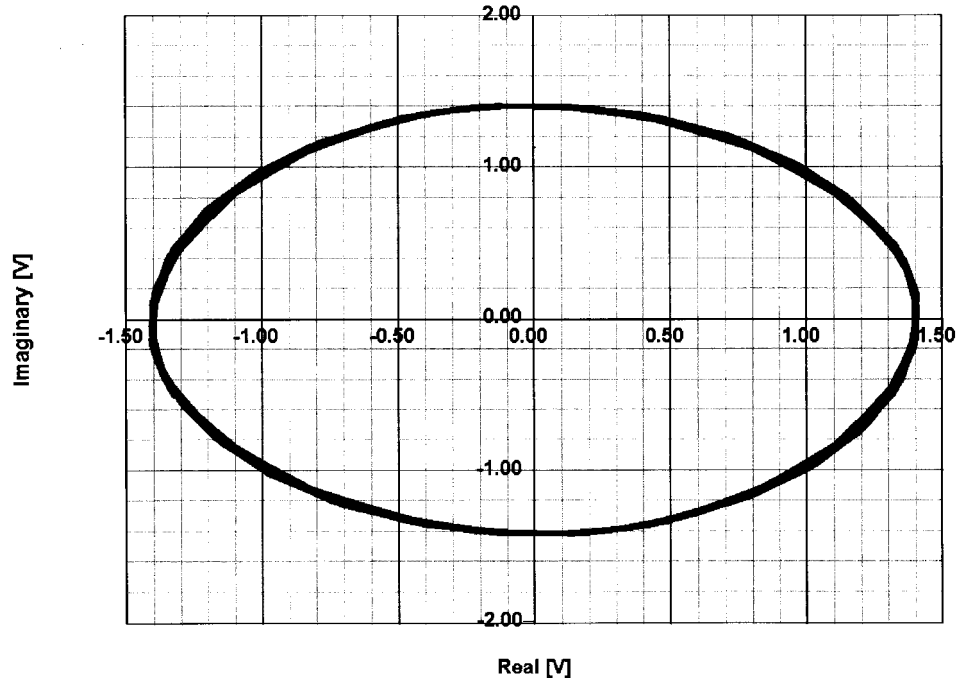


Figure 3-76 Output constellation at 10-dBm input.

BFP420 Matched Amplifier. The NE68133 has an attractive f_T of 8 GHz, but if more gain is needed, one may play with the concept of using a much "hotter" device, such as the BFP420. Recall that the NEC device was unconditionally stable, and our first worry is how the Siemens BFP420 will behave.

A study of Figure 3-77 shows that indeed the NEC device had to be stable all over and therefore was a good basis for a risk-free design. Things change dramatically with the BFP420, for which the K factor is essentially less than 1 up to a corner frequency of 2.75 GHz, where the transistor becomes unconditionally stable. Around 1.75 GHz, the K factor for the NEC transistor barely touches 1. We explained earlier that $G_{\max}$ deviating from the maximum stable gain is a more realistic number, since the conditions for maximum stable gain involve feedback and conjugate matching.

To improve the stability of the BFP420, we will now set out to use the same topology as the previous amplifier but with a resistive feedback between base and collector (Figure 3-78). A good value for this is always somewhere around 200–500 Ω. It must be remembered that this voltage feedback also stabilizes the input and output impedance and reduces the Miller effect because of the feedback being resistive and more dominant.

While the first example covered a fairly narrow frequency band, we are now attempting to cover a much wider frequency range; specifically, from 1.5 to 3 GHz reasonable gain and noise. This, by definition, requires a much higher gain-bandwidth product, and the simple matching network at the input and output provides a nonoptimal input and output matching (Figure 3-79). The feedback introduces some additional noise. This application shows the limits of single-stage UHF/SHF amplifiers with feedback.

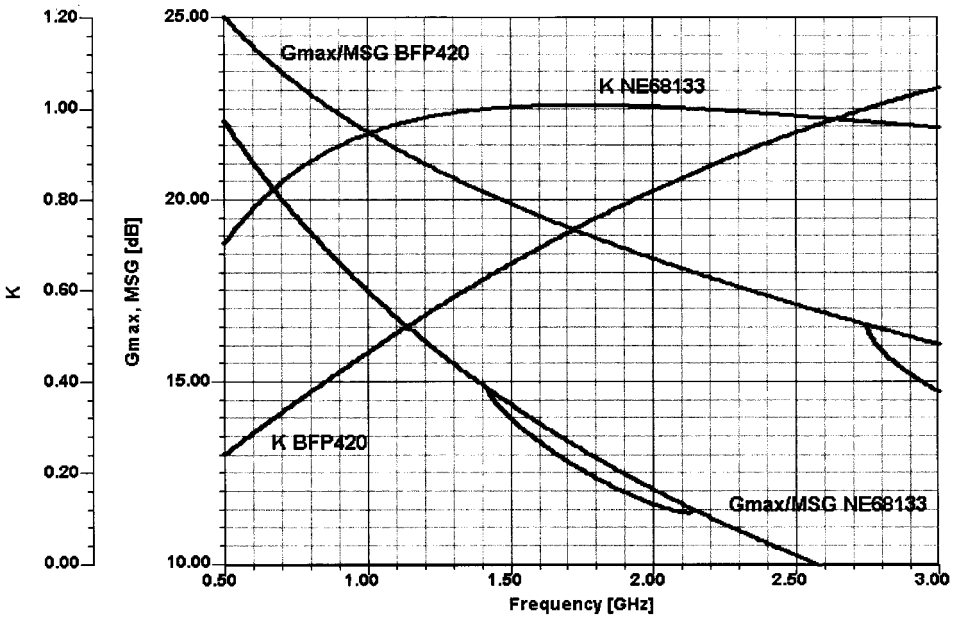


Figure 3-77 Comparison of K, $G_{\max}$, and maximum stable gain (MSG) for the NE68133 and BPF420 BJTs.

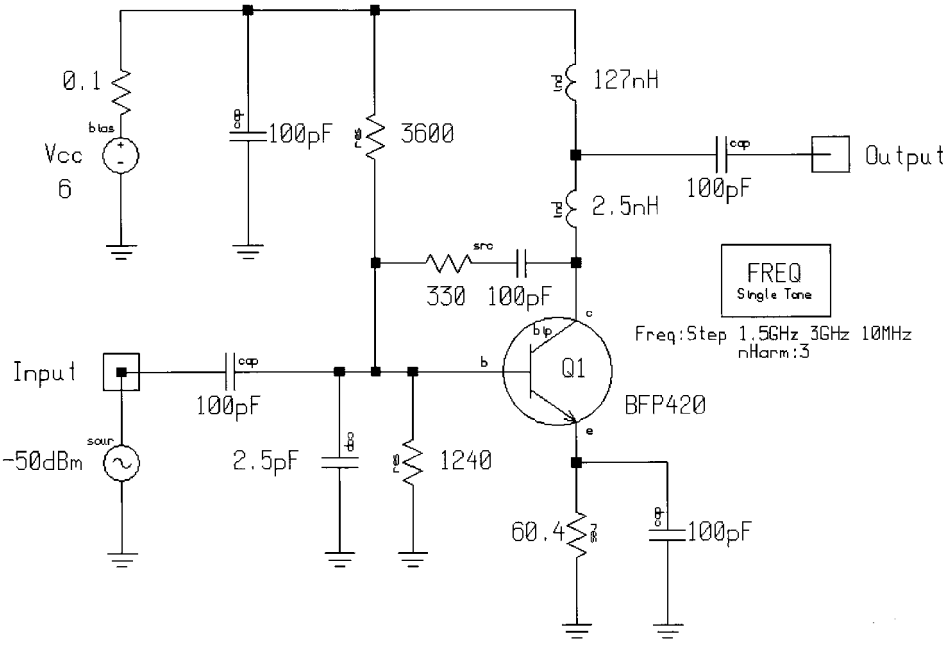


Figure 3-78 BFP420 amplifier with resistive voltage feedback from collector to base.

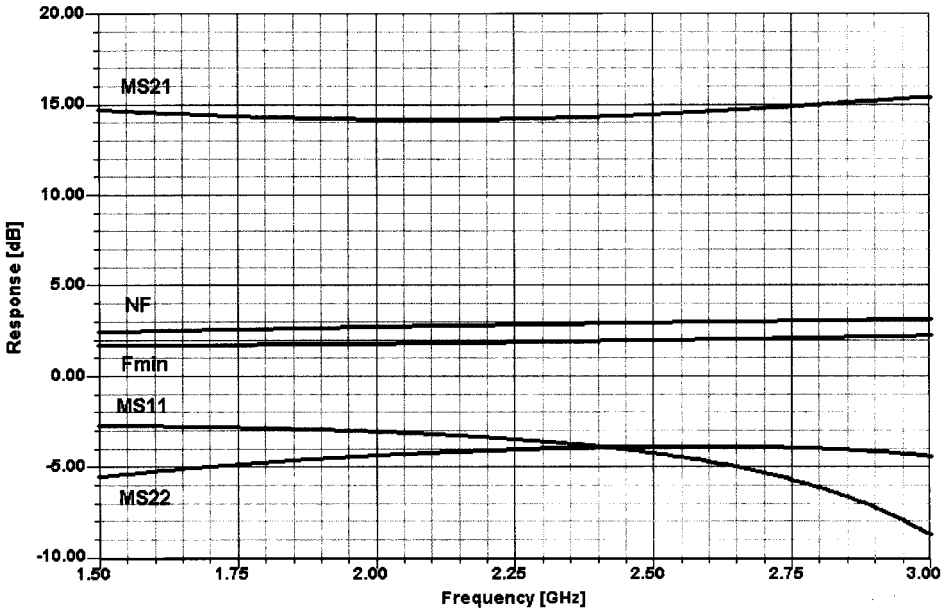


Figure 3-79 Frequency-dependent gain, matching, and noise performance of the BFP420 feedback amplifier.

Narrowband BFP420 Amplifier. Our next attempt will be to design a narrowband input stage that combines very high selectivity with good noise figure. This can only be achieved by using two tuned single-stage resonators with capacitive/magnetic coupling. The following amplifier is tailored to have a center frequency of 900 MHz. The transistor will operate at 12 V/10 mA to have a sufficient dynamic range. The input selectivity is achieved by the use of two top-coupled tuned circuits, followed by a BFP420 (Figure 3-80). The first order of business is to determine the input tuned circuit. The tuning capacitance for the two tuned circuits is arbitrarily set to a manageable 2 pF (small values have too much tolerances and larger values will require smaller inductances, resulting in lower Q). The reactance of the first tuned circuit is

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi \times 9\text{E}8 \times 2\text{E}{-}12} = 88.42\ \Omega$$

The necessary inductance for resonance will be

$$\omega L = 88.42$$

or

$$L = 88.42/\omega = 88.42/(2\pi \times 9\text{E}8) = 15.64\ \text{nH}$$

We will make both tuned circuits identical, but with different loaded Qs. By multiplying 88.42 times the loaded Q (20 for the input resonator and 50 for the output resonator), the resulting parallel-resonant resistance is $88.42 \times 20 = 1768\ \Omega$ at 900 MHz (for the input resonator) and $88.42 \times 50 = 4421\ \Omega$ (for the output resonator). These values are achieved by loading the input with the 50-Ω source and by loading the output with the input of the transistor. Therefore, the input transformation ratio $m_1 = 1768.4/50 = 35.37$. At the output, the equivalent number $m_2 = 4421/50 = 88.42$. For the purpose of simulation, and because it is extremely difficult to find the physical location of the appropriate tap on an air-wound coil for the necessary impedance, we do this "on paper" by incorporating an ideal transformer. We are also going to show later the actual transmission-line-based filter and its modeling. This filter will have a higher insertion loss based on the PC-board material used.

Using our nonlinear simulator, in small-signal ac mode, we obtain the frequency response and noise figure shown in Figure 3-81. In actually building this circuit, there is another problem besides these transformers—the top-coupling capacitor. We have used a grounded-T configuration to obtain a practically realizable value of 36 pF, whose lead inductance will not influence the circuit. (The actual coupling capacitance value for top coupling without the T configuration would have 0.1 pF. We would like to see this value repeatably realized.)

If the actual noise would be calculated relative to the output tuned circuit, the equivalent noise resistor R_n of 5 Ω would now be multiplied by m_2, resulting in 442 Ω. Fortunately, the Serenade simulator has a highly accurate BJT noise model that is both bias and temperature dependent, so the simulator will give us the answer directly.

We have not put any matching circuit at the output and leave this task to the reader. In practice, the slight tilt in the MS_{21} curve in Figure 3-81 would be tuned out. Also note that the minimum noise figure ($F_{\min}$) and the actual 50-Ω noise figure (NF) agree somewhat off center. This has to do with the Miller effect, which results in the well-known difference between noise matching and input-gain matching. By varying the coupling capacitor to move from the overcoupled response, which gives a low noise figure, to a critically coupled

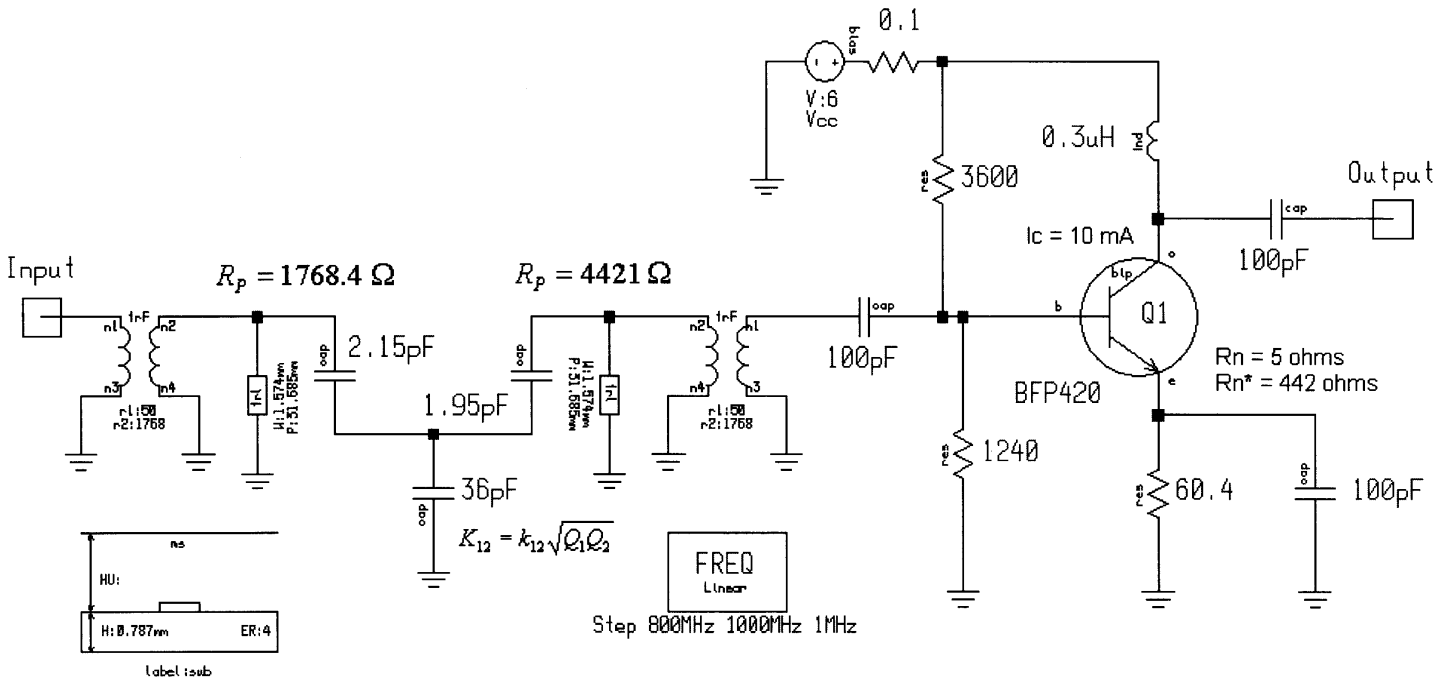


Figure 3-80 BFP420 amplifier with narrowband input filtering. Ideal transformers are used to match the filter's input and output to the input port and the BFP420 base, respectively.

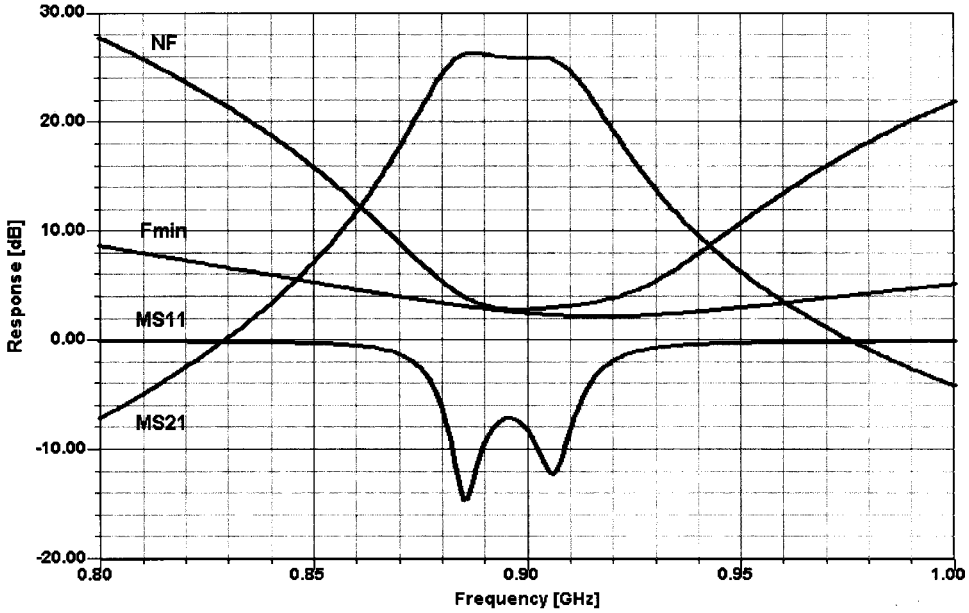


Figure 3-81 Frequency-dependent gain, matching, and noise performance of the narrowband BFP420 amplifier.

response ($K = 1$), we get significantly better selectivity and narrow bandwidth (at the obvious expense of the noise figure) as shown in Figure 3-82.

Since we somewhat compromised by not providing the reader with an actual input and output coupling that can be built, and furthermore, because the inductance in question (15.26 nH) is hard to realize with a high Q, we now show the printed circuit approach for the input filter. By using a Teflon-based material for the PC board, the losses would be less; however, the radiation would be more because ε would go down to 2.1 from the value we used (4). It is also convenient to introduce this filter to show that at higher frequencies, where lumped elements are really no longer available, one has to switch to distributed elements. A list of the commonly used distributed elements will be shown later; the transmission line and T junction are probably the most frequently used elements, followed by spiral and rectangular inductors, and bends.

The actual implementation of the input filter on a printed circuit board requires distributed elements (Figure 3-83). In practice, one has two parallel coupled lines with a tap somewhere toward the ground connection and, for modeling accuracy, we have to introduce a T junction with the appropriate impedance of 50 Ω for the input and a different impedance (if necessary) for the output. The coupling mechanism now is magnetic coupling, since no one can physically change the distance between the transmission lines by tuning some mechanical elements, one has to get it right the first time. This can only be achieved with a high-precision simulator. The reason why these simulators are so expensive is the investment of research time and personnel to produce valid models over a huge frequency and impedance range. Since we are now curious about the frequency response, we have plotted in Figure 3-84 the frequency response, which shows a higher (expected) insertion loss. The PC-board material is partially responsible for this.

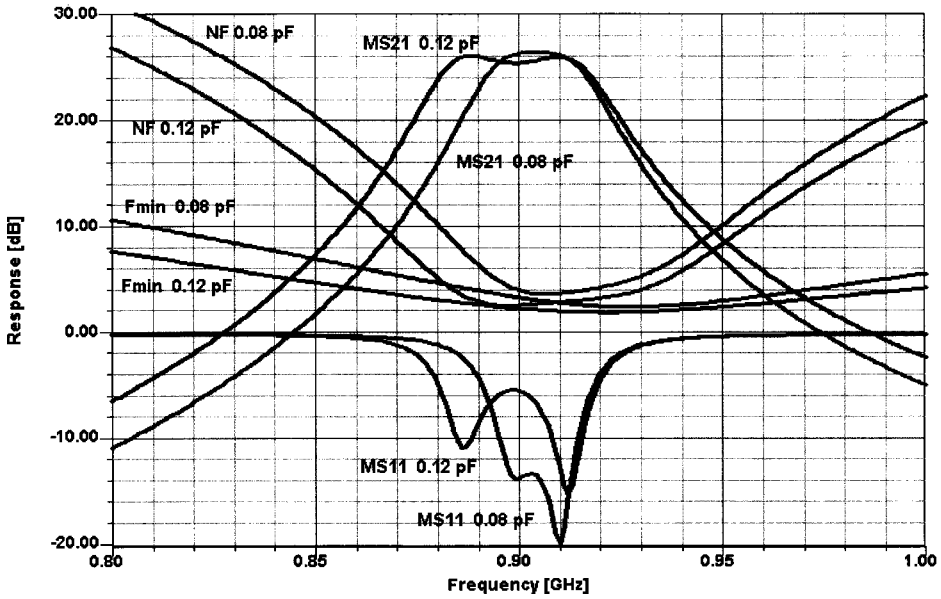


Figure 3-82 Comparison of filter responses with critical coupling (0.08-pF response) and overcoupling (0.12-pF response).

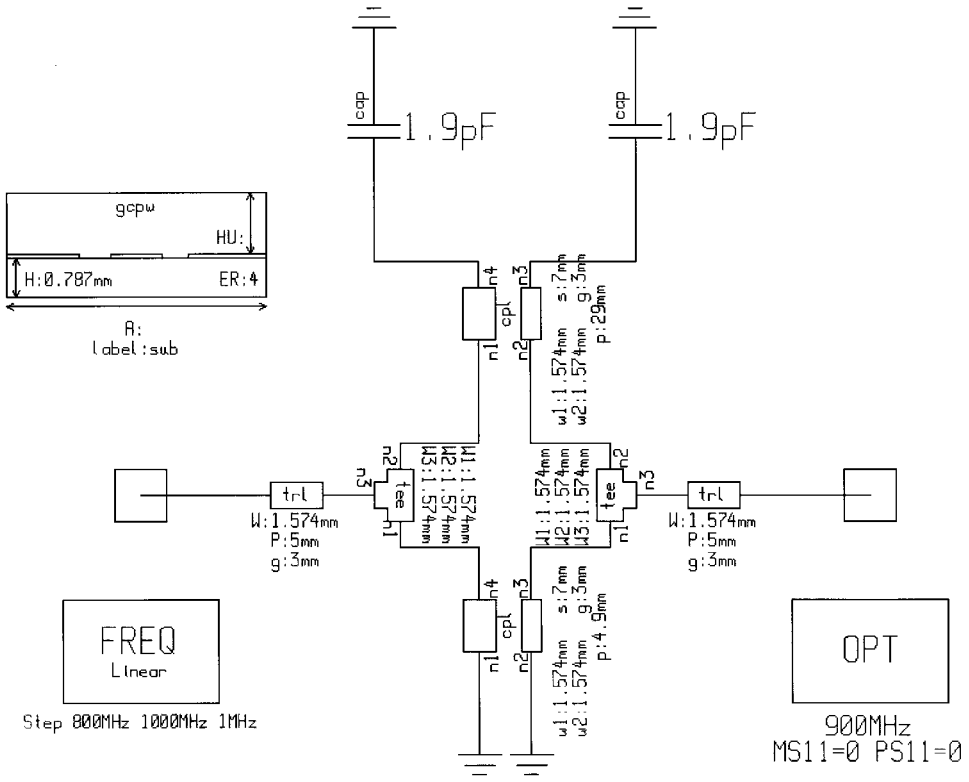


Figure 3-83 Input filter using distributed-element resonators and matching.

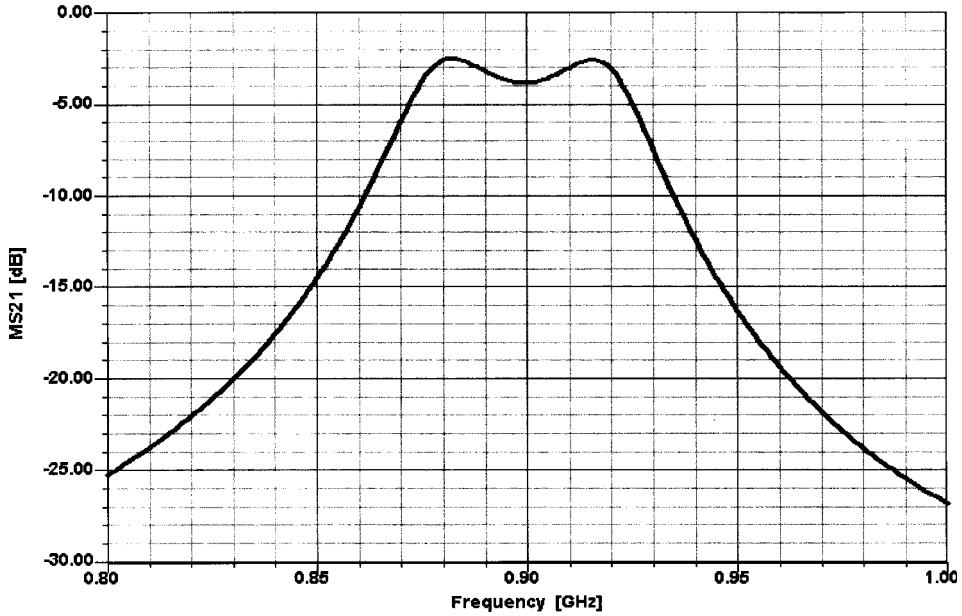


Figure 3-84 Gain versus frequency response for the distributed-element filter.

At the end of this chapter we will have a chance to discuss passive elements as needed for microwave frequencies, and one should consider frequencies above 1500 MHz as microwave frequencies.

GaAsFET Feedback Amplifier. The implementations we have shown so far have been based on bipolar transistors. Figure 3-85 shows a single-stage feedback amplifier using a GaAsFET. The feedback is based around an *LRC* network and some attempts at matching at the input and output are made. Figure 3-86 shows its frequency-dependent gain, matching, and noise performance.

3-2-3 High-Gain Amplifiers

In most cases, high-gain amplifiers receive an already amplified signal, which means we are less concerned about the signal-to-noise ratio but remain concerned about distortion. To reduce distortion, we have two options: (1) increase the dc current in the active device, which always improves the current-related distortion; and (2) add some voltage feedback, which reduces the voltage distortion. In simple terms again, only current feedback can reduce the current-related distortion of a square-law or exponential transfer characteristic and only voltage feedback can reduce the effect of overdrive as a function of the load, but only to some degree. Its main purpose remains to reduce the input impedance to useful values, specifically for FETs, since current feedback increases the input and output impedances. Some ultrasensitive circuits still require Class A operation, but now at higher current levels; these are typically CATV applications, and there are several transistors available that can be operated at 5–15 V but at 60 mA or higher. Because of the higher power consumption, this excludes handheld cordless or cellular telephones. The same circuit as used before (Figure 3-85) for

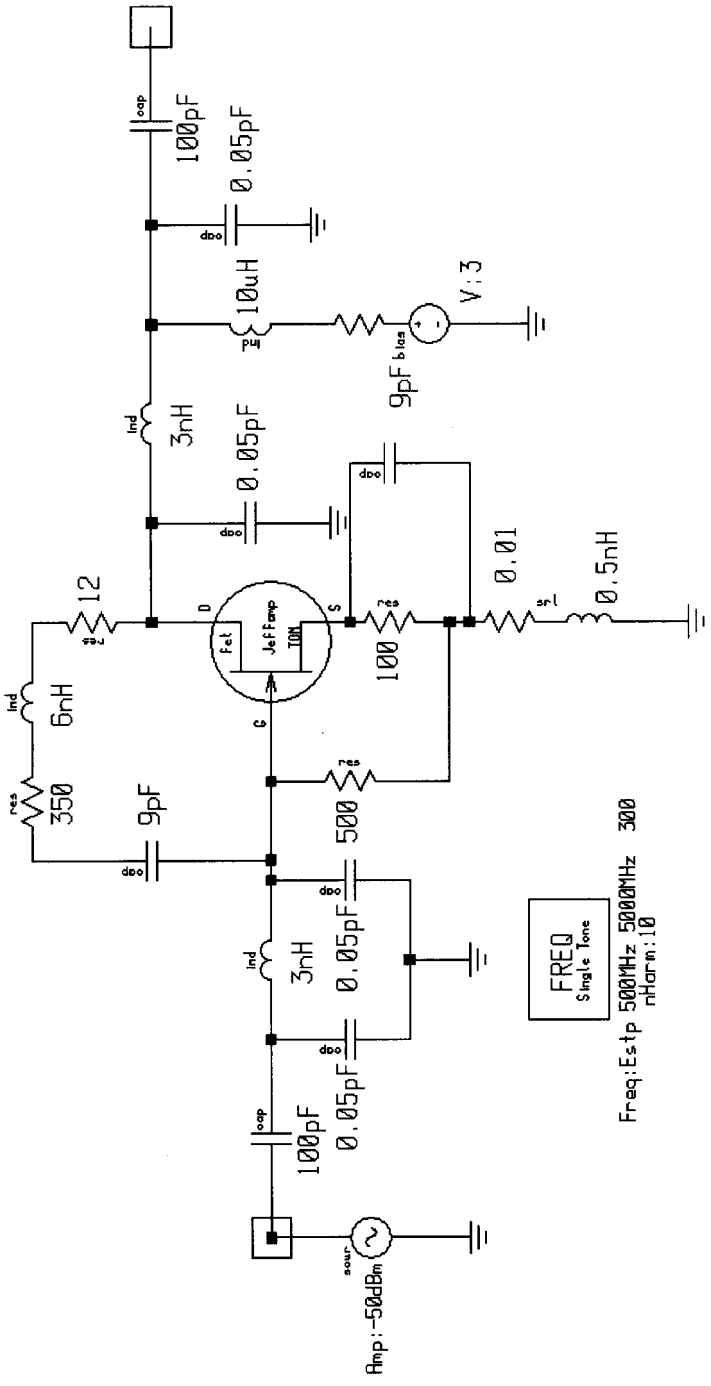


Figure 3-85 GaAsFET feedback amplifier.

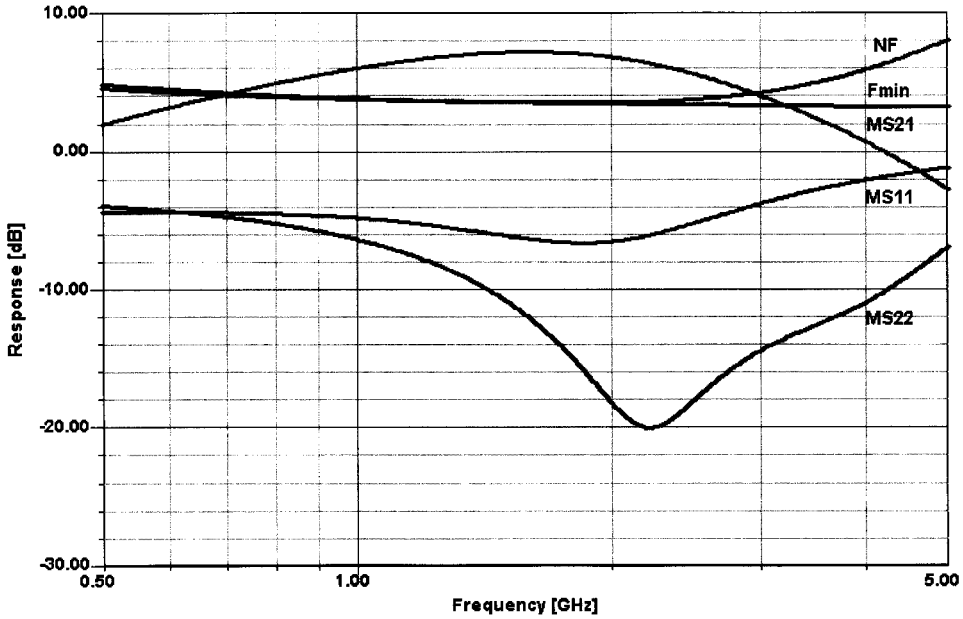


Figure 3-86 Frequency-dependent gain, matching, and noise responses for the GaAsFET amplifier.

Class A operation now can be operated at a higher current, which automatically increases the gain and the noise figure and reduces distortion. This is based on the fact that the transconductance (g_m) increases at higher currents. The second transistor in the TI circuit achieves this by being in reality a larger device, which could be modeled by several transistors, like the first stage, in parallel. This automatically increases the dc current by the scaling factor. The same is even more true for the output stage. Likewise, bipolar transistors can also be scaled. A visual inspection of some transistors in their packages will reveal that they are really not just one transistor, but several operating in parallel, each with its own ballast resistors (for bipolar transistors) or other means (for FETs).

Using tuned circuits rather than wideband matching, it is possible to go to Class AB or B operation. Class AB is really a hybrid between A and B, and Class B is probably better known. Figure 3-63 shows the Class B operating point. Most Class B stages are operated in push–pull because, together with the tuned circuit, one-half of the cycle is supplied by the other transistor. The definition of Class B operation is a conducting angle of 90° (see Figure 3-63). At this point it should be noted that some older texts define the conducting angle as being twice the value we have defined in this chapter.

Without going into further details, the highest efficiency in Class B operation is 78.5%. Based on the fact that modern wireless applications generally use a hybrid of amplitude and angle modulation, pure Class C operation finds little application today because of the severe nonlinear distortion it introduces. A Class B stage has less "power gain" than a Class A stage but provides much more output power. We have already mentioned the problem associated with characterizing a Class A stage in terms of power gain. That's the only time Kirchhoff's equations are valid, meaning that the sums of all currents and voltages have to be zero. Because Kirchhoff's equations do not consider nonlinearities and harmonic contributions,

the very moment we move into considering medium- and higher-power amplifiers, the nonlinearities will dominate and the performance of the stage is best analyzed with a high-quality circuit simulator with the appropriate dynamic range and good modeling capability. Without question, this applies to Class C operation also. As pointed out previously, the high-gain amplifier is optimized for gain rather than noise figure and typically requires more collector or drain current. This can easily be seen by rebiasing the FET amplifier as shown above to a significantly higher value of 33 mA. This is accomplished by changing the source resistor down to 3 Ω. As a result of this, we obtain the gain, matching, and noise figure responses shown in Figure 3-87.

It may be a surprise to the reader that the actual noise figure now is less than that shown in Figure 3-86. This device is now operating at 33 mA; the reason for this has to do with the fact that if the transconductance is too low, the noise figure also increases, and the NF curve for the FET has a similar-looking current-dependent minimum—a curve that is frequently not supplied by the manufacturer. This amplifier, now operating at 33 mA, should have quite good large-signal capabilities; we will examine this shortly.

Another interesting piece of information is the display of the dc I–V curves with the "load line." Since we have reactances in the circuit, we do not see a line but rather an ellipse, indicating the storage of energy. (The actual load line would be a line through the two foci of the ellipse.) This curve allows us to see immediately that at this operating point the transistor is neither going into current saturation nor becoming voltage limiting. The input power can be increased further until such a condition exists. Figure 3-88 shows the dc I–V curve for the amplifier shown in Figure 3-85.

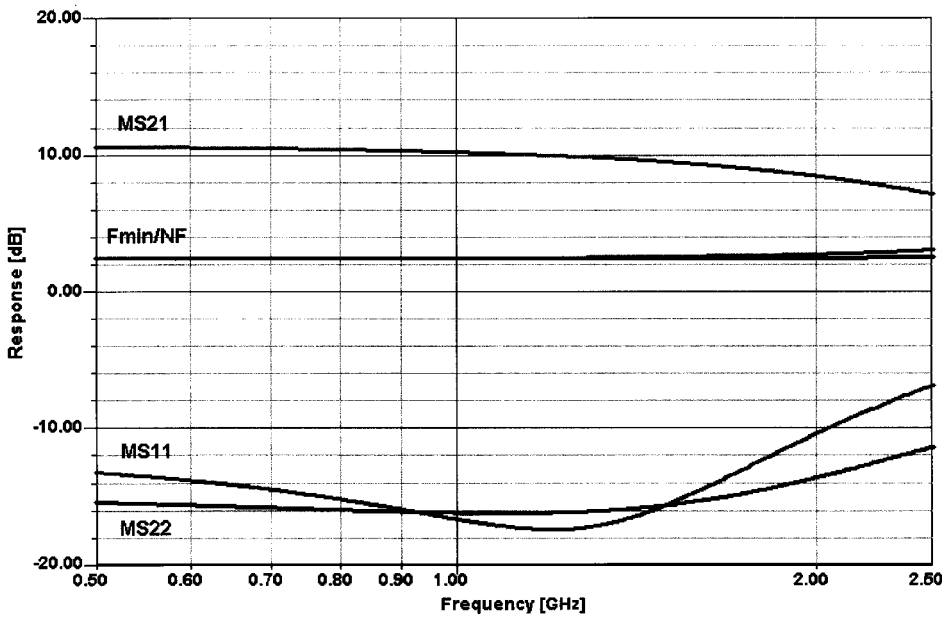


Figure 3-87 Frequency-dependent gain, matching, and noise performance for the GaAsFET amplifier of Figure 3-85 rebiased for a drain current of 33 mA.

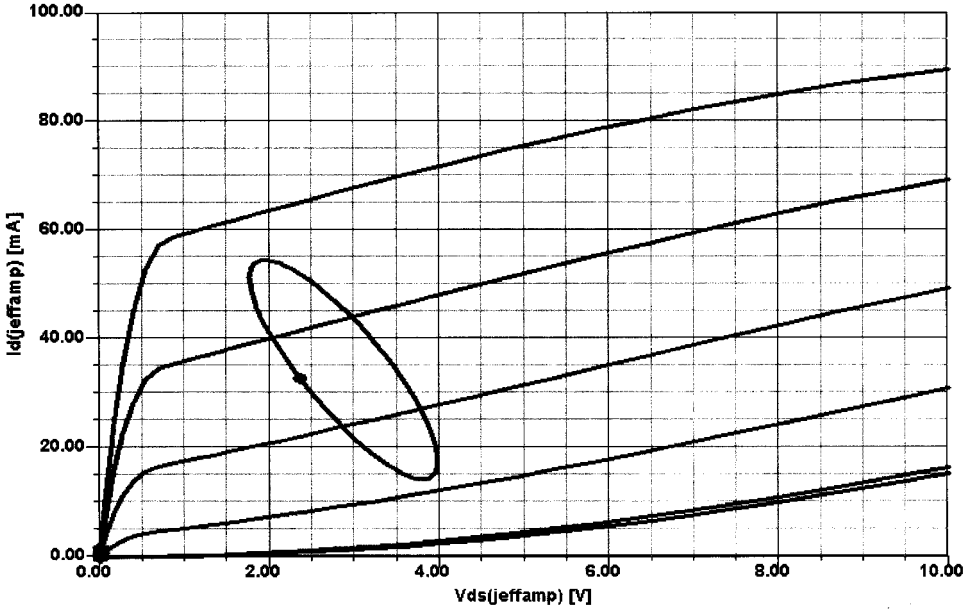


Figure 3-88 The dc *I–V* curves for the GaAsFET operating at I_D = 33 mA, including the ac load line. The load line is an ellipse rather than a straight line for the reason described in the text.

Besides looking for the adjacent-channel power ratio (ACPR), the other traditionally important points are third-order intercept point and the equivalent of a multitone simulation, as used for CATV application but relevant to the multisignal environment of wireless applications. Figure 3-89 shows the determination of the amplifier's third-order intercept point. It is most important for the reader to understand that very low values for input and output level have to be used to determine the crossover point. If this advice is not followed, totally erroneous numbers can occur.

Since an antenna supplies a large number of signals (this is true of handheld radios as well as base-station sites), a three-tone standard has been adopted as a good approximation of multisignal operation. This standard is in accordance with the German DIN 45004E,3.3 three-tone measurement. The requirement is that the intermodulation distortion products at the output are better than 60 dB suppressed relative to the largest of the three tones. One tone is used as the reference, and the other two tones are 6 dB less. This measurement has been derived from television, whereby the video transmitter, based on the vestigial sideband power, is 6 dB above the two sidebands produced by the FM sound transmitter using a modulation index of 5. The result of a three-tone test can be seen in Figure 3-90. This amplifier clearly meets the DIN specification for linearity at input levels up to –9/–15 dBm at the input.

Bipolar transistors are also quite popular for this purpose. Figure 3-91 shows a single-stage amplifier with an intrinsic transistor and all the other "hidden" intrinsics, normally not visible, added. The three diodes belong to those intrinsic elements, which are needed for accurate microwave modeling but which are not provided by the standard nonlinear models. The reader will see that we went into great detail to model all the necessary parts. This feedback amplifier, whose frequency response is shown in Figure 3-92, has been quite

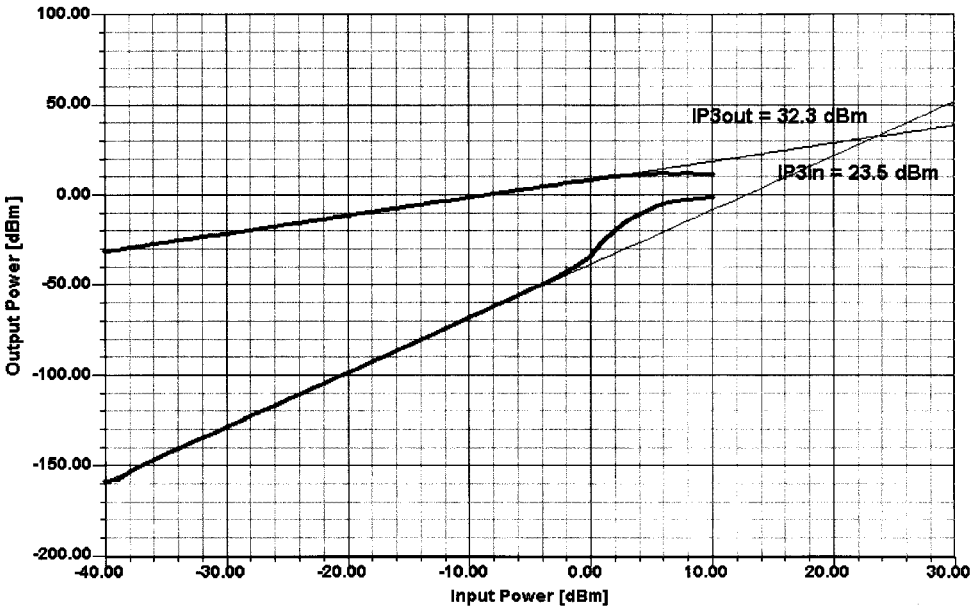


Figure 3-89 Determination of the third-order intercept point based on a two-tone measurement using tones spaced at 5% of the operating frequency (1.8 GHz).

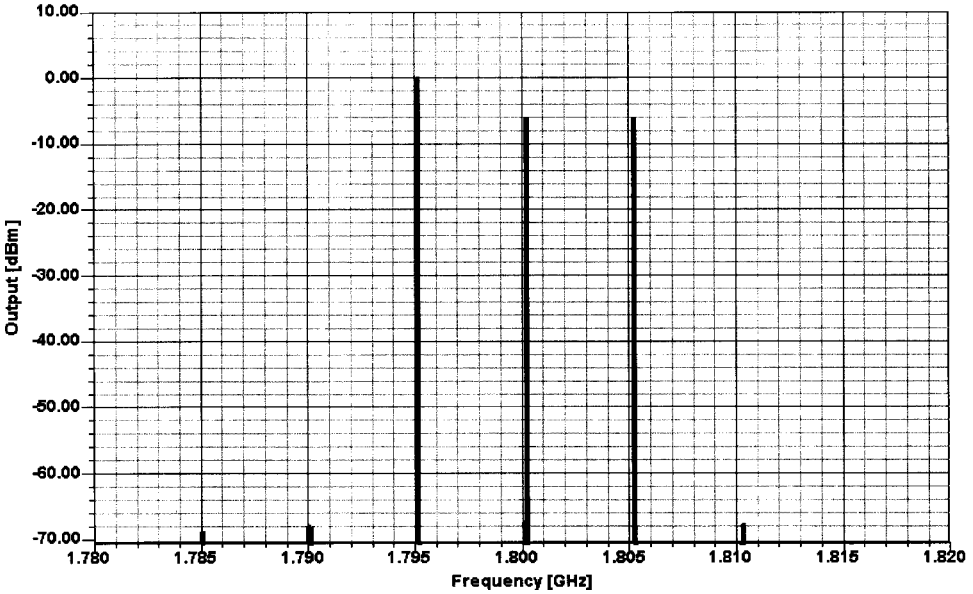


Figure 3-90 Simulated three-tone analysis of the GaAsFET amplifier with $I_D = 33$ mA.

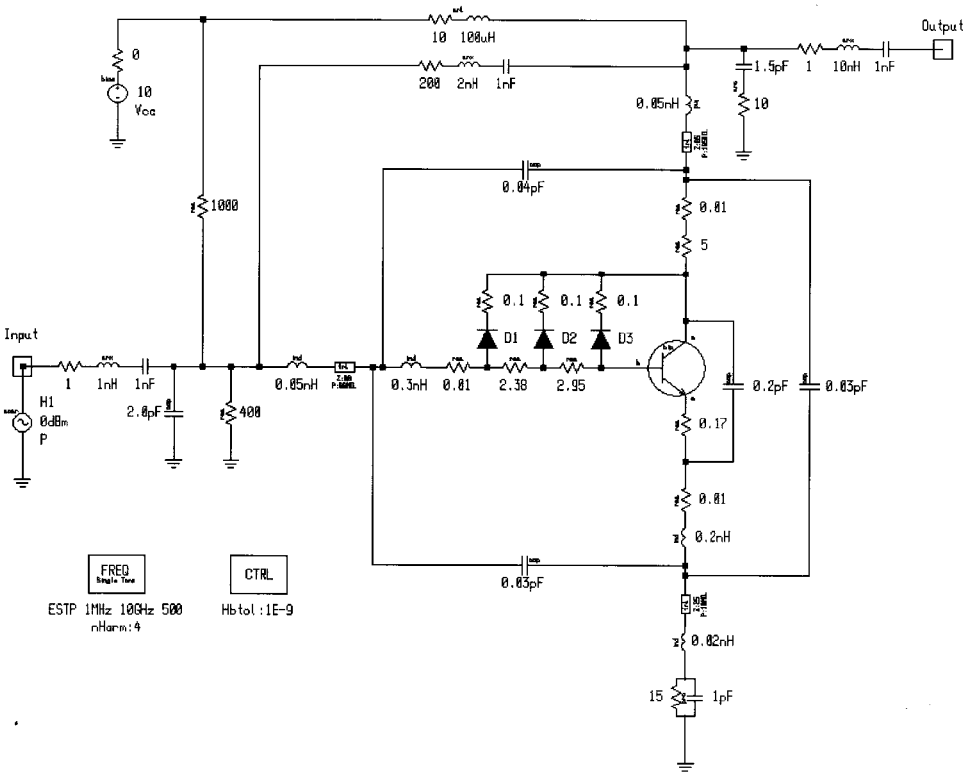


Figure 3-91 A high-linearity single-transistor amplifier modeled in great detail.

difficult to design, but there are monolithic circuits available where the higher gain–bandwidth product is being used to advantage.

Figure 3-92 shows the predicted frequency-dependent gain, matching, and noise figure at a bias point of 108 mA at 10 V. Using the same DIN three-tone test procedure, Figure 3-93 shows the expected good result.

Since this test requires two equal signals and one larger signal, we need to expect one IMD product larger than the rest. We have set our window for a dynamic range of 70 dB, and yes, in the lower left corner, we can see the product between the first two signals, while the other products are below the –50 dB level.

As a little task, Figure 3-94 shows a simple amplifier, based on a modern Siemens transistor, which is a feedback amplifier with fewer modeling details surrounding the transistor. Its associated frequency-dependent performance is shown in Figure 3-95. Interested readers are invited to improve its performance by using their skills and intuition.

Permitting higher currents and actually combining two transistors leads to our first MMIC. The following simulation is based on the MGA64135. Needless to say, there are some compensation networks inside the MMIC and distributed elements that affect the frequency response. However, with this arrangement, it is possible to obtain an appreciable gain at our wireless frequencies. The noise figure, based on the feedback, however, it not too impressive. This is a problem that occurs with most feedback amplifiers, which try to minimize the distortion and yet would like to keep the noise figure down. Figure 3-96 shows the schematic

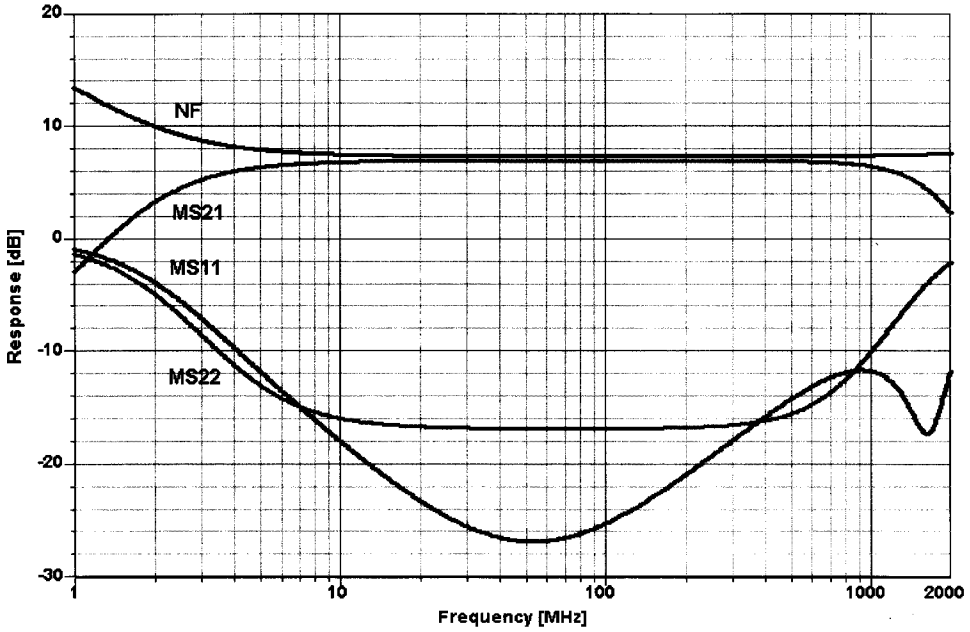


Figure 3-92 Simulated frequency-dependent gain, matching, and noise figure performance of the circuit in Figure 3-91.

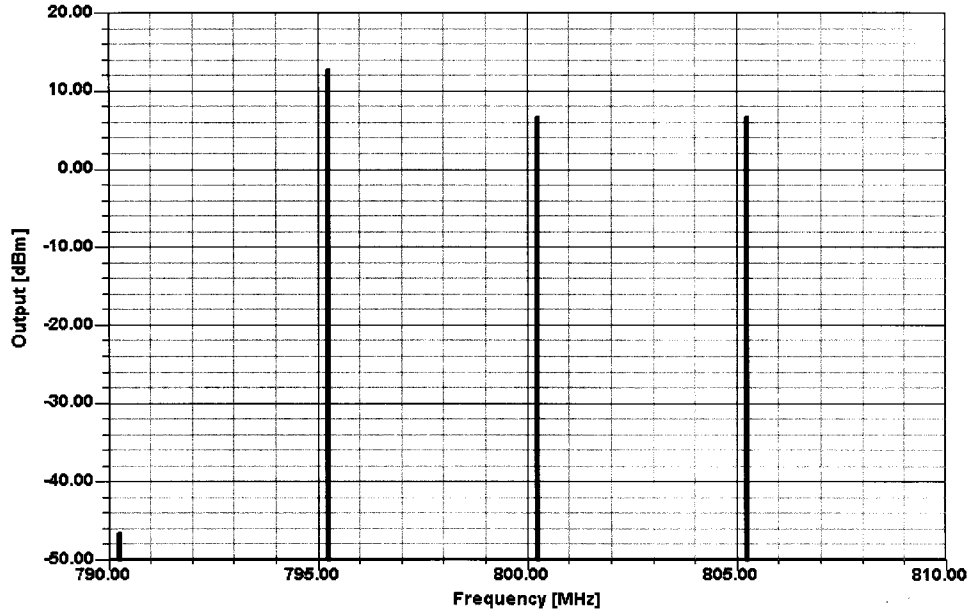


Figure 3-93 Simulated three-tone output spectrum of the amplifier in Figure 3-91.

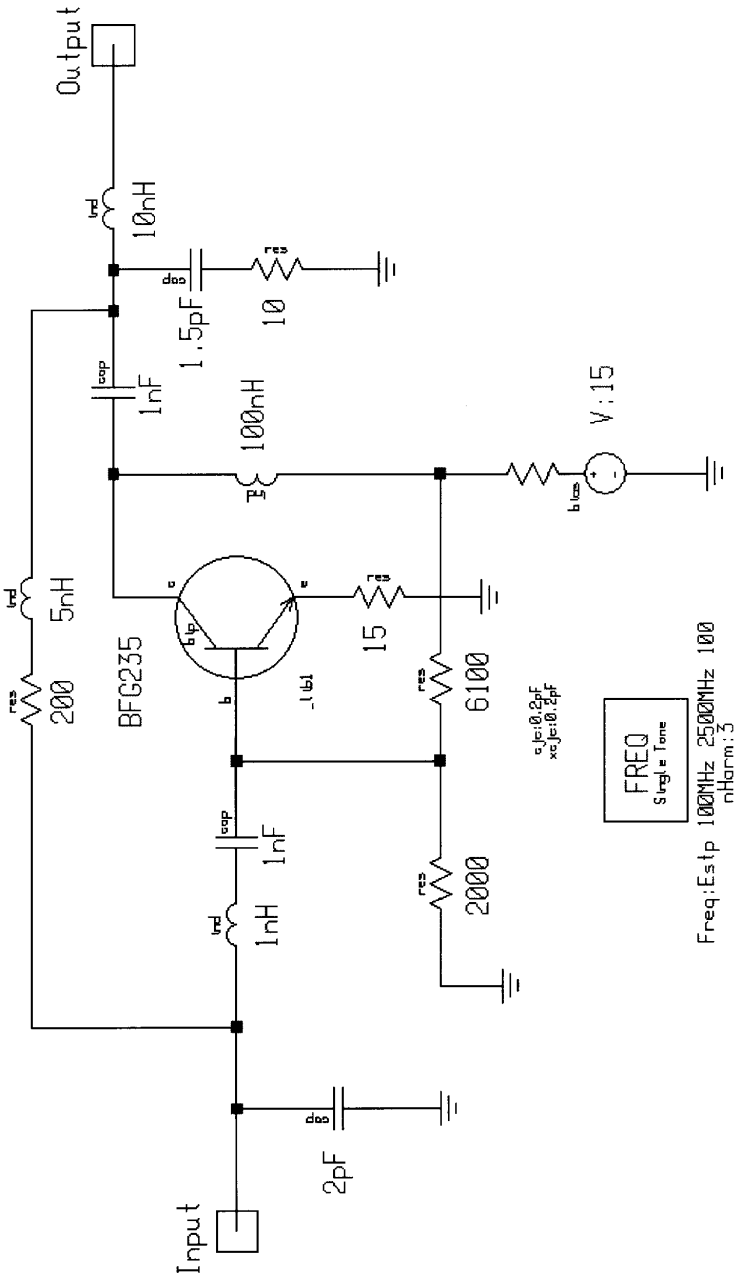


Figure 3-94 A high-linearity amplifier using the Siemens BFG235.

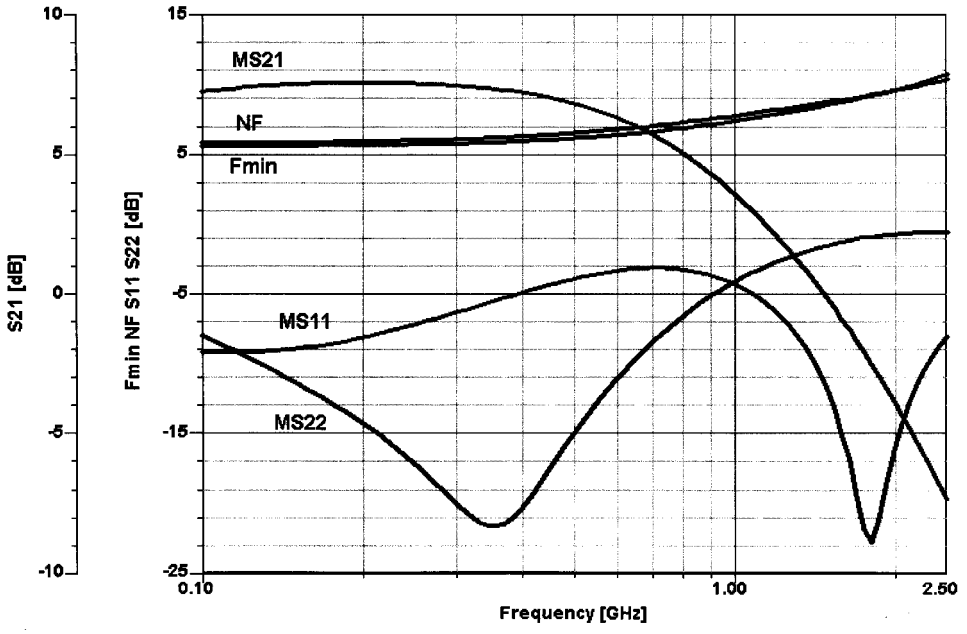


Figure 3-95 Frequency-dependent gain, matching, and noise performance of the amplifier in Figure 3-94.

of the GaAs MMIC; Figure 3-97 shows its frequency-dependent gain, matching, and noise performance; Figure 3-98 shows its two-tone response, which is used to obtain the intercept point ($IP_{3,in}$ = 8.5 dBm, $IP_{3,out}$ = 22 dBm); and, finally, Figure 3-99 shows the three-tone test results at an input RF level of –21/–27 dBm. While the gain–bandwidth product is considerably larger than the single-stage amplifier we examined before, the intercept point and three-tone performance do not reach the same values. As a general rule, the dynamic

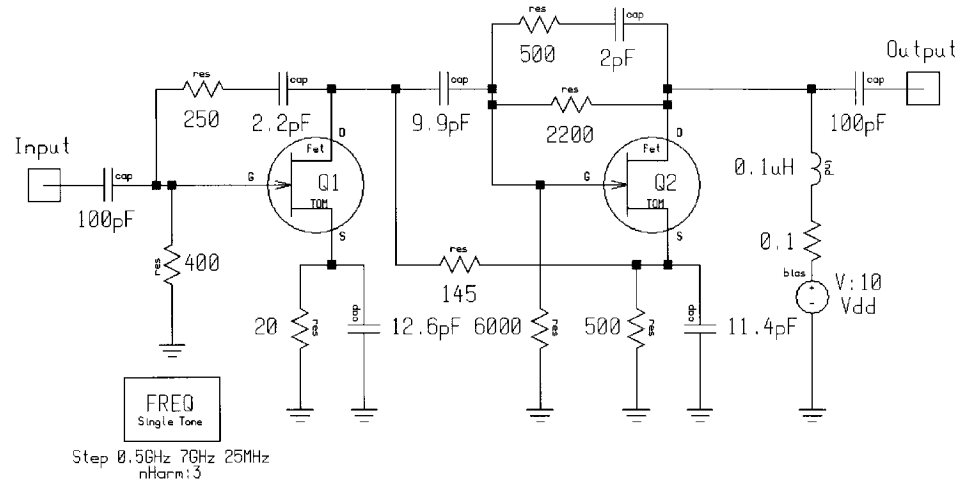


Figure 3-96 Simulation schematic for the MGA64135 MMIC amplifier. Q1 runs at 17.4 mA; Q2 at 32.2 mA.

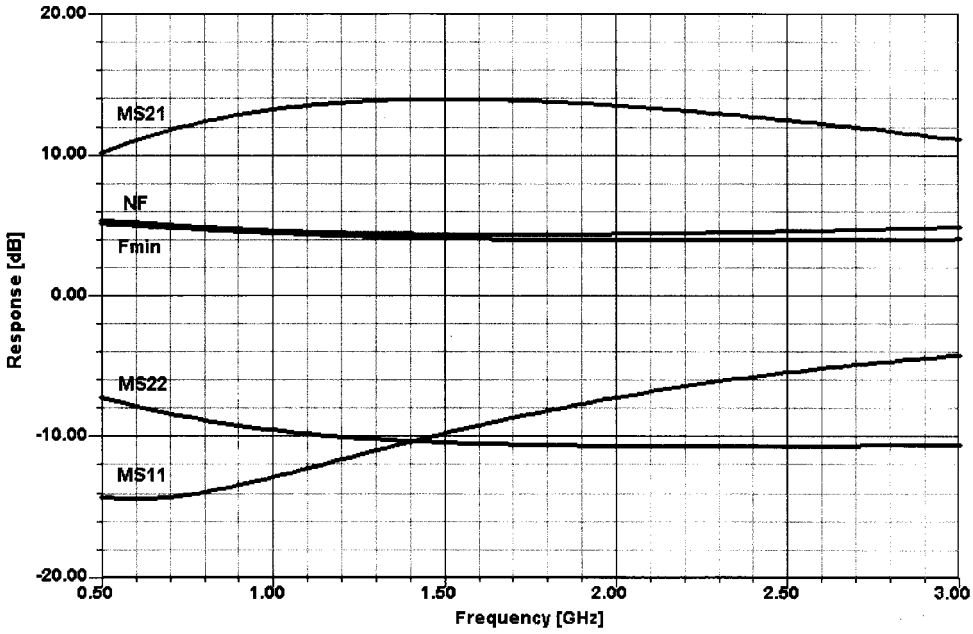


Figure 3-97 Frequency-dependent gain, matching, and noise figure performance for the MGA64135 MMIC amplifier.

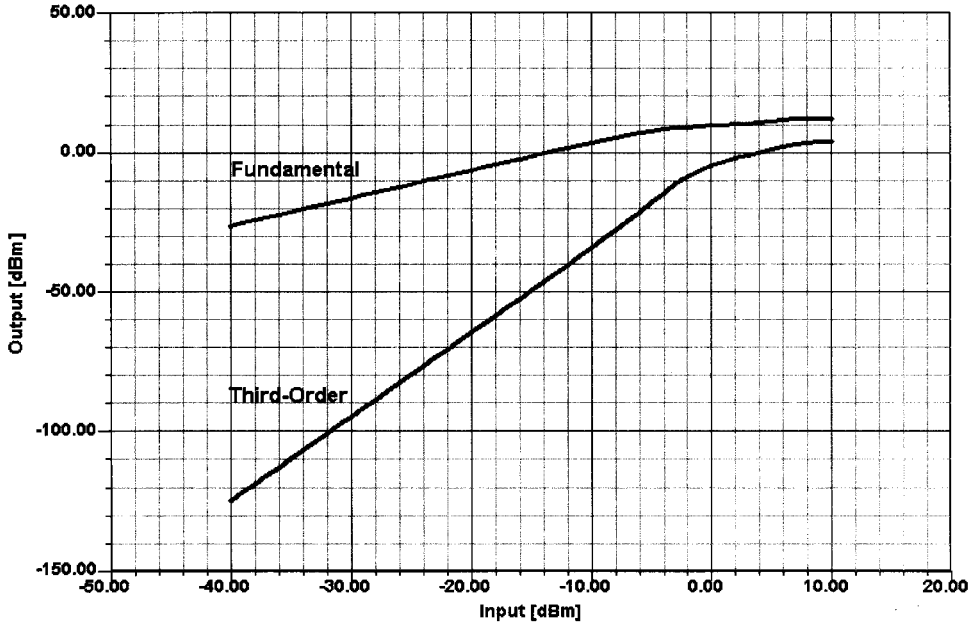


Figure 3-98 Simulated two-tone testing of the MGA64135 MMIC amplifier. $IP_{3,in}$ = 8.5 dBm and $IP_{3,out}$ = 22 dBm.

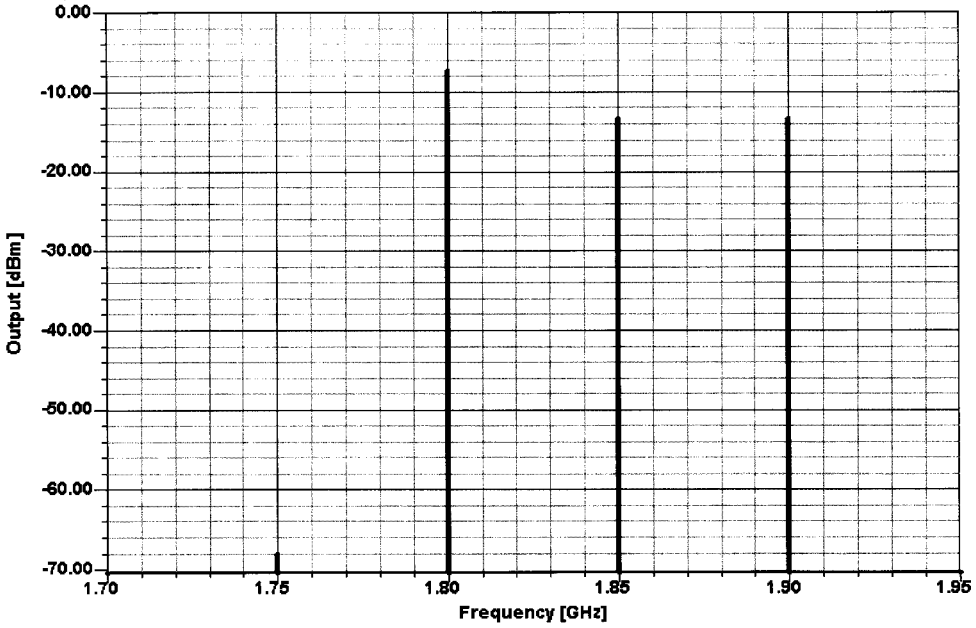


Figure 3-99 Simulated three-tone test of the MGA64135 MMIC. The test tones were 1.80 GHz (–21 dBm), 1.85 GHz (–27 dBm), and 1.90 GHz (–27 dBm). The 1.75-GHz IMD product is down 60.5 dB relative to the 1.80-GHz signal.

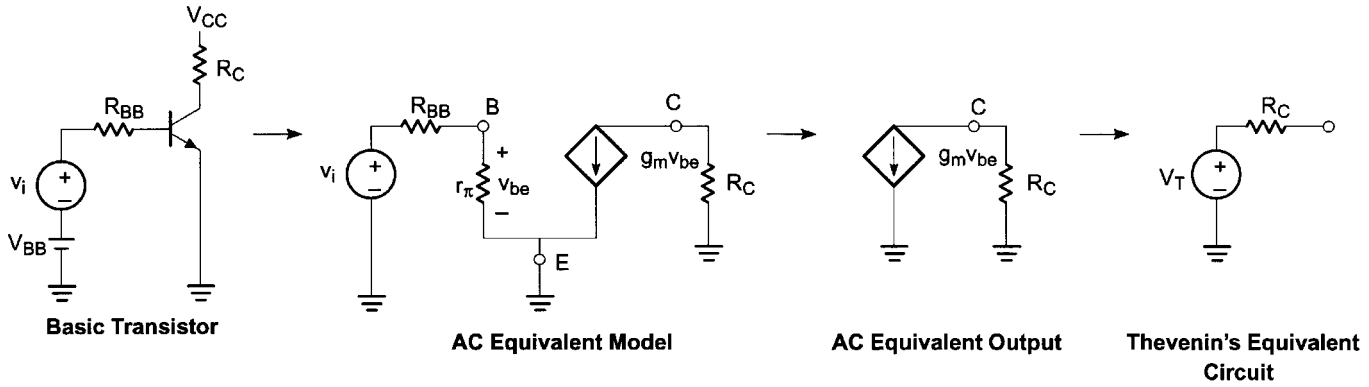
range is directly proportional to the power dissipated in the devices: specifically, the current. As explained earlier several times, the first-order nonlinearity comes from the voltage-to-current transfer characteristic, such as the transconductance itself, and can only be improved by a combination of resistive feedback in the emitter/source and by high currents. At these dc levels, the base–emitter (gate–source) junction gets linearized because the differential values are getting smaller compared to the parasitic loss resistances, including external dc/RF feedback. Once again, the voltage feedback reduces the voltage clipping but is initially more responsible for the input and output matching.

3-2-4 Low-Voltage Open-Collector Design [22–25]*

Why R_C Acts Like a Source Resistor. An open-collector output allows a designer the flexibility to choose the value of the R_C resistor. Choice of this resistor value not only sets the dc bias point of the device but also defines the source impedance value. Figure 3-100 shows the ac model of the transistor. Converting the output structure by applying a Norton and Thevenin transformation, one can conclude that R_C becomes the source resistance. Thus, by choosing R_C to be equal to the load, maximum power transfer will then occur.

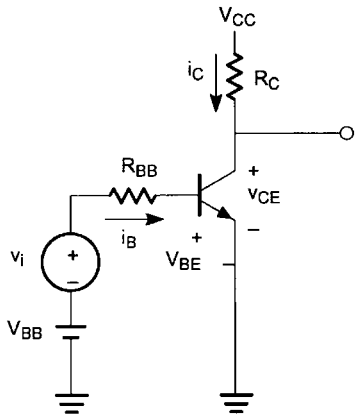
Figure 3-101 shows an active transistor with collector and base resistors. From basic transistor theory, Eq. A in Figure 3-101 is generated and has the same form as the general equation for a straight line ($y = mx + b$).

*Based on portions of the Philips Semiconductors/Signetics RF Communications Products Application Note AN1777, "Low-Voltage Front End Circuits: SA601, SA620," August 20, 1997. Used with permission.



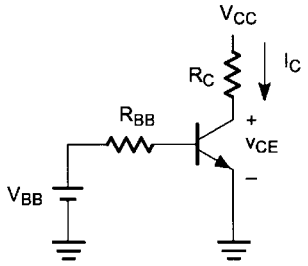
Symbol Convention:
DC: represented by uppercase symbols with uppercase subscripts, i.e., I_C
AC: represented by lowercase symbols with lowercase subscripts, i.e., i_C
Combined AC & DC: represented by lowercase symbol with uppercase subscript, i.e., i_C

Figure 3-100 *Rc* as source resistor.



DC Model

- Kill AC sources
- Open capacitors
- Short inductors

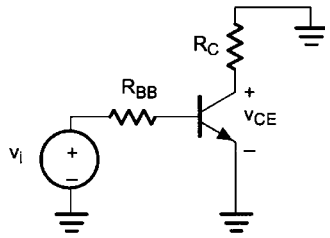


$$I_C = \left(\frac{-1}{R_C}\right) V_{CE} + \frac{V_{CC}}{R_C}$$

Eq. A

AC Model

- Kill DC sources
- Short capacitors
- Open inductors



$$i_C = \frac{-1}{R_C} (v_{CE})$$

Eq. B

Figure 3-101 Basic transistor analysis.

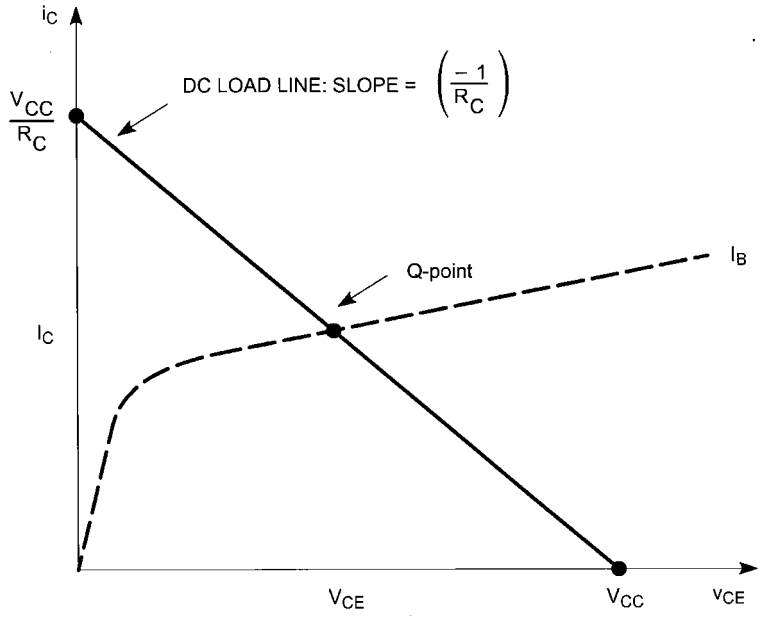


Figure 3-102 Load line and Q-point graph.

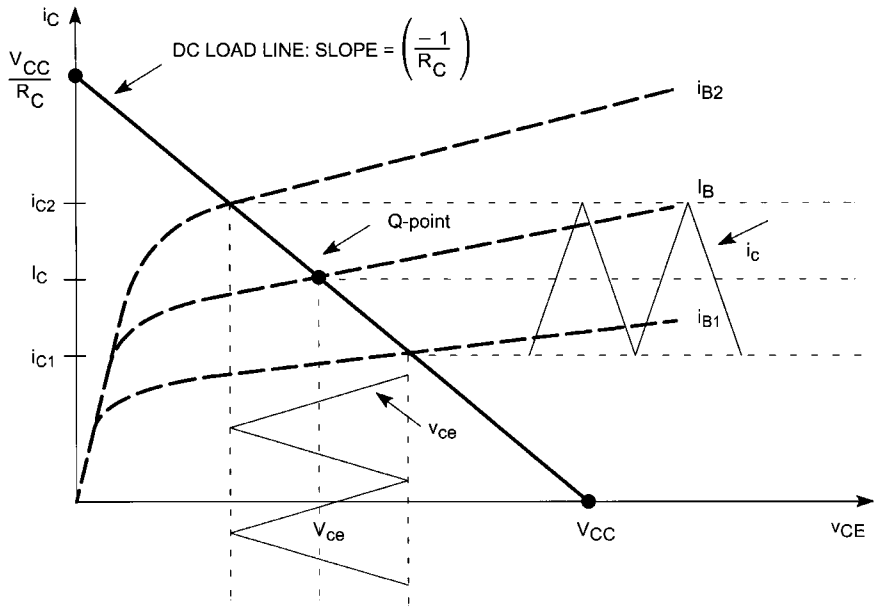


Figure 3-103 Graphical analysis for the circuitry in Figure 3-101.

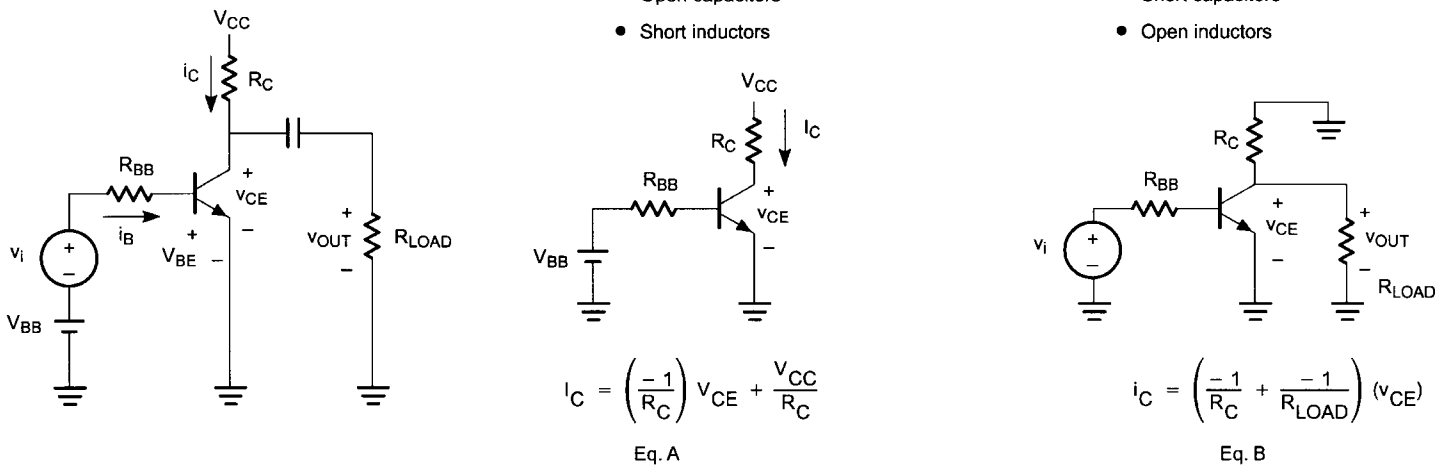


Figure 3-104 Basic transistor analysis with R_{LOAD}.

The slope of the dc load line is generated by the value of the collector resistor ($m = (1/R_C)$ and is shown in Figure 3-102. For a given small-signal base current, the collector current is shown by the dotted curve. The intersection of the dotted curve and the dc load line is called the *quiescent point* (Q-point) and dc bias determinant. The location of the Q-point is important because it determines whether the transistor is operated in the cutoff, active, or saturation region(s). In most cases, the Q-point should be in the active region because this is where the transistor acts like an amplifier.

Figure 3-103 shows the ac collector–emitter voltage (V_{CE}) output swing with respect to an ac collector current (I_C). Collector current is determined by the ac voltage presented to the input transistor's base (v_i) because it affects the base current (i_b), which then affects i_c. This is how v_i is amplified and seen at the output. Recall that this is with no external load (R_{LOAD}) present at the collector. Since no load is present, the slope of the ac load line is identical to that of the dc load line as seen in Eqs. A and B of Figure 3-101 ($m = -1/R_C$).

Open Collector with R_{LOAD}. A filter with some known input impedance is a typical load for the output of the transistor. For simplicity, we will assume a resistive load (R_{LOAD}) and neglect any reactance. Since a resistive load is used (see Figure 3-104), the ac output swing is measured at V_{OUT} or V_{CE}.

A dc blocking capacitor is used between the R_{LOAD} and the V_{CE} output to assure that the Q-point is not influenced by R_{LOAD}. It is also necessary to avoid passing dc to the load in applications where the load is a SAW filter. However, R_{LOAD} will affect the ac load line, which is seen in Eq. B in Figure 3-104. Note that the V_{CE} voltage swing is reduced; thus, the V_{OUT} signal is reduced (see Figure 3-105).

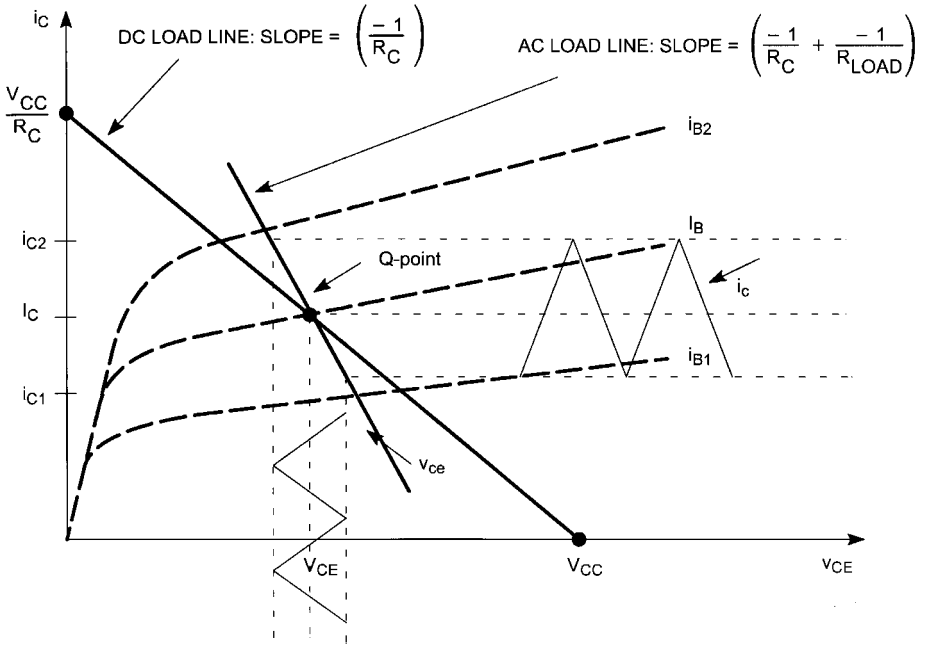


Figure 3-105 Graphical analysis for the circuitry in Figure 3-104. The ac load line intersects the Q point, and the slope is steeper than the dc load line due to R_{LOAD}, which makes V_{CE} smaller.

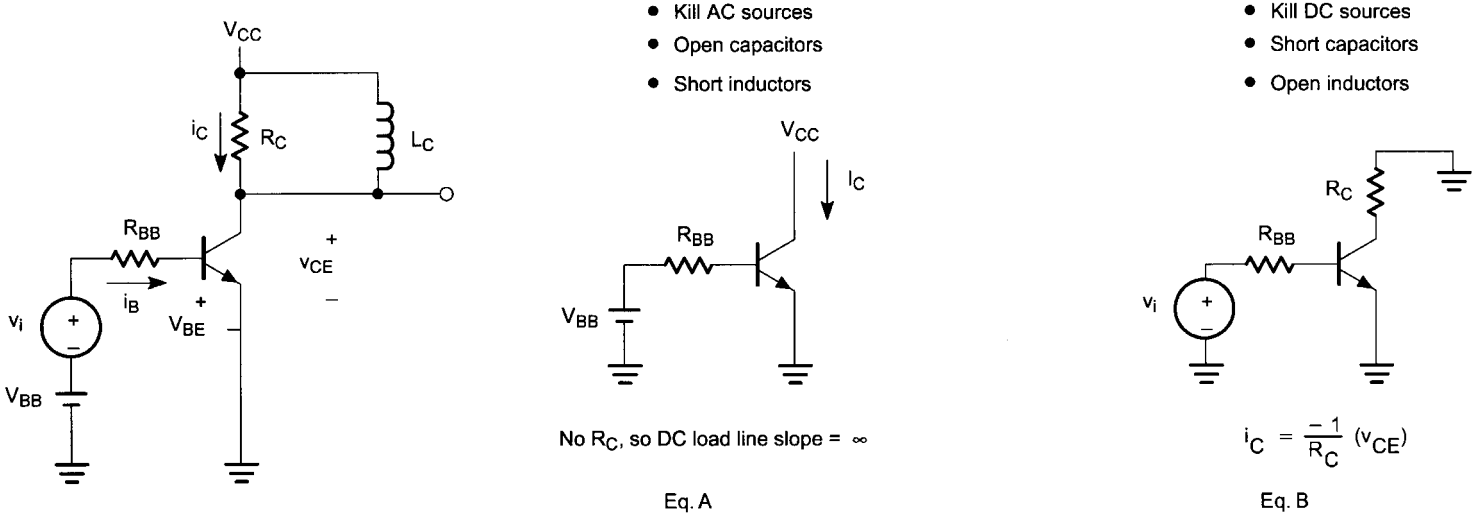


Figure 3-106 Basic transistor analysis with inductor added to collector.

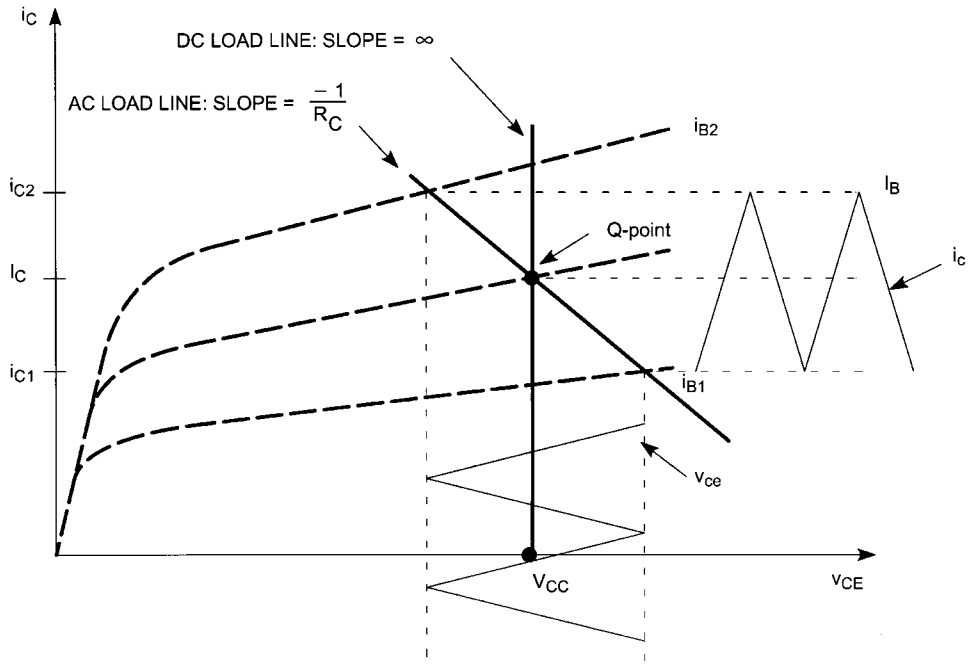


Figure 3-107 Graphical analysis for the circuitry in Figure 3-106.

Since the values of R_C and R_{LOAD} affect the ac load line slope, the value chosen is important. The higher the impedance of R_{LOAD} and R_C, the greater the ac output swing will be at the output, which means more conversion gain in a mixer. This is due to the slope getting flatter, thus allowing for more output swing.

Open Collector with Inductor (L_C). Adding an inductor in parallel with R_C can increase the ac output signal V_{CE}. Figure 3-106 shows the dc and ac analyses of this circuit configuration. In Eq. A in Figure 3-106, there is no R_C influence because the inductor acts like a short in the dc condition. This means the slope of the dc load is infinite and causes the Q-point to be centered around $V_{CC,}$ thus moving it to the right of the curve. The ac load line slope is set only by R_C because no load is present. Note that it has the same ac load line slope as the first condition in Eq. B of Figure 3-101.

Referring to Figure 3-107, one might notice that the base current (ac and dc) curves spread open as V_{CE} increases. This is caused by a noninfinite Early voltage (see Figure 3-108), which causes the collector current to be dependent on V_{CE}. Taking advantage of this nonideal condition, the peak-to-peak ac output swing V_{CE} can thereby be increased by moving the dc Q-point to the right due to the wider spreading between the curves corresponding to different base currents. Figure 3-109 combines Figures 3-103 and 3-107 to show the different ac output signals with different Q-points.

Looking at the ac output level, one might ask how the V_{CE} peak voltage can exceed the supply voltage V_{CC}. Recall that the inductor is an energy-storing device ($v = L\ di/dt$). Therefore, total instantaneous voltage is V_{CC} plus the voltage contribution of the inductor.

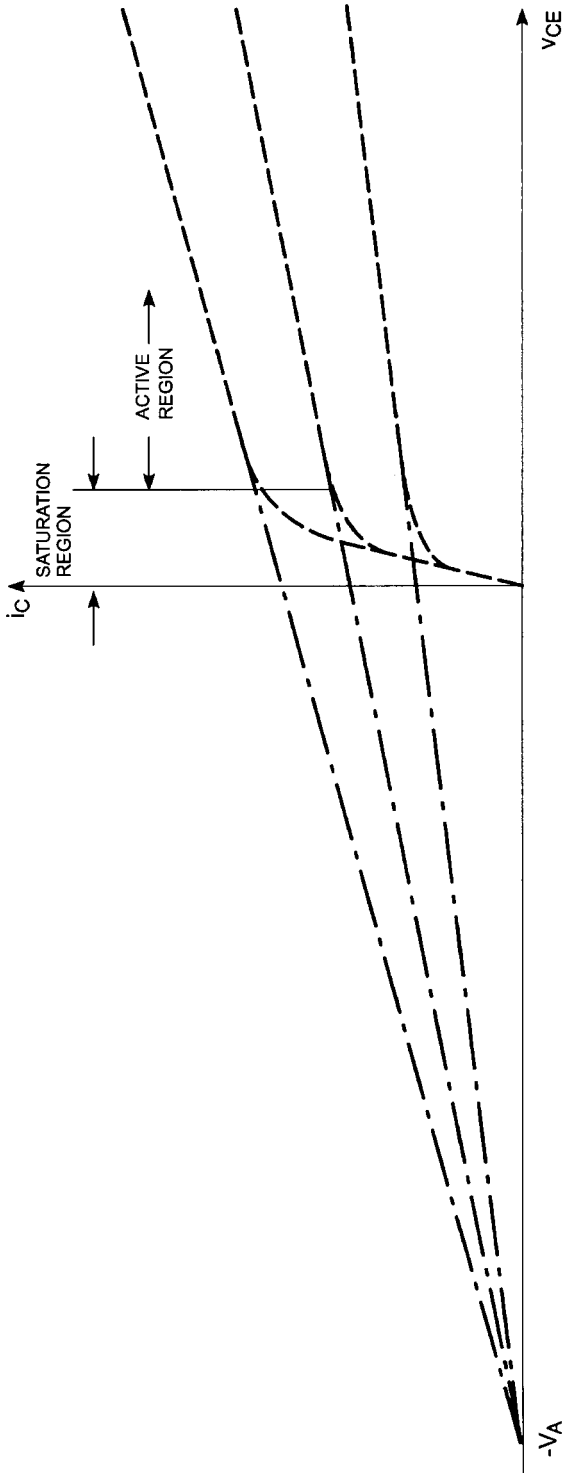


Figure 3-108 Graphical representation of Early voltage effect.

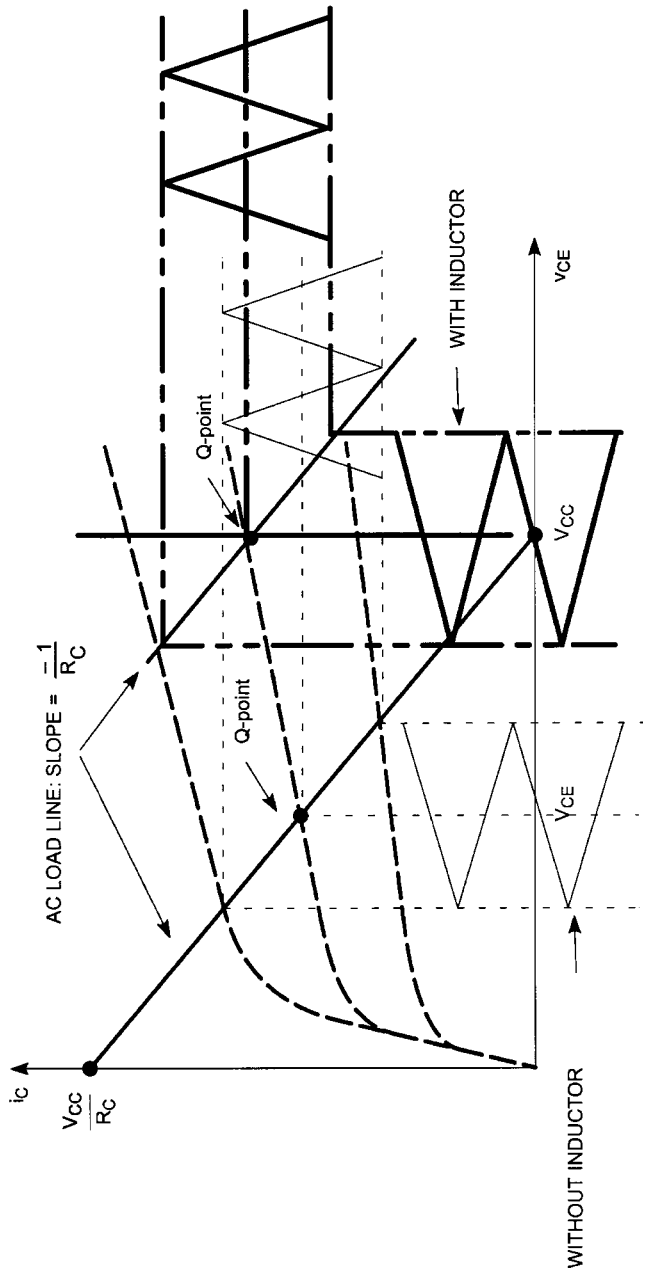


Figure 3-109 Comparison of open-collector circuit with inductor versus without inductor.

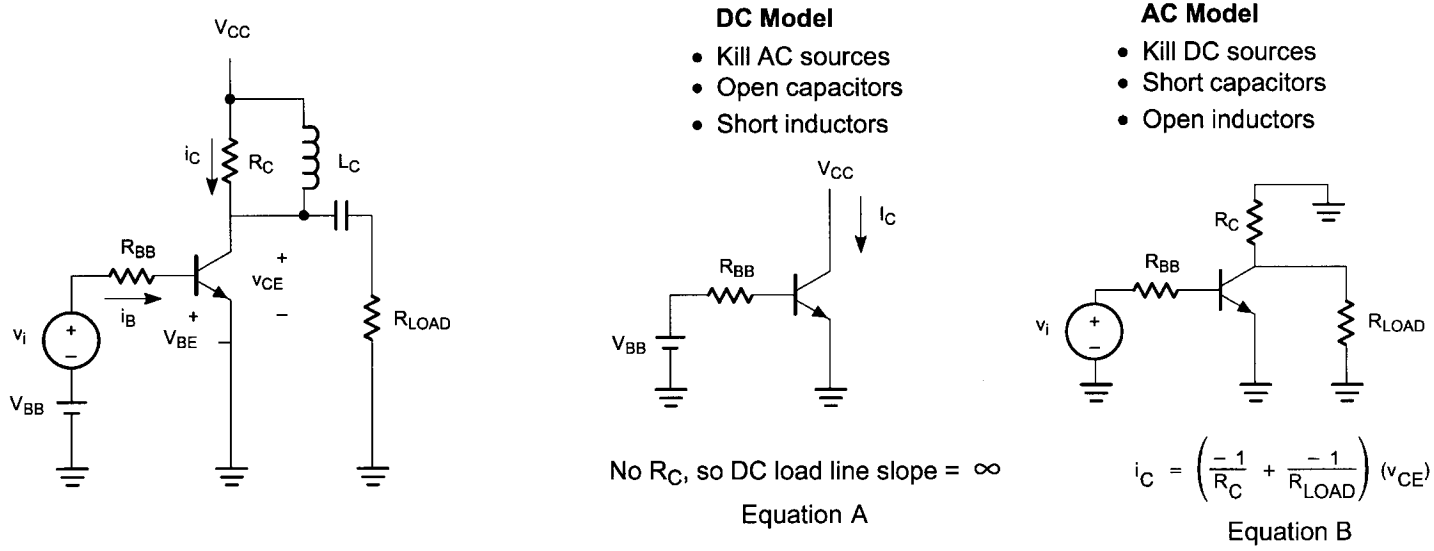


Figure 3-110 Basic transistor analysis with inductor and R_{LOAD}.

Open Collector with Inductor (L_C) and R_{LOAD}. Figure 3-110 shows the dc and ac analyses with the inductor and load resistor. Again, from the dc analysis, the inductor causes R_C to be nonexistent so the dc load line is vertical. In the ac analysis, the ac load line slope is influenced by both R_C and R_{LOAD} (see Eq. B in Figure 3-110). The ac load line slope is the same as the example of the open collector without the inductor. Figure 3-111 shows the response for the open collector with L_C and R_{LOAD}. Figure 3-112 combines Figures 3-105 and 3-111 showing the increase in ac output swing.

In conclusion, for the load line, R_C plays a role in setting up the bias-determined Q-point as well as the ac source impedance. However, when an inductor is placed in parallel with R_C, a different Q-point is set and the ac source impedance is altered. Moving the Q-point takes advantage of the transistor's nonideal i_c dependence on V_{CE} to get more signal output without having to change the base current. Since R_C is in parallel with R_{LOAD} in the ac condition, it influences the ac load line slope.

Flexible Matching Circuit [25]. A useful variation of the open-collector matching concepts previously outlined provided the capability of delivering equal power to two unequal resistive loads. This allows the power delivered to the load to be measured indirectly at another test point in the circuit where the impedance can be defined arbitrarily. If this impedance is defined to be 50 Ω, a spectrum analyzer can easily be placed directly into the circuit. This is an excellent troubleshooting technique and a valuable option to have available in high production environments.

Figure 3-113 shows the schematic for this flexible matching circuit. In this circuit, C_B functions only as a dc blocking capacitor and presents a negligible impedance at the frequency of interest. Recall from the previous open-collector matching discussions that,

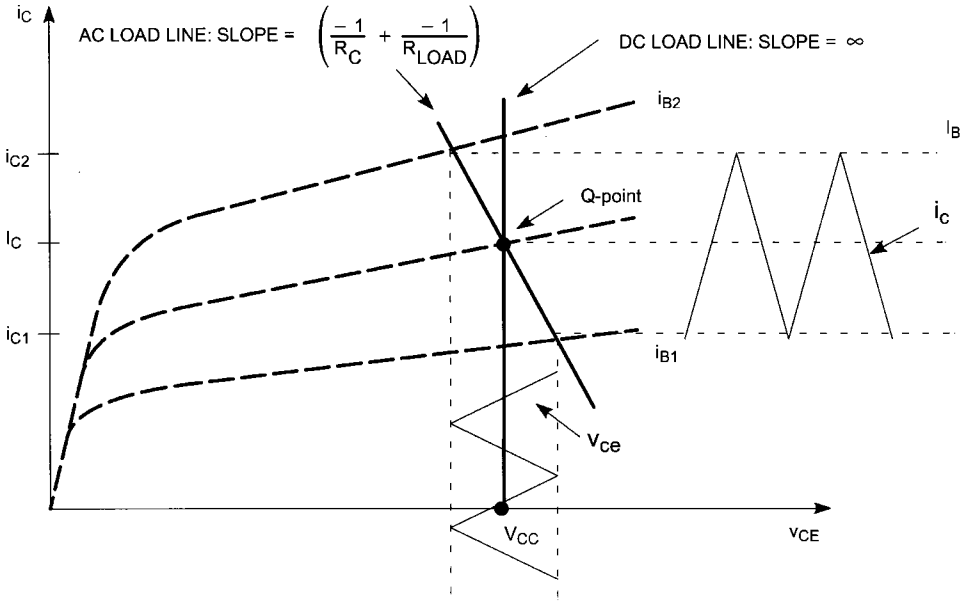


Figure 3-111 Graphical analysis for the circuitry in Figure 3-110.

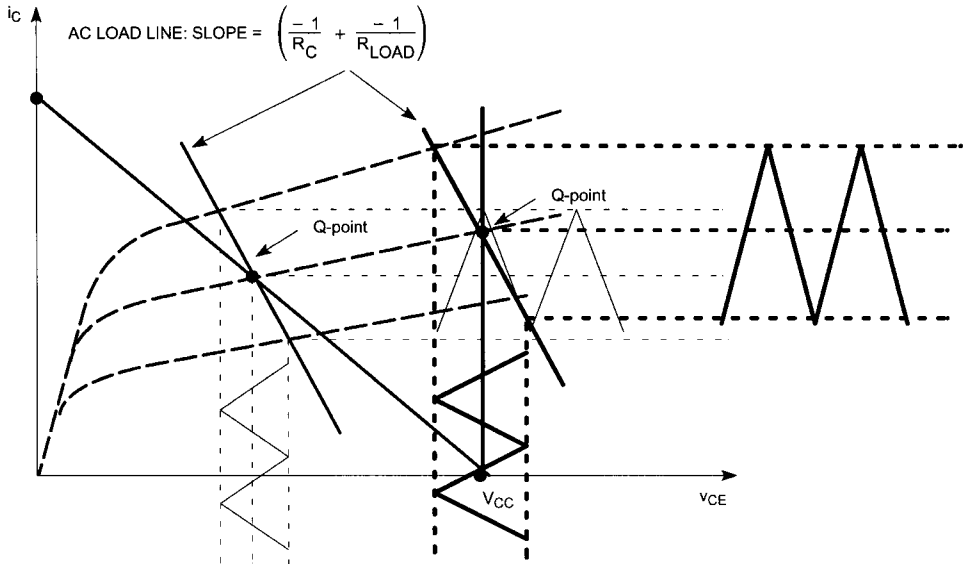


Figure 3-112 Comparison of open-collector R_{LOAD} circuit with inductor versus without inductor.

when R_C is placed in parallel with an inductor, it has no effect on the Q-point but does influence the slope of the ac load line as

$$\text{Slope} = \frac{-1}{R_C} + \frac{-1}{R_{\text{LOAD}}} \tag{3-166}$$

The capacitor C_S not only functions as a dc blocking capacitor but is also chosen such that the impedance presented by the combination of L, R_C, and C_S is equal to R_{LOAD} for optimum power transfer. The analysis is done in the following manner. First, note that inductor L is connected to V_{CC}, which is an affective ac ground. So L can be redrawn to ground. Next, R_C and C_S are converted to their parallel equivalent values as shown in Figure 3-114.

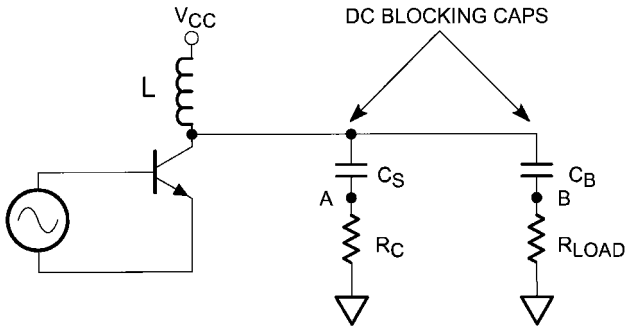


Figure 3-113 Flexible matching circuit.

Figure 3-114 Converting from series to parallel configuration.

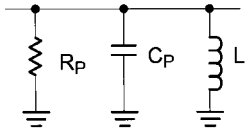


Figure 3-115 Converting from series to parallel configuration. Goal: Find L_C and C_p such that R_p looks like the R_{LOAD} value at the resonant frequency.

The resulting parallel *LCR* circuit is shown in Figure 3-115. At resonance, the parallel L–C_P combination will be an effective open circuit leaving only R_P. R_P is then simply chosen to be equal to R_{LOAD}.

Starting with the most simple transistor configuration and using *RC* types of dc-coupled amplifier circuits, we next present an overview of possible configurations for this application.

3-3 SINGLE-STAGE FEEDBACK AMPLIFIERS

This amplifier (Figure 3-116) is a standard minimal configuration amplifier with reduction of current distortion because of the emitter feedback. The voltage gain is

$$\text{VG} = \frac{R_L}{(26\text{mV})/I_C + R_e} \tag{3-167}$$

where R_L is the total load seen at the collector.

The feedback amplifier in Figure 3-117 uses voltage feedback to "correct" the input and output impedance to be close to 50 Ω but does not improve the current distortion. The voltage gain of this circuit is

$$\text{VG} = \frac{R_F}{R_g} \tag{3-168}$$

$$R_{\text{IN}} = R_{\text{OUT}}$$

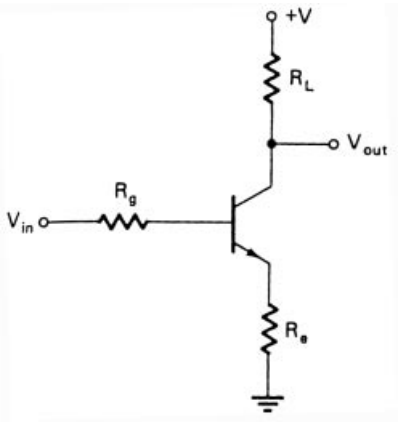


Figure 3-116 Single-stage amplifier with current feedback (provided by R_e).

$$= R_F R_D \quad \text{with} \quad R_D = \frac{26\text{mV}}{I_C} \tag{3-169}$$

Example. For $I_C = 10$ mA $\rightarrow R_D = 2.6\ \Omega$ and assuming $R_F = 200\ \Omega$, the input resistance $R_{\text{IN}} = 22.63\ \Omega$. The gain is calculated by getting the inverse of $R_D = g_m = 0.39$ A/V. Once this is multiplied with the output impedance of 22.63 Ω, the resulting gain becomes 10 log (8.6) or 9.3 dB. The gain can be increased by either using a larger collector current or a larger feedback resistor. It should be reminded that this circuit does not improve the distortion caused by the base–emitter junction.

To reduce the current distortion as well, we use the circuit shown in Figure 3-118, which contains both types of feedback. In reality, the collector–base resistor is frequently split into an unbypassed section responsible for the feedback and a larger resistor responsible for the appropriate biasing. The example circuit in Figure 3-119 shows this.

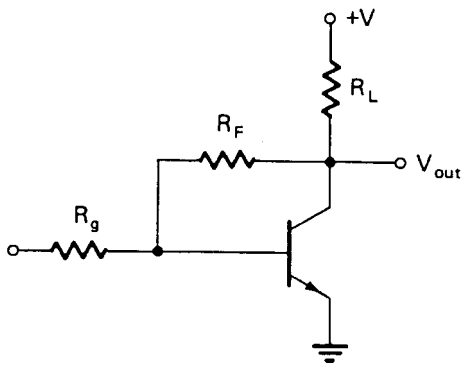


Figure 3-117 Single-stage amplifier with voltage feedback (provided by R_F).

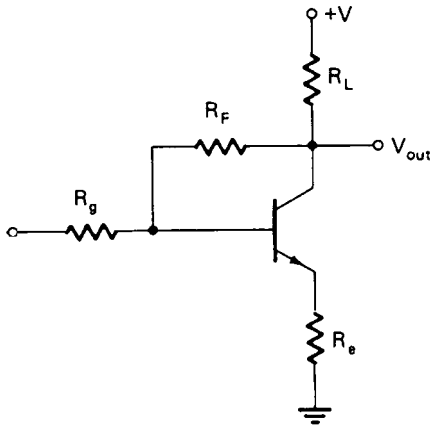


Figure 3-118 Single-stage amplifier with voltage and current feedback.

By using two of the fairly recent BFP420 transistors in parallel, and balancing the transistors by using the appropriate emitter resistors (which in total will result in 10 Ω) plus the bypassed value, we obtain the frequency-dependent gain, matching, and noise responses shown in Figure 3-120. To determine the appropriate feedback resistor R_F, we recommend the formula

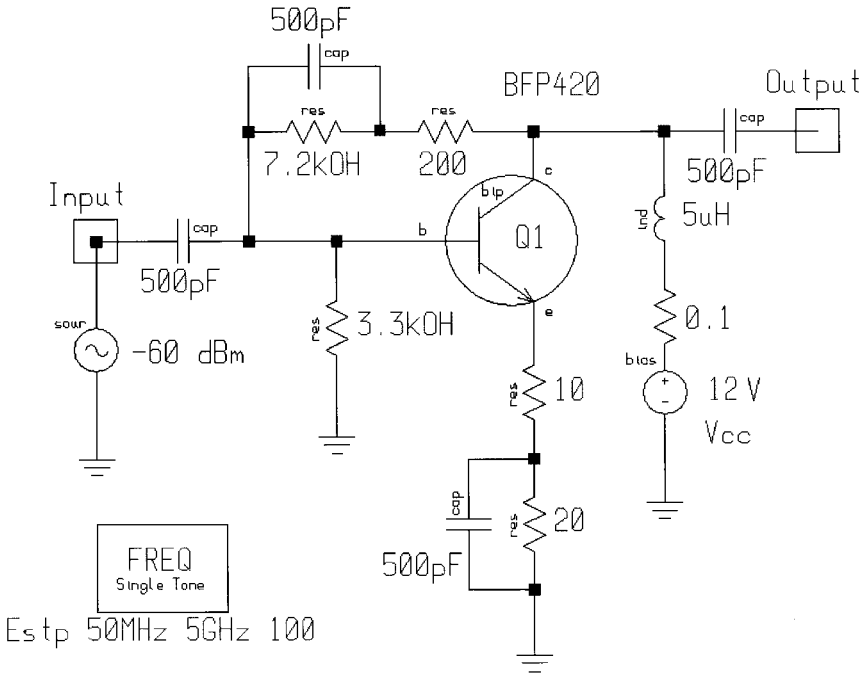


Figure 3-119 Practical amplifier circuit showing voltage and current feedback, and illustrating how reactances (in this case, 500-pF bypass capacitors) and multiple resistors can be used to provide different ac and dc resistances for the paths from collector to base and emitter to ground.

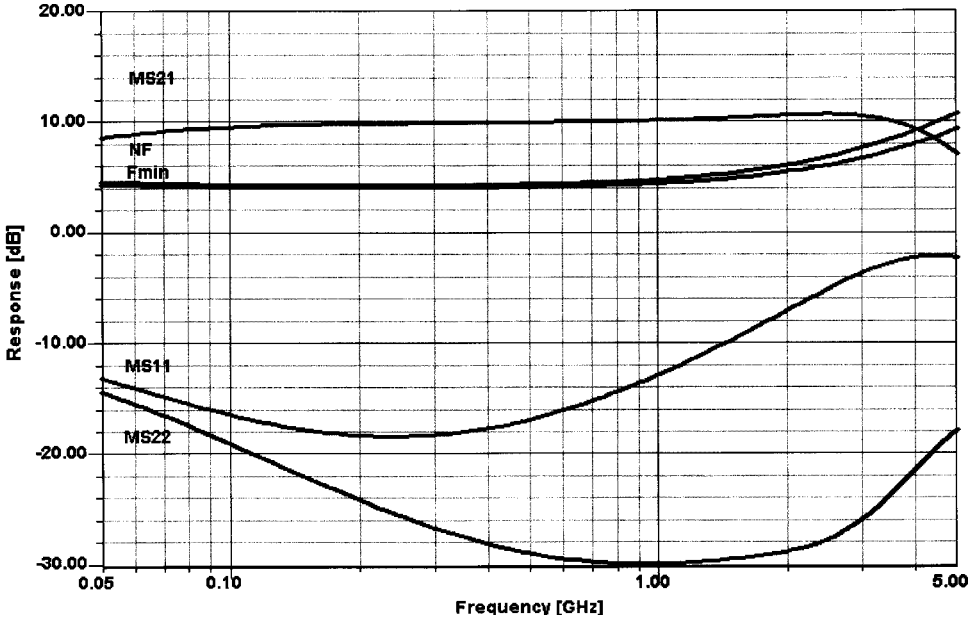


Figure 3-120 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-119.

$$R_F = \frac{(Z_{\text{in}}\, Z_{\text{out}})}{R_E} \tag{3-170}$$

This equation requires some additional comments. The combined shunt and series feedback only work as long as a sufficient transconductance is provided. In other words, there must be enough open-loop gain that will be reduced to obtain more bandwidth. A necessary condition for this approach, which really comes from the frequency range of less than 30 MHz, is that

$$Z_{\text{in}} = \frac{26\ \text{mV}}{I_C}\,\beta >> R_F \tag{3-171}$$

This requires a minimum transconductance of

$$g_m = \frac{1 + S_{21}}{Z_{\text{in}}} \tag{3-172}$$

where S_{21} must be a negative number that reflects the 180° phase shift. Of course, life becomes interesting when S_{21} goes through 1, but it (hopefully) never becomes negative and larger than 1 at the same time.

Example. The BFP420 has been used here frequently. A 10-dB gain would require an S_{21} of 3.16 ($\sqrt{10}$); therefore, we need to inspect the datasheet, which shows that the magnitude

of S_{21} is 3.16 around 4.5 GHz, yet the phase shift has changed from 180° to around 20°. In order to design an amplifier, we now set

$$R_F = Z_{\text{in}}(1 + |S_{21}|) \tag{3-173}$$

which is valid until S_{21} becomes too small. For a given 10 dB or 3.16 for S_{21}, the resulting RF is 4.16×50, or 208 Ω. The transconductance under these conditions becomes

$$g_m = \frac{1 + |S_{21}|}{Z_0} \tag{3-174}$$

where $Z_0 = 50$ Ω (the desired input impedance). g_m now becomes 83.2 mS. We now test the input and find

$$\frac{\beta}{g_m} = Z_{\text{in}} = \frac{50}{0.0832} = 600\ \Omega \tag{3-175}$$

Therefore, the emitter feedback resistor is not necessary to meet the previously mentioned condition whereby the input impedance needs to be higher than RF. Assuming the transconductance now is 0.00832, or 10% of the previous value, the input impedance would be 60 Ω, no longer meeting the prescribed requirement. Therefore, we need an unbypassed emitter resistor that fulfills the equation

$$R_E = \frac{Z_0^2}{R_F} - \frac{1}{g_{m_{\text{dc}}}} \tag{3-176}$$

where

$$g_{m_{\text{dc}}} = \frac{1 - S_{21_{\text{max}}}}{Z_0} \tag{3-177}$$

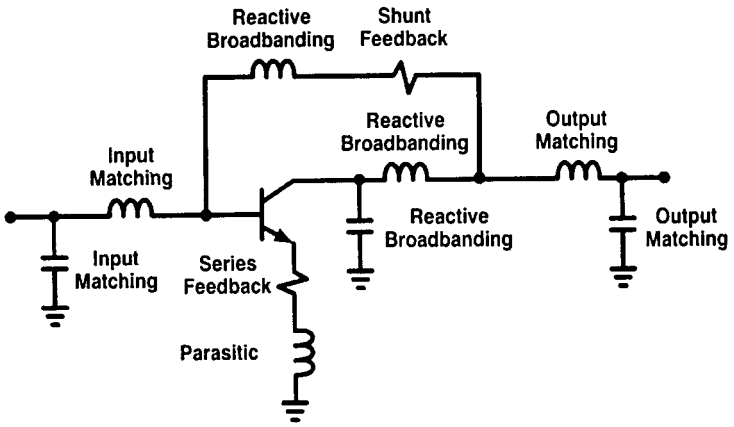


Figure 3-121 Elements of a feedback amplifier.

Again, the reader should be cautioned that in the common-emitter circuit, there is a 180° phase shift, and therefore $S_{21_{\max}}$ is a negative quantity, resulting in reality in

$$g_{m_{\text{dc}}} = \frac{1 + S_{21_{\max}}}{Z_0}$$

In our case,

$$g_{m_{\text{dc}}} = \frac{-S_{21}}{2Z_0} = \frac{38}{2 \times 50} = \frac{38}{100} = 0.38$$

so

$$R_E = \frac{50^2}{208} - \frac{1}{0.38} = 12 - 2.6 = 9.4\ \Omega \tag{3-178}$$

In practice, the designer would use a 10-Ω resistor.

Since the gain of the transistor continues to fall as a function of frequency, one can compensate to a degree for the cutoff frequency by putting an inductance in series that follows the equation

$$L_F = \frac{R_F}{2\pi f_{3\text{dB}}} \tag{3-179}$$

Assuming $f_{3\text{dB}}$ = 2 GHz, the resulting series inductance L_F = 16.6 nH. Basically, the same applies to circuits where the second transistor is used as an emitter follower to feed the signal back from the collector to the base of the first transistor.

In summary (see Figure 3-121):

- Resistive negative feedback at lower frequencies reduces open-loop gain and simultaneously matches input and output.
- Reactive elements set the gain and matching at high frequencies. Feedback is reduced at high frequencies.
- Emitter bypass capacitors are not used, as they would destroy the impedance match.

This is equally valid for FETs.

The next level of complexity can be reached by using two transistors with dc feedback.

3-3-1 Lossless or Noiseless Feedback

All the amplifiers shown above are based on a combination of resistive and reactive feedback. The very moment resistive feedback is used, the noise figure increases, as does the intercept point. As long as the intercept point increases overproportionally compared to the noise figure, this is an acceptable method. However, there are many cases in which an extremely low noise figure and a higher than normally achievable intercept point are required. To avoid this degradation of performance, Norton of Anzac developed a negative feedback structure based on the capabilities of the ferrite transformer. This patented technique (U.S. Patent Nos. 3,624,536 and 3,891,934, since expired), known as *lossless feedback*, provides lower noise figure and higher linear output than can be achieved using resistive feedback with the same transistor. Transformer feedback is used in which the transformer in the feedback network acts as a directional coupler. The coupling ratio between the input and output of the coupler

determines the magnitude of the feedback, and hence the gain. Across the frequency band over which the transformer operates as intended, the feedback is negative due to the in-phase coupling and the inversion of the active device.

This technique was used in amplifiers ranging from 5 to 500 MHz, 20 to 1000 MHz, and 300 to 1800 MHz. The difficulty associated with this type of amplifier is the fact that it is based on transformers that use twisted pairs of wires as transmission lines. In some implementations used at higher frequencies, actual directional couplers were used at the expense of bandwidth; in these circuits, the coupling was determined by the spacing of the coupler lines rather than transformer turns ratio. These microstrip couplers have much narrower bandwidths than transformers and are fairly large.

Another headache with the lossless feedback design is the fact that many of these transistor configurations are used in either grounded-base or grounded-emitter with transistors that have f_T values of 25 GHz and higher. Such devices are very hot, and achieving freedom from oscillation over a wide frequency range is a continuing problem, even for the circuits we have shown so far. A similar clever technique has been implemented in the termination-insensitive mixer, which we will examine in Chapter 4.

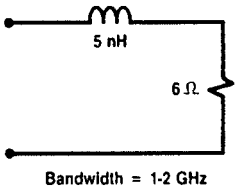
3-3-2 Broadband Matching

The circuitry above has enabled us to extend the frequency range of the amplifier and avoid the normal sharp cutoff. According to Fano [26], there is a limit to the bandwidth of matching. The concept is that impedance plus bandwidth determines the ability to match, since there is a Q of the circuit and there is a Q of the device. The following equations are valid:

$$|\rho_{\text{Bode}}| = e^{-(\pi Q_T / Q_{BP})} \tag{3-180}$$

$$Q_T = \frac{\sqrt{f_1 f_2}}{f_2 - f_1} \tag{3-181}$$

$$Q_{BP} = \frac{\text{Im}(Z)}{\text{Re}(Z)} = \frac{X_s}{R_s} \tag{3-182}$$



$$Q_{BP} = \frac{2\pi\sqrt{(2\text{ GHz})(1\text{ GHz})(5\text{ nH})}}{6\,\Omega} = 7.40$$

$$Q_T = \frac{\sqrt{(2\text{ GHz})(1\text{ GHz})}}{(2\text{ GHz} - 1\text{ GHz})} = 1.414$$

$$|\rho_{\text{Bode}}| = e^{-\left(\frac{\pi(1.414)}{7.40}\right)} = .549$$

$$\text{VSWR}_{\text{Bode}} = \frac{1 + .549}{1 - .549} = 3.43 : 1$$

Figure 3-122 Broadband matching example showing the Bode limit.

Figure 3-122 shows an example. This combination results in a VSWR of 3.43, which for practical purposes is already too high. The workaround is special circuitry (such as exponentially staggered impedance jumps in the matching network like 50 Ω > 20 Ω > 7 Ω > 3 Ω). This means we first transform the 50-Ω source of load impedance down to 3 Ω, and then look for a matching network that matches the transistor output/input, which is within a few ohms of 3 Ω. A detailed discussion of such matching can be found in the "Numerical Design" portion of Section 3-12-2 under the title "Broadband Matching Using Bandpass-Filter Networks—High-Q Case."

3-4 TWO-STAGE AMPLIFIERS

There are two popular configurations that have been used mostly for low-frequency applications. The first one (Figure 3-123, implemented in the circuit shown in Figure 3-124) shows a nice high-frequency performance assuming the emitter and collector currents are essentially the same. As for distortion products, this circuit behaves equivalently to the shunt impedance example from Figures 3-118 and 3-119, with the additional gain of the second transistor. Its frequency matching and noise performance are also shown, in Figure 3-125.

The next two-stage circuit (Figure 3-126, implemented in Figure 3-127) uses the collector current of the second stage as the feedback to the first. Proper biasing of this circuit is quite difficult and interactive. Figure 3-128 shows the frequency-dependent gain, matching, and noise figure. We believe that this circuit can be optimized further, but we leave such investigations to the reader.

A *cascode* amplifier is a special form of two-stage circuit. Although a cascaded pair acts like a single transistor, its merits are in its reduced feedback and resulting higher gain–bandwidth product. Figure 3-129 shows a simple cascode application that assumes the presence

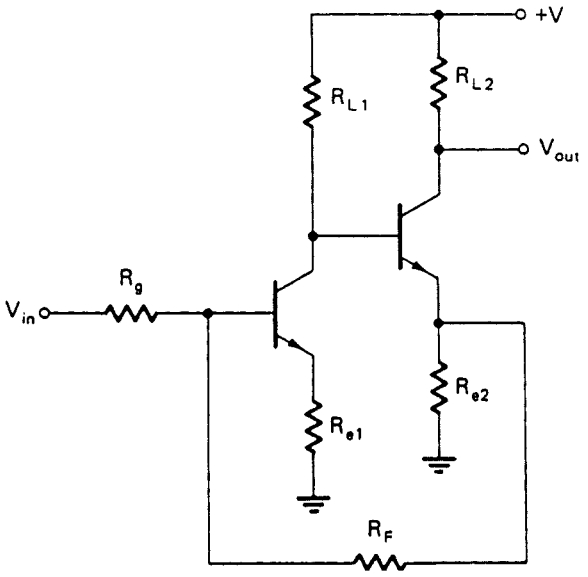


Figure 3-123 Two-stage amplifier with voltage feedback.

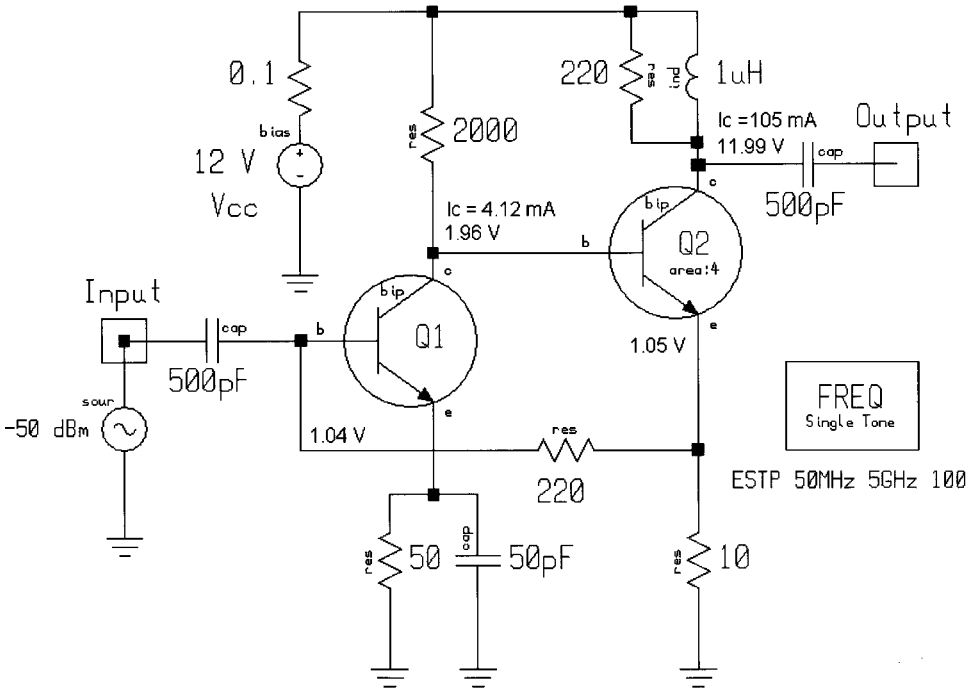


Figure 3-124 Implementation of the circuit in Figure 3-123.

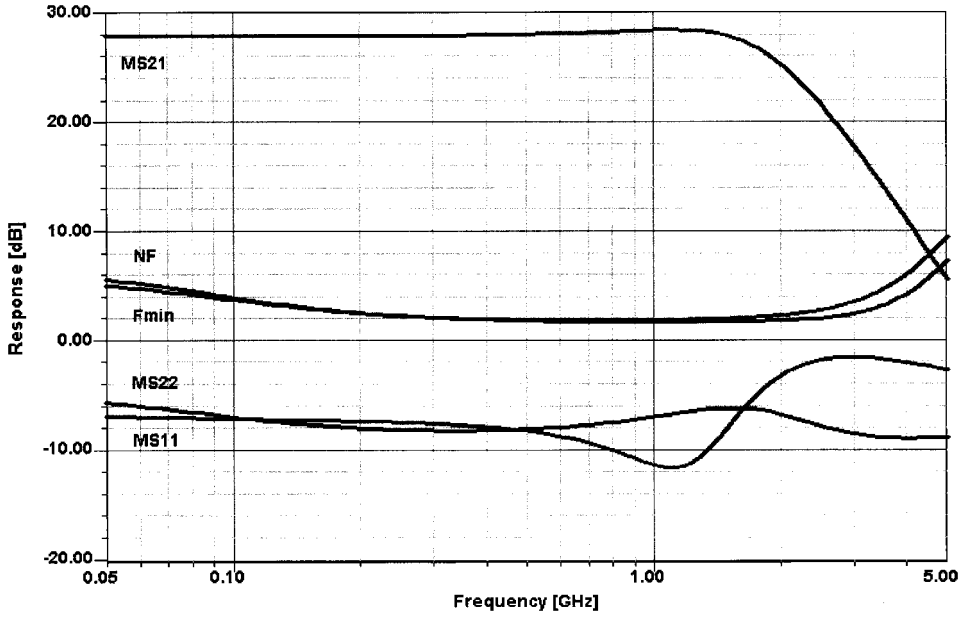


Figure 3-125 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-124.

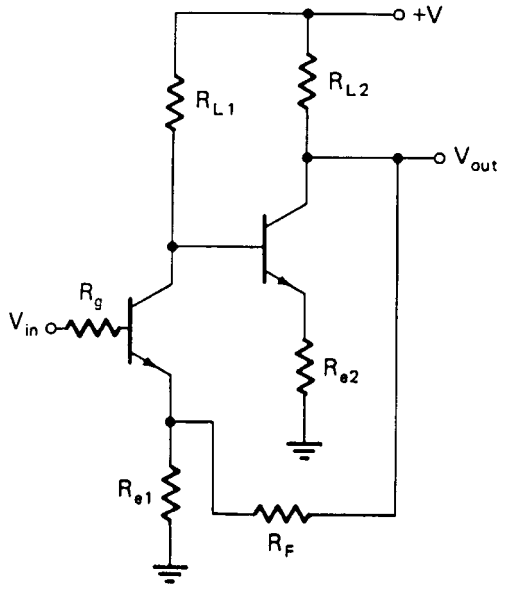


Figure 3-126 Two-stage amplifier with voltage feedback from Stage 2 to Stage 1, and current feedback in both stages.

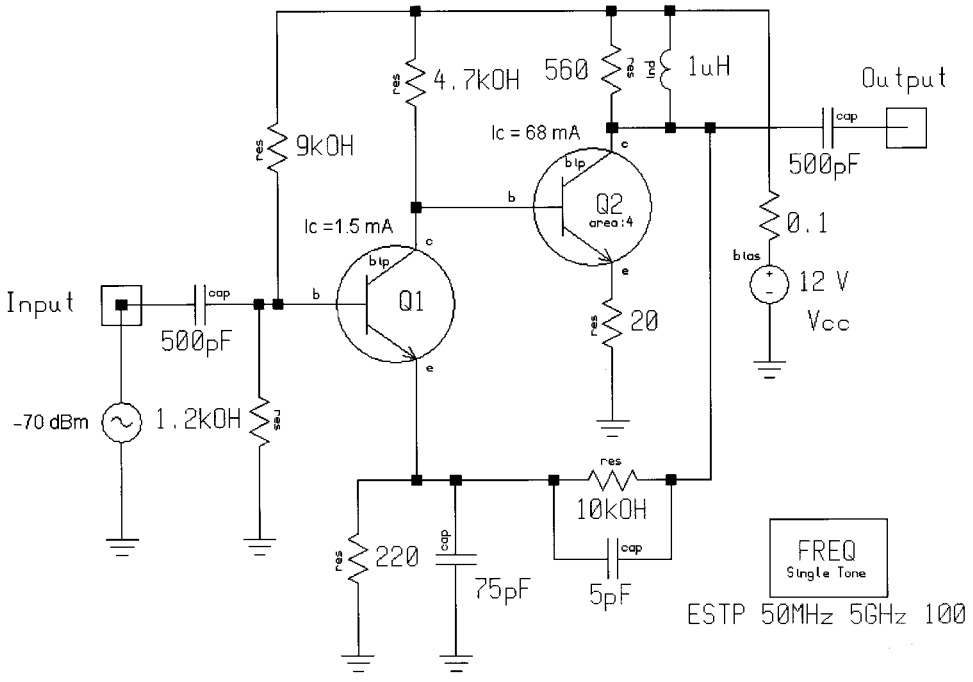


Figure 3-127 Simulated two-stage amplifier with feedback as shown in Figure 3-126.

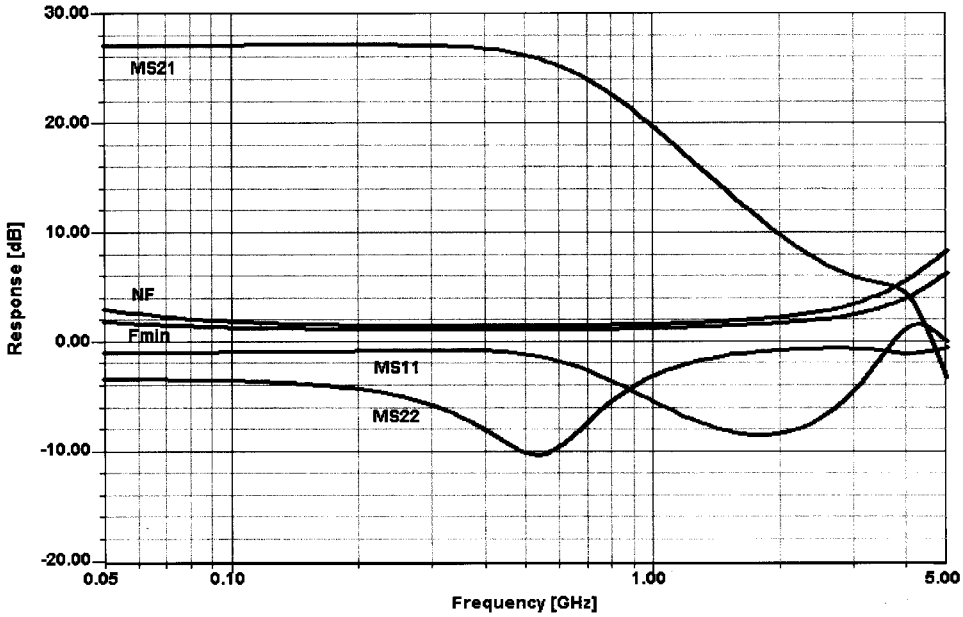


Figure 3-128 Gain, matching, and noise performance of the amplifier in Figure 3-127.

of an output transformer that absorbs the output capacitance. Figure 3-130 shows this circuit's frequency-dependent gain, matching, and noise performance.

It has been obvious to us that most published wireless application circuits are evaluated on a no-input-selectivity basis; that is, their reported performance does not take into account the filtering that is usually present in practical implementations. Evaluated in this way, a low-noise design may achieve a noise figure of less than 2 dB—1.2 dB or so is typical—but this is misleading because filters and connectors introduce insertion loss, resulting in a higher noise figure.

The cascode is a nice preamplifier combination. It combines (depending on the transistor) a low noise figure with essentially no feedback. Since the output of a grounded-base transistor provides a high-impedance source, this configuration is ideal to work into the SAW filters frequently used in wireless applications at the operating frequency. Special transistor pairs are available for use in cascode; as can be seen from this example, not every transistor combination is ideal. This one gives a higher noise figure than wanted and at frequencies above 2 GHz it tends to show possible instabilities. Both Siemens and Motorola are now offering transistors appropriate for these cascodes. Mostly, one would resort to transistors of less-exotic cutoff frequencies to guarantee stability. Typically, grounded-base/gate transistors tend to become oscillators, specifically because this configuration, like the emitter follower, allows a higher operating frequency while maintaining the same gain–bandwidth product; the actual gain for the last two circuit configurations is less. A grounded-base/gate stage has a high voltage gain and the emitter/source follower has power gain at low impedance levels. We invite our readers to experiment with the capacitor from the base of Q1 to ground in Figure 3-129 to get an interesting education about these transistors and their willingness to oscillate.

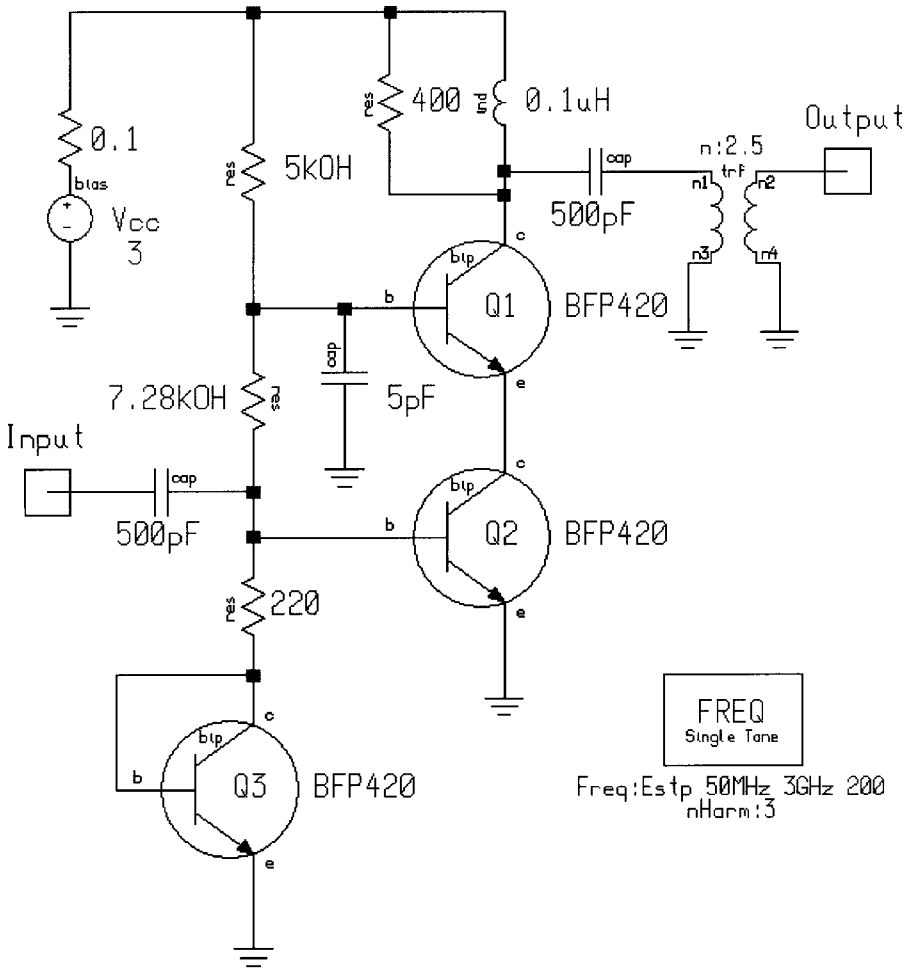


Figure 3-129 Schematic of the cascode amplifier.

The MC13144 is an example of a commercially available VHF to 2-GHz low-noise amplifier with programmable bias. As can be seen from Table 3-11, it is ideal for the 900-MHz range and its programmable gain/bias makes it very attractive. This design (Figure 3-131) is based on the cascode mentioned above and, because of high integration, uses many transistors for bias and level shifting.

The 900-MHz application circuit in Figure 3-132 shows the required external components for a successful design. Input and output are made tunable. The 5.6- and 39-nH inductances can be made from small air-wound coils. The output ceramic filter provides the necessary selectivity ahead of the mixer. With such a highly selective device in the input, the noise figure would deteriorate too much. The antenna is rarely in the vicinity of sufficiently strong transmitters to require such strong input filtering.

There are two more interesting applications for two-stage amplifiers: The one implemented in the Hewlett-Packard MSA-0735 MMIC has become a quite common one, and its usefulness up to higher frequencies depends mostly on the cutoff frequency of the transistors,

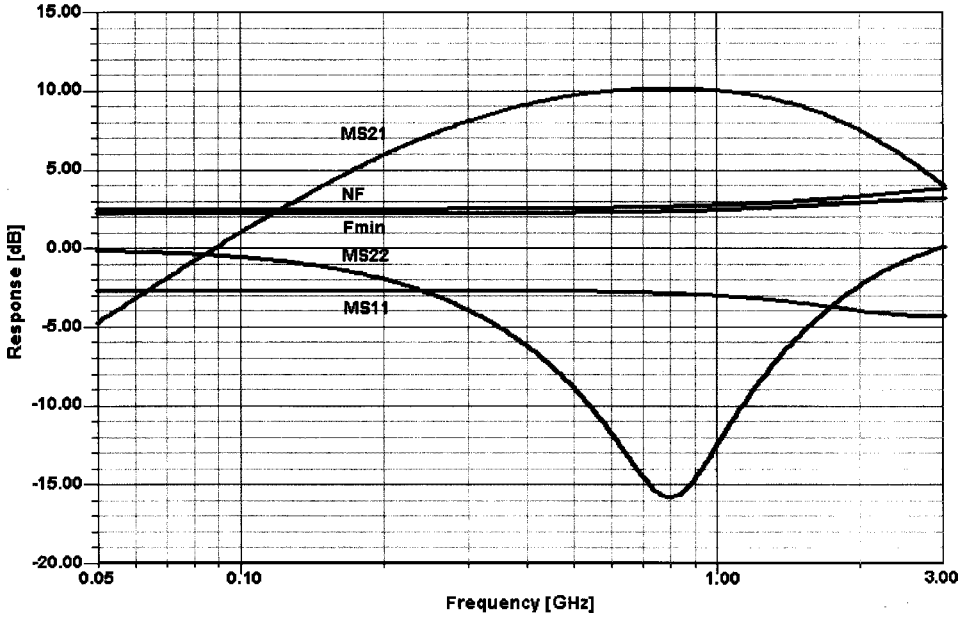


Figure 3-130 Frequency-dependent gain, matching, and noise performance of the cascode amplifier.

their parasitics, and the phase shift of the gain. The typical configuration is shown in Figure 3-133, and its frequency-dependent gain, match, and noise figure in Figure 3-134. There is an increase of S_{22} at higher frequencies; this indicates unwanted phase shift in the circuit. Some manufacturers hide the compensation circuits to overcome such peaking; most of the schematics shown by the manufacturers are really intended only to show operating principles, with proprietary details omitted.

Thanks to the properties of GaAs transistors, we can extend the frequency range significantly. The amplifier we are about to look at goes back several years and is described in great detail in Goyal [27]. Essentially what we have, seen in Figure 3-135, is an FET version of Figure 3-127. The main goal of this circuit was to have an all-monolithic circuit, avoiding all ac coupling and accomplishing all matching and interaction with dc coupling. The transistor Q1 is the main gain stage, which has the transistor Q2 as an active load. The output impedance of Q2 is $1/g_m$—typically in the area of 50–200 Ω, depending on the biasing,

Table 3-11 MC13144 features

Gain at 900 MHz, 17 dB
Noise figure at 900 MHz, 1.4 dB
P_{-1dB}, −7.0 dBm
$IP_{3,in}$, −5.0 dBm
Low operating supply voltage (1.8–6.0 Vdc)
Programmable bias with Enable 1 and Enable 2. [Enable 1 and Enable 2 programmed high for optimal noise figure and gain associated with NF. Enable can be overridden for external bias programming (up to 15 mA.)]

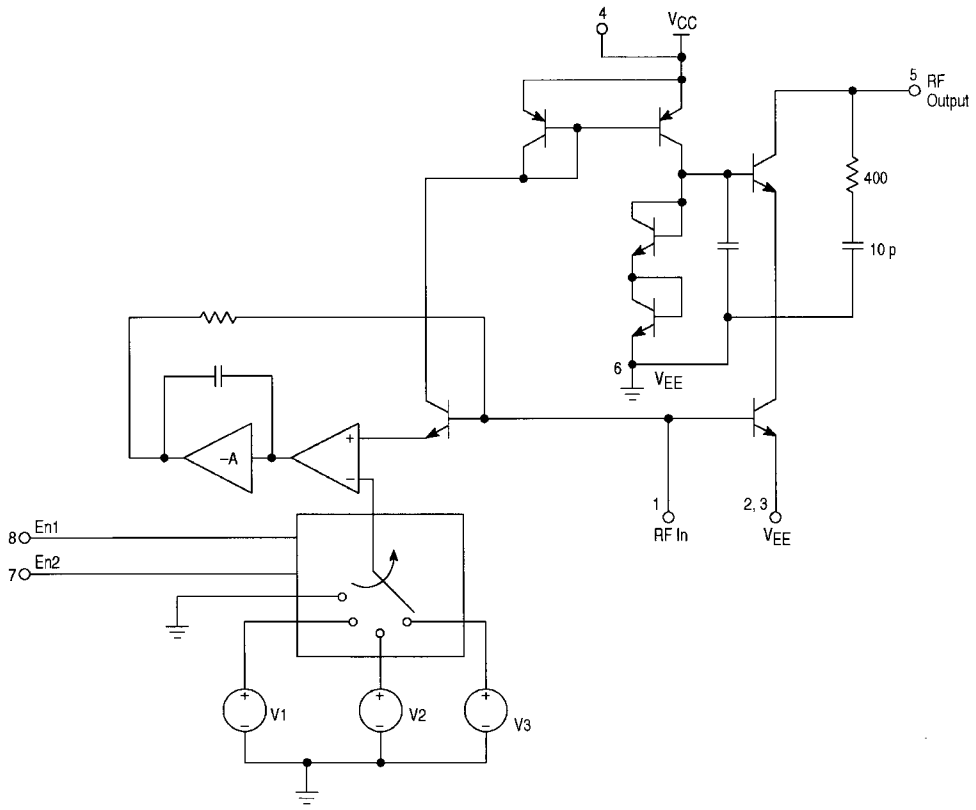


Figure 3-131 Schematic of the Motorola MC13144 cascode LNA.

which affects the transconductance. Transistor Q3 has the same function as an emitter/source follower; again, its dc bias determines its transconductance and therefore its output impedance. (Given this, a device that has a transconductance of 20 mS, or close to it, would establish a very good output match to 50 Ω.) The drains of Q1 and Q5 are at approximately 3.5 V. The diodes at the source of Q3 are level shifters that must shift the dc voltage at the gate of Q5 and at the drain of Q4 to approximately −1.5 V relative to ground—the necessary bias condition for Q1 and Q5. The standard shunt feedback circuit would show resistive feedback between the drain of Q1 (gate of Q2) and the gate of Q1; in the previous examples, we calculated the magic value of this part to be about 200 Ω. Adding in parallel to Q1 the transistor Q5 allows the designer to use a feedback scheme that isolates the feedback loop from the input. Dc-wise, the two transistors are tied together via a 10-kΩ resistor, which can be included in the actual packaged device. Insufficient information was published to allow us to duplicate the design, considering all the intricacies involved (we somewhat object to the fact that even in the *IEEE MTT-S Transactions*, so little information is given about the actual implementation of this design—using space constraints as an excuse—that its educational value is usually minimal); it shows a new principle without revealing the design steps necessary to obtain a practical circuit. Several variations of this type of circuit are available; the literature is full of discussions about them. The magnitude of the feedback is set by the ratio of the width of Q5 to that of Q1. Typical values for the Q5/Q1 ratio range

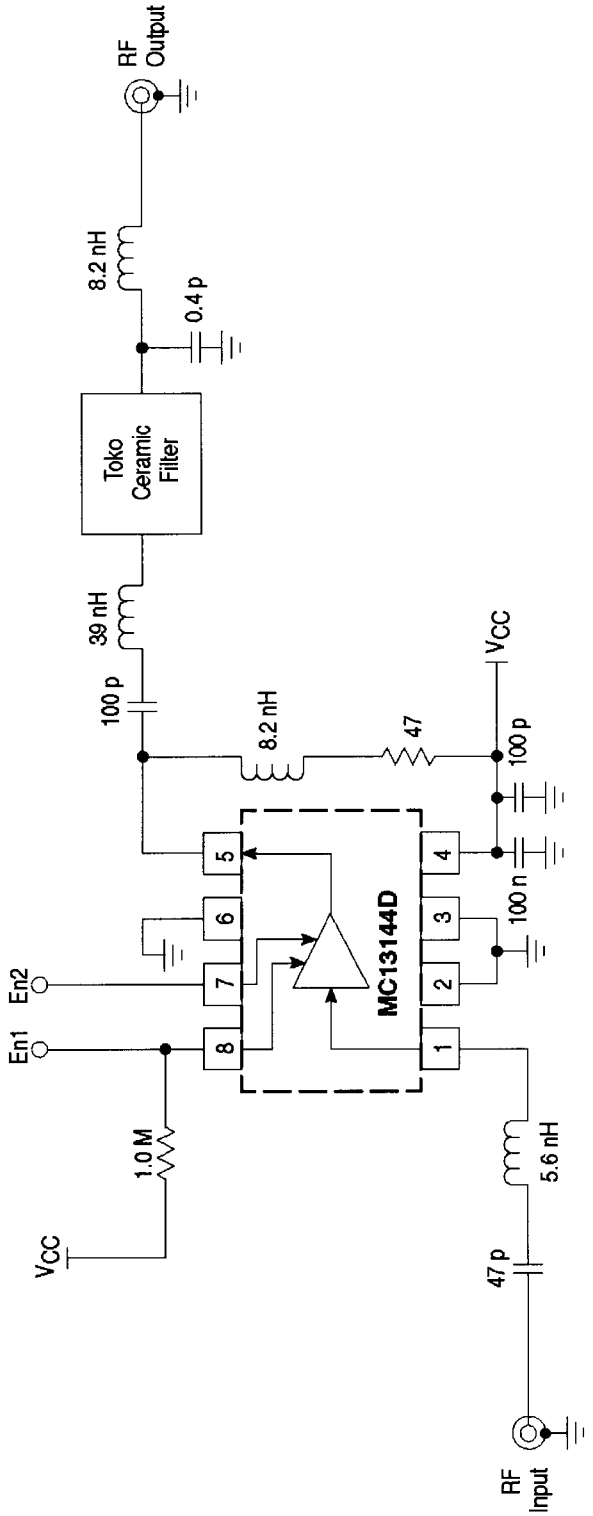


Figure 3-132 An MC13144 application circuit for 926.5 MHz.

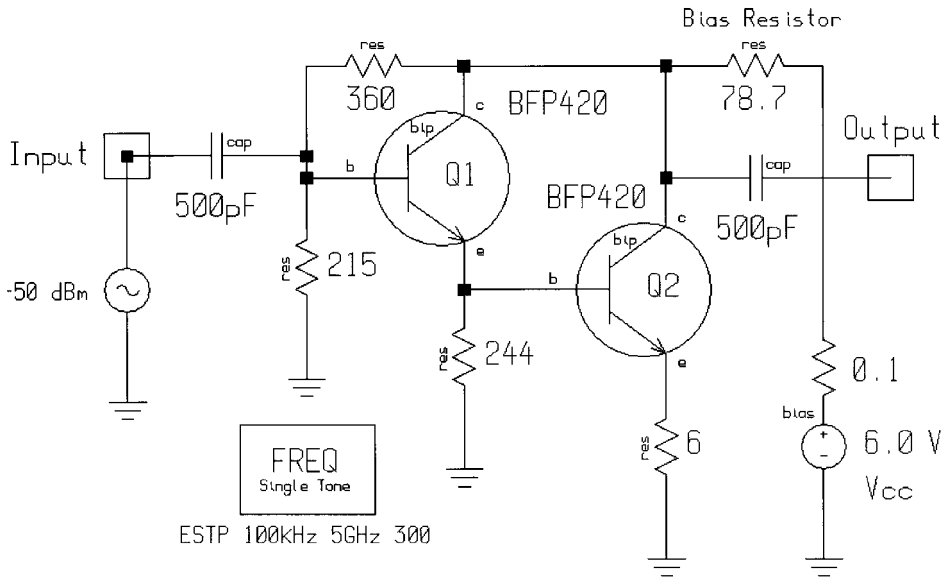


Figure 3-133 Schematic of the MSA-0375 MMIC amplifier.

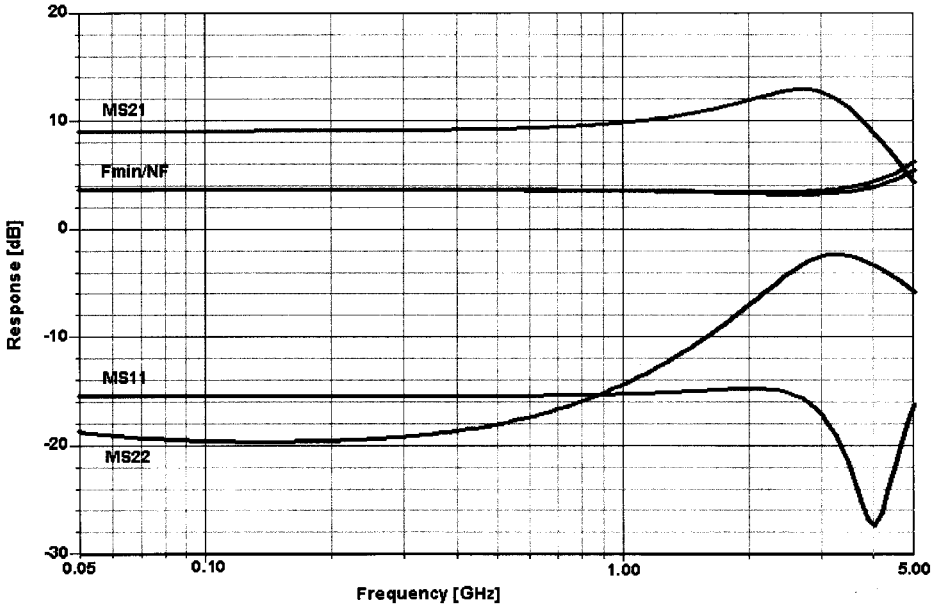


Figure 3-134 Frequency-dependent gain, matching, and noise performance of the MSA-0735 MMIC amplifier.

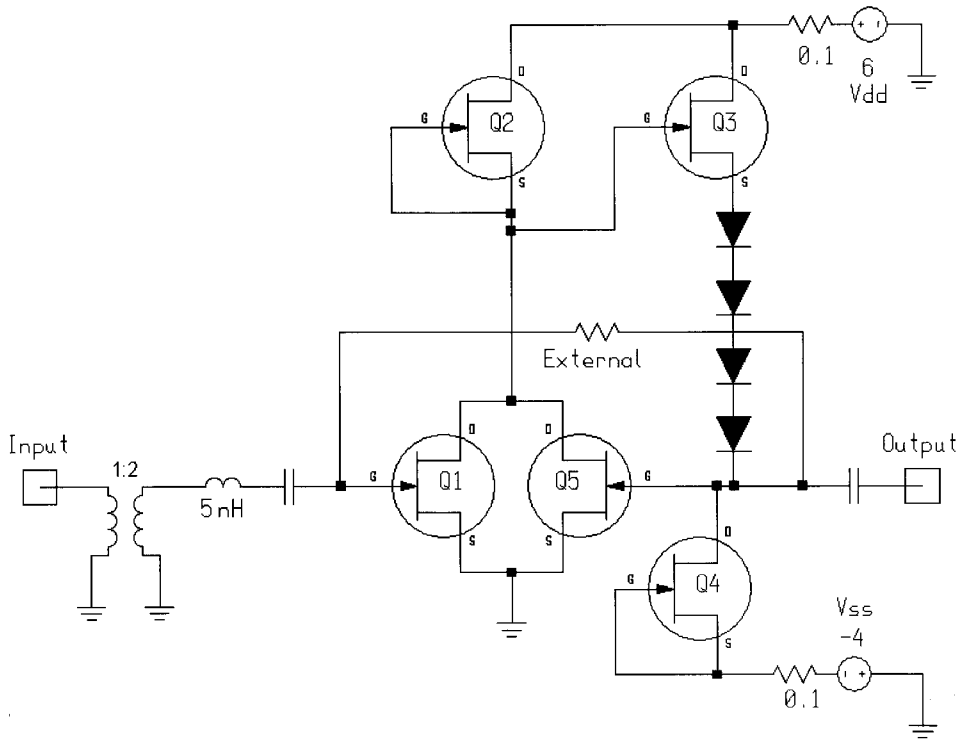


Figure 3-135 Schematic of the dc-coupled GaAsFET amplifier.

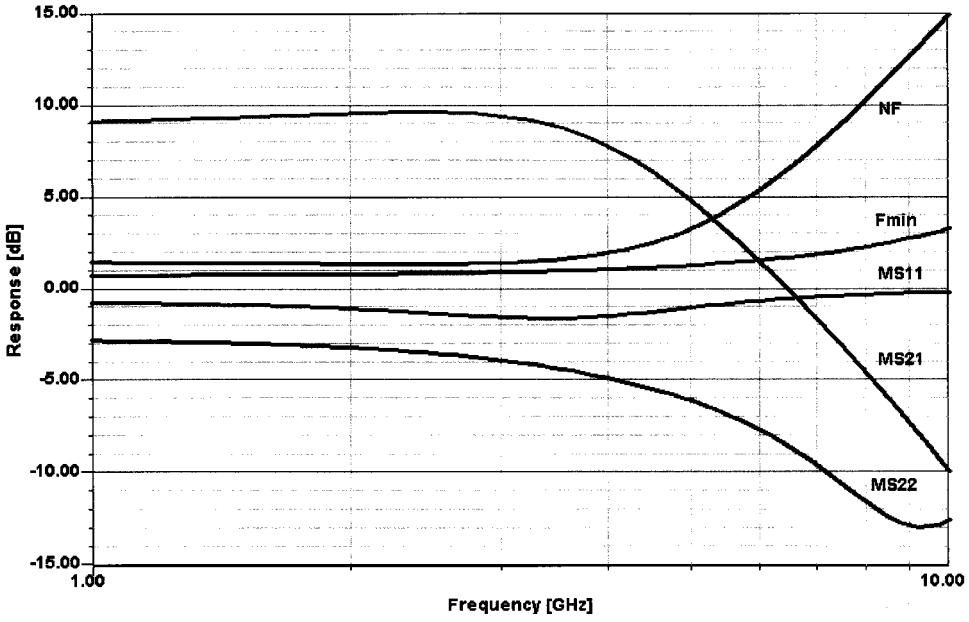


Figure 3-136 Frequency-dependent gain, matching, and noise performance of the dc-coupled GaAsFET amplifier.

from 0.15 to 0.30. In simulation, sizing the devices can best be accomplished by using the scaling factor for the FET model.

One of our questions was: "What would the perfect input termination for Q1 be?" By using an ideal transformer and playing with its turns ratio (needless to say, we did this with CAD), we determined experimentally that, relative to the gate, 200 Ω and a series inductor of 5–8 nH—the sort of details that designers often avoid telling their audiences—gave the ideal frequency response. Overall, we believe that the inventor of this circuit can be quite proud of the overall performance. The FETs we chose were taken from the Texas Instruments nonlinear foundry library; however, any similar device from another company would have worked equally well. The feedback loop used in this IC is an example of an approach frequently referred to as *active feedback*. The performance of this amplifier can be evaluated from Figure 3-136, which shows the frequency-dependent gain, matching, and noise figures (even including F_{min}).

3-5 AMPLIFIERS WITH THREE OR MORE STAGES

Most three-stage amplifiers are obtained by adding another buffer or power-gain stage to the previously described principal designs. As an example, we decided to evaluate the recent NEC UPC2749 IC, a silicon MMIC intended for application as a low-noise 1900-MHz amplifier operating at 3 V. Table 3-12 briefly summarizes its electrical specifications.

Table 3-12 NEC UPC2749 electrical characteristics

(T_A = 25 °C, $Z_L = Z_S$ = 50 Ω, V_{CC} = 3.0 V)			
	Part Number Package Outline		UPC2749T TO6
Symbols	Parameters and Conditions	Units	Typical
I_{CC}	Circuit current (no signal)	mA	6
G_S	Small-signal gain, f = 900 MHz dB	14.5	
	f = 1900 MHz	dB	16
f_{corner}	–3-dB Gain corner frequency	GHz	2.9
P_{1dB}	1 dB Compressed output power at 1900 MHz	dBm	–12.5
P_{SAT}	Saturated output power, f = 1900 MHz	dBm	–6
NF	Noise figure, f = 900 MHz	dB	3.2
	f = 1900 MHz	dB	4.0
RL_{IN}	Input return loss, f = 1900 MHz	dB	10
RL_{OUT}	Output return loss, f = 1900 MHz	dB	12.5
ISOL	Isolation, f = 1900 MHz	dB	30
OIP_3	SSB output third-order intercept, f_1 = 500 MHz, f_2 = 510 MHz	dBm	–3
	f_1 = 1000 MHz, f_2 = 1010 MHz	dBm	–3
	f_1 = 1900 MHz, f_2 = 1902 MHz	dBm	–3.5
	f_1 = 2000 MHz, f_2 = 2010 MHz	dBm	–4
$R_{TH(J-A)}$	Thermal resistance (junction to ambient)		
	Free air	°C/W	
	Mounted on a 50 × 50 × 1.6 mm epoxy glass PWB	°C/W	

In modeling the UPC2749's three transistors, we chose to use our favorite, the BFP420, which is not without its headaches at the higher frequencies. Figure 3-137 shows the circuit diagram, which closely resembles the one published by NEC; again, the only hard facts we had were the IC's basic circuit topology, gain, and dc current. Given those numbers, we staggered the current in the three devices by setting them at 0.64 mA for the first stage, 1.85 mA for the second stage, and 3 mA for the third stage. It should be noted that the two output transistors are not unlike the approach used for the Hewlett-Packard MSA-0735 (Figure 3-133). Taking (hopefully) all the right assumptions, we end up with a design that closely resembles the manufacturer's specifications for both power consumption, gain, and noise. This can be verified by looking at the results plotted in Figure 3-138.

The final example is a more complex derivative of the Philips NE/SA5204A amplifier, which was intended for standard 6-V operation and has a fairly "normal" circuit arrangement. This can be seen from Figure 3-139. A much more elaborate cousin of this is the low-voltage NEC UPC2710 amplifier. According to its specifications, it operates at 5 V but has significantly more gain than the NE/SA5204. Table 3-13 shows its electrical specifications and Figure 3-140 shows its measured gain as published by the manufacturer.

The actual interior of the circuit is shown in Figure 3-141. It consists of the standard preamplifier and then a Darlington configuration with the collectors tied together. The middle section of the circuit acts as a power supply responsible for the bias relative to the emitter of the first transistor influencing the rest of the circuit as well. Since NEC mentioned that they use their NESAT III process (f_T of 20 GHz at V_{ce} = 3) in producing the UPC2710T and we

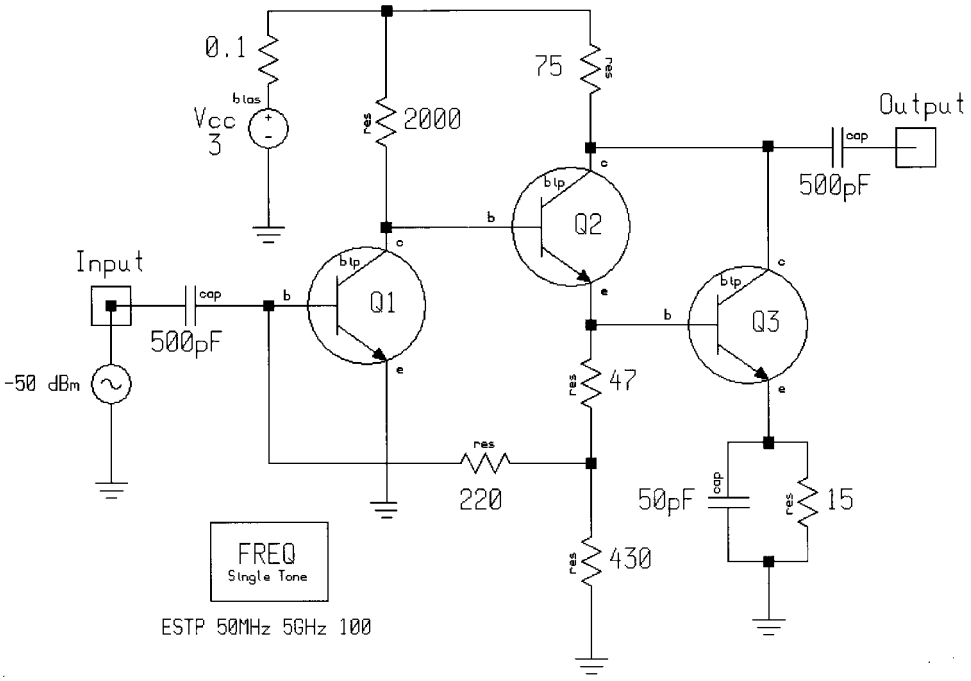


Figure 3-137 Circuit of the UPC2749 MMIC. The component values were not supplied by the manufacturer.

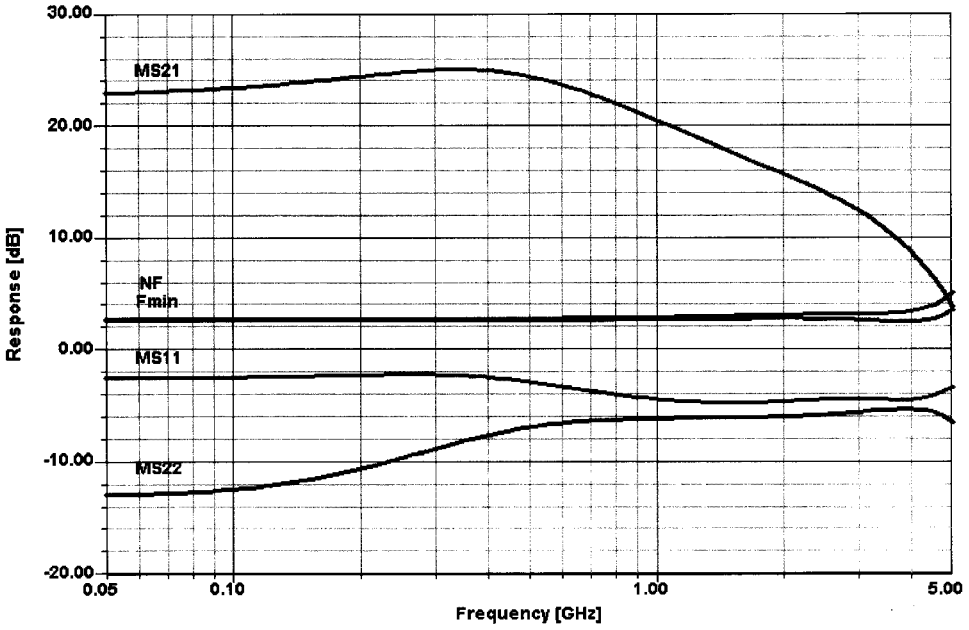


Figure 3-138 Simulated gain, matching, and noise performance of the NEC UP2749 MMIC.

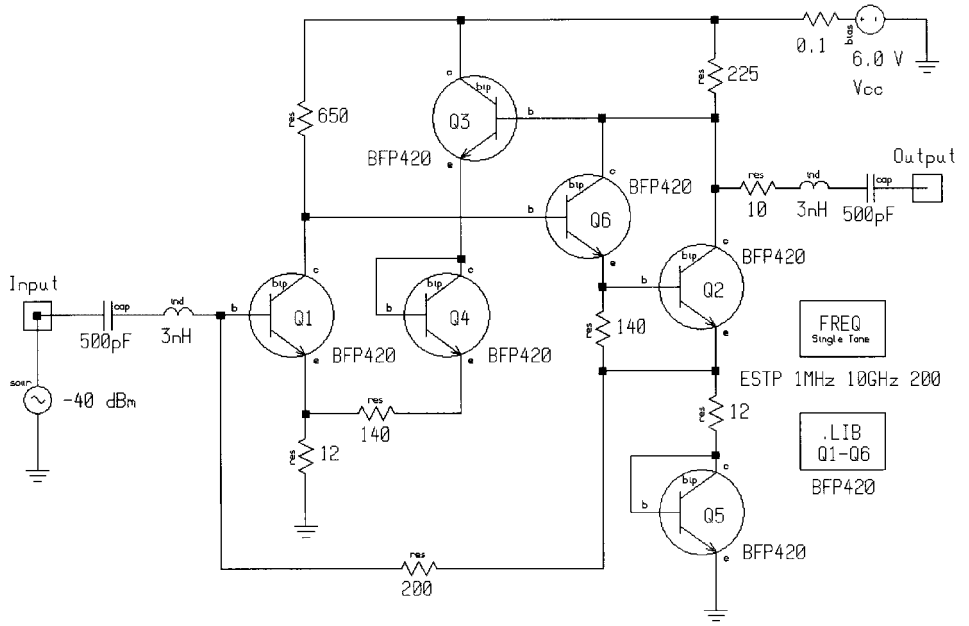


Figure 3-139 Schematic of the Philips NE/SA5204A amplifier IC entered for simulation.

Table 3-13 UPC2710 electrical specifications

	T_A = 25 °C, f = 0.5 GHz, V_{CC} = 5 V)		
	Part Number Package Outline		UPC2710T TO6
Symbols	Parameters and Conditions	Units	Typical
I_{CC}	Circuit current (no signal)	mA	22
G_S	Small-signal gain	dB	33
f_{corner}	–3-dB Gain corner frequency	GHz	1.0
ΔG_S	Gain flatness, f = 0.1–0.6 GHz f = 0.1–0.8 GHz	dB	±0.8
P_{SAT}	Saturated output power	dBm	13.5
P_{1dB}	Output power at 1dB compression point	dBm	7.5
NF	Noise figure	dB	3.5
RL_{IN}	Input return loss	dB	6
RL_{OUT}	Output return loss	dB	12
ISOL	Isolation	dB	39
ΔG_T	Gain–temperature coefficient	dB/°C	–0.006
R_{TH}	Thermal resistance (junction to ambient)	°C/W	

obtained some information regarding it, we decided to use our standard BFP420 (f_T of 25 GHz; see the datasheet in Chapter 2) for the purpose of simulation.

In our effort to simulate this circuit, we chose a combination of values that pretty much complies with the manufacturer's published specifications, but S_{22} is not close, as shown in Figure 3-142. We stopped at this point and invite our readers with CAD capabilities to provide the final touch in meeting all the published specifications for the circuit. We are particularly

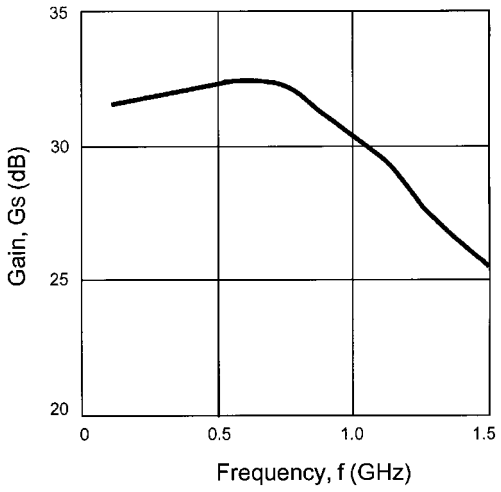


Figure 3-140 Manufacturer-supplied gain versus frequency for the NEC UPC2710T MMIC.

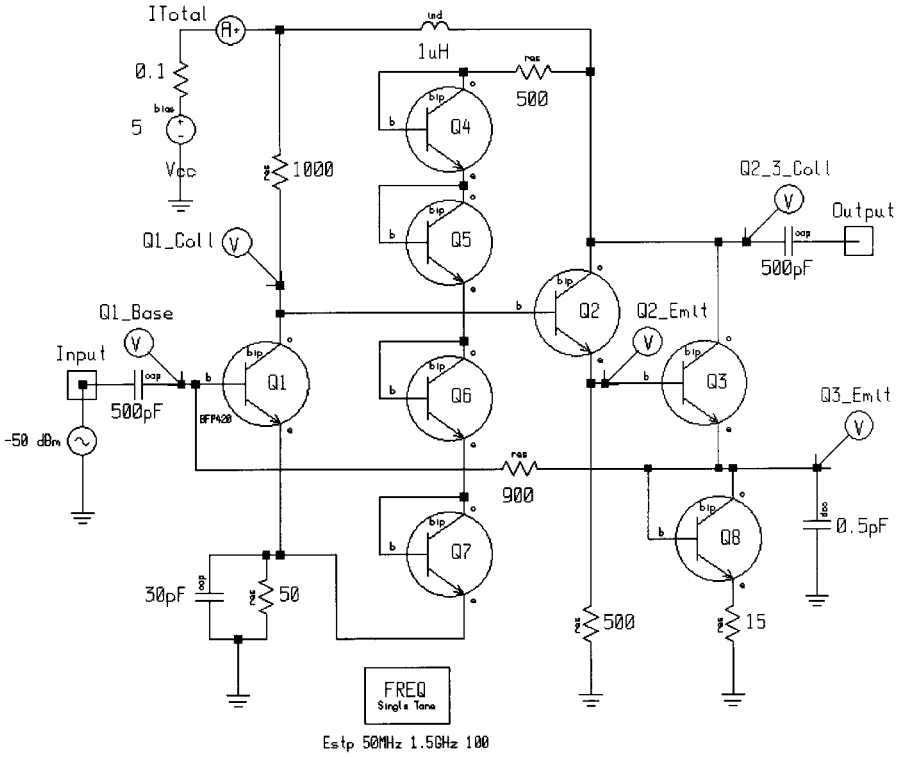


Figure 3-141 Schematic of the UPC2710T silicon MMIC entered for simulation.

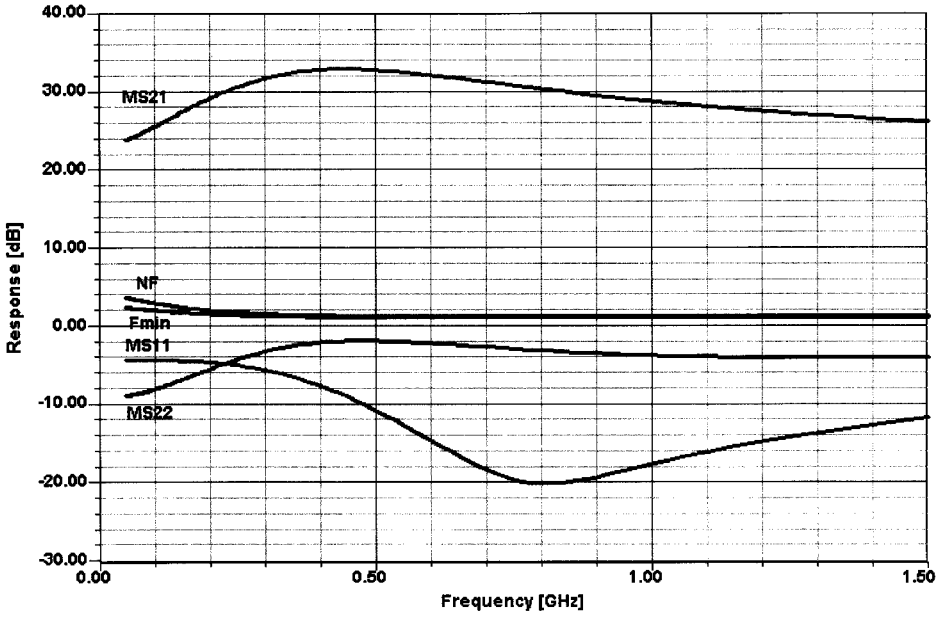


Figure 3-142 Simulated gain, match, and noise performance of the UPC2710T.

recommending such an effort because it shows very nicely the interaction of all the circuit elements in a fairly drastic way. Because of this, we admire the IC design even more, since the manufacturer has to sell a pretty consistent amplifier as a function of different production runs.

3-5-1 Stability of Multistage Amplifiers

Stability, needless to say, is a big issue, and the delay of the various stages adds to the phase shift, resulting in possible instabilities. The fact that S_{11} and/or S_{22} can become larger than 0 (in dB) means there is the potential of oscillating, depending on the reactances available. If a cable of inappropriate length is added, hell will break loose. Figure 3-143 shows the dangerous peaking of our simulation of the Philips NE5204A, in which the internal compensation was intentionally not assumed correctly. We will all agree that the high-end performance of this amplifier can generate high anxiety. As the production process of the aging device becomes more modern—meaning that transistors with much higher f_T will be the basis for the production—it will become quite difficult to maintain a previously achieved stability. This, by the way, applies also to discrete transistors, which for reasons of economics and multiple manufacturers will be improved over time. Such evolution results in generally smaller base-spreading resistances and capacitances, and other parameter changes, which, overall, cause headaches in manufacturing. We will leave the topic of three-stage amplifiers on this note.

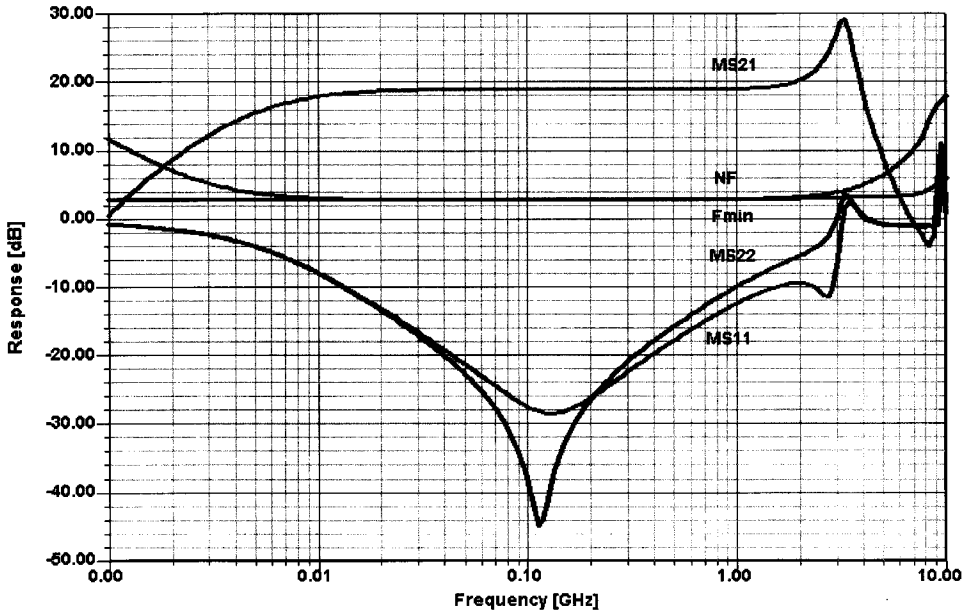


Figure 3-143 Simulated gain, match, and noise performance of the NE5204A IC.

3-6 A NOVEL APPROACH TO VOLTAGE-CONTROLLED TUNED FILTERS INCLUDING CAD VALIDATION [28]

Modern receivers control input stages as well as the oscillator band and frequency by electrical rather than mechanical means. Tuning is accomplished by voltage-sensitive capacitors (varactor diodes), and band switching by diodes with low forward conductance. Since the wireless band (essentially 400 MHz to 2.4 GHz) is so full of strong signals, the use of a tracking filter is desired as a solution to improve the performance and prevent second-order IMD products or other undesired overload effects. The dc control voltage needed for the filter can easily be derived from the VCO control voltage. There may be a small dc offset, depending on the IF used.

3-6-1 Diode Performance

The capacitance versus voltage curves of a varactor diode depend on the variation of the impurity density with the distance from the junction. When the distribution is constant, there is an "abrupt junction" and capacitance follows the law

$$C = \frac{K}{(V_d + V)^{1/2}} \tag{3-183}$$

where V_d is the contact potential of the diode and V is applied voltage.

Such a junction is well approximated by an alloyed junction diode. Other impurity distribution profiles give rise to other variations, and the above equation is usually modified to

$$C = \frac{K}{(V_d + V)^n} \tag{3-184}$$

where n depends on the diffusion profile and $C_0 = K/V_d^n$.

A so-called graded junction, having a linear decrease in impurity density with the distance from the junction, has a value of n. This is approximated in a diffused junction.

In all cases these are theoretical equations, and limitations on the control of the impurity pattern can result in a curve that does not have such a simple expression. In this case, the coefficient n is thought of as varying with voltage. If the impurity density increases away from the junction, a value of n higher than 0.5 can be obtained. Such junctions are called hyperabrupt. A typical n value for a hyperabrupt junction is about 0.75. Such capacitors are used primarily to achieve a large tuning range for a given voltage change. Figure 3-144 shows the capacitance–voltage variation for the abrupt and graded junctions as well as for a particular hyperabrupt-junction diode. Varactor diodes are available from a number of manufacturers, such as Motorola, Siemens, and Philips. Maximum values range from a few to several hundred picofarads, and useful capacitance ratios range from about 5 to 15.

Figure 3-145 shows three typical circuits that are used with varactor tuning diodes. In all cases, the voltage is applied through a large resistor R_e or, better, an RF choke in series with a small resistor. The resistance is shunted across the lower diode and may be converted to a shunt load resistor across the inductance to estimate Q. The diode also has losses that may result in lowering the circuit Q at high capacitance, when the frequency is sufficiently high. This must be considered in the circuit design.

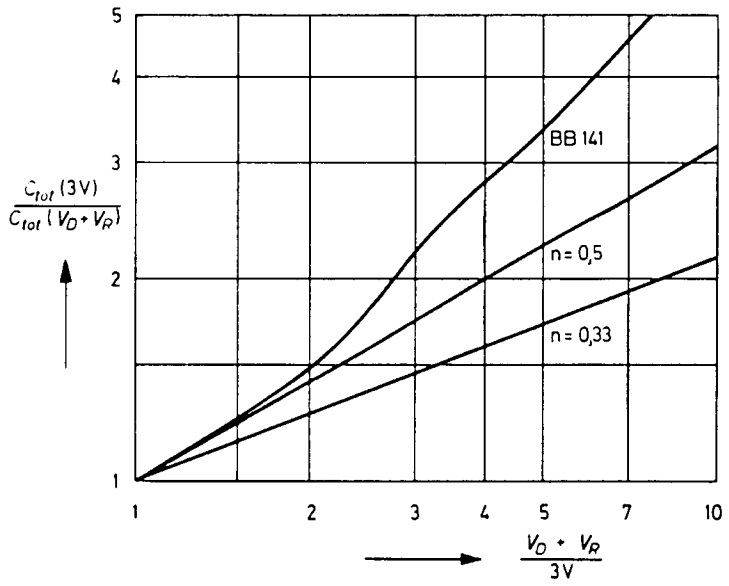


Figure 3-144 Voltage-dependent change of capacitance of different types of diodes. The BB141 is a hyperabrupt-junction diode with $n = 0.75$.

The frequency-dependent performance is not determined solely by applying the dc tuning voltage to Eq. (3-184). If the RF voltage is sufficient to drive the diode into conduction on peaks, an average current will flow in the circuits of Figure 3-145, which will increase the bias voltage. The current is impulsive, giving rise to various harmonics of the circuit. Even in the absence of conduction, Eq. (3-184) deals only with the small-signal capacitance. When the RF voltage varies over a relatively large range, the capacitance changes. In this case, Eq. (3-184) must be changed to

$$\frac{dQ}{dV} = \frac{K}{(V + V_d)^n} \tag{3-185}$$

Here Q is the charge on the capacitor. When this relation is substituted in the circuit differential equation, it produces a nonlinear differential equation, dependent on the parameter n. Thus, the varactor my generate direct current and harmonics of the fundamental frequency. Unless the diodes are driven into conduction at some point in the cycle, the direct current must remain zero.

The current of Figure 3-145c can be shown to eliminate the even harmonics, and it permits a substantially larger RF voltage without conduction than either circuit in Figure 3-145a or 3-145b. When $n = 0.5$, only the second harmonic is generated by the capacitor, and this can be eliminated by the back-to-back connection of the diode pair. Integrating Eq. (3-185), we find

$$Q + Q_A = \frac{K}{1 - n}(V + V_d)^{1-n}$$

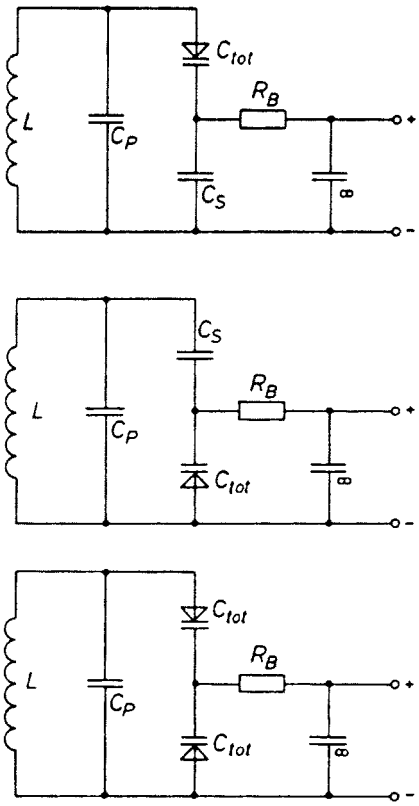


Figure 3-145 Various configurations to apply tuning diodes in a tuned circuit. The bottom version shows the lowest distortion.

$$= \frac{C_v}{1-n}(V + V_d) \tag{3-186}$$

C_v is the value of Eq. (3-184) for applied voltage V, and Q_A is a constant of integration. By letting $V = V_1 + v$ and $Q = Q_1 + q$, where the lowercase letters represent the varying RF and the uppercase letters indicate the values of bias when RF is absent, it follows that

$$q + Q_1 + Q_A = \frac{K}{1-n}[v + (V_1 + V_d)]^{1-n} \tag{3-187}$$

$$1 + \frac{v}{V'} = \left(1 + \frac{q}{Q'}\right)^{1/(1-n)} \tag{3-188}$$

where $V' = V_1 + V_d$ and $Q' = Q_1 + Q_A$. For the back-to-back connection of identical diodes, $K_{11} = K_{12} = K_1$, $V_1' = V_2' = V'$, $Q_1' = Q_2' = Q'$, $q = q_1 = -q_2$, and $v = v_1 - v_2$. Here the new subscripts 1 and 2 refer to the top and bottom diodes, respectively; v is the RF voltage across

the pair in series, and q is the charge transferred through the pair in series. This notation results in

$$\frac{v}{V'} \equiv \frac{v_1 - v_2}{V'} = \left(1 - \frac{q}{Q'}\right)^{1/(1-n)} - \left(1 - \frac{q}{Q'}\right)^{1/(1-n)} \tag{3-189}$$

For all n, this eliminates the even powers of q, and hence the even harmonics. This can be shown by expanding Eq. (3-189) in a series expansion and performing term by term combination of the equal powers of q. In the particular case $n = \frac{1}{2}$, $v/V' = 4q/Q'$, and the circuit becomes linear.

The equations hold as long as the absolute value of v_1/V' is less than unity, so that there is no conduction. At the point of conduction, the total value of v/V' may be calculated by noting that when $v_1/V' = 1$, $q/Q' = -1$, so $q_2/Q' = 1$, $v_2/V = 3$, and $v/V = -4$. The single-diode circuits conduct at $v/V' = -1$, so the peak RF voltage should not exceed this. The back-to-back configuration can provide a fourfold increase in RF voltage handling over the single diode. For all values of n, the back-to-back configuration allows an increase in the peak-to-peak voltage without conduction. For some hyperabrupt values of n, such that $1/(1 - n)$ is an integer, many of the higher-order odd harmonics are eliminated, although only $n = \frac{1}{2}$ provides elimination of the third harmonic. For example, $n = \frac{2}{3}$ results in $1/(1 - n) = 3$. The fifth harmonic and higher odd harmonics are eliminated, and the peak-to-peak RF without conduction is increased eightfold; for $n = \frac{3}{4}$ the harmonics 7 and above are eliminated, and the RF peak is increased 16 times. It must be noted in these cases that the RF peak at the fundamental may not increase so much, since the RF voltage includes the harmonic voltages.

Since the equations are only approximate, not all harmonics are eliminated, and the RF voltage at conduction, for the back-to-back circuit, may be different from that predicted. For example, abrupt-junction diodes tend to have n of about 0.46–0.48 rather than exactly 0.5. Hyperabrupt junctions tend to have substantial changes in n with voltage. The diode illustrated in Figure 3-144 shows a variation from about 0.6 at low bias to about 0.9 at higher voltages, with wiggles from 0.67 to 1.1 in the midrange. The value of V_d for varactor diodes tends to be in the vicinity of 0.7 V.

3-6-2 A VHF Example

The application of tuning diodes in double-tuned circuits in TV tuners has been common practice for many years. Figure 3-146 shows the circuit diagram. The input impedance of 50 Ω gets transformed up to 10 kΩ. The tuned circuits consist of the 0.3-μH inductor and two sets of antiparallel diodes. By dividing the RF current in the tuned circuit and using several diodes instead of just one pair, intermodulation distortion is reduced.

The coupling between the two tuned circuits is tuned via the 6-nH inductor that is common to both circuits. This type of inductance is usually printed on the circuit board. The diode parameters used for this application were equivalent to the Siemens BB515 diode. The frequency response of this circuit is shown in Figure 3-147.

The coupling is less than critical. This results in an insertion loss of about 2 dB and relatively steep passband sides. Once the circuit's large-signal performance (Figure 3-148) is seen, a third-order intercept point of about –2 dBm is not so unexpected. The reason for this poor performance is the high impedance (high L/C ratio), which provides a large RF voltage swing across the diodes. A better approach appears to be using even more diodes and at the same time changing the impedance ratio (L/C ratio).

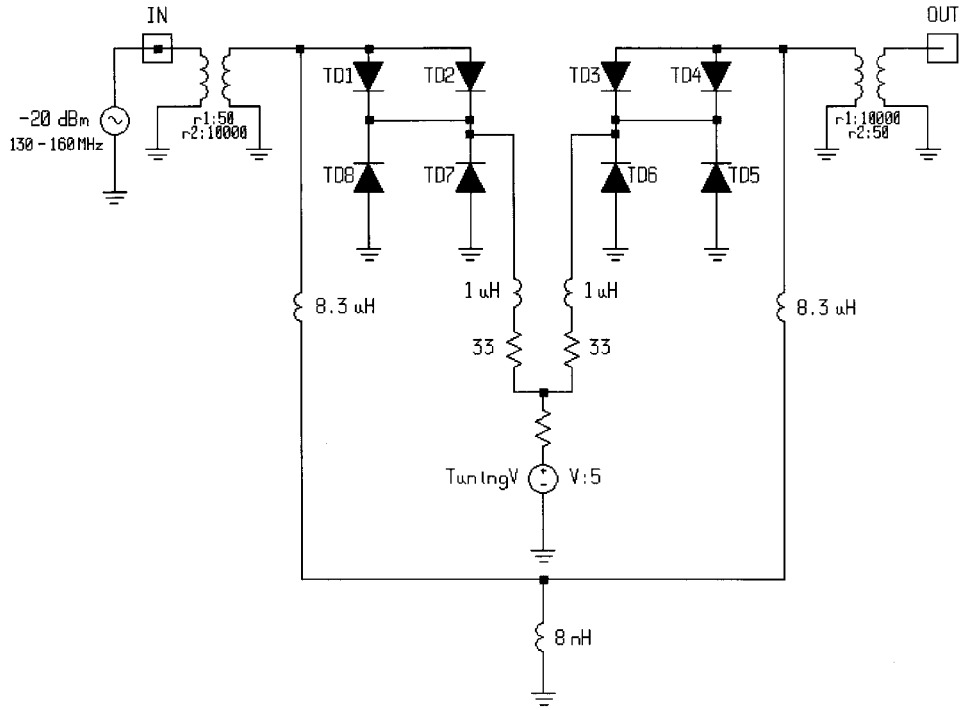


Figure 3-146 Double-tuned filter at 161 MHz using hyperabrupt-junction tuning diodes. By using several parallel diodes, the IMD performance improves.

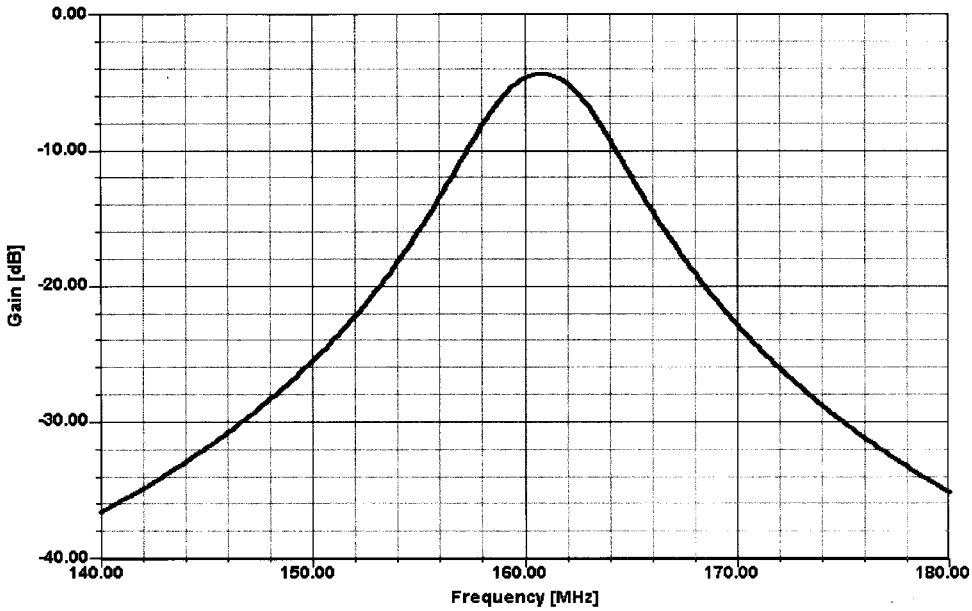


Figure 3-147 Frequency response of the tuned filter shown in Figure 3-146. The circuit is undercoupled (less than transitional coupling); $Q \times k < 1$.

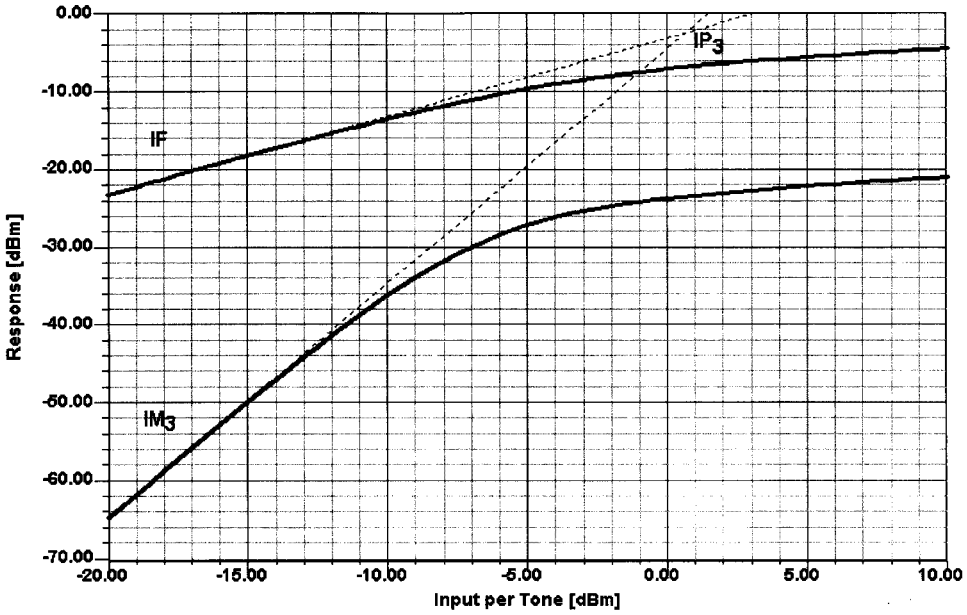


Figure 3-148 Prediction of intercept point of the double-tuned circuit shown in Figure 3-146. Note the compression of both the input signal and the IMD product.

3-6-3 An HF/VHF Voltage-Controlled Filter

The above example used a step-up procedure typically done by a tap at the input and output inductance. This method allows an impedance transformation; however, if one desires to change it into a series-tuned arrangement, it has to be done with a transformer. The large-signal conditions in the frequency range from 10 to 30 MHz on a medium- to large-sized antenna are equivalent to, if not worse than, the conditions for VHF operation. The only exception would be line of sight into the transmitter such as a tower in the middle of the city. Examples of such hostile conditions would be any large city, such as Munich, New York, Miami, Chicago, and San Francisco, where the authors had significant experience with intermodulation distortion problems.

By translating the circuit into a low-impedance-drive arrangement, building it symmetrically, and reducing the tuning range somewhat by capacitors and many series diode pairs, the filter's large-signal performance was significantly improved. The tuning diodes had about 125 pF at 1 Vdc per unit. Initial tests trying to use a high-capacitance diode, such as the BB112, resulted in much higher IMD products. The B112 has a 1-V capacitance of 470 pF down to about 20 pF at 8 V. Standard diodes show about 30 pF at 1 V, while the diodes used in the experiment had a 125-pF capacitance at 1 V. These experimental diodes were made available from a well-known manufacturer for this evaluation. Figure 3-149 shows the circuit diagram.

A step-down transformer drives the tuned circuit with a source impedance of 12.5 Ω. The circuit is symmetrical. There are two 0.8-μH inductors and 2×5 diodes in the loop. The output of the circuit is transformed back up to 50 Ω. The tuning voltage is supplied via a heavily filtered arrangement. This circuit has actually been implemented in Rohde & Schwarz field strength meter equipment. They guarantee an intercept point of about +20 dBm. The selectivity of these circuits is quite reasonable and mostly intended to reduce

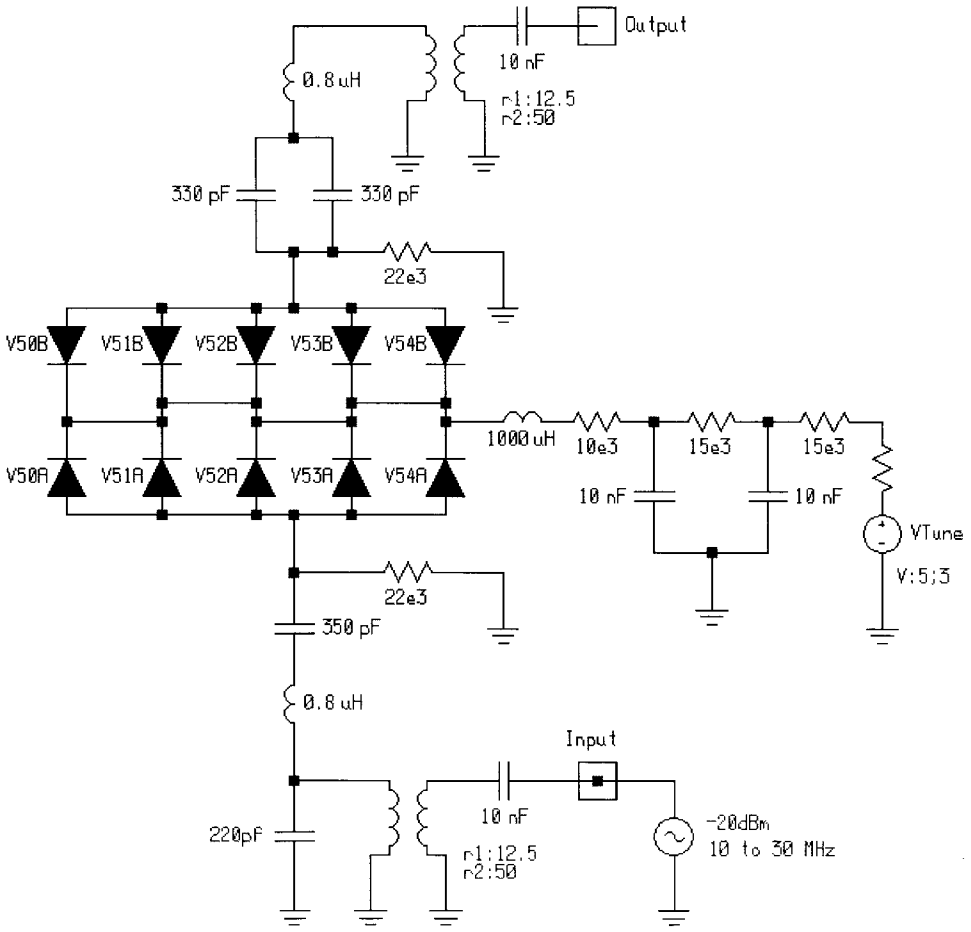


Figure 3-149 High dynamic range of a 10–20-MHz filter.

second-order intermodulation distortion products by about 10 dB, at the same time having a high third-order intercept point. Figure 3-150 shows the selectivity curves at two different tuning voltages.

The most interesting number, however, is the third-order intercept point (already determined as about 20 dBm), which still had to be simulated. This sometimes sounds like a contradiction: Once the measured values are available and are acceptable, why would one want—or, rather, need—to do a simulation besides the necessity to develop a high-input intercept filter? It was desirable to validate the nonlinear models and prove that the above-mentioned equations will hold true. This type of simulation is now more difficult because the number of nonlinear elements went from four to ten, and some numerical problems, such as convergence difficulties, can be expected, and the effect of harmonic frequency cancellation (compensation) can also be seen. By using diode combinations that result in $1(1 - n)$, n being an integer number, these IMD products can drastically be reduced. This implies that each of the five diode pairs has a selected value for n to meet this condition. Later, a final attempt will be made to improve the first VHF filter with the proper diode combinations.

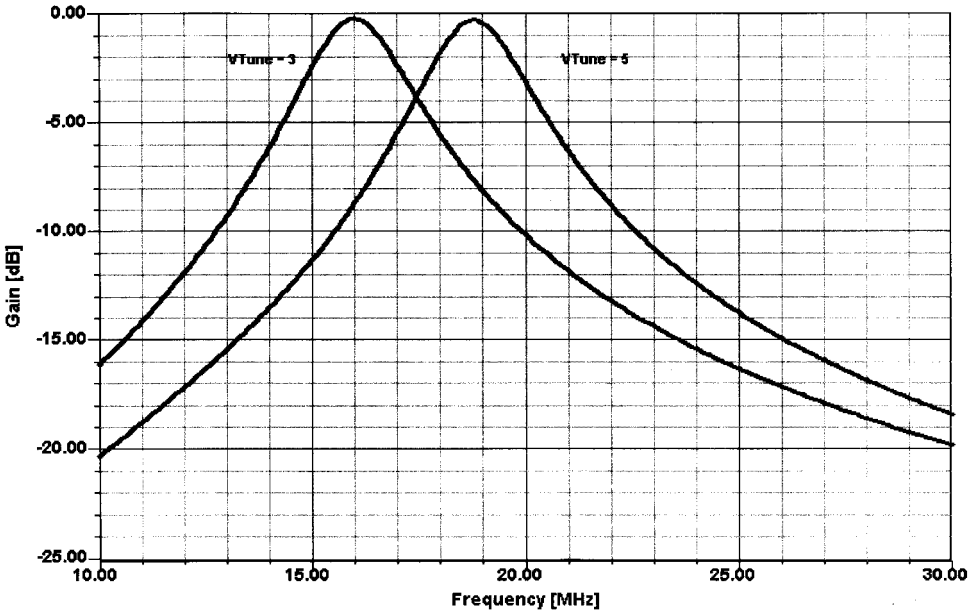


Figure 3-150 Frequency response of the filter shown in Figure 3-149. Filters of this type are intended to reduce second-order IMD by providing about 20-dB suppression at half the center frequency.

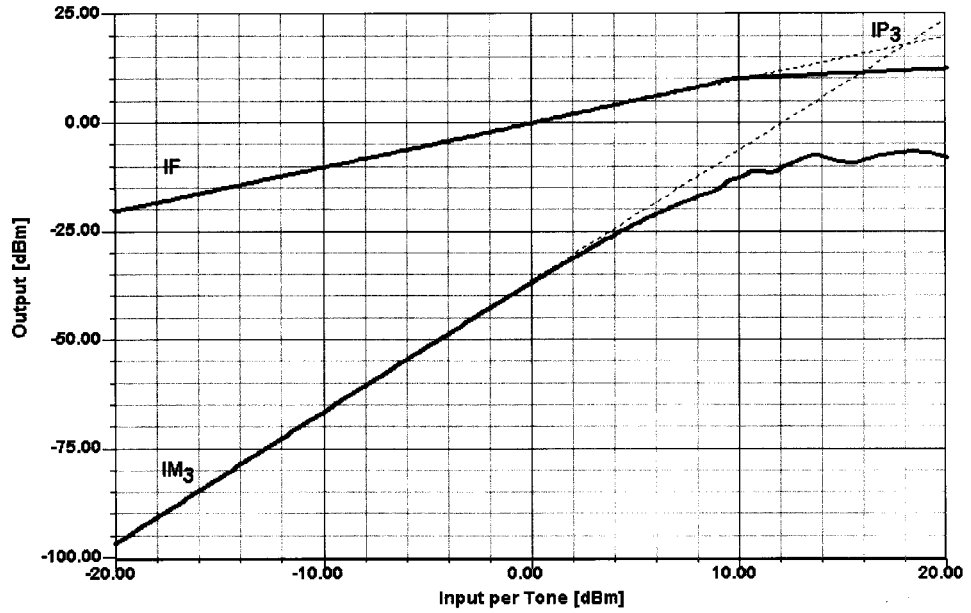


Figure 3-151 Prediction of third-order intercept point of the filter shown in Figure 3-149. The reason for the curves on the IMD plot lies in the interaction between the ten nonlinear devices.

Figure 3-151 shows the calculation of third-order intercept point for the arrangement shown in Figure 3-149. This number has now increased to +18 dBm. The difference between selection and measurement is approximately 2 dB. The reason for this more pessimistic value compared to the measured value probably has to do with slight variations of the exponent of the diodes' voltage-dependent capacitance. Therefore, the authors consider both the circuit performance and the simulation accuracy to be extremely good. This type of circuit, as mentioned earlier, has wide application in oscillator circuits. Danzeisen of Rohde & Schwarz was probably the first to have used this type of circuit by paralleling many diodes for improved performance.

3-6-4 Improving the VHF Filter

By selecting the appropriate diode combinations and circuit modifications as shown in Figure 3-152, a significant IMD improvement of the circuit shown in Figure 3-146 is obtained. A special shunt arrangement of several diodes with different exponents has been developed, which allows its value to be "adjusted." This circuit, for which a patent has been applied, showed an improvement from –2 dBm to 32 dBm. The same high operating Q was maintained. Figure 3-153 shows the frequency response after the modification. Note that a coupling slightly greater than critical ($Q \times k = 1.1$) has been selected.

3-6-5 Conclusion

After explaining some of the nonlinearities in mathematical terms, we have given some examples of voltage-tuned circuits and discussed their large-signal performance. This section has also shown that modern CAD tools can accurately predict the performance of such circuits.

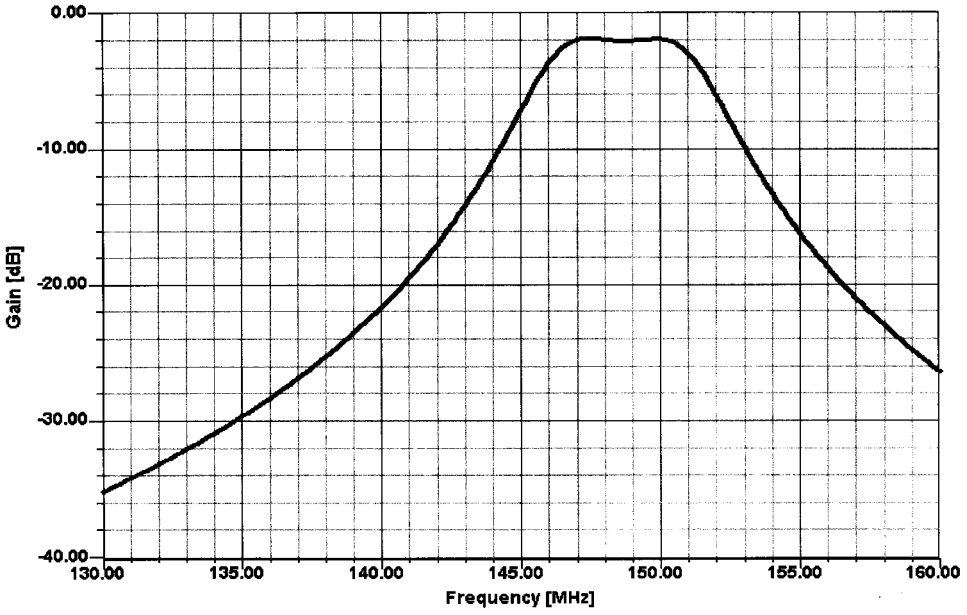


Figure 3-152 The frequency response of the improved circuit derived from Figure 3-146.

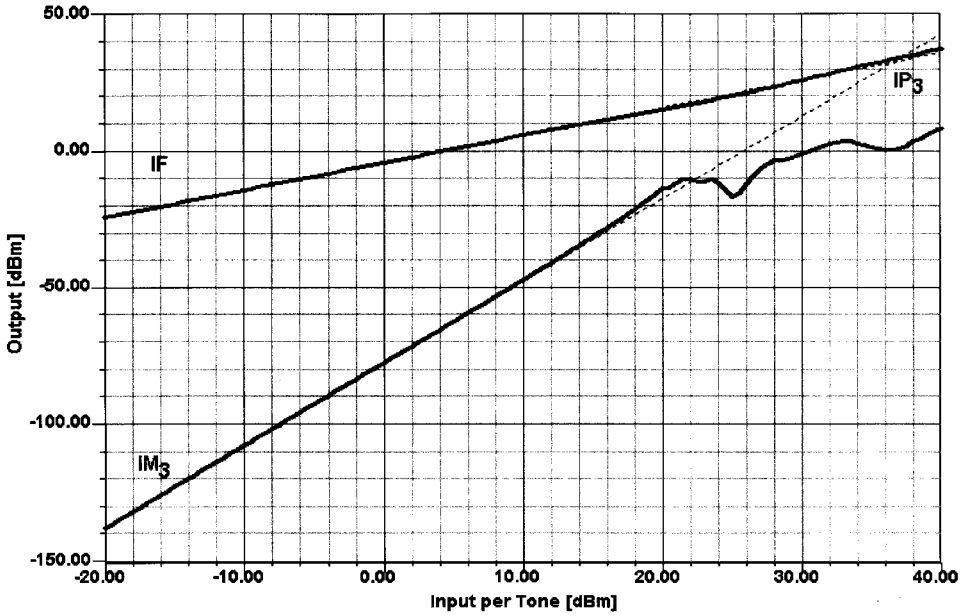


Figure 3-153 Predicted IMD performance of the improved version of the circuit shown in Figure 3-146. This type of improvement is significant for all applications. A patent is in the process of being obtained for this; therefore, the circuit cannot be disclosed at present.

3-7 DIFFERENTIAL AMPLIFIERS

The differential amplifier (Figure 3-154) goes back many years and also was the first step from TTL to ECL. The differential amplifier consists of an emitter follower (Q1) and a grounded-base stage (Q2). The constant-current generator (Q3) between ground and the emitters of Q1 and Q2 determines the current for these stages; the dc operating currents of Q1 and Q1 are the same. Most modern integrated circuits use this differential amplifier in one way or another.

If the circuit is operated about its point of symmetry, the dc component of the current through either Q1 or Q2 remains half of the emitter constant current through Q3 for all symmetrical driving signals. In addition, for input signals that are aperiodic, no even harmonics are generated. This assumes that the stage is driven symmetrically at both inputs. In our case, the differential amplifier has its second stage grounded, meaning that there is a base-to-ground capacitor, and the stage is asymmetrically fed. As the input voltage is increased, the differential stage becomes a limiter, which limits the current between plus and minus I_S, with I_S being the source current. The resulting output will become a square wave. This stage can also be used as a line receiver, translating a sine wave to an output logic suitable for driving emitter-coupled logic (ECL) stages.

One can assume that the differential amplifier has a transconductance of

$$g_m = \frac{\alpha g_{\text{in}}}{2} \tag{3-190}$$

where α is the ratio of collector to emitter current. Within the limits of the output being a sine wave, this stage has very low distortion. The output current now is determined by

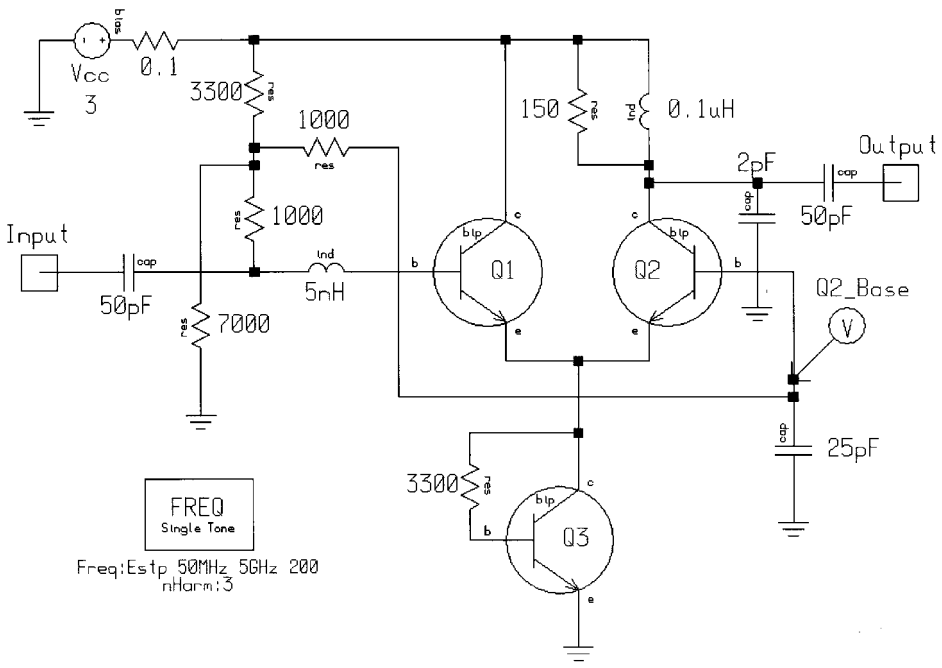


Figure 3-154 The differential amplifier schematic. The transistors are Siemens BFP420s. Figure 3-155 shows the circuit's frequency-dependent gain, matching, and noise performance.

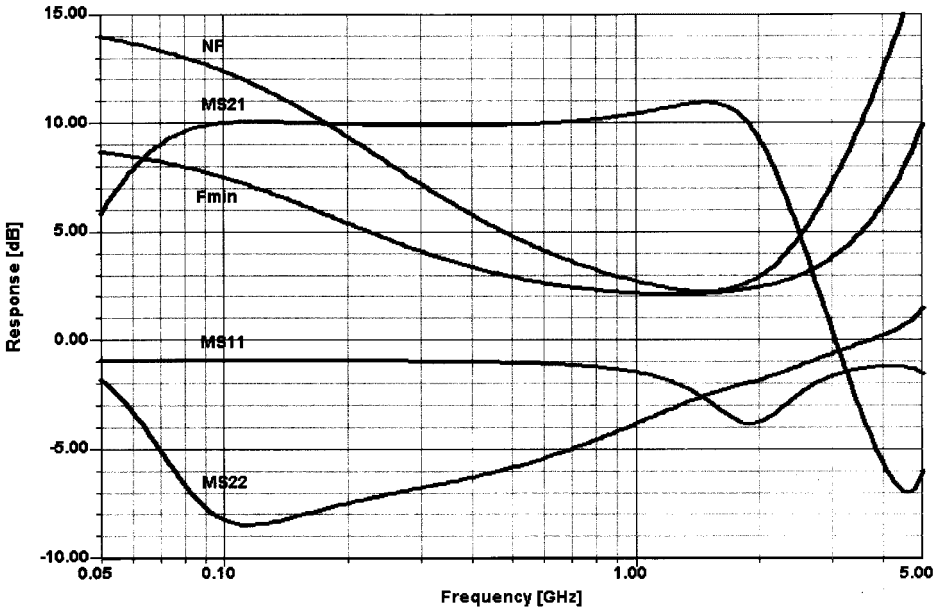


Figure 3-155 Frequency-dependent gain, matching, and noise performance of the differential amplifier.

$$i = \frac{I_{\text{in}}}{2} \tanh\left(\frac{x}{2} \cos \omega t\right) \tag{3-191}$$

where ωt is the input frequency and I_{in} is the current forced by Q3. The reason why the transconductance has to be divided by 2 comes from the fact that the current is split into two equal components for Q1 and Q2. Multiplying the new transconductance with the output load gives the voltage gain of this amplifier; its power gain can be obtained by multiplying this with $\sqrt{\text{output resistance} \div \text{load resistance}}$.

Having done the linear analysis and having learned that because of the hyperbolic tangent functions this amplifier can also be used as a line receiver, and converting sinusoidal waves to square waves, it may be useful to take a closer look at these issues. If we take a CAD oscilloscope, we see the resulting output waveform being a nonequal duty cycle output voltage, which results in high harmonic content. The amount of harmonic content can be controlled by the drive power and by the actual bias. Figure 3-156 shows the resulting output waveform.

If we now switch to a CAD spectrum analyzer, we can evaluate the harmonic content (Figure 3-157). Figure 3-158 shows the output transistor's dc I–V curves and ac load line under these conditions. Based on the previously defined conducting angle of the differential amplifier, one can optimize for harmonic content. Therefore, it is logical at this point to look at frequency multipliers, as they are sometimes needed. Finally, we are curious to see the so-called load line of the output transistor, which is, based on the interaction of all the stages, not looking the same way as the expected one, which we know from previous amplifier examples. Part of the reason for this is that for a certain voltage range, the collector voltage is in the negative region, and all kinds of saturation effects result in this surprising waveform. By reducing the input power, one would obtain a more "expected" load line. We invite readers with CAD capabilities to play with this example, as it gives a lot of insight into the operation of ECL stages.

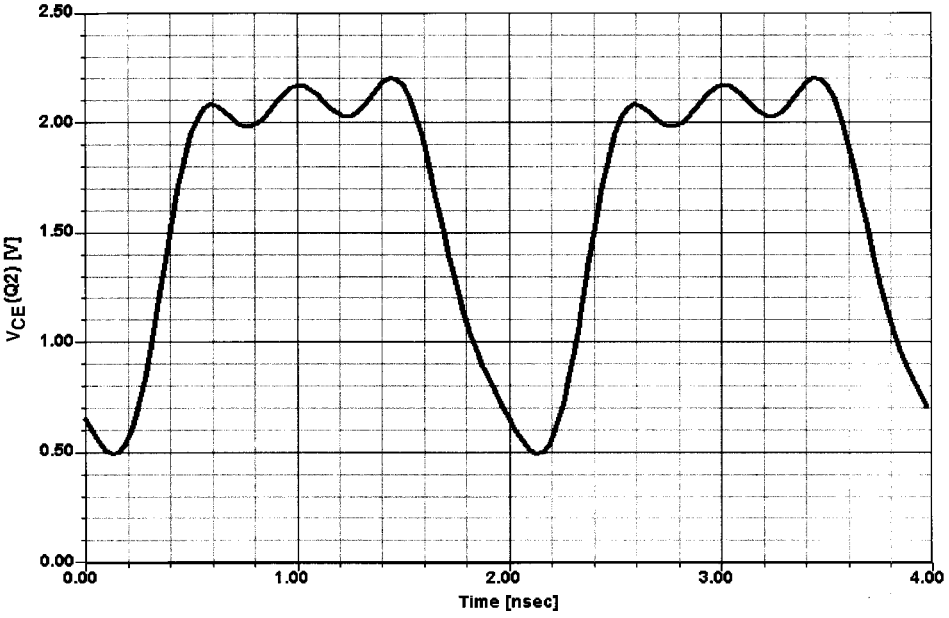


Figure 3-156 V_{CE} versus time for output transistor Q2 with a drive of 10 dBm.

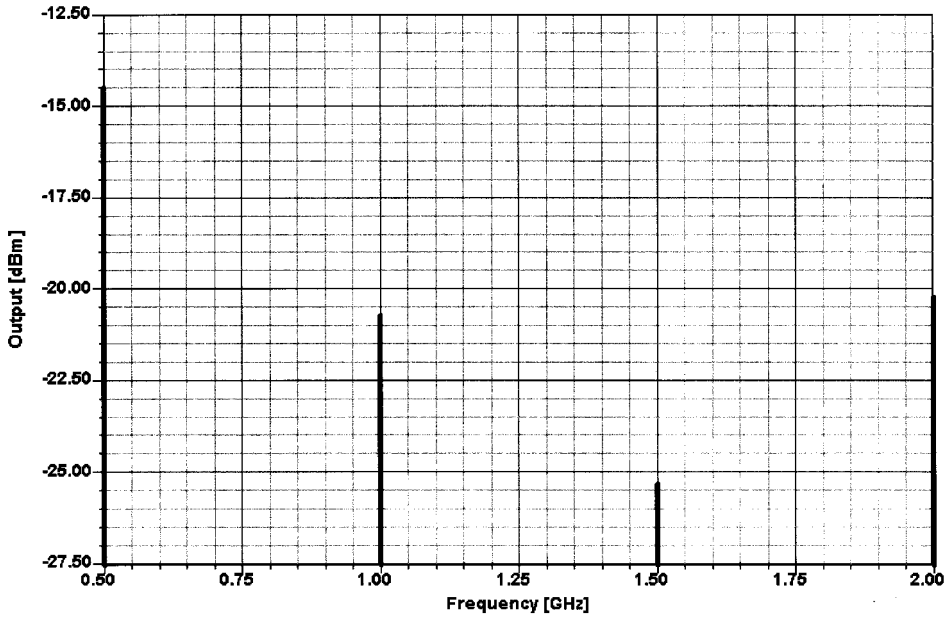


Figure 3-157 Output spectrum of the differential amplifier with a drive of 10 dBm. The biasing is not optimal for frequency multiplication.

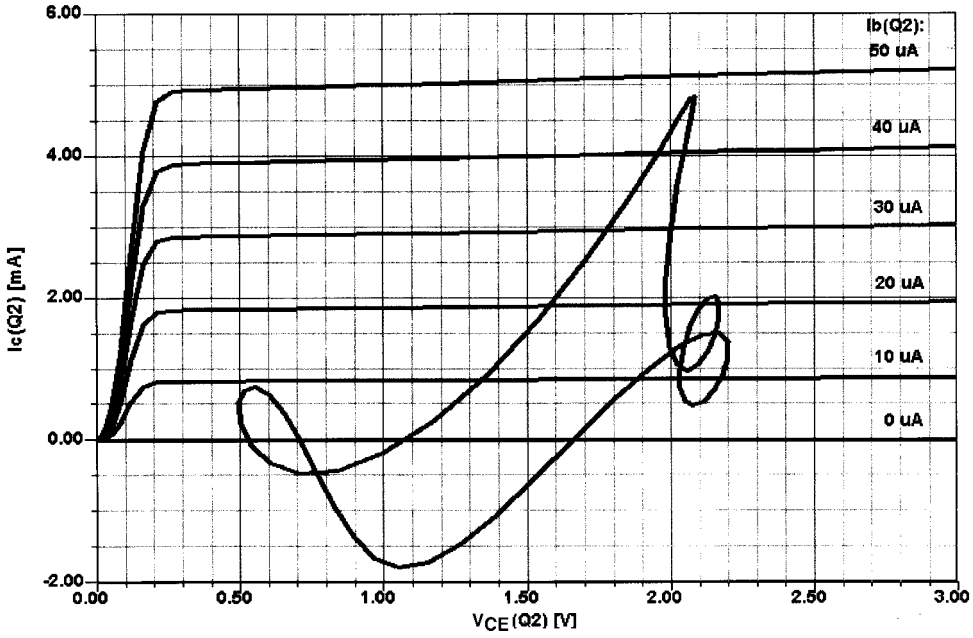


Figure 3-158 The dc *I–V* curves and ac load line for output transistor Q2 with a drive of 10 dBm.

3-8 FREQUENCY DOUBLERS

Having seen that the harmonic content invites us to bias a stage from Class A into Class B or C if our goal is a harmonically rich current, we will now evaluate a single stage that will be biased close to cutoff and driven by a relatively low frequency (500 MHz). To obtain the best results, we will use a parallel notch filter in the output, followed by a high-pass matching circuit to "catch" the desired harmonic. The input frequency is heavily suppressed by the notch filter. Based on its finite Q, there is a limited range over which this frequency can be varied with constant subharmonic attenuation. Figure 3-159 will be our test circuit.

This circuit consists of a bias transistor, Q2. This method has been discussed before and solves the temperature-dependent bias emitter thermal runaway problems through "temperature compensation." The purists among our readers may point out that the currents drawn by Q1 and Q2 are not quite identical (which would give 100% cancellation), but the approach shown is sufficient to do its job. Also, we have chosen a Class AB operation point, which is defined by Q1's 1-mA dc standing current. The conducting angle has not been optimized for the second harmonic—a task that the interested reader can do by using the conducting angle relationship shown at the beginning of this chapter. To give readers with access to nonlinear CAD some homework, we have not set this at the optimum point.

The second transistor, here called Q1 (because of its importance at RF), has a matching network at the input. This is recommended for optimum energy transfer. The matching condition can be validated by examining Figure 3-160. If we look at the higher frequency response, the input shows a possible trend of oscillation (negative loss); this could actually be eliminated by using a low-pass filter at the input instead of the high-pass filter used now. Findings of this type always make us nervous, and we hope that the nonlinear models are sufficiently accurate at higher frequencies to predict such behavior correctly. In the case of the FET doubler mentioned in Chapter 2, we found that the device modeling done by Compact (now Ansoft) gave the best results, the university answers were fairly close, and the device manufacturer data were worst. We have since become highly skeptical about device modeling unless we can validate it ourselves. Synergy Microwave has a complete set of test equipment in the laboratory, including network and spectrum analyzers with special software created by Rohde & Schwarz that can be used in addition to the Scout program to do in-house validation. This was achieved by the use of nonlinear optimization at a drive level of 0 dBm. One could say that this is similar to using a load-pull approach at the input, and then resorting to the large-signal S parameters obtained. Another note of caution: Capacitors at these low values tend to have a very large percentage error; for instance, a 1-pF capacitor, if not selected carefully, can vary between 0.7 and 1.3 pF—a 50% error. Needless to say, this is not acceptable for useful production.

The output consists of a high-pass filter and a notch filter with the purpose of reducing the subharmonic significantly. We have again done this on purpose, leaving room for the reader to come up with a more inventive output circuit like a Cauer-type (same as elliptical) low-pass filter with discrete notches at higher harmonics of the input frequency, such as 1.5 GHz, 2 GHz, and so on. In addition, we have shown three ways how to implement the inductance for the notch filter. This is equally applicable for other inductors in this value range. As frequency increases, this becomes more interesting but one needs to remember that many materials like silicon and gallium arsenide are extremely lossy. The same inductor approach has been used in Figure 5-61 of Chapter 5. For silicon, even the manufacturer becomes an issue. Figure 3-161 shows a top view of a "rectangular" inductor that actually approaches the spiral inductor; its implementation can be seen in Figure 3-162. Modeling this was a

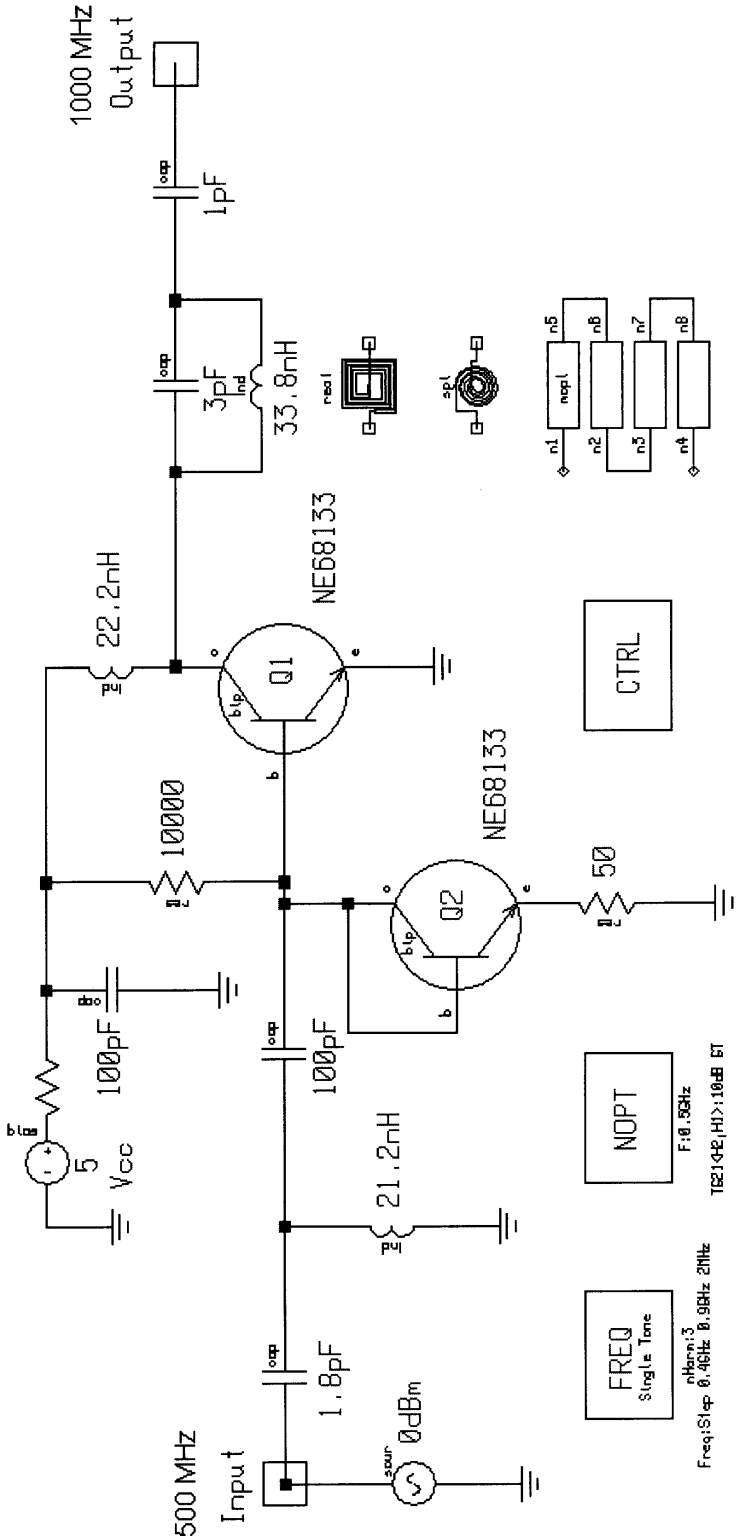


Figure 3-159 BJT frequency doubler schematic. A parallel-tuned trap in Q1's collector attenuates fundamental feedthrough.

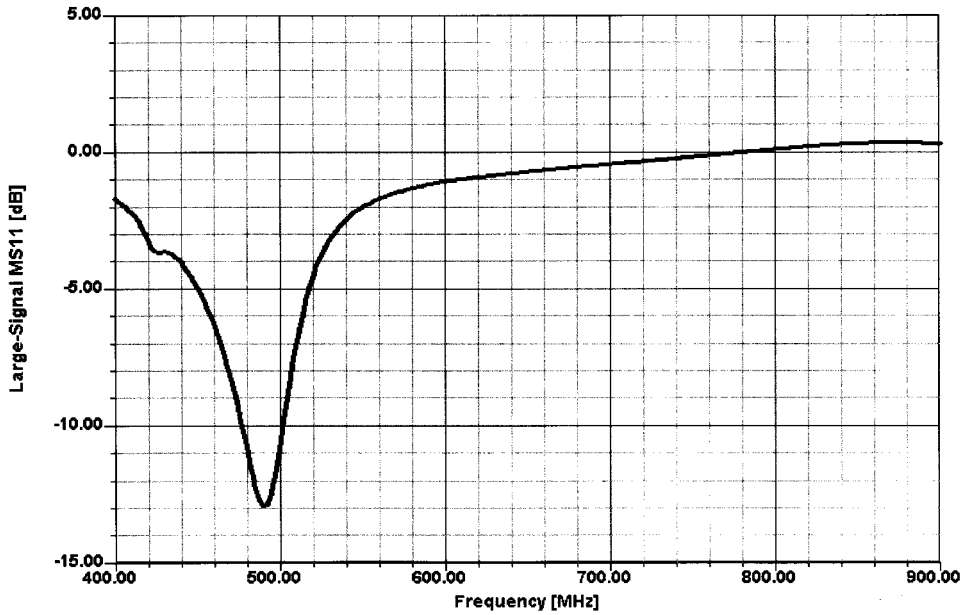


Figure 3-160 Large-signal MS_{11} for the BJT doubler. Note the positive region above 800 MHz.

nightmare, and during our Compact days, we actually had to modify the electromagnetic simulator to accommodate this and get results where measurements and simulation agreed. Figure 3-163 shows the measured and simulated Q as a function of frequency of a silicon-based inductor. In the case of other materials, such as PC board or GaAs, the models in the CAD tools are sufficient to predict the losses of the inductors. Following this initial exercise, Motorola then developed an improved silicon spiral inductor; its frequency-dependent Q is shown in Figure 3-164. Note that this Q was measured and not simulated.

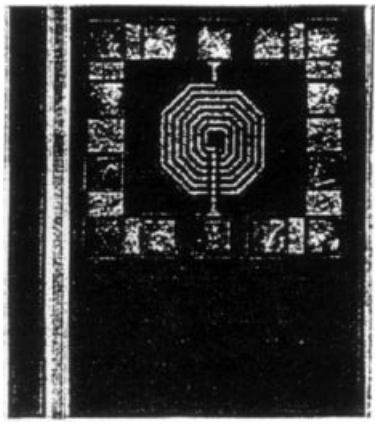


Figure 3-161 A 6.5-turn spiral inductor in silicon.

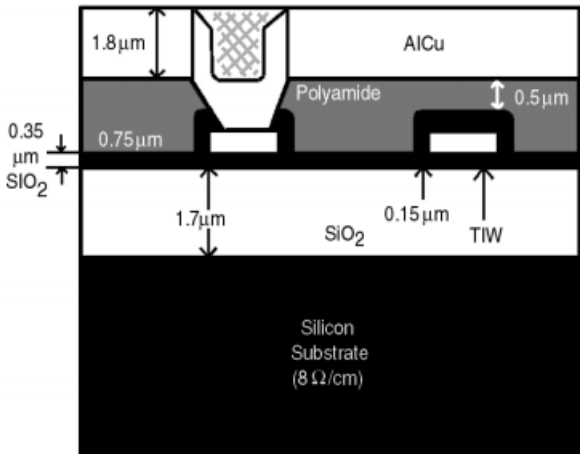


Figure 3-162 Cross section of the silicon inductor.

One of the reasons why we use a transistor multiplier is that we expect gain (because of the square-law characteristic, an FET would have been better tailored for a doubler); therefore, we have "measured" the gain of this circuit while driving it with 0 dBm. Another important factor is that, because of the way the CAD tool displays it, the gain is referenced to the 500-MHz input. Mentioning this is important; otherwise, the reader may look for the 1-GHz gain, which does not exist. Figure 3-165 shows the doubler gain.

By now we are curious to see the doubler in action, and by reviewing Figure 3-166 we can see the output spectrum. The output spectrum validates the statement that the output circuit, while doing a marvelous job of removing the 500-MHz input frequency, does not cure the problem of unwanted multiples of the input frequency. Again, a Cauer low-pass filter, while introducing some possible additional loss, should be considered for this circuit to have a bandpass response. On the other hand, many oscillators have very little harmonic suppression, and the same type of postfiltering needs to be considered. There are two-transistor solutions available also, such as push–push for frequency doubling (push–pull would

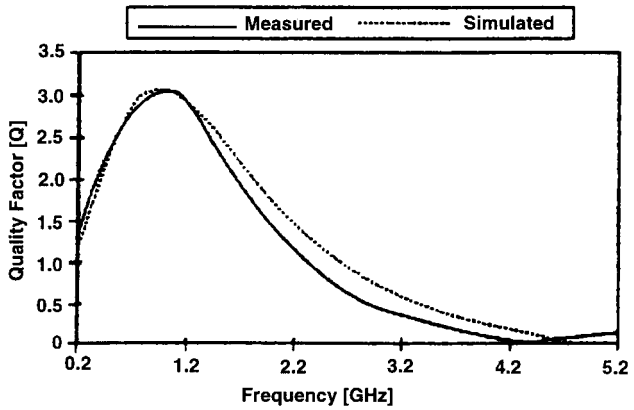


Figure 3-163 Measured and simulated *Q* versus frequency for the simulated inductor.

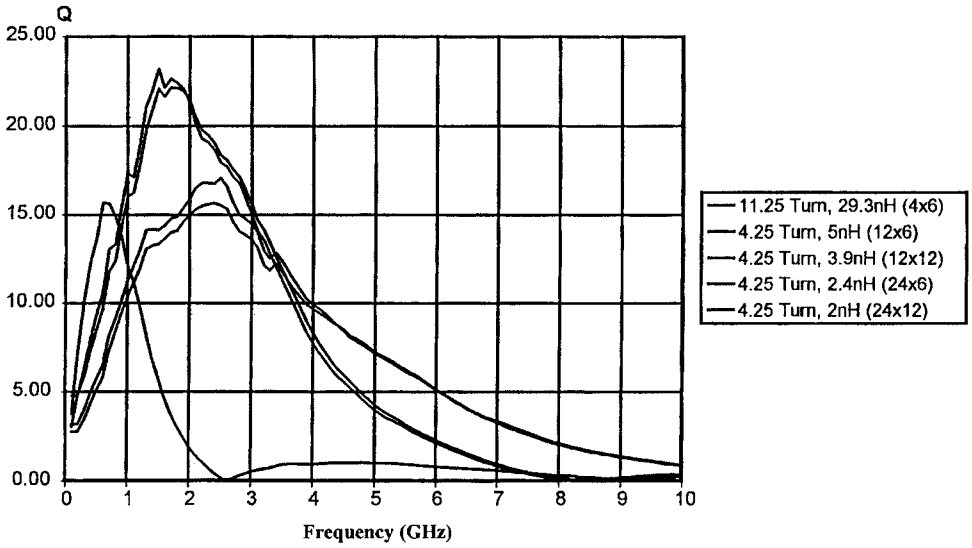


Figure 3-164 *Q* versus frequency for Motorola improved inductor.

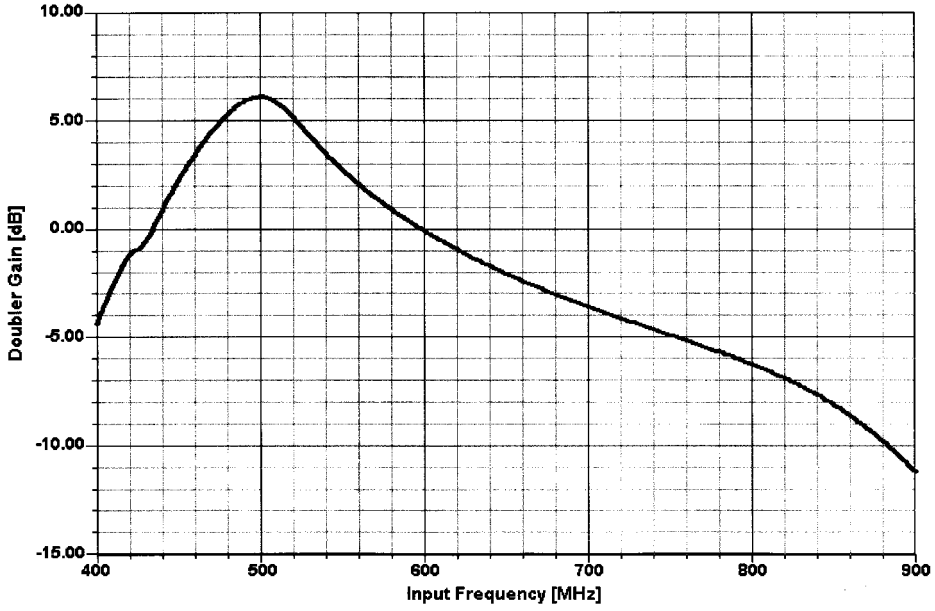


Figure 3-165 Frequency-dependent gain of the BJT doubler.

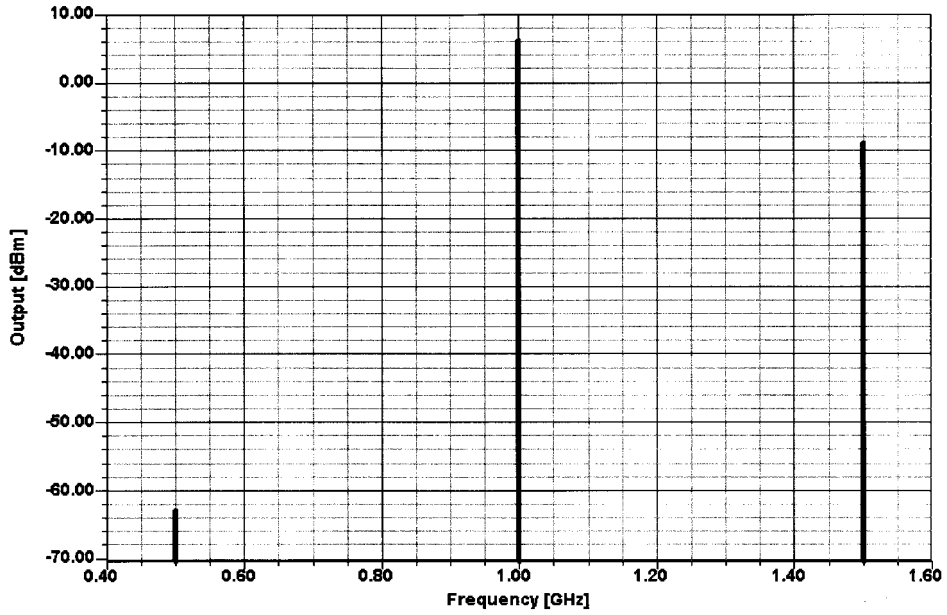


Figure 3-166 Output spectrum of the BJT doubler.

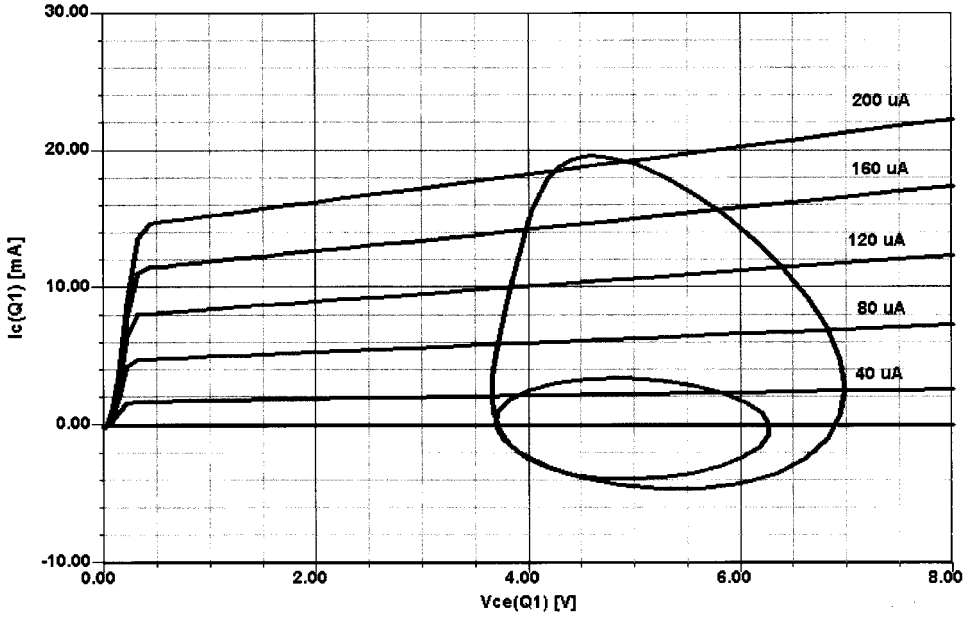


Figure 3-167 The dc *I–V* curves and ac load line for RF transistor Q1 in the BJT frequency doubler.

eliminate the desired second harmonic quite well). Actually, two diodes could also have been used for doubling, but since this requires transformers and a postamplifier, we have shown the preferred application. The only advantage of the diode application is that it is wideband, while this BJT circuit, as we have just learned, must be made narrowband for reasons of spectral purity.

Finally, we are curious to see the actual operating state of this amplifier, and this can be done by analyzing the load line (Figure 3-167). This is as complicated as the analysis for the differential amplifier. Because of the stored energy in the various tuned circuits, the load line looks quite impressive and would probably be difficult to predict without a nonlinear CAD tool.

3-9 MULTISTAGE AMPLIFIERS WITH AUTOMATIC GAIN CONTROL (AGC)

Having discussed the basic differential amplifier, we will now evaluate the performance of the MC1350/1490 video amplifier chip from Motorola. Most of the circuitry in the chip is self-explanatory. Figure 3-168 shows its inner circuit, and Figure 3-169 show a recommended application. By grounding one side of the input, we effectively get the same performance as the differential amplifier just discussed, with one difference being that the input is now fed to the constant-current generator. To minimize the distortion at the two lower transistors, they are decoupled by a 66-Ω resistor. Figure 3-170 shows the frequency-dependent gain, matching, and noise figure. Since this particular device has been specifically designed for AGC purposes, we also show in Figure 3-171 its gain reduction versus applied AGC voltage.

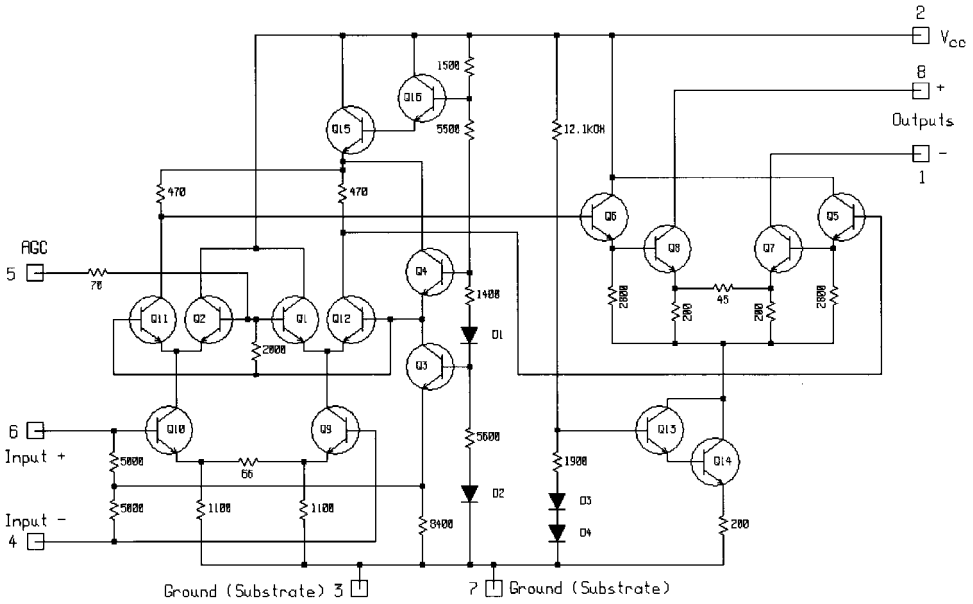


Figure 3-168 Schematic of the Motorola MC1350/1490 entered for simulation.

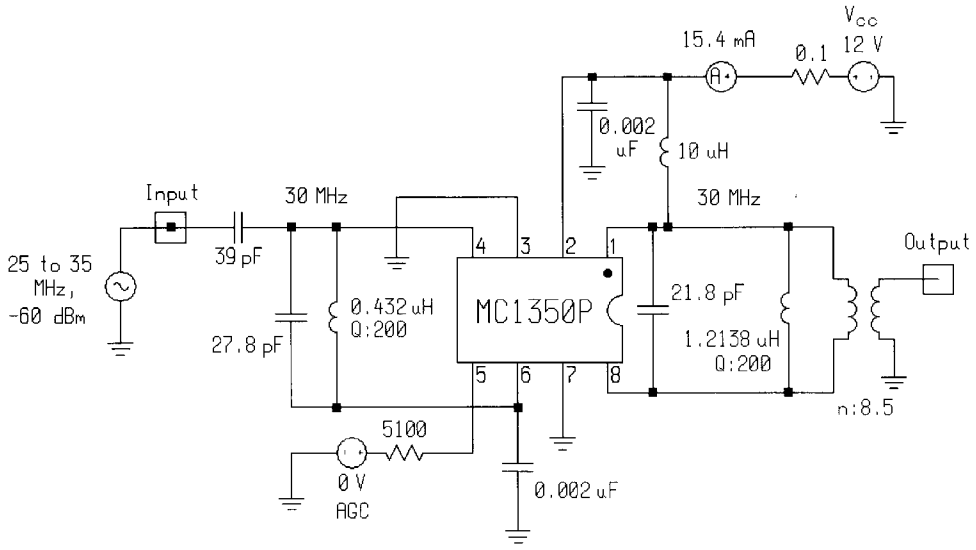


Figure 3-169 MC1350/1490 30-MHz application circuit.

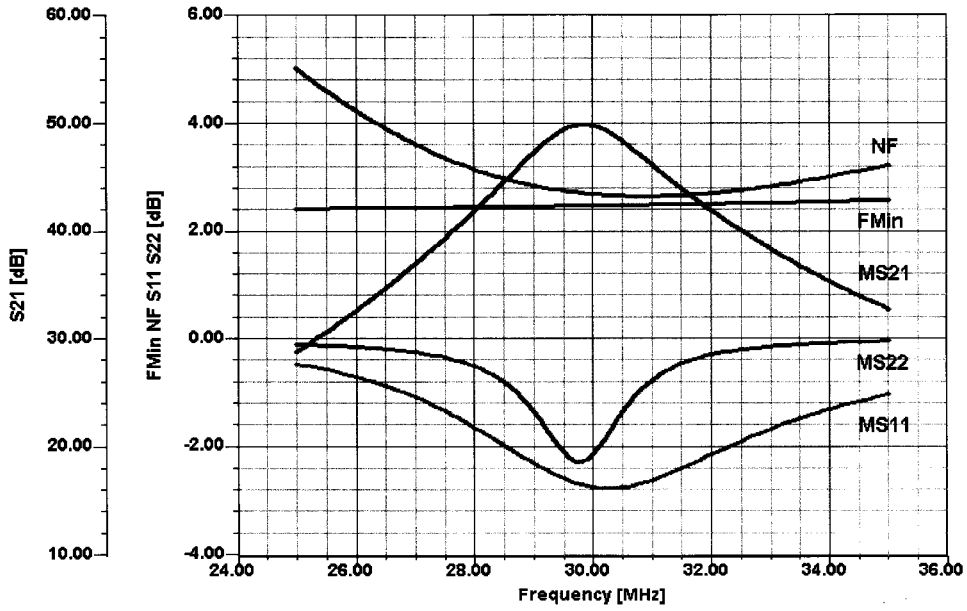


Figure 3-170 Frequency-dependent gain, matching, and noise performance of the MC1350/1490 in the circuit of Figure 3-169.

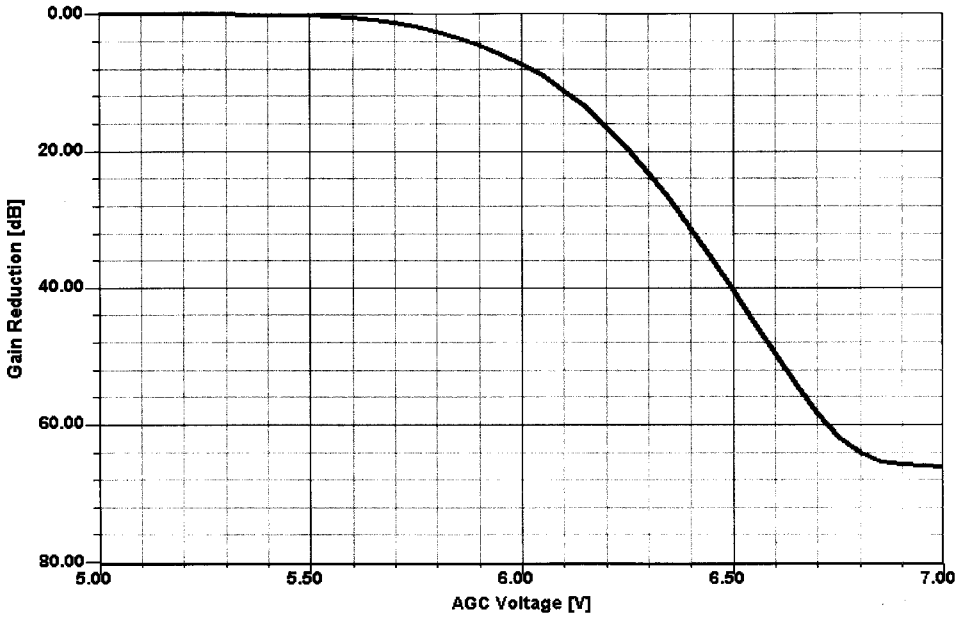


Figure 3-171 Gain reduction versus applied AGC voltage for the MC1350/1490 IC.

3-10 BIASING

A lot has been written about biasing, and since the purpose of this book is not to focus too much on dc, the following examples including RF consequences have been selected and considered sufficient. Early application reports by Motorola and others have provided a wealth of data. As a side effect of the introduction of the wireless era, we find that the engineers are so pressed for time that the number of complete application reports has approached zero. This is one of the reasons why we have tried to minimize the mathematical basis of this book (some is still needed), while still providing practical assistance. Figure 3-172 shows a textbook type of approach for RF.

In order to decouple the transistor from the biasing, we use RF chokes and dc decoupling (bypass) capacitors. This is the technique used in previous examples, so one might ask, "What's new?" As frequency increases, these inductors are either not manufacturable or have such a low Q that their use becomes questionable. This is the point where one may introduce the so-called distributed elements.

Figure 3-173 shows the very same circuit but resorting to distributed rather than lumped elements. The newly introduced elements (part of any good, up-to-date CAD tool; see element library for how to use their physical element descriptions) are described next.

- *Transmission Line.* Any printed connection between two points on a circuit board is a transmission line (Figure 3-174). Its characteristic electrical impedance is a function of the square root of the dielectric constant (ε_r), the width, metallization, thickness and height above substrate of the line, and the loss tangent of the substrate. Since lines frequently have to be laid out in the form of curved connections or have a bend in their direction, we have to add elements capable of describing the high-frequency conse-

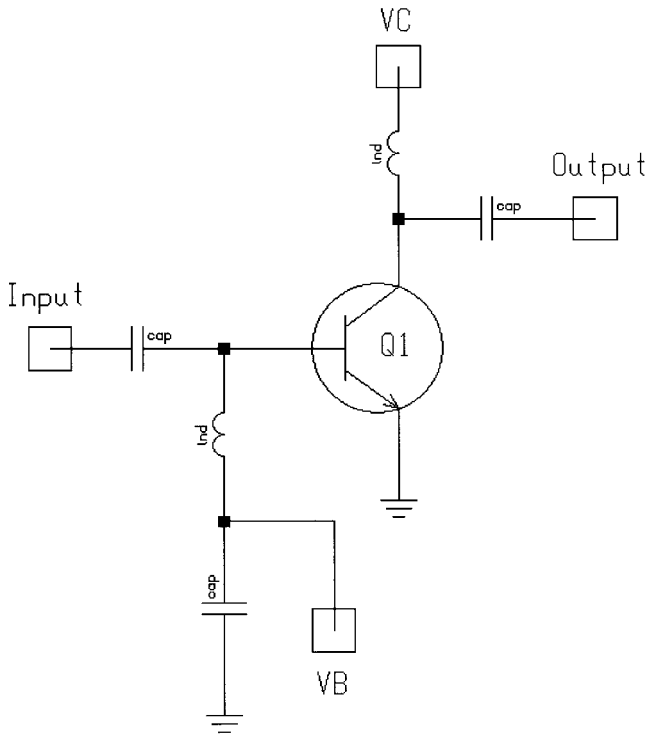


Figure 3-172 Simple BJT RF amplifier with lumped elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

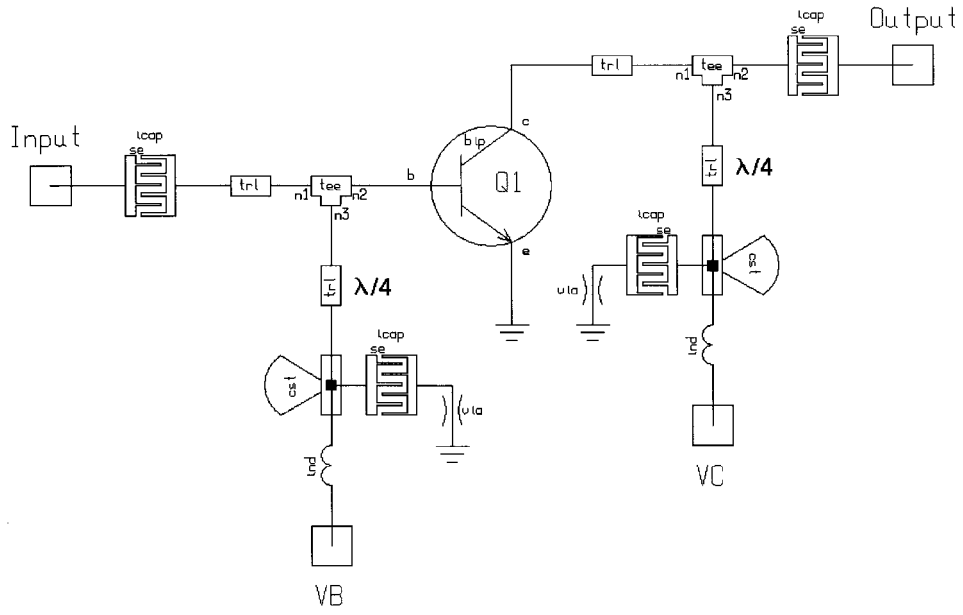


Figure 3-173 Simple BJT RF amplifier with distributed elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

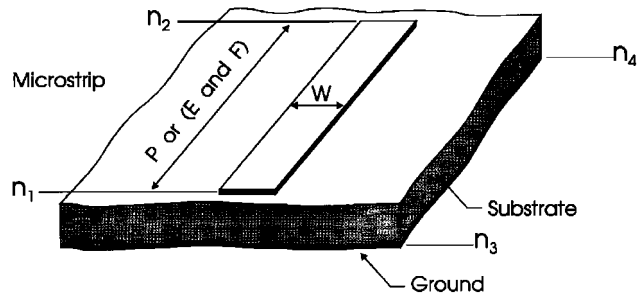


Figure 3-174 Transmission line in microstrip.

quences of such connections. Figures 3-175 and 3-176 show mitered and radial bend elements that perform this function.

- T *Junction, Cross, and* Y *Junction.* By the time a point like a collector or base, or its FET equivalent, spreads out into connections with other elements, we need additional modeling capability to describe ⊤ connections, crossings, and Y junctions. Figures 3-177, 3-178, and 3-179 show the way in which these connections must be modeled.

If the need exists, the standard inductances must be replaced with a transmission line whose length is λ/8 at the operating frequency. At higher frequencies, these transmission lines, however, then go into λ/4 resonant mode and later become capacitive. This type of design makes it fairly narrowband. A way around this is the use of printed inductors, as shown in Figures 3-180 and 3-181.

These inductors have a self-resonant frequency similar to the transmission line mentioned above, but the safety margin is significantly higher.

Talking about printed inductors, a logical extension of this is the printed transmission-line-based transformer, as shown in Figure 3-182. One can consider this as two interlaced rectangular inductors and, based on the substrate material, they are useful over a wide frequency range. Besides being used as a transformer, they can also be used to transit from

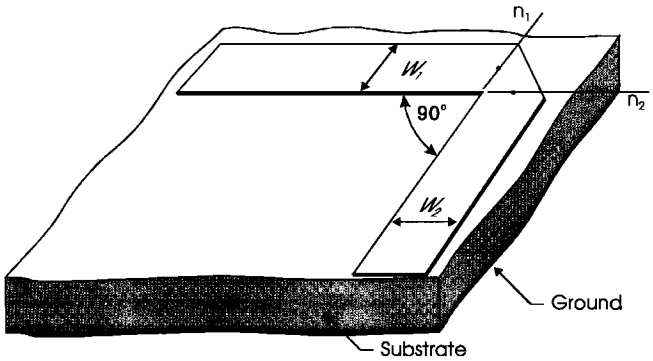


Figure 3-175 Mitered bend.

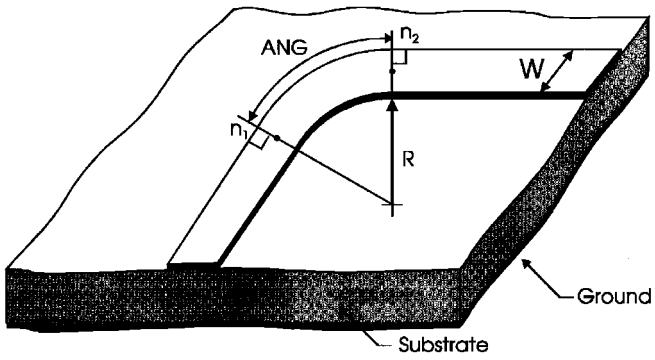


Figure 3-176 Radial bend.

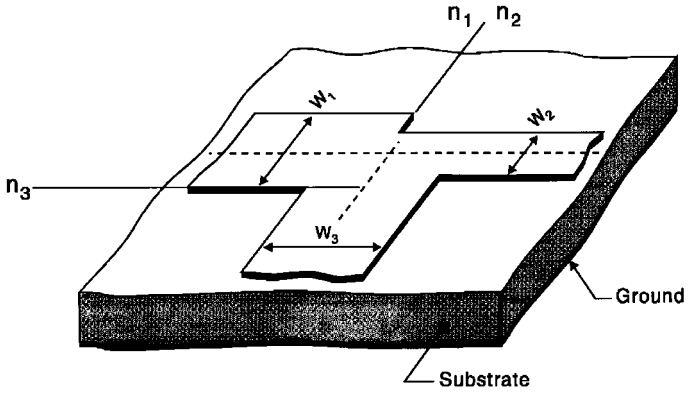


Figure 3-177 T junction.

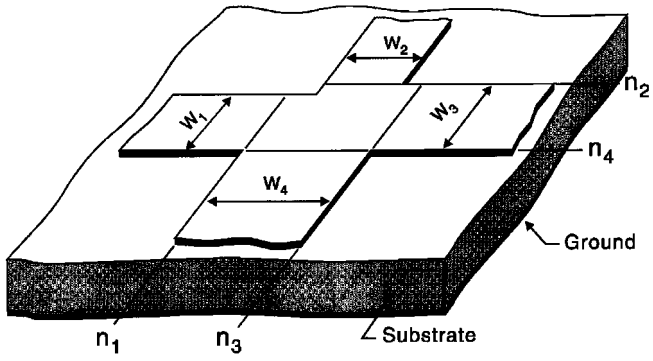


Figure 3-178 Cross.

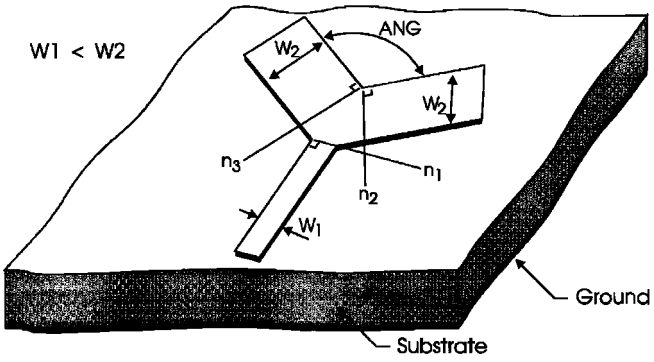


Figure 3-179 Y junction.

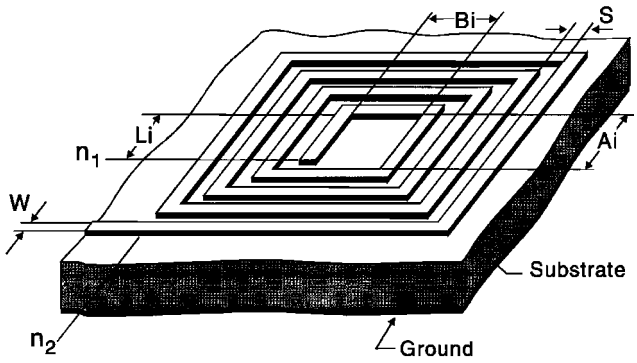


Figure 3-180 Rectangular inductor.

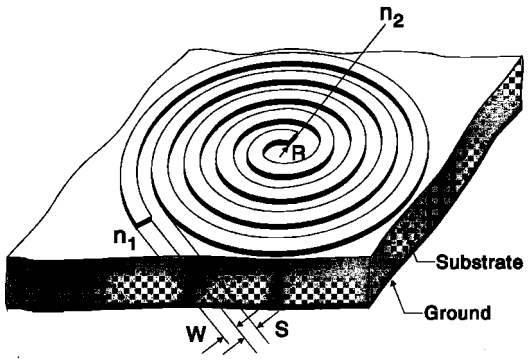


Figure 3-181 Spiral inductor.

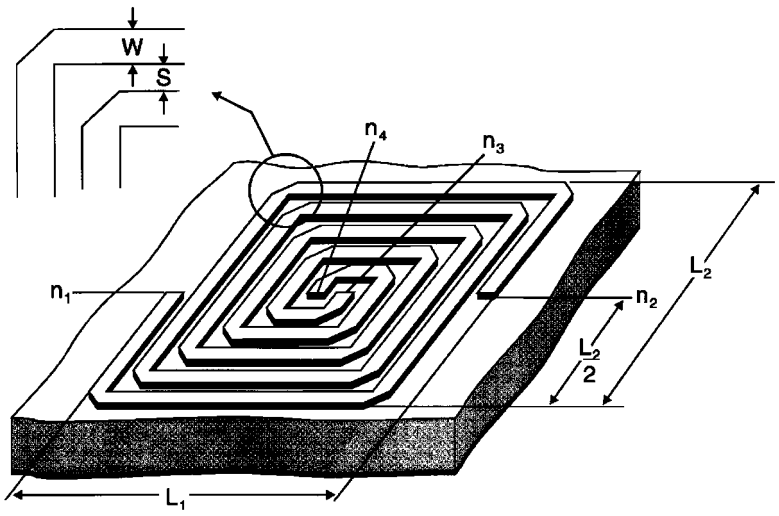


Figure 3-182 Transformer in microstrip.

unbalanced to balanced transmission provided that the difference in length from a connection point of view does not cause any problems (this is a layout issue).

A popular form of combining stages is the so-called Lange coupler (Figure 3-183) invented by the German Julius Lange. It is one of the major contributions in wideband applications.

Useful application of the Lange coupler probably starts at 4 GHz. It consists of parallel transmission lines with the appropriate connections as shown. Lange couplers are typically built with four, six, and eight fingers. The Ansoft Serenade product has a Lange coupler synthesis program that can be used to gain more insight into this coupler's application. We assume that other modern software has similar capabilities.

Where meander-type inductors are necessary, a neat way to implement and simulate them is to use the multiple coupled lines element (Figure 3-184) of the Serenade product, which both quickly and accurately calculates the behavior of the meander, including self-resonances and losses. We made use of this arrangement in our previous examples.

- *Interdigital Capacitors.* The issue of tolerances of small capacitors already has been brought up. The interdigital capacitor can be made on printed circuit board material as

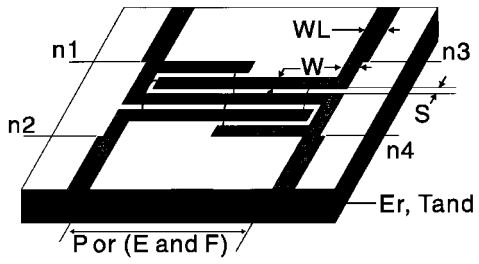


Figure 3-183 Four-strip Lange coupler.

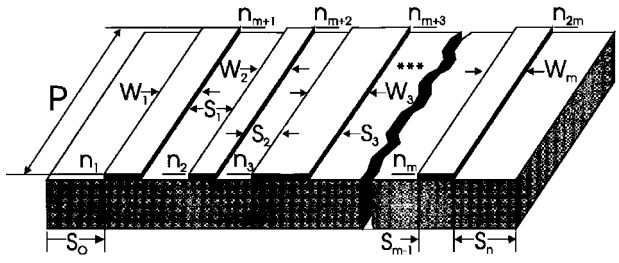


Figure 3-184 Multiple coupled lines element in microstrip.

well as gallium arsenide, and if its dimensions are continuous with the transmission-line width it does not cause any abrupt changes in impedance. This type of capacitor permits one to obtain very small values. An alternative to this is the use of transmission lines with a length of $\frac{3}{8}\lambda$. We learned earlier that a transmission line is inductive at frequencies below its fundamental resonance, is resistive at below resonance, and becomes capacitive above resonance. Again, bandwidth is also an issue. An interdigital capacitor consists of a number of parallel fingers as shown in Figure 3-185, and its capacitance can be varied by adjusting the number of fingers and their spacing. The advantage of the interdigital capacitor compared to discrete components is its low variation in value.

- *Radial Stubs.* The radial stub (Figure 3-186) is not much different from a $\lambda/4$ resonator, but its bandwidth is much greater than a simple transmission line. This is another way to ground the "cold" side of a transmission line or part of a circuit that needs to be grounded for RF. Of course, the interdigital capacitor comes in a version that is a combination of a capacitor and a via hole.
- *Via Holes.* The "cold" end of the transmission line, being considered either an inductor or a capacitor, needs to be connected to a sufficiently large copper backplane. One very efficient way to do this, especially if there is not enough copper left on the top of the board, is the use of via holes (Figure 3-187). One could theoretically generate

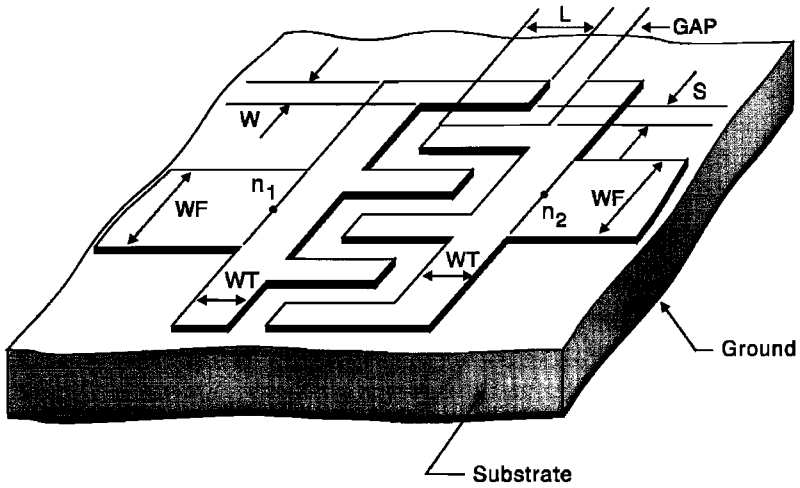


Figure 3-185 Interdigital capacitor.

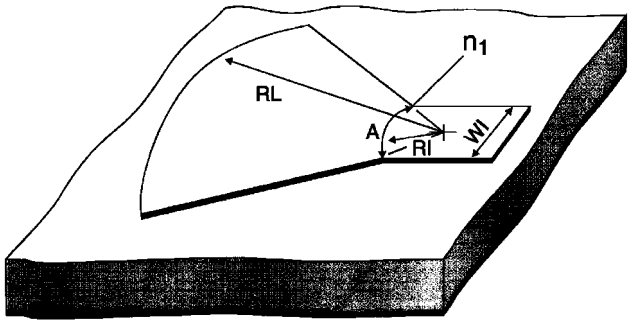


Figure 3-186 Radial stub.

a via hole with a rivet, but most manufacturing processes don't allow this; the normal solution is to use plated-through holes left open. In PC boards, via holes are typically cylindrical; on substrates like GaAs, they may be conical.

- *Correction Elements.* Although the behavior of an actual circuit proceeds regardless of our ability to measure and describe it, we do not enjoy this luxury in simulating circuit behavior in software. In a simulator, effects that are insufficiently described will be inaccurately simulated. For instance, segments of high-impedance transmission line (e.g., 120 Ω) are frequently used for dc feeds and RF chokes. In predicting the effect of a transition from 120-Ω line to 50-Ω line, a simulator must be alerted to the discontinuity so it can do the necessary mathematical corrections to account for the impedance jump. To do this, a specific circuit element, the STEP (Figure 3-188), must be inserted between the 120-Ω and 50-Ω line elements in the simulation circuit file. In addition to substrate data, we characterize a STEP by providing the widths of its input and output lines. The element itself has no physical length.

A similar correction is necessary if a transmission line is used as a resonator or just "left open" at one end. Such a transmission line tends to radiate and, because of its high-impedance properties, it reacts differently as far as its electrical length is concerned. A zero-length one-port element, the OPEN (Figure 3-189), must be added to such a line for mathematically correct calculation.

We end this excursion into distributed elements here. Interested readers may want to obtain a CD containing the element library of their favorite CAD tools that combine capabilities in the microwave area with a full set of distributed circuit elements. It is most important to keep

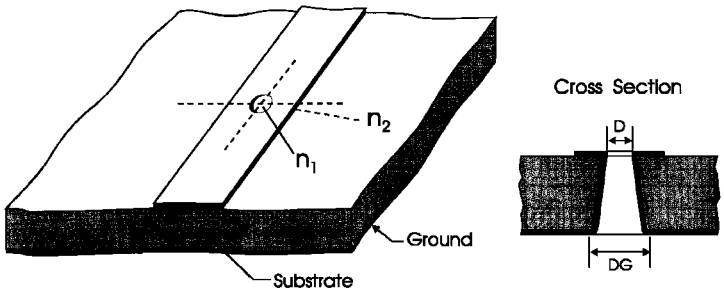


Figure 3-187 Via hole.

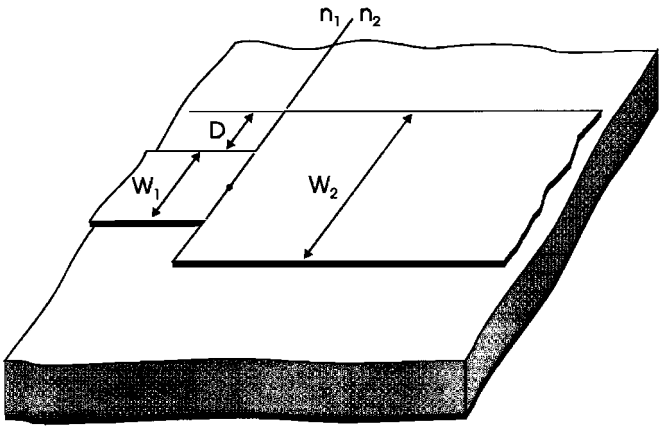


Figure 3-188 The STEP element tells the simulator to calculate the effects of joining transmission lines of differing characteristics.

in mind that, as frequency increases, we rapidly move into the area where we must consider all these distributed elements to achieve accurate simulations—even if doing so makes simulation a painful and time-consuming effort.

Finally, anyone who adventures into this area *must* obtain a foundry manual from the company that will build the integrated circuit or hybrid under design. There are basically two foundries, one applicable for MOS technology and one for GaAs. In the case of the bipolar transistor, the foundry service is not yet well-established. While we were able to interest one German company and one U.S. company in generating a custom bipolar IC within the activities of this book, the only "open" foundry we know of for bipolar (actually HBT) technology is the one operated by TRW. As of 1999, this seems to be the leading company for BJT-related products, and many currently available ICs have been developed with this foundry service.

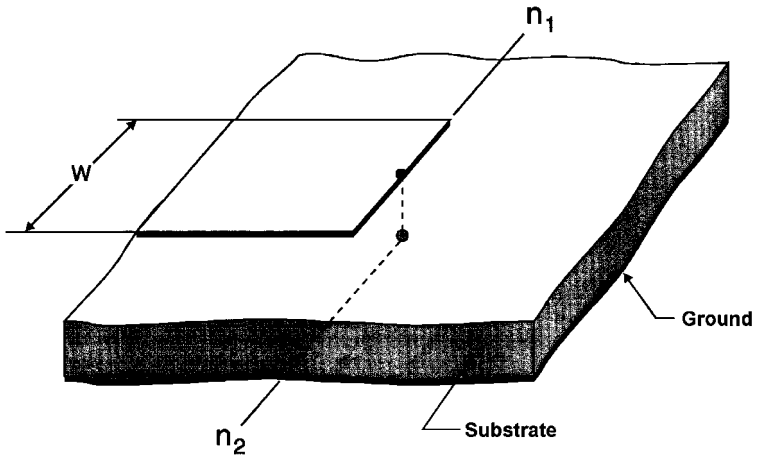


Figure 3-189 The OPEN element tells the simulator to calculate the effects of leaving the end of a transmission line unconnected.

We mention the issue of foundries here again because each foundry has its own proprietary approach to modeling discontinuities. The availability of a foundry service somewhat eases the requirement that a designer be fully up to speed on the nuances of discontinuities, because a foundry's designer service will help customers account for all relevant parasitics or discontinuities in their designs. In addition, there are tables of *S* parameters for standard cells of either capacitors, resistors, or inductors. The designer may then be forced to adjust the circuit so that it will work with a particular inductance value or value of another component within the resolution of the table that describes these elements. Information on the active parts, such as diodes and transistors, was given in Chapter 2.

3-10-1 RF Biasing

Applying the knowledge we have just acquired about distributed elements, Figure 3-190 shows a somewhat exaggerated case of using the various elements. While it is consistent with the simple RF case with which we started, modeling of this type is certainly necessary at higher frequencies. Many but not all of the elements we have just described are used, but we have not solved the question of the actual dc biasing.

3-10-2 dc Biasing

The following discussion is the recommended procedure for biasing a transistor. We will start with the collector current (in our example, Figure 3-191, 10 mA). To separate RF and dc, we have a combination of a load resistor (R_L) and an inductance (L_C). Information about this operation was given already in Section 3-2-6. The emitter resistor should have a voltage

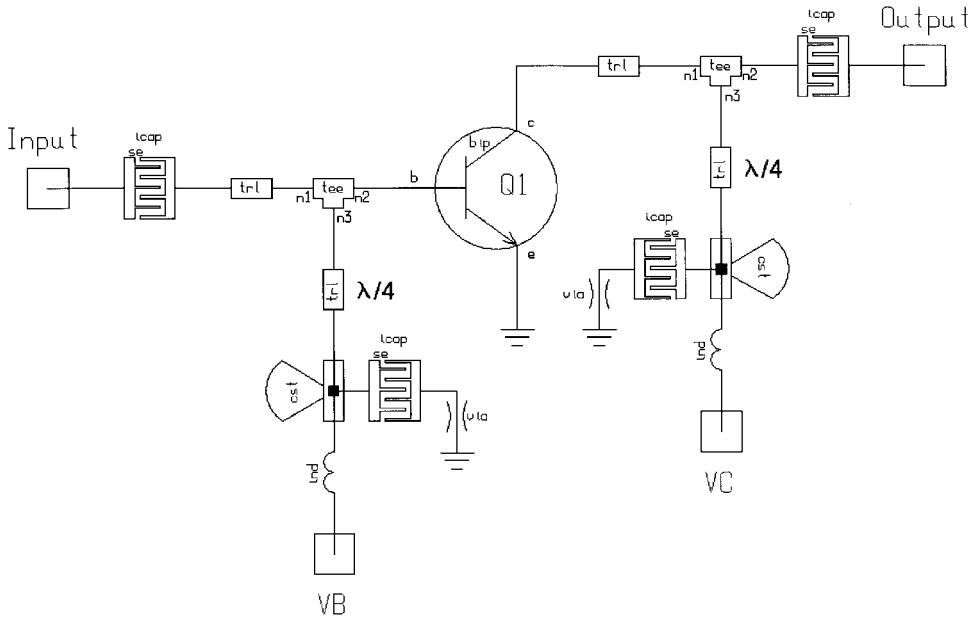


Figure 3-190 Simple BJT RF amplifier with distributed elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

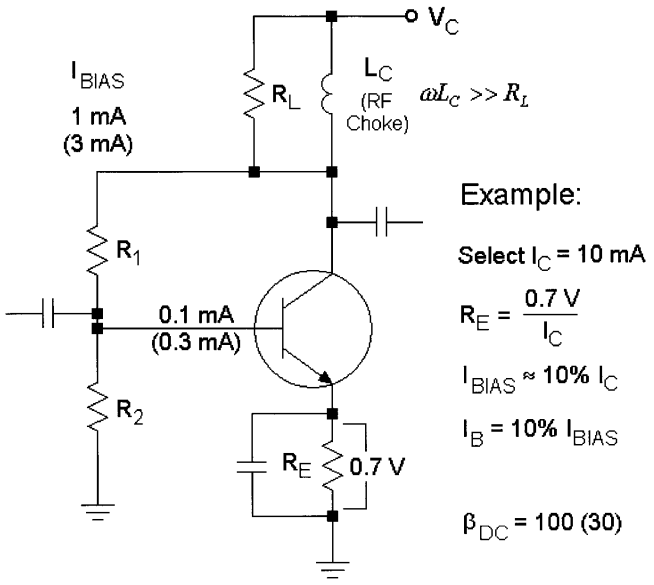


Figure 3-191 A dc biasing example.

drop of about 0.7 V, and therefore the emitter resistor is $0.7/I_C$. The next assumption is that the transistor has a dc current gain of 100, and therefore we decide to have the bias resistor chain (R_1 and R_2) draw 10% of I_C, or 1 mA. Assuming a collector voltage of 5 V, the value of $R_1 + R_2$ has to be 5 kΩ to result in the required 1 mA. Because the base current now is 10% of the divider current, the base will draw 0.1 mA. The voltage across R_2 has to be the voltage drop across R_E (0.7 V) + V_{BE}, which is also 0.7 V, or 1.4 V. Given the 1 mA, R_2 = 1.4/1E–3 = 1.4 kΩ. R_1 therefore is 5 kΩ – 1.4 kΩ = 3.6 kΩ. This concludes the calculation.

In the case of a dc beta of only 30, we need to raise I_{BIAS} from 1 mA to 3 mA since the base current now will be 0.3 mA. This changes the resistor values. $R_1 + R_2$ = 5 V/3 mA = (rounded) 1.6 kΩ. Following the same approach, we need 1.4/3E–3 = (rounded) 470 Ω. This makes R_1 = 1600 Ω – 470 Ω = (rounded) 1100 Ω. These roundings were necessary because resistors can only be bought in certain resistance steps.

However, for low battery voltages, the 0.7 V between the emitter and ground cannot be wasted. Therefore, a dc stabilization circuit, to which we have already referred, will be considered next. Figure 3-192 shows a combination of an RF transistor, Q1, and a dc stabilization transistor, Q2. The most important thing about this circuit is the fact that the difference between the supply voltage and the collector of Q1 is 0.3 V, which is approximately half of the voltage we had to "waste" across R_E in the above-mentioned circuit. Plus, the voltage drop is now in the collector, and given the fact that we feed Q1's collector through the inductor of a quasi-tuned circuit rather than through only a resistor, we get a much higher voltage swing. The base–collector diode temperature dependency of Q2 is compensated for with the silicon diode feeding the base. It is also fascinating to see that, because of the low currents, both the voltage drop of the diode and V_{BE} of Q2 are about the same, and yet Q1, operating at a reasonable current, shows the expected 0.7–0.8 V V_{BE}. We have supplied all the necessary dc voltages to give the reader an incentive to calculate the mechanism of this stabilization. The key equation to this is to determine the current in the voltage divider feeding

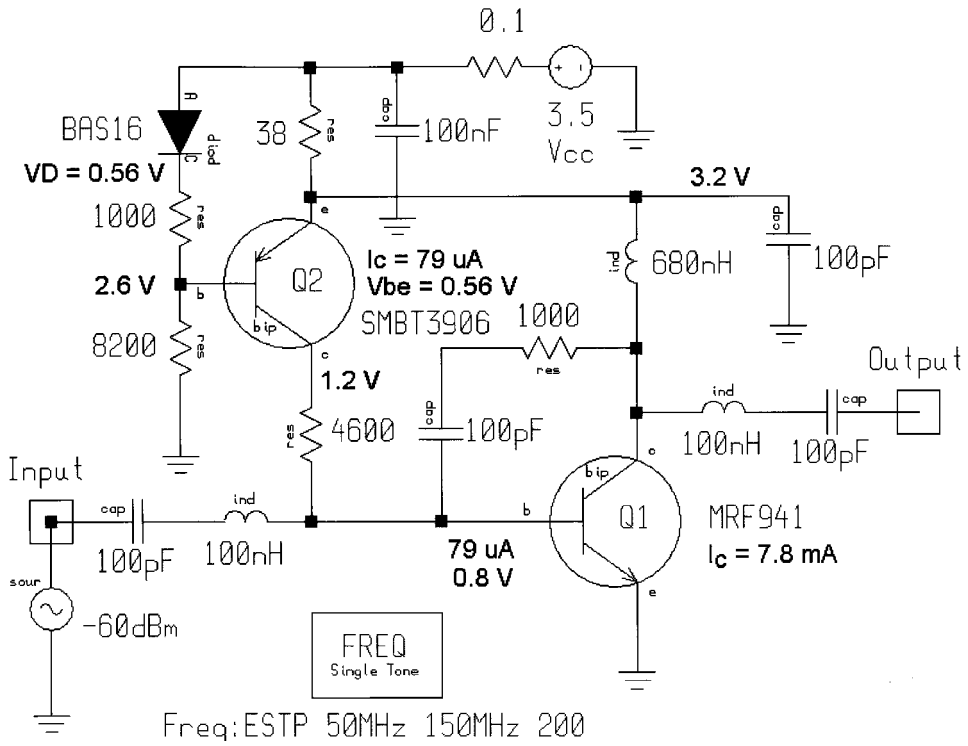


Figure 3-192 An RF amplifier with active biasing. (After the low-noise amplifier circuit in [29].)

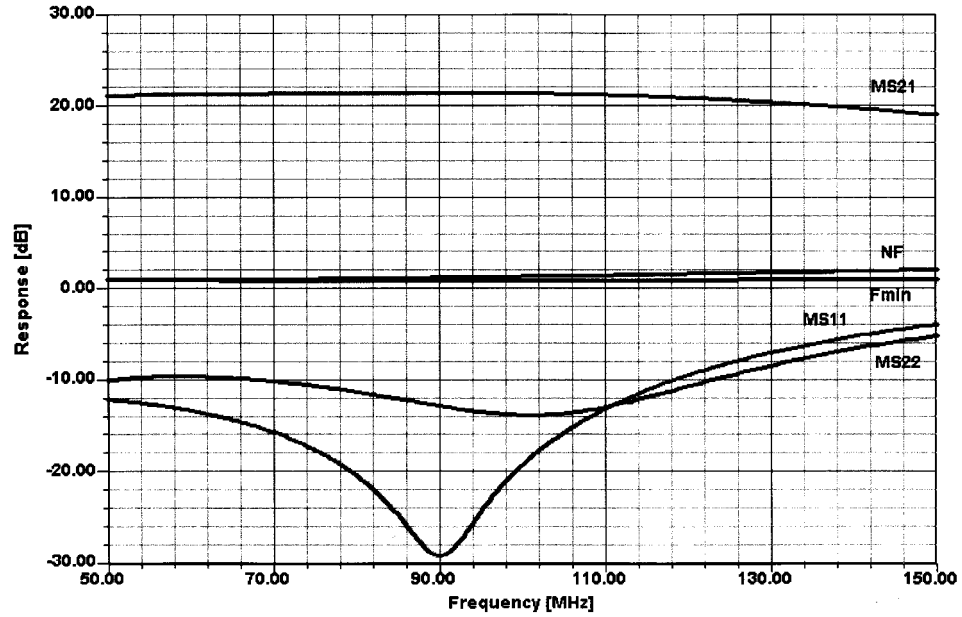


Figure 3-193 Frequency-dependent gain, matching, and noise responses for the amplifier in Figure 3-192.

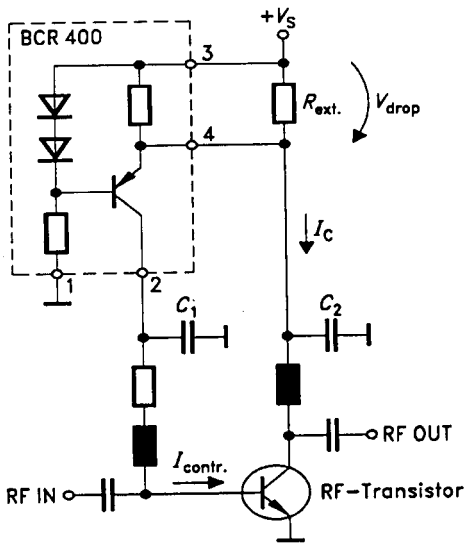


Figure 3-194 The Siemens BCR400 active bias controller can be applied to FETs as well as BJTs.

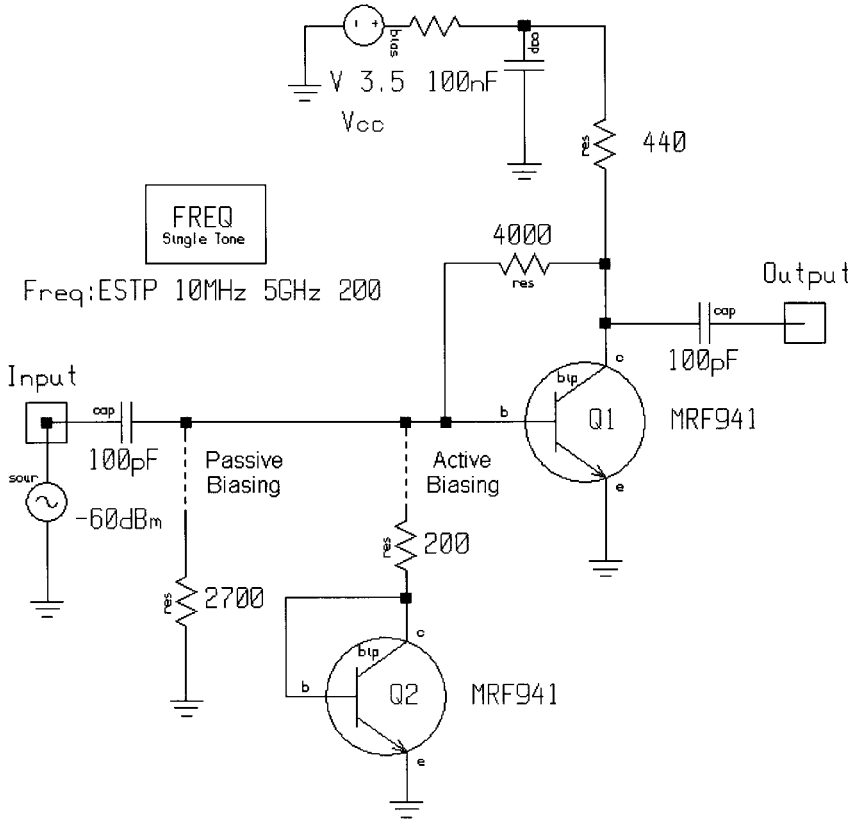


Figure 3-195 Amplifier stage (I_C = 3 mA) showing active and passive bias alternatives.

Q2, and proceeding from there. On the other hand, having a nice CAD tool that provides insight into the dc voltages and allows the addition of dc voltage probes makes life much easier. To prove that this circuit actually works, Figure 3-193 shows its frequency-dependent gain, matching, and noise figure performance. It is based on a Motorola application, which had excessive voltage drop for Q1 and which we have improved to be suitable for low-voltage operation [29].

Integrated active bias solutions are also available. Figure 3-194 shows the Siemens BCR400 active bias controller applied to a BJT; it can also be applied to FETs and TR switching diodes, as we saw in Figures 3-57 and 3-59, respectively.

3-10-3 dc Biasing of IC-Type Amplifiers

In integrated circuits, we do not have the unlimited flexibility described earlier. This results frequently in constant-current sources or transistors being used as diodes for biasing purposes. Our first suspicion is that this is going to cause a lot of noise in the IC; Figure 3-195 shows a test example.

As suspected, choosing the same bias point of 3 mA, the noise figure using active bias is quite a bit higher than the passive one, as shown in Figure 3-196.

3-11 PUSH–PULL/PARALLEL AMPLIFIERS

Since it is practically impossible to obtain $S_{11} = S_{22} = 0$, using a combiner at the input and output reduces the problem. Figure 3-197 shows the 3-dB hybrid approach. It has a practical

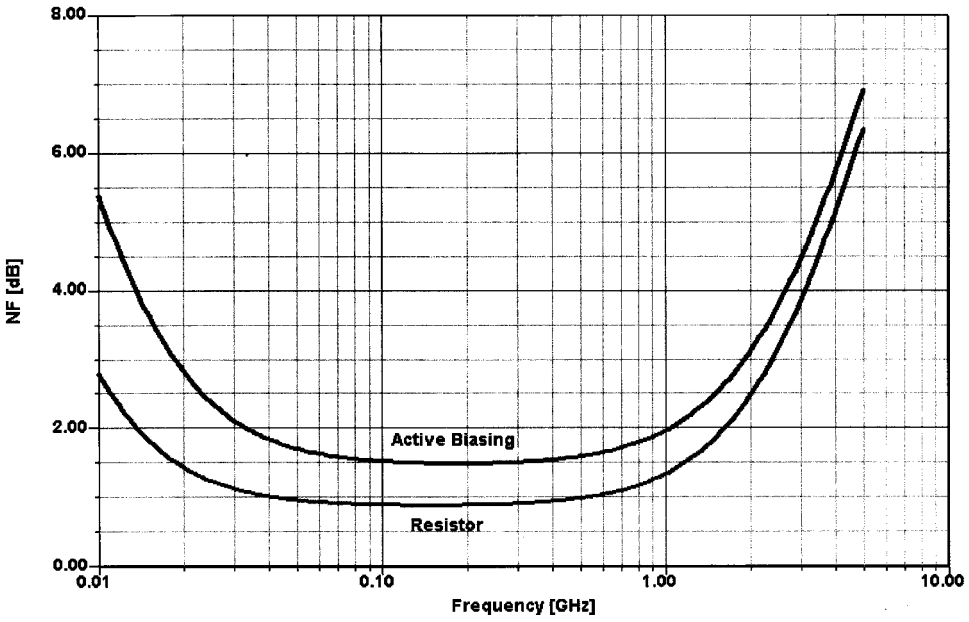


Figure 3-196 Frequency-dependent noise figure of the amplifier circuit shown in Figure 3-195. The active biasing curve is considerably noisier.

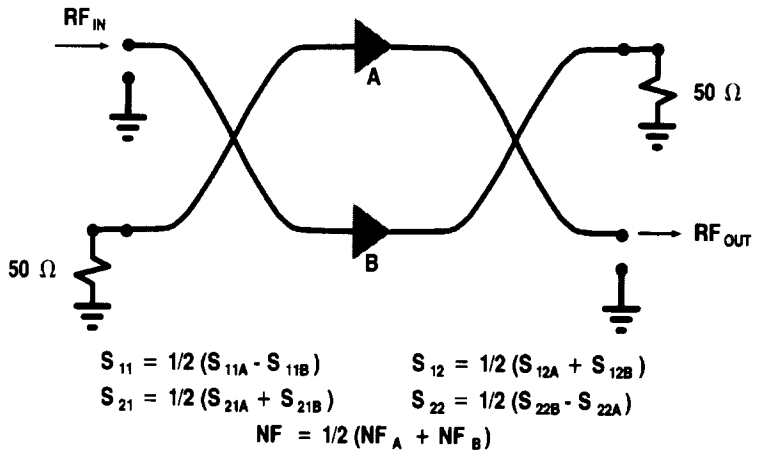


Figure 3-197 Combining power with 3-dB hybrid couplers.

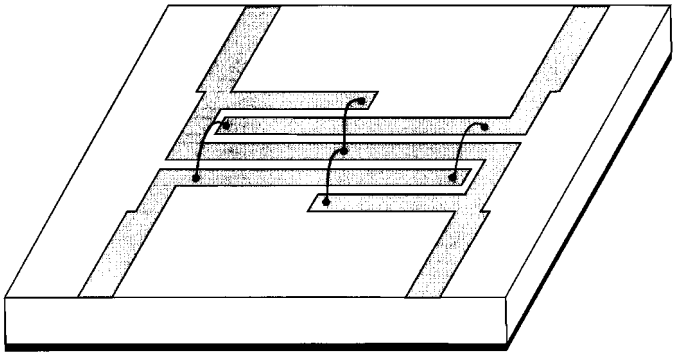


Figure 3-198 Four-strip version of the Lange coupler, a broadband 90° hybrid.

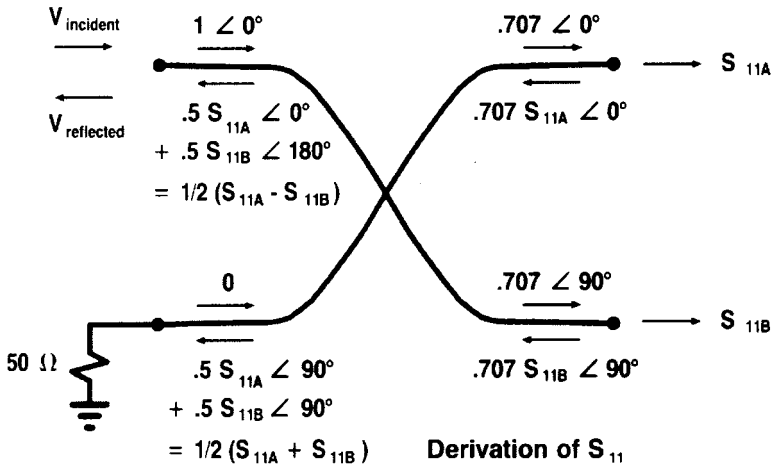


Figure 3-199 3-dB hybrids. This analysis method may be applied to any coupler.

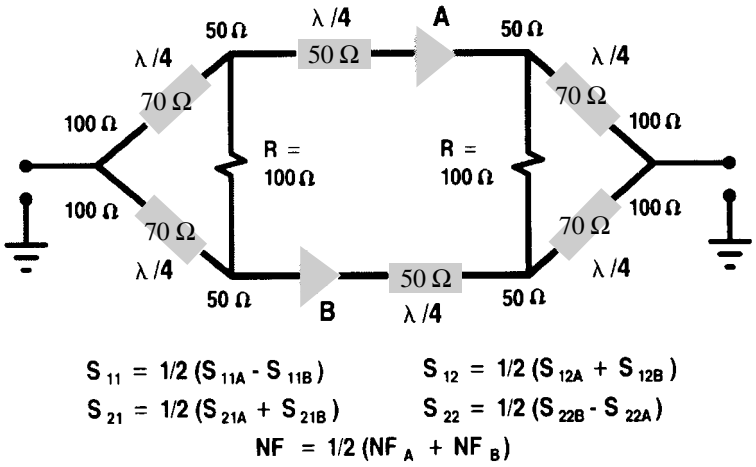


Figure 3-200 Wilkinson divider/combiners.

bandwidth limit of about 4:1; its advantage is that it is compact and ideal for cascading. If one stage opens, P_{OUT} and gain drop only 6 dB; however, it does not protect against load mismatch. A recommended distributed form is the Lange coupler (Figure 3-198). To analyze this approach, Figure 3-199 shows the resulting impedances and necessary phase shifts. If a Wilkinson coupler [30] is used instead of the 3-dB hybrid, the bandwidth is reduced. The Wilkinson coupler is larger in size than the 3-dB hybrid but more easily realized, leading to lower cost. Figure 3-200 shows the actual impedances necessary for matching, with A1 and A2 being the actual amplifiers combined.

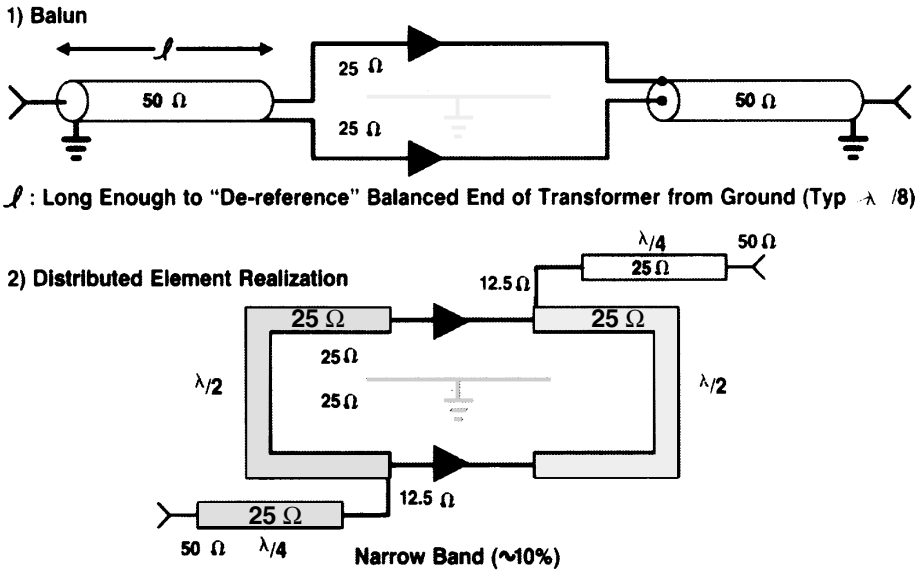


Figure 3-201 Push–pull amplifiers.

The solution for push–pull amplifiers resorts to a balun with the following results obtained: Bandwidths of greater than one decade are achievable; the balun structure gives 4:1 impedance advantage; there is no isolation; $S_{11} \neq 0$, $S_{22} \neq 0$; and cancellation of even-order products is provided.

Since most amplifiers have 50-Ω rather than 25-Ω impedances, Figure 3-201 shows a circuit that matches 25 Ω to 50 Ω on both sides. Since the standard amplifier operates at 50 Ω rather than 25 Ω, an additional network may be necessary to transform the impedance up to 50 Ω again. A similar problem occurs at the amplifier output.

3-12 POWER AMPLIFIERS

Power amplifiers are devices that are tasked to transform as much dc power into RF/microwave power as possible. Depending on the technology, devices are available that can produce at least 100 W up to 2 GHz. Since the transconductance of FETs at conducting angles of less than 180° becomes very small, the actual power gain sharply decreases compared to Class A operation while the power-added efficiency is optimized to values of 55% (Class B) or so. The dissipation of these devices is, of course, quite high, while in Class B operation this is even more of an issue. Power devices in Class C are dying to have big heat sinks to survive. One of the nice things about FETs is a much reduced thermal runaway effect compared to their bipolar brothers.

Power amplifier classes are defined in terms of the device conduction angle and/or the type of device operation:

Class A	"Linear" operation; 180° conduction angle
Class AB	"Linear" operation; conduction angle between 90° and 180°
Class B	"Linear" operation; 90° conduction angle
Class C	Fixed drive; less than 90° conduction angle
Class D	Switched operation, conduction range may vary with time from 0° to 180° or may be fixed.

There are also Classes E and F, which are currently under evaluation and showing some promising results.

Figure 3-202 shows Class C operation. The output of the amplifier is a current pulse train, which frequently leads to misunderstanding as to the calculation of its load. Given the fact that the amplifier is a current generator, it is silly to assume one has to do a conjugate matching looking backward into the amplifier. This is the condition for Class A operation. Unfortunately, this is still not understood by some engineers, but we hope we can clear this up today. The theoretical efficiency is 89.7% at a conducting angle of less than 180°; in practice, efficiencies more than 85% are not obtainable. (All of these efficiencies are those of the device itself; losses in matching networks are not taken into account in these efficiency values.) Considering optimal gain and optimal output power, the conducting angle actually settles somewhere between 100° and 140°. Again, the power gain in Class C operation is less than in Class A and Class B operation.

3-12-1 Example 1: 7-W Class C BJT Amplifier for 1.6 GHz

The following demonstrates how to design a 1.6-GHz, 7-W amplifier operating in Class C. We need to point out again that all modulation forms that contain the information in both

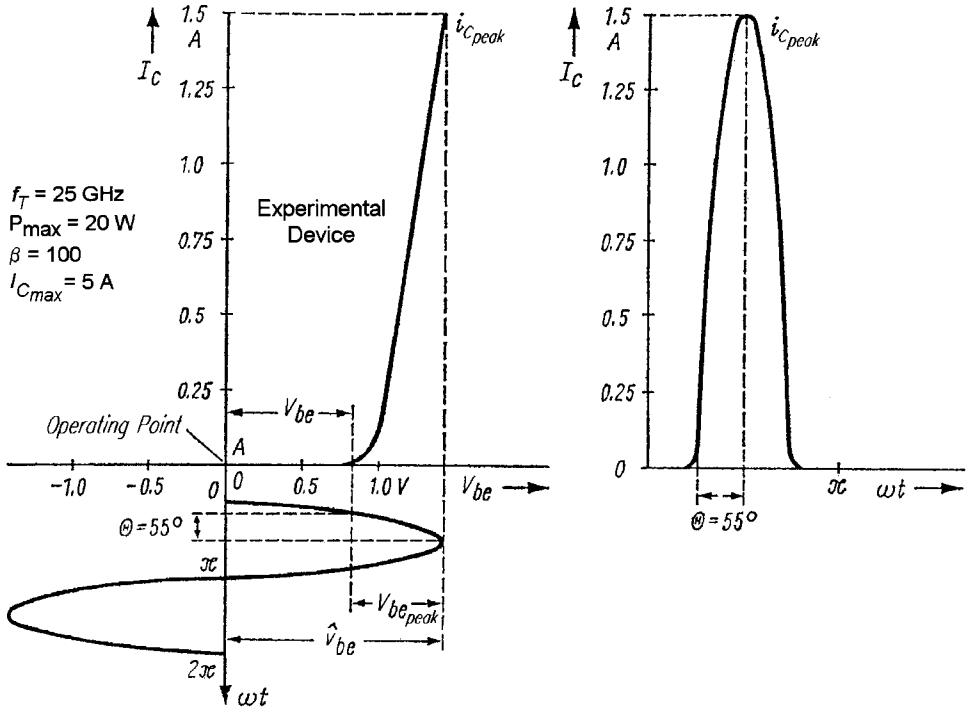


Figure 3-202 Output current as a function of drive voltage, taking into consideration the knee voltage of the base–emitter *pn* junction. The output is a series of current pulses with a duty cycle determined by the conduction angle.

amplitude and phase require more linearity than Class C affords. Needless to say, this affects the efficiency; therefore, at a later point we will also show a Class AB power amplifier. This is a tutorial, so we will present a UHF example and walk the reader through all the problems involved in the design. The selection of the transistor is done by first looking at the f_T, which in our case should be at least 4–5 GHz. The first order of business is to define the point of operation for the output amplifier. To do this, we need to determine the maximum voltage and current values supplied by the manufacturer. We have chosen a device that operates at 28 V and can sustain a peak collector current of 1.5 A. Figure 3-202 lets us determine the conducting angle Θ from the I_C/V_{BE} plane. This figure shows on the left side the I_C/V_{BE} plane and the necessary voltage to drive the amplifier.

Initially, we assume that the base–emitter junction is operated without any bias ($V_{BE} = 0$). The dc I–V curve shows that the collector current starts at about 650 mV. The amplifier itself, therefore, is already in Class C operation. For full output power, the drive voltage $V_{BE,\text{peak}}$ has to be 1.14 V. Therefore, the conducting angle can be determined by

$$\cos \Theta = \frac{V_{BE}}{\hat{\boldsymbol{v}}_{BE}} = \frac{0.65}{1.14} = 0.57$$

$$\cos 0.57 = 55^\circ$$

V_{BE} is the dc voltage at which the transistor starts conducting. From Figure 3-65 and Eqs. (3-160), (3-161), and (3-162), the fundamental current i_{c1} can be determined from

$$i_{c1} = i_{\text{peak}} \cdot f_1(\Theta)$$

or

$$i_{c1} = 1.5 \text{ A} \times 0.36 = 0.54 \text{ A}$$

The average collector dc current can be determined from

$$i_{c2} = i_{\text{peak}} \cdot f_2(\Theta)$$

or

$$i_{c2} = 1.5 \text{ A} \times 0.21 = 0.315 \text{ A}$$

The collector saturation voltage $V_{CE(\text{sat})}$, unfortunately, is around 2–3 V; for the purpose of the calculation, we assume that the saturation voltage is 3 V, and, therefore, we can assume that the maximum collector ac voltage $\hat{v}_{CE} \approx (V_B - 3) = 25$.

The appropriate collector load has to be

$$R_C = \frac{\hat{v}_{CE}}{i_{C1}} = \frac{25}{0.54} = 46.3 \ \Omega$$

The maximum output power is

$$P_{\text{out}} = \frac{(V_{CE} - V_{CE(\text{sat})})^2}{2R_C} = \frac{625}{92.6} = 6.75 \text{ W}$$

The only way to increase the power would be to have a smaller collector load at a higher current. In this family of Motorola transistors, there was no 10-W transistor; that's why we have chosen a 20-W experimental device, which for this purpose is overkill. The theoretical efficiency of the amplifier is defined as

$$\eta = \frac{P_{\text{out}}}{P_{\text{in}}} = \frac{6.75 \text{ W}}{0.315 \times 28} = 85\%$$

If the saturation voltage would be less, our efficiency would automatically increase, since for the given load, the actual output power would increase (6.75 W $\rightarrow$ 7.5 W). Assuming that this stage would operate at 1900 MHz with a gain of roughly 22 dB, the required drive would be significantly less than the output power, or roughly 40 mW. For the purpose of the following calculations, we will assume that the drive power is 0.7 W. The total dissipation of the device now adds up to be

$$P_T = P_{\text{in(dc)}} - P_{\text{out}} + P_{\text{drive}}$$

$$P_T = (28 \times 0.315) - 6.75 + 0.7 = 2.77$$

Power-added efficiency is derived from this. Knowing the power dissipation, one can calculate a thermal resistance R_{therm} of the heat sink and the transistor:

$$R_{\text{therm}} = \frac{T_J - T_A}{P_T}$$

Where T_J is the device junction temperature and T_A is the ambient temperature. For a maximum ambient temperature of 60 °C and a maximum junction temperature of 150 °C, the thermal resistance is

$$R_{\text{therm}} = \frac{120 - 60}{2.77} = 21.7\ ^\circ\text{C/W}$$

R_{therm} is the sum of thermal resistance of the transistor and its package, and the thermal resistance between the transistor package and the heat sink, and the thermal resistance between the heat sink and the ambient air. Therefore,

$$R_{\text{therm}} = R_{\text{th(case)}} + R_{\text{th(case/heat sink)}} + R_{\text{th(heat sink)}}$$

The required thermal resistance of the heat sink is calculated from

$$R_{\text{th(heat sink)}} = 21.7 - 10 - 1.4 = 15.6\ ^\circ\text{C/W}$$

Assuming an aluminum heat sink (a 2-mm-thick plate), the required heat-sink surface area A (cm^2) is

$$A = \frac{1}{\alpha R_{\text{th(heat sink)}}}; \quad \alpha \approx 1.5\ \frac{\text{mW}}{^\circ\text{Ccm}^2}$$

Therefore, the required area is

$$A = \frac{1}{1.5 \times 0.0156} = 42.7\ \text{cm}^2$$

The next task is to match both the input and output to the (most likely) 50-Ω source and termination. The following section will guide us through the procedure. The input impedance Z_{11} (after some lengthy search) was found to be $1 + j10\ \Omega$. This means that the input consists of the base-spreading resistance (approximately 1 Ω) and an input capacitance, which because of phase shift now turns out to be inductive. Since the popular method for input matching these days is a pi configuration, we are going to use a simple pi filter both for input and output matching. There are several equations in the literature that give the values for the two capacitors and inductors; most of them result in an incorrect resonance frequency. Since we have access to software tools, we decided to take three approximate values and use the optimizer to give us the appropriate matching. The only remaining variable in such an approach is the bandwidth. If we specify $S_{21} = 1$ and $S_{11} = 0$ (100% energy transfer and perfect return loss), we will realize that this is possible at only one frequency. We therefore need to identify a frequency band or bandwidth over which an acceptable match is needed. As the bandwidth increases, the match for such a simple circuit will get poorer; there are not enough elements available to do better. In this case, we took the center frequency ±100 MHz.

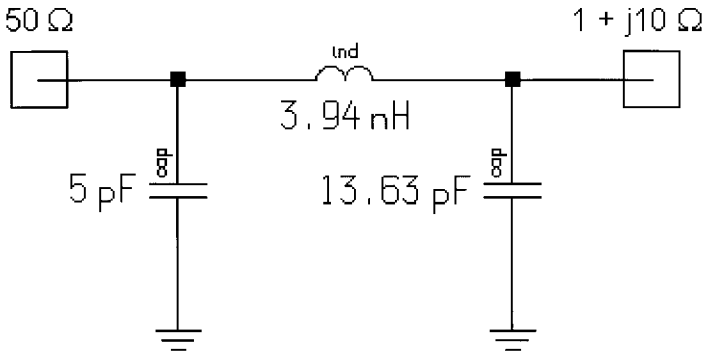


Figure 3-203 Input matching network for the amplifier used in this tutorial.

Figures 3-203 and 3-204 show the actual circuit with the element values and the resulting input match. Likewise, we did the same thing at the output, but the actual resulting dynamic output capacitance is not known *a priori*. For low frequencies, the integral equation gives a value of $2C_{CE}$; however, practical tests in frequencies above 500 MHz show that the actual increase is only 1.3 times the dc output capacitance. Therefore, we decided to design the filter with no output capacitance assumed, and the input capacitance of about 5 pF must be reduced for proper reactive matching. Mathematically, this could result in a negative number, specifically if the resulting output capacitance exceeds the 5-pF value. To compensate for this, one needs to use an RF choke that tunes out the resulting dynamic output capacitance and then one can revert to the original 5 pF as calculated.

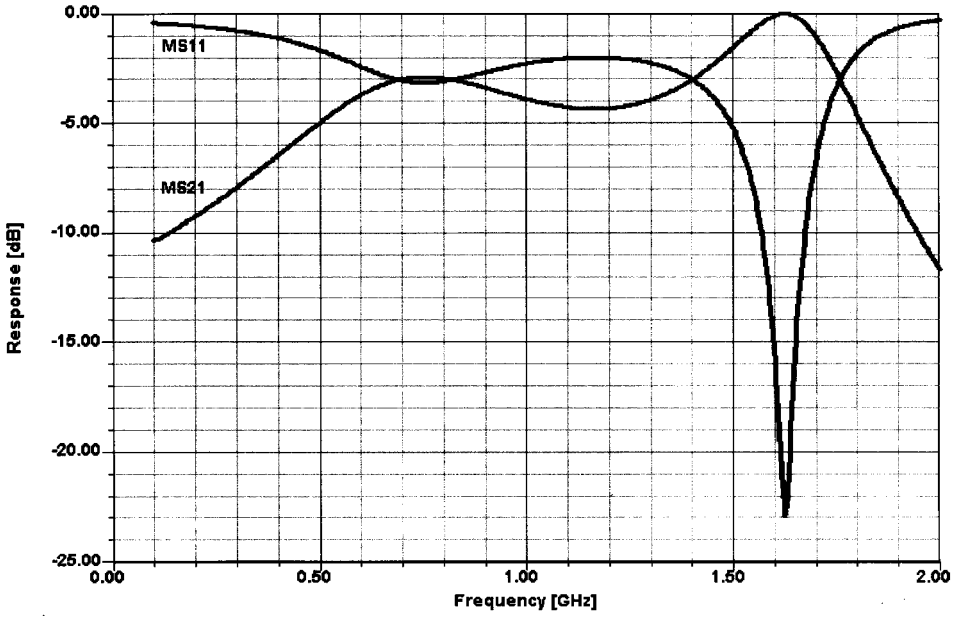


Figure 3-204 Frequency response of the input matching network.

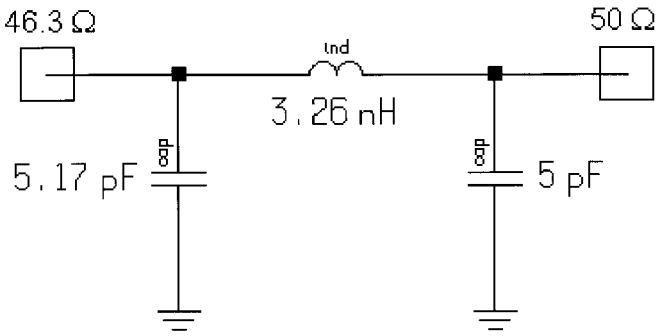


Figure 3-205 Output matching network for the amplifier.

The selection of the actual component is another issue, because these capacitors need to handle several amperes of RF current. It is not always easy to find capacitors with the appropriate value that can handle the current. The inductors can be printed. The paragraph following this section therefore will deal with the issue of how to translate the lumped values into distributed values, which will make the current handling much easier (we hope).

For this, we did the same optimizing approach by using a 50-Ω source and a 46.3-Ω load. Figures 3-205 and 3-206 show the circuit with the appropriate values and its simulated frequency-dependent responses.

The next step is to connect the input and output matching network to an "actual" transistor. This is a point where difficulties arise, since very few, if any, companies provide SPICE-type

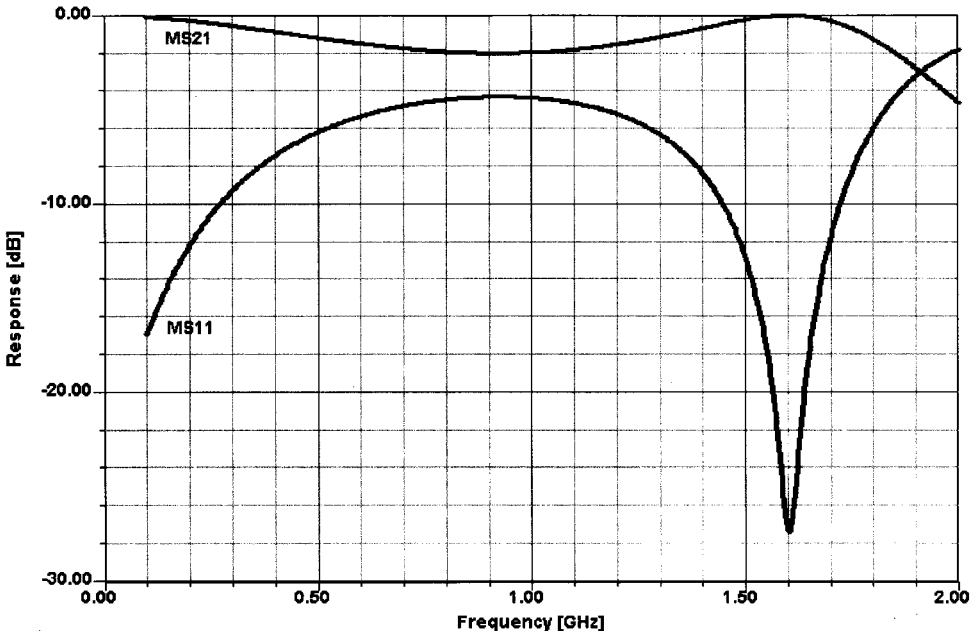


Figure 3-206 Frequency response of the output matching network.

parameters for transistors operating at currents of more than 200 mA. Figures 3-207 and 3-208 show the overall schematic and frequency response of the actual amplifier with all filters connected. Figure 3-209 details the amplifier's peak gain response. If more bandwidth is needed, the matching circuit must be increased in complexity, as will be shown in the following section.

Now is a good time to look at the final results of our simulation, which is based on the various assumptions as outlined above. Figure 3-210 shows the dc *I–V* curves with a straight load line and no reactances. This was accomplished by using a wideband transformer instead of an output matching circuit. Figure 3-211, however, shows the "load line" that results when the transformer is replaced with a real matching network. Based on the harmonics content and its improper termination, one can see how the output load line opens and shows ringing in the low-current area. If we expect to see the saturation voltage, we will be disappointed because the saturation voltage is frequency dependent (increases with frequency) and this figure gives an erroneous impression. Hence, we show dc *I–V* curves. To display the conduction angle, which we showed previously, we examine the output current as a function of time at a drive level of +16.5 dBm (Figure 3-212). During the negative half, the transistor is not conducting as predicted, so this validates our assumptions. Next, we are interested in determining the power gain. Therefore, we look at the output spectrum for a drive level of +16.5 dBm (Figure 3-213). It shows an output level of 38.6 dBm. If we subtract the drive level (16.5 dBm) from this, we see that the power gain is 22.1 dB at this drive level. Since we are used to thinking in watts, Figure 3-214 shows the absolute output power and the harmonics prior to filtering. Figure 3-215 shows the output spectrum after replacing the transformer with a simple pi output network.

During the conduction period, there will be a voltage across the transistor 180° out of phase, as shown in Figure 3-216. Finally, to highlight the nonlinearity of a Class C amplifier, Figure 3-217 shows the amplifier's output power versus drive, and Figure 3-218 shows the amplifier's gain versus drive. When the transistor turns on is clearly evident.

A particular difficulty arises in the case of FETs, where the charge energy makes for a difference between dc and pulsed measurements. Figure 3-219 shows this. Some of the more modern transistors, such as SiGe types, may need some pulsing also; as can be seen in Figure 3-219, even depending on the pulsing, one may get different dc *I–V* curves. This area is still under investigation, and even the "experts" do not agree on all aspects of this. In real life, it turns out that a lot of experimentation is the answer. Anyone who thinks that a pure CAD approach will give the right design answers immediately will encounter many surprises! This is not a fault of the CAD tools—assuming that they are properly engineered—but rather the availability of appropriately "tweaked" models for high-power applications, and the willingness of device suppliers to standardize on certain models and support them.

Because of the difficulties involved in cooling transistors under test and simulating RF pulse and CW conditions, depending on the application, manufacturers have so far stayed away from investing money in this area. The designer must therefore resort to application notes published by device manufacturers. Because even slight layer changes can have a large impact on the frequency response of a given device, application-note-based designs can run into difficulty if a device's fabrication has significantly evolved since the publication of its application notes.

Given the multitude of models outlined in Chapter 2, the CAD user is sometimes left hanging as to which is the best model to use. The latest indications are that the MEXTRAM (*m*ost *e*xquisite *t*ransistor *m*odel), developed by the University of Delft, Holland, with Philips of Holland, shows the most promising result, specifically in the area of IMD products. While

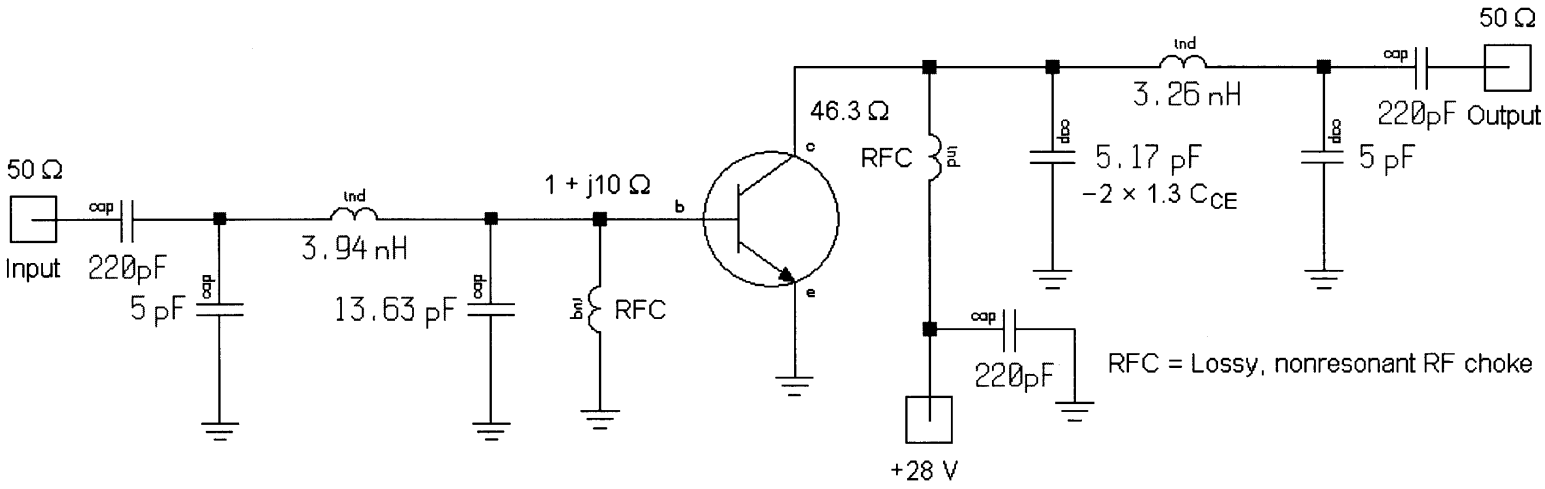


Figure 3-207 Actual schematic of our experimental 1.6-GHz amplifier. The RF chokes have enough RF losses not to be resonant and yet don't load the circuit. This is frequently accomplished with either ferrite beads or resistors in parallel with the chokes.

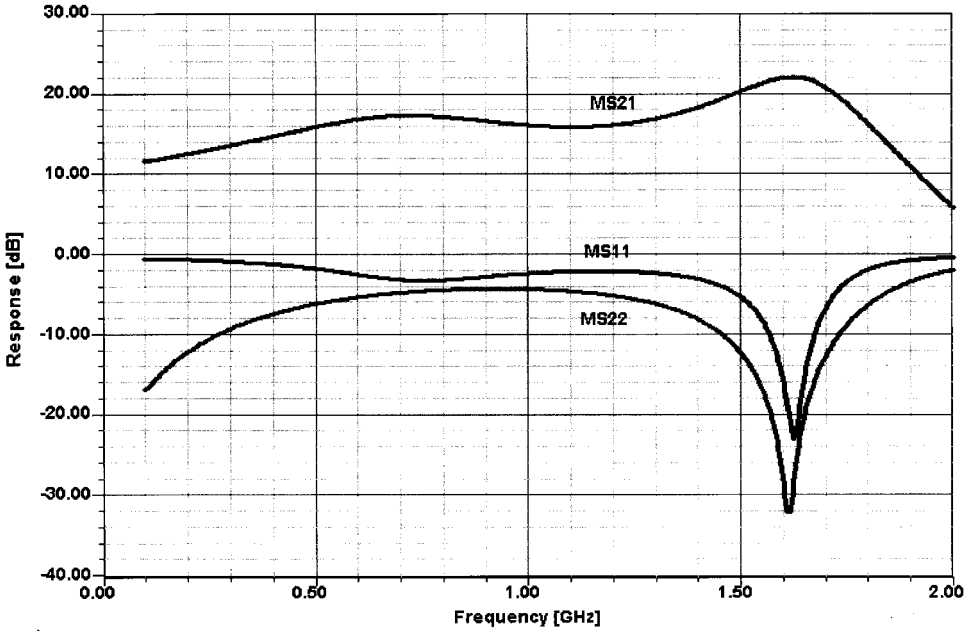


Figure 3-208 Overall frequency response of our experimental amplifier. Input and output matching, and gain, are shown as a function of frequency.

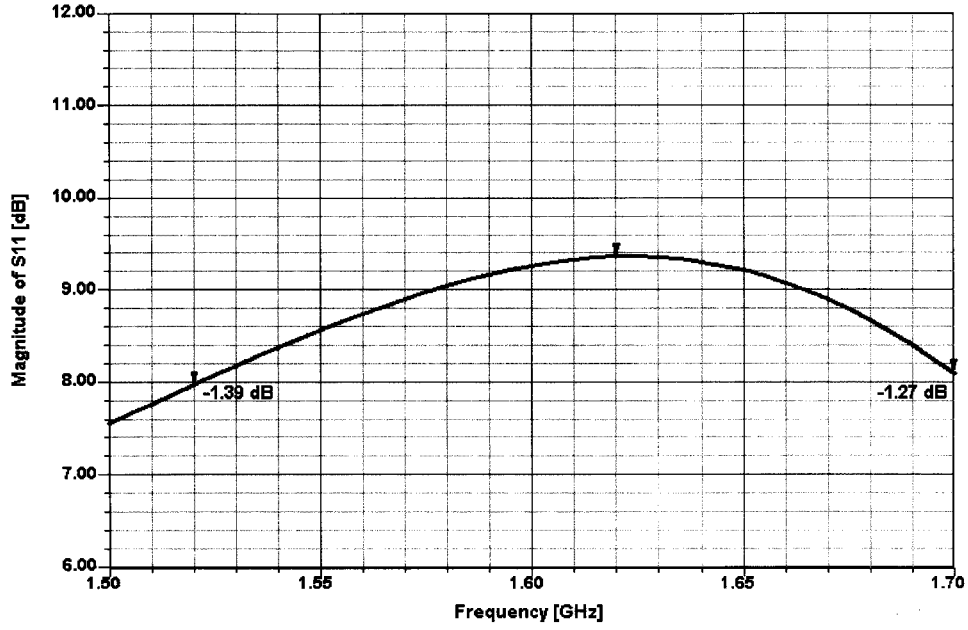


Figure 3-209 Close-in display of the frequency-dependent gain.

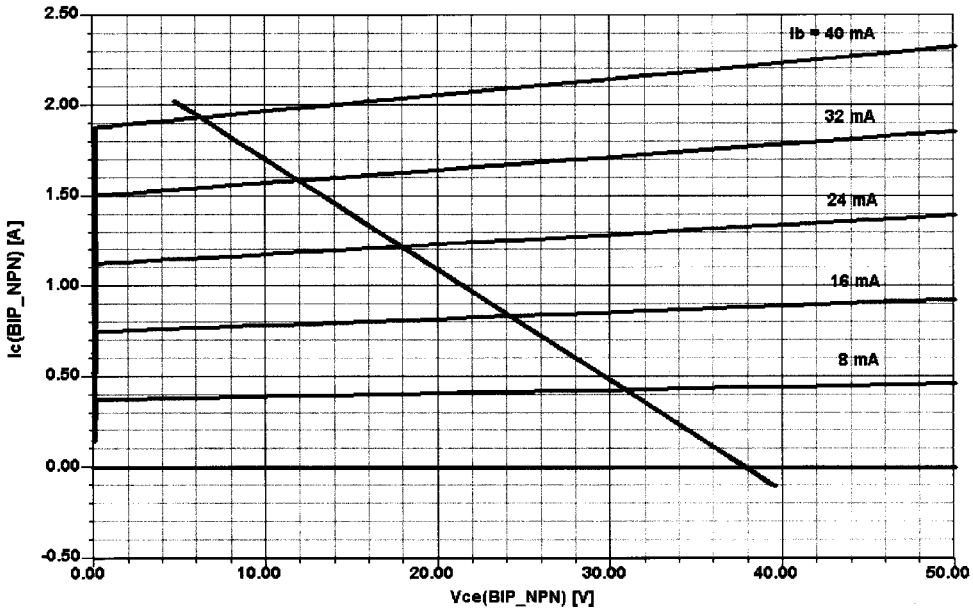


Figure 3-210 The dc *I–V* curves assuming a perfect resistive load. No RF saturation voltage or other RF saturation effects are evident.

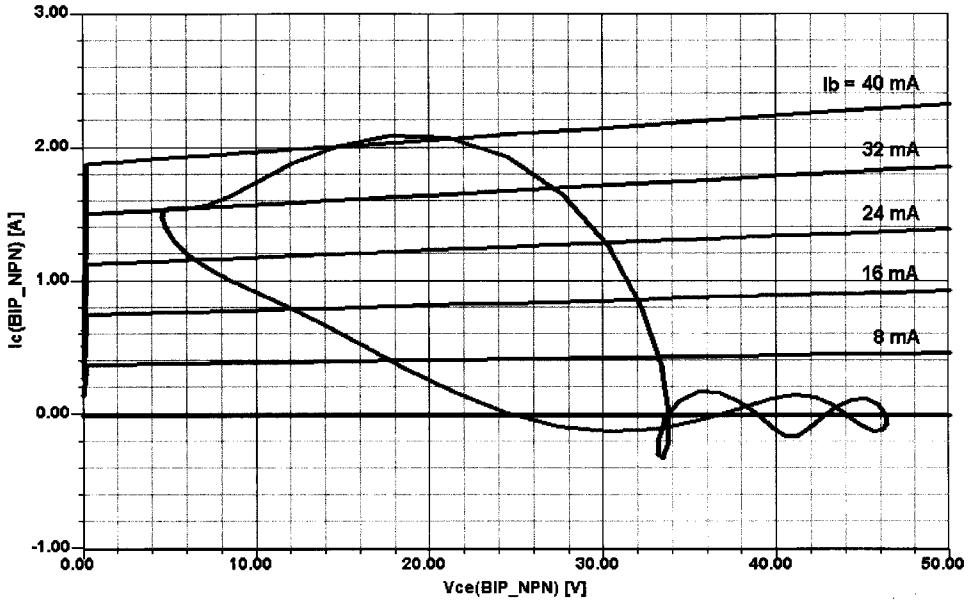


Figure 3-211 Load line with the output transformer replaced with a real matching circuit.

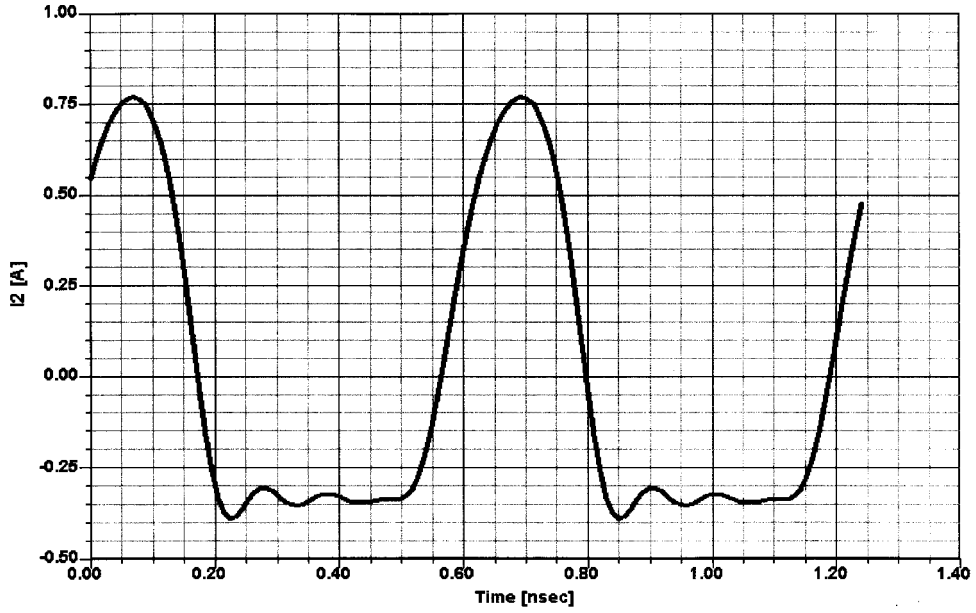


Figure 3-212 Output current as a function of time. The duty cycle is determined by the device conduction angle.

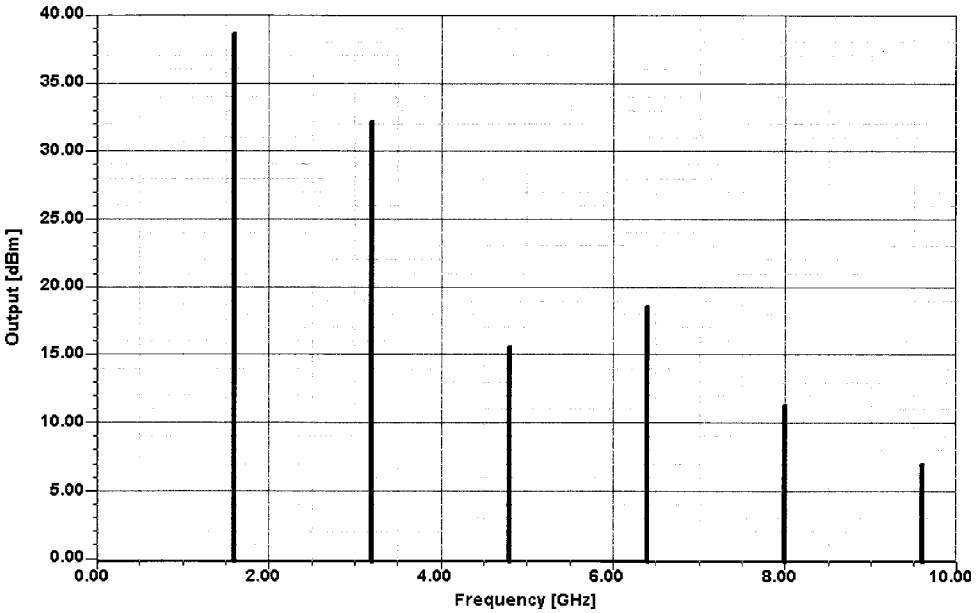


Figure 3-213 Power output at the collector prior to any filtering, assuming a real output termination.

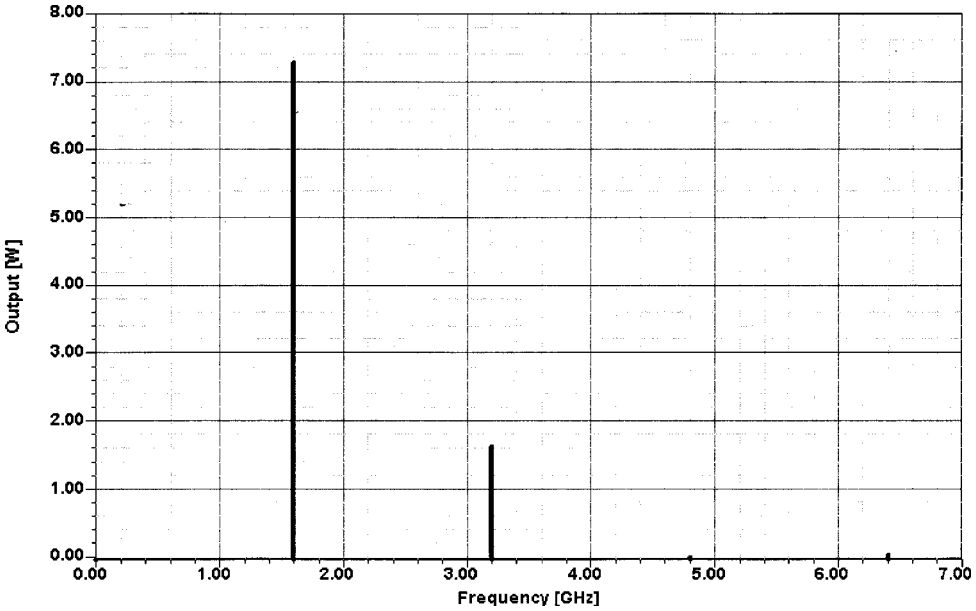


Figure 3-214 Available output power in watts, prior to any filtering, assuming a real output termination.

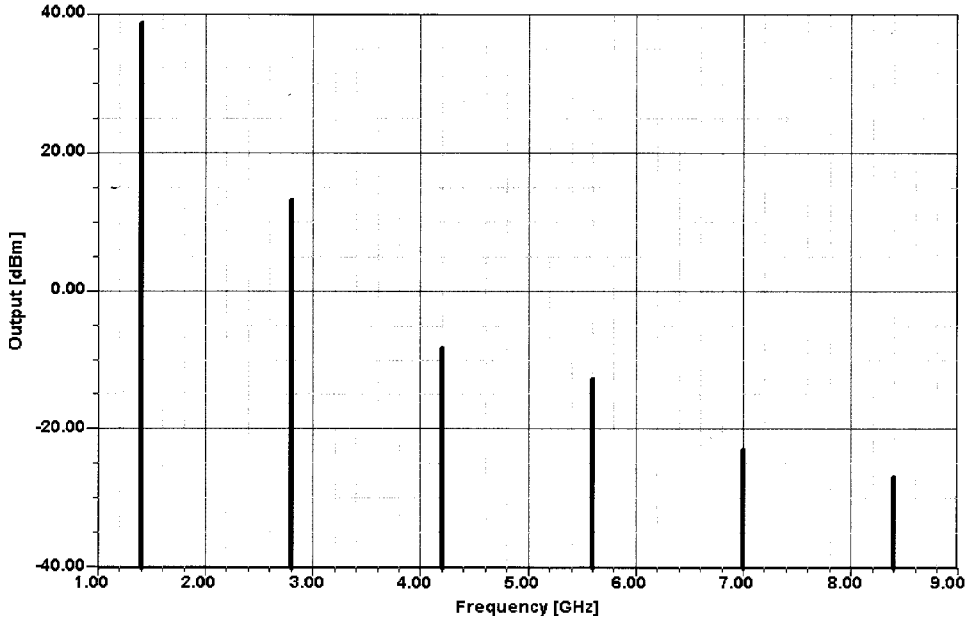


Figure 3-215 The output spectrum after replacing the transformer with a simple pi output network.

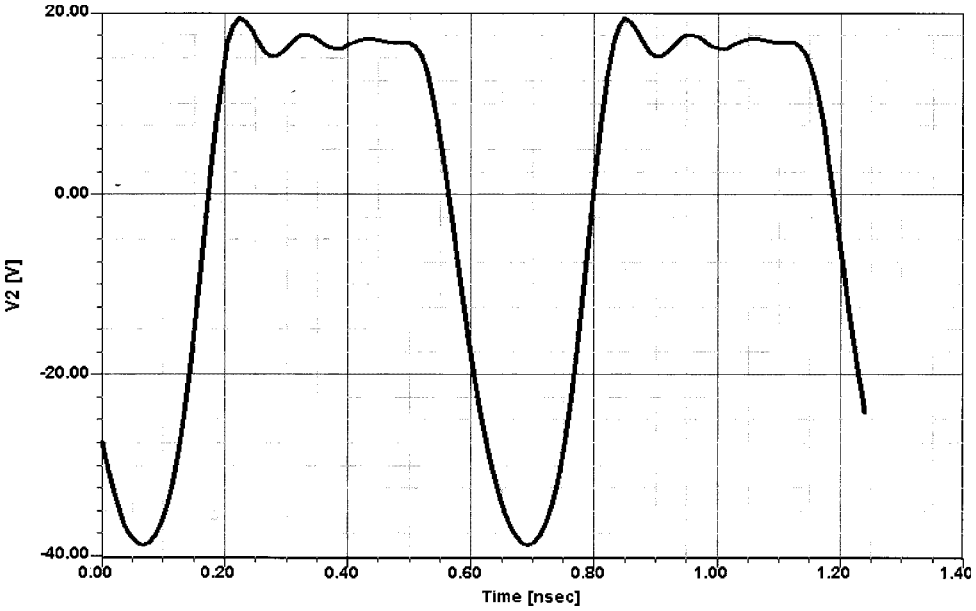


Figure 3-216 Output voltage at the collector. The voltage is asymmetrical compared to 0; this is important in the selection of a device capable of handling these voltages.

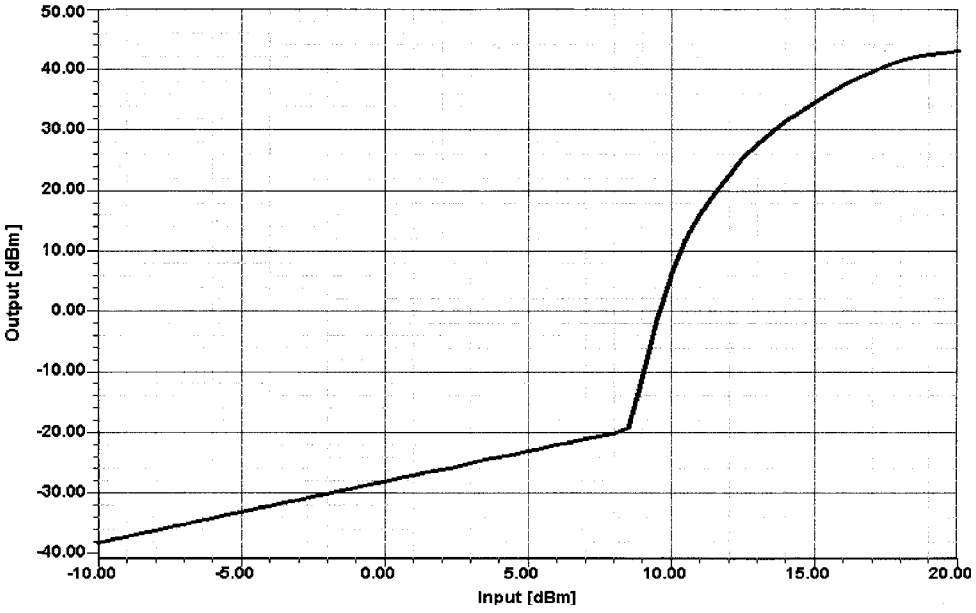


Figure 3-217 Output power versus input power for the Class C amplifier. Until the transistor turns on (at a drive level of about 8.5 dBm), only fedthrough power appears at the output.

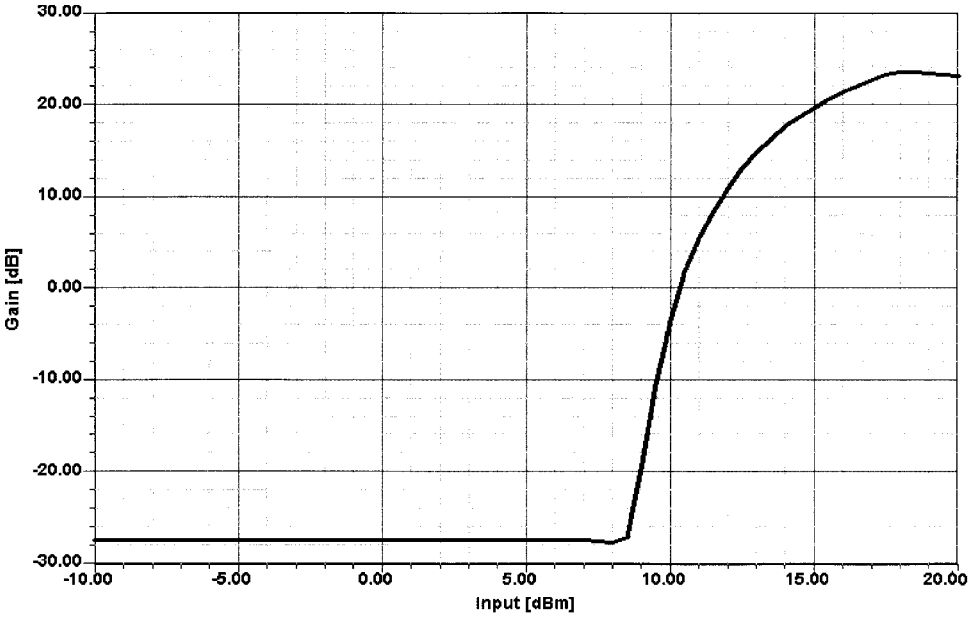


Figure 3-218 Amplifier gain as a function of drive. It is evident that one needs to cross the 0.65-V point for the transistor to become conductive and show gain. The initial amplification as a function of drive shows a very steep curve that then levels off and drops as the amplifier moves into saturation.

other later models have up to seven internal nodes, MEXTRAM fortunately uses only five. (The MEXTRAM model actually consists of three transistors, which makes modeling fun.) Power-added efficiency simulation is not available, since the MEXTRAM model was developed for small-signal applications, meaning circuits that operate at less than 100 mW of RF output power.

As if the absence of trustworthy high-power device models is not challenge enough, in developing this 1.6-GHz Class C amplifier example, we ran into another complication: Using a CAD optimizer to simultaneously match the circuit's input and output networks failed because of the device's input-impedance variation with drive. Every matching adjustment that improved the drive to the device resulted in an input-impedance change that reduced the drive to the device! In the end, the optimizer couldn't keep up. In actually building such a circuit, one would make the matching-network values adjustable on the initial breadboard and then, with the necessary input drive level applied, iteratively adjust the input and output matching for the desired output level. It appears that the CAD program could not handle such a degree of nonlinearity.

We then went to the approach of using large-signal *S* parameters at a drive level slightly less than 16 dBm, but this detuned the input filter, resulting in a poorer input match than the one initially predicted. Likewise, the output matching, which in a power stage must result in maximum output power rather than a conjugate match, was less impressive. The large-signal *S* parameters at the output for a power amplifier are somewhat dubious because Class C operation generates so many harmonics that the output termination will reflect a lot of energy and, depending on the output filter topology, different results can be expected. We firmly believe that there will be further mathematical improvement in the CAD tools, and we very much invite

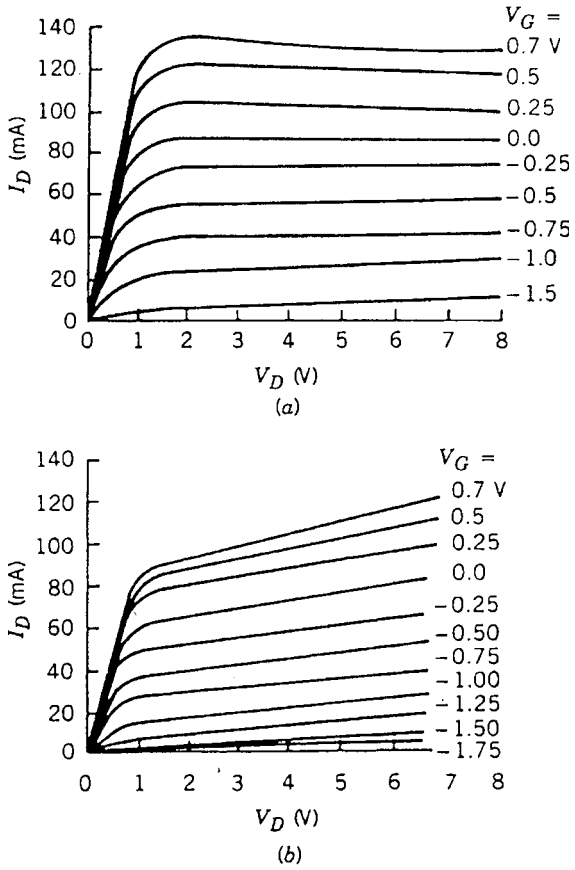


Figure 3-219 Comparison of FET drain current characteristics obtained from (a) pulsed *I–V* and (b) rectified sine-wave evaluation. The rectified sine-wave curves can be considered equivalent to curves obtained at dc.

our readers to use such a nonlinear example to test their tools, provided that a high-power model becomes available in this power range.

It has been rumored that the load-pull technique, rather than the application of large-signal *S* parameters, is a vehicle to solve such matching problems. However, the load-pull technique is really applied only to the fundamental frequency and does not deal with the issue of harmonics of the output current. More details about the load-pull technique and its capabilities are found in Appendix B.

Actual measurements of amplifiers of this type show that our simulated efficiency and drive level come quite close to reality, but at a frequency somewhat offset from that chosen for the design. As an example, the input and output filters tuned at 1.4 GHz instead of 1.6 GHz using the large-signal *S* parameters. Again, an iterative process, similar to manually trimming a prototype, will give a similar result. This is consistent with our earlier statement that most final designs are really hand-tweaked, and that CAD can only bring us close, but not all the way, to a good solution. For passive problems, the CAD does a significantly better job; the nonlinear cases are the ones that still cause headaches.

3-12-2 Impedance Matching Networks Applied to RF Power Transistors*

Introduction. Some graphical and numerical methods of impedance matching will be reviewed here. The examples given refer to high-frequency power amplifiers.

Although matching networks normally take the form of filters and therefore are also useful to provide frequency discrimination, this aspect will be considered only as a corollary of the matching circuit.

Matching is necessary for the best possible energy transfer from stage to stage. In RF power transistors, the input impedance is of low value, decreasing as the power increases or as the chip size becomes larger. This impedance must be matched either to a generator—of generally 50-Ω internal impedance—or to a preceding stage. Impedance matching has to be made between two complex impedances, which makes the design still more difficult, especially if matching must be accomplished over a wide frequency band.

Device Parameters

Input Impedance. The general shape of the input impedance of RF power transistors is as shown in Figure 3-220. It is a large-signal parameter, expressed here by the parallel combination of a resistance R_p and a reactance X_p [31].

The equivalent circuit shown in Figure 3-221 accounts for the behavior illustrated in Figure 3-220. With the presently used stripline or flange packaging, most of the power devices for VHF low band will have their R_p and X_p values below the series resonant point f_s. The input impedance will be essentially capacitive.

Most of the VHF high-band transistors will have the series-resonant frequency within their operating range; that is, they will be purely resistive at one single frequency f_s, which the parallel-resonant frequency f_p will be outside.

Parameters for 1- or 2-GHz transistors will be beyond f_s and approach f_p. They show a high value of R_p and X_p with inductive character.

A parameter that is very often used to judge on the broadband capabilities of a device is the input Q or Q_{IN}, defined simply as the ratio R_p/X_p. Practically, Q_{IN} ranges around 1 or less for VHF devices and around 5 or more for microwave transistors.

Q_{IN} is an important parameter to consider for broadband matching. Matching networks normally are low-pass or pseudo-low-pass filters. If Q_{IN} is high, it can be necessary to use bandpass filter-type matching networks and to allow insertion losses. But broadband matching is still possible. This will be discussed later.

Output Impedance. The output impedance of RF power transistors, as given by all manufacturers' datasheets, generally consists of only a capacitance C_{OUT}. The internal resistance of the transistor is supposed to be much higher than the load and is normally neglected. In the case of a relatively low internal resistance, the efficiency of the device would decrease by the factor

$$1 + R_L/R_T \tag{3-192}$$

*Based on Motorola application note AN-721 ("Impedance Matching Networks Applied to RF Power Transistors"). Copyright of Motorola; used by permission.

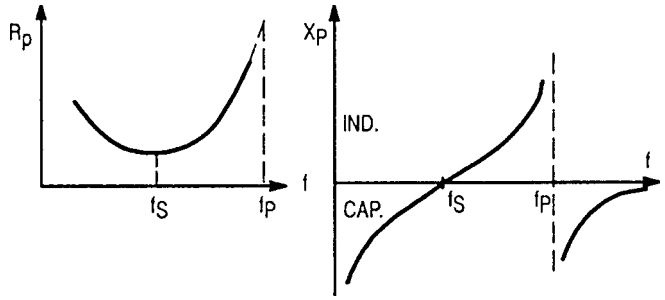


Figure 3-220 Input impedance of RF power transistors as a function of frequency.

where R_L is the load resistance, seen at the collector–emitter terminals, and R_T is the internal transistor resistance equal to

$$\frac{1}{\omega_T (C_{TC} + C_{DC})} \tag{3-193}$$

defined as a single parameter, where ω_T is the transit angular frequency and C_{TC} and C_{DC} are the transition and diffusion capacitances, respectively, at the collector junction.

The output capacitance, C_{OUT}, which is a large-signal parameter, is related to the small-signal parameter C_{CB}, the collector–base transition capacitance.

Since a junction capacitance varies with the applied voltage, C_{OUT} differs from C_{CB} in that it has to be averaged over the total voltage swing. For an abrupt junction and assuming certain simplifications, $C_{\text{OUT}} = 2C_{CB}$.

Figure 3-222 shows the variation of C_{OUT} with frequency. C_{OUT} decreases partly due to the presence of the collector lead inductance, but mainly because of the fact that the base–emitter diode does not shut off anymore when the operating frequency approaches the transit frequency, f_T.

Output Load. In the absence of a more precise indication, the output load R_L is taken as

$$R_L = \frac{[V_{CC} - V_{CE(\text{sat})}]^2}{2P_{\text{OUT}}} \tag{3-194}$$

with $V_{CE(\text{sat})}$ equal to 2 or 3 V, increasing with frequency.

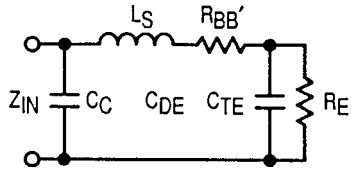


Figure 3-221 Equivalent circuit for the input impedance. R_E is the emitter diffusion resistance; C_{DE} and C_{TE} are the diffusion and transition capacitances, respectively, of the emitter junction; $R_{BB'}$ is the base-spreading resistance; C_C is the package capacitance; and L_S is the base lead inductance.

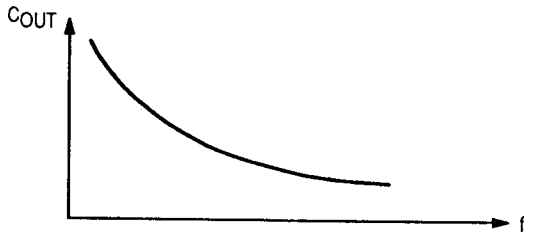


Figure 3-222 Output capacitance C_{OUT} as a function of frequency.

The above equation just expresses a well-known relation, but it also shows that the load, in first approximation, is not related to the device, except for $V_{CE(sat)}$. The load value is primarily dictated by the required output power and the peak voltage; it is not matched to the output impedance of the device.

At high frequencies this approximation becomes less exact, and for microwave devices the load that must be presented to the device is indicated on the datasheet.

Strictly speaking, impedance matching is accomplished only at the input. Interstage and load matching are more impedance transformations of the device input impedance and of the load into a value R_L (sometimes with additional reactive component) that depends essentially on the power demanded and the supply voltage.

Matching Networks. In the following, matching networks will be described by the order of complexity. These are ladder-type reactance networks. The different reactance values will be calculated and determined graphically. Increasing the number of reactances broadens the bandwidth. However, matching networks consisting of more than four reactances are rare. Above four reactances, the improvement is small.

Numerical Design

TWO-RESISTANCE NETWORKS. Resistance terminations will first be considered. Figure 3-223 shows the reactive L section and the terminations to be matched.

Matching or exact transformation from R_2 into R_1 occurs at a single frequency f_0. At f_0, X_1 and X_2 are equal to

$$X_1 = \pm R_1 \sqrt{\frac{R_2}{R_1 - R_2}} = R_1 \frac{1}{\sqrt{n-1}} \tag{3-195}$$

$$X_2 = \mp \sqrt{R_2(R_1 - R_{2)}} = R_1 \frac{\sqrt{n-1}}{1} \tag{3-196}$$

At f_0, $X_1X_2 = R_1R_2$. X_1 and X_2 must be of opposite sign. The shunt reactance is in parallel with the larger resistance.

The frequency response of the L section is shown in Figure 3-224, where the normalized current is plotted as a function of the normalized frequency.

If X_1 is capacitive and consequently X_2 is inductive, then

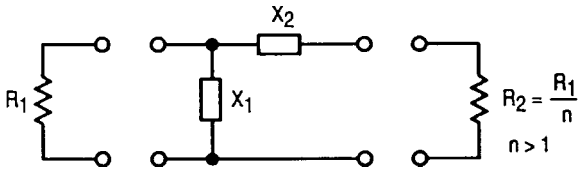


Figure 3-223 Two-reactance matching network.

$$X_1 = -\frac{f_0}{f} R_1 \sqrt{\frac{R_2}{R_1 - R_2}} = -\frac{f_0}{f} R_1 \frac{1}{\sqrt{n-1}} \tag{3-197}$$

and

$$X_2 = \frac{f}{f_0} \sqrt{R_2(R_1 - R_2)} = \frac{f}{f_0} R_1 \frac{\sqrt{n-1}}{1} \tag{3-198}$$

The normalized current absolute value is equal to

$$\left|\frac{I_2}{I_0}\right| = \frac{2\sqrt{n}}{\sqrt{(n-1)^2 \cdot (f/f_0)^4 - 2\,(f/f_0)^2 + (n+1)^2}} \tag{3-199}$$

where $I_0 = \sqrt{n}E/2R_1$ and is plotted in Figure 3-224 [32].

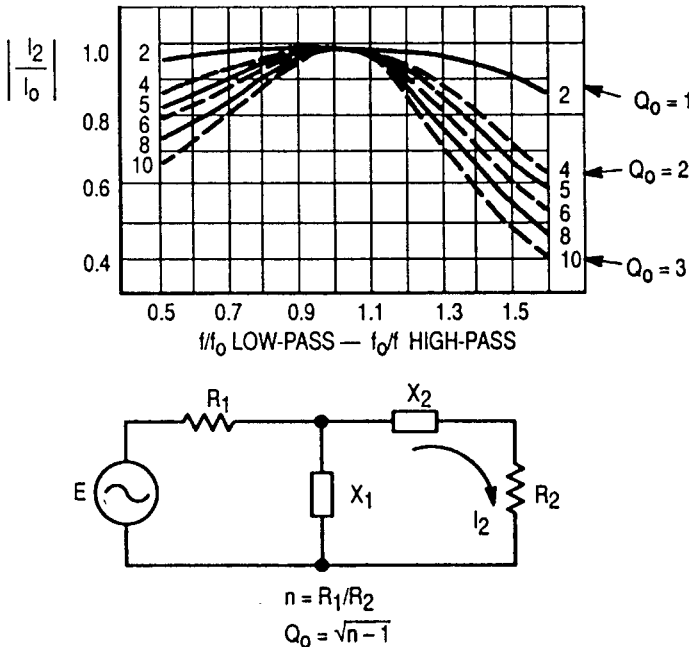


Figure 3-224 Normalized frequency response for the L section in low-pass or high-pass form.

If X_1 is inductive and consequently X_2 is capacitive, the only change required is a replacement of f by f_0 and vice versa. The L section has low-pass form in the first case and high-pass form in the second case.

The Q of the circuit at f_0 is equal to

$$Q_0 = \frac{X_2}{R_2} = \frac{R_1}{X_1} = \sqrt{n-1} \tag{3-200}$$

For a given transformation ratio n, there is only one possible value of Q. On the other hand, there are two symmetrical solutions for the network that can be either a low-pass filter or a high-pass filter.

The frequency f_0 does not need to be the center frequency, $(f_1 + f_2) / 2$, of the desired band limited by f_1 and f_2. In fact, as can be seen from the low-pass configuration of Figure 3-224, it may be interesting to shift f_0 toward the high-band edge frequency f_2 to obtain a larger bandwidth w, where

$$w = \frac{2(f_1 + f_2)}{f_2 - f_1} \tag{3-201}$$

This will, however, be at the expense of poor harmonic rejection.

Example. For a transformation ratio of $n = 4$, we can determine the following values from the above relations:

Bandwidth w	0.1	0.3
Maximum insertion losses	0.025	0.2
X_1/R_1	1.730	1.712

If the terminations R_1 and R_2 have a reactive component X, the latter may be taken as part of the external resistance as shown in Figure 3-225. This compensation is applicable as long as

$$Q_{int} = \frac{X_{\text{int}}}{R_2} \quad \text{or} \quad \frac{R_1}{X_{\text{int}}} < n - 1 \tag{3-202}$$

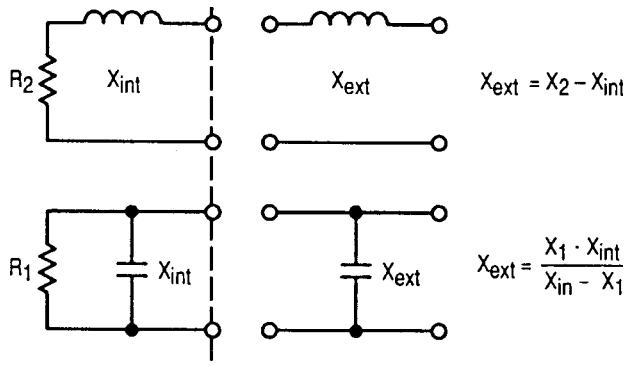


Figure 3-225 Termination reactance compensation.

Tables giving reactance values can be found in Matthaei et al. [33, 34].

USE OF TRANSMISSION LINES AND INDUCTORS. In the preceding section, the inductance was expected to be realized by a lumped element. A transmission line can be used instead (Figure 3-226).

As can be seen from the computed selectivity curves (Figure 3-227) for the two configurations, transmission lines result in a larger bandwidth. The gain is important for a transmission line having a length $L = \lambda/4$ ($\Theta = 90°$) and a characteristic impedance $Z_0 = \sqrt{R_1 R_2}$. It is not significant for lines short with respect to $\lambda/4$. One will notice that there is an infinity of solutions, one for each value of C, when using transmission lines.

THREE-REACTANCE MATCHING NETWORKS. The networks that will be investigated are shown in Figure 3-228. They are made of three reactances alternately connected in series and shunt.

A three-reactance configuration allows the designer to make the quality factor (Q) of the circuit and the transformation ratio $n = R_2/R_1$ independent of each other and consequently to choose the selectivity between certain limits.

For narrowband designs, one can use the following formulas (see [35], where tables are given):

Network (a)

$$X_{C1} = R_1/Q \quad Q \text{ must be selected first} \tag{3-203}$$

$$X_{C2} = R_2 \sqrt{\frac{R_1 R_2}{(Q^2 + 1) - R_1/R_2}} \tag{3-204}$$

$$X_L = \frac{QR_1 + (R_1 R_2/X_{C2})}{Q^2 + 1} \tag{3-205}$$

Network (b)

$$X_{L1} = R_1 Q \quad Q \text{ must be selected first} \tag{3-206}$$

$$X_{L2} = R_2 B \tag{3-207}$$

$$A = R_1(1 + Q^2) \tag{3-208}$$

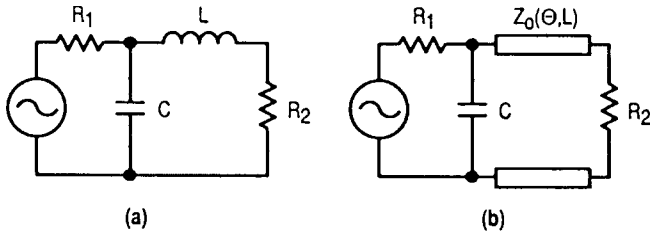


Figure 3-226 Use of a transmission line in the L section.

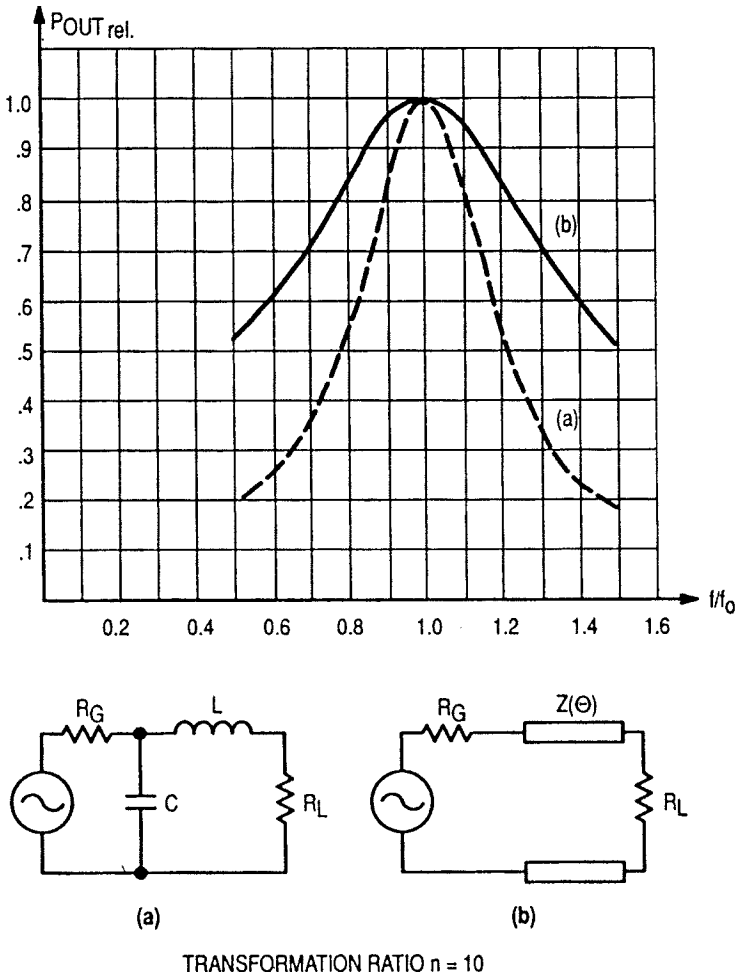


Figure 3-227 Bandwidth of the L section for $n = 10$ (a) with lumped constants and (b) with transmission line (λ/4).

$$X_{C1} = \frac{A}{Q + B} \tag{3-209}$$

$$B = \sqrt{\frac{A}{R_2} - 1} \tag{3-210}$$

Network (c)

$$X_{L1} = QR_1 \quad Q \text{ must be selected first} \tag{3-211}$$

$$X_{C2} = AR_2 \tag{3-212}$$

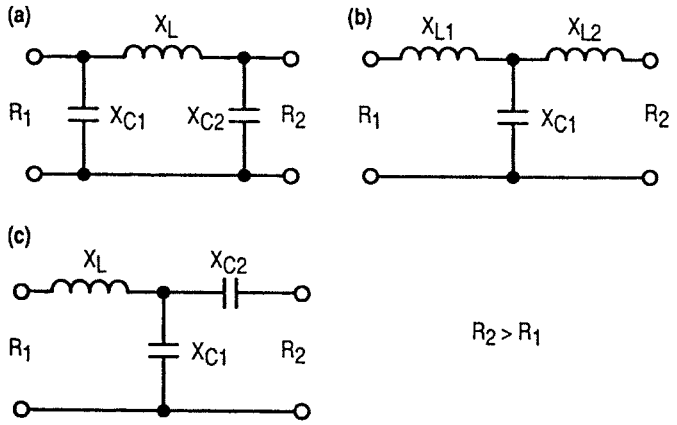


Figure 3-228 Three-reactance matching networks.

$$A = \sqrt{\frac{R_1(1+Q^2)}{R_2} - 1} \tag{3-213}$$

$$X_{C1} = \frac{B}{Q - A} \tag{3-214}$$

$$B = R_1(1+Q^2) \tag{3-215}$$

The network that yields the most-practical component values should be selected for a given application.

The three-reactance networks can be thought of as being formed of an L section (two reactances) and of a compensation reactance. The L section essentially performs the impedance transformation, while the additional reactance compensates for the reactive part of the transformed impedance over a certain frequency band.

Figure 3-229 shows a representation in the Z plane of the circuit of Figure 3-228a split into two parts: $R_1 - C_1 - L_1$ and $C_2 - R_2$. Exact transformation from R_1 into R_2 occurs at the points of intersection M and N. Impedances are then conjugate, or $Z' = R' + jX'$ and $Z'' = R'' + jX''$ with $R' = R''$ and $X' = -X''$.

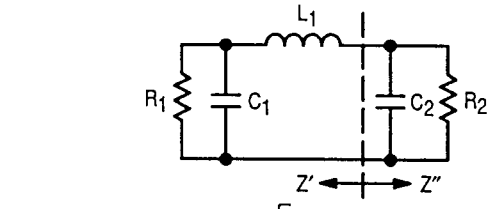
The only possible solution is obtained when X' and $-X''$ are tangential to each other. For the dashed curve, representing another value of L_1 or C_1, a wider frequency band could be expected at the expense of some ripple inside the band. However, this can only be reached with four reactances, as will be shown below.

With a three-reactance configuration, there are not enough degrees of freedom to permit $X' = -X''$ and simultaneously obtain the same variation of frequency on both curves from point M' to point N'. Exact transformation can, therefore, be obtained at only one frequency.

The values of the three reactances can be calculated by making

$$X' = -X'', \quad R' = R'', \quad \text{and} \frac{dX'}{dR'} = -\frac{dX''}{dR''} \tag{3-216}$$

The general solution of these equations leads to complicated calculations. Therefore, computed tables should be used.



$$Z' = \frac{R_1}{1 + \omega^2 R_1^2 C_1^2} + j\omega \left[L1 - \frac{R_1^2 C_1}{1 + \omega^2 R_1^2 C_1^2} \right] = R' + jX'$$

$$Z'' = \frac{R_2}{1 + \omega^2 C_2^2 R_2^2} - \frac{j\omega^2 C_2 R_2^2}{1 + \omega^2 C_2^2 R_2^2} = R'' + jX''$$

$$|x'| = \left(\frac{L_1}{C_1 R_1^2} - \frac{R'}{R_1} \right) \sqrt{\frac{R_1}{R'} - 1}$$

$$|x''| = \sqrt{R' R_2 - R'^2}$$

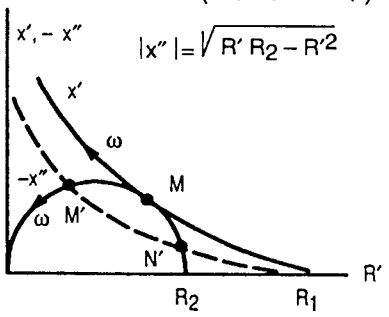


Figure 3-229 *Z*-plane representation of the circuit of Figure 3-228a.

One will note from Figure 3-229 that the compensation reactance contributes somewhat to the impedance transformation; that is, R' varies when going from M to R_2.

The circuit of Figure 3-228b is dual with respect to the first one and gives exactly the same results in a *Y*-plane representation. The circuit of Figure 3-228c is somewhat different since only one intersection M exists, as shown in Figure 3-230. Narrower frequency bands must be expected from this configuration. The widest band is obtained for $C_1 = \infty$.

Again, if one of the terminations has a reactive component, the latter can be taken as a part of the matching network, provided that it is not too large (see Figure 3-227).

FOUR-REACTANCE NETWORKS. Four-reactance networks are used essentially for broadband matching. The networks that will be considered in the following consist of two two-reactance sections in cascade. Some networks have pseudo-low-pass filter character; others, bandpass filter character. In principle, the former show narrower bandwidth since they extend the impedance transformation to very low frequencies unnecessarily, while the latter ensure good matching over a wide frequency band around the center frequency only (see Figure 3-233).

The two-reactance sections used in the networks in Figure 3-231 have either transformation properties or compensation properties. Impedance transformation is obtained with one series reactance and one shunt reactance. Compensation is made with both reactances in series or in shunt. If two cascaded transformation networks are used, transformation is accomplished partly by each one.

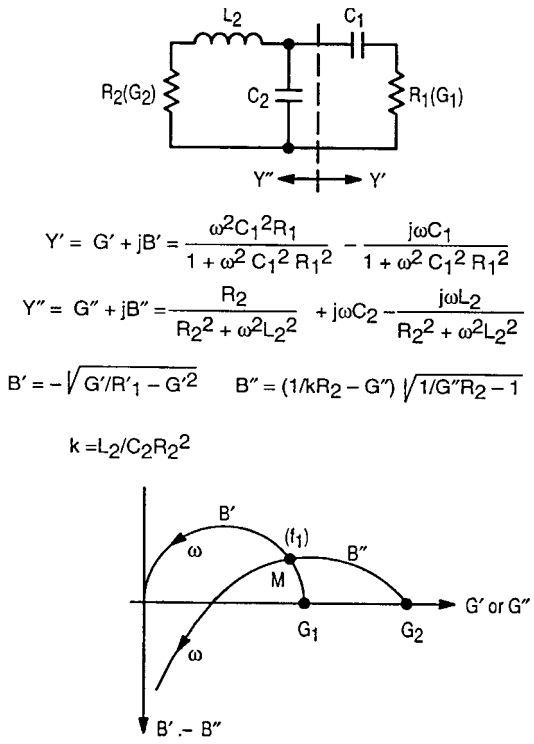


Figure 3-230 *Y*-plane representation of the circuit of Figure 3-228c.

With four-reactance networks there are two frequencies, f_1 and f_2, at which the transformation from R_1 into R_2 is exact. These frequencies may also coincide. For the network in Figure 3-231b, for instance, at point M, R_1 or R_2 is transformed into $\sqrt{R_1R_2}$ when both frequencies fall together. At all points (M), Z_1 and Z_2 are conjugate if the transformation is exact.

In the case of Figure 3-231b, the reactances are easily calculated for equal frequencies:

$$X_1 = \frac{R_1}{\sqrt{(n-1)^{1/2}}} \tag{3-217}$$

$$X_2 = R_1\sqrt{\frac{n^{1/2}-1}{n}} \tag{3-218}$$

$$X_1X_4 = R_1R_2 = X_2X_3 \tag{3-219}$$

$$X_3 = \frac{R_1}{\sqrt{n(n^{1/2}-1)}} \tag{3-220}$$

$$X_4 = \frac{R_1}{n}\sqrt{n^{1/2}-1} \tag{3-221}$$

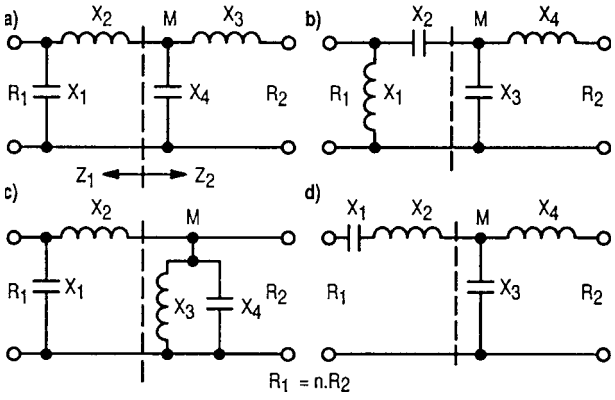


Figure 3-231 Four-reactance networks.

For the network in Figure 3-231a, normally at point M, Z_1 and Z_2 are complex. This pseudo-low-pass filter has been computed elsewhere [33]. Many tables can be found in the literature for networks of four and more reactances having Chebyshev character or maximally flat response [33, 34, 36].

Figure 3-232 shows the transformation path from R_1 to R_2 for networks (a) and (b) in Figure 3-231 on a Smith Chart.

Case (a) of Figure 3-232 has been calculated using tables in Matthaei [34]. Case (b) has been obtained from the relationship given above for $X_1, \ldots, X_4$. Both apply for a transformation ratio equal to 10 and for $R_1 = 1$. There is no simple relationship for $X_1', \ldots, X_4'$ of network (b) if f_1 is made different from f_2 for larger bandwidth. Figure 3-233 shows the respective bandwidths of networks (a) and (b) for the circuits shown in Figure 3-232.

If the terminations include a reactive component, the computed values for X_1 or X_4 may be adjusted to compensate for this.

For configuration (a), it can be seen from Figure 3-232 that, in the considered case, the Q values are equal to 1.6. For configuration (b), Q_1', which is equal to Q_2', is fixed for each transformation ratio

n	2	4	8	10	16
$Q'_1 = Q'_2$	0.65	1	1.35	1.46	1.73

$$Q' = \sqrt{n^{1/2} - 1} \qquad (3\text{-}222)$$

The maximum value of reactance that the terminations may have for use in this configuration can be determined from the above values of Q'.

If R_1 is the load resistance of a transistor, the internal transistor resistance may not be equal to R_1. In this case, the selectivity curve will be different from the curves given in Figure 3-233. Figure 3-234 shows the selectivity for networks (a) and (b) when the source resistance R_1 is infinite. From Figure 3-234 it can be seen that network (a) is more sensitive to changes in R_1 than network (b).

As mentioned earlier, the four-reactance network can also be thought of as two cascaded two-reactance sections: one used for transformation, the other for compensation. Figure 3-235 shows commonly used compensation networks, together with the associated L section.

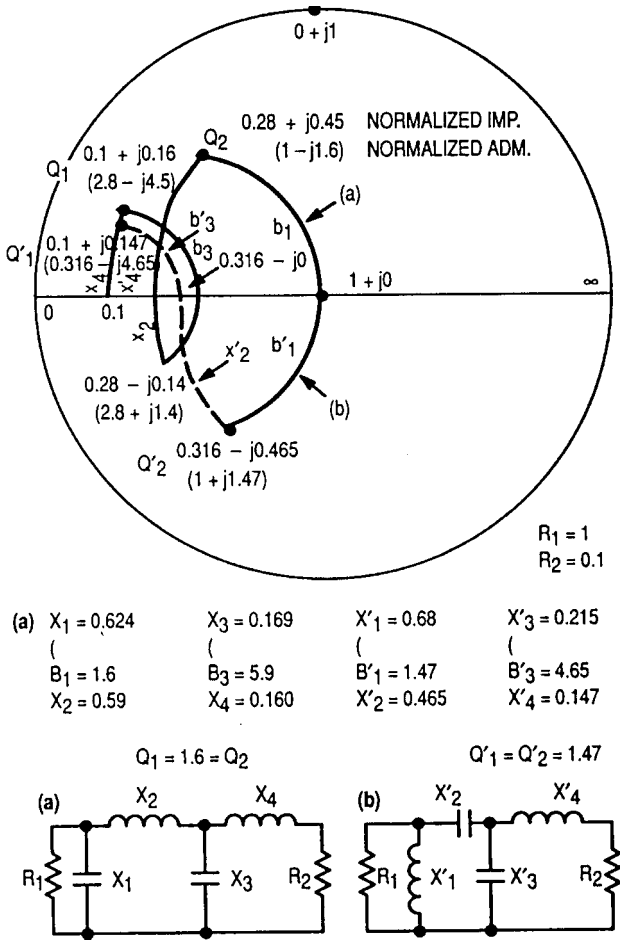


Figure 3-232 Transformation paths for networks (a) and (b).

The circuit of Figure 3-235a can be compared to the three-reactance network shown in Figure 3-228c. The difference is that capacitor C_2 of that circuit has been replaced by an *LC* circuit. The resulting improvement can be seen by comparing Figure 3-236 with Figure 3-230.

By adding one reactance, exact impedance transformation is achieved at two frequencies. It is now possible to choose component values such that the point of intersection M' occurs at the same frequency f_1 on both curves and simultaneously that N' occurs at the same frequency f_2 on both curves. Among the infinite number of possible intersections, only one allows the achievement of this.

When M' and N' coincide in M, the new $dX'/df = dX''/df$ condition can be added to the condition $X' = -X''$ (for three-reactance networks) and similarly $R' = R''$ and $dR'/df = dR''/df$.

Again, a general solution of the above equations leads to still more complicated calculations than in the case of three-reactance networks. Therefore, tables are preferable [33, 34, 36].

The circuit of Figure 3-235b is the dual of the circuit of Figure 3-233a and does not need to be treated separately. It gives exactly the same results in the *Z* plane. Figure 3-235c shows a higher-order compensation requiring six reactive elements.

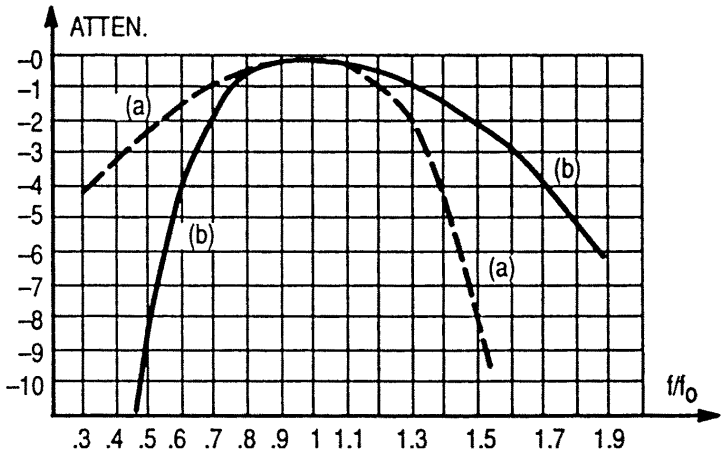


Figure 3-233 Selectivity curves for networks (a) and (b) of Figure 3-232.

The above-discussed matching networks employing compensation circuits result in narrower bandwidths than the former solutions (see p. 573) using two transformation sections. A matching arrangement with higher-order compensation, such as in Figure 3-235c, is not recommended. Better use can be made of the large number of reactive elements by using them all for transformation.

When the above configurations are realized using short portions of transmission lines, the equations or the usual tables no longer apply. The calculations must be carried out on a computer because of their complexity. However, a graphical method can be used (see the next section) that will consist essentially in tracing a transformation path on the Z–Y chart using the computed lumped element values and replacing it by the closest path obtained with distributed constants. The bandwidth change is not significant as long as short portions of lines are used.

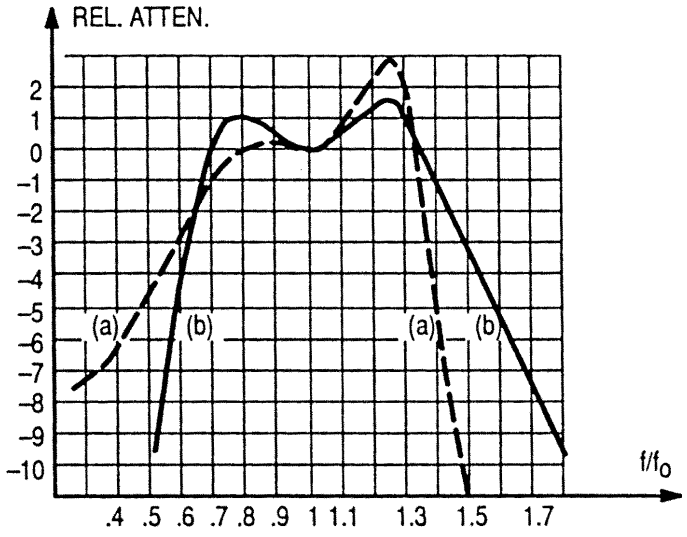


Figure 3-234 Selectivity curves for networks (a) and (b) of Figure 3-232 with infinite R_1.

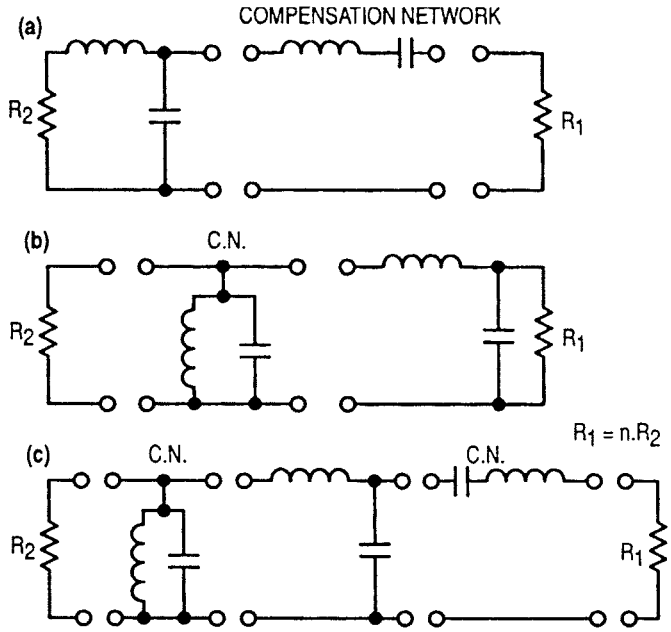


Figure 3-235 Compensation networks used with an L section.

MATCHING NETWORKS USING QUARTER-WAVE TRANSFORMERS. At sufficiently high frequencies, where λ/4-long lines of practical size can be realized, broadband transformations can easily be accomplished by the use of one or more λ/4 sections. Figure 3-237 summarizes the main relations for (a) one-section and (b) two-section transformations. A compensation network can be realized using a λ/2-long transmission line. Figures 3-238 and 3-239 show the selectivity curves for different transformation ratios and section numbers.

EXPONENTIAL LINES. Exponential lines have largely frequency-independent transformation properties. The characteristic impedance of such lines varies exponentially with their length l:

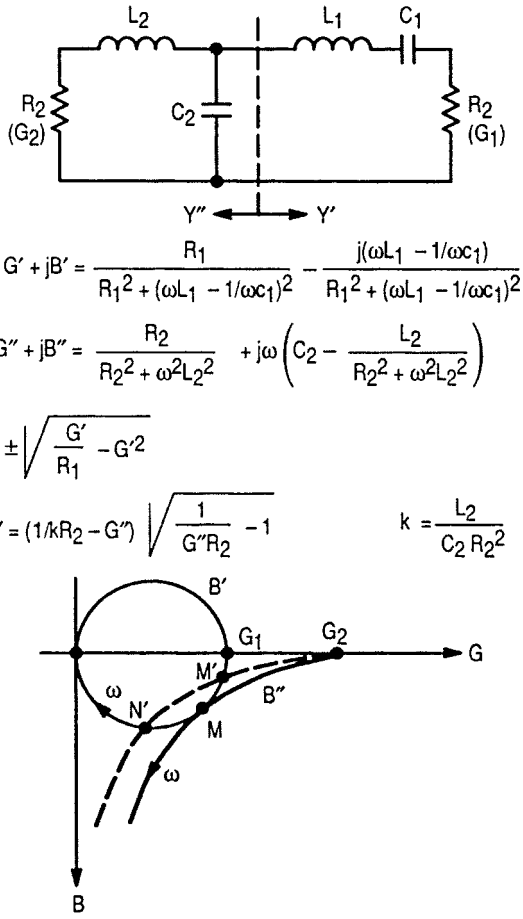
$$Z = Z_0 e^{kl} \tag{3-223}$$

where k is a constant, but these properties are preserved only if k is small.

BROADBAND MATCHING USING BANDPASS FILTER NETWORKS—HIGHQ CASE. The above circuits are applicable to devices having low input or output Q, if broadband matching is required. Generally, if the impedances to be matched can be represented, for instance, by a resistor R in series with an inductor L (sometimes a capacitor C) within the band of interest and if L is sufficiently low, the latter can be incorporated into the first inductor in the matching network. This is also valid if the representation consists of a shunt combination of a resistor and a reactance.

Practically, this is feasible for Q values around 1 or 2. For higher Q values or for input impedances consisting of a series or parallel resonant circuit (see Figure 3-221), as it appears to be for large bandwidths, a different treatment must be followed.

Let us first recall that, as shown by Bode [37] and Fano [38], limitations exist on the impedance matching of a complex load. In the example of Figure 3-240, the load to be matched consists of a capacitor C and a resistor R in shunt.



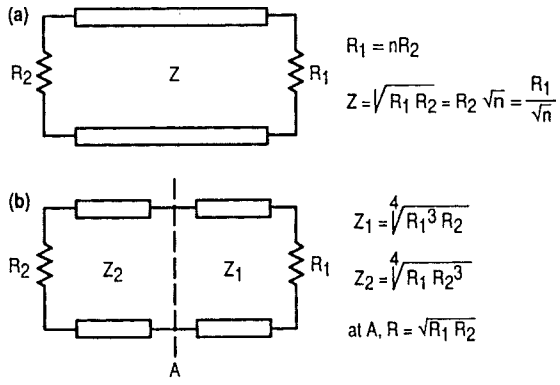
$$Y' = G' + jB' = \frac{R_1}{R_1^2 + (\omega L_1 - 1/\omega c_1)^2} - \frac{j(\omega L_1 - 1/\omega c_1)}{R_1^2 + (\omega L_1 - 1/\omega c_1)^2}$$

$$Y'' = G'' + jB'' = \frac{R_2}{R_2^2 + \omega^2 L_2^2} + j\omega\left(C_2 - \frac{L_2}{R_2^2 + \omega^2 L_2^2}\right)$$

$$B' = \pm\sqrt{\frac{G'}{R_1} - G'^2}$$

$$B'' = (1/kR_2 - G'')\sqrt{\frac{1}{G''R_2} - 1} \qquad k = \frac{L_2}{C_2 R_2^2}$$

Figure 3-236 *Y*-plane representation of the circuit of Figure 3-235a.



$$R_1 = nR_2$$

$$Z = \sqrt{R_1 R_2} = R_2\sqrt{n} = \frac{R_1}{\sqrt{n}}$$

$$Z_1 = \sqrt[4]{R_1^3 R_2}$$

$$Z_2 = \sqrt[4]{R_1 R_2^3}$$

$$\text{at A, } R = \sqrt{R_1 R_2}$$

Figure 3-237 Transformation networks using λ/4-long transmission lines.

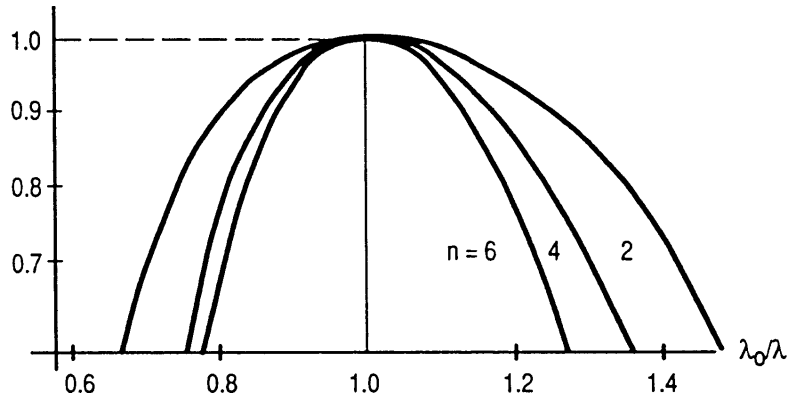


Figure 3-238 Selectivity curves for two λ/4-section networks at different transformation ratios.

The reflection coefficient between the transformed load and generator is equal to

$$\Gamma = \frac{Z_{\text{IN}} - R_G}{Z_{\text{IN}} + R_G} \tag{3-224}$$

When $\Gamma = 0$, there is perfect matching; when $\Gamma = 1$, there is total reflection. The ratio of reflected to incident power is

$$\frac{P_r}{P_i} = |\Gamma|^2 \tag{3-225}$$

The fundamental limitation on the matching takes the form

$$\int_{\omega=0}^{\infty} \ln\left(\frac{1}{|\Gamma|}\right) d\omega \le \frac{\pi}{RC} \qquad \text{Bode equation} \tag{3-226}$$

and is represented in Figure 3-241.

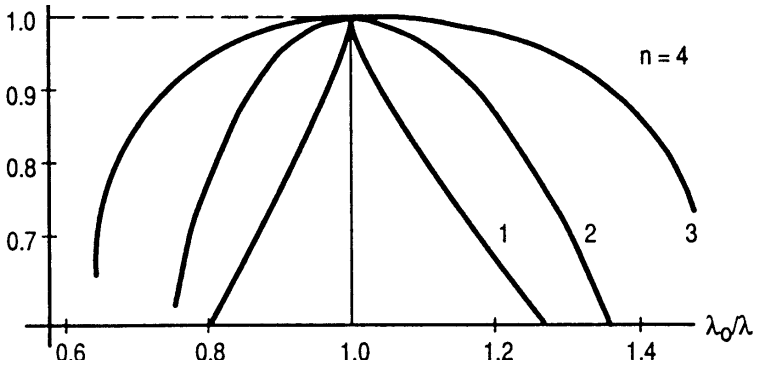


Figure 3-239 Selectivity curves for one, two, and three λ/4 sections.

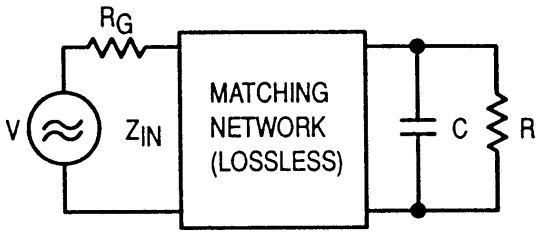


Figure 3-240 General matching conditions.

The meaning of the Bode equation is that the area S under the curve cannot be greater than π/RC and, therefore, if matching is required over a certain bandwidth, this can only be done at the expense of less power transfer within the band. Thus, power transfer and bandwidth appear as interchangeable quantities.

It is evident that the best utilization of the area S is obtained when $|\Gamma|$ is kept constant over the desired band ω_c and made equal to 1 over the rest of the spectrum. Then

$$|\Gamma| = e^{-\pi/\omega_c RC} \tag{3-227}$$

within the band and no power transfer happens outside. A network fulfilling this requirement cannot be obtained in practice because an infinite number of reactive elements would be necessary.

If the attenuation is plotted versus frequency for practical cases, one may expect to have curves like the ones shown in Figure 3-242 for a low-pass filter having Chebyshev character.

For a given complex load, an extension of the bandwidth from ω_1 to ω_2 is possible only with a simultaneous increase of the attenuation a. This is especially noticeable for Q values exceeding 1 or 2 (see Figure 3-243). Thus, devices having relatively high input Q values are usable for broadband operation, provided the consequent higher attenuation or reflection introduced is acceptable.

For average insertion losses or attenuation a (neglecting the ripple), the attenuation decreases if the number n of the network element increases. But above $n = 4$, the improvement is small. For a given attenuation a and bandwidth, the larger the n the smaller the ripple; for a given attenuation, the larger the n the larger the bandwidth.

Computations show that for $Q < 1$ and $n \leq 3$, the attenuation is below 0.1 dB (approximately). The impedance transformation ratio is not free here. The network is a

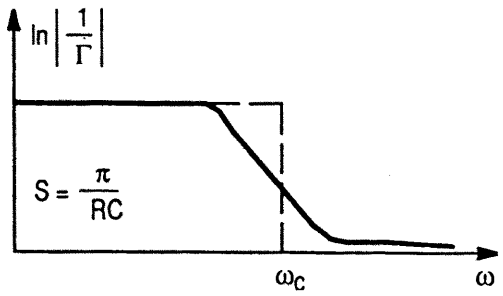


Figure 3-241 Representation of the Bode equation.

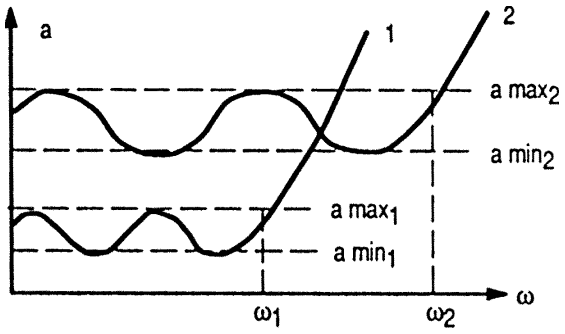


Figure 3-242 Attenuation versus angular frequency for different bandwidths with the same load.

true low-pass filter. For a given load, the optimum generator impedance will result from the computation.

Before impedance transformation is introduced, a conversion of the low-pass prototype into a bandpass filter network must be made. Figure 3-244 summarizes the main relations for this conversion. r is the conversion factor. For the bandpass filter, $Q_{\text{IN max}}$, or the maximum possible input Q of the device to be matched, has been increased by the factor r (from Figure 3-244, $Q'_{\text{IN max}} = rQ_{\text{IN max}}$).

Impedance inverters will be used for impedance transformation. These networks are suitable for insertion into a bandpass filter without affecting the transmission characteristics.

Figure 3-245 shows four impedance inverters. Note that one of the reactances is negative and must be combined in the bandpass network with a reactance of at least equal positive value. Insertion of the inverter can be made at any convenient place [32, 39].

When using the bandpass filter for matching the input impedance of a transistor, reactances $L'_1C'_1$ should be made to resonate at ω_0 by addition of a convenient series reactance.

As stated earlier, the series combination of R_0, L'_1 and C'_1 normally constitutes the equivalent input network of a transistor when considered over a large bandwidth. This is a good approximation up to about 500 MHz.

In practice, the normal procedure for using a bandpass filter matching network will be the following:

1. For a given bandwidth center frequency and input impedance of a device to be matched—for example, 50 Ω—first determine Q'_{IN} from the datasheet as

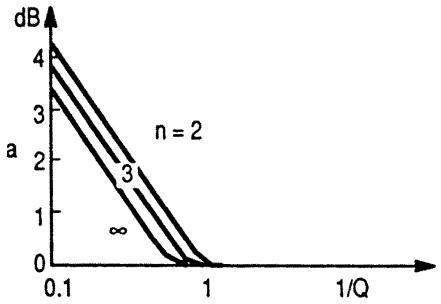


Figure 3-243 Insertion losses as a function of 1/*Q*.

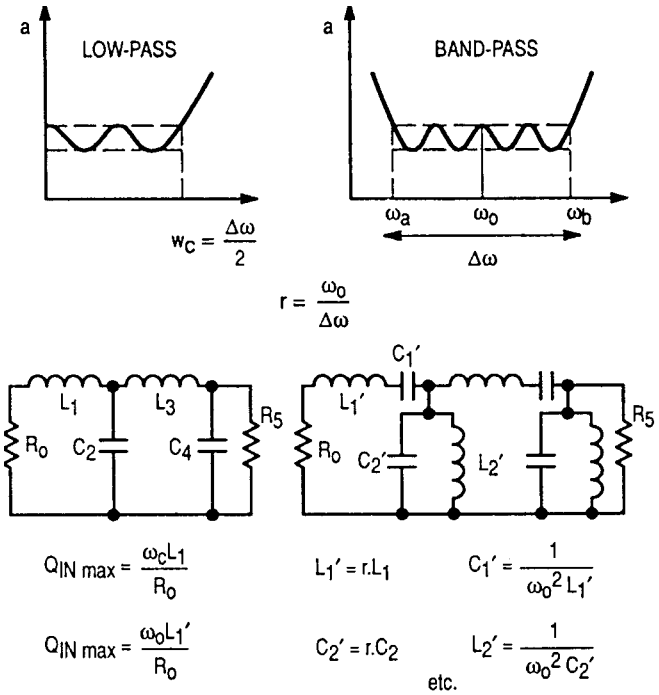


Figure 3-244 Conversion from low-pass filter into bandpass filter.

$$\frac{\omega_0 L_1'}{R_0} \tag{3-228}$$

after having eventually added a series reactor for centering.

2. Convert the equivalent circuit $R_0 L_1' C_1'$ into a low-pass prototype $R_0 L_1$ and calculate Q_{IN} using the formulas of Figure 3-244.
3. Determine the other reactance values from tables [33] for the desired bandwidth.
4. Convert the element values found by Step 3 into series- or parallel-resonant circuit parameters.
5. Insert the impedance inverter in any convenient place.

In the above discussions, the gain roll-off has not been taken into account. This is of normal use for moderate bandwidths (e.g., 30%). However, several methods can be employed to obtain a constant gain within the band despite the intrinsic gain decrease of a transistor with frequency. Tables have been computed elsewhere [40] for matching networks approximating 6 dB/octave attenuation versus frequency.

Another method consists of using the above-mentioned network and then adding a compensation circuit, as shown in Figure 3-246.

Resonance ω_b is placed at the high edge of the frequency band. Choosing Q correctly, roll-off can be made 6 dB/octave. The response of the circuit of Figure 3-246 is expressed by

$$\frac{1}{1 + Q^2(\omega/\omega_b - \omega_b/\omega)^2} \tag{3-229}$$

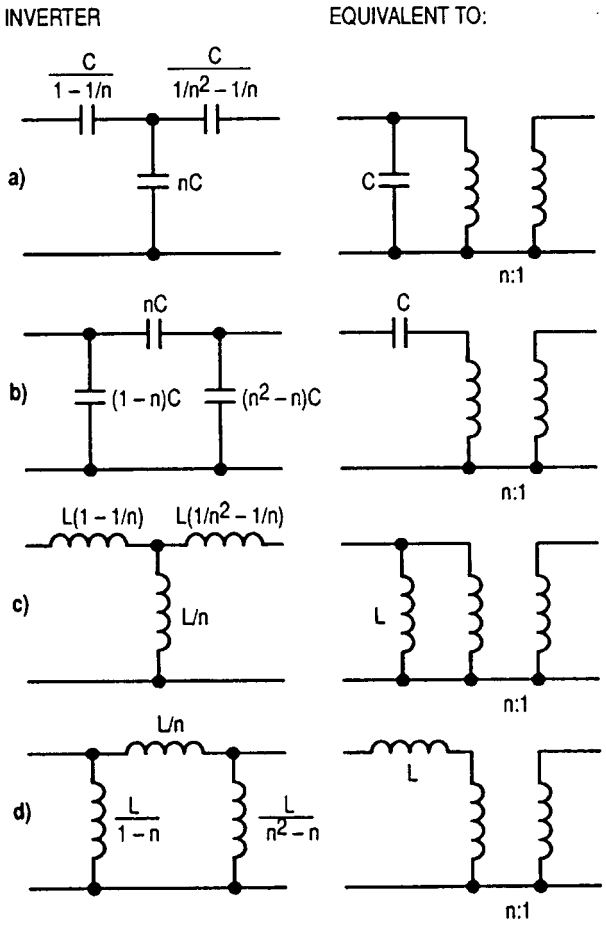


Figure 3-245 Impedance inverters.

where $\omega < \omega_b$. This must be equal to ω/ω_b for 6 dB/octave attenuation.

At the other band edge a, exact compensation can be obtained if

$$Q = \frac{(\omega_b/\omega_a)^2 - 1}{\omega_a/\omega_b - (\omega_b/\omega_a)^2} \tag{3-230}$$

There are several methods available for synthesis of matching networks. They should not be confused with filters because they can be either high-pass, low-pass, or bandpass, and their main purpose is to match one complex impedance to another. In many cases, one of the terminations happens to be 50 Ω and real, but this is just a "lucky" case of the general approach. An excellent paper on this [41] refers to the "real frequency" technique. Another equally important publication is the one by Cuthbert [42], since followed by his book [43]. Two more good references are Vendelin et al. [44] and Potter [45].

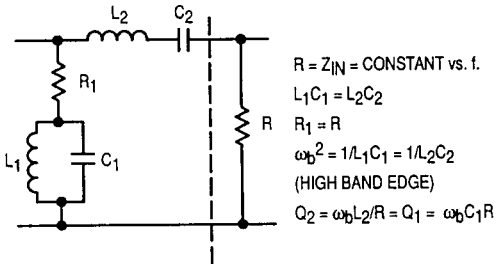


Figure 3-246 Roll-off compensation network.

3-12-3 Example 2: Low-Noise Amplifier Using Distributed Elements

In our previous examples, the amplifiers were built around lumped elements. Since a transmission line length of less than $\lambda/4$ is an inductor and $\frac{3}{4}\lambda$ is a capacitor, the whole amplifier can be built around microstrip technology.

If anyone wonders why we are going to show a low-noise amplifier in the middle of the power section, the reason is that we will select a low-power Siemens transistor, which, being operated in Class A, has an intercept point of +30 dBm and can easily achieve a noise figure of less than 2 dB. Such an amplifier is attractive when feeding an input stage from an antenna that sees many signals and curtails the bandwidth with an input filter of good selectivity that needs a good output termination, which in many cases works against the achievement of a good noise figure because when designing for best noise figure the input impedance of the transistor stage is far from 50 Ω.

First, we will start with the textbook approach and use the Smith diagram to obtain the best noise figure of which the device is capable. Figure 3-247 shows the actual approach to input and output matching.

We start our design by querying the Smith Tool about the minimum noise figure and the gain that comes with it. We mark the point of the lowest noise figure (labeled Γ_{opt}). The small

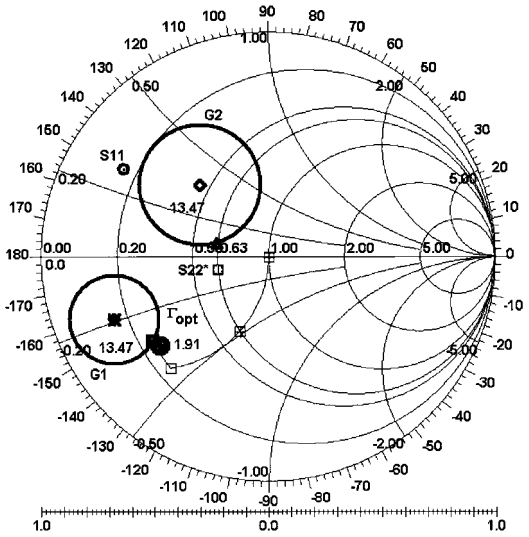


Figure 3-247 Input and output matches for 13.47 dB of gain and NF = 1.91 dB as displayed by the interactive Smith Tool.

noise circle yields the location for Γ_{opt} and tells us the minimum noise figure is going to be 1.91 dB. (By the way, this assumes no further losses, even in soldering the transistor to the PC board.) Then we construct the input gain circle (G1), which touches Γ_{opt} and turns out to represent 13.47 dB of available gain.

The next step is to find a matching network between Γ_{opt} and the 50-Ω source impedance. Immediately there is a conflict: If we look at S_{11}, which is the input reflection coefficient shown in the Smith diagram, it is on the other side of the complex plane and sufficiently far away, indicating that there is going to be quite a difference between noise matching and power-gain matching. We have two options for matching the input: Either we start from the origin and move toward the position of Γ_{opt}, or we start with Γ_{opt} and go to 50 Ω. Depending on the direction in which we do this, the result can be different topologies, such as a high-pass or low-pass filter. The high-pass filter is obtained by starting with Γ_{opt}; the low-pass topology is obtained by starting at the 50-Ω point. The Smith Tool then provides us with the corresponding values of capacitors and inductors after the transformation has been accomplished.

The final step is to do the same type of matching at the output. We start with S_{22}, and either select S_{22}^* (the complex conjugate for highest gain) or S_{22load}. As a rule, the second one is always a lower impedance (in terms of the real part) than the conjugate match. This graphic procedure is somewhat trial and error. As to the topology, a high-pass match should be preferred at the input and a low-pass match at the output. The reason for this is that the high-pass filter will tend to protect the stage from the enormous number of signals present below 1 GHz; using a low-pass filter at the output suppresses unwanted harmonics that the amplifier may generate. Figure 3-248 shows one of the approaches we like best simply because it gives a good input termination and we were trying for a simultaneous match for best noise figure and power gain.

The way we were able to achieve this was to take advantage of the existing feedback in the transistor and select a less-than-optimal output termination but gaining at the input. Taking advantage of the Miller effect, we rotate the phase angle of the feedback with the load to move the optimum noise and gain matching points closer. The danger in this is that the circuit is now highly sensitive to both input and output termination. The output *must* see 50 Ω to maintain this good input match, but since the second transistor stage can now be optimized for best input match, a two-stage combination can be configured to meet this

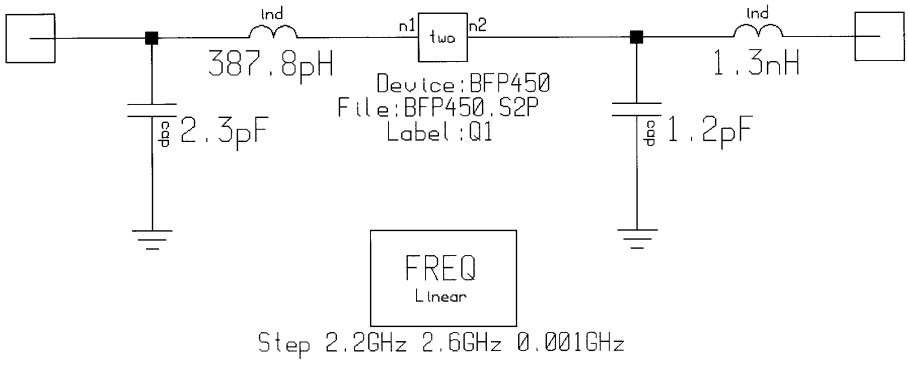


Figure 3-248 A 2.4-GHz amplifier using a BFP450 transistor and lumped-element input and output matching.

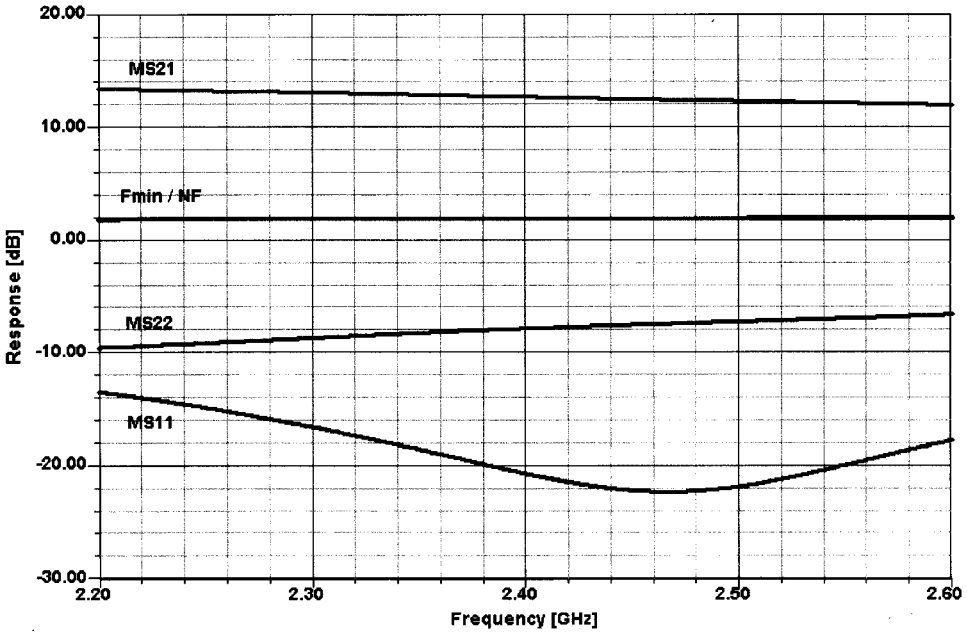


Figure 3-249 Frequency-dependent gain, matching, and noise characteristics of the matched high-dynamic-range BFP450 2.4-GHz amplifier.

requirement. Also, let us not forget that the device is running at 50 mA and as far as bipolar transistors are concerned, this is probably the best amplifier combination we have seen for getting the highest dynamic range and lowest noise figure. The resulting gain, matching, and noise figure characteristics are shown in Figure 3-249.

Now we turn to using distributed elements and at the same time use a shunt feedback at the output; to be specific, we will load it with 150 Ω. The reason for this is to be able to obtain a good compromise at the input that will depend less on the output and at the same

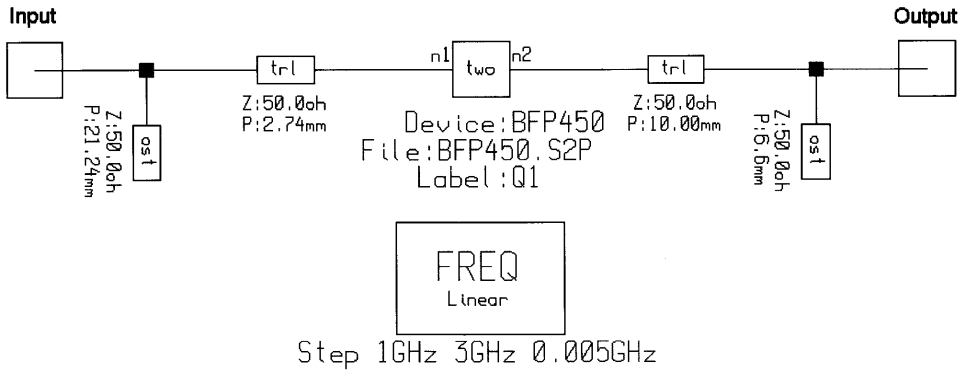


Figure 3-250 Distributed-element input and output matching for the 2.4-GHz BFP450 amplifier.

time improve the S_{22} matching at the output. The procedure for doing the matching with a transmission line offers only certain degrees of freedom, while doing the matching with lumped elements allows the choice of series or parallel inductors or capacitors. In the case of a transmission line, we have open stubs, shorted stubs, and transmission lines. This may result in somewhat more complex matching schemes. On the positive side, they can easily be trimmed, are reproducible, and certainly cost less than exotic capacitors, which need to have small tolerances and to be able to handle high RF currents.

The resulting input and output matching is shown in Figure 3-250, and its frequency response in Figure 3-251. We were able to maintain a good input and output match with an insignificant deterioration of the noise figure.

Since this is a high-dynamic-range amplifier, we also show its predicted two-tone response in Figure 3-252. The manufacturer specifies an $IP_{3,in}$ of 29 dB for a bias point of $V_{CE} = 3$ V and $I_C = 50$ mA with $Z_S = Z_{S,opt}$ and $Z_L = Z_{L,opt}$ at 1.8 GHz. Significantly, however, the manufacturer-supplied nonlinear BFP450 library we used exhibits considerably different characteristics at $V_{CE} = 4$ V and $I_C = 50$ mA than those reflected in the corresponding S parameters supplied by the same manufacturer. In our circuit, the S parameters resulted in about 13.5 dB of gain (based on a compromise between matching for optimum noise and maximum gain) versus about 9.5 dB of gain for the nonlinear library (match optimized for maximum gain). The difference between the 2.4-GHz F_{min} values returned for the two modeling approaches is similarly striking: 1.90 dB for the bare S parameter set and 7.76 dB for the bare nonlinear library biased to $V_{CE} = 4$ V and $I_C = 50$ mA. Clearly, the nonlinear modeling of the BFP450 needs work!

Having fully evaluated the amplifier at 2.4 GHz as a working example, we note that this frequency is used by radio amateurs for some applications. We will therefore supply

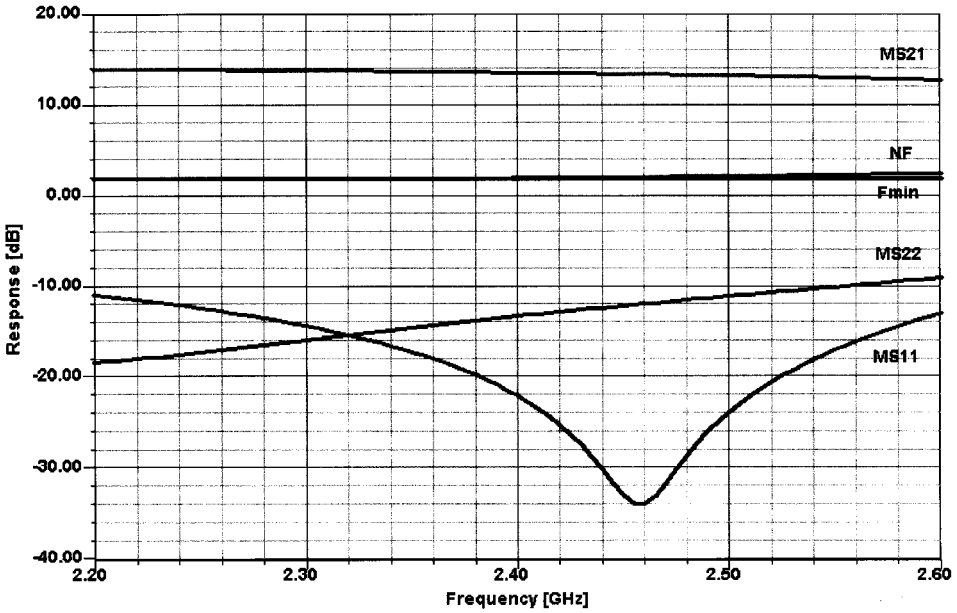


Figure 3-251 Predicted frequency-dependent gain, matching, and noise performance for the BFP450 amplifier with distributed-element matching.

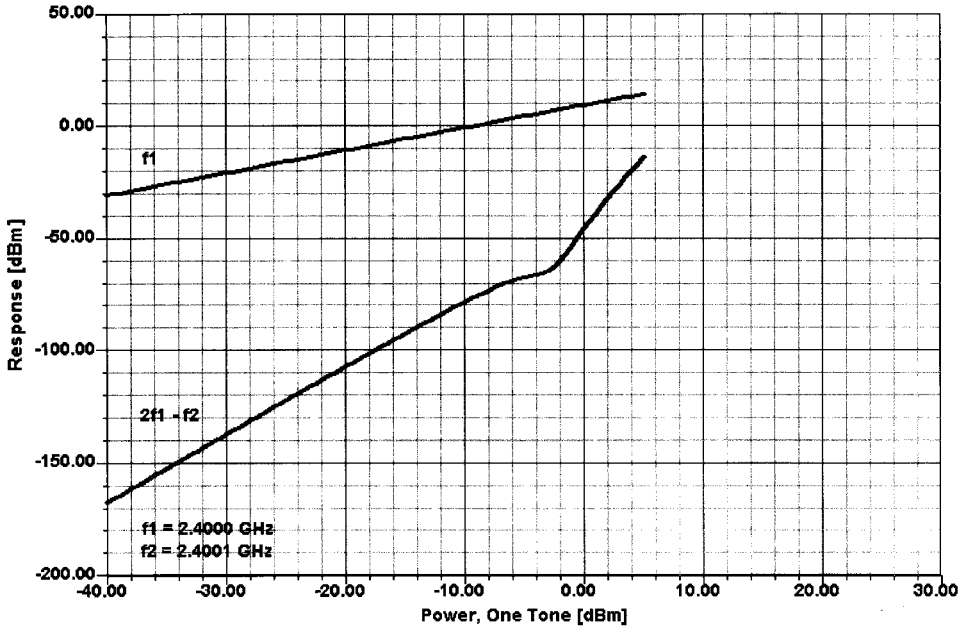


Figure 3-252 Predicted two-tone response of the BFP450 amplifier with the circuit adjusted for maximum gain (9.5 dB): $IP_{3,in} \approx 28.5$ dBm and $IP_{3,out} \approx 38$ dBm. That this version's gain is significantly lower than that of the *S*-parameter-based equivalent illustrates the need for better accuracy in parameter extraction.

examples for 432 and 1296 MHz, popular spot frequencies in the amateur 70- and 23-cm bands, respectively. Figures 3-253 and 3-254 show the 432-MHz version's schematic and frequency-dependent responses. Figures 3-255 and 3-256 show the schematic and frequency-dependent responses for the 1296-MHz version.

Using a bipolar transistor that operates at much less current, the noise figure can be improved at the expense of the dynamic range/intercept point. It is possible to reduce the noise figure by about 1 dB or so, but the reduction in intercept point will be much more dramatic. Another choice is the use of GaAsFETs for this application, but their matching in comparable circuits has been described as a nightmare simply because the devices have too much gain and very high input impedances (essentially 0.2 pF in series with 3 Ω) combined with reasonable output impedances in the vicinity of 240 Ω. Here is a wide field in which engineers can experiment and examine price/performance and performance/real estate issues. GaAsFETs also tend to require more than 30–40 mA to be alive, and even the low operating voltage of 4–5 V makes the device consume a lot of dc power.

3-12-4 Example 3: 1-W Amplifier Using the CLY15

The following is an application showing a fairly wideband amplifier with +30 dBm (1 W) output. We need to point out that in accordance with the various digital modulation modes used with this amplifier, this is not continuous operation, but a much less than 1:1 duty cycle operation. The device would probably not survive operating at full output under CW conditions. This is subject to a discussion with the manufacturer, Siemens, who has not

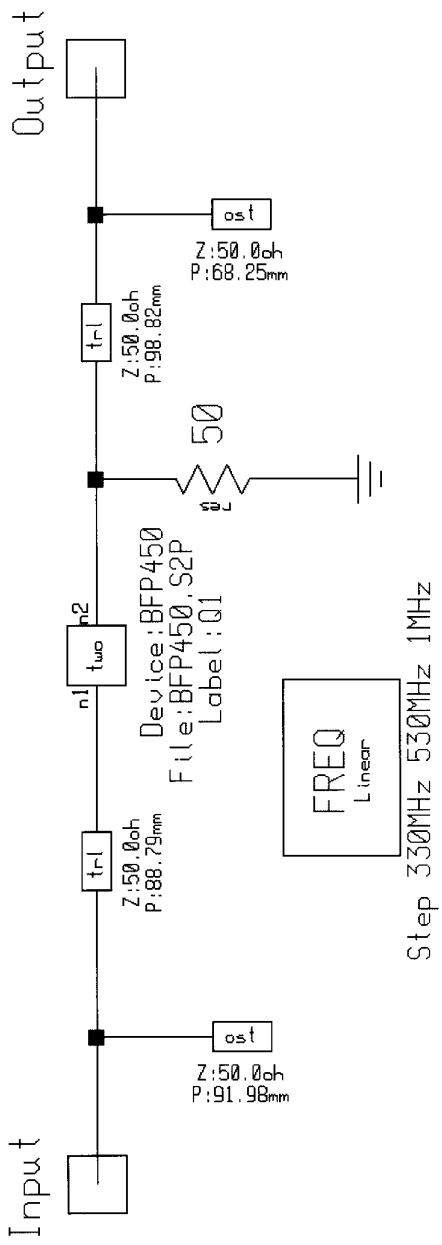


Figure 3-253 Schematic of BPF450 amplifier for 432 MHz. Because f_T = 25 GHz for the BFP450, at this lower frequency we must shunt the device output with a 50-Ω resistor to maintain stability.

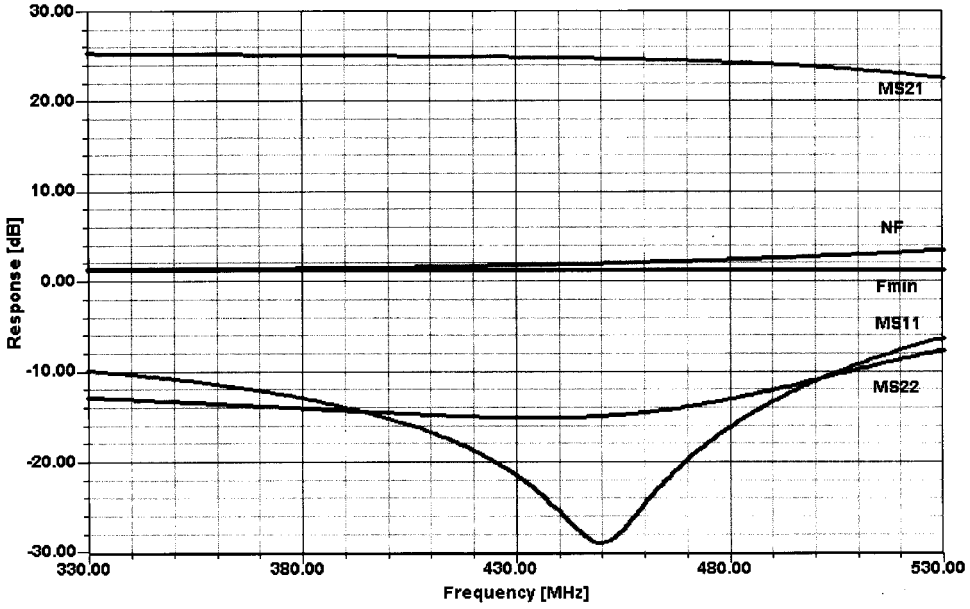


Figure 3-254 Frequency-dependent gain, matching, and noise performance for the BPF450 amplifier at 432 MHz. As with the circuit in Figure 3-253, the input and output network values have been chosen to achieve a low NF and good input matching at the expense of the output match.

released sufficient data for us to make a judgment as to the CLY15's continuous power capability. Since there is no useful large-signal model available at the time of this writing, we will design this Class A amplifier using large dc bias *S* parameters. In the actual design there will be some cut-and-try, but we find that in this type of application this is still the best way of handling the design.

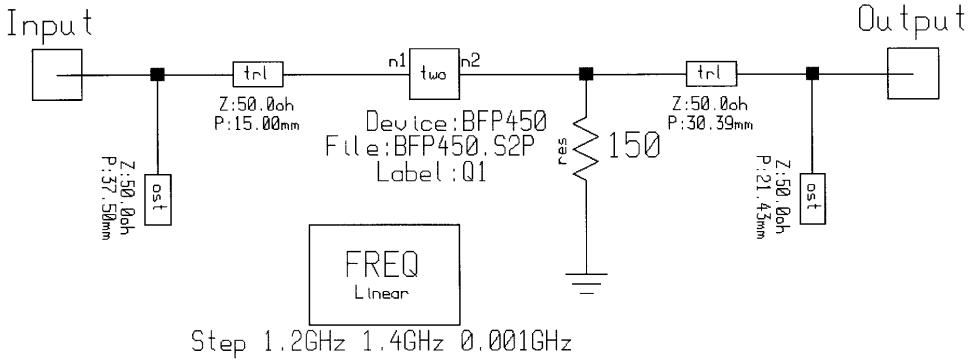


Figure 3-255 Schematic of the BPF450 amplifier for 1296 MHz. A loading resistor is still necessary to ensure stability, but because the transistor exhibits less gain at this frequency than at 432 MHz, we can increase the loading resistor to 150 Ω.

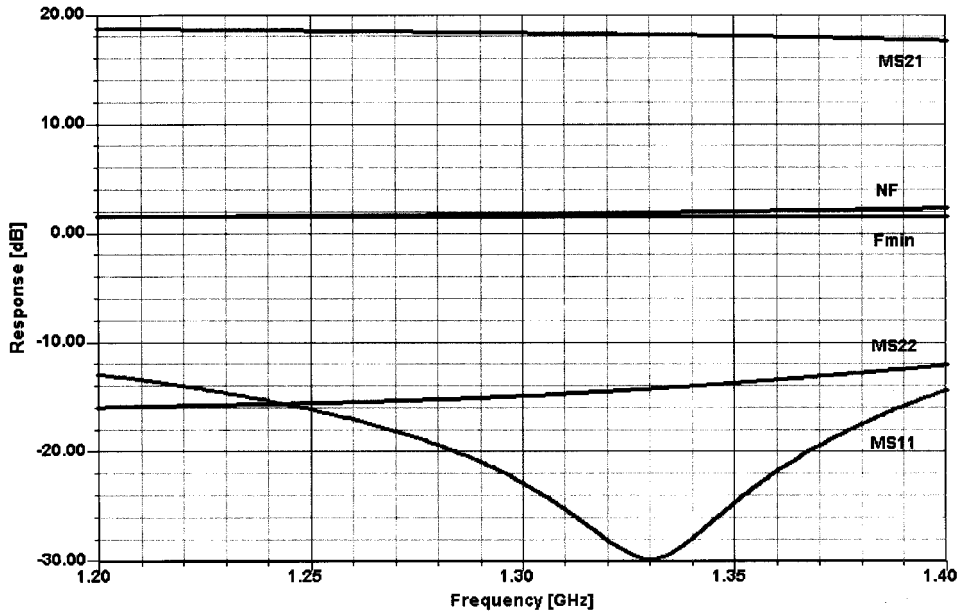


Figure 3-256 Frequency-dependent gain, matching, and noise responses of the BFP450 amplifier at 1296 MHz. As with the two previous versions, the input and output network values have been chosen to achieve a low NF and good input matching at the expense of the output match.

The CLY15 (gate width 16 mm) is the largest of a series of Siemens GaAsFETs; it was already mentioned in Chapter 2, where its datasheet is reproduced. Its device family also includes the CLY2 (gate width 2 mm), CLY5 (4 mm), and CLY10 (8 mm). It is possible that bigger devices will be offered in the future. One of the most critical parameters will be the saturation voltage; from the datasheet of the CLY15, we can see that at 1.5 A the saturation voltage is approximately 1 V. This drop of 1 V includes losses that result from other parasitic elements; the dc *I–V* curve (Figure 3-257a) indicates a saturation voltage of 0.75. Since we aim for 1-W output power, the resistive portion of the load will be $(5 - 1)^2 / (1 \times 2) = 8\ \Omega$. This assumes a supply voltage of 5 V.

The next step is to design the input and output matching network. Previewing the *S* parameters, it will be noticed that they are essentially inductive, which means that the input match, using lumped elements, will result in a combination of capacitors at both the input and the output. Figures 3-258 and 3-259 show the ways in which the input and output matches were done on the Smith diagram resulting in lumped matching elements. Inspecting the lumped input matching, we find the following. We start with S_{11}, which is in the lower part of the Smith diagram, and determine its conjugate match, S_{11}^*. We already pointed out that this device is inductive, which means that matching it from the conjugate point will require two capacitors. The first one is a series capacitor, as identified in the Smith diagram, and followed by a shunt capacitor to transform the impedance to 50 Ω. A close inspection of the Smith diagram shows that the resulting point is not quite on S_{11}; it is an iterative process to exactly find the crossover point, yet the input match of better than 20 dB is more than acceptable.

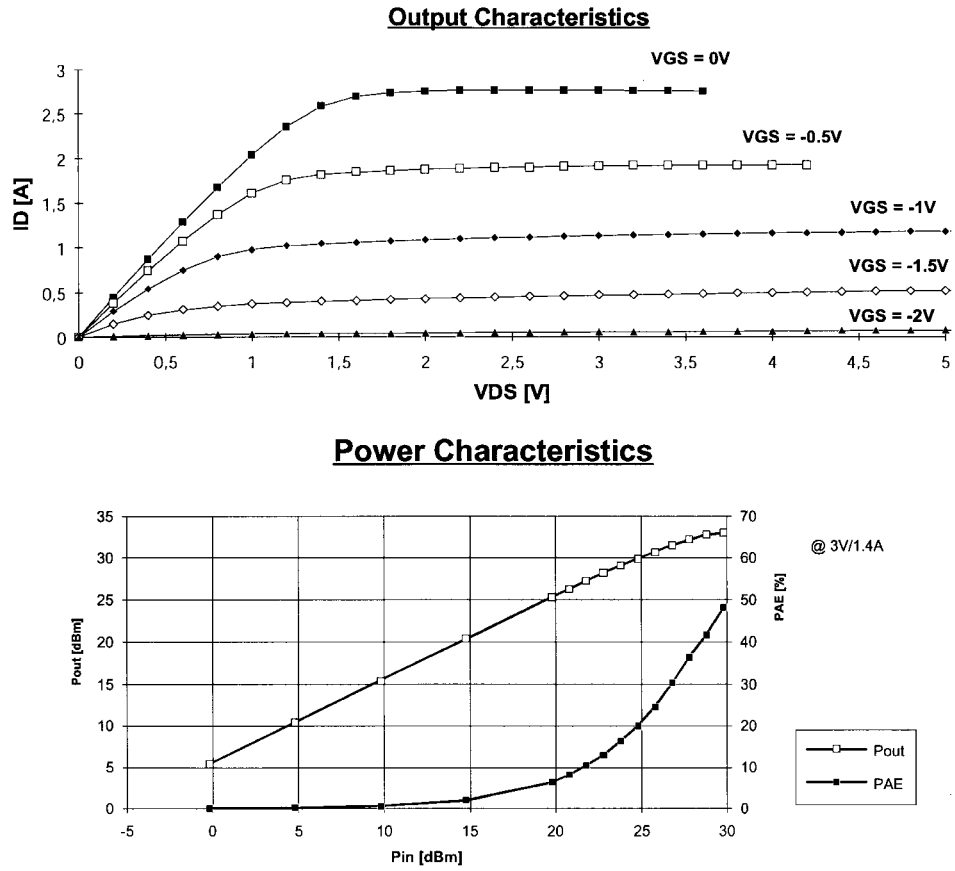


Figure 3-257 Output and power characteristics for the CLY15 medium-power GaAsFET.

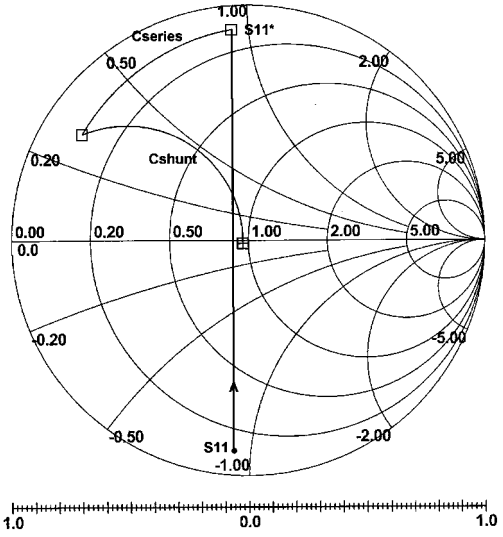


Figure 3-258 Input match for S_{11}^{*} using lumped elements.

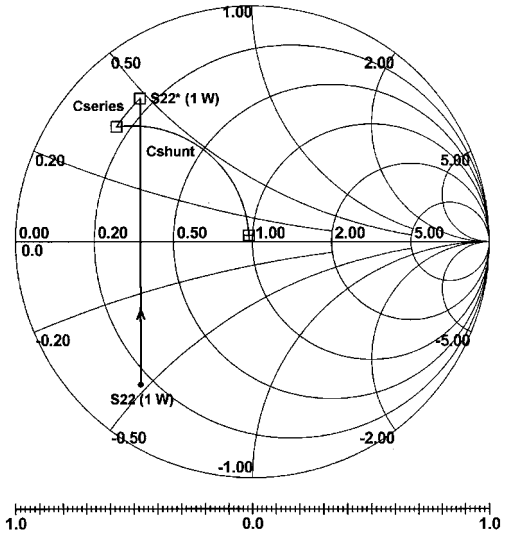


Figure 3-259 Output match for S_{22}^* (1 W) using ideal lumped elements.

As far as the output is concerned, we use the same technique by moving from S_{22} to S_{22}^*. S_{22} has been determined from the required output power; as we showed above, the real portion is 8 Ω. Using the lumped elements, the matching technique is identical; we first need to use a series capacitor and then a shunt capacitor for the transformation. The arc for the series capacitor this time is much longer, resulting in a series capacitance of 11.2 pF instead of the 2.1-pF capacitance used at the input. The actual schematic is shown Figure 3-260. The drawback of lumped elements is that it is not practical to buy components that can handle the heavy RF currents that flow in the input and output networks. Taking the values obtained

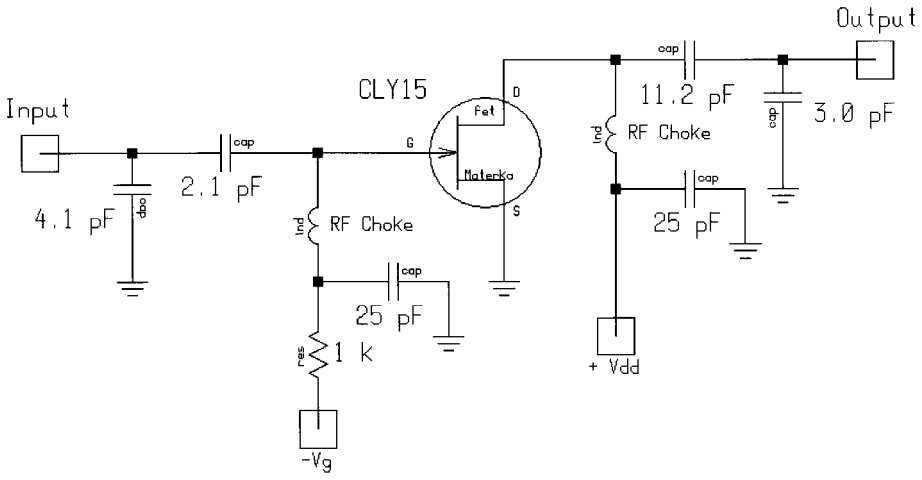


Figure 3-260 The complete amplifier using the CLY15 medium-power GaAsFET, with lumped-element input and output matching.

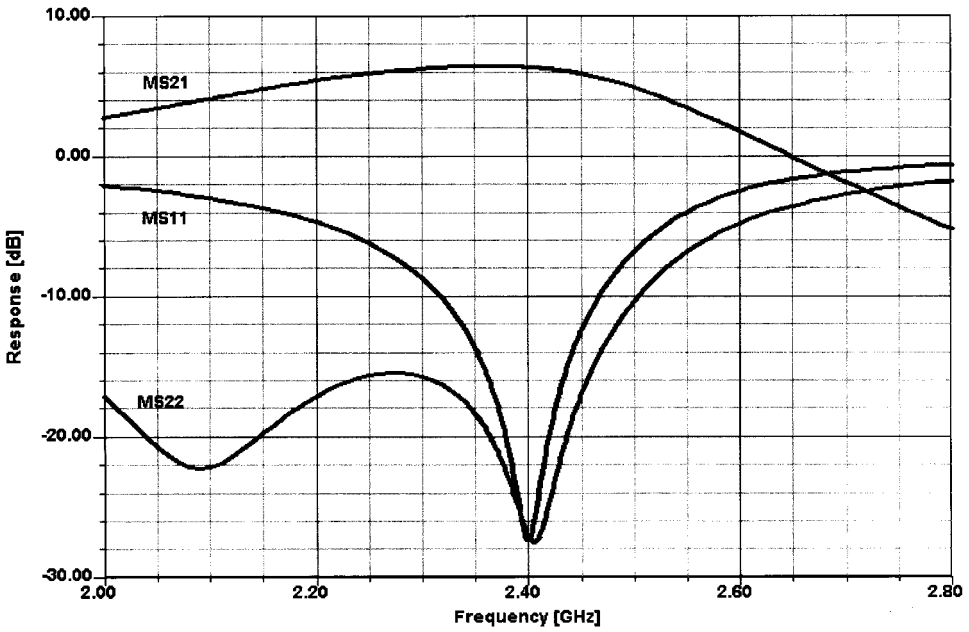


Figure 3-261 Frequency-dependent gain and matching responses of the CLY15 amplifier using ideal lumped elements.

from the lumped input and output matching approach, Figure 3-261 shows the resulting frequency response.

The next step is to provide the same matching condition using distributed elements, which of course have the advantage that they can be printed on the PC board. Instead of using ideal elements, we used a loss tangent of 2E–3, or a Q of 500. The manufacturer of the R4000 material (Rogers) provided this specification. Now we start with the same approach by moving from S_{11} to S_{11}^*. From there, we use an open stub with a transmission-line impedance of 50 Ω, followed by a transmission line to the point in the left lower corner, and, finally, a shorted transmission-line stub allows the 50-Ω matching. For this amplifier, we are not looking for noise matching but for input and output power matching. Figure 3-262 shows this match on the Smith Chart.

At the output, the procedure is again similar: We move from the S_{11} for 1 W to a conjugate value, then use an open stub that acts like a capacitor until we cross the real axis; finally, a λ/4 transmission line is used to match from approximately 75 Ω to 50 Ω. Figure 3-263 shows this match on the Smith Chart. This concludes the distributed matching design for the CLY15 amplifier.

Finally, the complete circuit using distributed elements is shown in Figure 3-264. From a dc point of view, we took advantage of the incoming shorted stub by using a transmission line of the appropriate length and a bypass capacitor to ground. This allows us to feed the gate with the required negative voltage without interfering with the frequency response of the amplifier. When evaluating the total frequency response (Figure 3-265), we noticed that this circuit seems to have a higher Q than the lumped equivalent. Again, to obtain a wider bandwidth more transformation stages will be needed. We leave this task to the interested reader.

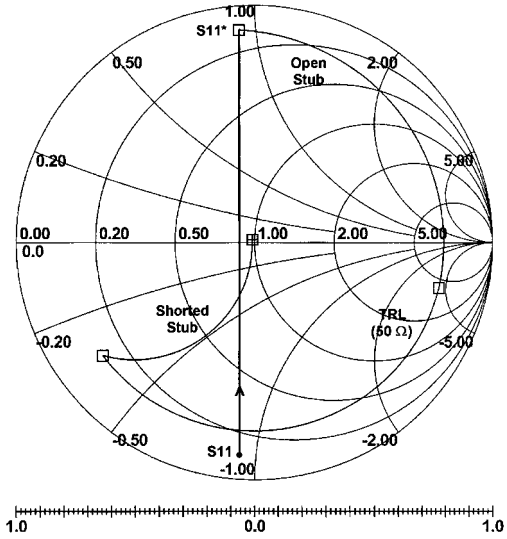


Figure 3-262 Input match for S_{11}^* (1 W) using distributed elements.

The 1-dB bandwidth of the amplifier using distributed, lossy elements is 120 MHz. If a wider bandwidth is required, matching networks with more elements are required. Because of the gain–bandwidth product, the peak gain for such a wider-bandwidth arrangement will be less than this narrowband example.

Having used a lumped to distributed transformation, we next present a power amplifier that uses coaxial lines, baluns, and other transformation steps capable of handling the increased power requirement.

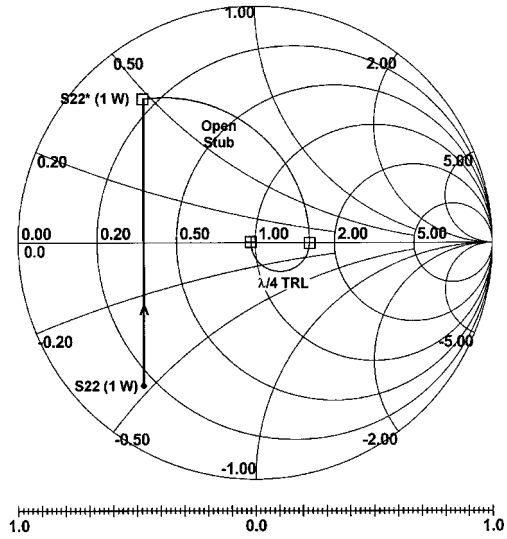


Figure 3-263 Output match for S_{22}^* (1 W) using distributed elements.

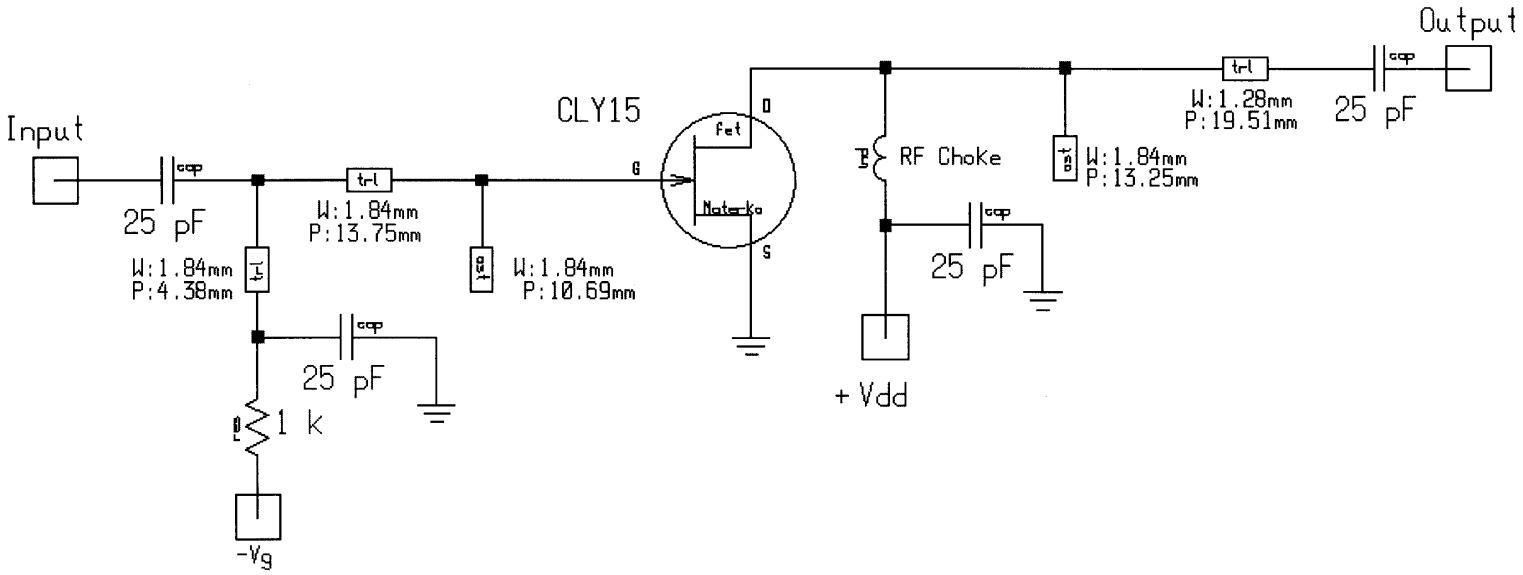


Figure 3-264 The complete amplifier using the CLY15 medium-power GaAsFET, with distributed-element input and output matching.

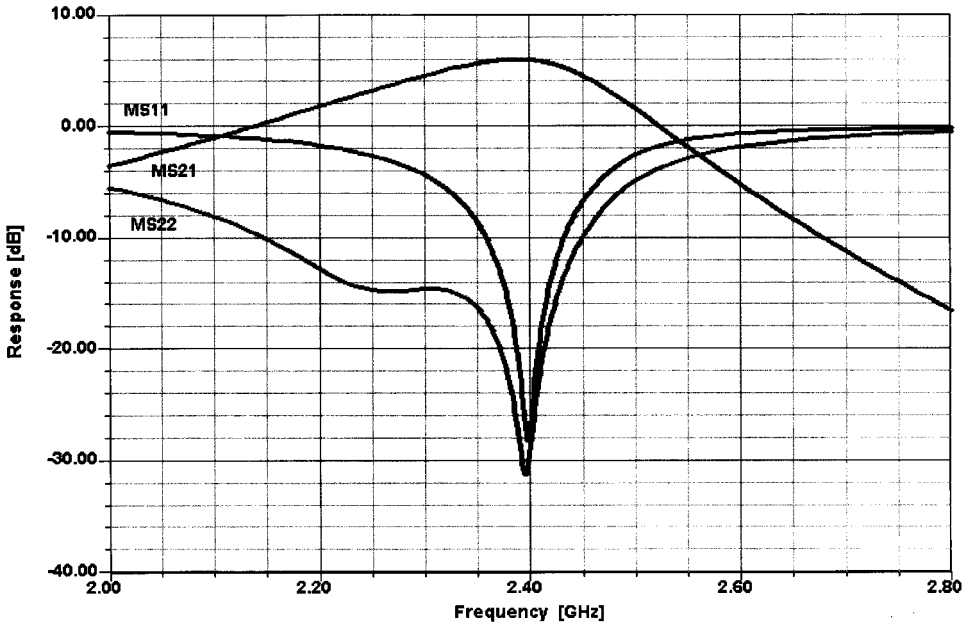


Figure 3-265 Frequency-dependent gain and matching using distributed, lossy elements.

3-12-5 Example 4: 90-W Push–Pull BJT Amplifier at 430 MHz

The last complete application we want to examine is a 90-W output, push–pull amplifier using two TRW transistors (that are no longer made—an all-too-frequent occurrence after a design is complete, resulting in a scramble for replacements). Figure 3-266 shows the circuit diagram of this UHF transistor amplifier, and Figure 3-267 shows its mechanical design. The amplifier requires 28 Vdc and will draw approximately 6 A, requiring 15 W of drive at 438 MHz. This amplifier operates in Class AB, which means it is not limited to amplifying FM and other angle-modulated signals, and can accommodate even single-sideband (SSB) operation.

The drive power coming from the source connected to the input N connector (15 W) is fed through a short piece of coaxial cable to the PC board sitting on top of the heat sink. To change from single ended to push–pull, a 50-Ω coaxial cable is coiled to provide sufficient electrical length to act as a balun that provides two symmetrical paths through the amplifier. The RC section of 10 Ω in parallel with a 220-pF capacitor stabilizes the amplifier at lower frequencies, while at higher frequencies the 10-Ω resistor has no contribution. There are two independent 4:1 transformers, one for each section; they are built using 25-Ω coaxial cable with a set of two binocular ferrite cores to force the transformation. The dc supply is also fed to the transistors at this point. The stabilization circuit consists of a temperature monitor using the *pn* junction of a transistor, side-mounted on the heat sink, and a combination of operational amplifiers and adjustments. This guarantees proper, temperature-independent biasing. The 4:1 input baluns result in 12.5 Ω impedance, or 3.125 Ω per transistor. Since the input impedance at this drive level is somewhere in the vicinity of $[(1 \ldots 2) + j(3 \ldots 5)]$ Ω; the final matching is accomplished by an *RC* section in which the inductance is actually printed

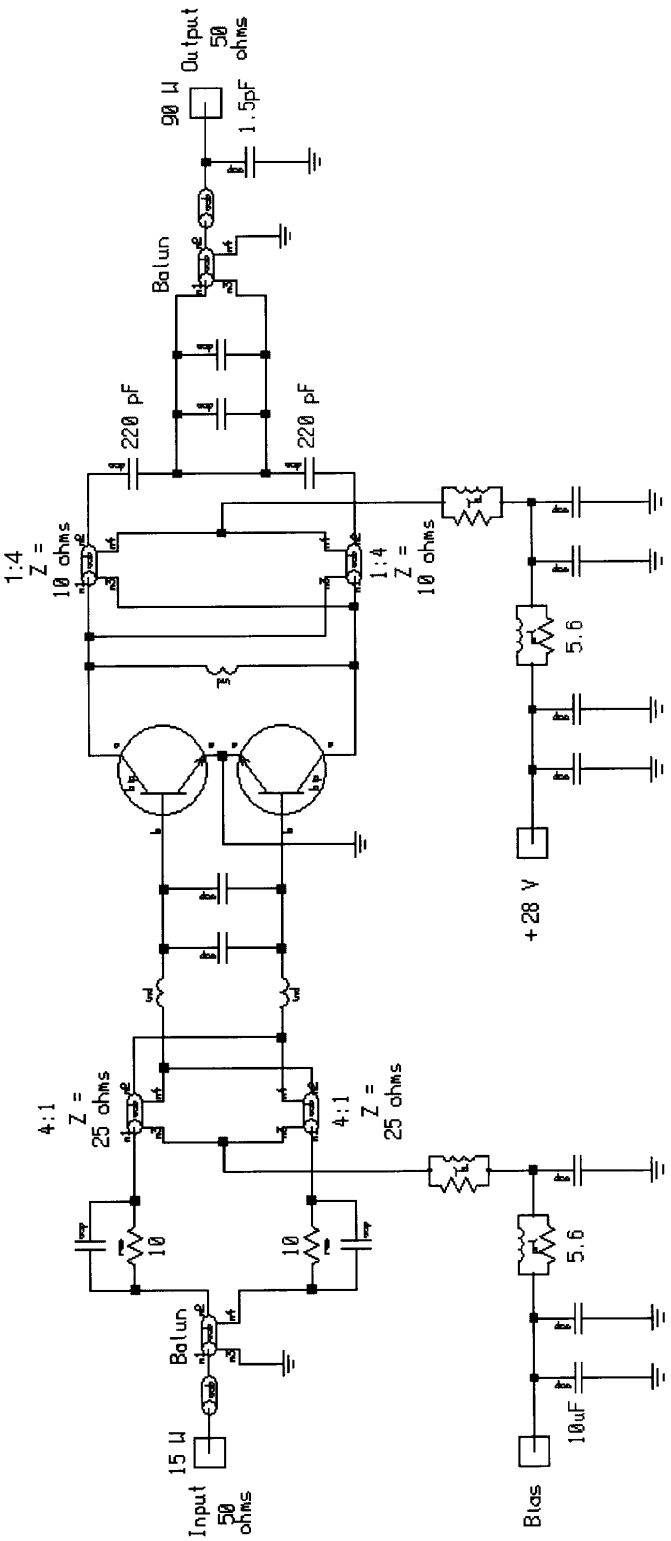


Figure 3-266 Circuit diagram of the 90-W, push–pull, 438-MHz linear amplifier.

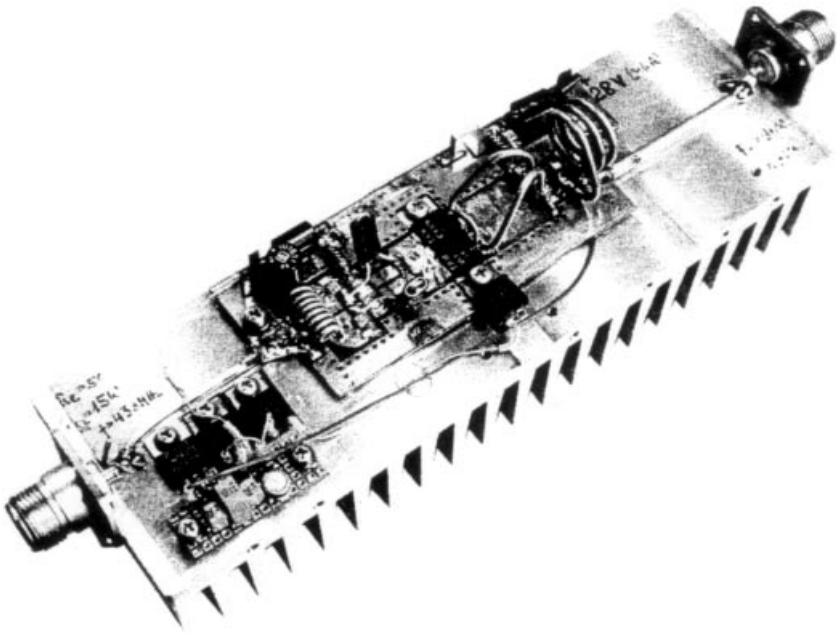


Figure 3-267 Mechanical assembly of the 90-W amplifier shown in Figure 3-266.

and the shunt capacitor across is soldered from base to base as close to the transistor bodies as possible. A ceramic-dielectric variable capacitor in parallel with this allows fine tuning.

The output consists of an inductor from collector to collector that tunes out the capacitive reactance and makes the transistor output impedance real. We then follow the same pattern by taking a 1:4 transformer, but this time built using 10-Ω rigid line instead of 25-Ω rigid line. Together with the output capacitance, prior to the output balun, this is a medium-bandwidth matching arrangement that transforms up to 50 Ω. The collector impedance per collector had to be $(28 - 2)^2/2 \times 90$, resulting in roughly 3.8 Ω—not that different from the input match at the balun point. This way, the matching networks at the input and output can be identical. At the output of the matching network, another coiled 50-Ω coaxial cable serves as balun; because it must operate at the 90-W level, larger-diameter line was used than in the input balun. Finally, another section of rigid line is used to connect the output balun on the PC board to the output N connector. The output capacitor again helps in the transformation.

Both the base and collector voltage supplies are heavily filtered using ferrite-bead-based RF chokes and heavy wire inductors wound on low-impedance resistors. A fairly large number of these amplifiers were actually built and are still in use.

3-12-6 Quasiparallel Transistors for Improved Linearity

Some applications require really low intermodulation distortion. One of the authors (Rohde), during his tenure at RCA Government Systems Division, and his team developed a parallel-device topology that is somewhat similar to the conventional parallel approach but shows vastly increased linearity. In creating higher-power devices, manufacturers parallel several transistors in one package, connecting them in a clever way so the overall drive, dissipation, and output power are equally divided between all the devices. Special nickel–chromium

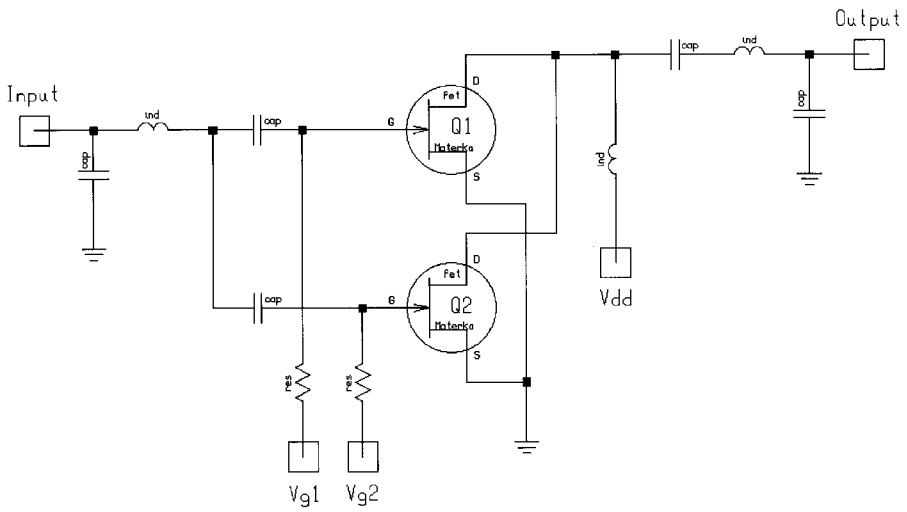


Figure 3-268 The quasiparallel amplifier topology looks like a standard parallel amplifier. Biasing the devices to different operating points nets an improvement in IMD performance.

ballasting resistors in the many fingers of the emitter connections of the multiple internal devices equalize any differences among the transistors while increasing the saturation voltage by only a small amount. In our high-linearity design, we used two transistors in parallel, one being biased somewhere between Class A and B, and the second one somewhere between Class B and C (Figure 3-268). The exact bias voltages and operating points had to be found through experiment. In practice, Q1, the "initially on" transistor, is driven to a point close to that at which its IMD products would increase. Q2 is biased such that at this same drive level it adds to the output power.

We frequently referred to this configuration as "turbocharged." The IMD products at full output power were down 40–45 dB, compared to 26–30 dB for a single transistor. While such an arrangement shows superior IMD performance, because of the somewhat abrupt kick-in of the second transistor, the transconductance response is semismooth, a characteristic that might make alert readers think that IMD will occur at this point. On the other hand, the transconductance for Q1, in effect, saturates or stays constant, so when the transconductance responses of both transistors are overlaid, the result is a smooth response overall.

This technique has been validated up to 100 W and up to at least 50 MHz. Its drawbacks include the need for two transistors with their associated costs, additional real estate, and higher input and output capacitances. Nonetheless, considering that MOS transistors capable of operation to 1 GHz were unavailable when this technique was developed, we believe that it should be revisited.

The Class D and higher operating modes require a large number of additional components and will make the efficiency higher (selective in frequency), but not necessarily reduce the IMD products. A further point of consideration is that it is incorrect to assume that reducing the input level to a given Class A will necessarily reduce its IMD products. As with tubes, the closer one gets to the knee, even in Class A operation, the worse the IMD. (This phenomenon was particularly troublesome in systems using 1-kW amplifiers up to 150 MHz;

underdrive with the intent of getting better IMD performance produced just the opposite result.) If a wide power range is considered, we recommend using a bias scheme that includes sufficient intelligence to find the best operating point.

Another hot amplifier issue is that of IMD as a function of load impedance. If an amplifier is used at output power considerably below its design value, its load impedance will be too low and, depending on the circuit configuration, the result may be additional IMD.

3-12-7 Distribution Amplifiers

It is not uncommon that a signal from a particular source like an antenna has to be distributed to several places. A condominium is a typical situation in which energy from a main signal source must be split for distribution to a larger number of users—there must be enough voltage or power to produce a noise-free picture. An amplifer capable of doing this is a *distribution* amplifier, not to be confused with a *distributed* (or traveling-wave) amplifier, such as that shown earlier in this chapter in Figure 3-6. An interesting means of achieving this is to use multiple Wilkinson couplers following an amplifier. Figure 3-269 shows an arrangement that makes heavy use of multiple power dividers to provide multiple outputs. Theoretically, the reverse is also possible: If we want to collect energy from a number of sources and combine them, as on a cable, we can use Wilkinson couplers at the *input* of a gain block. Note, however, that depending on the level and number of the combined signals, this may put enormous strain on the linearity of the amplifier. What we want to avoid is a lot of distortion, which, in the case of television, will show up as color confetti in the noise and signal.

3-12-8 Stability Analysis of a Power Amplifier

One of the hidden pitfalls in narrowband circuit design is that a circuit may be unstable outside the frequency range of interest. Circuits containing structures built from distributed elements, which exhibit multiple resonances, can be particularly susceptible to such instabilities. In this example, we analyze the stability of a simple FET power amplifier designed using small-signal concepts and the interactive Smith Tool utility in Ansoft's Serenade

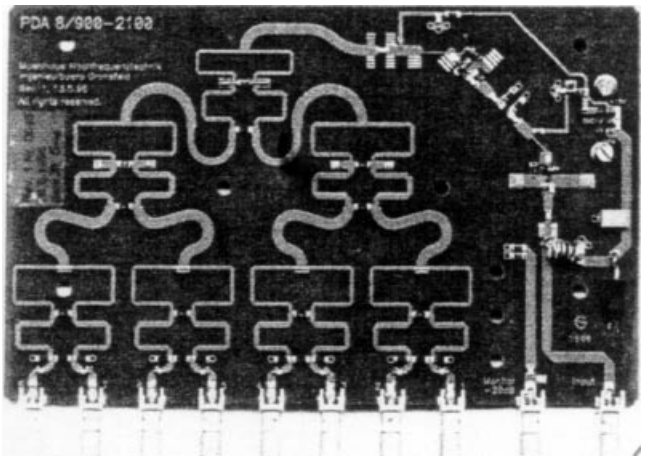


Figure 3-269 Distribution amplifier showing the use of multiple Wilkinson power dividers.

Design Environment. The evaluation shows how K-factor analysis alone is not sufficient to check for stability of a design.

Small-Signal ac Analysis. We begin by taking the Figure 3-270 amplifier, which was designed for about 16 dB of gain at 2.2 GHz. Due to the very low, capacitive input impedance of the FET and the proximity of the source-plane gain circles and stability circles to the edge of the Smith Chart, the matching network attempts to transform 50 Ω to an impedance of approximately $(2.5 + j15)$ Ω. The impedance lies close to the stability circles at 2 GHz but remains in a stable region. The bias point selected (–3.5 V) is chosen for Class AB operation. Higher linear gain can be achieved by reducing the gate voltage to –2.0 V; at this bias point, the amplifier gain will increase to about 19 dB.

First, we will do a small-signal ac analysis and examine the S parameters and K and B_1 factors [see Eqs. (3-154) and (3-157), respectively] from 1 GHz to 6 GHz—well beyond the bandwidth of the design. Figure 3-271 shows the circuit's frequency-dependent gain and matching characteristics.

If we examine B_1 and the stability factor K, we see that the amplifier is not unconditionally stable in the 2.2-GHz passband as shown in Figure 3-272. K dips slightly below 1, although B_1 remains above zero. In practice, a slight change in design would be required to bring K above 1 and remove this potential instability. However, for our example, this is of no consequence, as we shall see below.

Nyquist Stability Analysis. Now we will utilize a Nyquist stability analysis to check the stability of the amplifier. To fully realize the advantages of a Nyquist stability analysis, we must sweep frequency over a very wide range. We want to be able to pick up instabilities at *any* frequency, not just near the design frequency. In this example we sweep from 1 kHz to 20 GHz. A start frequency of 1 kHz is chosen so we can determine the starting point of the Nyquist plot and we ensure that any low-frequency resonances are picked up. The stop frequency is chosen past the f_{max} of the transistor (where we know it cannot oscillate).

The analysis results are shown in polar plot form in Figure 3-273 and in magnitude–angle form in Figure 3-274. The polar graph shows the real and imaginary parts of the system determinant as frequency is swept. The magnitude–angle graph shows the magnitude and cumulative angle (no cut at ±180°) of the system determinant as frequency is swept. The key to interpreting these graphs lies in the following statement drawn from the Nyquist criterion:

The circuit will be unstable if the Nyquist plot encircles the origin in a clockwise direction.

The term "unstable" refers to the existence of natural frequencies lying in the right-half of the complex frequency plane (RHP). These natural frequencies on the RHP will cause oscillations. To encircle the origin in a clockwise direction means that the Nyquist plot will travel in a clockwise path, cross the negative real axis, and continue to completely surround the origin. If the path appears to encircle the origin but then "unravels" itself in a counter-clockwise direction, the origin is not actually encircled. Figure 3-273 shows that the origin is not encircled and therefore this circuit is stable.

The magnitude–angle plot can assist in determining if a clockwise encirclement has actually been made with the help of the following:

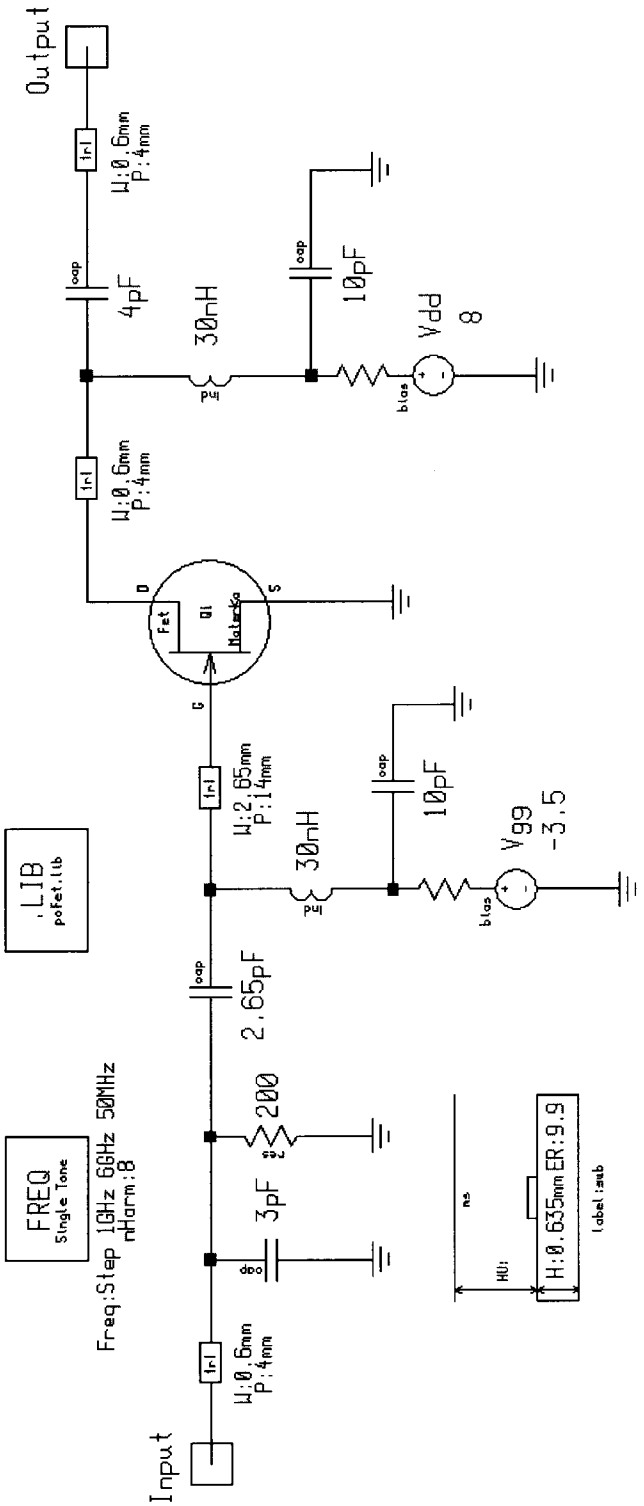


Figure 3-270 A 2.2-GHz amplifier design for stability analysis.

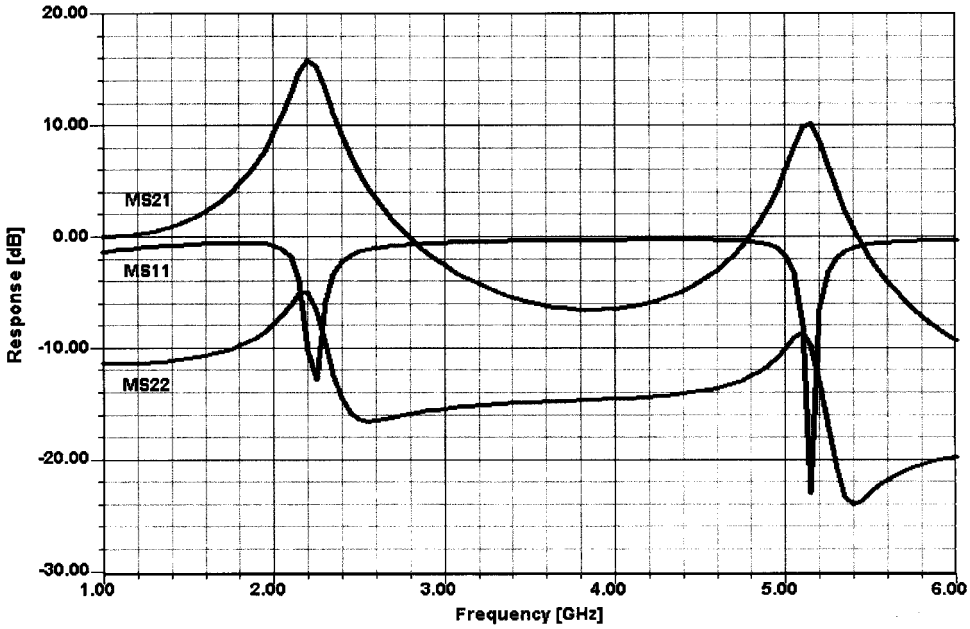


Figure 3-271 Frequency-dependent gain and matching characteristics of the amplifier. Although the circuit is designed for 2.2-GHz operation, the periodic responses of its distributed-element-based input and output matching networks result in a second gain peak near 5.1 GHz. As we shall see, however, instabilities may result at frequencies that cannot be intuited from this graph.

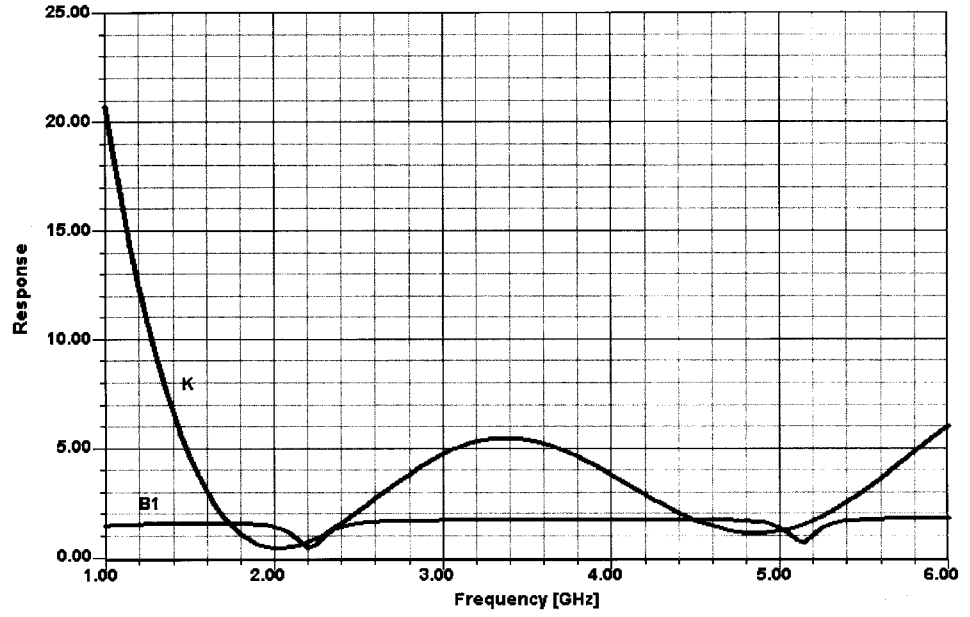


Figure 3-272 K and B_1 for the 2.2-GHz amplifier.

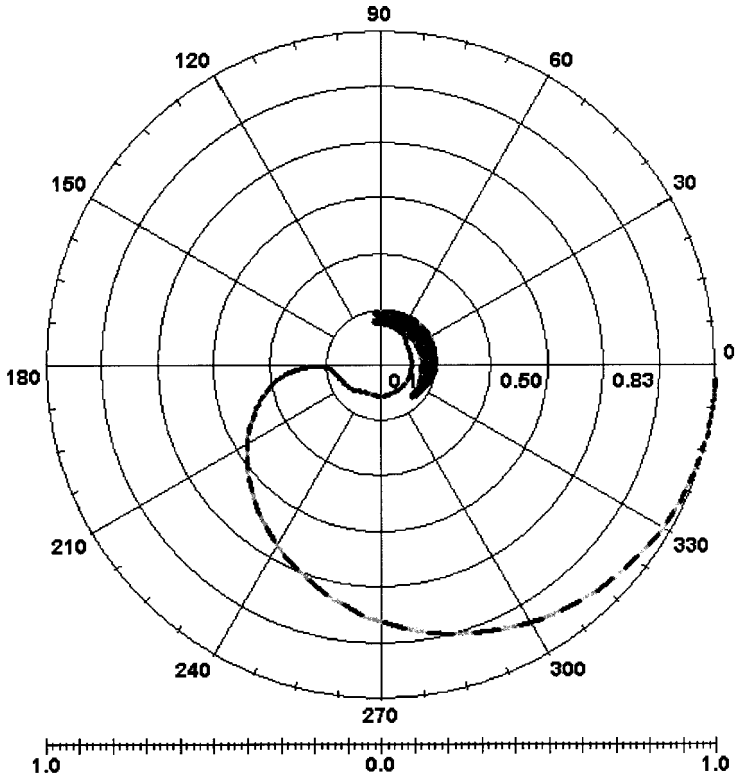


Figure 3-273 Polar Nyquist plot for the 2.2-GHz amplifier covering 1 kHz to 20 GHz.

- To travel in a clockwise direction on a polar chart means to decrease phase angle. On a magnitude–angle plot, the phase will become more negative as frequency is increased (although it may not be monotonic).
- If the path crosses the negative real axis on the polar chart, it will cross –180° on the magnitude–angle plot.
- Encirclement will occur if the phase angle continues to decrease to, say, –360°. If "unraveling" occurs, the phase angle will not stay below –360°. Figure 3-274 clearly shows that the phase angle decreases to about –180° but does not decrease any more. Rather, the angle returns to a positive value and oscillates between about –40° and +100°. Therefore an encirclement did not occur, and the circuit is stable.

An Unstable Case. Let us now change the gate bias to –2 V—a value that increases the amplifier gain enough to cause instability. Figures 3-275 and 3-276 show the results. From Figure 3-275 we can see that the origin is encircled in a clockwise direction. Figure 3-276 confirms this since the phase angle goes through –180° and continues without unraveling.

Approximate Frequency of Oscillation. The frequency where the Nyquist plot crosses –180° is often the approximate frequency of the unstable right-half-plane zeros and provides an estimate of the potential oscillation. For this example, there is a sharp crossing at 200 MHz that is characteristic of a circuit resonance. Note that the frequency is not related

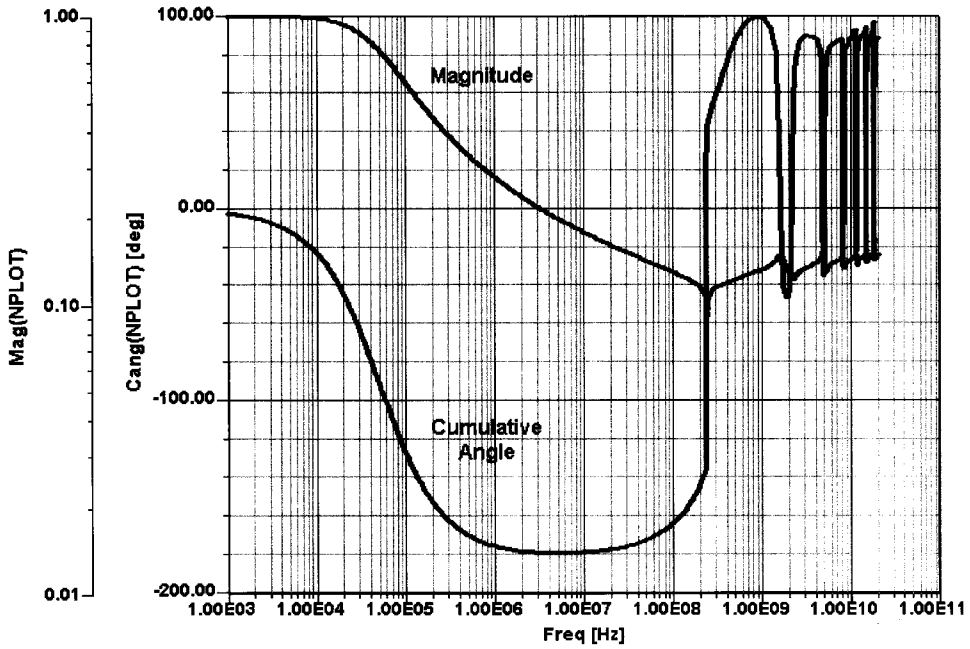


Figure 3-274 Magnitude and cumulative angle Nyquist plot for the 2.2-GHz amplifier covering 1 kHz to 20 GHz.

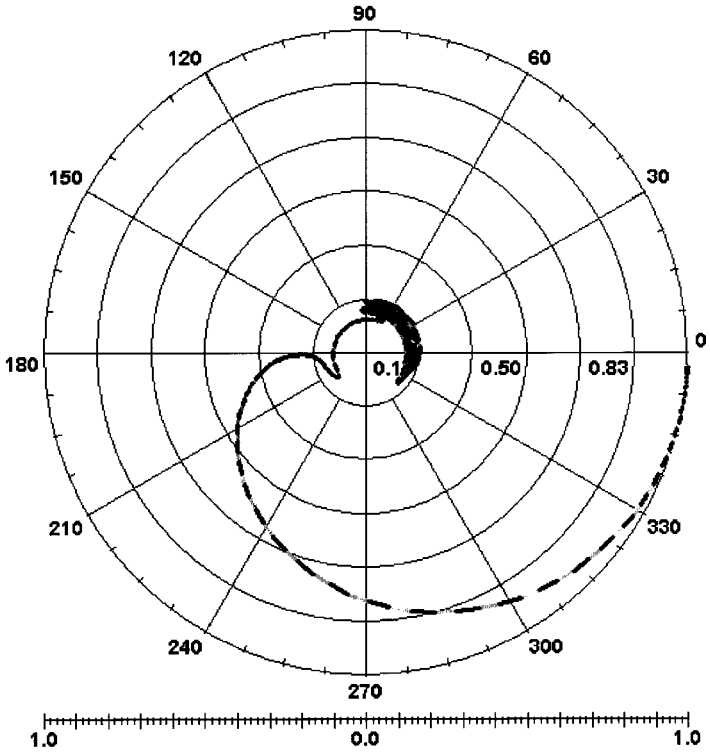


Figure 3-275 Polar Nyquist plot of the unstable amplifier.

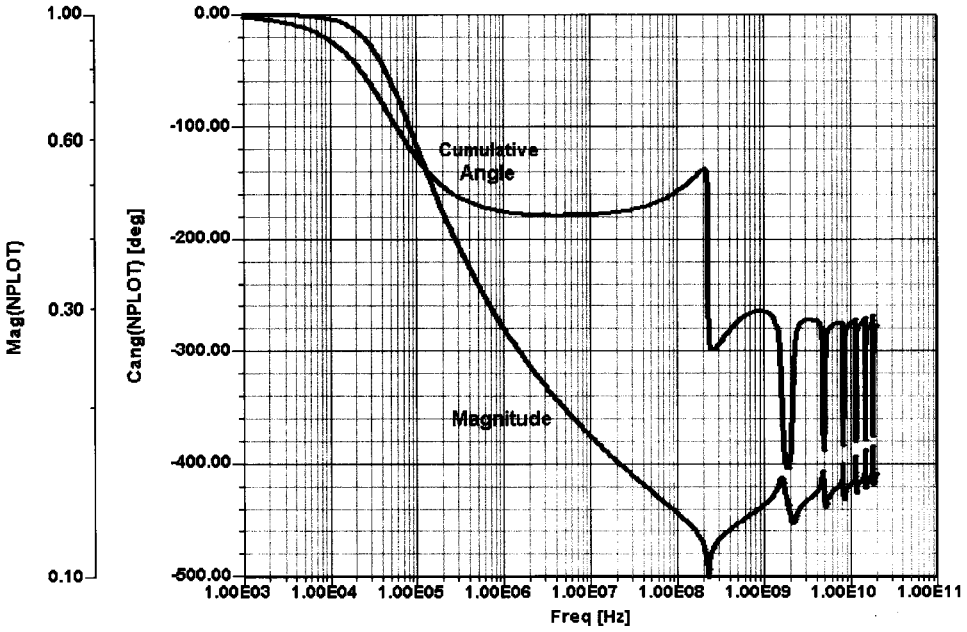


Figure 3-276 Magnitude–cumulative angle Nyquist plot of the unstable amplifier.

to a resonant frequency in a rigorous mathematical fashion, but it has been our experience that this frequency approximates the resonant frequency. To confirm the circuit's readiness for oscillation in the vicinity of this frequency, we evaluate it with Serenade's Oscillator Design Aid from 10 MHz to 1 GHz. The Design Aid finds an initial resonant frequency near 230.6 MHz—not too far from the –180° crossing frequency of 200 MHz of the Nyquist plot. Figure 3-277 shows this point.

Oscillator analysis finds a final oscillation frequency of 228.48 MHz. Figure 3-278 shows the resulting output spectrum at Port 2. At this point we decided to do a quick jump into Chapter 5 and take advantage of the CAD capability to predict the resulting phase noise, which is shown in Figure 3-279. As we compare this with other measured results, the approximately 228-MHz oscillator has a poor phase noise since the same phase noise is achievable at 1 GHz and higher. We remind the reader that this is an unwanted effect, but all of us who have built amplifiers will have experienced either low-frequency oscillations (motorboating) or very-high-frequency oscillation, sometimes caused by the components in the immediate vicinity of the transistor itself and found somewhere above 3 GHz. This case is really a demonstration case; a closer look reveals that the transconductance required for this condition is unrealistically high; but nonetheless it is a good example to demonstrate the stability issue.

Where Does Oscillation Begin? An interesting question to solve now is, "What gate bias voltage is needed for oscillation to begin?" Ansoft's Serenade Design Environment includes oscillator synchronous stability analysis, which uses a modification of the harmonic-balance technique to determine critical points, such as the voltage needed to "turn on" an oscillator. Mathematically, this is called a *Hopf bifurcation* because the behavior of

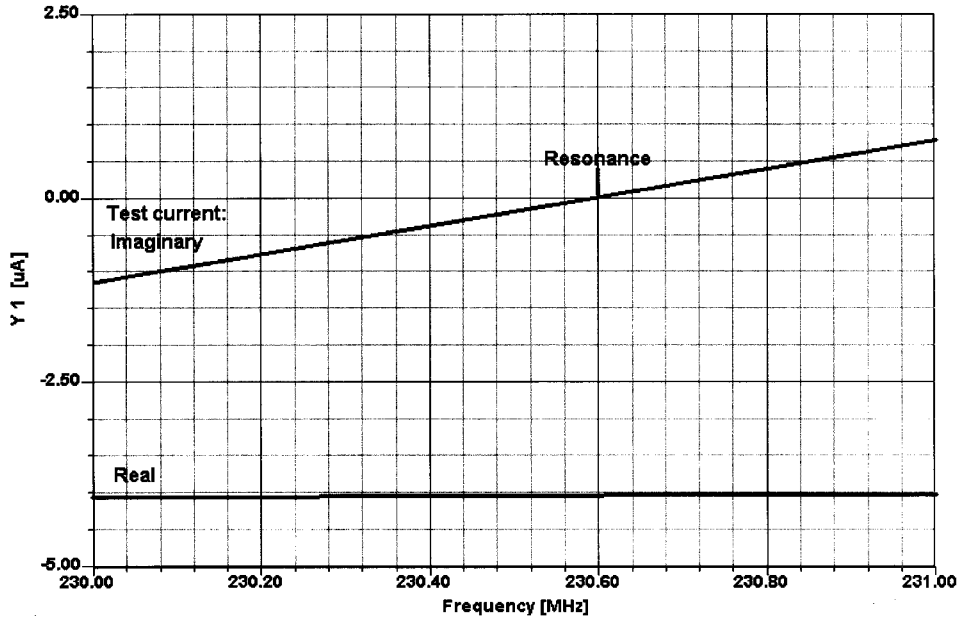


Figure 3-277 Graph from Serenade Oscillator Design Aid analysis of the amplifier. A resonant frequency is indicated where the imaginary part of the test current equals zero and the real part is negative.

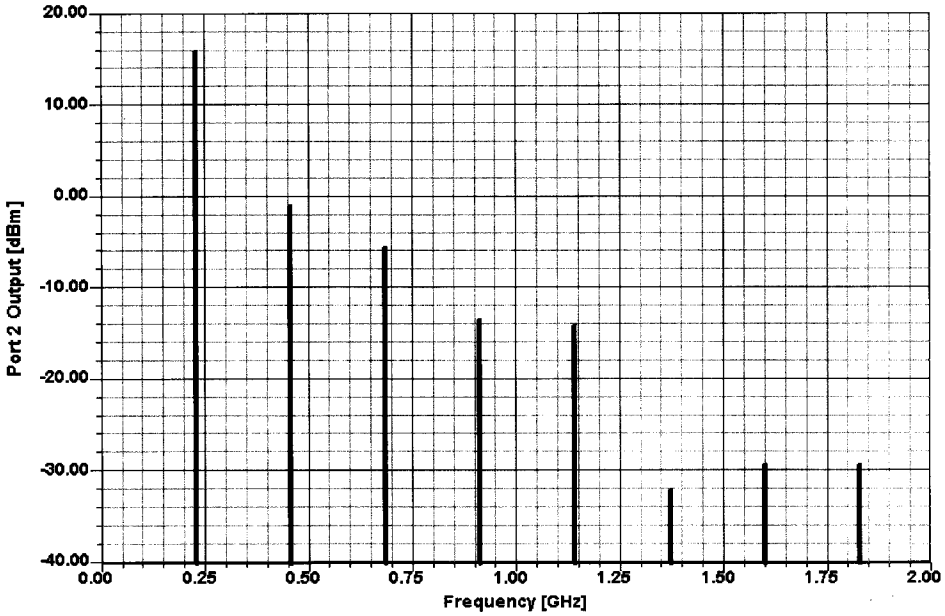


Figure 3-278 Predicted output spectrum of the unintended oscillator.

the circuit changes dramatically from a dc steady state to an oscillatory response. Figure 3-280 shows this effect by plotting output power versus gate bias. The turning point at –3.2 V (22 mW) is clearly visible, as is the bifurcation at –3.1 V (0 mW output).

Figure 3-280 can be interpreted as follows:

1. If the bias voltage is increased from, say, –4.0 V, the circuit will begin to oscillate at the bifurcation voltage and the output power will jump to 26 mW.
2. If the circuit is oscillating and the voltage is decreased from –2.0 V, the circuit will cease oscillating at the turning voltage of –3.2 V.
3. The branch between the turning point and the bifurcation point is unstable and cannot be realized in practice. This forms a hysteresis loop around points (–3.1 V, 0 mW), (–3.1 V, 26 mW), (–3.2 V, 22 mW), and (–3.2 V, 0 mW).

This amplifier stability analysis section is based on the stability analysis examples provided by Ansoft for the Serenade Design Environment.

An interesting example that combines both CAD and the feed-forward amplifier technique can be found in the Serenade Design Environment system simulator examples manual. This technique has been under evaluation for the last two years, but our feeling is that not enough measured data at frequencies above 500 MHz with reliable performance information have yet been made available in the literature. Along these lines, work done at Ansoft under a U.S. Air Force contract has provided very good insight into high-efficiency amplifiers at microwave frequencies [46].

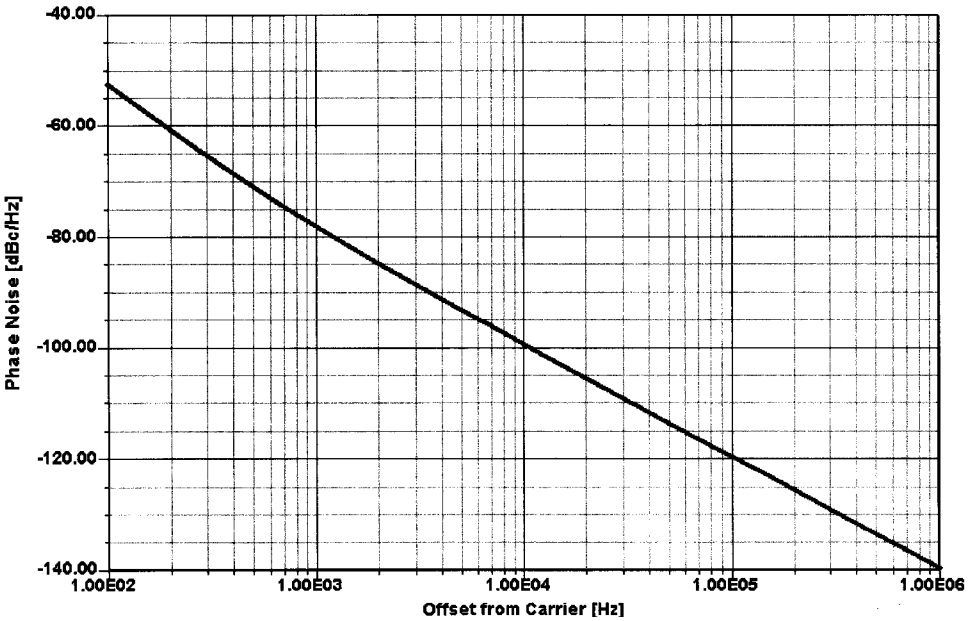


Figure 3-279 Predicted phase noise of the oscillating amplifier.

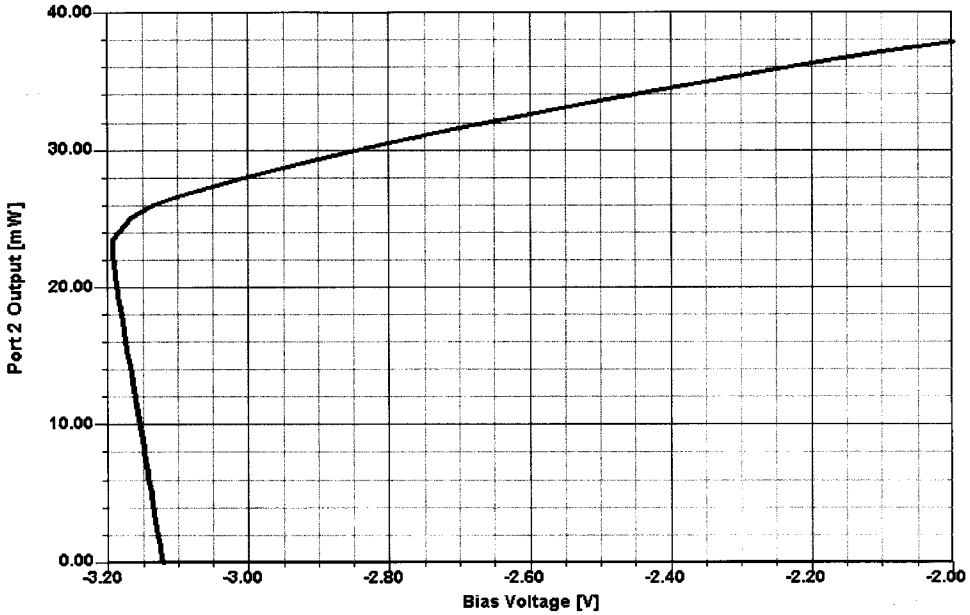


Figure 3-280 Output power versus gate bias for the unstable amplifier.

3-13 POWER AMPLIFIER DATASHEETS AND MANUFACTURER-RECOMMENDED APPLICATIONS

Having discussed the various aspects of BJTs and FETs, we have decided to use MOS examples from Ericsson and Motorola, and a bipolar example from Motorola, as good examples to show the technology required to build high-power transistor amplifiers in this frequency range. In the following pages, we reproduce, with permission, datasheets for Ericsson's PTF 10009 FET and Motorola's MRF186 FET, followed by Motorola's MRF899 BJT.

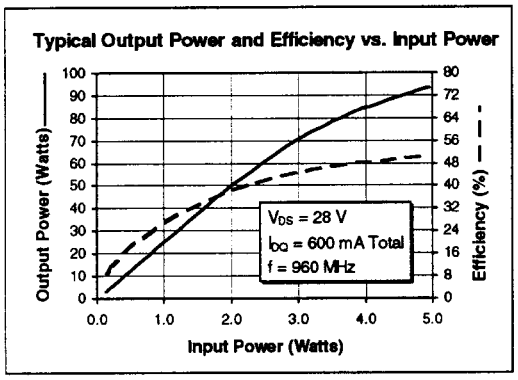
ERICSSON

PTF 10009
85 Watts, 1.0 GHz
LDMOS Field Effect Transistor

Description

The 10009 is a common source N-channel enhancement-mode lateral MOSFET intended for large signal amplifier applications to 1.0 GHz. It is rated at 85 watts minimum output power. Nitride surface passivation and gold metallization are used to ensure excellent device lifetime and reliability. 100% lot traceability is standard.

- Performance at 960 MHz, 28 Volts
  - Output Power = 85 Watts
  - Power Gain = 13.0 dB Typ
  - Efficiency = 50% Typ
- Gold Metallized
- Silicon Nitride Passivated
- Excellent Thermal Stability



Package 20230

Maximum Ratings

Parameter	Symbol	Value	Unit
Drain-Source Voltage	V_{DSS}	65	Vdc
Gate-Source Voltage	V_{GS}	±20	Vdc
Operating Junction Temperature	T_J	200	°C
Total Device Dissipation at T_{flange} = 25°C Above 25°C derate by	P_D	270 1.54	Watts W/°C
Storage Temperature Range	T_{STG}	-65 to 150	°C
Thermal Resistance (T_{flange} = 70°C)	$R_{\theta JC}$	0.65	°C/W

PTF 10009

ERICSSON

Electrical Characteristics (100% Tested)

Characteristic (per side)	Conditions	Symbol	Min	Typ	Max	Units
Drain-Source Breakdown Voltage	V_{GS} = 0 V, I_D = 25 mA	$V_{(BR)DSS}$	65	—	—	Volts
Drain-Source Leakage Current	V_{DS} = 28 V, V_{GS} = 0 V	I_{DSS}	—	—	1.0	mA
Gate Threshold Voltage	V_{DS} = 10 V, I_D = 75 mA	$V_{GS(th)}$	—	2.0	—	Volts
Forward Transconductance	V_{DS} = 10 V, I_D = 3 A	g_{fs}	—	2.8	—	Siemens

Dynamic Characteristics

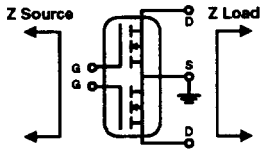
Characteristic (per side)	Symbol	Min	Typ	Max	Units
Input Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{iss}	—	90	—	pF
Output Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{oss}	—	36	—	pF
Reverse Transfer Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{rss}	—	1.9	—	pF

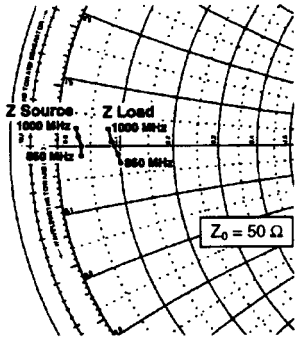
RF Specifications (100% Tested)

Characteristic	Symbol	Min	Typ	Max	Units
Gain (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz)	G_{ps}	12.0	13.0	—	dB
Drain Efficiency (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz)	η	47	50	—	%
Load Mismatch Tolerance (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz— all phase angles at frequency of test)	Ψ	—	—	5:1	—

Impedance Data (data shown for fixed-tuned broadband circuit)

(V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA)



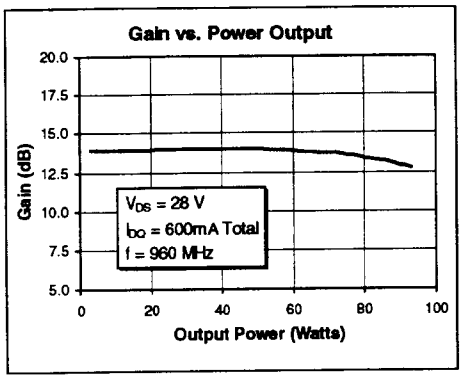


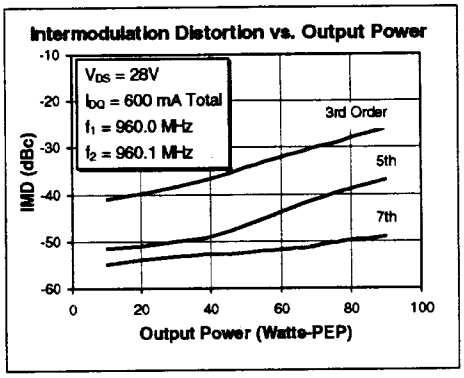
Frequency	Z Source		Z Load	
MHz	R	jX	R	jX
860	1.76	-0.78	5.00	-1.50
900	1.80	-0.05	4.80	-0.78
960	1.58	0.69	4.24	0.36
1000	1.39	1.35	3.95	1.41

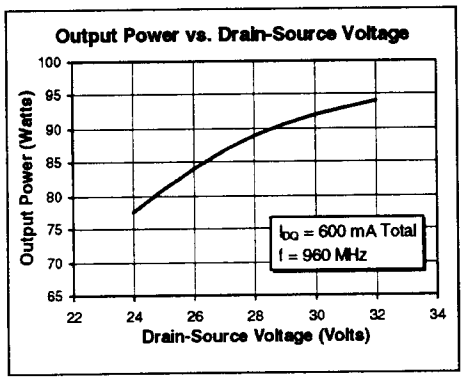
ERICSSON

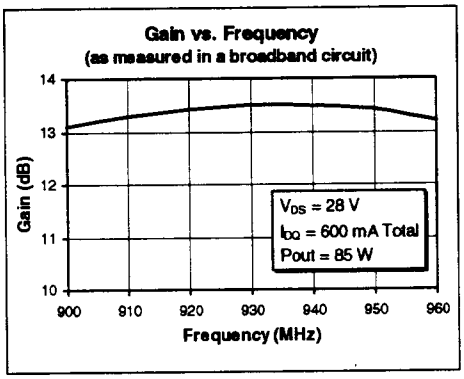
PTF 10009

Typical Performance





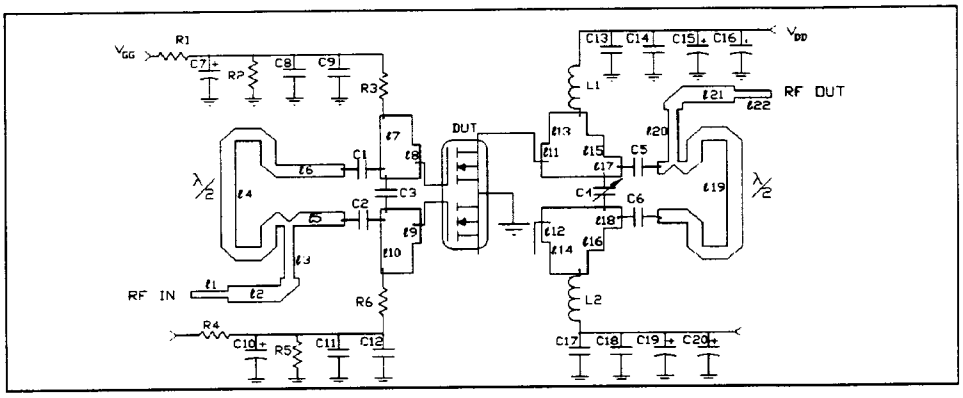




PTF 10009

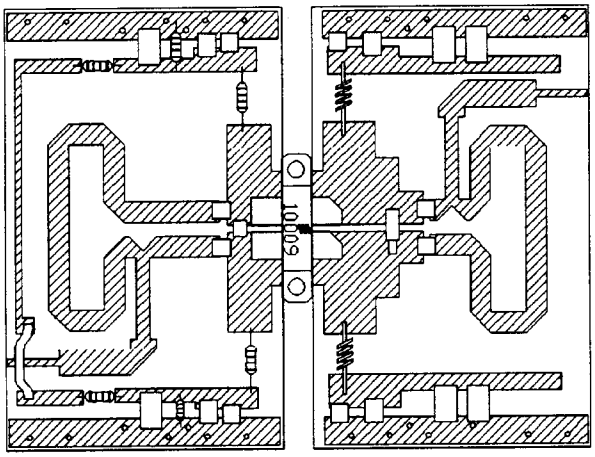
ERICSSON

Test Circuit



Schematic for f = 960 MHz

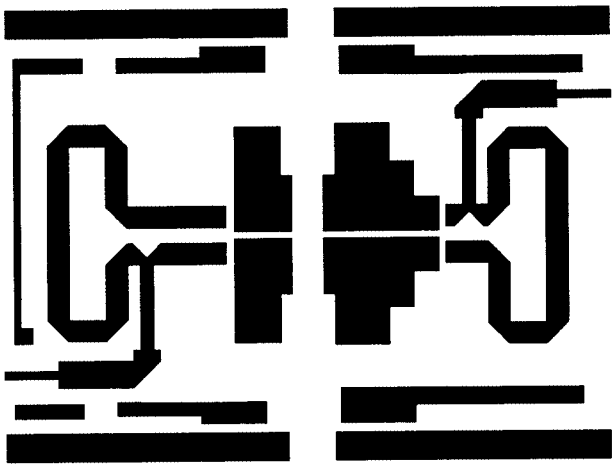
DUT	10009
C1-2, C5-6, C9, C12-13, C17	33 pF, Capacitor ATC 100 B
C3	11 pF, Capacitor ATC 100 B
C4	6.0 pF, Variable Capacitor, JMC 5701
C7, C10	10 µF, +10 V Electrolytic Capacitor
C8, C11, C14, C18	0.01 µF, Capacitor ATC 100 B
C15, C16, C19, C20	10 µF, +30 V Electrolytic Capacitor
L1. L2	4 Turn, #20 AWG, .120" I.D.
R1, R2, R4, R5	1.0 K, Ω Resistor
R3, R6	5.1 K, 1/4 Ω Resistor
ℓ1, ℓ22	50 Ω, .030 λ
ℓ2, ℓ21	20 Ω, .080 λ
ℓ3, ℓ20	32 Ω, .191 λ
ℓ4, ℓ19	25 Ω, .500 λ
ℓ5, ℓ6	25 Ω, .091 λ
ℓ7, ℓ10	7 Ω, .056 λ
ℓ8, ℓ9	13.0 Ω, .017 λ
ℓ11, ℓ12	13.0 Ω, .017 λ
ℓ13, ℓ14	7.0 Ω, .064 λ
ℓ15, ℓ16	10.0 Ω, .029 λ
ℓ17, ℓ18	19.0 Ω, .028 λ
Circuit Board	.031" Thick, ε_r = 4.0, AlliedSignal, G200



Parts Layout (not to scale)

ERICSSON

PTF 10009



Artwork (1 inch ——— *)*

Ericsson Components
RF Power Products
675 Jarvis Drive
Morgan Hill, CA 95037 USA
Telephone: 408-778-9434

1-877-GOLDMOS
(1-877-465-3667)
e-mail: rfpower@ericsson.com
www.ericsson.com/rfpower

Specifications subject to change
without notice.
L1
© 1997 Ericsson Inc.
EUS/KR 1301-PTF 10009 Uen Rev. B 09-29-98

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF186/D

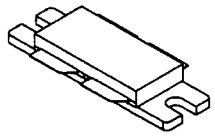
The RF MOSFET Line
RF Power Field-Effect Transistor
N–Channel Enhancement–Mode Lateral MOSFET

MRF186

120 W, 1.0 GHz
28 VOLTS
LATERAL N–CHANNEL
BROADBAND
RF POWER MOSFET

Designed for broadband commercial and industrial applications at frequencies from 800 MHz to 1.0 GHz. The high gain and broadband performance of this device makes it ideal for large–signal, common source amplifier applications in 28 volt base station equipment.

- Guaranteed Performance @ 960 MHz, 28 Volts
  Output Power — 120 Watts (PEP)
  Power Gain — 11 dB
  Efficiency — 30%
  Intermodulation Distortion — –28 dBc
- Excellent Thermal Stability
- 100% Tested for Load Mismatch Stress at all Phase Angles with 5:1 VSWR @ 28 Vdc, 960 MHz, 120 Watts CW



CASE 375B–02, STYLE 2

MAXIMUM RATINGS (2)

Rating	Symbol	Value	Unit
Drain–Source Voltage	V_{DSS}	65	Vdc
Drain–Gate Voltage (R_{GS} = 1 MΩ)	V_{DGR}	65	Vdc
Gate–Source Voltage	V_{GS}	±20	Vdc
Drain Current — Continuous	I_D	14	Adc
Total Device Dissipation @ T_C = 70°C Derate above 70°C	P_D	162.5 1.25	Watts W/°C
Storage Temperature Range	T_{stg}	– 65 to +150	°C
Operating Junction Temperature	T_J	200	°C

THERMAL CHARACTERISTICS (2)

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	0.8	°C/W

NOTE – **CAUTION** – MOS devices are susceptible to damage from electrostatic charge. Reasonable precautions in handling and packaging MOS devices should be observed.

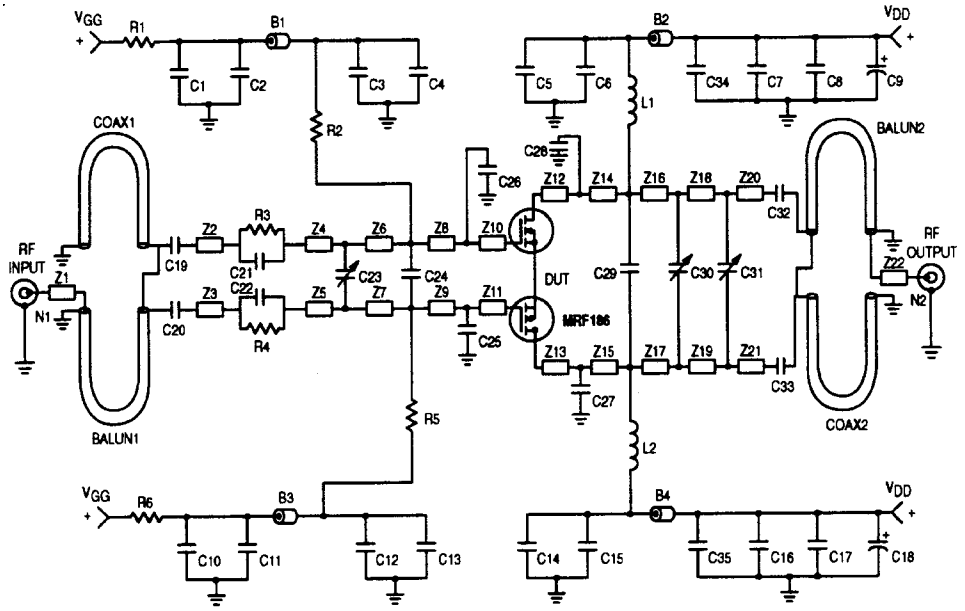
© Motorola, Inc. 1999

MOTOROLA

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted)

Characteristic	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS (1)					
Drain–Source Breakdown Voltage (V_{GS} = 0 Vdc, I_D = 50 μAdc)	$V_{(BR)DSS}$	65	—	—	Vdc
Zero Gate Voltage Drain Current (V_{DS} = 28 Vdc, V_{GS} = 0)	I_{DSS}	—	—	1	μAdc
Gate–Source Leakage Current (V_{GS} = 20 Vdc, V_{DS} = 0)	I_{GSS}	—	—	1	μAdc
ON CHARACTERISTICS (1)					
Gate Quiescent Voltage (V_{DS} = 26 Vdc, I_D = 300 μAdc Per Side)	$V_{GS(th)}$	2.5	3	4	Vdc
Gate Quiescent Voltage (V_{DS} = 26 Vdc, I_D = 300 mAdc Per Side)	$V_{GS(Q)}$	3.3	4.2	5	Vdc
Delta Gate Threshold Voltage (Side to Side) (V_{DS} = 28 V, I_D = 300 mA Per Side)	$\Delta V_{GS(Q)}$	—	—	0.3	Vdc
Drain–Source On–Voltage (V_{GS} = 10 Vdc, I_D = 3 Adc Per Side)	$V_{DS(on)}$	—	0.58	0.7	Vdc
Forward Transconductance (V_{DS} = 10 Vdc, I_D = 3 Adc Per Side)	g_{fs}	2.4	2.8	—	S
DYNAMIC CHARACTERISTICS (1)					
Input Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{iss}	—	177	—	pF
Output Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{oss}	—	45	—	pF
Reverse Transfer Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{rss}	—	3.4	—	pF
FUNCTIONAL CHARACTERISTICS (In Motorola Test Fixture) (2)					
Two–Tone Common Source Amplifier Power Gain (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	G_{ps}	11	12.2	—	dB
Two–Tone Drain Efficiency (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	η	30	35	—	%
3rd Order Intermodulation Distortion (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	IMD	—	–32	–28	dBc
Input Return Loss (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	IRL	9	16	—	dB
Two–Tone Common Source Amplifier Power Gain (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	G_{ps}	—	12	—	dB
Two–Tone Drain Efficiency (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	η	—	33	—	%
3rd Order Intermodulation Distortion (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	IMD	—	–32	—	dBc
Input Return Loss (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	IRL	—	16	—	dB
Output Mismatch Stress (V_{DD} = 28 Vdc, P_{out} = 120 W CW, I_{DQ} = 2 x 400 mA, f = 960 MHz, VSWR = 5:1, All Phase Angles at Frequency of Tests)	Ψ	No Degradation In Output Power Before and After Test			

(1) Each side of device measured separately.
(2) Device measured in push–pull configuration.



B1 – B4	Fair Rite Products Short Ferrit Bead, 2743021446	C31	0.8 – 8.0 pF, Variable Capacitor, Johanson Gigatrim
C1, C7, C8, C10, C16, C17	10 µF, 50 V, Tantalum	L1, L2	3 Turns, #20 AWG, IDIA 0.126", 24.7 nH
C2, C11, C34, C35	0.1 µF, Chip Capacitor	N1, N2	Type N Connectors
C3, C6, C12, C15	330 pF, Chip Capacitor	R1, R6	1 kΩ, 1/4 W, Carbon Resistor
C4, C5, C13, C14, C19, C20, C32, C33	47 pF, Chip Capacitor	R2, R5	1.2 kΩ, 0.1 W, Chip Resistor
C9, C18	250 µF, 50 V, Electrolytic Capacitor	R3, R4	75 Ω, 0.1 W, Chip Resistor
C21, C22	12 pF, Chip Capacitor	Z1 – Z22	Microstrip (See Component Placement)
C23, C30	0.6 – 4.5 pF, Variable Capacitor, Johanson Gigatrim	Balun1, Balun2, Coax1, Coax2	2.20" 50 Ω, 0.086" OD Semi–Rigid Coax
C24, C25, C26	5.1 pF, Chip Capacitor	Board	1/32" Glass Teflon®, ε_r = 2.55
C27, C28	3.9 pF, Chip Capacitor		

Figure 1. 930 – 960 MHz Test Circuit Schematic

TYPICAL CHARACTERISTICS

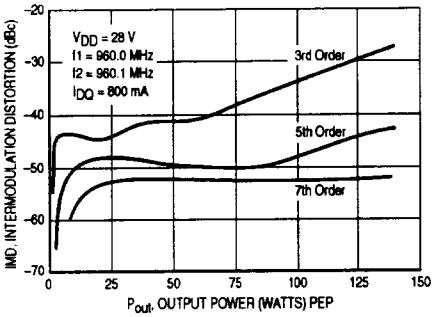


Figure 2. Intermodulation Distortion versus Output Power

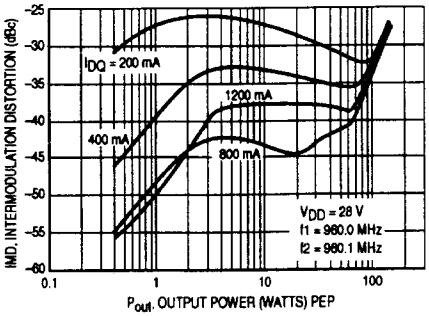


Figure 3. Intermodulation Distortion versus Output Power

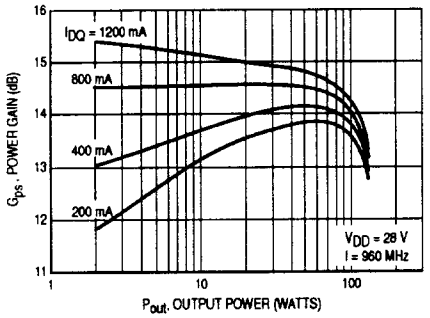


Figure 4. Power Gain versus Output Power

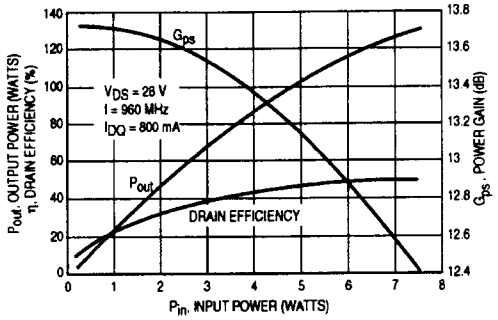


Figure 5. Output Power versus Input Power

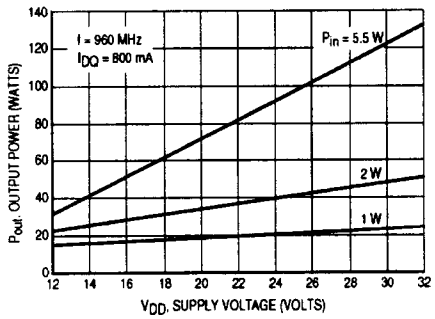


Figure 6. Output Power versus Supply Voltage

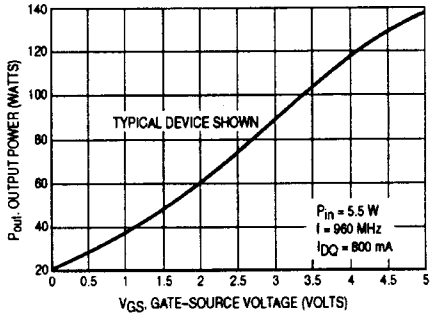


Figure 7. Output Power versus Gate Voltage

TYPICAL CHARACTERISTICS

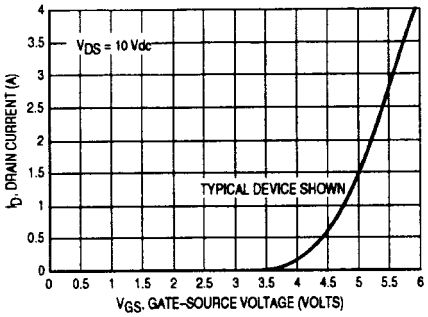


Figure 8. Drain Current versus Gate Voltage

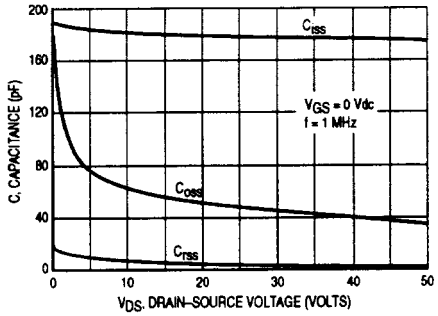


Figure 9. Capacitance versus Voltage

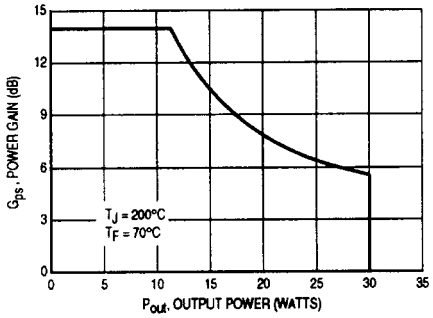


Figure 10. DC Safe Operating Area

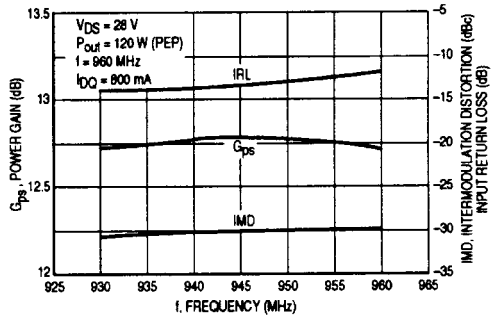
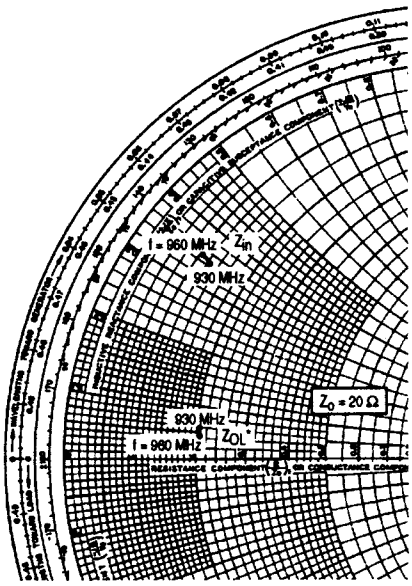


Figure 11. Broadband Circuit Performance



V_{CC} = 28 V, I_{DQ} = 2 × 400 mA, P_{out} = 120 Watts (PEP)

f MHz	Z_{in} Ω	Z_{OL}^* Ω
930	2.5 + j6.9	4.3 + j1.2
945	2.5 + j7.0	4.3 + j1.0
960	2.2 + j7.1	4.3 + j0.9

Z_{in} = Conjugate of impedance.

Z_{OL}^* = Conjugate of the optimum load impedance at a given output power, voltage, IMD, bias current, efficiency and frequency.

Note: Z_{OL}^* was chosen based on tradeoffs between gain, output power, drain efficiency and intermodulation performance. Impedances shown represent a single channel (1/2 of MRF186) impedance measurement.

Figure 12. Series Equivalent Input and Output Impedance

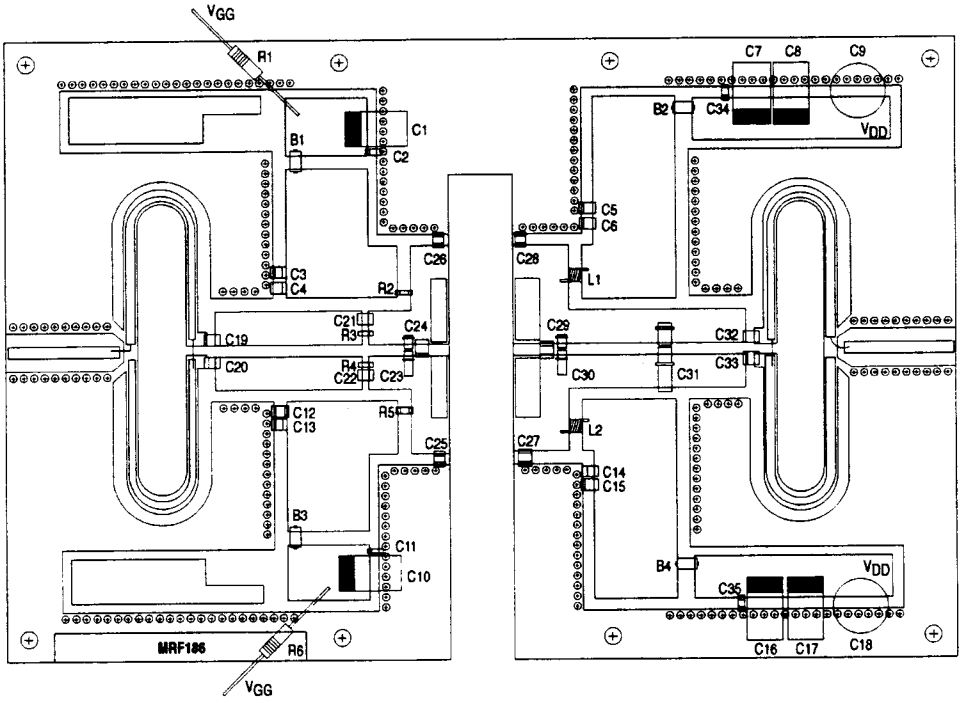
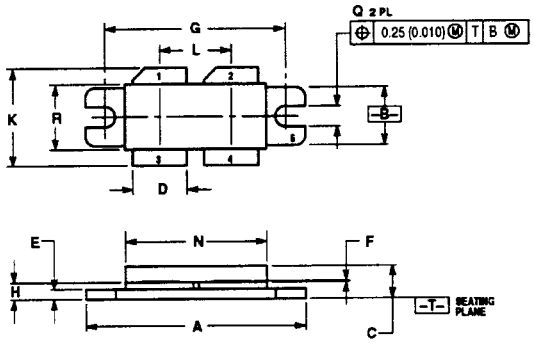


Figure 13. Component Placement Diagram of 930 – 960 MHz Broadband Test Fixture

PACKAGE DIMENSIONS



NOTES:
1. DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2. CONTROLLING DIMENSION: INCH.

DIM	INCHES MIN	INCHES MAX	MILLIMETERS MIN	MILLIMETERS MAX
A	1.330	1.350	33.78	34.29
B	0.375	0.395	9.52	10.03
C	0.180	0.205	4.57	5.21
D	0.320	0.340	8.13	8.64
E	0.060	0.070	1.52	1.77
F	0.004	0.006	0.11	0.15
G	1.100 BSC		27.94 BSC	
H	0.082	0.097	2.09	2.46
K	0.580	0.620	14.73	15.75
L	0.435 BSC		11.05 BSC	
N	0.845	0.875	21.46	22.23
Q	0.118	0.130	3.00	3.30
R	0.390	0.410	9.91	10.41

STYLE 1:
PIN 1. COLLECTOR
2. COLLECTOR
3. BASE
4. BASE
5. EMITTER

CASE 375A–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Motorola data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and Ⓜ are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

Mfax is a trademark of Motorola, Inc.

How to reach us:
USA/EUROPE/Locations Not Listed: Motorola Literature Distribution; P.O. Box 5405, Denver, Colorado 80217. 303–675–2140 or 1–800–441–2447

Mfax™: RMFAX0@email.sps.mot.com – TOUCHTONE 602–244–6609 – US & Canada ONLY 1–800–774–1848

INTERNET: http://motorola.com/sps

JAPAN: Nippon Motorola Ltd.: SPD, Strategic Planning Office, 4–32–1, Nishi–Gotanda, Shinagawa–ku, Tokyo 141, Japan. 81–3–5487–8488

ASIA/PACIFIC: Motorola Semiconductors H.K. Ltd.; 8B Tai Ping Industrial Park, 51 Ting Kok Road, Tai Po, N.T., Hong Kong. 852–26629298

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF899/D

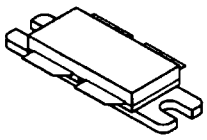
The RF Line
NPN Silicon RF Power Transistor

MRF899

150 W, 900 MHz
RF POWER
TRANSISTOR
NPN SILICON

Designed for 26 Volt UHF large–signal, common emitter, Class AB linear amplifier applications in industrial and commercial FM/AM equipment operating in the range 800–960 MHz.

- Specified 26 Volt, 900 MHz Characteristics
  Output Power = 150 Watts (PEP)
  Minimum Gain = 8.0 dB @ 900 MHz, Class AB
  Minimum Efficiency = 35% @ 900 MHz, 150 Watts (PEP)
  Maximum Intermodulation Distortion –28 dBc @ 150 Watts (PEP)
- Characterized with Series Equivalent Large–Signal Parameters from 800 to 960 MHz
- Silicon Nitride Passivated
- 100% Tested for Load Mismatch Stress at all Phase Angles with 5:1 VSWR @ 26 Vdc, and Rated Output Power
- Gold Metallized, Emitter Ballasted for Long Life and Resistance to Metal Migration
- Circuit board photomaster available upon request by contacting RF Tactical Marketing in Phoenix, AZ.



CASE 375A–01, STYLE 1

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Collector–Emitter Voltage	V_{CEO}	28	Vdc
Collector–Emitter Voltage	V_{CES}	60	Vdc
Emitter–Base Voltage	V_{EBO}	4.0	Vdc
Collector–Current — Continuous	I_C	25	Adc
Total Device Dissipation @ T_C = 25°C Derate above 25°C	P_D	230 1.33	Watts W/°C
Storage Temperature Range	T_{stg}	–65 to +150	°C

THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	0.75	°C/W

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS					
Collector–Emitter Breakdown Voltage (I_C = 100 mAdc, I_B = 0)	$V_{(BR)CEO}$	28	37	—	Vdc
Collector–Emitter Breakdown Voltage (I_C = 50 mAdc, V_{BE} = 0)	$V_{(BR)CES}$	60	85	—	Vdc
Emitter–Base Breakdown Voltage (I_E = 10 mAdc, I_C = 0)	$V_{(BR)EBO}$	4.0	4.9	—	Vdc
Collector Cutoff Current (V_{CE} = 30 Vdc, V_{BE} = 0)	I_{CES}	—	—	10	mAdc
ON CHARACTERISTICS					
DC Current Gain (I_{CE} = 1.0 Adc, V_{CE} = 5.0 Vdc)	h_{FE}	30	75	120	—
DYNAMIC CHARACTERISTICS					
Output Capacitance (V_{CB} = 26 Vdc, I_E = 0, f = 1.0 MHz) (1)	C_{ob}	—	75	—	pF

(1) For information only. This part is collector matched.

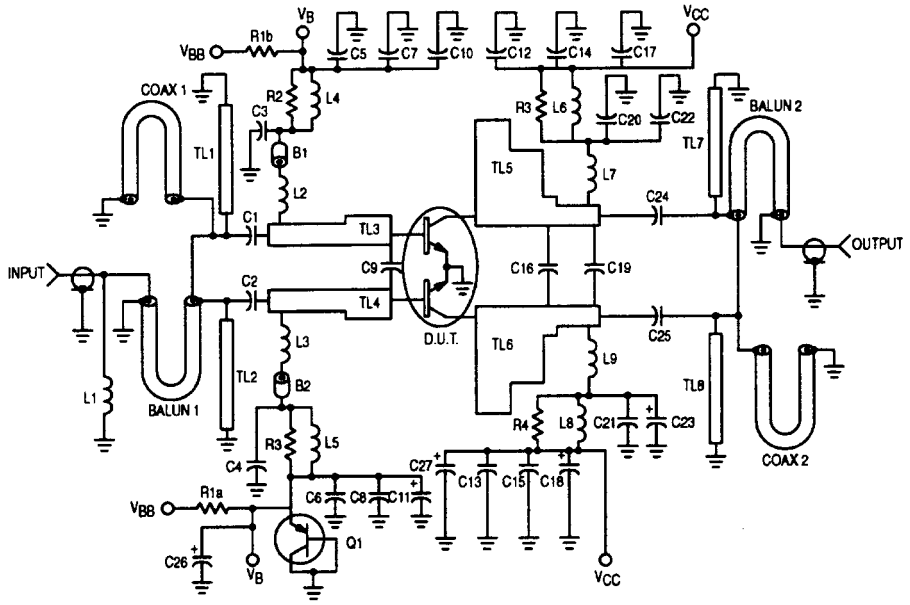
(continued)

© Motorola, Inc. 1997

MOTOROLA

ELECTRICAL CHARACTERISTICS — continued (T_C = 25°C unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
FUNCTIONAL CHARACTERISTICS					
Common–Emitter Amplifier Power Gain V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	G_{pe}	8.0	9.0	—	dB
Collector Efficiency V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	η	30	40	—	%
3rd Order Intermodulation Distortion V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	IMD	—	–32	–28	dBc
Output Mismatch Stress V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz, VSWR = 5:1 (all phase angles)	Ψ	No Degradation in Output Power Before and After Test			



B1, B2 — Ferrite Bead, Ferroxcube #56–590–65–3B
C1, C2, C24, C25 — 43 pF, B Case, ATC Chip Capacitor
C3, C4, C20, C21 — 100 pF, B Case, ATC Chip Capacitor
C5, C6, C12, C13 — 1000 pF, B Case, ATC Chip Capacitor
C7, C8, C14, C15 — 1800 pF, AVX Chip Capacitor
C9 — 9.1 pF, A Case, ATC Chip Capacitor
C10, C11, C17, C18, C22, C23 — 10 µF, Electrolytic Capacitor Panasonic
C16 — 3.9 pF, B Case, ATC Chip Capacitor
C19 — 0.8 pF, B Case, ATC Chip Capacitor
C26 — 200 µF, Electrolytic Capacitor Mallory Sprague
C27 — 500 µF Electrolytic Capacitor

L1 — 5 Turns 24 AWG IDIA 0.059″ Choke, 19.8 nH
L2, L3, L7, L9 — 4 Turns 20 AWG IDIA 0.163″ Choke
L4, L5, L6, L8 — 12 Turns 22 AWG IDIA 0.140″ Choke
N1, N2 — Type N Flange Mount, Omni Spectra
Q1 — Bias Transistor BD136 PNP
R2, R3, R4, R5 — 4.0 x 39 Ohm 1/8 W Chips in Parallel
R1a, R1b — 56 Ohm 1.0 W
TL1–TL8 — See Photomaster
Balun1, Balun2, Coax 1, Coax 2 — 2.20″ 50 Ohm 0.086″ o.d. Semi–rigid Coax, Micro Coax
Board — 1/32″ Glass Teflon, ε_r = 2.55″ Arlon (GX–0300–55–22)

Figure 1. 900 MHz Power Gain Test Circuit

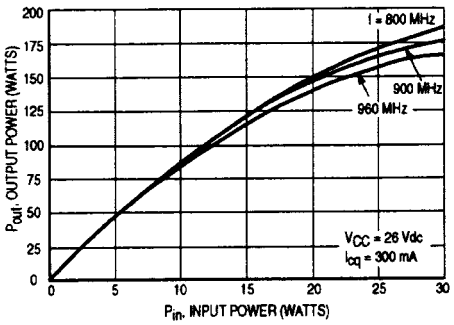


Figure 2. Output Power versus Input Power

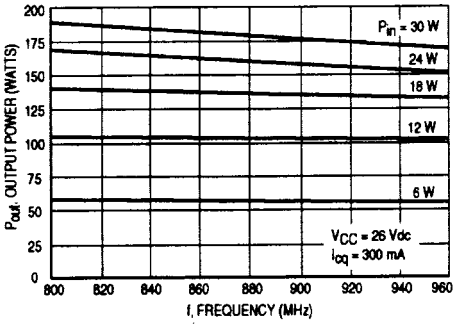


Figure 3. Output Power versus Frequency

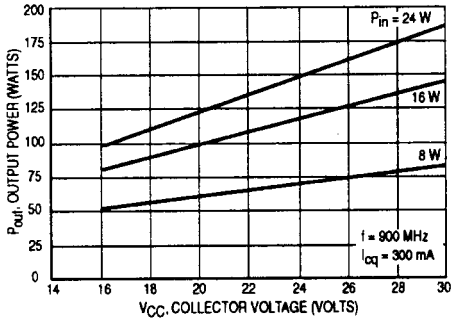


Figure 4. Output Power versus Supply Voltage

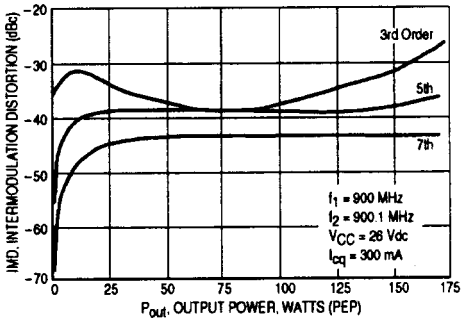


Figure 5. Intermodulation versus Output Power

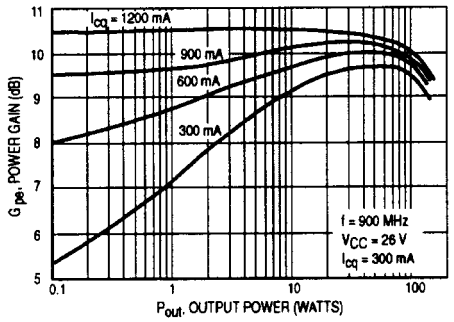


Figure 6. Power Gain versus Output Power

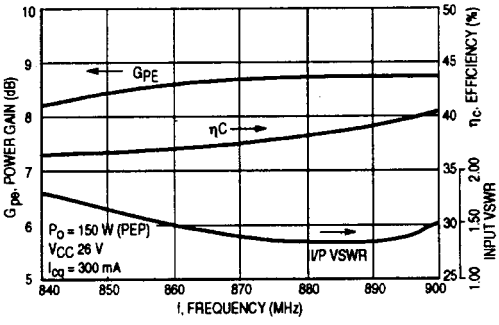
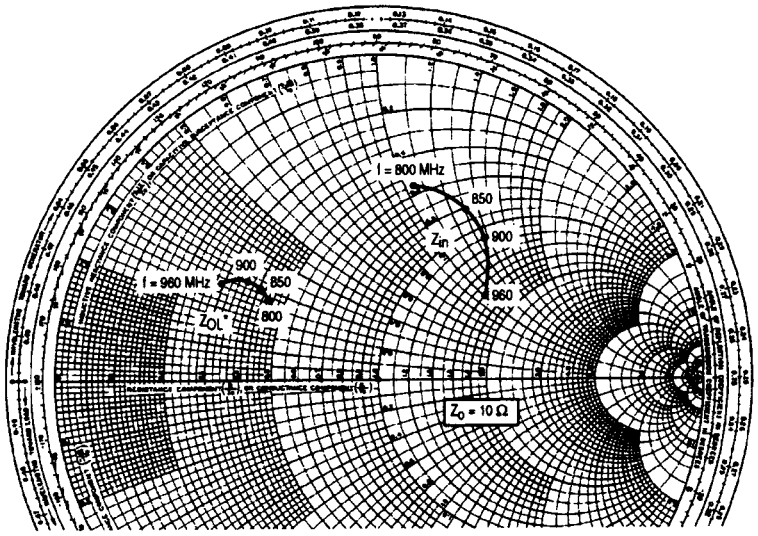


Figure 7. Broadband Test Fixture Performance



f MHz	Z_{in} Ohms	Z_{OL}^* Ohms
800	5.51 + j10.6	4.52 + j2.64
850	8.17 + j13.2	4.21 + j2.98
900	11.2 + j13.8	3.68 + j2.97
960	16.8 + j10.1	2.98 + j2.71

Z_{OL}^* = Conjugate of optimum load impedance into which the device operates at a given output power, voltage and frequency.

NOTE: Z_{in} & Z_{OL}^* are given from base–to–base and collector–to–collector respectively

Figure 8. Input and Output Impedances with Circuit Tuned for Maximum Gain @ P_O = 150 W (PEP), V_{CC} = 26 V

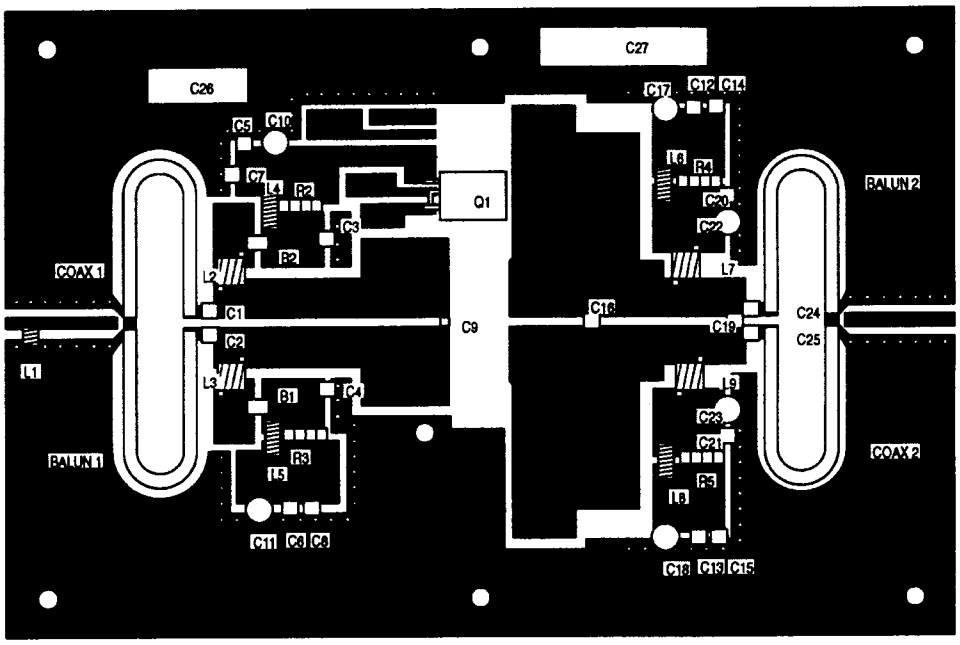
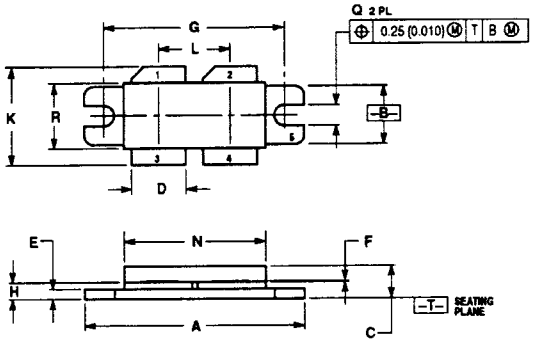


Figure 9. MRF899 Test Fixture Component Layout

PACKAGE DIMENSIONS



NOTES:
1. DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2. CONTROLLING DIMENSION: INCH.

DIM	INCHES MIN	INCHES MAX	MILLIMETERS MIN	MILLIMETERS MAX
A	1.330	1.350	33.79	34.29
B	0.375	0.395	9.52	10.03
C	0.160	0.205	4.57	5.21
D	0.320	0.340	8.13	8.64
E	0.060	0.070	1.52	1.77
F	0.004	0.006	0.11	0.15
G	1.100 BSC		27.94 BSC	
H	0.062	0.097	2.08	2.46
K	0.580	0.620	14.73	15.75
L	0.435 BSC		11.05 BSC	
N	0.845	0.875	21.46	22.23
Q	0.116	0.130	3.00	3.30
R	0.390	0.410	9.91	10.41

STYLE 1:
PIN 1. COLLECTOR
2. COLLECTOR
3. BASE
4. BASE
5. EMITTER

CASE 375A–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Motorola data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and M are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

Mfax is a trademark of Motorola, Inc.

How to reach us:
USA/EUROPE/Locations Not Listed: Motorola Literature Distribution;
P.O. Box 5405, Denver, Colorado 80217. 303–675–2140 or 1–800–441–2447

Mfax™: RMFAX0@email.sps.mot.com – TOUCHTONE 602–244–6609
– US & Canada ONLY 1–800–774–1848

INTERNET: http://motorola.com/sps

JAPAN: Nippon Motorola Ltd.: SPD, Strategic Planning Office, 4–32–1, Nishi–Gotanda, Shinagawa–ku, Tokyo 141, Japan. 81–3–5487–8488

ASIA/PACIFIC: Motorola Semiconductors H.K. Ltd.; 8B Tai Ping Industrial Park, 51 Ting Kok Road, Tai Po, N.T., Hong Kong. 852–26629298

REFERENCES

1. Ulrich L. Rohde, Anthony M. Pavio, and Robert A. Pucel, "Accurate Noise Simulation of Microwave Amplifiers Using CAD," *Microwave J.*, **31**(12), 130–141, December 1988.
2. H. T. Friis, "Noise Figures of Radio Receivers," *Proc. IRE*, **32**, 419–422, July 1944.
3. "IRE Standards on Methods of Measuring Noise in Linear Twoports," *Proc. IRE*, **48**, 60–68, January 1960.
4. H. Rothe and W. Dhalke, "Theory of Noisy Fourpoles," *Proc. IRE*, **44**, 811–818, June 1956.
5. "Representation of Noise in Linear Twoports," *Proc. IRE*, **48**, 69–74, January 1960.
6. H. A. Haus and R. B. Adler, *Circuit Theory of Linear Noisy Networks*, Cambridge, MA: MIT Press, 1959/New York: Wiley, 1959.
7. H. Fukui, "Noise Performance of Microwave Transistors," *IEEE Trans. Electron Devices*, **ED-13**, 329–341, March 1966.
8. S. Iversen, "The Effect of Feedback on Noise Figure," *Proc. IEEE*, **63**, 540–542, March 1975.
9. L. Besser, "Stability Considerations of Low-Noise Transistor Amplifiers with Simultaneous Noise and Power Match," *IEEE MTT-S International Microwave Symposium Digest*, 1975, pp. 327–329.
10. G. D. Vendelin, "Feedback Effects on the Noise Performance of GaAs Mesfets," *IEEE MTT-S International Microwave Symposium Digest*, 1975, pp. 324–326.
11. W. A. Suter, "Feedback and Parasitic Effects on Noise," *Microwave J.* **26**(2), 123–129, February 1983.
12. K. Hartmann and M. J. O. Strutt, "Changes in Four Noise Parameters Due to General Changes in Linear Two-Port Circuits," *IEEE Trans. Electron Devices*, **ED-20**(10), 874–877, October 1973.
13. H. Fukui, "Available Power Gain, Noise Figure, and Noise Measure of Two-Ports and Their Graphical Representation," *IEEE Trans. Circuit Theory*, **CT-13**, 137–142, June 1966.
14. C. R. Poole and D. K. Paul, "Optimum Noise Measure Terminations for Microwave Transistor Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**(11), 1254–1257, November 1985.
15. H. Hillbrand and P. H. Russer, "An Efficient Method for Computer Aided Noise Analysis of Linear Amplifier Networks," *IEEE Trans. Circuits Syst.*, **CAS-23**(4), 235–238, April 1976.
16. W. E. Pastori, *Topics on Noise*, Eaton Electronics Instrumentation Division (EID) Seminar Notes, Los Angeles, 1983.
17. W. E. Pastori, "Image and Second Stage Corrections Resolve Noise Figure Measurement Confusion," *Microwave Syst. News*, **13**(5), 67–86, May 1983.
18. R. Q. Lane, "The Determination of Device Noise Parameters," *Proc. IEEE*, **57**, 1461–1462, August 1969.
19. R. Q. Lane, "A Microwave Noise and Gain Parameter Test Set," *IEEE International Solid-State Circuits Conference*, 1978, pp. 172–173.
20. N. Kuhn, "Accurate and Automatic Noise Figure Measurements with Standard Equipment," *IEEE MTT-S International Microwave Symposium Digest*, 1980, pp. 425–427.
21. E. W. Strid, "Measurement of Losses in Noise-Matching Networks," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**, 247–252, March 1981.
22. Chris Bowick, *RF Circuit Design*. Indianapolis: Howard W. Sams & Co., 1982.
23. Adel S. Sedra and Kenneth C. Smith, *Microelectronic Circuits*. Chicago: Saunders College Publishing, 1991.
24. Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed. New York: Wiley, 1984.

25. S. H. Lee, A. K. Wong, and M. G. Wong, *A Current-Combiner Circuit for Better Mixer Conversion Gain*, Philips Semiconductors applications document.
26. F. M. Fano, "Theoretical Limitations in the Broadband Matching Arbitrary Impedances," *J. Franklin Inst.*, **249**, 57–85, 139–155, 1950.
27. Ravender Goyal, ed., *High-Frequency Analog Integrated-Circuit Design*. New York: Wiley, 1995, Chap. 6.
28. Stefan Georgi and Ulrich L. Rohde, "A Novel Approach to Voltage Controlled Tuned Filters Using CAD Validation," *Microwave Eng. Eur.*, 35–42, October 1997.
29. *Analog/Interface ICs Device Data*, Vol. II, Q4/96, DL128, Rev 6, Motorola, Inc., MC13158 Wideband FM IF Subsystem specifications, pp. 8-308 to 8-329.
30. E. J. Wilkinson, "An *N*-Way Power Divider," *IRE Trans. Microwave Theory Tech.*, **MTT-8**, 116, January 1960.
31. Motorola Application Note AN-282A, *Systemizing RF Power Amplifier Design*.
32. W. E. Everitt and G. E. Anner, *Communication Engineering*. New York: McGraw-Hill, 1956.
33. G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks and Coupling Structures*. New York: McGraw-Hill, 1980.
34. G. L. Matthaei, "Tables of Chebyshev Impedance-Transforming Networks of Low-Pass Filter Form," *Proc. IEEE*, August 1964.
35. Motorola Application Note AN-267, *Matching Network Designs with Computer Solutions*.
36. E. G. Cristal, "Tables of Maximally Flat Impedance-Transforming Networks of Low-Pass Filter Form," *IEEE Trans. Microwave Theory Tech.*, **MTT 3**(5), September 1965.
37. H. W. Bode, *Network Analysis and Feedback Amplifier Design*. New York: Van Nostrand, 1943.
38. R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," *J. Franklin Inst.*, 139–154, January-February 1950.
39. J. H. Horwitz, "Design Wideband UHF-Power Amplifiers," *Electron. Des.* **11**, x–x, May 24, 1969.
40. O. Pitzalis and R. A. Gilson, "Tables of Impedance Matching Networks Which Approximate Prescribed Attenuation Versus Frequency Slopes," *IEEE Trans. Microwave Theory Tech.*, **MTT-9**(4), April 1971.
41. Anthony N. Gerkis, "Broadband Impedance Matching Using the 'Real Frequency' Network Synthesis Technique," *Appl. Microwave & Wireless*, **10**(6), 26–36, July/August 1998.
42. T. Cuthbert, "Broadband Impedance Matching Methods," *RF Des.*, **17**(8), 64–71, August 1994.
43. T. Cuthbert, *Broadband Direct-Coupled and Matching RF Networks*. Greenwood, AR, T. R. Cuthbert, 1999.
44. George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990, pp. 279–283.
45. Andy Potter, "HP RF Compiler Automates Schematic Capture and Extends Capabilities of Circuit Synthesis." *Appl. Microwave & Wireless*, **11**(6), 106–118, June 1999.
46. Aron Kain, *Final Report for an Integrated CAD Tool for High-Power, High-Efficiency RF and Microwave Amplifiers*, STTR Phase II, Contract no. F33615-95-C-174, Dept. of the Air Force, Small Business Technology Transfer (STTR) Program.
47. Kenneth K. Clarke and Donald T. Hess, *Communication Circuits: Analysis and Design.* Reading, MA: Addison-Wesley, 1971.

FURTHER READING

P. L. D. Abrie, *The Design of Impedance-Matching Networks for Radio-Frequency and Microwave Amplifiers*. Norwood, MA: Artech House, 1985.

P. L. D. Abrie and P. Rademeyer, "A Method for Evaluating and the Evaluation of the Influence of the Reverse Transfer Gain on the Transducer Power Gain of Some Microwave Transistors," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**(8), 711–713, August 1985.

A. Ambrózy, *Electronic Noise*. New York: McGraw-Hill, 1982.

W. Baechtold and M. J. O. Strutt, "Noise in Microwave Transistors," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**(9), 578–585, September 1968.

A. Cappy, "Noise Modeling and Measurement Techniques," *IEEE Trans. Microwave Theory Tech.*, **MTT-36**, 1–10, January 1988.

R. S. Carson, *High Frequency Amplifiers*. New York: Wiley, 1975.

Cezar A. A. Carioca and Paulo H. P. de Carvalho, "Computer-Aided Design of Diode Frequency Multipliers," *Appl. Microwave & Wireless*, **11**(4), 38–46, April 1999.

H. J. Carlin, "A New Approach to Gain-Bandwidth Problems," *IEEE Trans. Circuits Syst.*, **CAS-23**, 170–175, April 1977.

H. J. Carlin and P. Amstutz, "On Optimum Broad-Band Matching," *IEEE Trans. Circuits Syst.*, **CAS-23**, 401–405, April 1977.

H. J. Carlin and P. Amstutz, "On Optimum Broadband Matching," *IEEE Trans. Circuits Syst.*, **CAS-28**(5), May 1981.

H. J. Carlin and J. J. Komiak, "A New Method of Broad-Band Equalization Applied to Microwave Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**, 93–99, February 1979.

H. J. Carlin and B. Yarman, "A Simplified Real Frequency Technique Applied to Broad-Band Multistage Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-30**(12), December 1982.

H. J. Carlin and B. S. Yarman, "The Double Matching Problem: Analytic and Real-Frequency Solutions," *IEEE Trans. Circuits Syst.*, **CAS-30**(1), 15–28, January 1983.

G. Caruso and M. Sannino, "Computer-Aided Determination of Microwave Two-Port Noise Parameters," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 639–642, September 1978.

Complex Math User's Guide, Compact Software, Paterson, NJ, 1988.

Steve C. Cripps, *RF Power Amplifiers for Wireless Communications*. Norwood, MA: Artech House, 1999.

T. Cuthbert, *Circuit Design Using Personal Computers*. New York: Wiley, 1983.

J. A. Eisenberg, "Designing Amplifiers for Optimum Noise Figure," *Microwaves*, **13**, 36–41, April 1974.

R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," *J. Franklin Inst.*, **249**, 57–83, January 1960/139–155, February 1960.

H. Fukui, *Low-Noise Microwave Transistors and Amplifiers*. New York: IEEE Press, 1981.

C. Gentili, *Microwave Amplifiers and Oscillators*. New York: McGraw-Hill, 1987.

Stefan Georgi and Ulrich L. Rohde, "A Novel Approach to Voltage-Controlled Tuned Filters Using CAD Validation," *Microwave Eng. Eur.*, 35–42, October 1997.

G. Gonzales, *Microwave Transistor Amplifier Analysis and Design*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

H. A. Haus and R. B. Adler, "Optimum Noise Performance of Linear Amplifiers," *Proc. IRE*, **46**, 1517–1533, August 1958.

Aron Kain, *Final Report for Bias Dependence Noise Modeling of Heterojunction Bipolar Transistors*, USAF SBIR Phase II (PIIN), F33615-95-C-1707, November 1997. Issued by USAF/AFMC/ASC, Wright Laboratory WL/AAKE BLD 7, 2530 C ST, Wright-Patterson AFB, OH 45433-7607.

M. J. Köppen, "A Single-Stage Power Amplifier for 17 watts Output in the 470-MHz Band Using the 266BLY, Operating from 12.5 volt Supply Voltage," C.A.B. Communications, Eindhoven, Holland, Electronic Application Laboratory Report, August 8, 1970.

William B. Kuhn and Naveen K. Yanduru, "Spiral Inductor Substrate Loss Modeling in SiliconRFICs," *Microwave J.*, **11**(3), 66–81, March 1999.

J. Lange, "Noise Characterization of Linear Two-Ports of Invariant Parameters," *IEEE J. Solid-State Circuits*, **SC-2**, 37–40, June 1967.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

David Lovelace, Natalino Camilleri, and George Kannell, "Silicon MMIC Inductor Modeling for High Volume, Low Cost Applications," *Microwave J.*, **39**(8), 60–71, August 1994.

Jarek Lucek and Robbin Damen, "Designing an LNA for a CDMA Front End," *RF Des.*, **22**(2), 20–30, February 1999.

Stephen A. Maas, "FET Models for Volterra-Series Analysis," *Microwave J.*, **42**(5), 260–270, May 1999.

Donald A. McClure, KB2Z, "Transmission-Line Transformers," *Commun. Q.*, 45–58, Summer 1997.

D. J. Mellor and J. G. Linvill, "Synthesis of Interstage Networks and Prescribed Gain Versus Frequency Slopes," *IEEE Trans. Microwave Theory Tech.*, **MTT-23**(12), 1013–1020, December 1975.

Sean Mercer, "Exact Simulation of LNAs Reduces Design Cycle Time," *Appl. Microwave & Wireless*, 74–82, January 1999.

R. P. Meys, "A Wave Approach to the Noise Properties of Linear Microwave Devices," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 34–37, January 1978.

C. K. S. Miller, W. C. Daywitt, and M. G. Arthur, "Noise Standards, Measurements and ReceiverNoise Definitions," *Proc. IEEE*, **55**, June 1967.

Brian J. Minnis, *Designing Microwave Circuits by Exact Synthesis* . Boston: Artech House, 1996.

R. Minton, *Design of Large Signal VHF Transistor Power Amplifiers* , RCA Application Note SMA-36, July 1964.

R. Minton and H. C. Lee, *Frequency Multiplication Using Overlay Transistors* , RCA Application Note SMA-40, September 1965.

M. Mitama and H. Katoh, "An Improved Computational Method for Noise Parameter Measurement," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**, 612–615, June 1979.

Motorola tutorial, "The Cascode Amplifier: A Useful RF Building Block," *Microwave Prod. Digest*, 39, 59, March 1999.

T. Mukaihata, "Applications and Analysis of Noise Generation in *N*-Cascaded Mismatched Two-Port Networks," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 699–708, September 1968.

W. W. Mumford and E. H. Scheibe, *Noise Performance Factors in Communications Systems* . Dedham, MA: Horizon House, 1968.

D. E. Norton, "High Dynamic Range Transistor Amplifier Using Lossless Feedback," *Microwave J.*, **19**(5), 53–57, May 1976.

D. E. Norton, "Strong Signal Survivability Design for Lossless Feedback Amplifiers," *Proceedings of the Eleventh Annual European Microwave Conference* , 1981.

H. S. Okean and P. O. Lombardo, "Noise Performance of M/W and MM-Wave Receivers," *Microwave J.*, **16**(1), January 1973.

T. Y. Otoshi, "The Effect of Mismatched Components on Microwave Noise-Temperature Calibrations," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 675–686, September 1968.

P. Penfeld, "Wave Representations of Amplifier Noise," *IRE Trans. Circuit Theory*, 84–86, March 1962.

Ulrich L. Rohde, Jerry C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*, 2nd ed. New York: McGraw-Hill, 1997, Chap. 3.

C. Rorabaugh, *Circuit Design and Analysis—Featuring C Routines* . New York: McGraw-Hill, 1992.

M. Sannino, "On the Determination of Device Noise and Gain Parameters," *Proc. IEEE*, **67**, 1364–1366, September 1979.

SYNMAT, matching synthesis program by Radom, Inc., developed by Douglas Mellor.

A. van der Ziel, *Noise in Solid State Devices and Circuits* . New York: Wiley, 1986.

Dave Vondran, "Noise Figure Measurement: Corrections Related to Match and Gain," *Microwave J.*, **42**(3), 22–38, March 1999.

D. F. Wait, "Thermal Noise from a Passive Linear Multiport," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 687–691, September 1968.

B. S. Yarman and H. J. Carlin, "A Simplified 'Real Frequency' Technique Applied to Broad-Band Multistage Microwave Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-30**(12), 2216–2222, December 1982.

G. P. Young and S. O. Scanlan, "Matching Network Design Studies for Microwave Transistor Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**(10), 1027–1035, October 1981.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

3

AMPLIFIER DESIGN WITH BJTs AND FETs

3-1 PROPERTIES OF AMPLIFIERS

3-1-1 Introduction

The goal of this chapter is to move from the details on the semiconductor devices themselves to our first practical application. One of the more interesting features of amplifiers is that they can easily become what will be covered in Chapter 5: *oscillators*. This is due to the high gain of the devices, unaccounted for parasitic elements, and other design flaws. The amplifiers we need will fall into three categories:

- Low-noise amplifiers
- High-gain amplifiers
- Medium- to high-power amplifiers

See Figure 3-1. The low-noise amplifier always operates in Class A, typically at 15–20% of its maximum useful current. The high-gain amplifier can operate in Class A, as well as Class B (mostly push–pull). The higher dc current for the same device results in a higher noise figure and more gain, and ultimately more output power. Class C operating mode is really reserved for either FM transmissions or constant-carrier modes like CW. Some of the modern digital modulation types are sensitive to phase distortion rather than amplitude changes and because of the resulting output spectrum designers have stayed away from Class C operation.

Some of these amplifiers, such as bipolar versions, can be dc-coupled with very few difficulties; others, like those in the FET families, will cause more headaches. An interesting example, although much too high in frequency for the purpose of this book, is shown in Figures 3-2, 3-3, 3-4, and 3-5. It shows a three-stage amplifier that has noise feedback for the first two stages and resistive feedback for the output stage. These are the three types of amplifiers, regardless of technology, we will be evaluating. The achievement of accurate noise analysis of this circuit, reported in Rohde et al. [1], marked the first time a complete

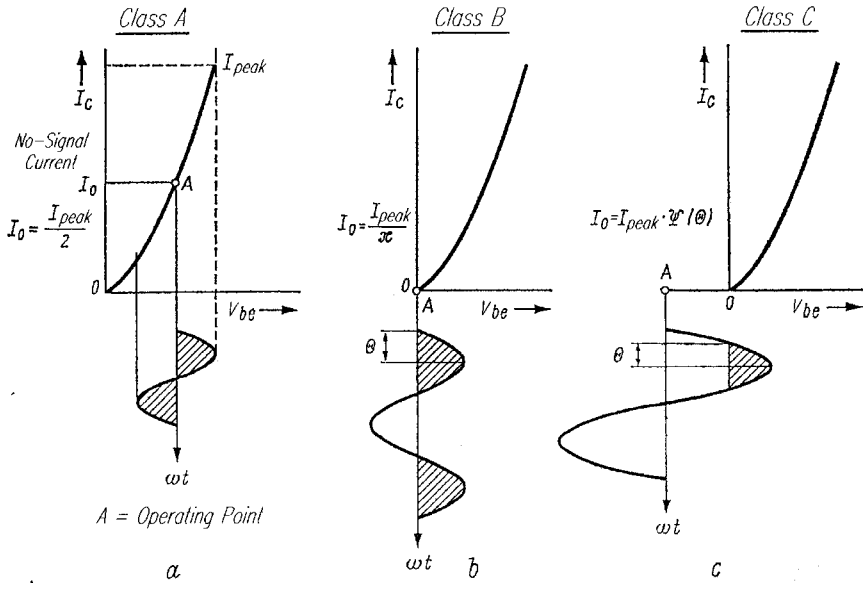


Figure 3-1 Definition of Class A, B, and C operation and resulting bias, including conduction angle (Θ, to be discussed in Section 3-2-2). The transfer characteristic can be either quadratic or exponential. This results in different distortion but does not change the basic operating mode.

linear noise model with essentially no frequency limitation had been developed. Its accuracy depends solely on the accuracy of the element values of the linear equivalent circuit, which can be obtained from measured *S*-parameters by an optimization process. Needless to say, this technique is also applicable at lower frequencies, such as 1.5–3 GHz, and its deviation from measured values has never been worse than 0.2 dB or, relative to the noise figure, 10% expressed in dB.

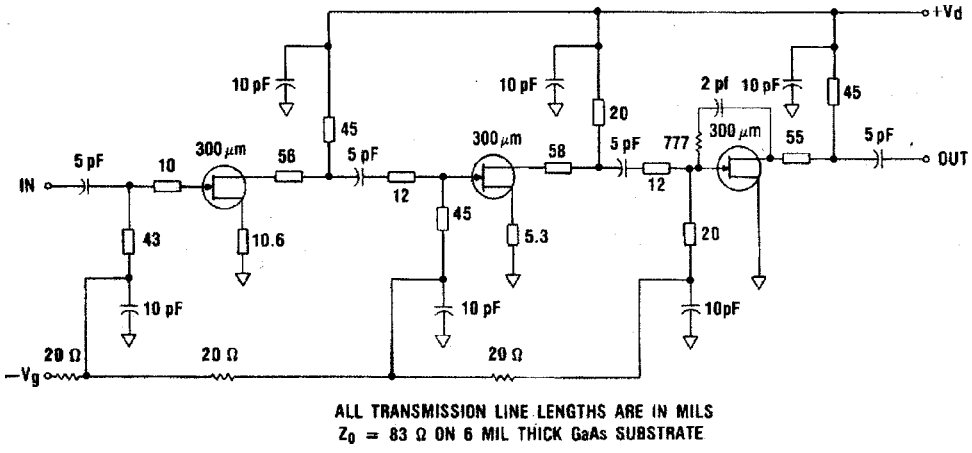


Figure 3-2 Schematic of the X-band GaAs monolithic low-noise amplifier (Texas Instruments EG8021). For reasons of linearity, all three stages operate in Class A.

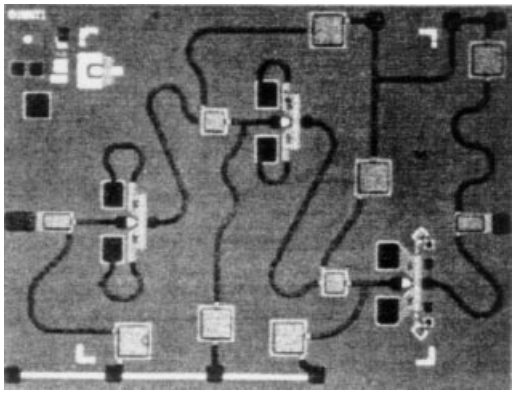


Figure 3-3 Photograph of the EG8021 monolithic amplifier chip. The area pictured is 0.09 inch by 0.12 inch in size.

The results of Figure 3-4 were obtained by using the linear FET model as supplied by Texas Instruments. It becomes somewhat obvious that a match for S_{11} of –60 dB is not likely. We then replaced the linear transistor with the SPICE-type nonlinear model, resulting in the responses shown in Figure 3-5. S_{11} now looks much more realistic in the center of the range, but at the same time exhibits a tendency toward instability at frequencies around 8 GHz. The actual circuit, when measured, did not exhibit this potential instability; on the other hand,

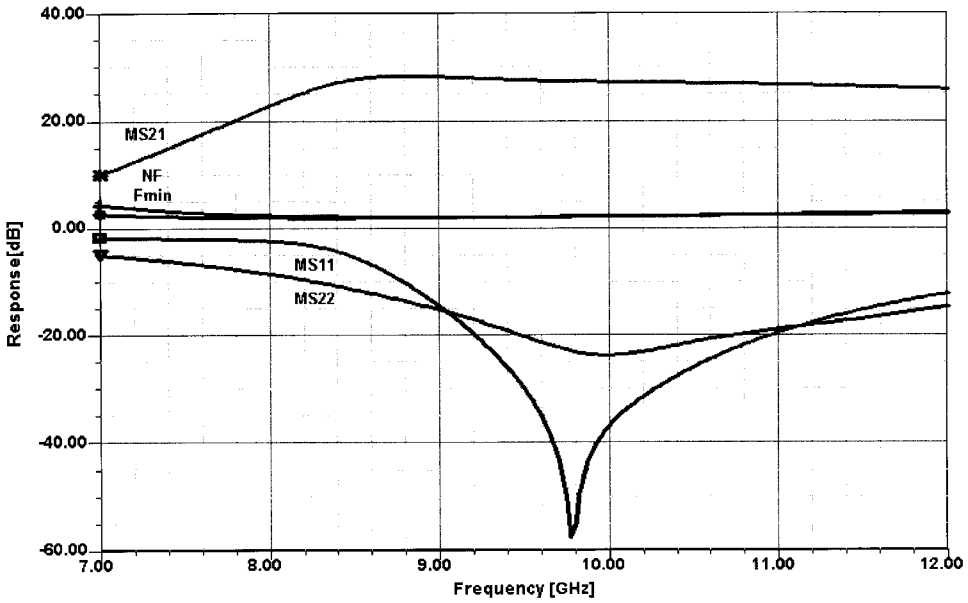


Figure 3-4 Simulated F_{min}, noise figure, S_{11}, S_{21}, and S_{22} responses for the three-stage GaAsFET amplifier using the TI linear FET model. The values at 10 GHz are: F_{min}, 2.20 dB; NF, 2.21 dB; S_{11}, –36.8 dB; S_{21}, 27.2 dB; and S_{22}, –23.7 dB.

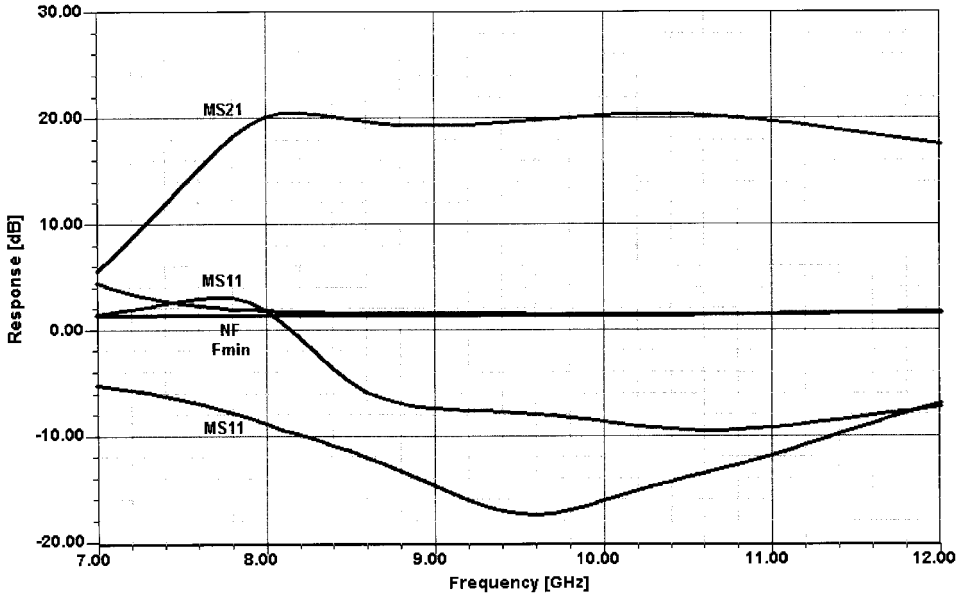


Figure 3-5 Simulated F_{min}, noise figure, S_{11}, S_{21}, and S_{22} responses for the three-stage GaAsFET amplifier using the nonlinear FET model. The values at 10 GHz are: F_{min}, 1.53 dB; NF, 1.65 dB; S_{11}, –8.5 dB; S_{21}, 20.3 dB; and S_{22}, –15.9 dB.

the noise figure and gain agreed quite well. The discrepancies between the results with the linear and nonlinear models emphasize the importance of accurate device modeling.

Figure 3-6 shows a wideband amplifier, specifically a distributed amplifier, that covers 1–20 GHz. Figure 3-7 shows its simulated frequency-dependent gain, matching, and noise performance. The example given here as the upper frequency for wireless systems is rapidly expanding and the technology to accommodate this is already in place. Modern CAD tools can handle such high frequencies accurately, which means one can trust them at the significantly lower wireless frequencies of 500 MHz to 3 GHz. Above 1500 MHz, one really should model using distributed elements. We will come to this later in this chapter.

Input/output and interstage matching will become another challenging task, specifically if one requires sufficient bandwidth. As circuits become more complex and as the operating frequency increases, passive components play an increasingly important role; the use of

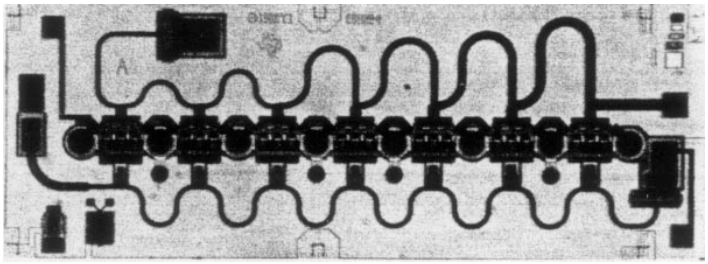


Figure 3-6 Photo of the 1–20-GHz distributed amplifier.

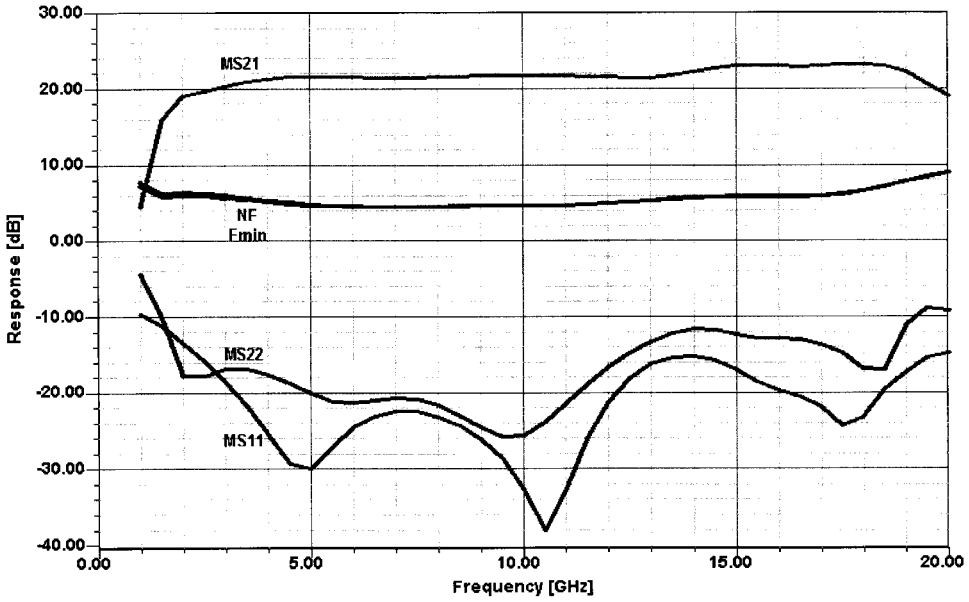


Figure 3-7 Simulated F_{min}, NF, S_{11}, S_{21}, and S_{22} responses. At the center of the range, the following numbers were found: F_{min}, 4.80 dB; NF, 4.80 dB; S_{11}, –32.8 dB; S_{21}, 21.8 dB; and S_{11}, –25.2 dB.

distributed elements is practically a necessity above 1500 MHz. For this reason, we begin this chapter by covering the application of transistors as amplifiers, following this with a discussion of the surrounding external circuitry, such as transmission lines, inductors, capacitors, and resistors. Monolithic ICs on silicon are a totally different ball game, and neither of the authors has sufficient expertise in this area. For this reason, we refer readers who will concentrate on MMIC design to the literature mentioned at the end of this chapter.

The goal of the applications presented here is to address, where possible, high-performance wireless stages that are less subject to constraints of space, power consumption, and cost, than portable/mobile applications. This is consistent with our preface, in which we have clearly stated our goal. The designers of handsets typically cannot consider the full variety of circuits we will present here. These types of stages (low noise, high gain and power) frequently will be used with filtering, and we will give some examples of input filters, including a tracking high-performance design.

A key topic besides the frequently mentioned gain is the whole noise issue. We will give the reader a complete introduction to this important area, followed by a similar introduction to the gain and matching of transistors. *Caution*: The majority of applications published go through the exercise of using *S*-parameters manually to design low-noise and small-signal amplifiers. While we have been tempted to repeat the information found in several textbooks, we feel that today's engineers will have access to CAD tools one way or another. It is highly unlikely that readers will undertake the inefficient process of going through all the design steps manually (a procedure that can take days); hopefully, engineers will resort to a linear/nonlinear CAD tool to do so. Only for the purpose of summarizing the necessary steps will we provide the relevant information regarding gain, stability, and related topics, demonstrating their usefulness or possible misuse with a modern CAD tool—in our case,

the Serenade Design Environment product from Ansoft. At this time, several additional companies beside Ansoft and Hewlett-Packard have emerged, but it is dangerous to list all of them as they come and go.

3-1-2 Gain

In practical terms, when we talk about gain we mean the output power relative to the input power. The industry standard now uses 50 Ω; however, at the point at which we have to match high-impedance devices such as monolithic crystal filters where the impedance jumps, things are not so obvious. The transformation equation that allows us to transform from one impedance to another is

$$m = \sqrt{\frac{R_2}{R_1}} \tag{3-1}$$

This correction factor, m, is required to do the transformation from R_1 (50 Ω) to R_2 (e.g., 1.2 kΩ). It needs to be understood that while the power gain remains the same, the voltage gain requires this correction factor. Since the power gain is expressed as 10 log A, A being the loop gain, 20 log A is the voltage gain, both expressed in dB. The reason for 10 versus 20 is that the power gain is always the voltage squared, which accounts for the 2. As a reminder of these relationships we present the following:

Distinction: Power Ratios–Voltage Ratios

$$n \text{ dB} = 10 \log_{10}(P_2/P_1) \quad \xrightarrow{PP = V^2/Z} \quad n = 10 \log_{10}\left(\frac{V_2^2/Z_2}{V_1^2/Z_1}\right)\text{dB}$$

$$Z_2 = Z_1 \rightarrow n \text{ dB} = 10 \log_{10}(V_2^2/V_1^2) = 10 \log_{10}(V_2/V_1)^2$$

$$\Longrightarrow \boldsymbol{n \text{ dB} = 20 \log_{10}(V_2/V_1) \text{ dB}}$$

The gain is typically defined under the condition of real resistive source and load terminations, which means that all the available power is being supplied to the amplifier, and likewise through proper matching at the output all the available power is fed to the termination. This gain figure is typically measured at a single spot frequency, but since most amplifiers are not totally flat, one needs to specify the gain tolerances or gain flatness. A good number is x dB ± 0.5 dB gain variation. In the case of a narrowband amplifier, the gain typically is shaped like a tuned circuit or like a critically coupled bandpass filter frequency response, while in the case of a real wideband application, the desired gain has some kind of ripple, which is due to the various compensation components in the circuit. An extreme case of a wideband amplifier is a 6–18-GHz single-stage amplifier that uses some very clever matching. This is shown in the following example. Figure 3-8 shows the circuit diagram, and Figure 3-9 shows its frequency-dependent performance.

While this is a cute approach for generating a single-stage amplifier over such a wide bandwidth, in real life it is not acceptable. The reason for this is that neither S_{11} nor S_{22} meets any useful specifications over the entire range. Typically, S_{11}/S_{22} values are only a few dB, while in reality more than 10 dB is required. We have chosen this example to show that gain

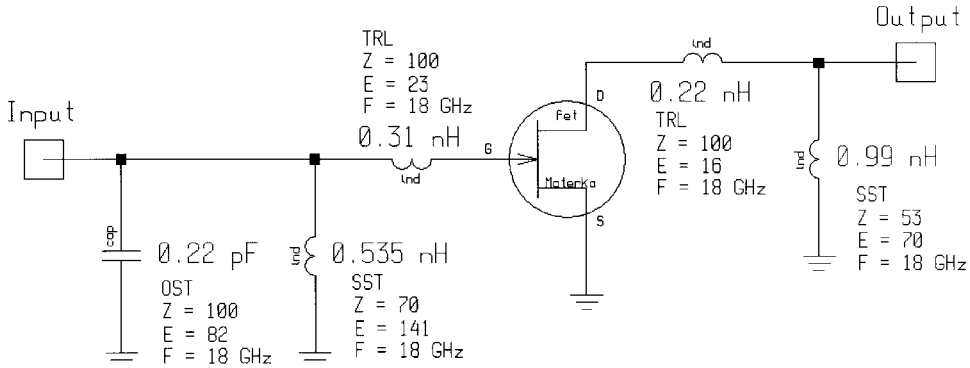


Figure 3-8 A 6–18-GHz FET amplifier.

and matching, depending on the circuit topology, are connected. We get similar results for the noise figure, whereby the best achievable noise figure ($F_{\min}$) and the actual 50-Ω noise figure are only a few dB apart; however, its frequency dependence in a wideband stage is certainly not acceptable.

Another parameter associated with gain is the stability factor K of the circuit. A value of $K > 1$ is mandatory to guarantee stability together with the S_{11} and S_{22} magnitudes being less than 1. Figure 3-10 plots the K stability factor, which for absolute stability must be >1, and maximum stable gain (MSG) as well as $G_{\max}$. It turns out that MSG is a highly artificial

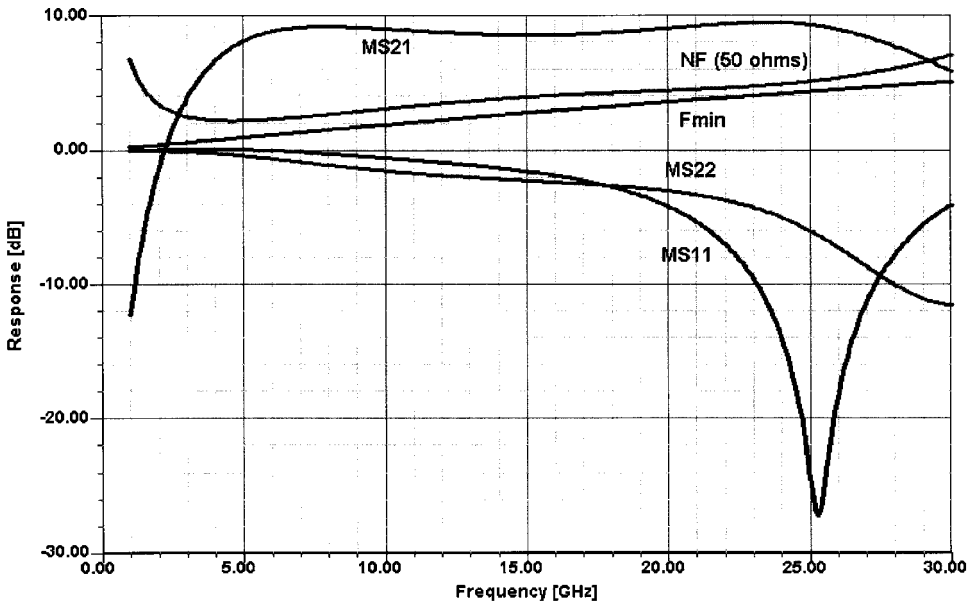


Figure 3-9 Frequency-dependent gain, matching, and noise performance of the amplifier shown in Figure 3-8.

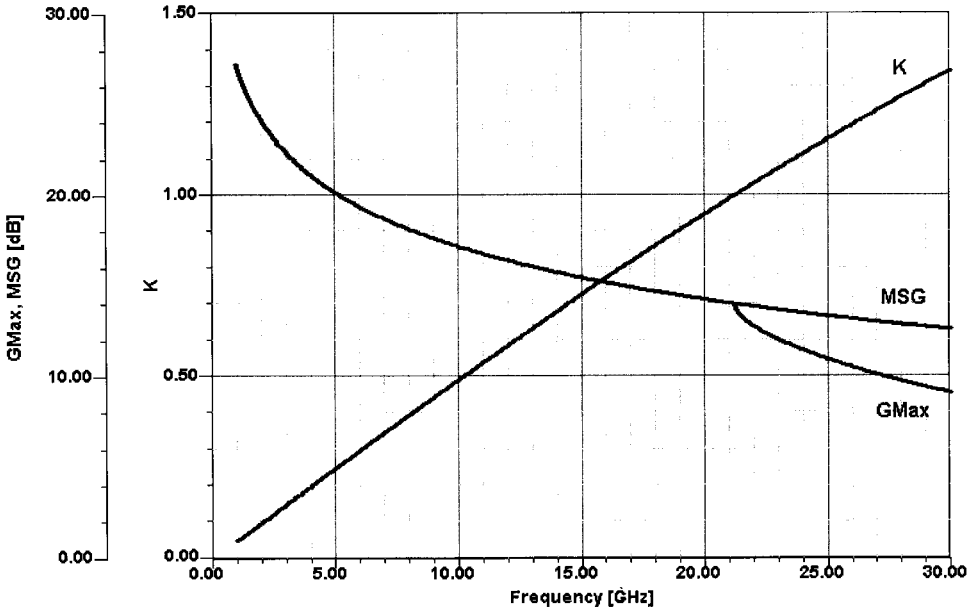


Figure 3-10 G_{max}, maximum stable gain (MSG), and K performance of the amplifier in Figure 3-8.

definition of stability that implies that input and output are terminated with pure resistances and then conjugately matched. There are four ways of achieving this resistive loading, either in the emitter as a shunt resistor between gate/drain or base/collector, or in parallel to the input and output to ground. Needless to say, this is not a real operating mode because it deteriorates both noise and gain. In our case, the entity G_{max} is a function of S_{12} of the circuit, not just the transistor, and depends highly on S_{12}. If we bring the system close to oscillation, then G_{max} will increase at the expense of bandwidth. Unfortunately, some of these microwave definitions of gain (see Table 3-1) have become widely used in industry without always being useful. Evaluating these equations for S_{12} is not legitimate because other quantities like K become infinite, and the entire system of equations falls apart. The power gain with input and output conjugately matched, or the transducer power gain for arbitrary source and load matching, still give more insight into the system.

In the above example, we specifically stated that the input and output VSWR was totally unacceptable, and this was due to a one-stage design. Figure 3-11 shows an amplifier chip built around two transistors with standard feedback techniques applied; Figures 3-12 and 3-13 show its frequency-dependent gain, matching, stability, and noise performance. This amplifier was one of the early examples of GaAsFET wideband amplifiers, but it is interesting because while it has wild (and wide) gain swings, it exhibits a fairly constant input and output impedance or reflection coefficient. A further evaluation of the circuit shows that not only is the circuit unconditionally stable, but that the maximum stable gain has an extremely high value based on essentially minimal S_{12}, the K factor is always significantly higher than 1, and G_{max} follows, to some degree, the actual gain response of the amplifier. (Ultimately, designers gave up on simple *RCL* feedback amplifiers for such a wide bandwidth, replacing them with distributed amplifiers previously called traveling wave amplifiers.) The problem of gain and stability starts at audiofrequency amplifiers, where the

Table 3-1 Nine power gains

Transducer power gain in 50-Ω system	$G_T = \lvert S_{21}\rvert^2$
Transducer power gain for arbitrary Γ_G and Γ_L	$G_T = \dfrac{(1-\lvert\Gamma_G\rvert^2)\,\lvert S_{21}\rvert^2\,(1-\lvert\Gamma_L\rvert^2)}{\lvert(1-S_{11}\Gamma_G)\,(1-S_{22}\Gamma_L)-S_{12}S_{21}\Gamma_G\Gamma_L\rvert^2}$
Unilateral transducer power gain	$G_{\mathrm{TU}} = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_G\rvert^2)\,(1-\lvert\Gamma_L\rvert^2)}{\lvert 1-S_{11}\Gamma_G\rvert^2\lvert 1-S_{22}\Gamma_L\rvert^2}$
Power gain with input conjugate matched	$G = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_L\rvert^2)}{\lvert 1-S_{22}\Gamma_L\rvert^2(1-\lvert S'_{11}\rvert^2)} = \dfrac{\lvert S_{21}\rvert^2}{1-\lvert S_{11}\rvert^2}$ (for $\Gamma_L = 0$)
Available power gain with output conjugate matched	$G_A = \dfrac{\lvert S_{21}\rvert^2(1-\lvert\Gamma_G\rvert^2)}{\lvert 1-S_{11}\Gamma_G\rvert^2(1-\lvert S'_{22}\rvert^2)} = \dfrac{\lvert S_{21}\rvert^2}{1-\lvert S_{22}\rvert^2}$ (for $\Gamma_G = 0$)
Maximum available power gain[a]	$G_{ma} = \left\lvert\dfrac{S_{21}}{S_{12}}\right\rvert\,(k-\sqrt{k^2-1})$
Maximum unilateral transducer power gain	$G_{\mathrm{TUmax}} = \dfrac{\lvert S_{21}\rvert^2}{(1-\lvert S_{11}\rvert^2)(1-\lvert S_{22}\rvert^2)}$
Maximum stable power gain	$G_{ms} = \dfrac{\lvert S_{21}\rvert}{\lvert S_{12}\rvert}$
Unilateral power gain	$U = \dfrac{1/2\lvert S_{21}/S_{12}-1\rvert^2}{k\lvert S_{21}/S_{12}\rvert-\mathrm{Re}(S_{21}/S_{12})}$

[a]Definition for $G_{\max}$: If $k > 1$, it is called G_{ma}, and if $k < 1$, it is called G_{ms}.

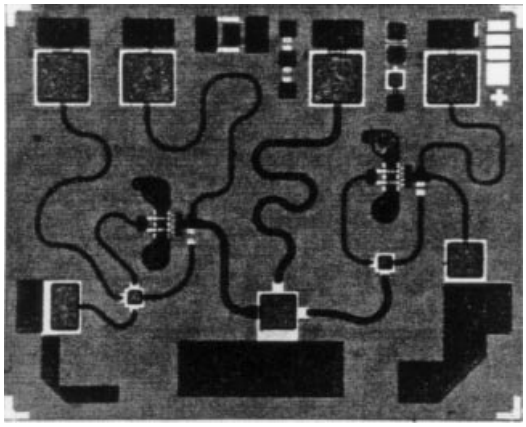


Figure 3-11 A two-stage, wideband feedback amplifier.

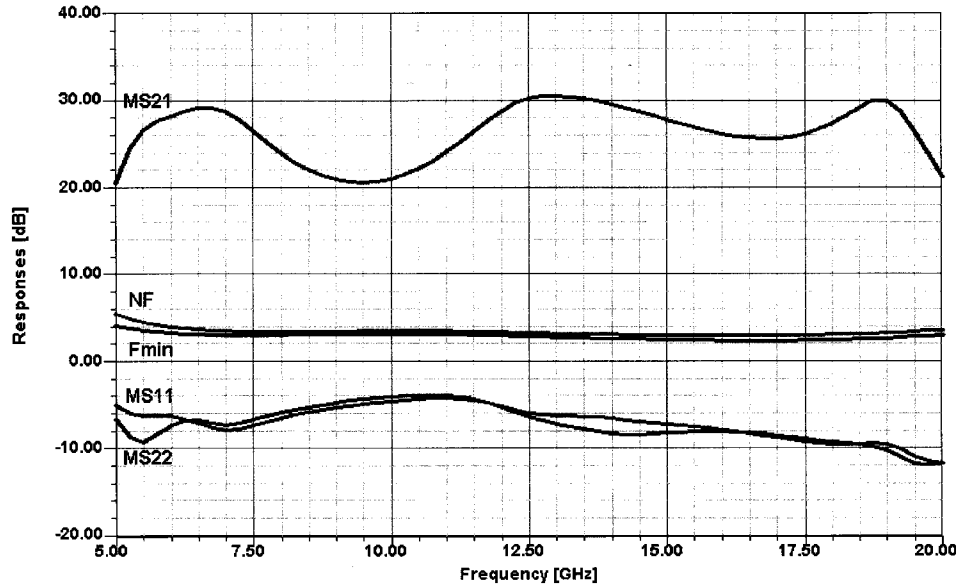


Figure 3-12 Frequency-dependent gain, matching, and noise performance for the amplifier in Figure 3-11.

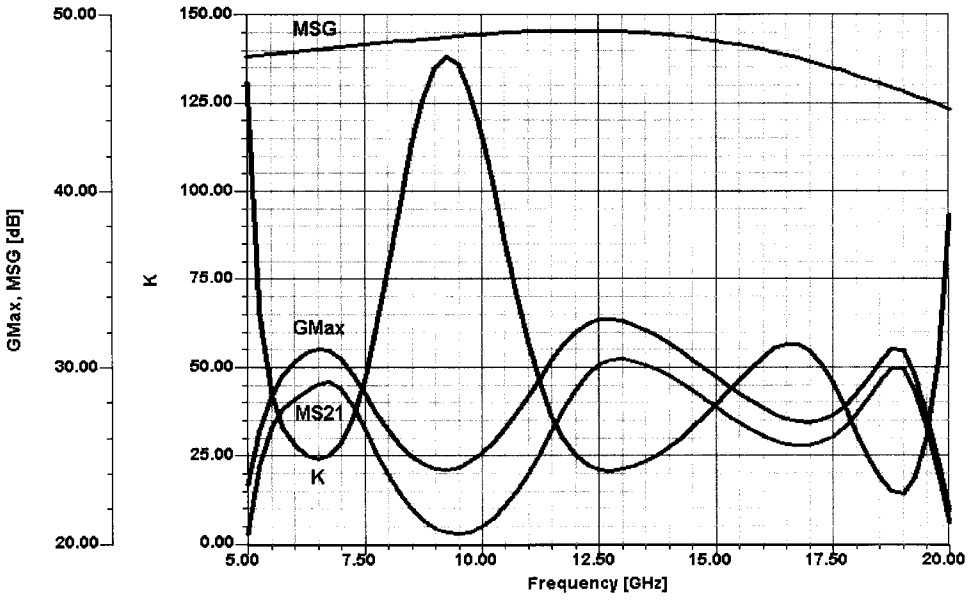


Figure 3-13 Frequency-dependent G_{max}, K, maximum stable gain, and gain performance for the amplifier in Figure 3-11.

motorboating phenomenon has been known for years, and the problem moves up in frequency with more modern technology and is largely dependent not only on the device and the circuit, but also on the layout. The motorboating effect has to do with ground loops and the fact that if the printed circuit traces from the power supply connection to the output power stage present too high a resistance, then at high currents the voltage will drop to a point below useful operation and then, as the volume is lowered, become stable again. In this case the amplitude of the signal through the amplifier controls the stability of the output stage. Needless to say, this most unpleasant effect applies to high-power stages for all output transistors, where we use Class B or C operation, meaning that the dc operating current as a function of output does not remain constant.

In terms of gain, there is also an important gain characteristic called *differential gain* and *differential phase*. These expressions were first used in television circuits, where linearity in both areas is an absolute must. The amplitude of the signal determines the black and white contrast, and the phase determines the color. As an example, transmitting an image of one of our presidents playing golf in California to the New York area is done by many microwave hops. Any multipath reception, such as that caused by reflection from airplanes, may have little effect on the signal amplitude, but the resulting phase shift can transform a green lawn to a red lawn if a total shift of 180° occurs. Given a particular operating point, if the 1-dB compression point of multiple amplifiers in the chain is approached or exceeded, their differential gain, which is defined by the IEEE as "the difference between (a) the ratio of the output of a small high-frequency sine-wave signal at two different levels of a low-frequency signal on which it is superimposed, and (b) unity," may increase. This mode of saturation charges all the capacitances of the amplifier transistors, causing huge phase shifts that result in the known problem of changing colors. The same applies, of course, in test equipment amplifiers, in which such changes cannot be tolerated. A somewhat similar, but not quite the same, problem is spectral regrowth [evaluated in terms of adjacent-channel power ratio (ACPR)], which is also due to nonlinearities. The question of linearity has already been covered in detail in Chapter 1. Needless to say, any overdrive condition also changes an amplifier's input and output SWR.

3-1-3 Noise Figure (NF)*

Introduction. In this section we look at the best way to describe noise in active devices (black boxes). Even when a two-port is linear, the output waveform will differ from the input, because of the failure to transmit all spectral components with equal gain (or attenuation) and delay. By careful design of the two-port, or by limitation of the bandwidth of the input waveform, such distortions can largely be avoided. However, noise generated within the two-port can still change the waveform of the output signal. In a linear passive two-port, noise arises only from the losses in the two-port; thermodynamic considerations indicate that such losses result in the random changes that we call noise. When the two-port contains active devices, such as transistors, there are other noise mechanisms present. A very important consideration in a system is the amount of noise that it adds to the transmitted signal. This is often judged by the ratio of the output signal power to the output noise power

*Section 3-1-3 is reprinted from Chapter 2 of George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*, copyright 1990 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

(S/N). The ratio of signal plus noise power to noise power [$(S + N)/N$] is generally easier to measure and approaches S/N when the signal is large.

In the evaluation of a two-port it is important to know the amount of noise added to a signal passing through it. An important parameter for expressing this characteristic is the noise factor. The signal energy coming from a generator or antenna is amplified or attenuated in passing from the input to the output of a two-port, as is the noise that accompanies the input signal energy. A system generally includes a cascade of two-port networks that constitute one overall two-port that amplifies the signal to a high-enough power level for its intended use. The *noise factor* of a system is defined as the ratio of signal-to-noise ratios available at input and output:

$$F = \frac{(S/N)_{\text{input}}}{(S/N)_{\text{output}}} \geq 1 \tag{3-2}$$

The noise figure (or factor) of a receiver is an easily measured quantity that describes the signal-to-noise ratio reduction of that receiver.

When this ratio of powers is converted to decibels, it is generally referred to as the *noise figure* rather than noise factor. Various conventions are used to distinguish the symbols used for noise factor and noise figure. Here we use F to represent the noise factor and NF to represent the noise figure, although the terms are usually used interchangeably.

For an amplifier with the power gain G, the noise factor can be rearranged as

$$F = \frac{S_i/N_i}{GS_i/G(N_i + N_a)} \tag{3-3}$$

when N_a is the additional noise power added by the amplifier referred to the input. This can be computed to be

$$F = 1 + N_a/N_i \tag{3-4}$$

The noise factor is often replaced by the noise figure (NF), which is defined in decibels as

$$\text{NF} = 10 \log_{10} F \tag{3-5}$$

In applications like satellite receivers, the noise factor becomes such a small number that it is inconvenient to work with. Many people have adopted the use of an equivalent noise temperature for a circuit to remedy this situation. Since the thermal noise power available from a resistor at temperature T_e is

$$N = kT_eB \tag{3-6}$$

where k is Boltzmann's constant (1.38×10^{-23} J/K), T_e is the effective temperature in kelvins, and B is the bandwidth in hertz. The equation above may be used to associate an effective noise temperature with circuits containing more than just thermal noise sources. This allows Eq. (3-4) to be written as

$$F = 1 + \frac{kT_eB}{kT_0B} = 1 + \frac{T_e}{T_0} \tag{3-7}$$

where T_e is the effective noise temperature of the circuit and T_0 is the temperature of the generator resistor in kelvins. The noise temperature T_e now characterizes our circuit noise contribution and can directly be related to the noise factor.

Assuming a reference noise temperature of 290 K (–273 + 290 = 17 °C), let us determine the noise temperature of the system with a noise factor of 2.6 (4.15 dB):

$$T_e = (2.6 - 1)\ (290) = 464 \text{ K}$$

This temperature T_e should not be confused with the environmental operating temperature T_0. It is quite common to operate low-noise amplifiers with T_e below 100 K at an ambient temperature of 290 K.

Signal-to-Noise Ratio. Let us consider the signal-to-noise ratio of power delivered from a generator to a load as shown in Figure 3-14. The signal power delivered to the input is given by

$$S_{\text{in}} = P_{\text{in}} = \frac{E_g^2 \text{Re}(Z_{\text{in}})}{|Z_g + Z_{\text{in}}|^2} \tag{3-8}$$

where E_g is the rms voltage of the input signal supplied to the system, and the noise power supplied to the input is expressed by

$$N_{\text{in}} = \frac{\overline{v_n^2} \text{Re}(Z_{\text{in}})}{|Z_g + Z_{\text{in}}|^2} \tag{3-9}$$

where the noise power at the input is provided by the noise energy of the real part of Z_g. The input impedance Z of the system in the form $Z = R_{\text{in}} + jX_{\text{in}}$ is assumed to be complex.

The Johnson noise of a resistor [here $\text{Re}(Z_g)$], given by the mean-square voltage

$$\overline{v_n^2} = 4kTRB \tag{3-10}$$

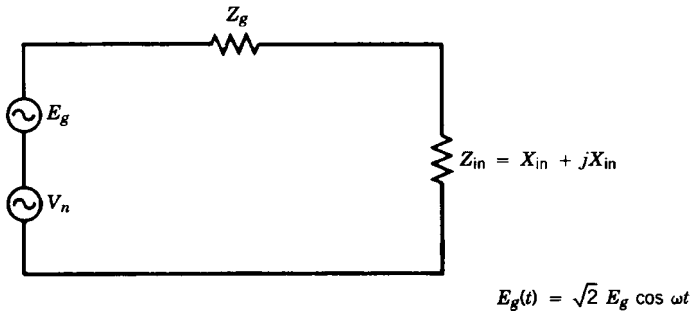


Figure 3-14 Combination of signal and noise voltages supplied to a complex termination.

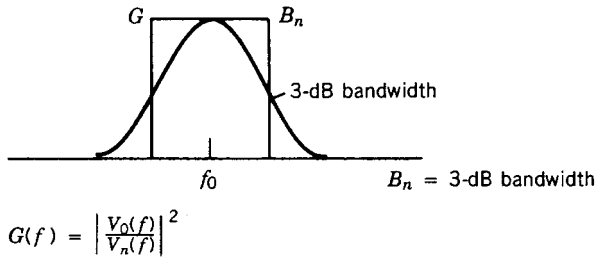
where k (Boltzmann's constant) = 1.38×10^{-23} J/K, T is the absolute temperature of the resistor, and B is the bandwidth, is sufficiently small that the resistive component of impedance does not change. The available signal power from the generator has a lower limit, even if the signal is attenuated by the highest possible attenuation. The generator resistor acts as a Johnson noise generator, its power being

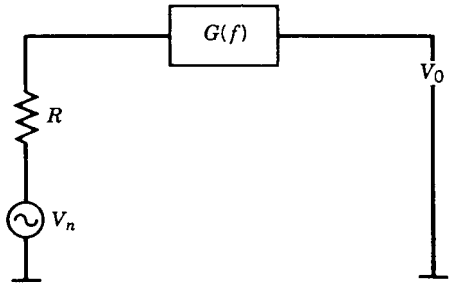
$$P_A = \frac{4kTRB}{4R} = kTB \tag{3-11}$$

where k is Boltzmann's constant, T is the absolute temperature, and B is the bandwidth. This power is the maximum available output power.

For an ambient temperature of 290 K, $kT = 4 \times 10^{-21}$ W/Hz. This expression is also given as kT = –204 dBW/Hz = –174 dBm/Hz = –114 dBm/MHz. We can combine Eqs. (3-8) to (3-10) to obtain

$$\left(\frac{S}{N}\right)_{\text{in}} = \frac{E_g^2}{4kT\text{Re}(Z_g)B} \tag{3-12}$$





$$V_0^2 = \int_0^\infty 4kT_0RG(f)\,dF$$
$$= 4kT_0R \int_0^\infty G(f)\,dF$$
$$B_n = \frac{1}{G} \int_0^\infty G(f)\,dF$$
B_n = noise bandwidth

Figure 3-15 Graphical and mathematical explanations of the noise bandwidth from a comparison of the Gaussian-shaped bandwidth to the rectangular filter response.

This is the value of *S/N* contributed by the generator, which does not include the noise generated by the load, in this case $\mathrm{Re}(Z_{\mathrm{in}})$, which would need to be included in the measurement of the total *S/N* across the input impedance.

A critical parameter is the noise bandwidth, B_n, which is defined as the equivalent bandwidth, as shown in Figure 3-15. For reasons of group delay correction, most practical filters have round rather than sharp corners. The noise figure measurements shown later can be used to determine the "integrated" bandwidth, which is B_n.

An active system such as a combination of amplifiers and mixers will add noise to the input signals, and the noise factor that describes this is defined as the *S/N* ratio at the input to the *S/N* ratio at the output, which is always greater than unity [2]. In practice, a certain minimum signal-to-noise ratio is required for operation. For example, in a communication system such a minimum is required for intelligible transmission, either voice or data. For high-performance TV reception, to provide a noise-free picture to the eye, a typical requirement is for a 60-dB *S/N*. In the case of a TV system, a large dynamic range is required, as well as a very large bandwidth to reproduce all colors truthfully and all shades from high-intensity white to black. Good systems will have a bandwidth of 8 MHz or more.

Noise Figure Measurements. Some of the noise equations are based on mathematical models and physics. To understand some of these expressions, it is useful to look at a practical system, with amplifiers, that has to be evaluated.

Let us look at Figure 3-16, which consists of a signal generator, the system or device under test (DUT), and a selective receiver with a built-in root-mean-square (rms) voltmeter to determine the signal and noise voltages. It is necessary that the system have enough gain so that the noise voltage supplied by the generator will be indicated [3].

If we assume that our selective receiver is a video noise meter calibrated in rms voltage levels, we can perform two measurements. With an input termination connected to the TV system (typically, 75 Ω for cable TV, 50 Ω for satellite TV), the noise receiver/meter will read a value for proper termination that can easily be calculated. Since one-half of the mean-square noise voltage appears across the input,

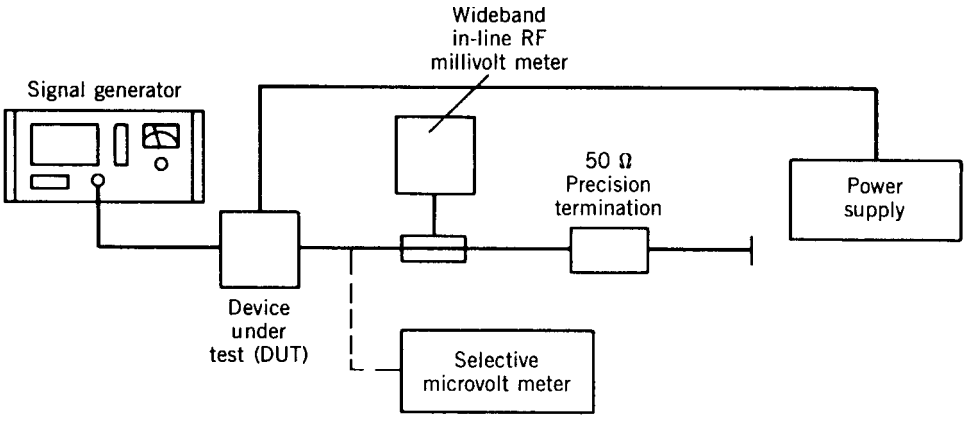


Figure 3-16 Test setup to measure signal-to-noise ratio.

$$v_{\text{in}} = \frac{v_n}{2} = \frac{\sqrt{4kTRB}}{2} \tag{3-13}$$

With $B = 10$ MHz, $T = 290$ K (T is always expressed in absolute temperature; $T_0 = -273$ °C), and $k = 1.38 \times 10^{-23}$ J/K, then for $R = 75\ \Omega$,

$$v_{\text{in}} = \frac{v_n}{2} = 1.73\ \mu\text{V} \tag{3-14}$$

where the rms noise voltage has been referred to the input port. We can verify this with our first measurement.

Now we increase the input voltage of the signal generator to a value that indicates a 60-dB *S/N* ratio at the output port. This should be about

$$E_g = \frac{v_n}{2}\sqrt{F} \times 1000 = 1.73\sqrt{F} \quad \text{mV}$$

where F is the noise factor of the receiver. For a receiver noise factor of 10, we would obtain $E_g = 5.48$ mV (rms value). If the noise energy equivalent to a noise factor of F is assumed, we need $\sqrt{F}$ times more voltage. For a 60-dB ratio, this means that $E_g = 1000 \times (v_n/2) \times \sqrt{F}$.

As they are done here, over a power range of 60 dB, the measurements can be performed over such a wide range only if special equipment is available. In cases where the internal detector of a piece of communication equipment is used, the signal-to-noise ratio measurements are performed over much smaller power ranges.

Let us assume that for the above-mentioned case ($F = 10$) we find a *S/N* ratio of 10 dB at the output for an input signal of 5.47 μV. By rewriting Eq. (3-7) as

$$E_g = \frac{v_n}{2}\sqrt{F} = \sqrt{KTRBF} \tag{3-15}$$

with F being the noise factor, we can solve for F with

$$F = \frac{P_s}{P_n} = \frac{E_g^2/R}{kTB} \tag{3-16}$$

While the input power from the thermal energy of the input termination resistor was $kTB = 4 \times 10^{-4}$ W, the input power required for the 10-dB *S/N* ratio was

$$P_s = \frac{(5.47 \times 10^{-6})^2}{75} = 3.98 \times 10^{-13}\ \text{W} \tag{3-17}$$

The noise factor is defined as the ratio P_s/P_n:

$$F = \frac{3.98 \times 10^{-13}}{4 \times 10^{-14}} = 10 \tag{3-18}$$

which is the proof.

This method is used more frequently at the 3-dB point, or double the input power if the dynamic range of the detector is small or only a linear indicator is available. Because of hum and other pickup, this is not an easy measurement. Using a signal generator is very expensive because in a laboratory or production environment a wide frequency range requires several generators.

Another method is the use of a wideband noise generator. Modern gas discharge diodes or avalanche diodes are available that provide essentially white noise energy over a large frequency range. These microwave diodes typically have an output of 30 dB above kT when switched on and kT when switched off. To provide good matching at microwave frequencies, a 15-dB attenuator is cascaded. This means that the noise power of the source in the ON condition is about 15 dB above kT.

In the early 1960s, low-cost noise figure test equipment was built around vacuum diodes whose operating range was limited to 1200 MHz due to the resonance effects of the structure. Today, automatic noise gain analyzers use calibrated solid-state noise sources up to 26.5 GHz. It appears that the upper frequency limit has to do with matching, and the lower-frequency limit with $1/f$ noise.

Noisy Two-Port Description. Based on the convention by Rothe and Dalke [4], any linear two-port can be in the form shown in Figure 3-17. This general case of a noisy two-port can be redrawn showing noise sources at the input and at the output. Figure 3-17b shows this in admittance form and Figure 3-17c in impedance form. The internal noise sources are assumed to produce very small currents and voltages, and we assume that linear two-port equations are valid. The internal noise contributions have been expressed by using external noise sources:

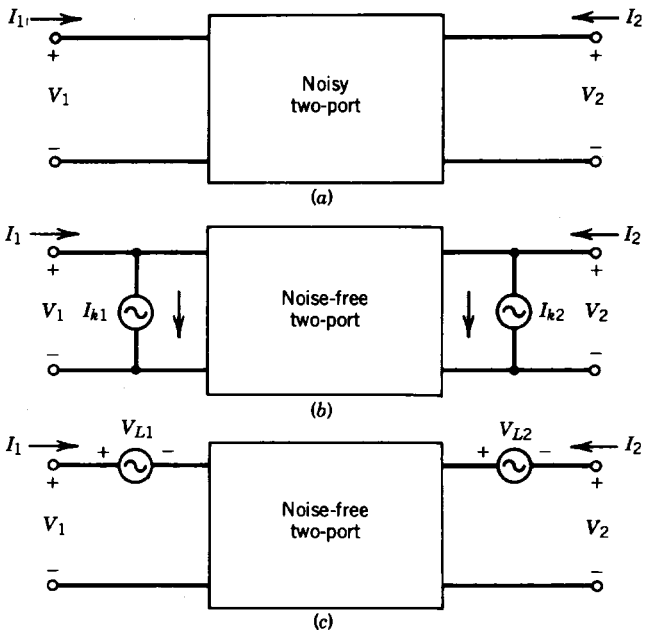


Figure 3-17 Noisy linear two-ports: (a) general form, (b) admittance form, and (c) impedance form.

$$I_1 = y_{11}V_1 + y_{12}V_2 + I_{K1}$$
$$I_2 = y_{21}V_1 + y_{22}V_2 + I_{K2} \tag{3-19}$$

$$V_1 = z_{11}I_1 + z_{12}I_2 + V_{L1}$$
$$V_2 = z_{21}I_1 + z_{22}I_2 + V_{L2} \tag{3-20}$$

where the external noise sources are I_{K1}, I_{K2}, V_{L1}, and V_{L2}.

Since we want to describe our noisy circuit in terms of the noise figure, the *ABCD*-matrix description will be more convenient since it refers both noise sources to the input of the two-port [5]. This representation is given below (note the change in direction of I_2):

$$V_1 = AV_2 + BI_2 + V_A$$
$$I_1 = CV_2 + DI_2 + I_A \tag{3-21}$$

where V_A and I_A are the external noise sources.

It is important to remember that all of these matrix representations are interrelated. For example, the noise sources for the *ABCD*-matrix description can be obtained from the *z*-matrix representation shown in Eq. (3-20). This transformation is

$$V_A = -\frac{I_{K2}}{y_{21}} = V_{L1} - \frac{V_{L2}z_{11}}{z_{21}} \tag{3-22}$$

$$I_A = I_{K1} - \frac{I_{K2}y_{11}}{y_{21}} = -\frac{V_{L2}}{z_{21}} \tag{3-23}$$

The *ABCD* representation is particularly useful because it allows us to define a noise temperature for the two-port referenced to the input. The two-port itself (shown in Figure 3-18) is assumed to be noise-free.

In the past, *z* and *y* parameters have been used, but in microwave applications it has become common to use *S*-parameter definitions. This is shown in Figure 3-19. The previous equations can be rewritten in their new form using *S*-parameters:

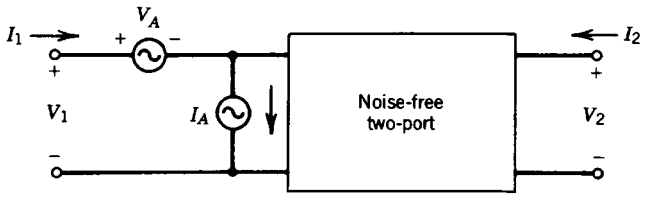


Figure 3-18 Chain-matrix form of linear noisy two-ports.

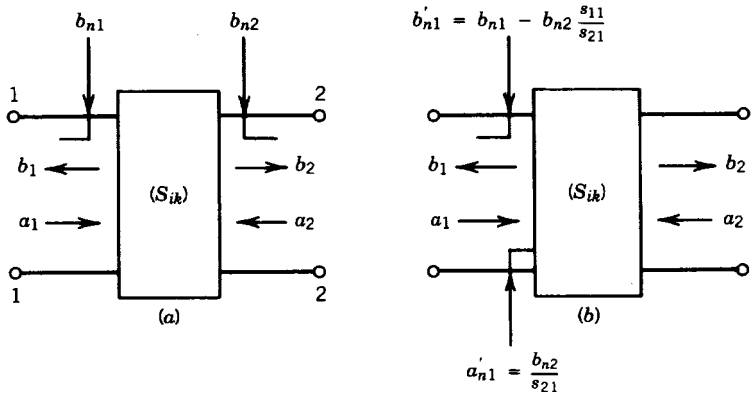


Figure 3-19 *S*-parameter form of linear noisy two-ports.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_{n1} \\ b_{n2} \end{bmatrix} \tag{3-24}$$

There are different physical origins for the various sources of noise. Typically, thermal noise is generated by resistances and loss in the circuit or transistor, whereas shot noise is generated by current flowing through semiconductor junctions and vacuum tubes. Since these many sources of noise are represented by only two noise sources at the device input, the two equivalent noise sources are often a complicated combination of the circuit internal noise sources. Often, some fraction of V_A and I_A is related to the same noise source. This means that V_A and I_A are not independent in general. Before we can use V_A and I_A to calculate the noise figure of the two-port, we must calculate the correlation between the V_A and I_A shown in Figure 3-18.

The noise source V_A represents all the device noise referred to the input when the generator impedance is zero; that is, when the input is short-circuited. The noise source I_A represents all the device noise referred to the input when the generator admittance is zero; that is, the input is open-circuited.

The correlation of these two noise sources considerably complicates analysis. By defining a correlation admittance, we can simplify the mathematics and get some physical intuition for the relationship between noise figure and generator admittance. Since some fraction of I_A will be correlated with V_A, we split I_A into correlated and uncorrelated parts as follows:

$$I_A = I_n + I_u \tag{3-25}$$

I_u is the part of I_A uncorrelated with V_A. Since I_n is correlated with V_A, we can say that I_n is proportional to V_A and the constant of proportionality is the correlation admittance:

$$I_n = Y_{\text{cor}} V_A \tag{3-26}$$

This leads us to

$$I_A = Y_{\text{cor}} V_A + I_u \tag{3-27}$$

The following derivation of noise figure will use the correlation admittance. Y_{cor} is not a physical component located somewhere in the circuit. Y_{cor} is a complex number derived by correlating the random variables I_A and V_A. To calculate Y_{cor}, we multiply each side of Eq. (3-27) by V_A^* and average the result. This gives

$$\overline{V_A^* I_A} = Y_{\text{cor}} \overline{V_A^2} \tag{3-28}$$

where the I_u term averaged to zero since it was uncorrelated with V_A. The correlation admittance is thus given by

$$Y_{\text{cor}} = \frac{\overline{V_A^* I_A}}{V_A^2} \tag{3-29}$$

Often, people use the term "correlation coefficient." This normalized quantity is defined as

$$c = \frac{\overline{V_A^* I_A}}{\sqrt{\overline{V_A^2} I_A^2}} = Y_{\text{cor}} \sqrt{\frac{\overline{V_A^2}}{I_A^2}} \tag{3-30}$$

Note that the dual of this admittance description is the impedance description. Thus, the impedance representation has the same equations as above with Y replaced by Z, I replaced by V, and V replaced by I.

V_A and I_A represent internal noise sources in the form of a voltage source acting in series with the input voltage and a source of current flowing in parallel with the input current. This representation conveniently leads to the four noise parameters needed to describe the noise performance of the two-port. Again using the Nyquist formula, the open-circuit voltage of a resistor at the temperature T is

$$\overline{V_A^2} = 4kTRB \tag{3-31}$$

This voltage is a mean-square fluctuation (or spectral density). It is the method used to calculate the noise identity. We could also define a noise equivalent resistance for a noise voltage as

$$R_n = \frac{\overline{V_A^2}}{4kTB} \tag{3-32}$$

The resistor R_n is not a physical resistor but can be used to simulate different portions of the noise equivalent circuit.

In a similar manner a mean-square current fluctuation can be represented in terms of an equivalent noise conductance G_n, which is defined by

$$G_n = \frac{\overline{I_A^2}}{4kTB} \tag{3-33}$$

and

$$G_u = \frac{\overline{I_u^2}}{4kTB} \tag{3-34}$$

for the case of the uncorrelated noise component. The input generator to the two-port has a similar contribution:

$$G_G = \frac{I_G^2}{4kTB} \tag{3-35}$$

with Y_G being the generator admittance and G_G being the real part. With the definition of F above, we can write

$$F = 1 + \left| \frac{I_A + Y_G V_A}{I_G} \right|^2 \tag{3-36}$$

The use of the voltage V_A and the current I_A has allowed us to combine all the effects of the internal noise sources.

We can use the previously defined [Eq. (3-29)] correlation admittance, $Y_{\text{cor}} = G_{\text{cor}} + jB_{\text{cor}}$, to simplify Eq. (3-36). First, we determine the total noise current:

$$\overline{I_A^2} = 4kT(|Y_{\text{cor}}|^2 R_n + G_u)B \tag{3-37}$$

where R_n and G_u are defined in Eqs. (3-32) and (3-34). The noise factor can now be determined:

$$F = 1 + \frac{G_u}{G_g} + \frac{R_n}{G_g}[(G_G + G_{\text{cor}})^2 + (B_G + B_{\text{cor}})^2] \tag{3-38}$$

$$F = 1 + \frac{R_u}{R_g} + \frac{G_n}{R_g}[(R_G + R_{\text{cor}})^2 + (X_G + X_{\text{cor}})^2] \tag{3-39}$$

The noise factor is a function of various elements, and the optimum impedance for the best noise figure can be determined by minimizing F with respect to generator reactance and resistance. This gives

$$R_{0n} = \sqrt{\frac{R_u}{G_n} + R_{\text{cor}}^2} \tag{3-40}$$

$$X_{0n} = -X_{\text{cor}} \tag{3-41}$$

and

$$F_{\min} = 1 + 2G_n R_{\text{cor}} + 2\sqrt{R_u G_n + (G_n R_{\text{cor}})^2} \tag{3-42}$$

(To distinguish between optimum noise and optimum power, we have introduced the convention 0*n* instead of the more familiar abbreviation *opt*.) At this point we see that the optimum condition for minimum noise figure is not a conjugate power match at the input port. We can explain this by recognizing that the noise source V_A and I_A represent all the

two-port noise, not just the thermal noise of the input port. We should observe that the optimum generator susceptance, $-X_{corr}$, will minimize the noise contribution of the two noise generators.

In rearranging for conversion to S parameters, we write

$$F = F_{\min} + \frac{g_n}{R_G}|Z_G - Z_{0n}|^2 \tag{3-43}$$

$$F = F_{\min} + \frac{R_n}{G_G}|Y_G - Y_{0n}|^2 \tag{3-44}$$

From the definition of the reflection coefficient,

$$\Gamma_G = \frac{Y_0 - Y_G}{Y_0 + Y_G} \tag{3-45}$$

and with

$$g_G = \frac{G_G}{Y_0} \tag{3-46}$$

$$r_n = \frac{R_n}{Z_0} \tag{3-47}$$

the normalized equivalent noise resistance,

$$F = F_{\min} + \frac{4r_n|\Gamma_G - \Gamma_{0n}|^2}{g_G(1 - |\Gamma|^2)|1 + \Gamma_{0n}|^2} \tag{3-48}$$

$$r_n = (F_{50} - F_{\min})\frac{|1 + \Gamma_{0n}|^2}{4|\Gamma_{0n}|^2} \tag{3-49}$$

$$\Gamma_{0n} = \frac{Z_{0n} - Z_0}{Z_{0n} + Z_0} \tag{3-50}$$

The noise performance of any linear two-port can now be determined if the values of the four noise parameters, $F_{\min}$, r_n (= $R_n/50$), and Γ_{0n} (magnitude and phase), are known.

Figure 3-20 shows the noise factor of a high-frequency transistor as a function of B_g for G_g = constant and as a function of G_g for $B_g = B_{opt}$.

Noise Figure of Cascaded Networks. In a system with many circuits connected in cascade (Figure 3-21), we must consider the contributions of the various circuits. In considering the equivalent noise resistor R_n in series with the input circuit,

$$F = \frac{R_G + R_n}{R_G} \tag{3-51}$$

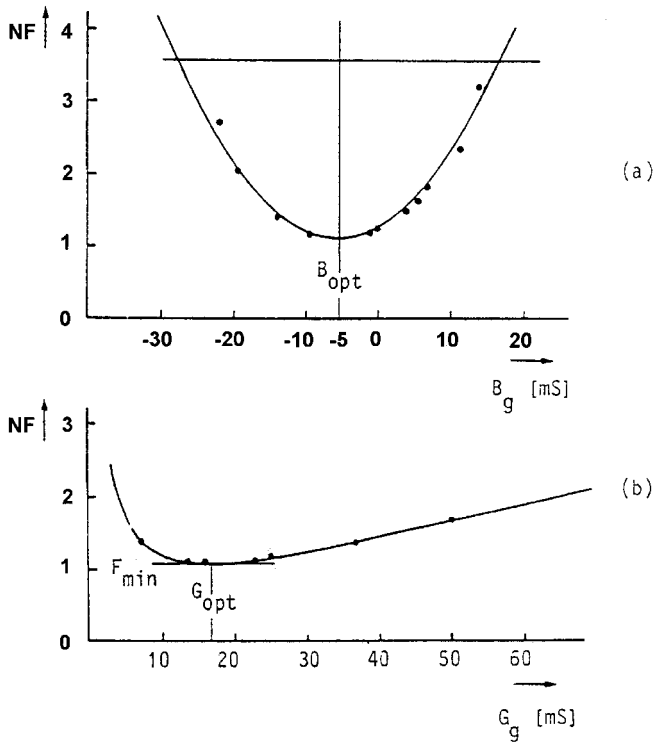


Figure 3-20 Noise factor in high-frequency BJTs for $f = 600$ MHz: (a) as a function of B_g for G_g = constant and (b) as a function of G_g for $B_g = B_{opt}$.

$$F = 1 + \frac{R_n}{R_G} \tag{3-52}$$

The excess noise added by the circuit is R_n/R_G.

In considering two cascaded circuits a and b, by definition the available noise at the output of b is

$$N_{ab} = F_{ab} G_{ab} kTB \tag{3-53}$$

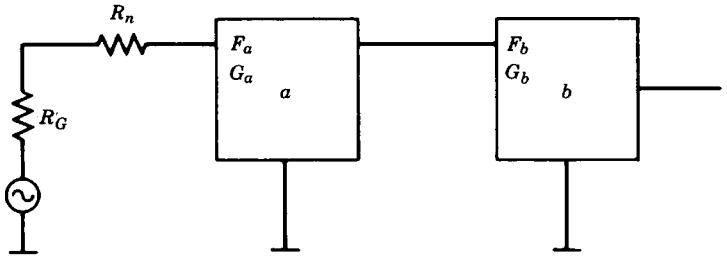


Figure 3-21 Cascaded noisy two-ports with the noise factors F_a and F_b and the gain figures G_a and G_b.

with B the equivalent noise bandwidth in which the noise is measured. The total available gain G is the product of the individual available gains, so

$$N_{ab} = F_{ab} G_a G_b kTB \tag{3-54}$$

The available noise from network a at the output of network b is

$$N_{a/b} = N_a G_b = F_a G_a G_b kTB \tag{3-55}$$

The available noise added by network b (its excess noise) is

$$N_{b/b} = (F_b - 1) G_b kTB \tag{3-56}$$

The total available noise N_{ab} is the sum of the available noise contributed by the two networks:

$$\begin{aligned} N_{ab} &= N_{a/b} + N_{b/b} = F_a G_a G_b kTB + (F_b - 1) G_b kTB \\ &= \left(F_a + \frac{F_b - 1}{G_a} \right) G_a G_b kTB \end{aligned} \tag{3-57}$$

$$F_{ab} = F_a + \frac{F_b - 1}{G_a} \tag{3-58}$$

For any number of circuits, this can be extended to

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \cdots \tag{3-59}$$

When considering a long chain of cascaded amplifiers, there will be a minimum noise figure achievable for this chain. This is a figure of merit and was proposed by Haus and Adler [6]. It is calculated by rearranging Eq. (3-59):

$$(F_{\text{tot}})_{\min} = (F_{\min} - 1) + \frac{F_{\min} - 1}{G_A} + \frac{F_{\min} - 1}{G_A^2} + \cdots + 1 \tag{3-60}$$

where $F_{\min}$ is the minimum noise figure for each stage and G_A is the available power gain of the identical stages. Using

$$\frac{1}{1 - X} = 1 + X + X^2 + \cdots \tag{3-61}$$

we find a quantity $(F_{\text{tot}} - 1)$, which is defined as the noise measure M. The minimum noise measure

$$(F_{\text{tot}})_{\min} - 1 = \frac{F_{\min} - 1}{1 - 1/G_A} = M_{\min} \tag{3-62}$$

refers to the noise of an infinite chain of optimally tuned, low-noise stages, so it represents a lower limit on the noise of an amplifier.

The minimum noise measure $M_{\min}$ is an invariant parameter and is not affected by feedback. It is somewhat similar to a gain–bandwidth product, in its use as a system invariant. The minimum noise measure is achieved when the amplifier is tuned for the available power gain and $\Gamma_G = \Gamma_{0n}$, given by Eq. (3-50).

Influence of External Parasitic Elements. Mounting an active two-port such as a transistor usually adds stray capacitance and lead inductance to the device, as shown in Figure 3-22. These external components consisting of transmission lines and parasitic reactances modify the noise parameters and the gain. Some researchers have published the results of these parasitic effects and have made manual computations or used some limited computer programs.

In a paper by Fukui [7], an attempt was made to determine the necessary equations, but the formulas are too involved even for pocket calculators. A more generic study by Iversen [8] is also very involved because of the various matrix manipulations and is more suitable for a computer. Besser's paper [9] in the *IEEE MTT-S* in 1975 and Vendelin's paper [10] in the same issue have shown for the first time some practical results using computers and even optimization methods, using an early version of COMPACT. The intention of these investigations was to find feedback that modifies the device noise and scattering parameters such that a noise match could also provide a low-input VSWR. It can be seen from these discussions that some feedback, besides resulting in some gain reduction, may improve the noise matching at the input for a limited frequency range.

A more recent paper by Suter [11] based on a report by Hartmann and Strutt [12] has given a simple transformation starting from the *S*-parameters and the noise parameters from common-source (or common-emitter) measurements. The noise parameters for the "packaged" device are calculated. This means that the parameters for the "new" device, including

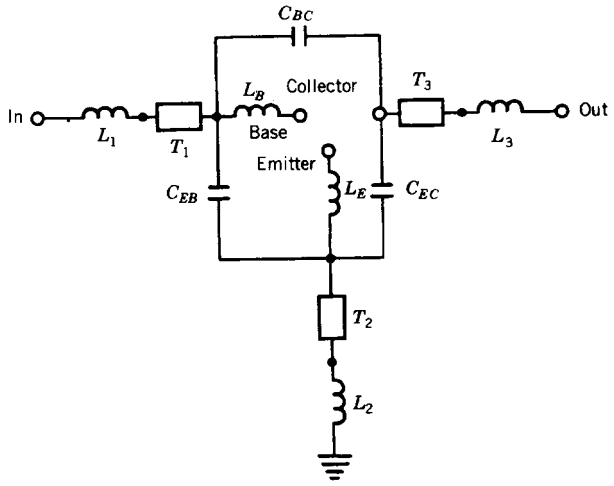
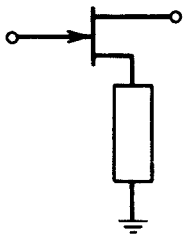


Figure 3-22 Equivalent circuit of the transistor package.

Table 3-2 Transformation matrix parameters

Series Feedback



$$Z_S = R_S + jX_S = Z_S$$

$$[n] = \begin{bmatrix} n_{11} = 1 & n_{12} = Z_0 \dfrac{S_{21}M - S'_{21}N}{S_{21}C'_1 + S'_{21}C_1}\ \Omega \\ n_{21} = 0 & n_{22} = \dfrac{S_{21}C'_1}{S_{21}C'_1 + S'_{21}C_1} \end{bmatrix}$$

where

$$S'_{11} = S'_{22} = \frac{-1}{1 + 2Z_S}$$

$$S'_{12} = S'_{21} = \frac{2Z_S}{1 + 2Z_S}$$

$$M = (1 + S'_{11})(1 - S'_{22}) + S'_{12}S'_{21}$$

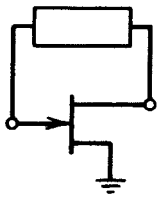
$$N = (1 + S_{11})(1 - S_{22}) + S_{12}S_{21}$$

$$C_1 = (1 - S_{11})(1 - S_{22}) - S_{12}S_{21}$$

$$C'_1 = (1 + S'_{11})(1 - S'_{22}) - S'_{12}S'_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

Shunt Feedback



$$Z_P = R_P + jX_P = Z_P$$

$$[n] = \begin{bmatrix} n_{11} = \dfrac{S_{21}C'_2}{S_{21}C'_2 + S'_{21}C_2} & n_{12} = 0\ \Omega \\ n_{21} = \dfrac{1}{Z_0}\dfrac{S_{21}P - S'_{21}Q}{S_{21}C'_2 + S'_{21}C_2}\,S & n_{22} = 1 \end{bmatrix}$$

where

$$S'_{11} = S'_{22} = \frac{Z_P}{2 + Z_P}$$

(*continued*)

Table 3-2 (*Continued*)

$$S'_{12} = S'_{21} = \frac{2}{2 + Z_P}$$

$$P = (1 - S'_{11})(1 + S'_{22}) + S'_{12}S'_{21}$$

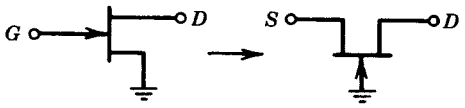
$$Q = (1 - S_{11})(1 + S_{22}) + S_{12}S_{21}$$

$$C_2 = (1 + S_{11})(1 + S_{22}) - S_{12}S_{21}$$

$$C'_2 = (1 + S'_{11})(1 + S'_{22}) - S'_{12}S'_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

Common Gate



$$[n] = \begin{bmatrix} n_{11} = \dfrac{2S_{21}}{-2S_{21} + C_4} & n_{12} = 0\ \Omega \\ n_{21} = \dfrac{1}{Z_0}\dfrac{C_3C_4 - 4S_{12}S_{21}}{V(-2S_{21} + C_4)}S & n_{22} = -1 \end{bmatrix}$$

where

$$V = (1 + S_{11})(1 + S_{22}) - S_{12}S_{21}$$

$$C_3 = (1 - S_{11})(1 + S_{22}) + S_{12}S_{21}$$

$$C_4 = (1 + S_{11})(1 - S_{22}) + S_{12}S_{21}$$

$$[S]_{\text{DEVICE}} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \text{common-source } S \text{ parameters}$$

the common-gate (or common-base) case, are calculated. The equations are device dependent. They are valid for any active two-port.

A transformation matrix, n, may be used to combine the noise sources of the various circuit configurations. The transformation matrix parameters are given in Table 3-2 for series feedback, shunt feedback, and the common-gate (base) case, which will be important for oscillator analysis. The transformation matrix gives the new noise parameters as follows:

$$R'_n = R_n|n_{11} + n_{12}Y_{\text{cor}}|^2 + G_n|n_{12}|^2 \tag{3-63}$$

$$G'_n = \frac{G_nR_n}{R'_n}|n_{11}n_{22} - n_{12}n_{21}|^2 \tag{3-64}$$

$$Y'_{\text{cor}} = \frac{R_n}{R'_n}(n_{21} + n_{22}Y_{\text{cor}})(n^*_{11} + n^*_{12}Y^*_{\text{cor}}) + \frac{G_n}{R'_n}n_{22}n^*_{12} \tag{3-65}$$

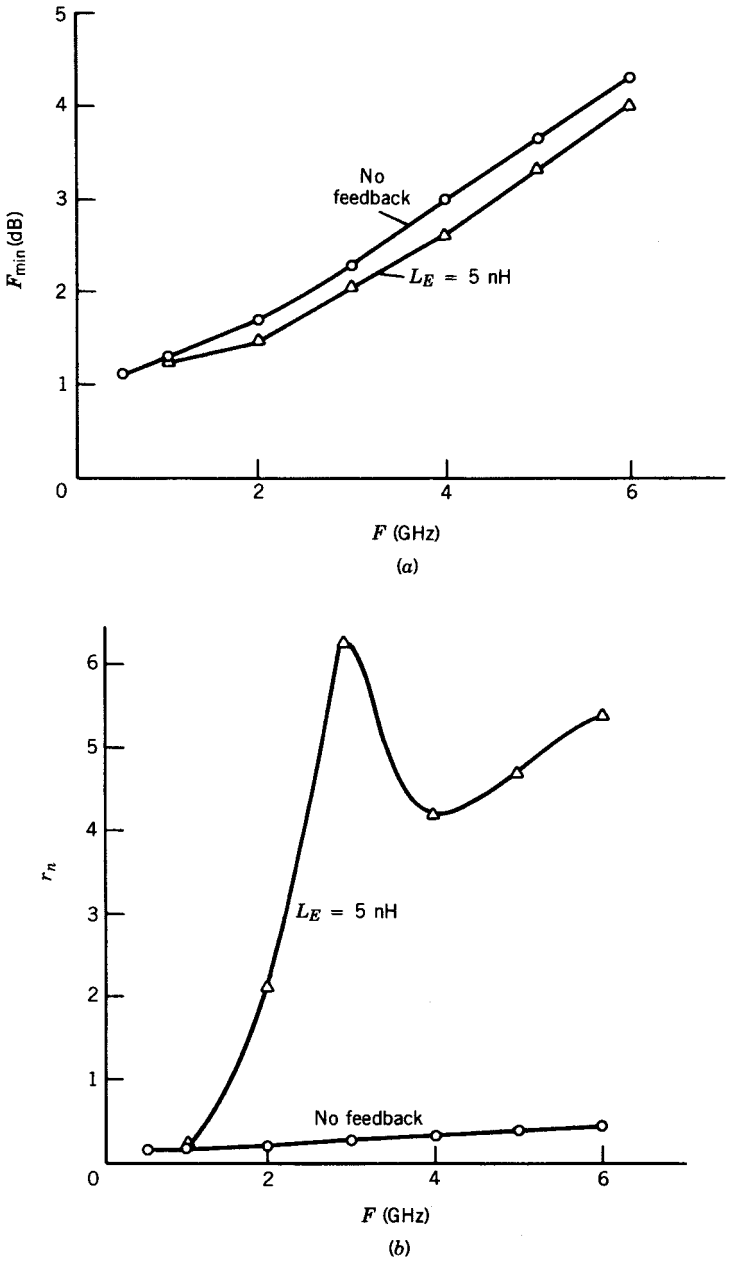


Figure 3-23 Noise parameters versus feedback for the AT-41435 silicon bipolar transistor: (a) F_{min} versus frequency and feedback; (b) r_n versus frequency and feedback; (c) Γ_{0n} versus frequency and feedback.

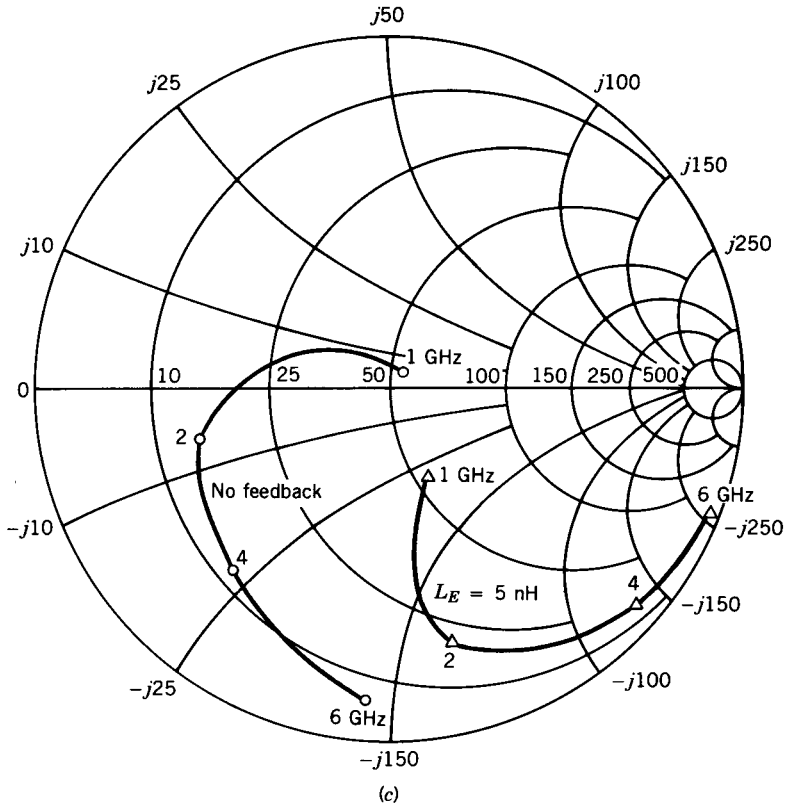


Figure 3-23 (*Continued*)

A final transformation to the more common noise-parameter format given by Eq. (3-44) is still needed [10]:

$$F_{\min} = 1 + 2R_n'(G_{\text{cor}}' + G_{0n}') \tag{3-66}$$

$$R_n = R_n' \tag{3-67}$$

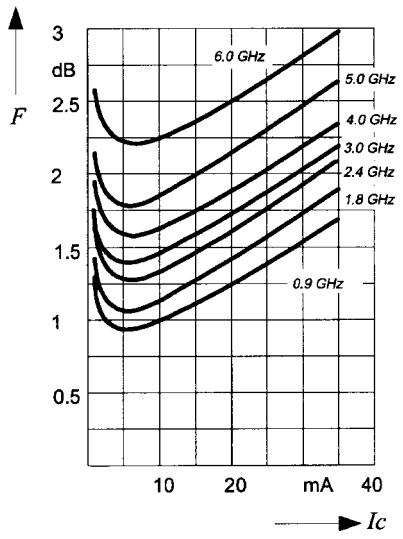
$$G_{0n} = \sqrt{\frac{G_n'}{R_n'} + G_{\text{cor}}'^2} \tag{3-68}$$

$$B_{0n} = -B_{\text{cor}}' \tag{3-69}$$

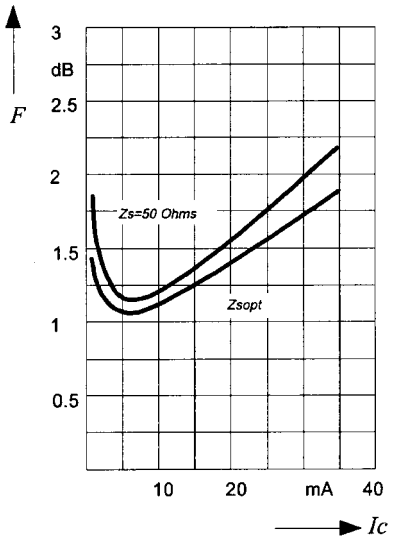
Figure 3-23 shows the noise figure as a function of external feedback for a low-noise microwave bipolar transistor, the AT-41435.

Bias-Dependent Noise Parameters. Since we have elaborated on the effect of parasitics, it is now useful to point out that noise is also a strong function of bias point and frequency. Figure 3-24 shows the noise figure expressed in dB as a function of different dc currents and frequencies for a BFP420 BJT by Siemens. A change in the collector voltage will also affect the noise parameters, not quite as strongly as shown here, but still significantly. The reason

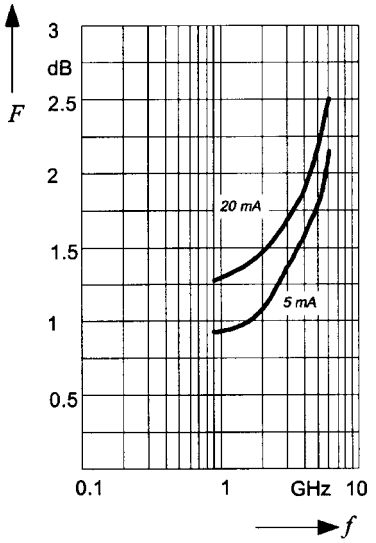
Noise Figure
versus Collector Current
V_{CE} = 2 V, $Z_S = Z_{Sopt}$



Noise Figure
versus Collector Current
VCE = 2 V, f = 1.8 GHz



Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA,
$Z_S = Z_{Sopt}$



Source Impedance for min.
Noise Figure versus Frequency
V_{CE} = 2 V, I_C = 5 mA / 20 mA

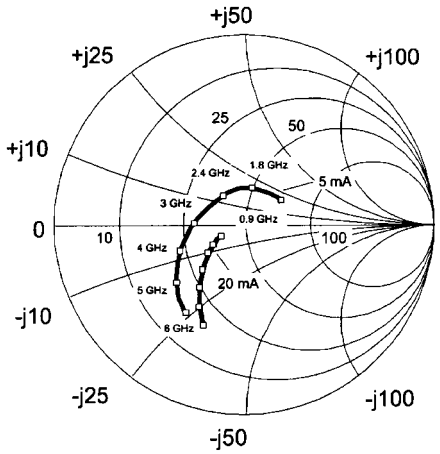


Figure 3-24 Noise figure versus collector current, noise figure versus frequency, and source impedance for minimum noise figure versus frequency for the Siemens BFP420 transistor.

Table 3-3 BFP420 common-emitter noise parameters

f	$F_{\min}{}^a$	$G_a{}^a$	Γ_{opt}		R_N	r_n	$F_{50\Omega}{}^b$	$\lvert S_{21} \rvert^{2b}$
(GHz)	dB	(dB)	(Magnitude)	(Angle)	(Ω)		(dB)	(dB)
V_{CE} = 2 V, I_C = 5 mA								
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	–175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	–157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	–142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	–123.0	15.0	0.30	3.30	5.3

[a]Input matched for minimum noise figure, output for maximum gain.
[b]$Z_S = Z_L = 50\ \Omega$.

for the turning point of the noise figure has to do with the fact that for lower currents the f_T cutoff frequency has not reached its peak yet, and therefore the noise increases. The increase of the noise at higher currents is due to the Schottky noise calculated from the equation

$$P_n = 2Iq \tag{3-70}$$

where P_n is the noise power, I is the saturation, collector, or drain current, and q is the charge of an electron (1.60×10^{-19} coulomb).

Table 3-3 shows the common-emitter noise parameters for the BFP420.

Noise Circles. From the section entitled "Influence of External Parasitic Elements," we see that the noise factor is a strong function of the generator admittance (or impedance) presented to the input terminals of the noisy two-port. Noise tuning is used to change the values of the input admittance to obtain the best noise performance. There is a range of values of input reflection coefficient over which the noise figure is constant. In plotting these points of constant noise figure, we obtain the so-called noise circles, which can be drawn on the Smith chart Γ_G plane [13]. Starting with the noise equation [see Eq. (3-44)] for a 50-Ω generator impedance, we find that

$$F_{50} = F_{\min} + 4r_n \frac{|\Gamma_{0n}|^2}{|1 + \Gamma_{0n}|^2} \tag{3-71}$$

We want to find the position of the reflection coefficient on the Smith chart, as in the case of the gain circles, for which F is constant. First we rearrange Eq. (3-67) to read

$$r_n = (F_{50} - F_{\min}) \frac{|1 + \Gamma_{0n}|^2}{4|\Gamma_{0n}|^2} \tag{3-72}$$

By introducing

$$N_i = \frac{F_i - F_{\min}}{4r_n} |1 + \Gamma_{0n}|^2 \tag{3-73}$$

we can find an expression for a circle of constant noise figure as introduced by Rothe and Dahlke [4] (see also Fukui [13]). The center for the noise circle is

$$C_i = \frac{|\Gamma_{0n}|}{1 + N_i} \tag{3-74}$$

and the radius is

$$r_i = \frac{\sqrt{N_i^2 + N_i(1 - |\Gamma_{0n}|^2)}}{1 + N_i} \tag{3-75}$$

with the definition of N used previously. However, if we consider only the minimum noise figure for a given device, we will not obtain the minimum noise figure for the multistage amplifier system. This was explained when the noise measure was introduced. (See Figure 3-25.) Therefore, a better way to design the amplifier would be to use circles at constant noise measure instead of circles of constant noise figure. This was recently done by Poole

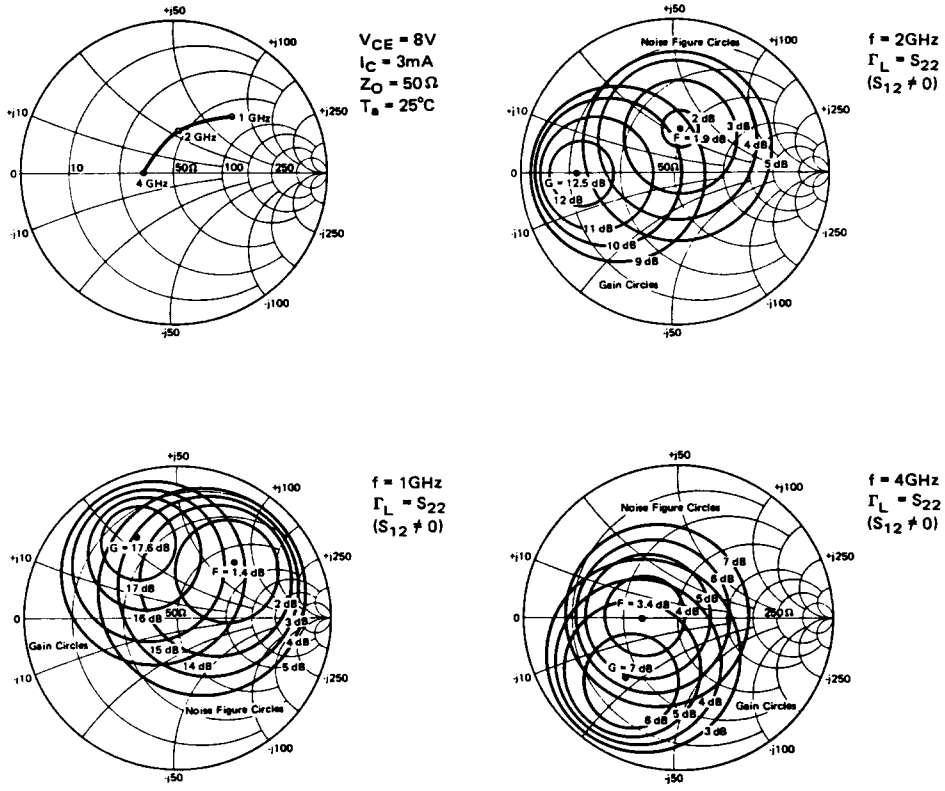


Figure 3-25 Typical noise figure circles and gain circles.

and Paul [14]. They derived the expressions for the noise measure circles as a function of S parameters, noise parameters, and Γ_G, using

$$G_A = \frac{|S_{21}|^2(1 - |\Gamma_G|^2)}{(1 - |S_{22}|^2) + |\Gamma_G|^2 (|S_{11}|^2 - |\Delta|^2) - 2\,\mathrm{Re}(\Gamma_G C_1)} \tag{3-76}$$

where

$$C_1 = S_{11} - S_{22}^* \Delta$$

$$\Delta = S_{11}S_{22} - S_{12}S_{11}$$

In terms of the generator reflection coefficient, the noise measure can be expressed as

$$M = [(F_{\min} - 1)(1 - |\Gamma_G|^2)|S_{21}|^2|1 + \Gamma_{0n}|^2 + 4r_n|S_{21}|^2|\Gamma_G - \Gamma_{0n}|^2]$$

$$\times \left\{|1 + \Gamma_{0n}|^2[|S_{21}|^2(1 - |\Gamma_G|^2) - (1 - |S_{22}|^2) - |\Gamma_G|^2(|S_{11}|^2 - |\Delta|^2) + 2\,\mathrm{Re}(\Gamma_G C_1)]\right\}^{-1} \tag{3-77}$$

Equation (3-77) can be shown to represent circles in the source reflection coefficient plane described by the following:

$$|\Gamma_G|^2 + |\Gamma_m|^2 - \Gamma_G\Gamma_m^* - \Gamma_G^*\Gamma_m = r_m^2 \tag{3-78}$$

The centers and radii of the constant-noise-measure circles are given by

$$C_m = \frac{M|1 + \Gamma_{0n}|^2 C_1^* + 4r_n|S_{21}|^2\Gamma_{0n}}{M|1 + \Gamma_{0n}|^2 P + |S_{21}|^2(4r_n - W)} \tag{3-79}$$

$$r_m = \frac{\sqrt{M^2M_a + MM_b + M_c}}{M|1 + \Gamma_{0n}|^2 P + |S_{21}|^2(4r_n - W)} \tag{3-80}$$

where

$$P = |S_{21}|^2 + |S_{11}|^2 - |\Delta|^2$$

$$Q = |S_{21}|^2 + |S_{22}|^2 - 1$$

$$W = |1 + \Gamma_{0n}|^2(F_{\min} - 1)$$

$$M_a = |1 + \Gamma_{0n}|^4(PQ + |C_1|^2)$$

$$M_b = |1 + \Gamma_{0n}|^2|S_{21}|^2[8r_n\,\mathrm{Re}(\Gamma_{0n}C_1)$$

$$- (4r_n|\Gamma_{0n}|^2 + W)P - (W - 4r_n)Q]$$

$$M_c = |S_{21}|^4[W - 4r_n(1 - |\Gamma_{0n}|^2)]$$

The value of the minimum noise measure can be found by considering the noise measure circle of zero radius; that is, set r_m equal to zero in Eq. (3-80). This results in

$$M_{\min} = \frac{-M_b \pm \sqrt{M_b^2 - 4M_a M_c}}{2M_a} \tag{3-81}$$

Equation (3-81) yields the same value of $M_{\min}$ as would have been obtained by using the immittance parameter equation given by Fukui [13] and can, therefore, be considered as the reflection coefficient plane analog of Fukui's expression. The elimination of the need for the parameter R_{eg}, however, results in a considerable simplification as compared with the earlier approach [13].

The value of the minimum noise measure is taken as the smallest nonnegative value of $M_{\min}$ given by Eq. (3-81). The source reflection coefficient that results in the minimum noise measure can now be obtained by employing Eq. (3-79):

$$\Gamma_{0m} = \frac{M_{\min}|1 + \Gamma_{0n}|^2 C_1^* + 4r_n|S_{21}|^2\Gamma_{0n}}{M_{\min}|1 + \Gamma_{0n}|^2 P + |S_{21}|^2\,(4r_n - W)} \tag{3-82}$$

The output reflection coefficient of the device, when Γ_{0m} is presented to the input port, is given by

$$S'_{22} = \frac{S_{22} - \Delta\Gamma_{0m}}{1 - S_{11}\Gamma_{0m}} \tag{3-83}$$

Noise Correlation in Linear Two-Ports Using Correlation Matrices. In the introduction to two-port noise theory, it was indicated that noise correlation matrices form a general technique for calculating noise in n-port networks. Haus and Adler [6] have described the theory behind this technique. In 1976, Hillbrand and Russer [15] published equations and transformations that aid in supplying this method to two-port CAD.

This method is useful because it forms a base from which we can rigorously calculate the noise of linear two-ports combined in arbitrary ways. For many representations, the method of combining the noise parameters is as simple as that for combining the circuit element matrices. In addition, noise correlation matrices can be used to calculate the noise in linear frequency conversion circuits. The following is an introduction to this subject.

A linear, noisy two-port can be modeled as a noise-free two-port with two additional noise sources. These noise sources must be chosen so that they add directly to the resulting vector of the representation, as shown in Eqs. (3-84) and (3-85) and Figure 3-17:

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \tag{3-84}$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{3-85}$$

where the i and v vectors indicate noise sources for the y and z representations, respectively. This two-port example can be extended to n-ports in a straightforward, obvious way.

Since the noise vector for any representation is a random variable, it is much more convenient to work with the noise correlation matrix. The correlation matrix gives us deterministic numbers with which to calculate. The correlation matrix is formed by taking the mean value of the outer product of the noise vector. This is equivalent to multiplying the noise vector by its adjoint (complex conjugate transpose) and averaging the result:

$$\langle \bar{i}\ \bar{i}^{+} \rangle = \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} [i_1^* \quad i_2^*] = \begin{bmatrix} \langle i_1 i_1^* \rangle & \langle i_1 i_2^* \rangle \\ \langle i_1^* i_2 \rangle & \langle i_2 i_2^* \rangle \end{bmatrix} + [C_y] \tag{3-86}$$

where the angular brackets denote the average value.

Note that the diagonal terms are the "power" spectrum of each noise source and the off-diagonal terms are complex conjugates of each other and represent the "cross-power" spectra of the noise sources. "Power" is used because these magnitude-squared quantities are proportional to power.

To use these correlation matrices in circuit analysis, we must know how to combine them and how to convert them between various representations. An example using y matrices will illustrate the method for combining two-ports and their correlation matrices. Given two matrices y and y', when we parallel them we have the same port voltages, and the terminal currents add (Figure 3-26):

$$\begin{aligned} I_1 &= y_{11}V_1 + y_{12}V_2 + y'_{11}V_1 + y'_{12}V_2 + i_1 + i'_1 \\ I_2 &= y_{21}V_1 + y_{22}V_2 + y'_{21}V_1 + y'_{22}V_2 + i_2 + i'_2 \end{aligned} \tag{3-87}$$

or

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} + y'_{11} & y_{12} + y'_{12} \\ y_{21} + y'_{21} & y_{22} + y'_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 + i'_1 \\ i_2 + i'_2 \end{bmatrix} \tag{3-88}$$

Here we can see that the noise current vectors add just as the y parameters add. Converting the new noise vector to a correlation matrix yields

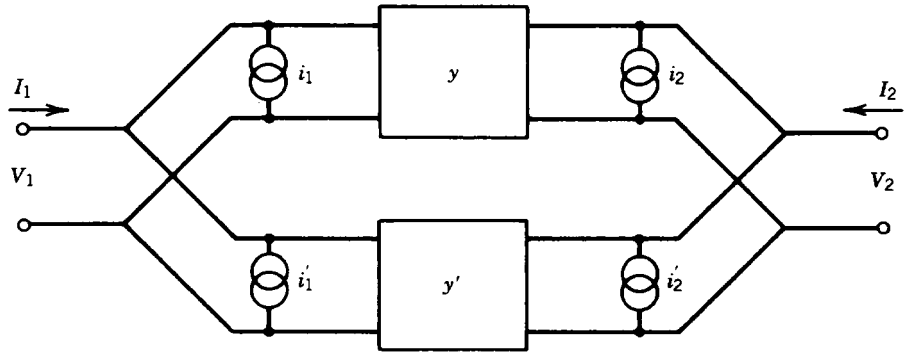


Figure 3-26 Parallel combination of two-ports using y parameters.

$$\langle \bar{i}_{\text{new}} \bar{i}^{+}_{\text{new}} \rangle = \left\langle \begin{bmatrix} i_1 + i_1' \\ i_2 + i_2' \end{bmatrix} [i_1^* + i_1'^* \quad i_2 i_2'^*] \right\rangle \tag{3-89}$$

$$= \begin{bmatrix} \langle i_1 i_1^* \rangle + \langle i_1' i_1'^* \rangle & \langle i_1 i_2^* \rangle + \langle i_1' i_2'^* \rangle \\ \langle i_2 i_1^* \rangle + \langle i_2' i_1'^* \rangle & \langle i_2 i_2^* \rangle + \langle i_2' i_2'^* \rangle \end{bmatrix} \tag{3-90}$$

The noise sources from different two-ports must be uncorrelated, so there are no cross products of different two-ports. By inspection, Eq. (3-90) is just the addition of the correlation matrices for the individual two-ports, so

$$[C_{y\text{new}}] = [C_y] + [C_y'] \tag{3-91}$$

The same holds true for g, h, and z parameters, but $ABCD$ parameters have the more complicated form shown below. If

$$[A_{\text{new}}] = [A][A'] \tag{3-92}$$

then

$$[C_{A\text{new}}] = \{C_A] + [A]\,[C_{A'}]\,[A]^{+} \tag{3-93}$$

The transformation of one representation to another is best illustrated by an example. Let us transform the correlation matrix for a Y representation to a Z representation. Starting with

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = [Y] \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \tag{3-94}$$

we can move the noise vector to the left side and invert y:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = [Y^{-1}] \begin{bmatrix} I_1 - i_1 \\ I_2 - i_2 \end{bmatrix} = [Y^{-1}] \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + [Y^{-1}] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-95}$$

Table 3-4 π Noise matrix $T_{\alpha\beta}$ transformation

		Original Form (α Form)					
		Y		Z		A	
Resulting Form (β Form)	Y	1	0	y_{11}	y_{12}	$-y_{11}$	1
		0	1	y_{21}	y_{22}	$-y_{21}$	0
	Z	z_{11}	z_{12}	1	0	1	$-z_{11}$
		z_{21}	z_{22}	0	1	0	$-z_{21}$
	A	0	A_{12}	1	$-A_{11}$	1	0
		1	A_{22}	0	$-A_{21}$	0	1

Since $(Y)^{-1} = (Z)$, we have

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = [Z] \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + [Z] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-96}$$

so

$$= [Z] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} = [T_{yz}] \begin{bmatrix} -i_1 \\ -i_2 \end{bmatrix} \tag{3-97}$$

where the signs of i_1 and i_2 are superfluous since they will cancel when the correlation matrix is formed. Here the transformation of the Y noise current vector to the Z noise voltage vector is done simply by multiplying by (Z). Other transformations are shown in Table 3-4.

To obtain the noise correlation matrix, we again form the mean of the outer product:

$$\langle vv^+ \rangle = \begin{bmatrix} \langle v_1 v_1^* \rangle & \langle v_1 v_2^* \rangle \\ \rangle v_1^* v_2 \rangle & \langle v_2 v_2^* \rangle \end{bmatrix} = [Z] \left\langle \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} [i_1^* \quad i_2^*] \right\rangle [Z]^+ \tag{3-98}$$

or

$$[C_z] = [Z]\,[C_y]\,[Z]^+ \tag{3-99}$$

where

$$v^+ = [i_1^* \quad i_2^*]\,[Z]^+ \tag{3-100}$$

This is called a congruence transformation. The key to all of these derivations is the construction of a correlation matrix from the noise vector, as shown in Eq. (3-90).

These correlation matrices may easily be derived from the circuit matrices of passive circuits with only thermal noise sources. For example,

$$[C_z] = 2kT\Delta f\,\mathrm{Re}([Z]) \tag{3-101}$$

and

$$[C_y] = 2kT\Delta f\,\mathrm{Re}([Y]) \tag{3-102}$$

The $2kT$ factor comes from the double-sided spectrum of thermal noise. The correlation matrix from the $ABCD$ matrix may be related to the noise figure, as shown by Hillbrand and Russer [15]. We have

$$F = 1 + \frac{\overline{Y}[C_a]\overline{Y}^+}{2kT\,\mathrm{Re}(Y_G)} \tag{3-103}$$

where

$$\overline{Y} = \begin{bmatrix} Y_G \\ 1 \end{bmatrix}$$

Expressing the noise factor in terms of the correlation matrix, we obtain a complete formula:

$$F = 1 + \frac{C_{22}^{A}(f) + 2\,\mathrm{Re}\{Y_g(f)\,C_{12}^{A}(f)\} + |Y_g(f)|^2\,C_{11}^{A}(f)}{2kT_0\,\mathrm{Re}\{Y_g(f)\}} \tag{3-104}$$

Once we transform this in the Y parameter form, we obtain the following equation:

$$F(f) = F_{\min}(f) + \frac{R_n(f)|Y_{\mathrm{opt}}(f) - Y_g(f)|^2}{\mathrm{Re}\{Y_g(f)\}} \tag{3-105}$$

It should be noted that all these values are frequency dependent, as expressed in this equation.

The *ABCD* correlation matrix can be written in terms of the noise figure parameters as (double-sided spectrum)

$$[C_a] = 2kT\begin{bmatrix} R_n & \dfrac{F_0 - 1}{2} - R_n Y_{0n}^{*} \\ \dfrac{F_0 - 1}{2} - R_n Y_{0n} & R_n|Y_{0n}|^2 \end{bmatrix} \tag{3-106}$$

The noise correlation matrix method forms an easy and rigorous technique for handling noise in networks. This technique allows us to calculate the total noise for complicated networks by combining the noise matrices of subcircuits. It should be remembered that although noise correlation matrices apply to n-port networks, noise figure calculations apply only to pairs of ports. The parameters of the C_a matrix can be used to give the noise parameters:

$$Y_{0n} = \sqrt{\frac{C_{ii*}}{C_{uu*}} - \left[\mathrm{Im}\left(\frac{C_{ui*}}{C_{uu*}}\right)\right]^2} + j\,\mathrm{Im}\left(\frac{C_{ui*}}{C_{uu*}}\right) \tag{3-107}$$

$$F_0 = 1 + \frac{C_{ui*} + C_{uu*}Y_{0n}^{*}}{kT} \tag{3-108}$$

$$R_u = C_{uu*} \tag{3-109}$$

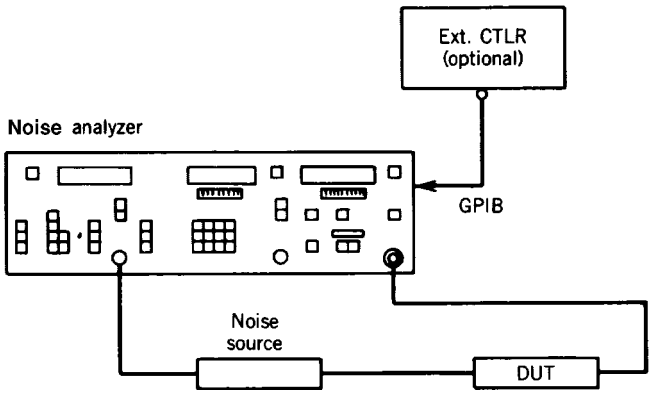


Figure 3-27 Noise figure measurement.

Noise Figure Test Equipment. Figure 3-27 shows the block diagram of a noise test setup. It includes the noise source and the other components. The metering unit has a special detector that is linear and over a certain dynamic range measures linear power. The tunable receiver covers a wide frequency range (e.g., 10–1800 MHz) and controls the noise source. The receiver is a double-conversion superheterodyne configuration with sufficient image rejection to avoid double-sideband noise measurements that would give the wrong results.

These receivers are microprocessor controlled and the measurement is a two-step procedure. The first is a calibration step that measures the noise figure of the receiver system and a reference power level. Then the device under test (DUT) is inserted and the system noise figure and total output power are measured. The noise factor is calculated by

$$F_1 = F_{\text{system}} - \frac{F_2 - 1}{G_1} \tag{3-110}$$

and the gain is given by the change in output power from the reference level [16]. The noise of the system is calculated by measuring the total noise power with the noise source on and off. With the excess noise ratio (ENR) known [16],

$$F_{\text{system}} = \frac{\text{ENR}}{Y - 1} \tag{3-111}$$

The noise bandwidth is usually set by the bandwidth of the receiver, which is assumed to be constant over the linear range. The ENR of the noise source is given by

$$\text{ENR} = \frac{T_{\text{hot}}}{T_{\text{cold}}} - 1 \tag{3-112}$$

where T_{cold} is usually room temperature (290 K). This ENR number is about 15 dB for noise sources with a 15-dB pad and 5 dB for noise sources with a 25-dB pad. Since both gain and noise were stored in the initial calibration, a noise/gain sweep can be performed.

For frequencies above 1800 MHz, we can extend the range with the help of the external signal generators, as shown in Figure 3-28. As shown, a filter ahead of the external mixer reduces the noise energy in the image band. If the DUT has a very broad frequency range

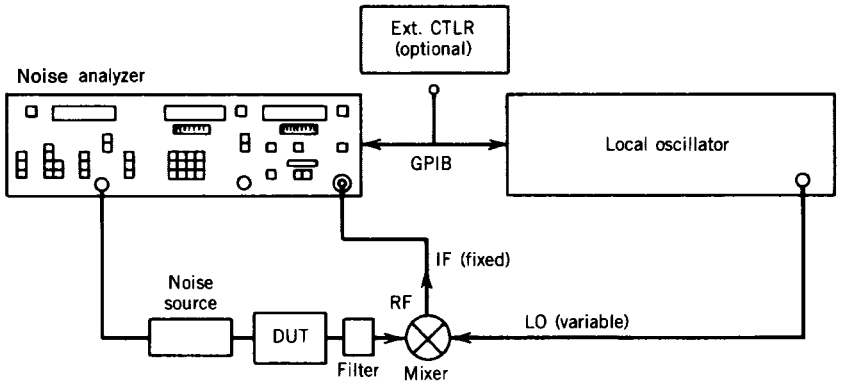


Figure 3-28 Single-sideband (SSB) noise figure measurements using an external mixer.

and has flat gain and noise over that range, a double-sideband (DSB) measurement is possible, with the image-rejection filter removed. However, a single-sideband (SSB) measurement is always more accurate [17].

How To Determine the Noise Parameters. The noise figure of a linear two-port network as a function of source admittance may be represented by

$$F = F_{\min} + \frac{R_n}{G_G}\left[(G_{0n} - G_G)^2 + (B_{0n} - B_G)^2\right] \tag{3-113}$$

where $G_G + jB_G$ = generator admittance presented to the input of the two-port
$G_{0n} + jB_{0n}$ = generator admittance at which optimum noise figure occurs
R_n = empirical constant relating sensitivity of the noise figure to generator admittance, with dimensions of resistance

It may be noted that for an arbitrary noise figure measurement with a known generator admittance, Eq. (3-113) has four unknowns—$F_{\min}$, R_n, G_{0n}, and B_{0n}. By choosing four known values of generator admittance, a set of four linear equations is formed and solutions of the four unknowns can be found [18, 19]. Eq. (3-113) may be transformed to

$$F = F_{\min} + \frac{R_n|Y_{0n}|^2}{G_G} - 2R_nG_{0n} + \frac{R_n|Y_G|^2}{G_G} - 2R_nB_{0n}\frac{B_G}{G_G} \tag{3-114}$$

or

$$F = F_{\min} + \frac{R_n}{G_G}|Y_G - Y_{0n}|^2 \tag{3-115}$$

Let

$$X_1 = F_{\min} - 2R_nG_{0n}$$

$$X_2 = R_n|Y_{0n}|^2$$

$$X_3 = R_n$$

$$X_4 = R_nB_{0n}$$

Then the generalized equation may be written

$$F_i = X_1 + \frac{1}{G_{si}}X_2 + \frac{|Y_{si}|^2}{G_{si}}X_3 - 2\frac{G_{si}}{B_{si}}X_4 \tag{3-116}$$

or, in matrix form,

$$[F] = [A]\,[X] \tag{3-117}$$

and the solution becomes

$$[X] = [A]^{-1}\,[F] \tag{3-118}$$

These parameters completely characterize the noise behavior of the linear two-port network. Direct measurement of these noise parameters by this method would be possible only if the receiver on the output of the two-port were noiseless and insensitive to its input admittance. In actual practice, the receiver itself behaves as a noisy two-port network and can be characterized in the same manner. What is actually being measured is the system noise figure of the two-port and the receiver. Thus, it becomes apparent that to do a complete two-port noise characterization, the gain of the two-port must be measured [20]. In addition, any losses in the input-matching networks must be carefully accounted for, because they add directly to the measured noise figure reading [21].

3-1-4 Linearity

Consistent with the elaborate discussion of linearity questions above, here is a quick refresher on what must be considered in amplifiers.

Dynamic Range, Compression, and IMD. An amplifier's dynamic range is related to the minimum discernible signal and the 1-dB compression point according to the equation

$$DR_n = \frac{(n-1)(IP_{n(\text{in})} - MDS)}{n} \tag{3-119}$$

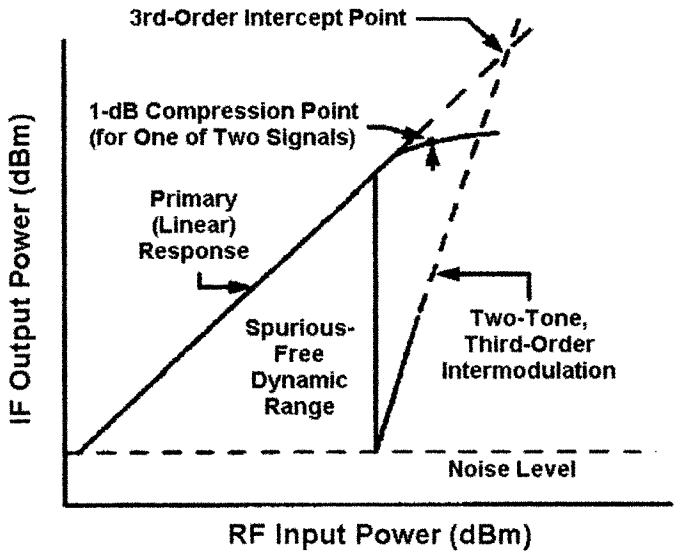


Figure 3-29 Amplifier linearity evaluation, including compression and two-tone IMD dynamic range. $P_{-1\text{dB}}$ for a single-tone cannot be read directly from this graph because the values shown are the result of two equal-power tones.

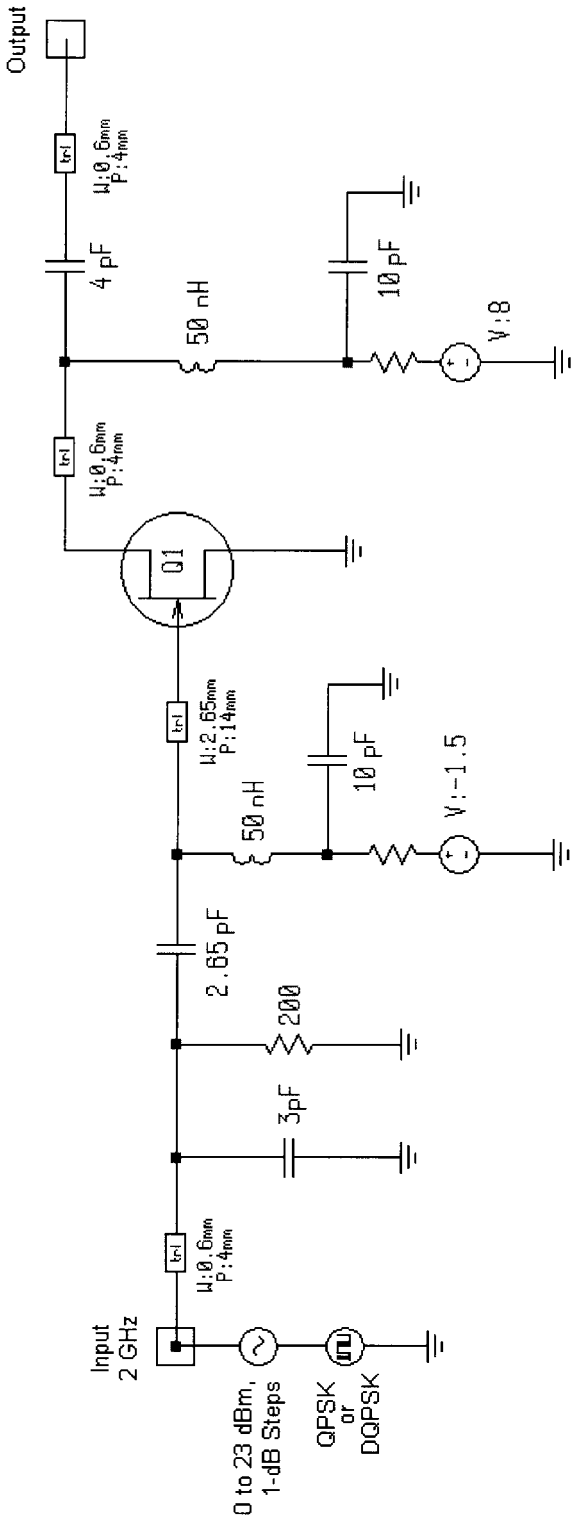


Figure 3-30 The power amplifier schematic.

where DR is the dynamic range in decibels, n is the order, $IP_{(in)}$ is the input intercept power in dBm, and MDS is the minimum detectable signal power in dBm. Once the 1-dB compression point is known, it is a fair assumption that the third-order intercept point, expressed in dBm, is 10 dB above this.

The intermodulation distortion gives the quality of an amplifier to withstand multiple signals without generating large intermodulation products or when using amplitude-modulated signals (cross-modulation). Figure 3-29 provides a quick reminder of IMD issues.

In some amplifiers, we would like to have control over the gain. While the better way to control gain while obtaining high linearity is a PIN-diode attenuator, special AGC circuits have become part of the amplifiers. Figure 3-48 shows a circuit of the SL610, made by Plessey, which has an AGC input provision. The noise figure of the PIN-diode attenuator is essentially equal to its attenuation. The AGC in multistage amplifiers is distributed over several stages and there is a lesser correlation between noise and AGC than in purely passive attenuators.

Special Case of Linearity Requirements for Digital Modulation. For cellular telephone systems we use digital modulation formats, such as QPSK and π/4-DQPSK, that combine phase and amplitude modulation. Amplifiers handling such signals must be carefully characterized and designed if adequate amplitude and phase linearity are to be maintained.

Examples of Power Amplifiers: Looking into the Effects of Distortion. This example presents modulation analysis using QPSK in a 2-GHz power amplifier (Figure 3-30). The simple amplifier was designed using the electronic Smith Tool of Serenade to determine the matching network for a narrowband design. The example demonstrates setup of modulation analysis and available results.

Siemens among others provides integrated amplifiers based on the concept depicted in Figure 3-30, the most popular ones being the CGY96 and CGY94. For the purpose of measurements, these amplifiers were actually put on a breadboard. Figure 3-31 shows the big brother, the CGY96. It is a GaAs MMIC intended as a power amplifier for GSM Class 4 phones with 3.2 W (35 dBm) at 3.5 V at an overall power added efficiency of 50%. Table 3-5 shows its electrical characteristics.

For the purpose of our experiments, we considered the CGY94 because it requires less external peripheral circuits. Table 3-6 shows its electrical characteristics. For those of us interested in seeing part of the internal circuit, Figure 3-32 shows the chip. This two-stage amplifier uses an interesting internal feedback scheme. Its schematic is shown in Figure 3-33.

For the purpose of looking at the waveforms, we will continue with the single-stage amplifier (Figure 3-30) and then compare the measured versus predicted performance of this CGY94-based amplifier.

A modulation source is connected to the RF source at the input port of the amplifier. The RF source specifies the carrier power (or voltage or current) and the modulation source specifies the modulation format and properties of the modulated signal.

A brief review of the modulation source used in this example indicates the following properties:

Property	Value	Description
NB	128	Number of bits
BR	1.2288E6	Bit rate
N	8	Number of samples per bit
BW1	590.E3	One-sided bandwidth of the main channel in the PIB or P1IB calculation
FS1	740.E3	Start baseband frequency of the fist adjacent channel
BW2	590.E3	One-sided bandwidth of the main channel in the P2IB calculation
FS2	1990.E3	Start baseband frequency of the second adjacent channel
BW3	590.E3	One-sided bandwidth of the main channel in the P3IB calculation
FS3	3240.E3	Start baseband frequency of the third adjacent channel
M	4	Order of the signal space; M = 4 is QPSK
DLY	0.0	Fractional bit delay
IASC	1.0	I channel amplitude scale; multiplier for the I waveform to model amplitude imbalance
QASC	1.0	Q channel amplitude scale; multiplier for the Q waveform to model amplitude imbalance
LPF	3	Baseband low-pass filter; 3 = root-raised cosine filter
LPFC	665.E3	–3 dB cutoff frequency of LPF
LPFN	3	Number of resonators if LPF = 1
LPFR	0.35	Roll-off factor if LPF = 3 or LPF = 4
FILE	N/A	File name of user-defined modulated signal file

The modulation source is a QPSK stream with a bit rate of 1.2288 Mbps. A total of 128 bits will be analyzed, and each bit is sampled eight times to construct an accurate analog waveform. The BW and FS parameters are defined here to specify the main (in-band) channel and adjacent channel for ACPR calculations. Imbalance of the I and Q amplitudes and phases can be described by the IASC, QASC, and DLY parameters, if desired. The modulation signal

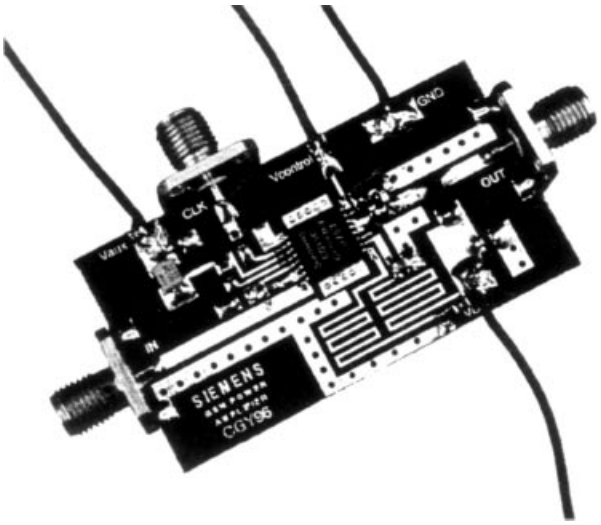


Figure 3-31 CGY96 GaAs MMIC power amplifier.

Table 3-5 CGY96 electrical characteristics

(T_A = 25 °C, V_{neg} = 5 V, $V_{control}$ = 2.2 V; duty cycle 12.5%, t_{on} = 577 µs)

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
Frequency range	f	880	—	915	MHz
Supply current P_{in} = 0 dBm	I_D	—	1.8	—	A
Supply current negative voltage generator V_{aux} = 3.5 V	I_{AUX}	—	10	—	mA
Gain (small signal)	G	—	40	—	dB
Power gain P_{in} = 0 dBm	G_P	—	35	—	dB
Output power P_{in} = 0 dBm, $V_{control}$ = 2.0–2.5 V	P_{OUT}	—	35	—	dBm
Overall power added efficiency P_{in} = 0 dBm	η	—	50	—	%
Dynamic range output power $V_{control}$ = 0.2–2.2 V		—	80	—	dB
Harmonics P_{in} = 0 dBm	$H(2f_0)$	—	–40	—	dBc
	$H(3f_0)$	—	–43	—	dBc
	$H(4f_0)$	—	–44	—	dBc
Noise power in RX (935–960 MHz) P_{in} = 0 dBm, P_{out} = 35 dBm, 100 kHz RBW	N_{RX}	—	–81	—	dBm
Stability *all spurious outputs* < –60 dBc, *VSWR load, all phase angles*		—	10:1	—	—
Input VSWR		—	1.7:1	—	—

Table 3-6 CGY94 electrical characteristics

T_A = 25 °C, f = 0.9 GHz, $Z_S = Z_L$ = 50 Ω, V_D = 3.6 V, V_G = 4 V, VTR pin connected to ground, unless otherwise specified; pulsed with a duty cycle of 10%, ton = 0.33 ms)

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
Supply current V_D = 3.0 V P_{in} = 10 dBm	I_{DD}	—	1.18	—	A
Negative supply current *(normal operation)*	I_G	—	2	—	mA
Shut-off current *VTR n.c.*	I_D	—	400	—	µA
Negative supply current (*shut off mode, VTR pin n.c.*)	I_G	—	10	—	µA
Gain P_{in} = *5* dBm	G	27.0	29.0	—	dB
Power gain V_D = 3.6 V; P_{in} = 10 dBm	G	22.8	23.6	—	dB
Output power V_D = 3.0 V*;* P_{in} = 10 dBm	P_O	31.5	32.3	—	dBm
Output power V_D = 3.6 V*;* P_{in} = 10 dBm	P_O	32.8	33.6	—	dBm
Output power V_D = 5 *V;* P_{in} = 10 dBm	P_O	34.5	35.5	—	dBm
Overall power added efficiency V_D = 3.0 V; P_{in} = 10 dBm	η	43	48	—	%
Overall power added efficiency V_D = 3.6 V; P_{in} = 10 dBm	η	42	47	—	%
Overall power added efficiency V_D = 5 V; P_{in} = 10 dBm	η	41	46	—	%
Harmonics (P_{in} = 10 dBm, *CW*) $2f_0$	—	—	–49	—	dBc
V_D = 3.6 V; (P_{out} = 33.1 dBm) $3f_0$	—	—	–45	—	dBc
Input VSWR V_D = 3.6 V	—	—	1.5:1	2.0:1	—

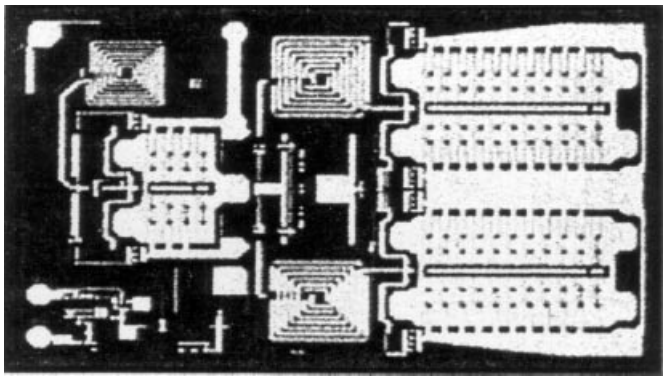


Figure 3-32 The CGY94 amplifier chip.

is often filtered, and several types of filters are available. Here a root-raised cosine filter is used with a cutoff frequency of 665 kHz and a roll-off factor of 0.35.

Analysis. Figure 3-34 shows the compression characteristics of the amplifier as the available RF source power (a sinusoid) is swept. Note that P_{-1dB} is 11.5 dBm referred to the input and 28.5 dBm referred to the output. We will look at modulation characteristics in the linear region of operation at a source power of 0 dBm, and the nonlinear region at a source power of 20 dBm.

Next, we will view the voltage across the transistor to examine its behavior under compression. Figure 3-35 shows the drain–source voltage of the FET as the source power is swept. It is clear from this graph that clipping due to pinchoff and forward-gate conduction

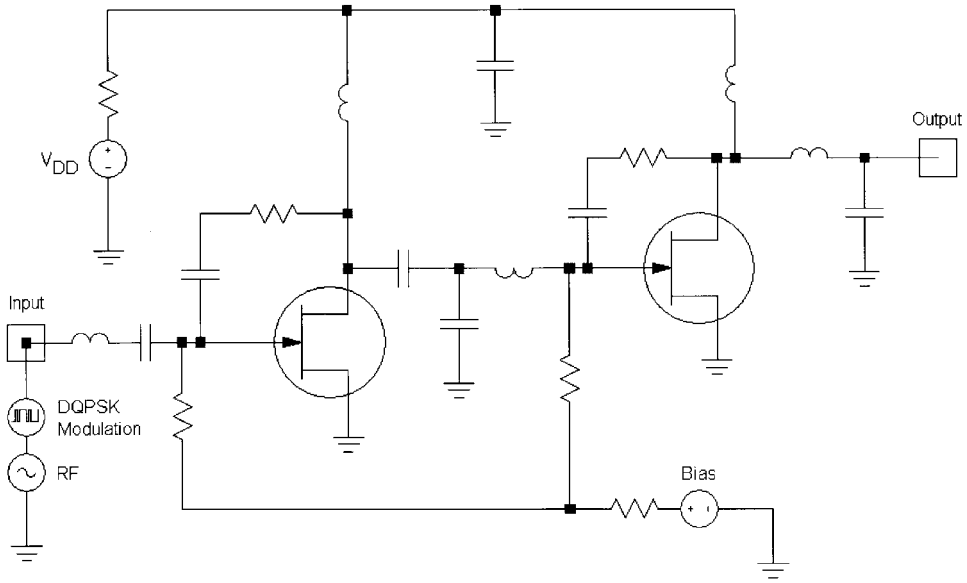


Figure 3-33 Schematic of the CGY94 GaAs MMIC power amplifier.

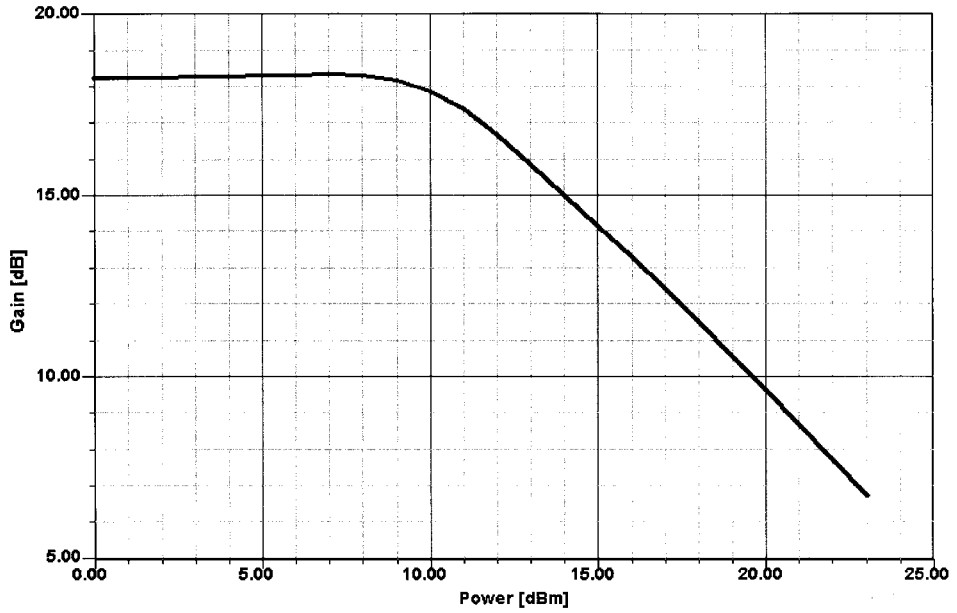


Figure 3-34 Single-tone RF power sweep analysis of the FET amplifier.

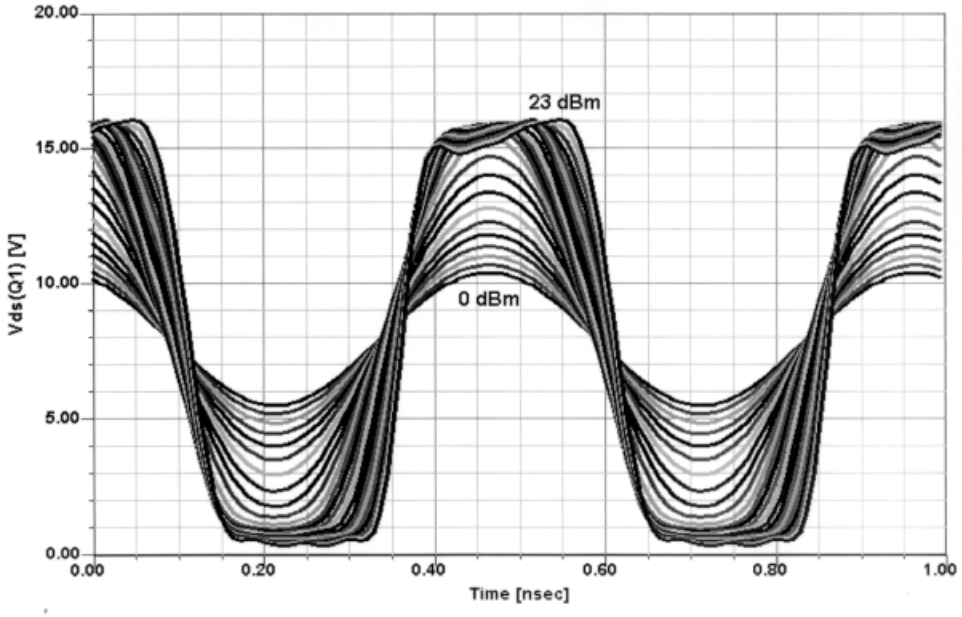


Figure 3-35 V_{DS} of the FET versus RF source power.

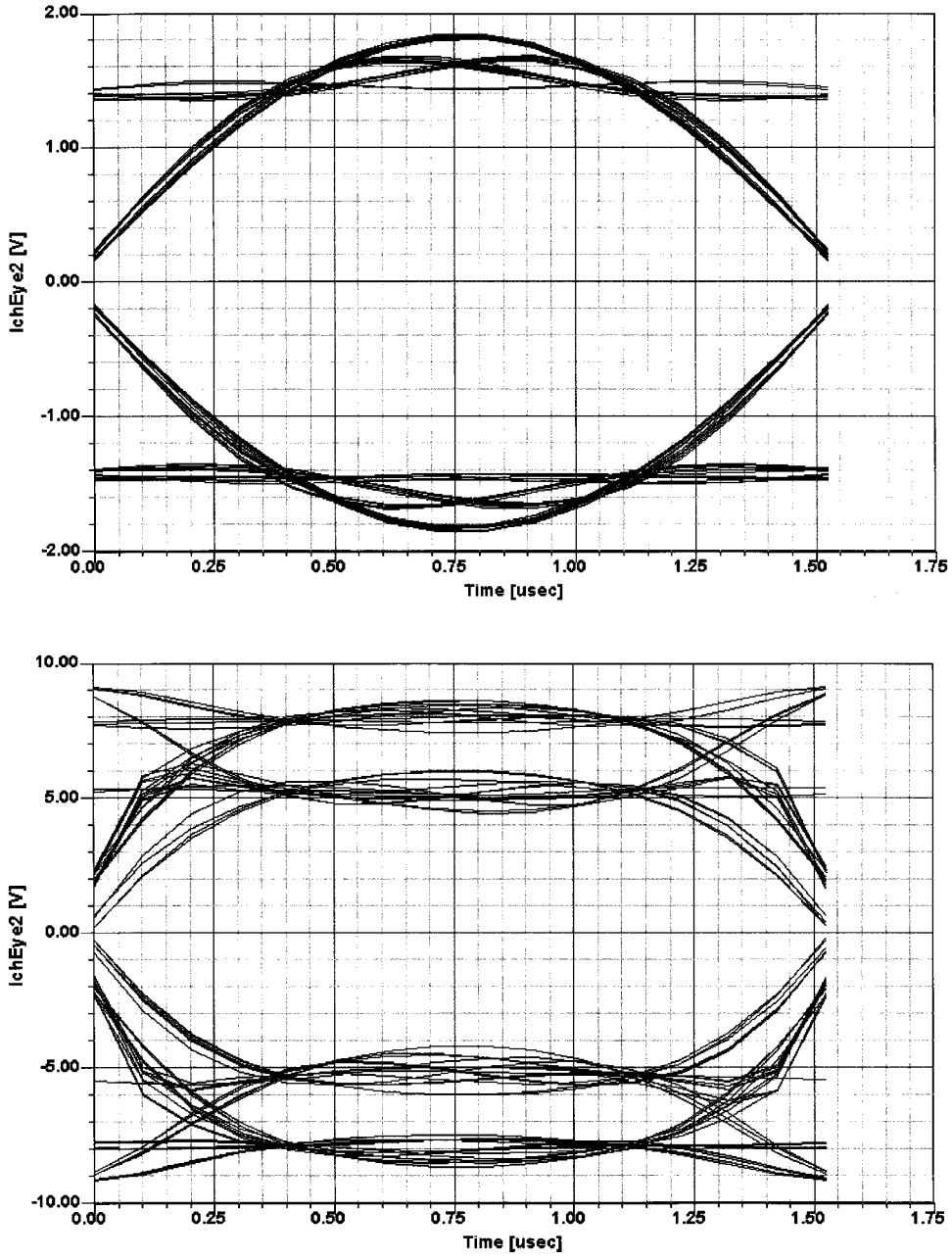


Figure 3-36 Eye diagrams of the I channel for source powers of (a) 0 dBm and (b) 20 dBm.

is limiting the performance of the device. A phase shift of the voltage waveform as power is increased is also apparent. The combination of the power compression and phase shift, or AM to PM conversion, will be used in the modulation analysis to determine the overall distortion of the modulated signal.

The output available from modulation analysis includes:

- I and Q channel waveforms
- Eye and constellation diagrams
- Spectral plots
- ACPR, in-band, and adjacent power

Figure 3-36 shows the eye diagrams for the 128-bit QPSK signal at two RF source powers, 0 and 20 dBm. Operating linearly at 0 dBm input, the amplifier doesn't distort the eye and it remains wide open. Note that in this example the filter bandwidth doesn't produce any intersymbol interference, as evidenced by the open eye. At an input power of 20 dBm, the amplifier compresses the signal, and distortion in the eye is evident.

We can also view the modulation spectrum at these two power levels to investigate the intermodulation distortion and spectral regrowth that takes place. Figure 3-37 shows the spectral plots. The lower trace corresponds to 0 dBm source power and shows almost no regrowth, while the upper trace at 20 dBm source power shows considerable regrowth.

The uncorrupted modulation source can also be shown. This corresponds to the source as a pure voltage without any circuit influence.

Now we evaluate the CGY94 board in a similar fashion. Figure 3-38 shows the waveform used for CAD purposes. Needless to say, such a signal cannot be produced by any signal

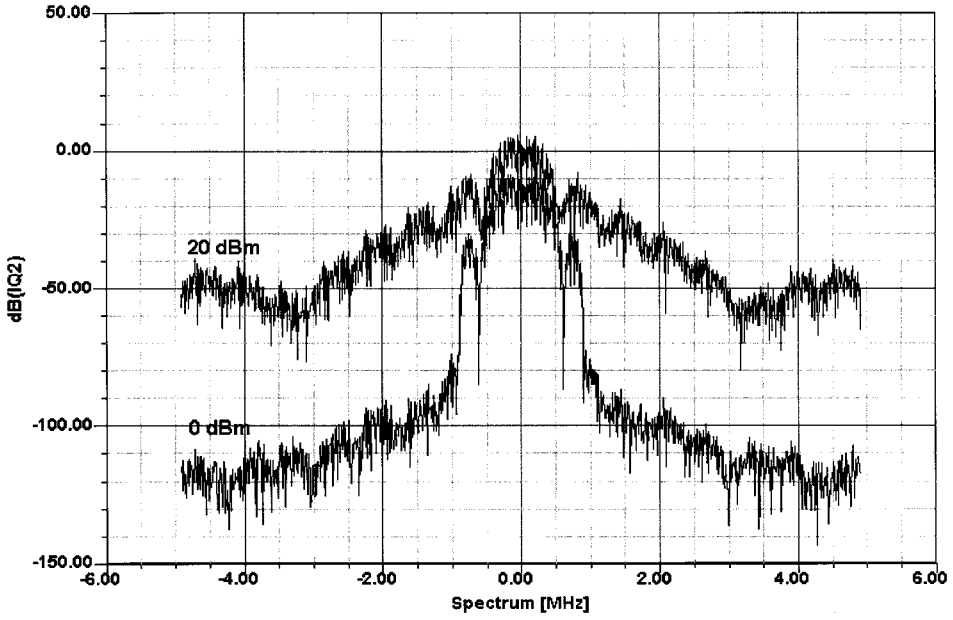


Figure 3-37 Spectrum of the modulation signal at the amplifier output. Source powers are 0 and 20 dBm.

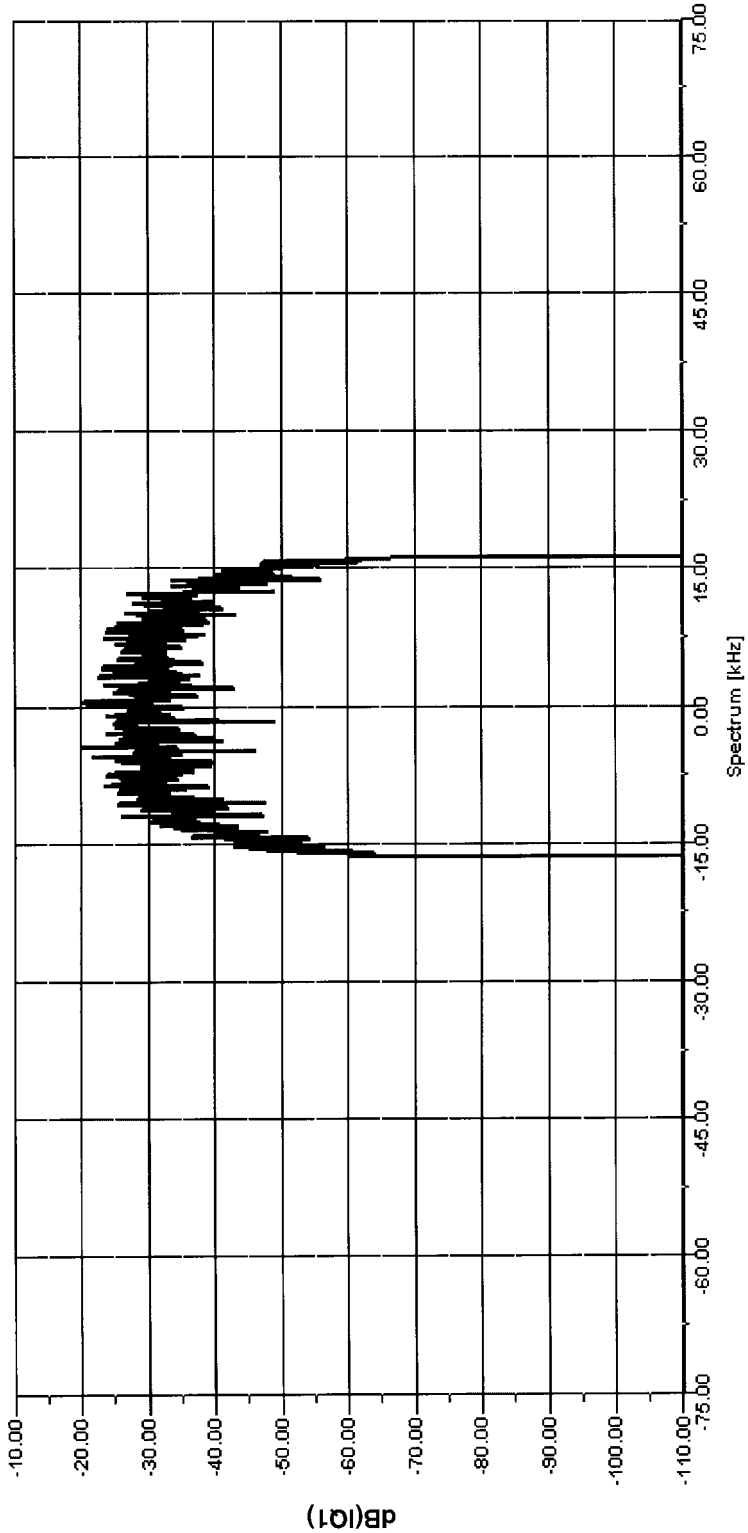


Figure 3-38 Simulated signal used for CAD analysis of the CGY94 amplifier.

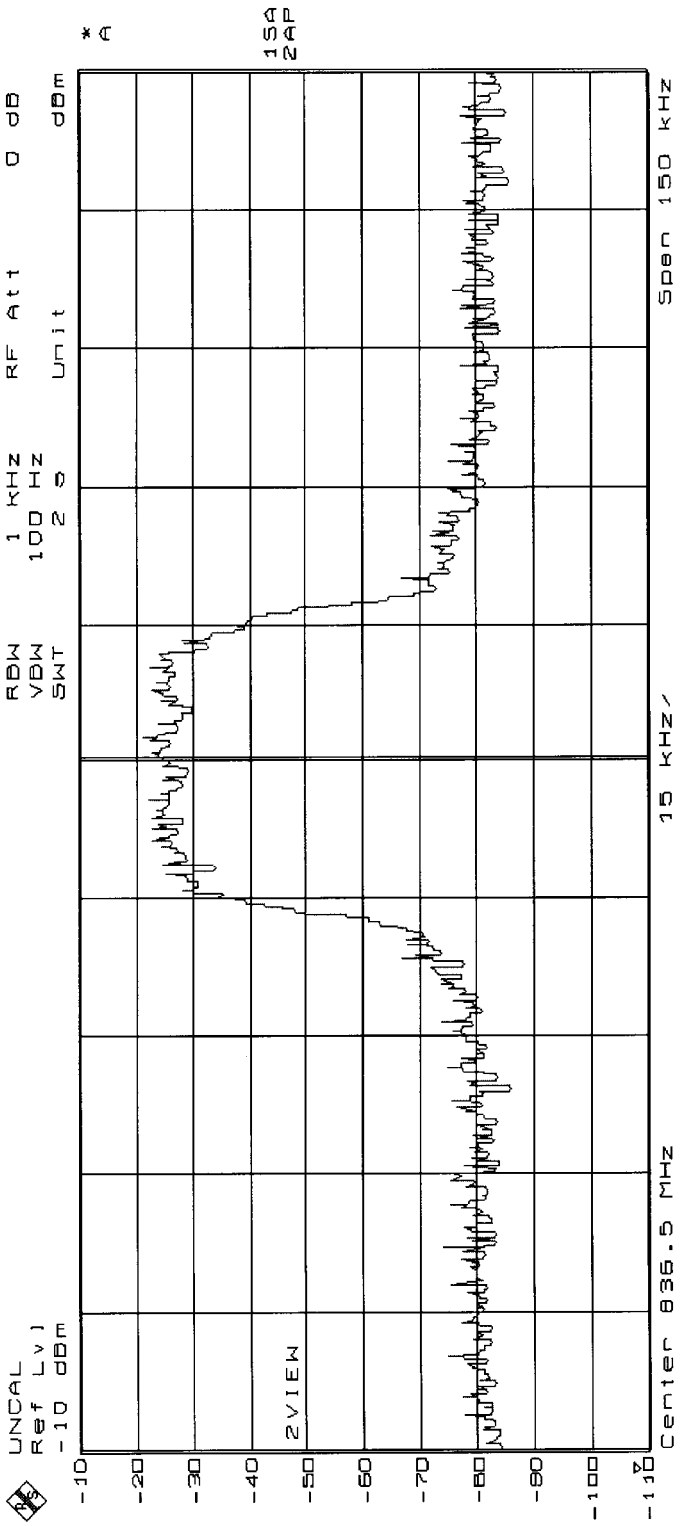


Figure 3-39 Real signal used for testing the actual amplifier.

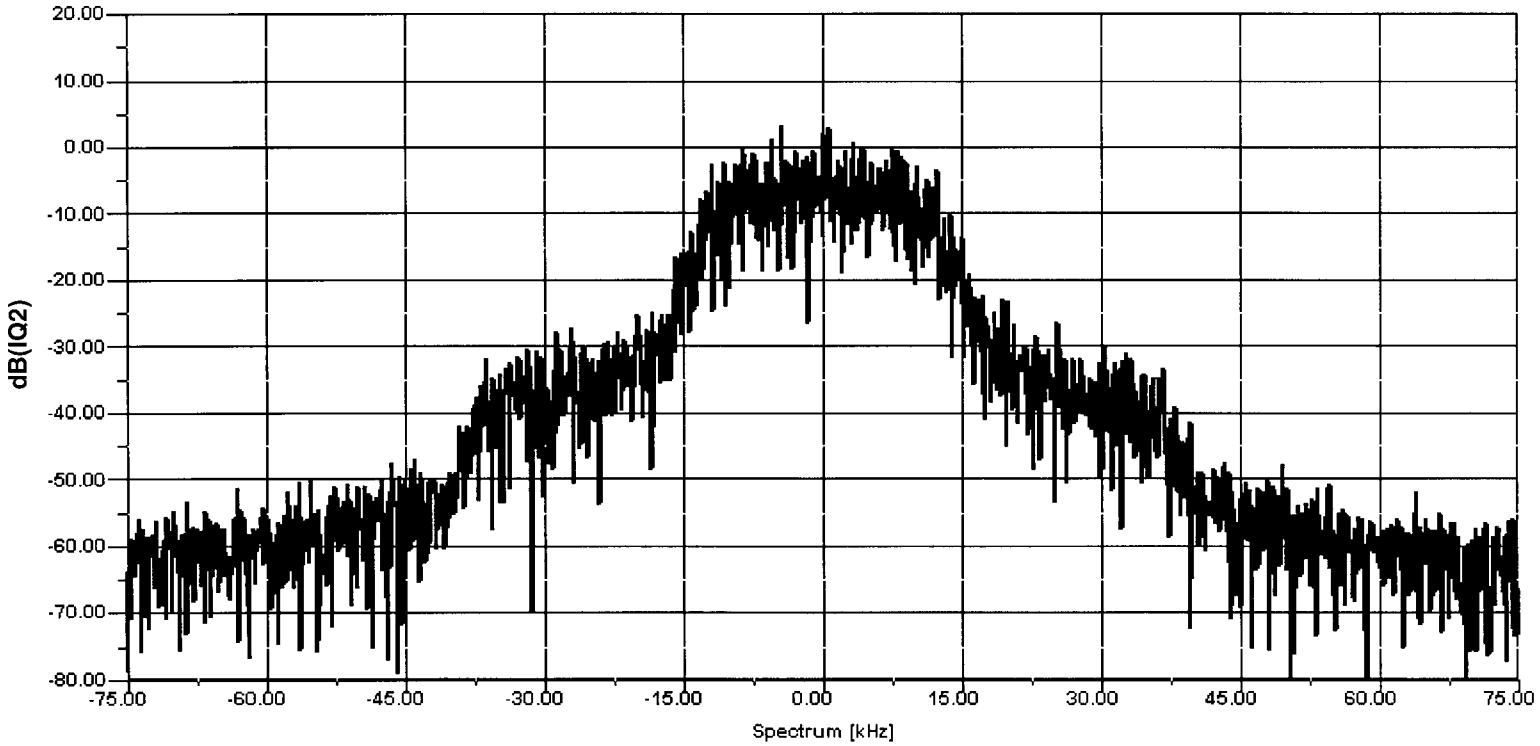


Figure 3-40 Spectral regrowth predicted by the simulation.

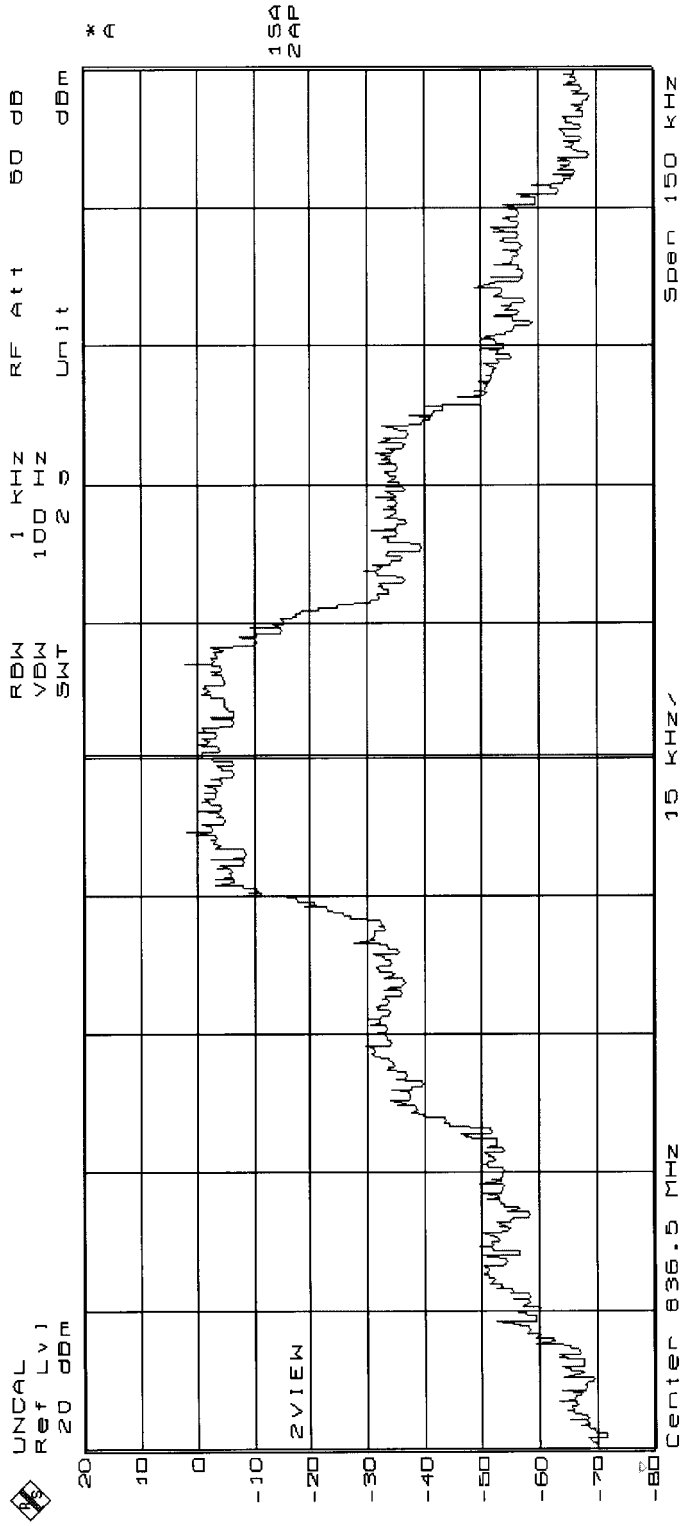


Figure 3-41 Measured spectral regrowth of the CGY94 amplifier.

generator. The one generated by an actual signal generator is shown in Figure 3-39. The wideband noise is around –80 dBm and the input level is about –25 dBm, consistent with the simulation signal. Looking at the output, Figure 3-40 shows the simulated signal with the expected increase in bandwidth. This is due to the nonlinearity of the amplifier and is consistent with Figure 3-37 for the single-stage amplifier. The actual measured output for the CGY94 shown in Figure 3-41 is quite close to its simulated output, indicating that the mathematical approach used for simulating is correct. This is only one example; in reality, many more waveforms can be analyzed and predicted using this method. Figure 3-41 shows essentially three distinct steps, which are sufficiently close to the response shown in the simulation. A dynamic range of 70 dB, while displayed in Figures 3-40 and 3-41, is not really necessary for good-quality transmission, and the discrepancy between the two pictures is in part due to the limited modeling quality of the active device as provided by the company that did the parameter extraction for this transistor.

Another useful view to aid in the understanding of the compression and distortion within the amplifier is to look at the magnitude of the complex waveform. Since the signal is composed of the in-phase and quadrature-phase components it is represented as

$$s(t) = i(t) + j \cdot q(t) \tag{3-120}$$

The time-domain magnitude is written

$$|s(t)| = \sqrt{i(t)^2 + q(t)^2} \tag{3-121}$$

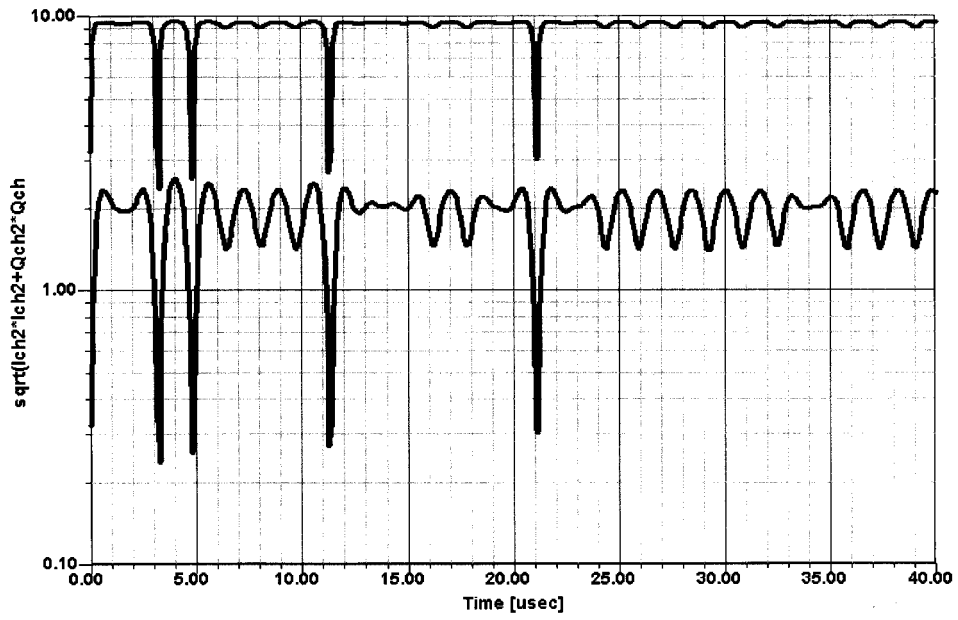


Figure 3-42 Time-domain magnitude of the complex modulation signal at source powers of 0 and 20 dBm.

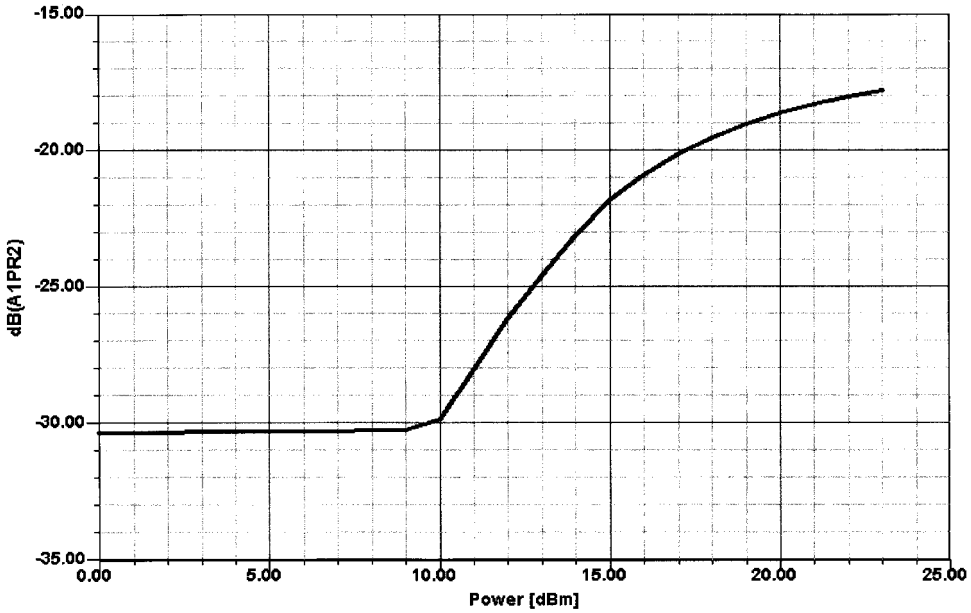


Figure 3-43 Adjacent-channel power ratio as function of RF source power.

Figure 3-42 shows the magnitude at source powers of 0 and 20 dBm. The *X* axis has been rescaled to 40 μs for improved viewing. Note the significant compression of the 20-dBm waveform at the higher signal levels. It is clear from this view that the signal is severely distorted.

We can also compute the adjacent-channel power ratio (ACPR) against the RF power sweep. ACPR is the ratio of the adjacent-channel power to the in-band channel power. The bandwidth (BW) and adjacent-channel start frequency (FS) are used for this calculation. For accurate ACPR calculation, a large number of bits are needed. This example uses 128 bits so the computation time is short, but 512 or more bits should be used for a more accurate computation.

Figure 3-43 shows the ACPR as a function of RF source power. The ACPR is nearly constant (and nonzero due to the gradual skirt of the baseband filter and inherent spillover to the adjacent channel) up to the P_{-1dB} compression point . As the intermodulation products spill into the adjacent channel and spectral regrowth occurs, the ACPR degrades.

π/4-DQPSK Circuit Analysis. Next, we will examine the same amplifier using a π/4-DQPSK modulation source. The task is identical to the previous one except for the source and the number of bits. The number of bits for this project has been increased from 128 to 256. Figure 3-44 shows the eye diagrams from this modulation format, at source powers of 0 and 20 dBm. The distortion and related intersymbol interference are clearly evident.

The constellation plot at 0 dBm is shown in Figure 3-45.

It may be interesting to take a look at the different waveforms currently in use. Figure 3-46 shows a spectrum analyzer picture of the most popular ones.

For reasons of time-division management, the European standard, GSM, which is now offered by Omnipoint at the 1.8-GHz band in the United States, uses gating to reduce the actual bandwidth of the spectrum. Figure 3-47 compares the GSM signal with and without gating.

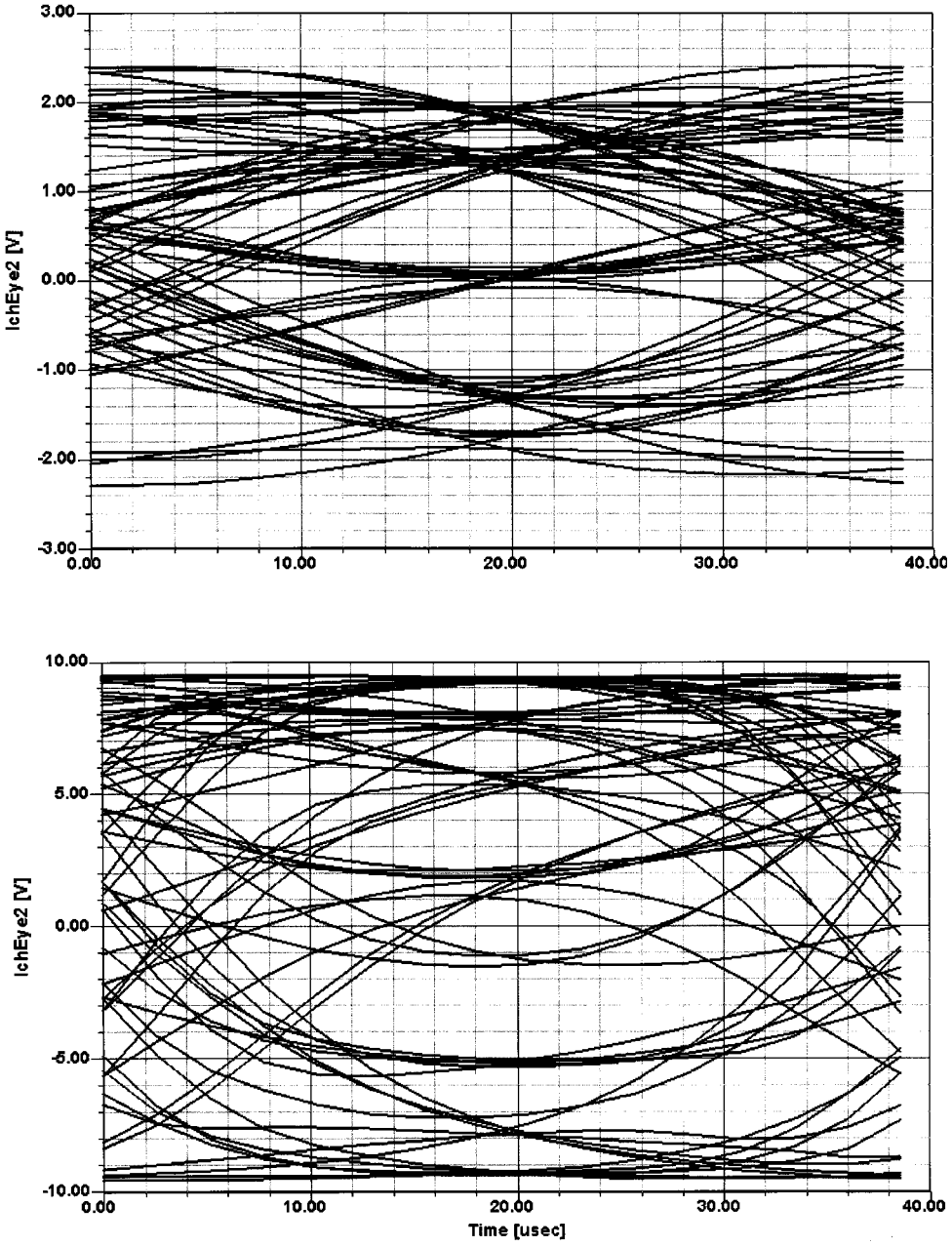


Figure 3-44 Eye diagrams for $\pi/4$-DQPSK at source powers of (a) 0 dBm and (b) 20 dBm. The distortion and related intersymbol interference are clearly visible.

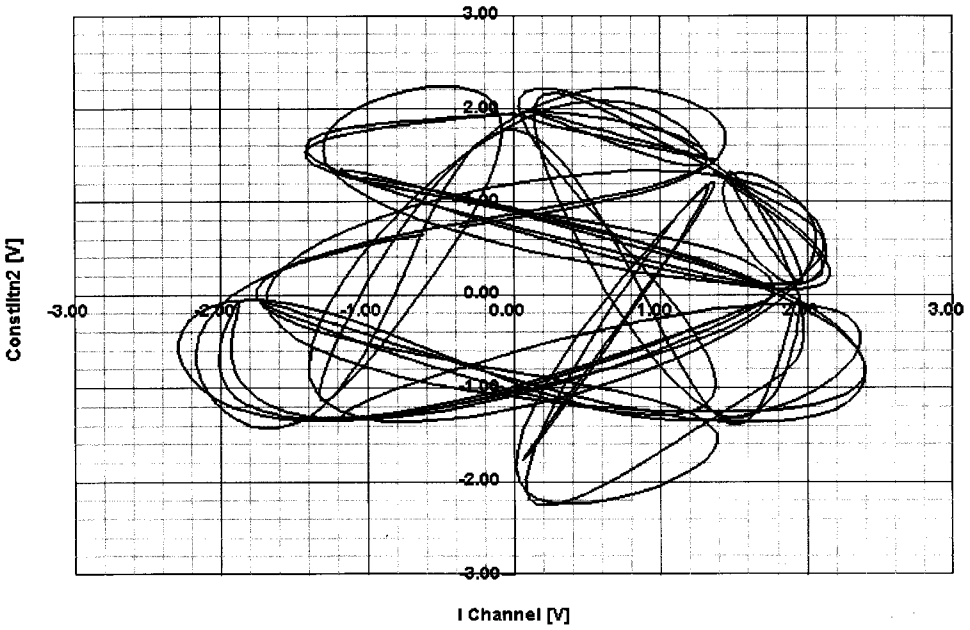


Figure 3-45 Constellation diagram for π/4-DQPSK at Port 2 and 0-dBm source power.

3-1-5 AGC

Many wireless systems consist of an up conversion or down conversion using an intermediate frequency, typically 10% or less of the receiving frequency (the FM broadcast IF has been set at 10.7 MHz for the same reason).

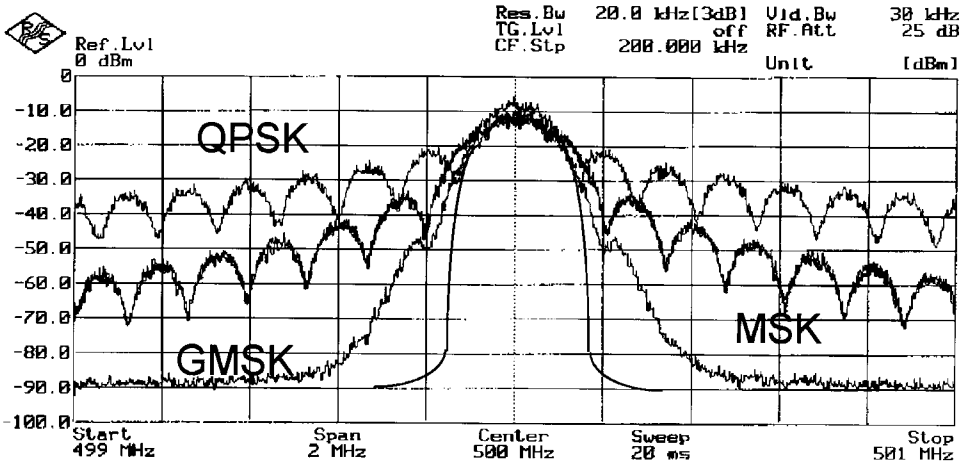


Figure 3-46 Spectrogram showing the characteristics of MSK, GMSK, and QPSK.

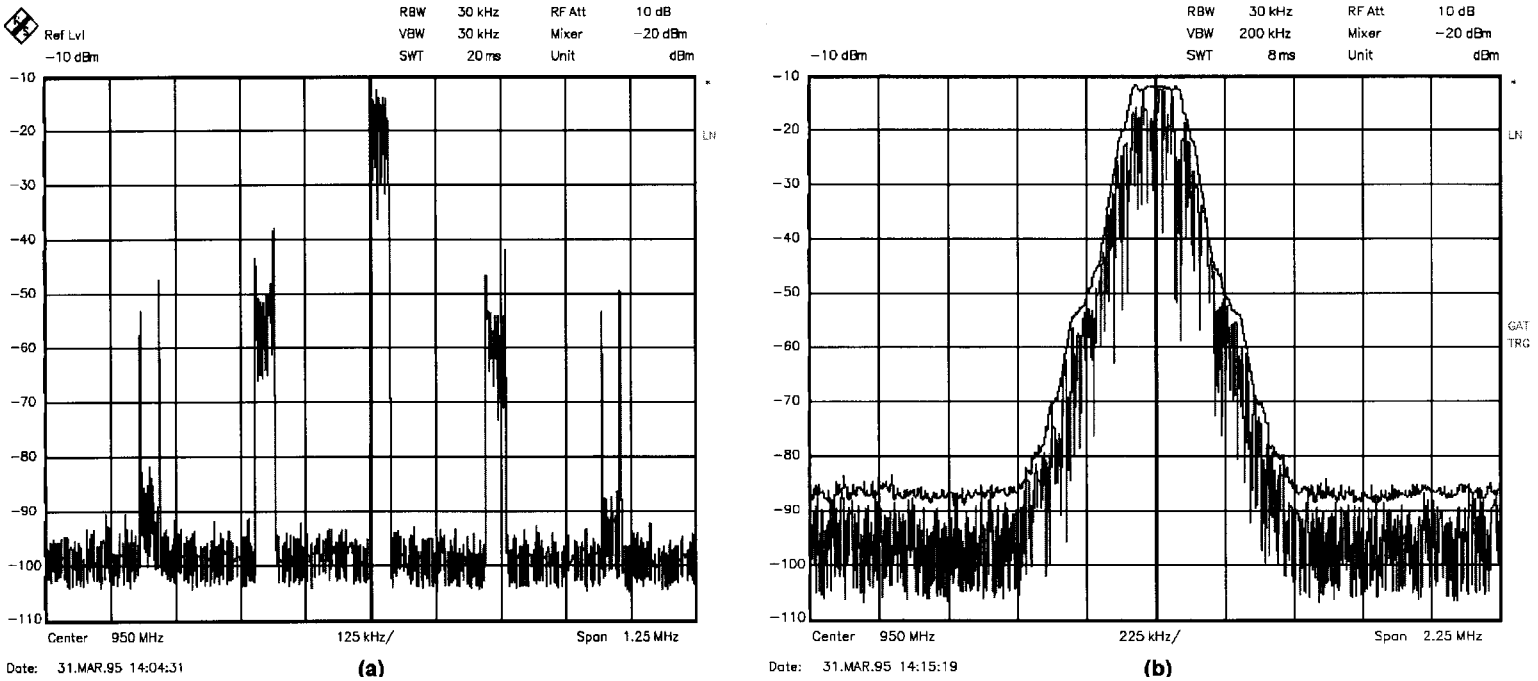


Figure 3-47 Pulsed GSM signal (a) without gating and (b) with gating.

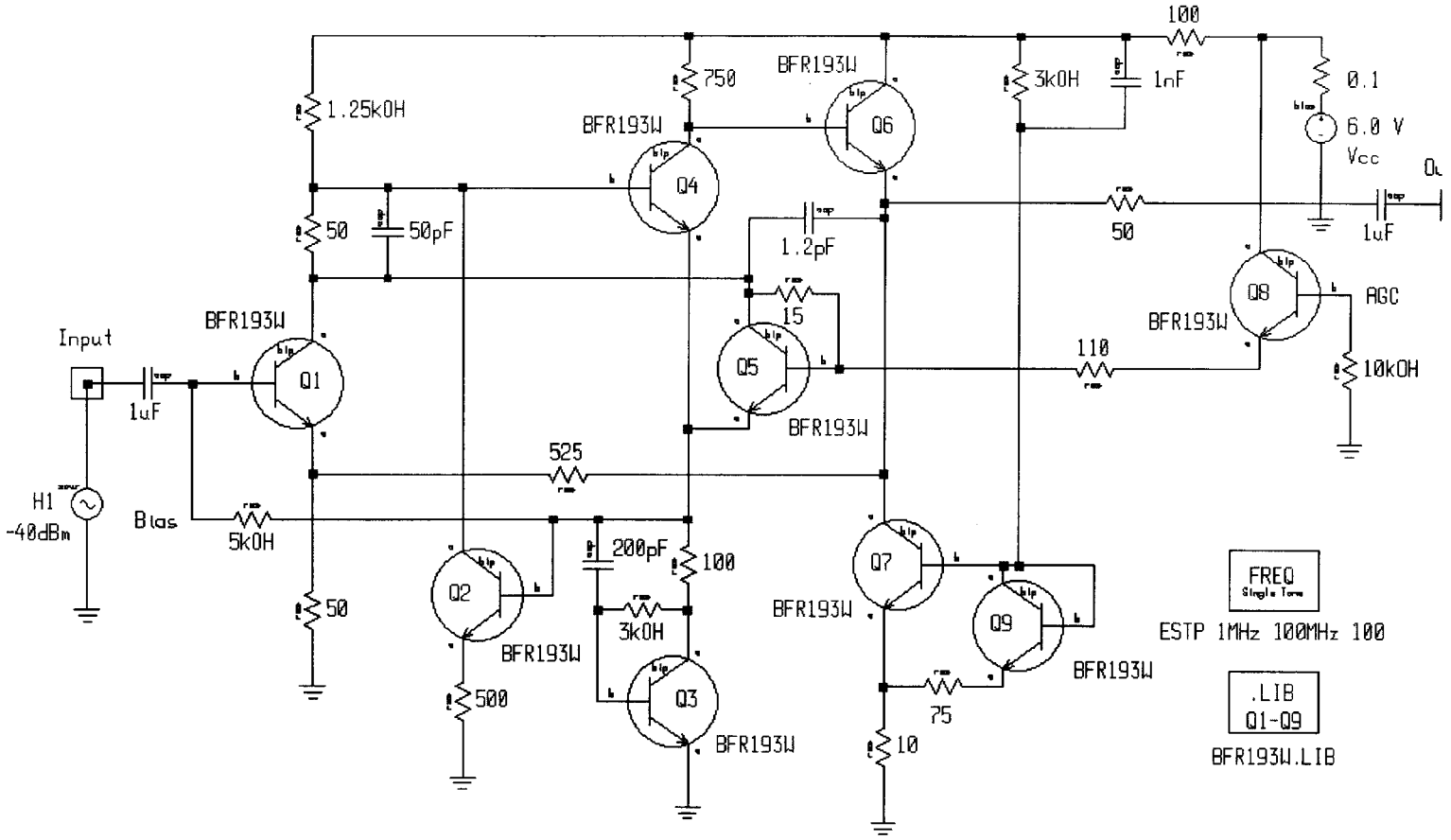


Figure 3-48 Schematic of the Plessey SL610 wideband amplifier with AGC. The 75-Ω resistor in the emitter of Q9 had to be added to obtain the appropriate dc current.

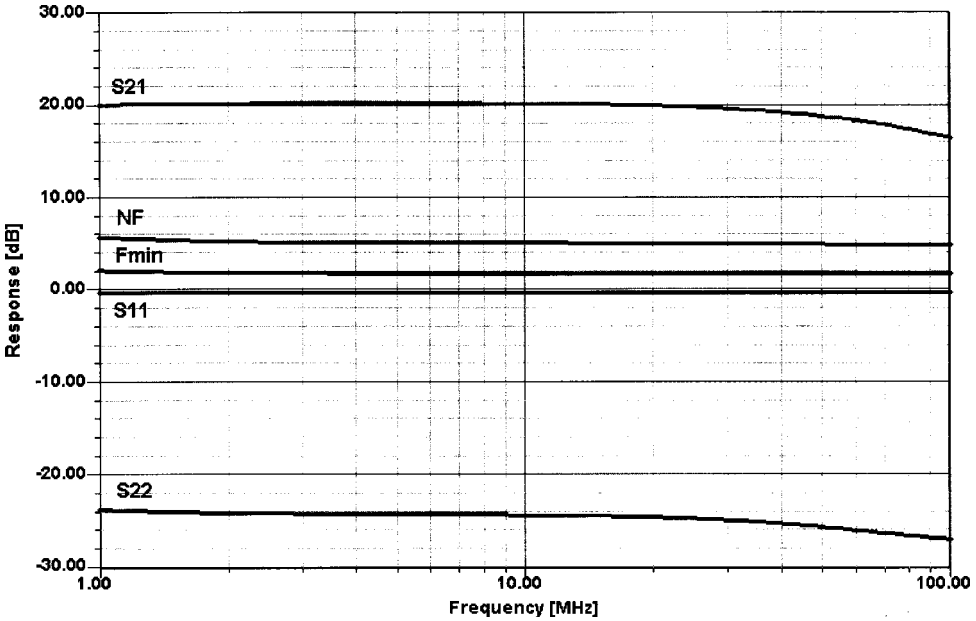


Figure 3-49 Simulated frequency response of the Plessey SL610 wideband amplifier IC with AGC = 0 V.

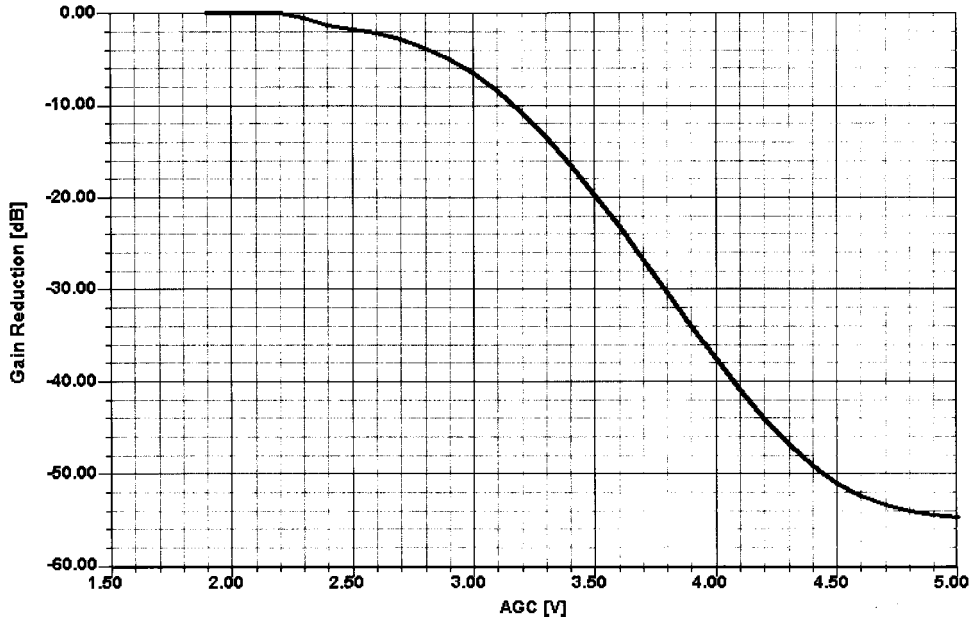


Figure 3-50 Simulated gain reduction versus AGC of the Plessey SL610 wideband amplifier IC.

Table 3-7 CGY121A features

- Variable gain amplifier (MMIC amplifier) for mobile communication
- Typical gain control range over 50 dB
- Positive control voltage
- 50-Ω input and output matched
- Low power consumption
- Operating voltage range: 2.7–6 V
- Frequency range 800 MHz to 2.5 GHz

We therefore will show two amplifiers with AGC and available data. The first one is a wideband amplifier type SL610, which has been manufactured by Plessey; its current fate is unknown. However, since we put the circuit into the simulator it provides significant insight into the operation of such AGC stages. Figure 3-48 shows the actual schematic of the SL610 entered in Ansoft's Serenade Design Environment for simulation.

The AGC action is derived from Q8 through Q5, which, together with Q4, forms a differential amplifier. Each consecutive stage (Q1, Q4, Q6) has increasing current, totaling about 20 mA; the last transistor takes about 13 mA. Based on the feedback surrounding the transistors, the actual transistor characteristics are less important. We have shown this by putting much higher f_T devices into the circuit than were available for the original design (20 years ago). The circuit is still interesting today because of its good large-signal-handling capability.

Table 3-8 CGY121A electrical characteristics at 900 MHz and 1.8 GHz

Characteristics	Symbol	Minimum	Typical	Maximum	Unit
(T_A 25 °C, f = 900 MHz, V_g = –4 V, $RS = RL$ = 50 Ω unless otherwise specified)					
Power gain V_d = 3 V; I = 45 mA; V_{con} = 3 V	G	17	19	—	dB
Input return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{in}	—	11	—	dB
Output return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{out}	—	10	—	dB
Gain control range V_{con} = 3 V ... 0 V; V_d = 3 V; I = 45 mA	dG	48	53	—	dB
1-dB gain compression V_d = 3 V; I = 45 mA; V_{con} = 3 V	P_{1dB}	—	14	—	dBm
(T_A 25 °C, f = 1800 MHz, V_g = –4 V, $RS = RL$ = 50 Ω unless otherwise specified)					
Power gain V_d = 3 V; I = 45 mA; V_{con} = 3 V	G	15.5	17.5	—	dB
Input return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{in}	—	10	—	dB
Output return loss V_d = 3 V; I = 45 mA; V_{con} = 3 V	RL_{out}	—	8.5	—	dB
Gain control range V_{con} = 3 V ... 0 V; V_d = 3 V; I = 45 mA	dG	48	53	—	dB
1-dB gain compression V_d = 3 V; I = 45 mA; V_{con} = 3 V	P_{1dB}	—	14	—	dBm

Figure 3-49 shows the amplifier's frequency response which is extremely close to the measured data. The good simulation also can be seen from comparing the published AGC curve with the one we found in the simulation. The bias of the first transistor remains essentially constant, resulting in a noise figure more independent of the AGC than the PIN-diode attenuator, and also stabilizes the input impedance as a function of AGC. Figure 3-50 shows the SL610's predicted AGC response.

The same requirement is applicable for the wireless frequency range. Tables 3-7 and 3-8 and Figures 3-51 through 3-54 present specifications for the recently introduced CGY121A GaAs MMIC by Infineon Technologies AG, formerly the Semiconductor Group part of Siemens AG.

3-1-6 Bias and Power Voltage and Current (Power Consumption)

Biasing of transistors operating in Class A or B follows some standard rules. The only deviation from this tends to be when the operating voltage is very low, below 3 V, because then the luxury of using voltage drops for good dc stability as a function of temperature disappears. (We will discuss low-voltage design in detail in Section 3-2-6.) In our simple example (Figure 3-55), we are operating from a 12-V power supply, and the first rule of thumb is to make sure that the voltage between the emitter and ground is at least 0.7 V. This is necessary to compensate the base–emitter junction threshold voltage, which decreases by about 2 mV/°C as the temperature increases. A larger voltage drop in the emitter–ground connection greatly reduces the influence. In this amplifier the emitter–ground path consists of an emitter unbypassed resistor of 10 Ω in series with a bypassed 501-Ω resistor, resulting in an overall voltage drop of 2.6. We could reduce the supply voltage by about 1.9 V at the collector side and still maintain a safety margin of 0.7 V at the emitter–ground connection. The current 2.6 V, so to speak, is overkill.

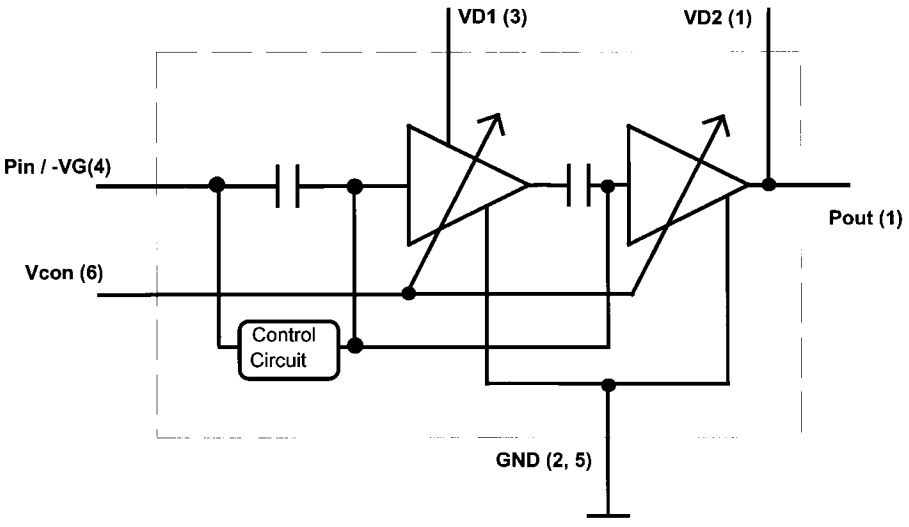
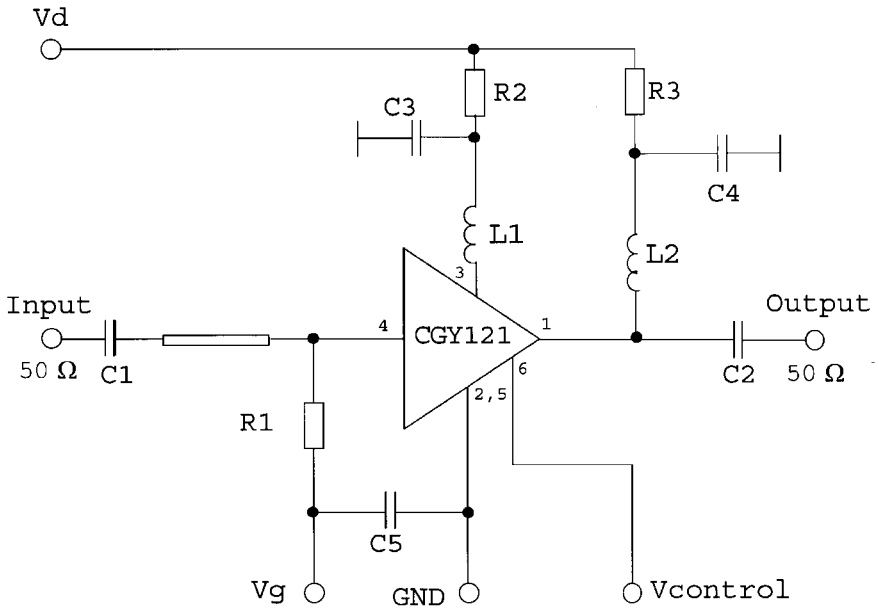


Figure 3-51 Functional block diagram of the gain-controllable CGY121A GaAs MMIC. Gain control is achieved by applying 0–3 V dc to the Vcon pin.



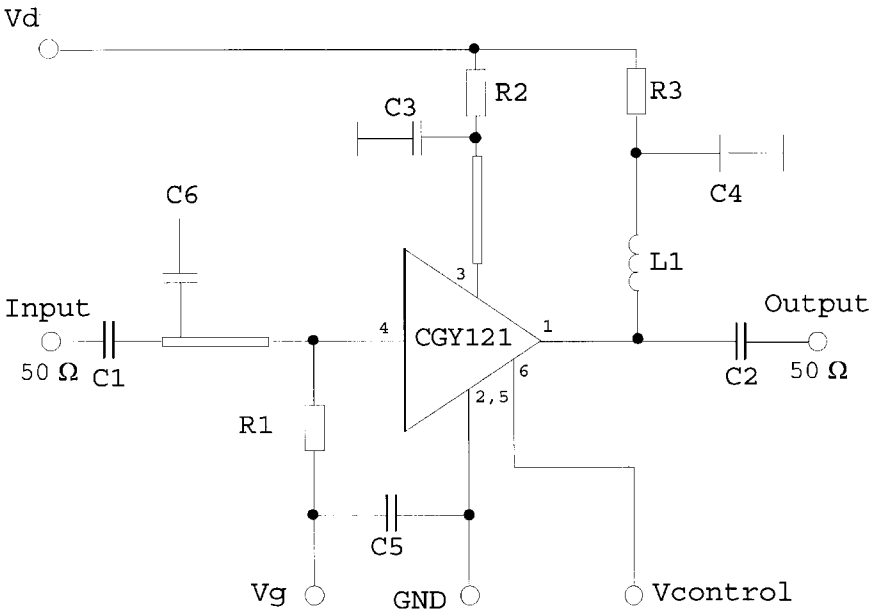
Parts List

Frequency	900 MHz
C1, C2 (Siemens Size 0603)	22 pF
C3, C4 (Siemens Size 0603)	100 nF
C5 (Siemens Size 0603)	47 nF
L1 (Coilcraft 0805CS-150XKBC)	15 nH
L2 (Coilcraft 0805CS-270XMBC)	27 nH
R1 (Siemens B 54102-A1271-J60)	270 Ω
R2 (Siemens B 54102-A1120-J60)	12 Ω
R3	6.8 Ω

Figure 3-52 CGY121A 900-MHz application circuit and parts list.

On the collector side, for an *RC* type of amplifier, it is not a bad idea to set the collector voltage at about half the supply, which allows the RF swing to be half the supply voltage. In this case, we have violated the rule because of our 2.6-V drop. Assuming we had chosen a 0.7-V drop, then the collector–emitter voltage would be 6.3 V available for the swing. The choice of 5 mA for the transistor has come from a combination of good linearity and low noise figure. While the 10-Ω unbypassed resistor will grade the noise figure somewhat, it reduces the emitter-current-dependent distortion. The base–emitter diffusion resistance equals 26 mV/5 mA, or roughly 5 Ω. The 10-Ω resistor is twofold larger and therefore adds to the linearity and to the noise.

To determine the base bias resistor it is a good assumption that the current going through the two resistors (in our case, 8.2 and 3.3 kΩ) is about 10% (or a slightly higher fraction) of



Parts List

Frequency	1.9 GHz
C1, C2 (Siemens size 0603)	12 pF
C3, C4 (Siemens size 0603)	100 nF
C5 (Siemens size 0603)	47 nF
C6 (Siemens size 0603)	1.2 pF
L1 (Coilcraft 0805CS-150XKBC)	15 nH
R1 (Siemens B 54102-A1271-J60)	270 Ω
R2 (Siemens B 54102-A1120-J60)	12 Ω
R3	2.7 Ω

Figure 3-53 CGY121A 1.9-GHz application circuit and parts list.

the collector current. Since the dc current gain of the transistor is typically between 50 and 100, it is sufficient if the current through the resistive network is between 5 and 10 times higher than the dc bias current taken off the network. This type of bias scheme is only valid for Class A to AB_1. Later, we will see that for monolithic circuits, additional transistors will be used to generate the required voltage drops. Lately, combinations of *npn* and *pnp* transistors have become popular to avoid using any resistance in the emitter for dc biasing purposes. Figure 3-56 shows such a circuit. This, of course, invites thermal runaway. To minimize the circuitry, there are bias ICs, such as the Siemens BCR400R, now available. Figures 3-57 and 3-58 show its application in biasing BJTs and FETs, respectively; it can also be used to control PIN-diode bias current for TR antenna switching, as shown in Figure 3-59.

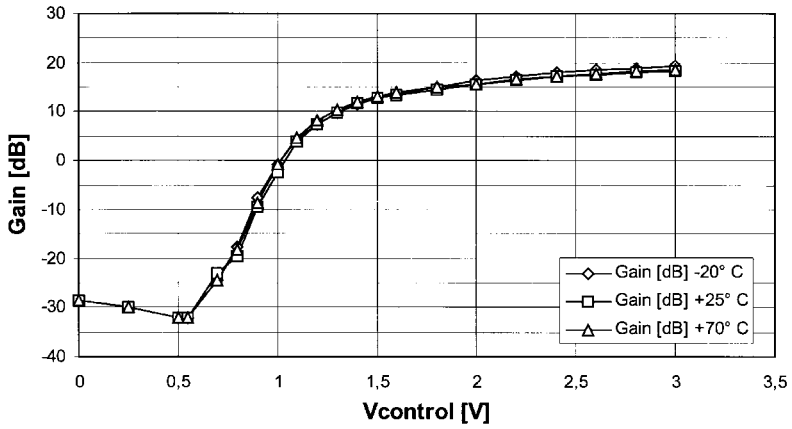


Figure 3-54 CGY121A gain versus $V_{control}$ at 1.9 GHz.

To validate this, we have used the Serenade 8.0 CAD tool, which has a built-in bias- and temperature-dependent noise model for BJTs and FETs. The SPICE-type dc analysis determines the actual currents, which were mentioned above. By maintaining the same bias and just bypassing both emitter resistors at the same time, the set of curves shown in Figure 3-60 will be obtained. As an exercise, we recommend the user to determine the optimum generator impedance for F_{min} with a 10-Ω resistor. Again, both the emitter current and collector voltage affect the noise figure, the intermodulation, and f_T. If a transistor is used

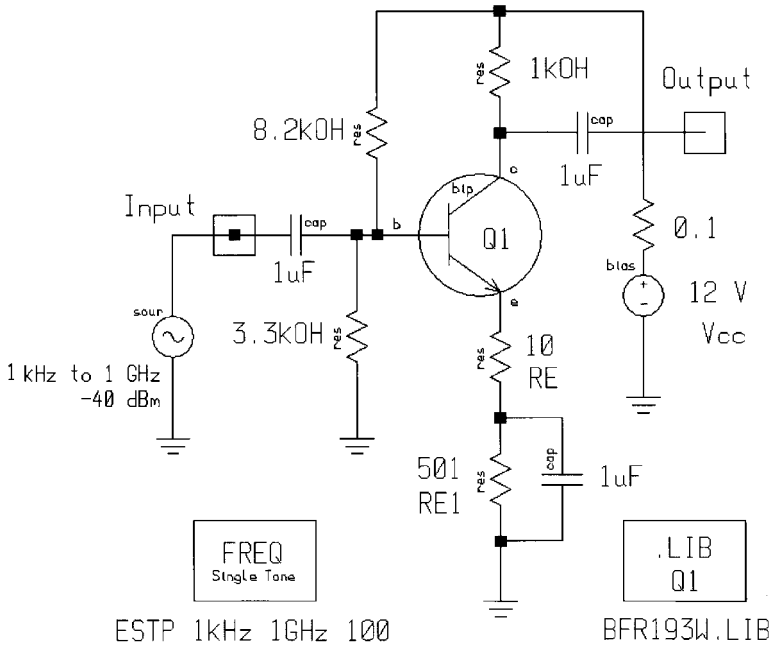


Figure 3-55 BJT amplifier example.

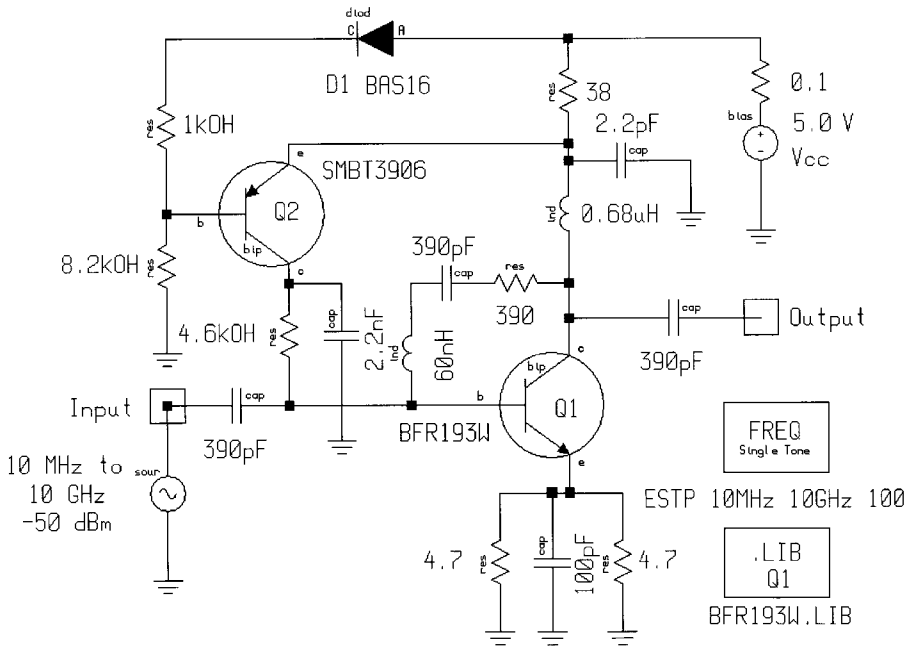


Figure 3-56 Wideband amplifier with active biasing. The 4.7-Ω resistors in the BFR193W emitter compensate the circuit's gain at low frequencies.

that is intended to be operated at significantly higher currents, then f_T will suffer while the other two parameters are less affected.

The dc input power translates into thermal dissipation, and Figure 3-60 is a plot of the noise figure as a function of temperature with and without emitter feedback. The heavy dc stabilization prevents a significant effect on the other parameters. The actual dc shift is from 5.0 to 5.2 mA for a temperature change of 70 K.

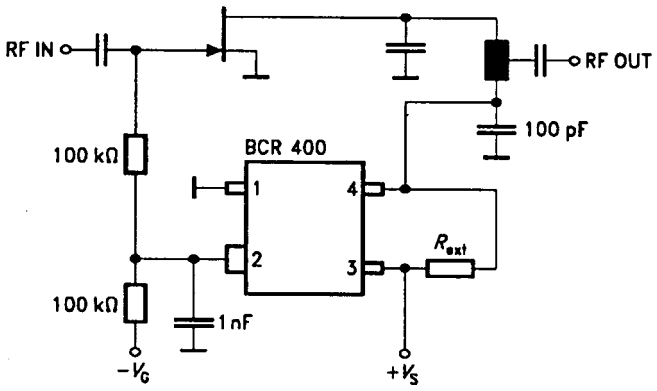


Figure 3-57 The Siemens BCR400 bias controller applied to a GaAsFET.

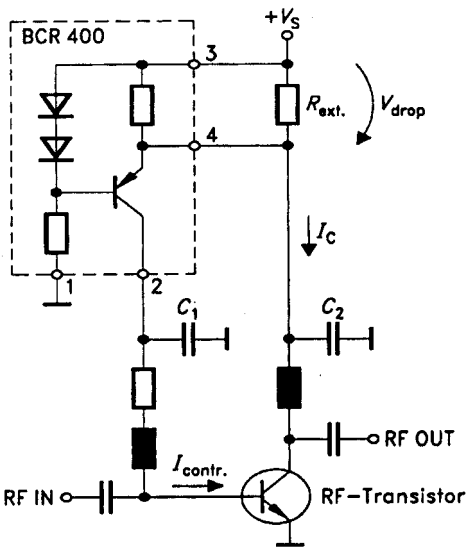


Figure 3-58 The Siemens BCR400 bias controller applied to a BJT.

As far as FETs are concerned, the enhancement types, typically PMOS, can use a similar type of bias, while the NMOS types complicate the designer's life as they require a negative gate voltage. For low-noise operation, these devices are operated somewhat close to the pinchoff voltage. Putting a source resistor in place causes mostly headaches simply because the gain is so high that the transistors are always ready to find a frequency at which they love to oscillate. The transistor will look for every possible parasitic to turn the amplifier into an oscillator.

3-2 AMPLIFIER GAIN, STABILITY, AND MATCHING

Most semiconductor manufacturers describe their devices these days by providing either their *S* parameters (small-signal) or SPICE-type parameters (large-signal). One of our headaches in writing this book has been that many of the published models don't translate

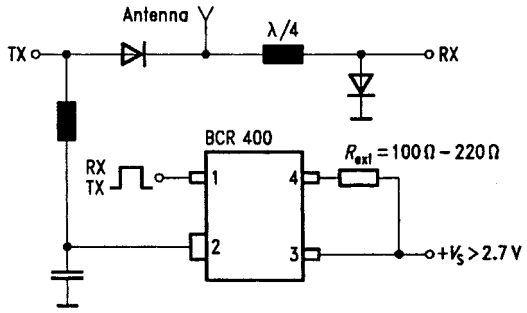


Figure 3-59 The Siemens BCR400 bias controller applied to TR switching.

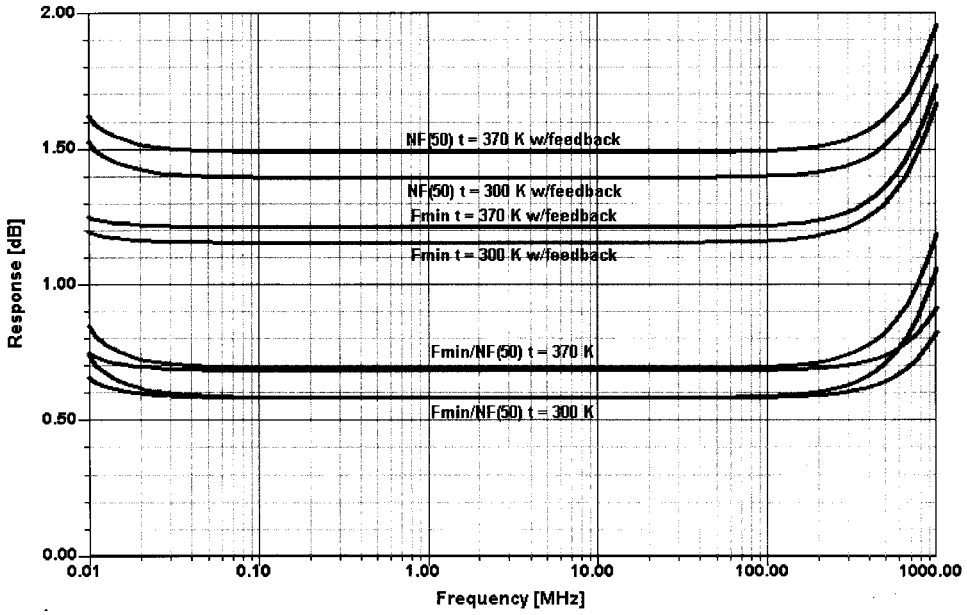


Figure 3-60 F_{min} and NF for the amplifier in Figure 3-55, with and without emitter feedback, at 300 and 370 K.

from large signal to small signal without giving unexpected or false results. Despite the availability of tools to do it right, many companies still can only provide data that are practically inadequate, and unfortunately the CAD manufacturers then get the blame for results for which they are not really responsible.

Tables 3-9 and 3-10 show typical data for a bipolar transistor (the BFP420 by Siemens). The next section presents a summary of the derivation and use of *S* parameters as published by Hewlett-Packard and others.

3-2-1 Scattering Parameter Relationships

Scattering (*S*) parameters characterize a network in terms of incident and reflected waves (Figures 3-61 and 3-62).

$$b_1 = s_{11}a_1 + s_{12}a_2 \tag{3-122}$$

$$b_2 = s_{21}a_1 + s_{22}a_2 \tag{3-123}$$

Notice that

$$s_{11} = \frac{b_1}{a_1} = \frac{V_1/I_1 - Z_0}{V_1/I_1 + Z_0} = \frac{Z_1 - Z_0}{Z_1 + Z_0} \tag{3-124}$$

and

Table 3-9 BFP420 common-emitter *S* parameters[a]

f	S_{11}		S_{21}		S_{12}		S_{22}	
(GHz)	(Magnitude)	(Angle)	(Magnitude)	(Angle)	(Magnitude)	(Angle)	(Magnitude)	(Angle)
				V_{CE} = 2 V, I_C = 20 mA				
0.01	0.452	–2.3	37.62	178.3	0.0011	94.4	0.956	–0.6
0.1	0.447	–52.1	36.30	164.7	0.0068	82.5	0.941	–12.4
0.5	0.386	–101.1	23.41	121.0	0.0262	61.7	0.632	–47.2
1.0	0.378	–146.2	13.99	96.0	0.0395	57.8	0.395	–63.9
2.0	0.405	173.5	7.18	70.8	0.0664	54.0	0.222	–87.3
3.0	0.446	149.4	4.77	52.6	0.0949	47.1	0.133	–111.3
4.0	0.501	130.0	3.52	36.8	0.1206	38.5	0.133	–158.5
6.0	0.599	104.8	2.27	8.2	0.1646	18.9	0.196	142.0
8.0	0.700	78.5	1.51	–20.8	0.1800	–2.4	0.289	99.3
9.0	0.758	67.6	1.25	–34.4	0.1820	–13.0	0.379	84.1
10.0	0.800	62.0	1.04	–43.5	0.1800	–19.3	0.465	76.6
				V_{CE} = 2 V, I_C = 5 mA				
0.01	0.790	–1.0	15.14	179.2	0.0012	83.4	0.988	–0.7
0.1	0.786	–11.6	14.98	171.8	0.0092	84.1	0.982	–6.5
0.5	0.702	–55.7	12.86	140.1	0.0398	62.8	0.857	–29.8
1.0	0.589	–99.1	9.63	112.6	0.0603	46.5	0.647	–48.6
2.0	0.507	–156.0	5.60	79.4	0.0798	34.6	0.401	–70.3
3.0	0.511	168.5	3.84	57.1	0.0957	29.8	0.267	–84.2
4.0	0.549	142.0	2.87	38.5	0.1121	25.1	0.207	–110.5
5.0	0.604	123.9	2.26	22.1	0.1285	19.4	0.150	–137.3
6.0	0.633	110.0	1.86	6.7	0.1442	13.1	0.173	–169.8

Table 3-10 BFP420 noise parameters

f	$F_{\min}$[a]	G_a[a]	Γ_{opt}		R_N	r_n	$F_{50\Omega}$[b]	$\lvert S_{21} \rvert^2$ [b]
(GHZ)	dB	(dB)	(Magnitude)	(Angle)	(Ω)		(dB)	(dB)
				V_{CE} = 2 V, I_C = 5 mA				
0.9	0.90	20.5	0.28	41.0	8.7	0.17	1.02	20.3
1.8	1.05	15.2	0.20	82.0	6.7	0.13	1.11	15.8
2.4	1.25	13.0	0.20	124.0	5.5	0.11	1.32	13.5
3.0	1.38	12.1	0.22	–175.0	5.0	0.10	1.48	11.6
4.0	1.55	10.3	0.33	–157.0	5.5	0.11	1.83	9.1
5.0	1.75	8.6	0.45	–142.0	5.0	0.10	2.20	7.0
6.0	2.20	6.4	0.53	–123.0	15.0	0.30	3.30	5.3

[a]Input matched for minimum noise figure, output for maximum gain.
[b]$Z_S = Z_L = 50\ \Omega$.

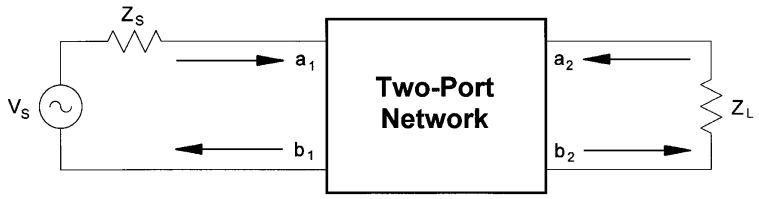


Figure 3-61 *S* parameters characterize a network (a two-port is shown) in terms of incident waves (a_1, a_2) and reflected waves (b_1, b_2). Figure 3-62 shows the flow graph for this network.

$$Z_1 = Z_0 \frac{(1 + s_{11})}{(1 - s_{11})} \tag{3-125}$$

where $Z_1 = V_1/I_1$ is the input impedance at Port 1.

This relationship between reflection coefficient and impedance is the basis of the Smith Chart transmission-line calculator. Consequently, the reflection coefficients S_{11} and S_{22} can be plotted on Smith charts, converted directly to impedance, and easily manipulated to determine matching networks for optimizing a circuit design.

The above equations show one of the important advantages of *S* parameters: They are simply gains and reflection coefficients, both familiar quantities to engineers. By comparison, some of the *Y* parameters described earlier in the article are not so familiar. For example, the *Y* parameter corresponding to insertion gain S_{21} is the "forward transadmittance," Y_{21}. Clearly, insertion gain gives by far the greater insight into the operation of the network.

S parameters are simply related to power gain and mismatch loss, quantities that are often of more interest than the corresponding voltage functions:

$$|s_{11}|^2 = \frac{\text{Power reflected from the network input}}{\text{Power incident on the network input}} \tag{3-126}$$

$$|s_{21}|^2 = \frac{\text{Power delivered to a } Z_0 \text{ load}}{\text{Power available from } Z_0 \text{ source}} \tag{3-127}$$

$$|s_{12}|^2 = \text{Reverse transducer power gain with } Z_0 \text{ load and source.} \tag{3-128}$$

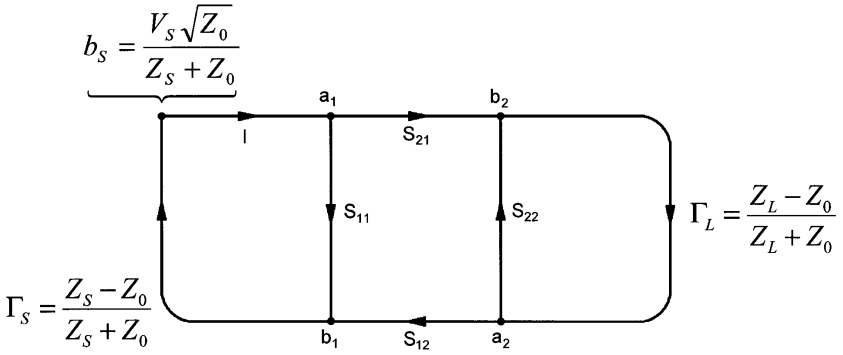


Figure 3-62 Flow graph for the network shown in Figure 3-61.

$$|s_{22}|^2 = \frac{\text{Power reflected from the network output}}{\text{Power incident on the network output}} \tag{3-129}$$

From this we obtain:
Input reflection coefficient with arbitrary Z_L:

$$s'_{11} = s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1 - s_{22}\Gamma_L} \tag{3-130}$$

Output reflection coefficient with arbitrary Z_S:

$$s'_{22} = s_{22} + \frac{s_{12}s_{21}\Gamma_S}{1 - s_{11}\Gamma_S} \tag{3-131}$$

Voltage gain with arbitrary Z_L and Z_S:

$$A_V = \frac{V_2}{V_1} = \frac{s_{21}(1 + \Gamma_L)}{(1 - s_{22}\Gamma_L)(1 + s'_{11})} \tag{3-132}$$

$$\text{Power gain} = \frac{\text{Power delivered to load}}{\text{Power input to network}}:$$

$$G = \frac{|s_{21}|^2 (1 - |\Gamma_L|^2)}{(1 - |s_{11}|^2) + |\Gamma_L|^2 (|s_{22}|^2 - |D|^2) - 2\,\text{Re}\,(\Gamma_L N)} \tag{3-133}$$

$$\text{Available power gain} = \frac{\text{Power available from network}}{\text{Power available from source}}:$$

$$G_A = \frac{|s_{21}|^2 (1 - |\Gamma_S|^2)}{(1 - |s_{22}|^2) + |\Gamma_S|^2 (|s_{11}|^2 - |D|^2) - 2\,\text{Re}\,(\Gamma_S M)} \tag{3-134}$$

$$\text{Transducer power gain} = \frac{\text{Power delivered to load}}{\text{Power available from source}}:$$

$$G_T = \frac{|s_{21}|^2 (1 - |\Gamma_S|^2)(1 - |\Gamma_L|^2)}{|(1 - s_{11}\Gamma_S)(1 - s_{22}\Gamma_L) - s_{12}s_{21}\Gamma_L\Gamma_S|^2} \tag{3-135}$$

Unilateral transducer power gain ($S_{12} = 0$):

$$G_{Tu} = \frac{|s_{21}|^2(1 - |\Gamma_S|^2)(1 - |\Gamma_L|^2)}{|1 - s_{11}\Gamma_S|^2|1 - s_{22}\Gamma_L|^2}$$

$$= G_0 G_1 G_2 \tag{3-136}$$

$$G_0 = |s_{21}|^2 \tag{3-137}$$

$$G_1 = \frac{1 - |\Gamma_S|^2}{|1 - s_{11}\Gamma_S|^2} \tag{3-138}$$

$$G_2 = \frac{1 - |\Gamma_L|^2}{|1 - s_{22}\Gamma_L|^2} \tag{3-139}$$

Maximum unilateral transducer power gain when $|S_{11}| < 1$ and $|S_{22}| < 1$:

$$G_u = \frac{|s_{21}|^2}{|(1 - |s_{11}|^2)(1 - |s_{22}|)^2|}$$

$$= G_0\, G_{1\max}\, G_{2\max} \tag{3-140}$$

$$G_{i\max} = \frac{1}{1 - |s_{ii}|^2} \quad i = 1, 2 \tag{3-141}$$

Constant-gain circles (unilateral case: $S_{12} = 0$):

- Center of constant-gain circle is on line between center of Smith Chart and point representing S_{ii}^*
- Distance of center of circle from center of Smith Chart:

$$r_i = \frac{g_i |s_{ii}|}{1 - |s_{ii}|^2(1 - g_i)} \tag{3-142}$$

- Radius of circle:

$$\rho_i = \frac{\sqrt{1 - g_i}(1 - |s_{ii}|^2)}{1 - |s_{ii}|^2(1 - g_i)} \tag{3-143}$$

where $i = 1, 2$ and

$$g_i = \frac{G_i}{G_{i\max}} = G_i(1 - |s_{ii}|^2) \tag{3-144}$$

Unilateral figure of merit:

$$u = \frac{|s_{11}s_{22}s_{12}s_{21}|}{|(1 - |s_{11}|^2)(1 - |s_{22}|^2)|} \tag{3-145}$$

Error limits of unilateral gain calculation:

$$\frac{1}{(1 + u^2)} < \frac{G_T}{G_{Tu}} < \frac{1}{(1 - u^2)} \tag{3-146}$$

Conditions for absolute stability. No passive source or load will cause a network to oscillate if a, b, and c are all satisfied:

$$\text{a.}\quad |s_{11}| < 1, |s_{22}| < 1 \tag{3-147}$$

$$\text{b.} \quad \left| \frac{|s_{12}s_{21}| - |M^*|}{|s_{11}|^2 - |D|^2} \right| > 1 \tag{3-148}$$

$$\text{c.} \quad \left| \frac{|s_{12}s_{21}| - |N^*|}{|s_{22}|^2 - |D|^2} \right| > 1 \tag{3-149}$$

where

$$D = s_{11}s_{22} - s_{12}s_{21} \tag{3-150}$$

$$M = s_{11} - Ds_{22}^* \tag{3-151}$$

$$N = s_{22} - Ds_{11}^* \tag{3-152}$$

Condition that a two-port network can be simultaneously matched with a positive real source and load:

$$K > 1 \tag{3-153}$$

$$K = \frac{1 + |D|^2 - |s_{11}|^2 - |s_{22}|^2}{2|s_{12}s_{21}|} \tag{3-154}$$

Source and load for simultaneous match

$$\Gamma_{mS} = M^* \left[\frac{B_1 \pm \sqrt{B_1^2 - 4|M|^2}}{2|M|^2} \right] \tag{3-155}$$

$$\Gamma_{mL} = N^* \left[\frac{B_2 \pm \sqrt{B_2^2 - 4|N|^2}}{2|N|^2} \right] \tag{3-156}$$

where

$$B_1 = 1 + |s_{11}|^2 - |s_{22}|^2 - |D|^2 \tag{3-157}$$

$$B_2 = 1 + |s_{22}|^2 - |s_{11}|^2 - |D|^2 \tag{3-158}$$

Maximum available power gain. If $K > 1$,

$$G_{A\max} = \left| \frac{s_{21}}{s_{12}} (K \pm \sqrt{K^2 - 1}) \right| \tag{3-159}$$

In conclusion, *a two-port network (active or passive) is unconditionally stable* when

$$K > 1$$

$$|S_{12}S_{21}| < 1 - |S_{11}|^2$$

and

$$|S_{12}S_{21}| < 1 - |S_{22}|^2$$

3-2-2 Low-Noise Amplifiers

A low-noise amplifier combines a low noise figure, reasonable gain, and stability without oscillation over its entire useful frequency range. These amplifiers are typically operated in Class A, which is characterized by a bias point more or less at the center of maximum current and voltage capability of the device used, and by RF current and voltages that are sufficiently small relative to the bias point that the bias point does not shift. This means that an amplifier operating at, say, 3 V and 1 mA for low-noise, high-gain operation will not be Class A if the RF current is larger than 0.1% of the current or the voltage swing is more than 0.1% of the current swing. These amplifiers can be designed using standard linear *S* parameters, which are (hopefully) available from the device manufacturer for the bias point of interest. If suitable *S* parameters are not available, we must resort to a CAD tool that allows us to determine a device's bias point as a function of its bias network and to generate small-signal *S* parameters for the device. With suitable *S* parameters in hand, and assuming the amplifier will be frequency selective, we then connect tuned circuits to the input and output to provide matching.

In designing amplifiers, the "power gain" of a circuit is mentioned frequently. This is actually deceptive because power gain is somewhat associated with dc power dissipation, and in considering it, we are on the verge of discussing large-signal performance. Therefore it would be better to define the power gain as a function of operating mode, such as "Class A power gain." We have also outlined that there is a difference between noise matching and gain matching. If we can be allowed to step back into the vacuum-tube era for a moment, the reader should be reminded that the difference between power gain and noise matching had to do with the grid-plate feedback capacitance which, via the Miller effect, showed up in parallel with the input. This actually detunes the input slightly off-center relative to the best gain. The reason for this is that as the feedback increases, Y_{12} or S_{12} starts playing a role, and the frequently assumed unilateral case no longer exists. The difference between noise matching and gain matching increases as a function of frequency. It can be brought closer together if appropriate reactive feedback is used. The two choices are: (1) increase the emitter or source inductance or (2) slightly increase the base-to-collector/gate-to-drain capacitance. Needless to say, the penalty for this is potential instability, and a sweep must be done from a few megahertz to several gigahertz to make sure the stage doesn't "take off."

The most efficiency this circuit provides is 50%; even this is a theoretical limit if one tries to keep distortion under control.

While the equation $P_{\text{OUT}} = (V_{\text{bat}} - V_{\text{sat}})^2/2R_L$ is valid, it really applies only to power amplifiers in the sense that for Class A, high-gain operation R_L can be made significantly higher as the absolute power output is not the important item. While the supply voltage V_{bat} is always known, the saturation voltage V_{sat} has to be obtained by the load line dc analysis. Note: The RF saturation voltage is always higher than the dc one. When looking at Figure 3-63, it becomes obvious that the conducting angle is 180°, which implies that the entire sine wave at the input is reproduced as an essentially undistorted sine wave at the output.

We can now calculate the components of the distorted current in separating them into the dc component, the fundamental frequency, and its first and second harmonics:

$$\text{Fundamental} \quad i_{c1} = i_{\text{peak}} \cdot f_1(\Theta) \tag{3-160}$$

$$\text{Second harmonic} \quad i_{c2} = i_{\text{peak}} \cdot f_2(\Theta) \tag{3-161}$$

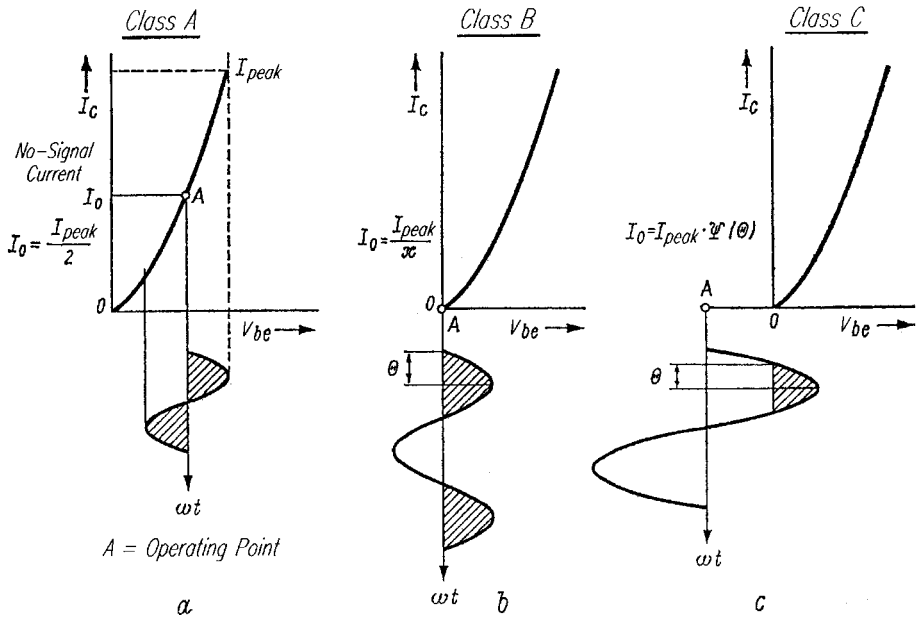


Figure 3-63 Definitions of Class A, B, and C operation, including conduction angle (Θ). Figure 3-64 shows the determination of Θ.

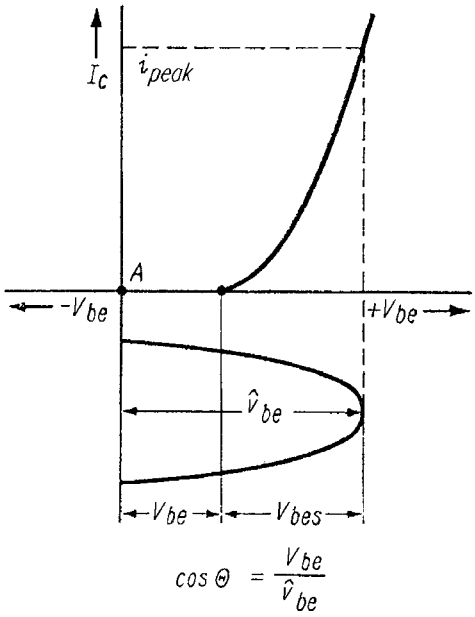


Figure 3-64 Determination of the conduction angle Θ from the base threshold voltage V_{be} and the base RF voltage $\hat{v}_{be}$.

$$\text{Third harmonic} \quad i_{c3} = i_{\text{peak}} \cdot f_3(\Theta) \tag{3-162}$$

For the collector dc current we obtain

$$I_c = i_{\text{peak}} \cdot \Psi(\Theta) \tag{3-163}$$

The frequency-dependent currents are a function of the conduction angle Θ. Figure 3-65 shows these transfer characteristics as a function of the conduction angle for a square-law dynamic transfer function; for higher-order transfer functions, the values don't change much. Tables for this can be found in various sources, including Clark and Hess [47]. Since the transistors exhibit a mixture of an exponential and a linear transfer characteristic (due to the unwanted parasitic emitter/source and collector/drain resistances when building the device), the plot in figure 3-65 is sufficiently valid for practical purposes. This function plot allows us to determine the amplitude of the current at various frequencies once the conducting angle has been set. For example, for Θ = 70°, the maximum collector current is $i_{\text{peak}} = I_c/0.253$, meaning that the peak current of the transistor is approximately four times larger than the dc value one would measure. Once the peak collector current is known, the fundamental and harmonic currents can be calculated:

$$i_{c1} = \frac{I_c}{0.253} f_1(\Theta)$$

$$i_{c1} = \frac{I_c}{0.253} 0.436$$

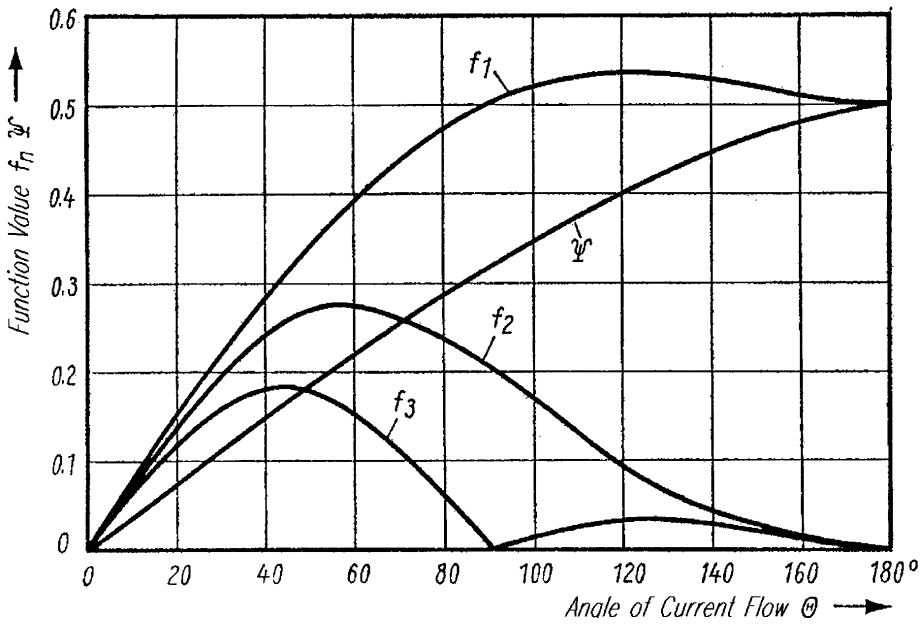


Figure 3-65 Function values for the fundamental, second harmonic, third harmonic, and dc components (f_1, f_2, f_3, and Ψ, respectively) versus device conduction angle.

In order to obtain i_{c2} and i_{c3} one needs to do the same calculations as above with the appropriate functions for f_2 and f_3.

It becomes obvious that the distorted collector current generates a lot of harmonics, which definitely requires sufficient low-pass filtering at the output, even for Class A operation. We already stated that for Class A operation the conduction angle is 180° while it is 90° for Class B operation. The conduction angle for Class C operation can be solved from the ratio of the knee voltage V_{be} and the peak value of the ac voltage at the base. We find

$$\cos \Theta = \frac{V_{be}}{\hat{V}_{be}} \tag{3-164}$$

where $\Theta = 90°$ and $V_{be} = 0$. Given the voltage V_{bes}, which determines the peak collector current (up to i_c peak), we can calculate the effective RF voltage at the base as

$$\hat{v}_{be} = \frac{v_{bes}}{1 - \cos \Theta} \tag{3-165}$$

Therefore, the peak value of v_b becomes v_{bes} for $\Theta = 90°$.

The necessary base voltage V_{be} for maximum current is obtained by the transistor model. Then one finds the threshold voltage at which the transistor becomes conductive and by using Eqs. (3-164) and (3-165) it can be checked whether this results in the appropriate conducting angle. An example for this will be shown in the high-power stage section.

Design Guidelines. The design of a one-stage amplifier consists of finding

- An input lossless matching network M_1
- An input lossless matching network M_2

so that the maximum or desired transistor gain is achieved over the operating bandwidth of the amplifier. Usually, the common-emitter or common-source configuration is chosen for highest gain per stage. If the stability factor K is greater than unity, these two networks can be found to give the maximum available gain $G_{\max}$. If the stability factor is less than or equal to unity, the amplifier could be terminated in a matching structure that causes oscillation; that is, $G_{\max}$ is infinite. This should be avoided by locating the regions of instability in the Γ_G and Γ_L planes. The input and output terminations (Γ_G and Γ_L) must be designed to avoid the instability regions. Usually, these unstable regions are near the conjugate match for S_{11} and S_{22}. Thus, a stable amplifier will require some input and/or output mismatch if K is less than or equal to unity.

There are at least two alternative approaches for potentially unstable amplifiers:

1. Add resistive matching elements to make $K \geq 1$ and $G_{\max} \approx G_{ms}$.
2. Add feedback to make $K \geq 1$ and $G_{\max \approx Gms}$.

For narrowband amplifiers, it is usually recommended that one accept a transistor with $K < 1$, design the amplifier for a gain approaching G_{ms}, and ensure that the Γ_G and Γ_L terminations provide stability at all frequencies, both inside and outside the amplifier passband.

The design of an amplifier would usually have the following specifications:

Gain and gain flatness
Bandwidth and center frequency ($f_2 - f_1, f_0$)
Noise figure
Linear output power
Input reflection coefficient ($VSWR_{in}$)
Output reflection coefficient ($VSWR_{out}$)
Bias voltage and current

For small-signal amplifiers, the small-signal S parameters are sufficient to complete the design. After selecting an appropriate transistor based on these specifications, a one-stage amplifier design should be considered if sufficient gain can be achieved; otherwise, a two-stage amplifier should be designed.

The circuit topology should be chosen to allow dc bias for the transistor. Usually, RF short-circuited stubs are placed near the transistor to allow dc biasing. If the topology does not allow dc bias, a broadband bias choke or high-resistance bias circuit that does not affect the amplifier performance must be used.

The following steps can be tabulated for the design of a one-stage amplifier (see Figure 3.66):

1. Select a transistor based on a datasheet description of the S parameters, noise figure, and linear output power.
2. Calculate K and G_{max} or G_{ms} versus frequency.
3. For $K > 1$, select the topologies that match the input and output (and allow dc biasing) at the upper band edge f_2. Ideally, this will give G_{max} and $S'_{11} = S'_{22} = 0$. Usually, $S_{12} = 0$ is assumed for the initial design. Next, the topology may be varied to flatten gain versus frequency at the expense of S'_{11} and S'_{22}.
4. For $K < 1$, plot the regions of instability on the Γ_G and Γ_L planes and select topologies that partially match the input and output at the upper band edge and avoid the unstable regions. The gain will approach G_{ms} as an upper limit. Next, the topology may be varied to flatten gain versus frequency.
5. After finding initial M_1 and M_2, plot the amplifier S parameters versus frequency; make adjustments in topology until the specifications for gain, input reflection coefficient, and output reflection are satisfied. Also plot Γ_G and Γ_L versus frequency to verify amplifier stability.
6. Design the dc bias circuit. Lay out the elements of the complete amplifier and check realizability.

As a beginning point in an amplifier design, the circuit topologies may be designed with the assumption that S_{12} is zero. Later, an exact design procedure will be described that includes the topologies with terminations in the stable region.

Design Examples. Throughout this book we always address the issue that we have to separate the battery-operated handheld applications and the more stationary designs like those for base stations. In the case of the handheld or low-voltage applications, the dynamic range requirements are significantly less than the front-end of a base station requires.

NE68133 Matched Amplifier. We will first consider the design of a low-voltage, low-current, low-noise amplifier based on the NEC BJT NE68133. This device was chosen

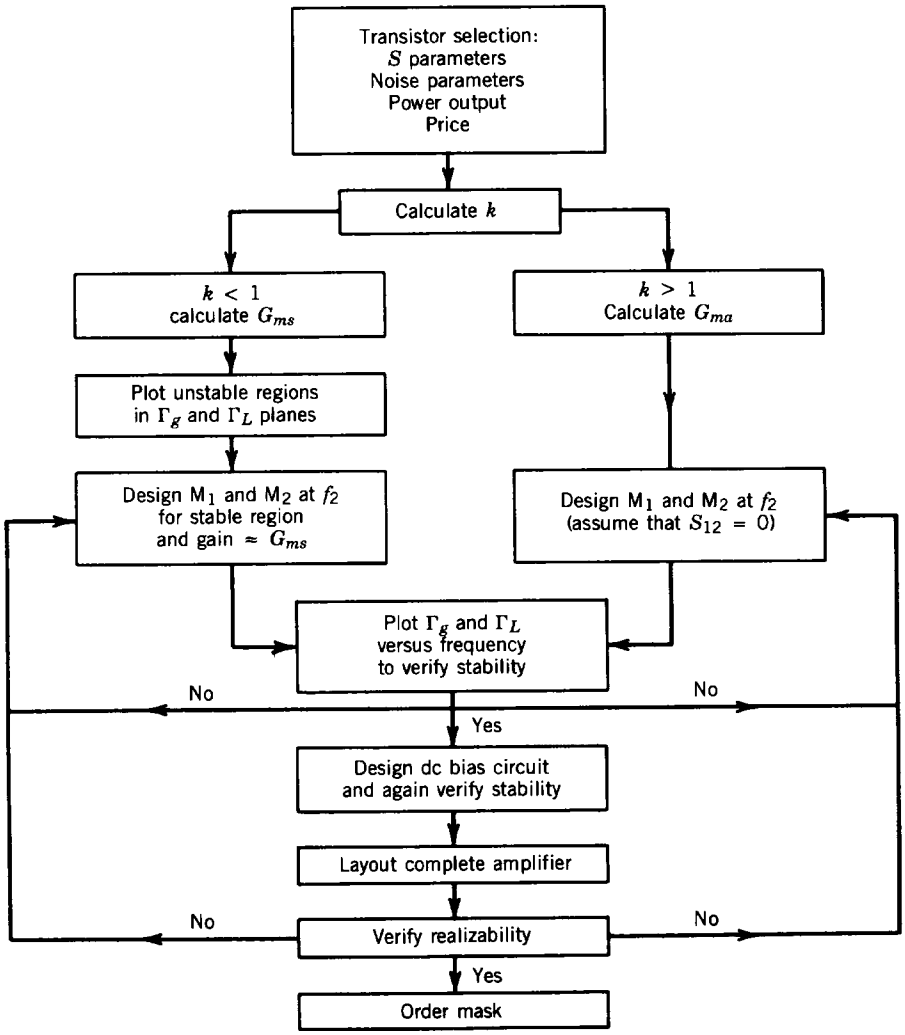


Figure 3-66 Simplified amplifier design procedure.

because it is unconditionally stable, has a minimum noise figure of about 1.2 dB at 7 mA, has 15.5 dB gain, and has an f_T of 8 GHz. While it is possible to start the design based on the published S and noise parameters, we will choose also to involve the nonlinear model, as this allows us more freedom in selecting the bias point including the collector voltage. The actual topology of the amplifier is shown in Figure 3-67.

Determining appropriate values for the input and output matching networks is essential to finalizing the design. The next several figures illustrate matching-network syntheses undertaken with the interactive Smith Tool in Ansoft's Serenade Design Environment. We evaluate the point of tangency as well as stability circles in the source and load planes, indicating that our impedance match will result in a stable amplifier.

In Figure 3-68, the center circle (with 1.1 marker) refers to the noise circle for which the noise figure of 1.1 dB can be obtained. The left circle (with the 14.3 marker) corresponds to

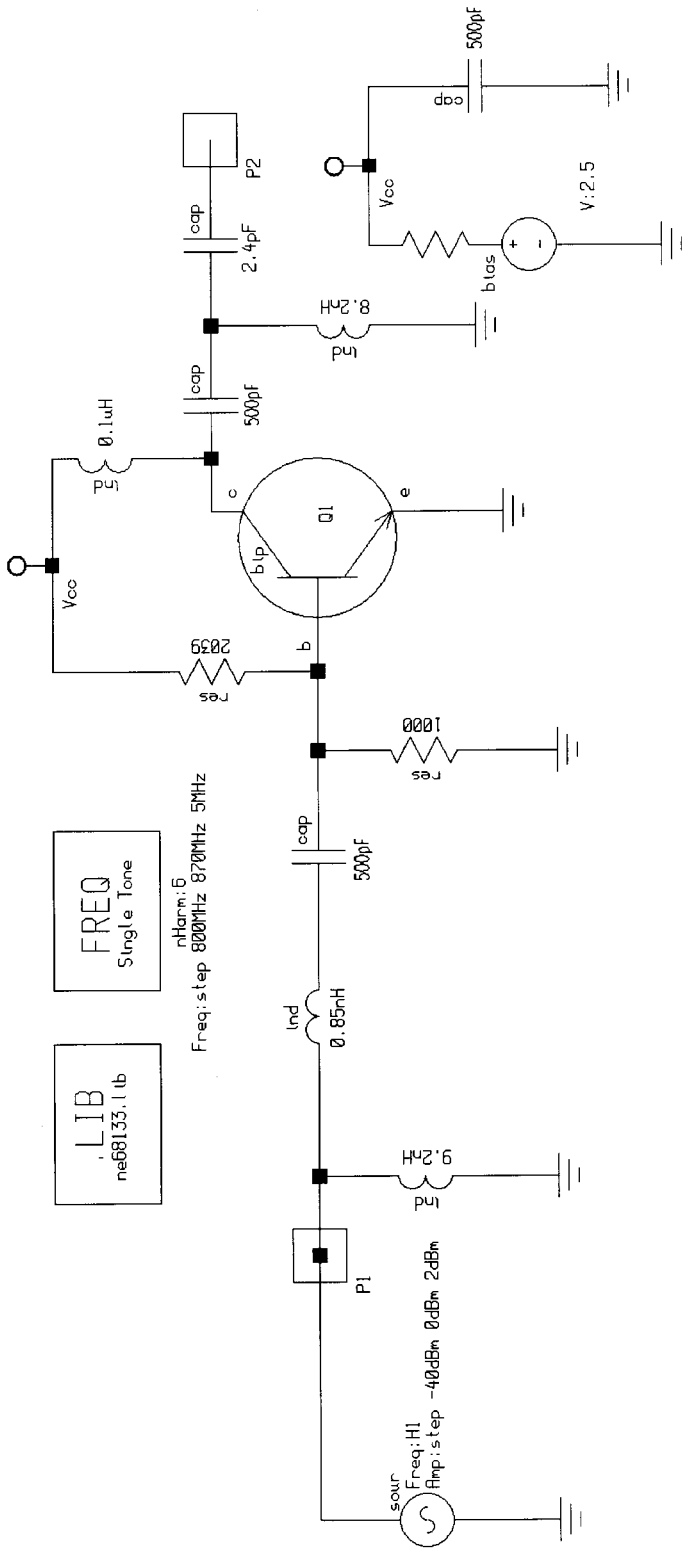


Figure 3-67 The NE68133 low-noise amplifier.

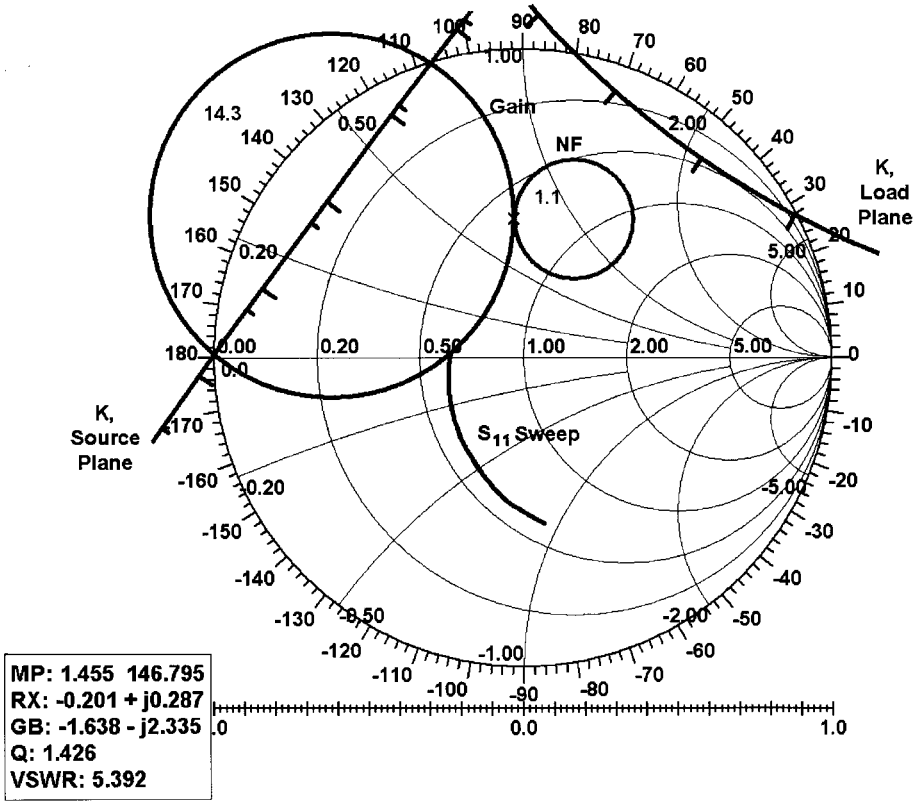


Figure 3-68 Display from the interactive Smith Tool, showing circles for gain, noise figure, and source- and load-plane stability, as well as a sweep of S_{11} versus frequency. In synthesizing an input matching circuit, we choose a single working frequency and work from the chart's origin to the point of tangency between the gain and noise figure circles. See Figure 3-69.

source-plane terminations that correspond to 14.3 dB of available gain. In order to get the required noise figure *and* gain, the input termination must be at the point of tangency of both circles. The arc touching the X axis close to 0.50 is the amplifier's S_{11} response from 800 to 875 MHz. The shallow arc that nearly bisects the 14.3-dB gain circle is actually a portion of the source-plane stability circle; the arc at the top right is a portion of the load-plane stability circle. Because we will be terminating the amplifier input with a value inside the source-plane stability circle, and terminating the amplifier output with a value outside the load-plane stability circle, the amplifier will be unconditionally stable at the frequency at which the match is obtained.

The next task is to obtain the values for the network circuit that matches the input for the 1.1-dB noise figure. A noise matching will not result in the best possible S_{11} matching (Figure 3-69). This translates into an inductor of 9.2 nH to ground and a series inductor of 0.85 nH based on the large-signal model of the NE68133 at the chosen bias point.

After we have determined our input matching network, it is time to find the output matching (Figure 3-70). First, find the termination point on the output plane that corresponds to 14.3-dB gain and 1.1-dB NF at the input plane. The output gain circle goes through the

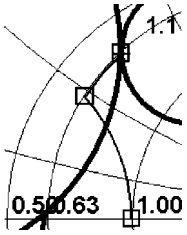


Figure 3-69 Input matching-network extraction, assuming lossless matching components and no parasitic effects.

1.1-dB NF circle at two points, and for reasons of easy matching we selected the lower point. The next step is to find the conjugate reflection coefficient $\Gamma_L = S_{22}^*$. This is the lower square marker in Figure 3-70. The remaining task is to find the matching network that brings this value to 50 Ω. By doing so, we first obtain a shunt inductance $L_{\text{shunt}} = 8.2$ nH and a series capacitor $C_{\text{series}} = 2.4$ pF. Again, these are the values applicable to the large-signal equivalent circuit.

The careful reader will have noticed that we have been avoiding the issue of small-signal exact matching using the manufacturer-supplied parameters. If this is done, the narrow band response shown in Figure 3-71 will be obtained. Therefore, we must optimize the circuit for a better frequency response to obtain the one shown in Figure 3-72.

After this optimization, we were well able to meet the specification of a center band gain of 14 dB and a corresponding NF of 1.1 dB. To show the flatness of the gain, Figure 3-73 shows a gain variation in the range of 0.25 dB.

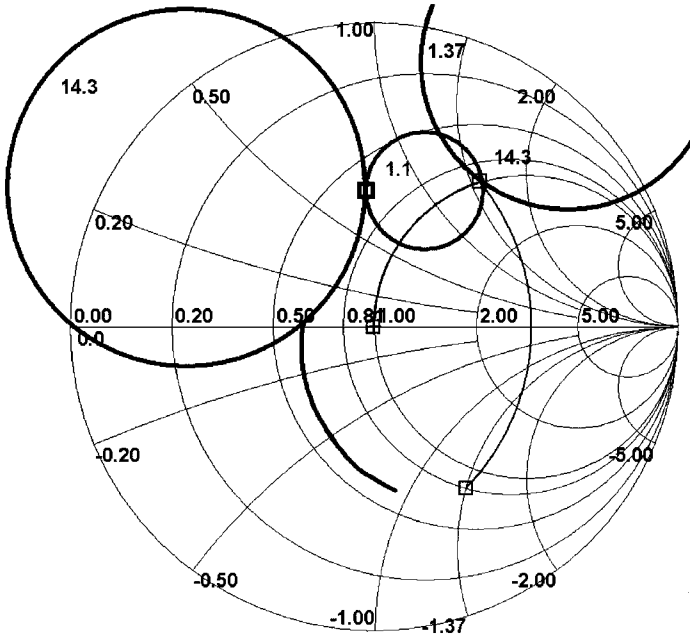


Figure 3-70 Output matching-network extraction.

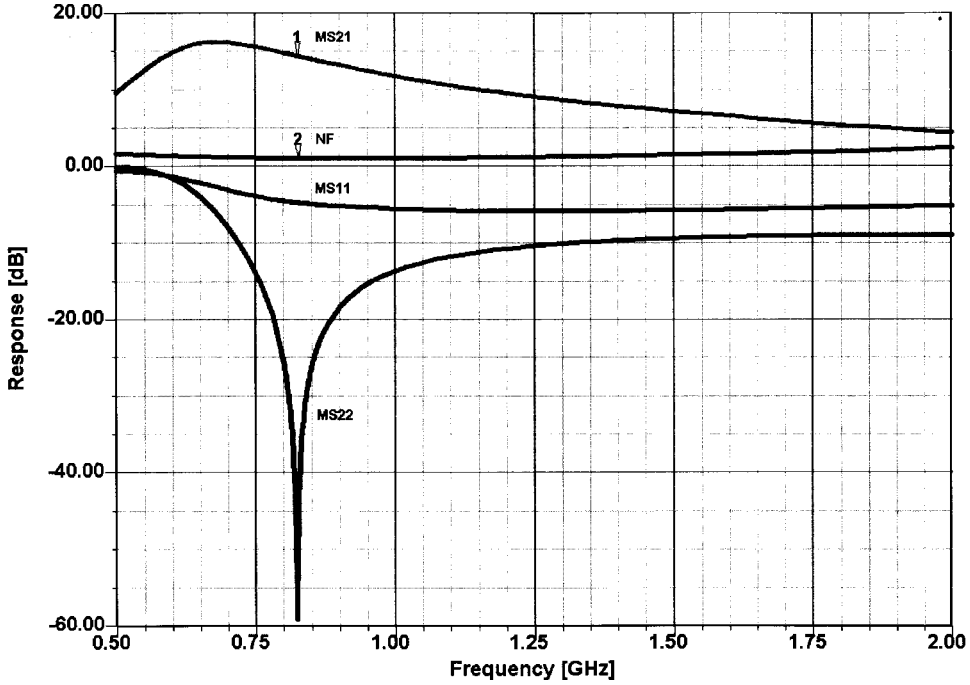


Figure 3-71 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-67.

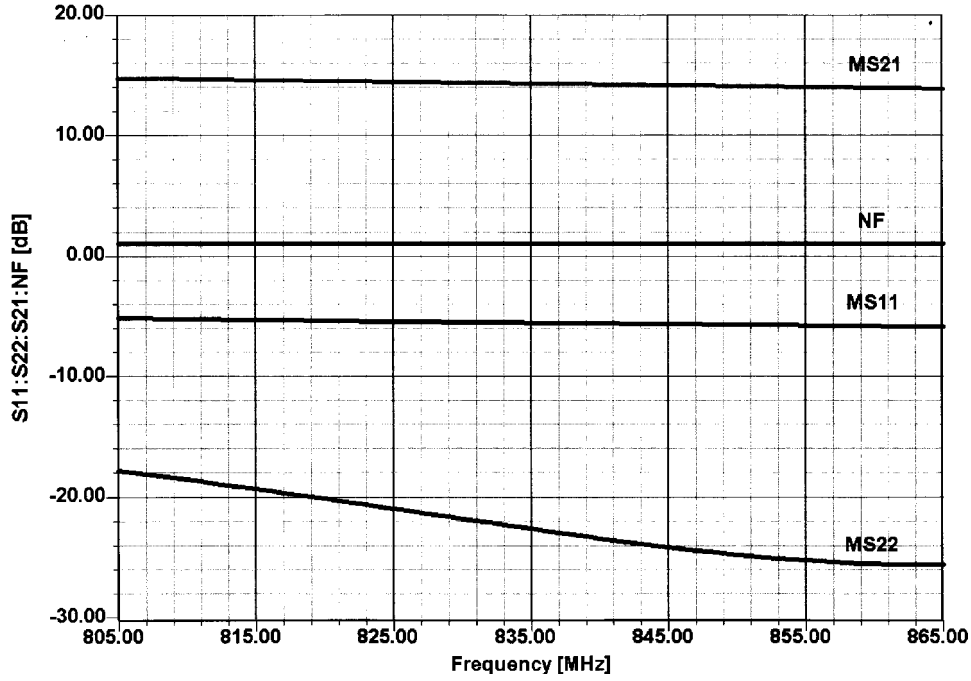


Figure 3-72 Optimized performance of input/output reflection, gain, and noise figure.

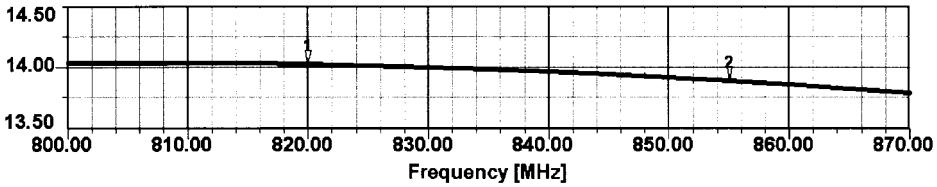


Figure 3-73 High-resolution display of gain as a function of frequency.

The next step is to look into the large-signal performance. The first order of business is to determine the third-order intercept point, which is based on a two-tone analysis. Given the fact that we run the device at low voltage and at a just-reasonable current, the actual intercept point of +13.5 dBm is quite a good number. Figure 3-74 shows the fundamental and third-order outputs. The crossover point between the two, as shown in Chapter 1, is defined as the third-order intercept point. If we had chosen the third harmonic output rather than the fundamental, we would have shown the fifth-order intercept point. The second-order IMD product ($f_1 \pm f_2$) typically can be minimized by input selectivity.

How an amplifier's linearity affects its ability to preserve the characteristics of digital signals is a major concern in wireless design. Figures 3-75 and 3-76 show constellation diagrams with the amplifier handling a PSK signal at drive levels of –40 dBm and +10 dBm, respectively. As expected, we see significant distortion at the higher drive levels due to circuit nonlinearities.

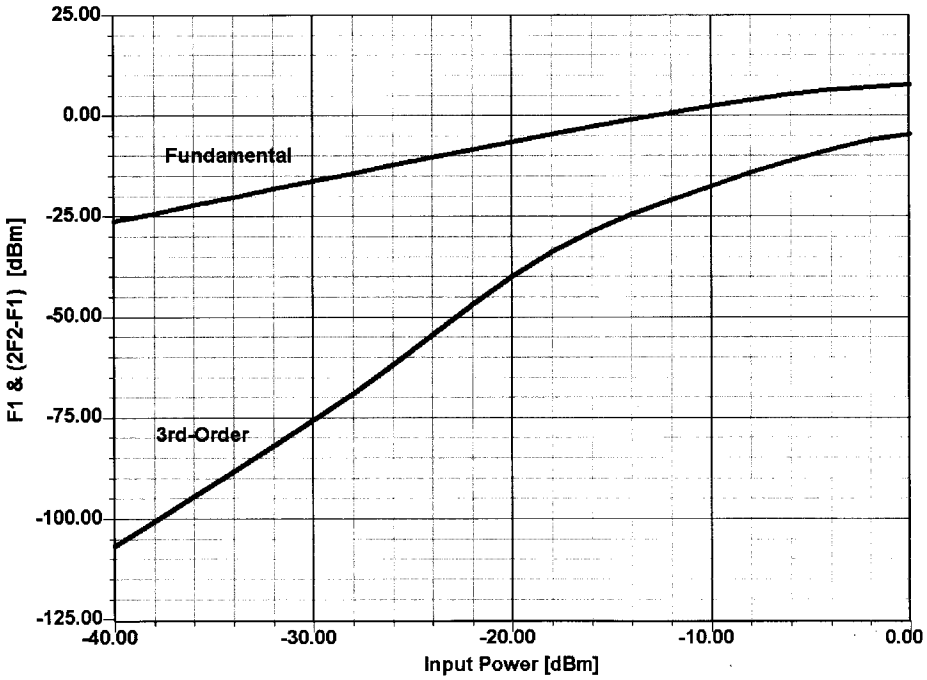


Figure 3-74 Simulated fundamental and third-order IMD outputs of the NE68133 amplifier.

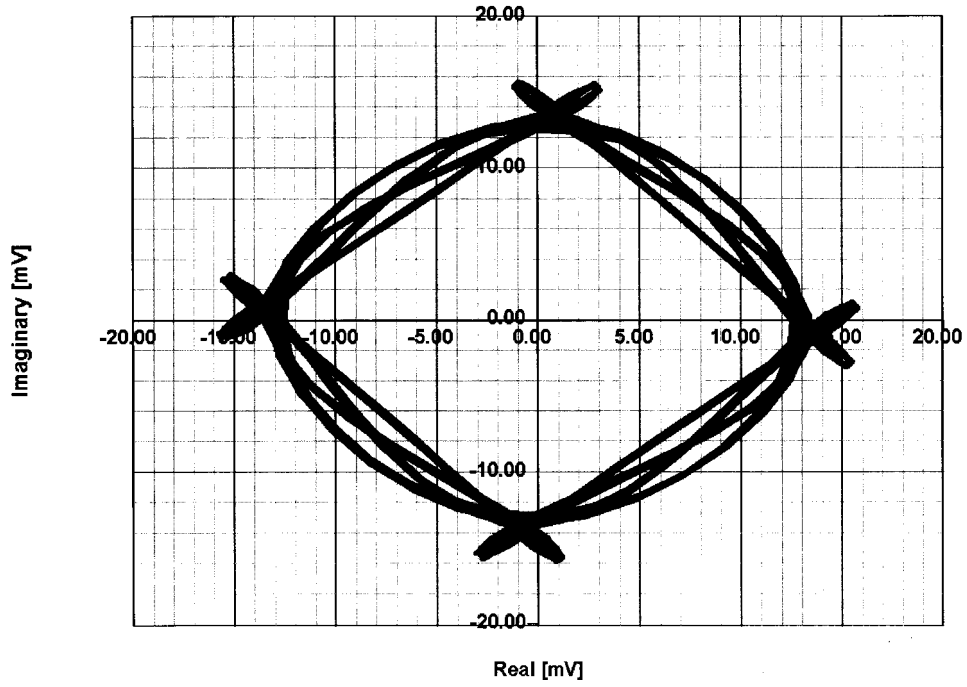


Figure 3-75 Output constellation at –40-dBm input.

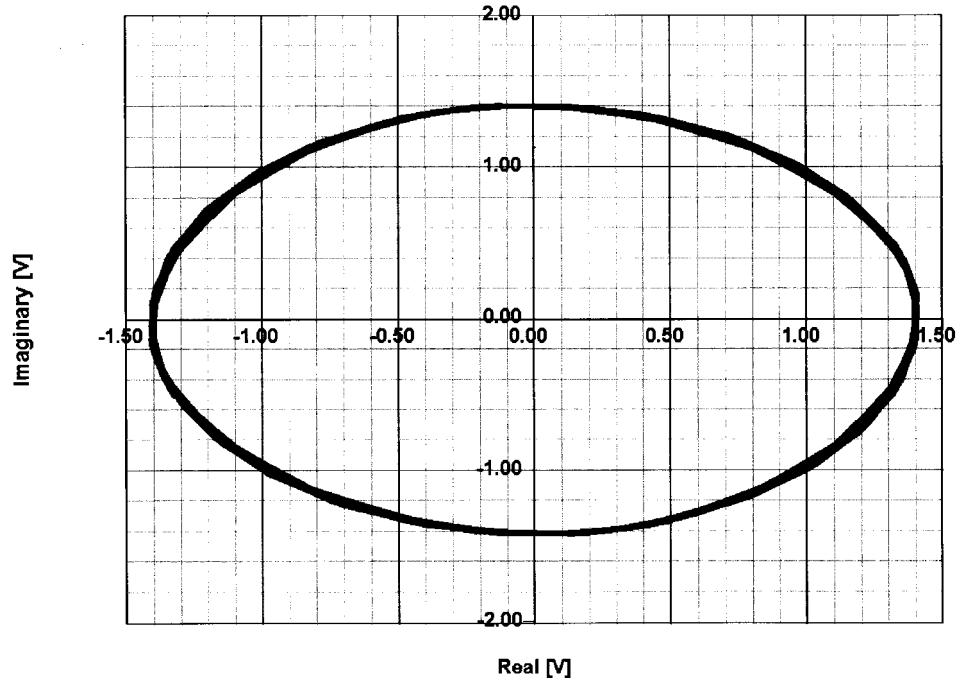


Figure 3-76 Output constellation at 10-dBm input.

BFP420 Matched Amplifier. The NE68133 has an attractive f_T of 8 GHz, but if more gain is needed, one may play with the concept of using a much "hotter" device, such as the BFP420. Recall that the NEC device was unconditionally stable, and our first worry is how the Siemens BFP420 will behave.

A study of Figure 3-77 shows that indeed the NEC device had to be stable all over and therefore was a good basis for a risk-free design. Things change dramatically with the BFP420, for which the K factor is essentially less than 1 up to a corner frequency of 2.75 GHz, where the transistor becomes unconditionally stable. Around 1.75 GHz, the K factor for the NEC transistor barely touches 1. We explained earlier that $G_{\max}$ deviating from the maximum stable gain is a more realistic number, since the conditions for maximum stable gain involve feedback and conjugate matching.

To improve the stability of the BFP420, we will now set out to use the same topology as the previous amplifier but with a resistive feedback between base and collector (Figure 3-78). A good value for this is always somewhere around 200–500 Ω. It must be remembered that this voltage feedback also stabilizes the input and output impedance and reduces the Miller effect because of the feedback being resistive and more dominant.

While the first example covered a fairly narrow frequency band, we are now attempting to cover a much wider frequency range; specifically, from 1.5 to 3 GHz reasonable gain and noise. This, by definition, requires a much higher gain-bandwidth product, and the simple matching network at the input and output provides a nonoptimal input and output matching (Figure 3-79). The feedback introduces some additional noise. This application shows the limits of single-stage UHF/SHF amplifiers with feedback.

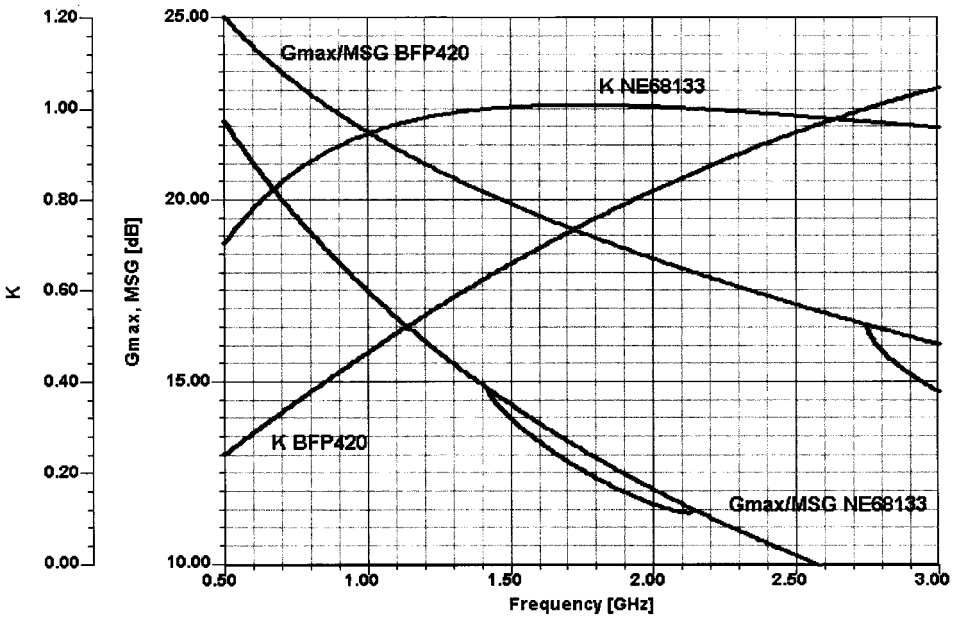


Figure 3-77 Comparison of K, $G_{\max}$, and maximum stable gain (MSG) for the NE68133 and BPF420 BJTs.

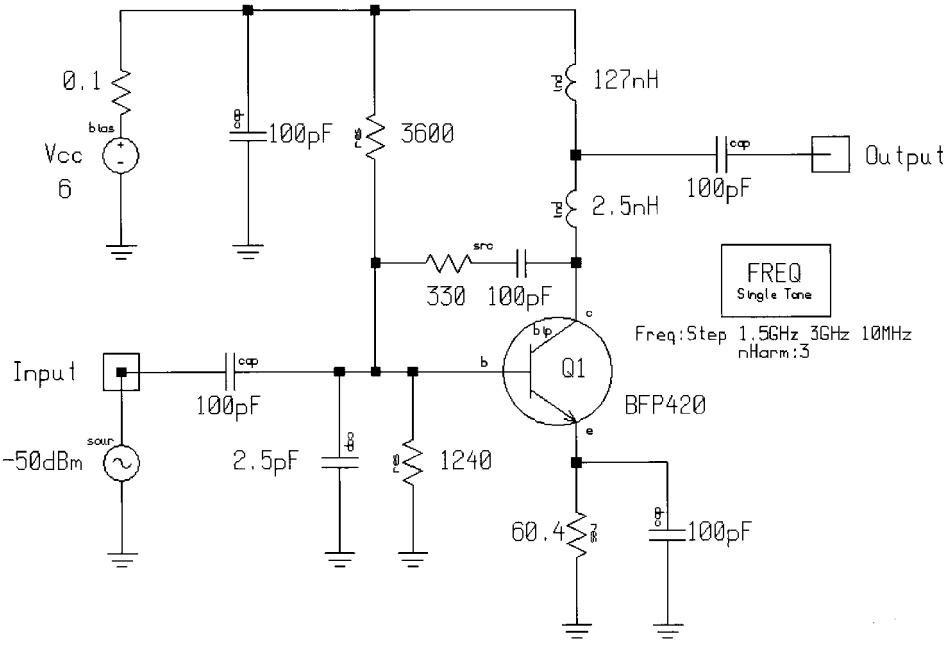


Figure 3-78 BFP420 amplifier with resistive voltage feedback from collector to base.

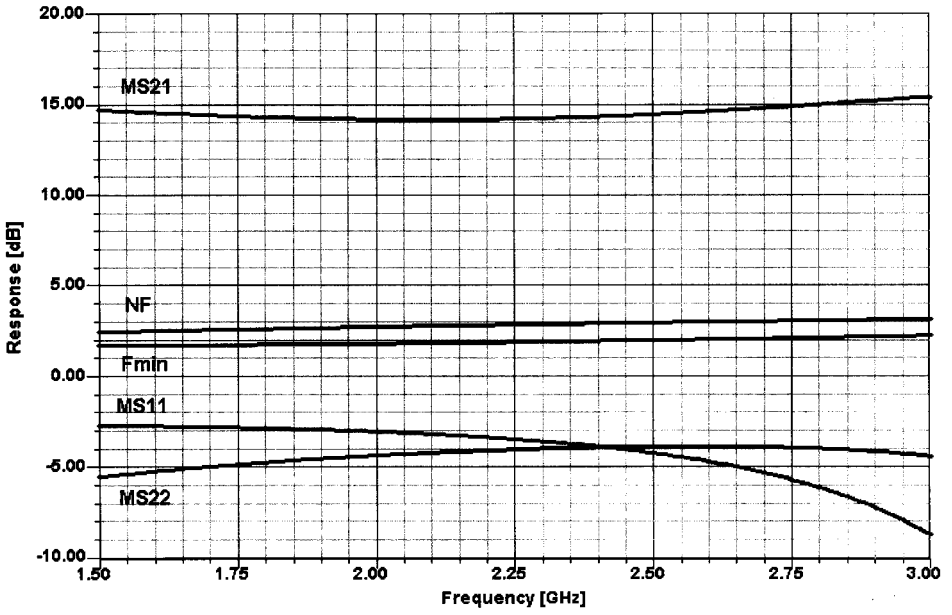


Figure 3-79 Frequency-dependent gain, matching, and noise performance of the BFP420 feedback amplifier.

Narrowband BFP420 Amplifier. Our next attempt will be to design a narrowband input stage that combines very high selectivity with good noise figure. This can only be achieved by using two tuned single-stage resonators with capacitive/magnetic coupling. The following amplifier is tailored to have a center frequency of 900 MHz. The transistor will operate at 12 V/10 mA to have a sufficient dynamic range. The input selectivity is achieved by the use of two top-coupled tuned circuits, followed by a BFP420 (Figure 3-80). The first order of business is to determine the input tuned circuit. The tuning capacitance for the two tuned circuits is arbitrarily set to a manageable 2 pF (small values have too much tolerances and larger values will require smaller inductances, resulting in lower Q). The reactance of the first tuned circuit is

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi \times 9\text{E}8 \times 2\text{E}{-}12} = 88.42\ \Omega$$

The necessary inductance for resonance will be

$$\omega L = 88.42$$

or

$$L = 88.42/\omega = 88.42/(2\pi \times 9\text{E}8) = 15.64\ \text{nH}$$

We will make both tuned circuits identical, but with different loaded Qs. By multiplying 88.42 times the loaded Q (20 for the input resonator and 50 for the output resonator), the resulting parallel-resonant resistance is $88.42 \times 20 = 1768\ \Omega$ at 900 MHz (for the input resonator) and $88.42 \times 50 = 4421\ \Omega$ (for the output resonator). These values are achieved by loading the input with the 50-Ω source and by loading the output with the input of the transistor. Therefore, the input transformation ratio $m_1 = 1768.4/50 = 35.37$. At the output, the equivalent number $m_2 = 4421/50 = 88.42$. For the purpose of simulation, and because it is extremely difficult to find the physical location of the appropriate tap on an air-wound coil for the necessary impedance, we do this "on paper" by incorporating an ideal transformer. We are also going to show later the actual transmission-line-based filter and its modeling. This filter will have a higher insertion loss based on the PC-board material used.

Using our nonlinear simulator, in small-signal ac mode, we obtain the frequency response and noise figure shown in Figure 3-81. In actually building this circuit, there is another problem besides these transformers—the top-coupling capacitor. We have used a grounded-T configuration to obtain a practically realizable value of 36 pF, whose lead inductance will not influence the circuit. (The actual coupling capacitance value for top coupling without the T configuration would have 0.1 pF. We would like to see this value repeatably realized.)

If the actual noise would be calculated relative to the output tuned circuit, the equivalent noise resistor R_n of 5 Ω would now be multiplied by m_2, resulting in 442 Ω. Fortunately, the Serenade simulator has a highly accurate BJT noise model that is both bias and temperature dependent, so the simulator will give us the answer directly.

We have not put any matching circuit at the output and leave this task to the reader. In practice, the slight tilt in the MS_{21} curve in Figure 3-81 would be tuned out. Also note that the minimum noise figure ($F_{\min}$) and the actual 50-Ω noise figure (NF) agree somewhat off center. This has to do with the Miller effect, which results in the well-known difference between noise matching and input-gain matching. By varying the coupling capacitor to move from the overcoupled response, which gives a low noise figure, to a critically coupled

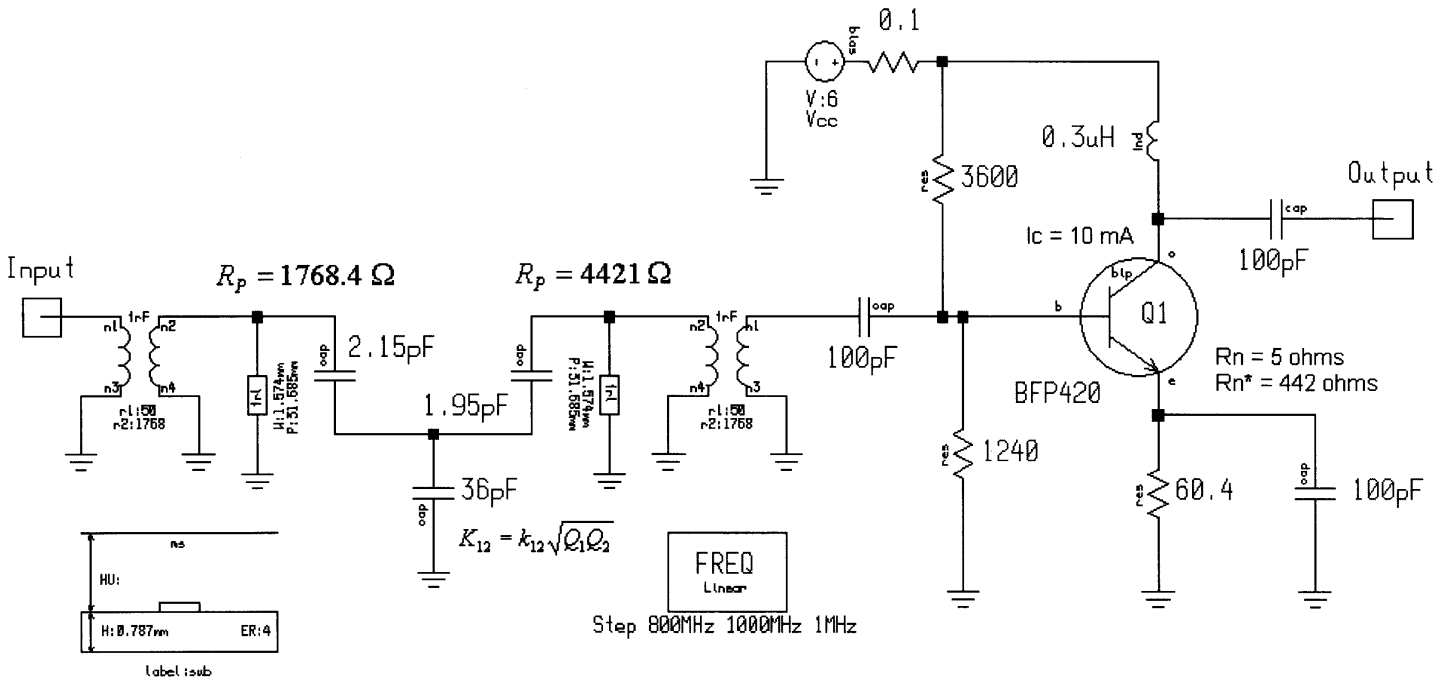


Figure 3-80 BFP420 amplifier with narrowband input filtering. Ideal transformers are used to match the filter's input and output to the input port and the BFP420 base, respectively.

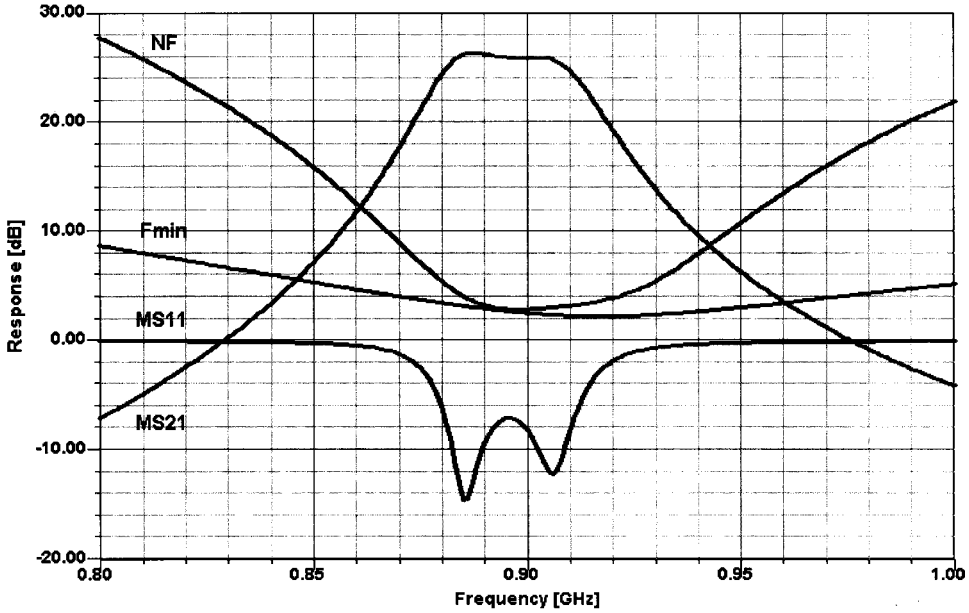


Figure 3-81 Frequency-dependent gain, matching, and noise performance of the narrowband BFP420 amplifier.

response ($K = 1$), we get significantly better selectivity and narrow bandwidth (at the obvious expense of the noise figure) as shown in Figure 3-82.

Since we somewhat compromised by not providing the reader with an actual input and output coupling that can be built, and furthermore, because the inductance in question (15.26 nH) is hard to realize with a high Q, we now show the printed circuit approach for the input filter. By using a Teflon-based material for the PC board, the losses would be less; however, the radiation would be more because ε would go down to 2.1 from the value we used (4). It is also convenient to introduce this filter to show that at higher frequencies, where lumped elements are really no longer available, one has to switch to distributed elements. A list of the commonly used distributed elements will be shown later; the transmission line and T junction are probably the most frequently used elements, followed by spiral and rectangular inductors, and bends.

The actual implementation of the input filter on a printed circuit board requires distributed elements (Figure 3-83). In practice, one has two parallel coupled lines with a tap somewhere toward the ground connection and, for modeling accuracy, we have to introduce a T junction with the appropriate impedance of 50 Ω for the input and a different impedance (if necessary) for the output. The coupling mechanism now is magnetic coupling, since no one can physically change the distance between the transmission lines by tuning some mechanical elements, one has to get it right the first time. This can only be achieved with a high-precision simulator. The reason why these simulators are so expensive is the investment of research time and personnel to produce valid models over a huge frequency and impedance range. Since we are now curious about the frequency response, we have plotted in Figure 3-84 the frequency response, which shows a higher (expected) insertion loss. The PC-board material is partially responsible for this.

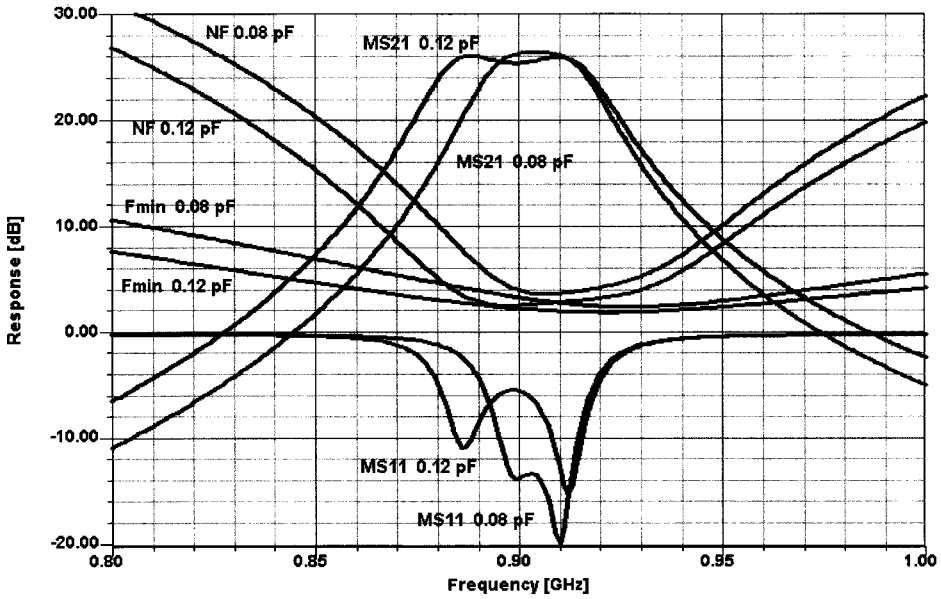


Figure 3-82 Comparison of filter responses with critical coupling (0.08-pF response) and overcoupling (0.12-pF response).

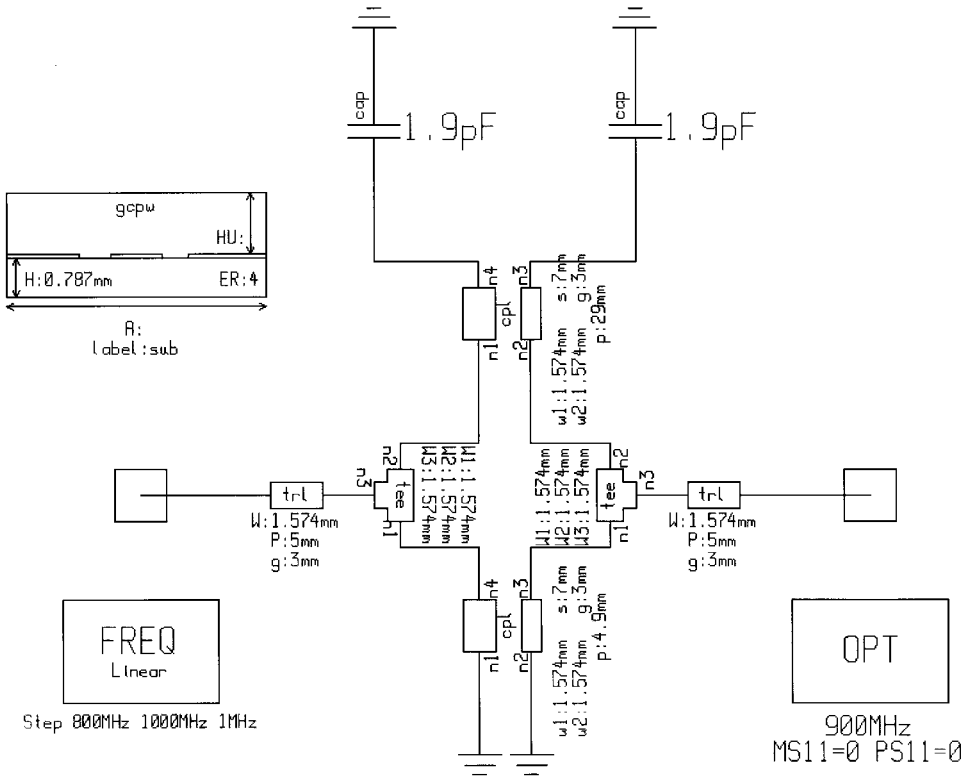


Figure 3-83 Input filter using distributed-element resonators and matching.

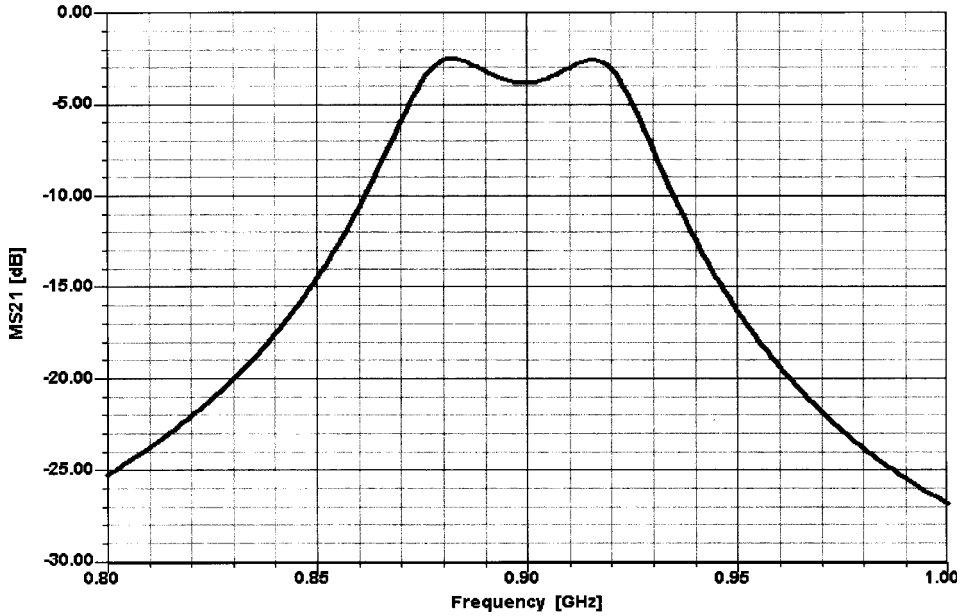


Figure 3-84 Gain versus frequency response for the distributed-element filter.

At the end of this chapter we will have a chance to discuss passive elements as needed for microwave frequencies, and one should consider frequencies above 1500 MHz as microwave frequencies.

GaAsFET Feedback Amplifier. The implementations we have shown so far have been based on bipolar transistors. Figure 3-85 shows a single-stage feedback amplifier using a GaAsFET. The feedback is based around an *LRC* network and some attempts at matching at the input and output are made. Figure 3-86 shows its frequency-dependent gain, matching, and noise performance.

3-2-3 High-Gain Amplifiers

In most cases, high-gain amplifiers receive an already amplified signal, which means we are less concerned about the signal-to-noise ratio but remain concerned about distortion. To reduce distortion, we have two options: (1) increase the dc current in the active device, which always improves the current-related distortion; and (2) add some voltage feedback, which reduces the voltage distortion. In simple terms again, only current feedback can reduce the current-related distortion of a square-law or exponential transfer characteristic and only voltage feedback can reduce the effect of overdrive as a function of the load, but only to some degree. Its main purpose remains to reduce the input impedance to useful values, specifically for FETs, since current feedback increases the input and output impedances. Some ultrasensitive circuits still require Class A operation, but now at higher current levels; these are typically CATV applications, and there are several transistors available that can be operated at 5–15 V but at 60 mA or higher. Because of the higher power consumption, this excludes handheld cordless or cellular telephones. The same circuit as used before (Figure 3-85) for

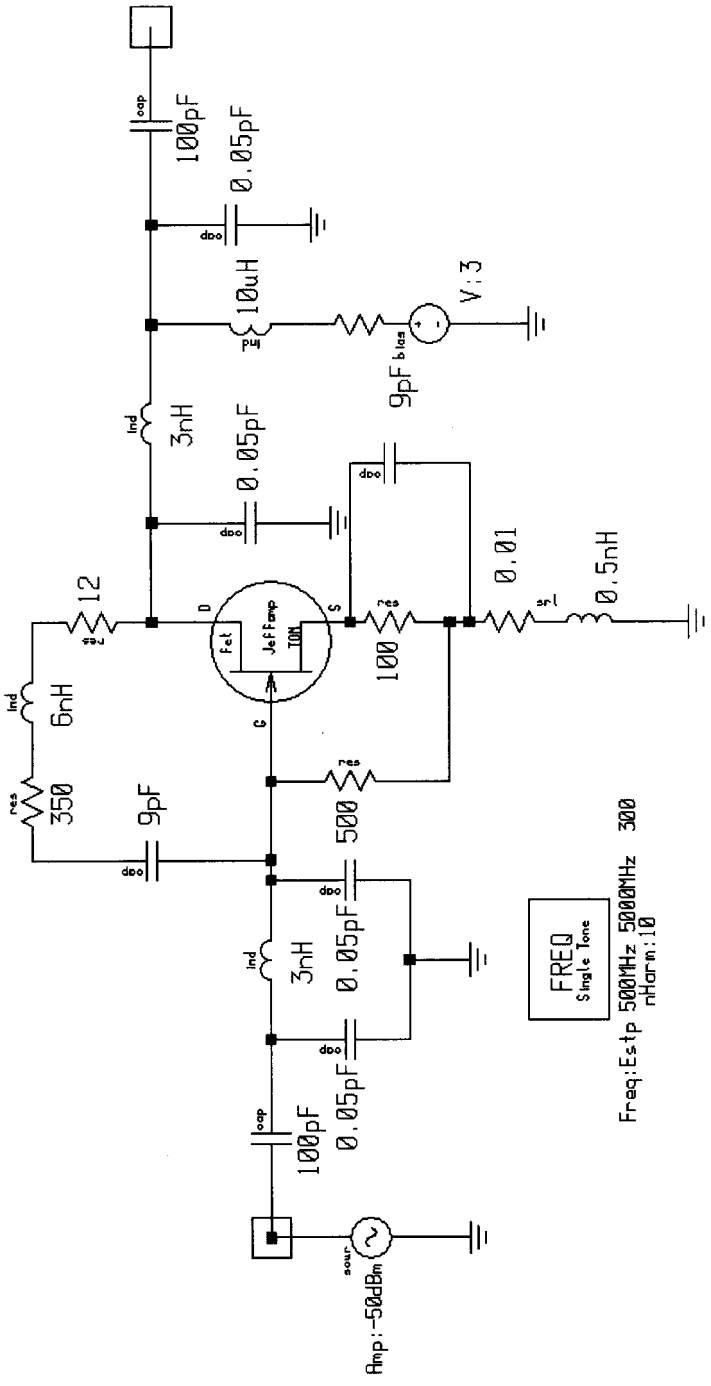


Figure 3-85 GaAsFET feedback amplifier.

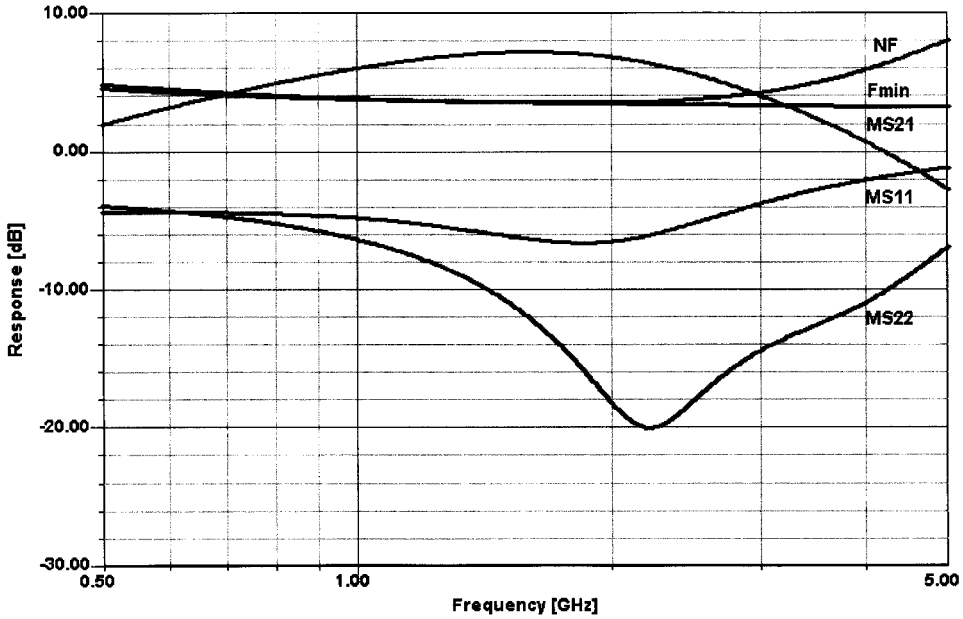


Figure 3-86 Frequency-dependent gain, matching, and noise responses for the GaAsFET amplifier.

Class A operation now can be operated at a higher current, which automatically increases the gain and the noise figure and reduces distortion. This is based on the fact that the transconductance (g_m) increases at higher currents. The second transistor in the TI circuit achieves this by being in reality a larger device, which could be modeled by several transistors, like the first stage, in parallel. This automatically increases the dc current by the scaling factor. The same is even more true for the output stage. Likewise, bipolar transistors can also be scaled. A visual inspection of some transistors in their packages will reveal that they are really not just one transistor, but several operating in parallel, each with its own ballast resistors (for bipolar transistors) or other means (for FETs).

Using tuned circuits rather than wideband matching, it is possible to go to Class AB or B operation. Class AB is really a hybrid between A and B, and Class B is probably better known. Figure 3-63 shows the Class B operating point. Most Class B stages are operated in push–pull because, together with the tuned circuit, one-half of the cycle is supplied by the other transistor. The definition of Class B operation is a conducting angle of 90° (see Figure 3-63). At this point it should be noted that some older texts define the conducting angle as being twice the value we have defined in this chapter.

Without going into further details, the highest efficiency in Class B operation is 78.5%. Based on the fact that modern wireless applications generally use a hybrid of amplitude and angle modulation, pure Class C operation finds little application today because of the severe nonlinear distortion it introduces. A Class B stage has less "power gain" than a Class A stage but provides much more output power. We have already mentioned the problem associated with characterizing a Class A stage in terms of power gain. That's the only time Kirchhoff's equations are valid, meaning that the sums of all currents and voltages have to be zero. Because Kirchhoff's equations do not consider nonlinearities and harmonic contributions,

the very moment we move into considering medium- and higher-power amplifiers, the nonlinearities will dominate and the performance of the stage is best analyzed with a high-quality circuit simulator with the appropriate dynamic range and good modeling capability. Without question, this applies to Class C operation also. As pointed out previously, the high-gain amplifier is optimized for gain rather than noise figure and typically requires more collector or drain current. This can easily be seen by rebiasing the FET amplifier as shown above to a significantly higher value of 33 mA. This is accomplished by changing the source resistor down to 3 Ω. As a result of this, we obtain the gain, matching, and noise figure responses shown in Figure 3-87.

It may be a surprise to the reader that the actual noise figure now is less than that shown in Figure 3-86. This device is now operating at 33 mA; the reason for this has to do with the fact that if the transconductance is too low, the noise figure also increases, and the NF curve for the FET has a similar-looking current-dependent minimum—a curve that is frequently not supplied by the manufacturer. This amplifier, now operating at 33 mA, should have quite good large-signal capabilities; we will examine this shortly.

Another interesting piece of information is the display of the dc I–V curves with the "load line." Since we have reactances in the circuit, we do not see a line but rather an ellipse, indicating the storage of energy. (The actual load line would be a line through the two foci of the ellipse.) This curve allows us to see immediately that at this operating point the transistor is neither going into current saturation nor becoming voltage limiting. The input power can be increased further until such a condition exists. Figure 3-88 shows the dc I–V curve for the amplifier shown in Figure 3-85.

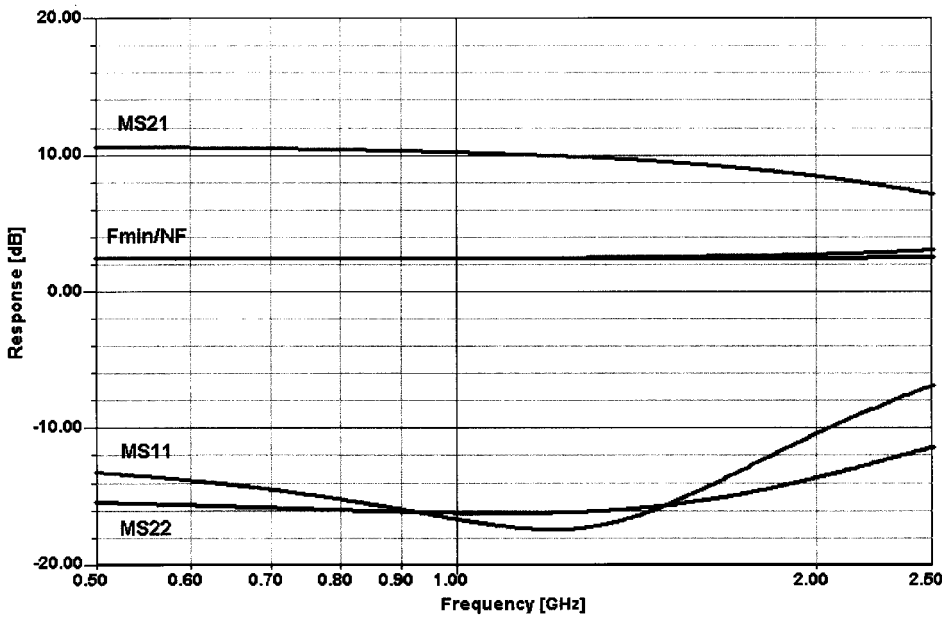


Figure 3-87 Frequency-dependent gain, matching, and noise performance for the GaAsFET amplifier of Figure 3-85 rebiased for a drain current of 33 mA.

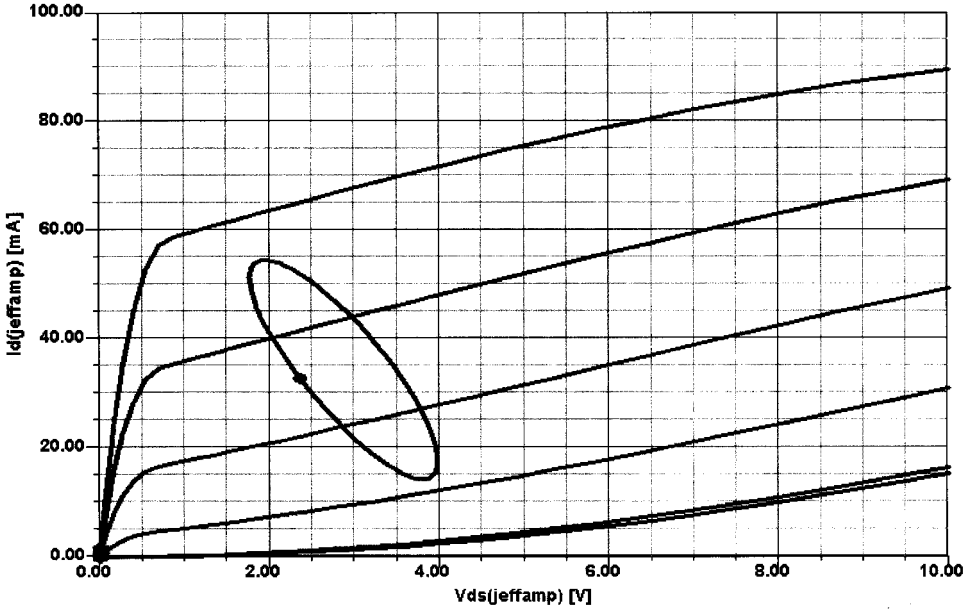


Figure 3-88 The dc *I–V* curves for the GaAsFET operating at I_D = 33 mA, including the ac load line. The load line is an ellipse rather than a straight line for the reason described in the text.

Besides looking for the adjacent-channel power ratio (ACPR), the other traditionally important points are third-order intercept point and the equivalent of a multitone simulation, as used for CATV application but relevant to the multisignal environment of wireless applications. Figure 3-89 shows the determination of the amplifier's third-order intercept point. It is most important for the reader to understand that very low values for input and output level have to be used to determine the crossover point. If this advice is not followed, totally erroneous numbers can occur.

Since an antenna supplies a large number of signals (this is true of handheld radios as well as base-station sites), a three-tone standard has been adopted as a good approximation of multisignal operation. This standard is in accordance with the German DIN 45004E,3.3 three-tone measurement. The requirement is that the intermodulation distortion products at the output are better than 60 dB suppressed relative to the largest of the three tones. One tone is used as the reference, and the other two tones are 6 dB less. This measurement has been derived from television, whereby the video transmitter, based on the vestigial sideband power, is 6 dB above the two sidebands produced by the FM sound transmitter using a modulation index of 5. The result of a three-tone test can be seen in Figure 3-90. This amplifier clearly meets the DIN specification for linearity at input levels up to –9/–15 dBm at the input.

Bipolar transistors are also quite popular for this purpose. Figure 3-91 shows a single-stage amplifier with an intrinsic transistor and all the other "hidden" intrinsics, normally not visible, added. The three diodes belong to those intrinsic elements, which are needed for accurate microwave modeling but which are not provided by the standard nonlinear models. The reader will see that we went into great detail to model all the necessary parts. This feedback amplifier, whose frequency response is shown in Figure 3-92, has been quite

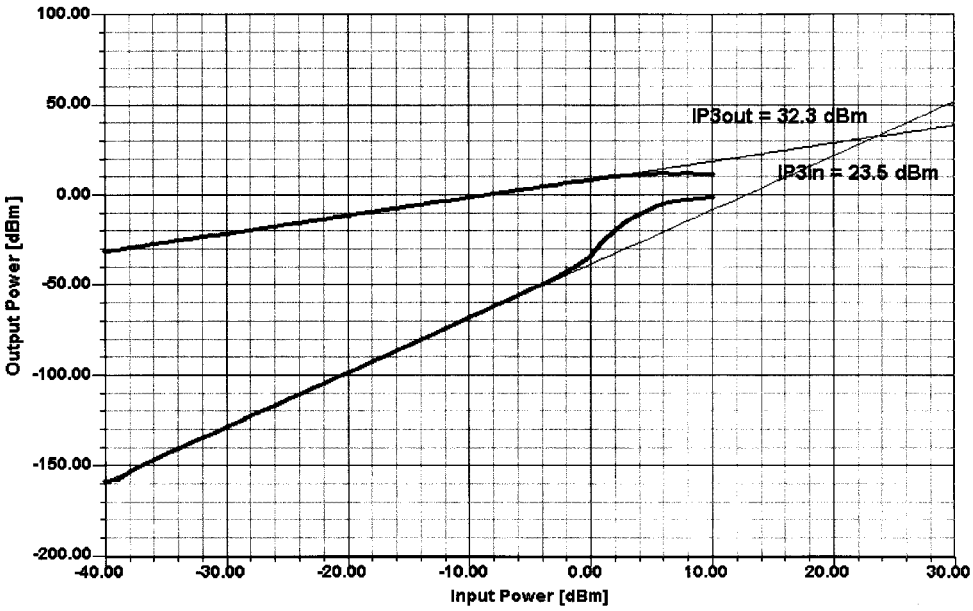


Figure 3-89 Determination of the third-order intercept point based on a two-tone measurement using tones spaced at 5% of the operating frequency (1.8 GHz).

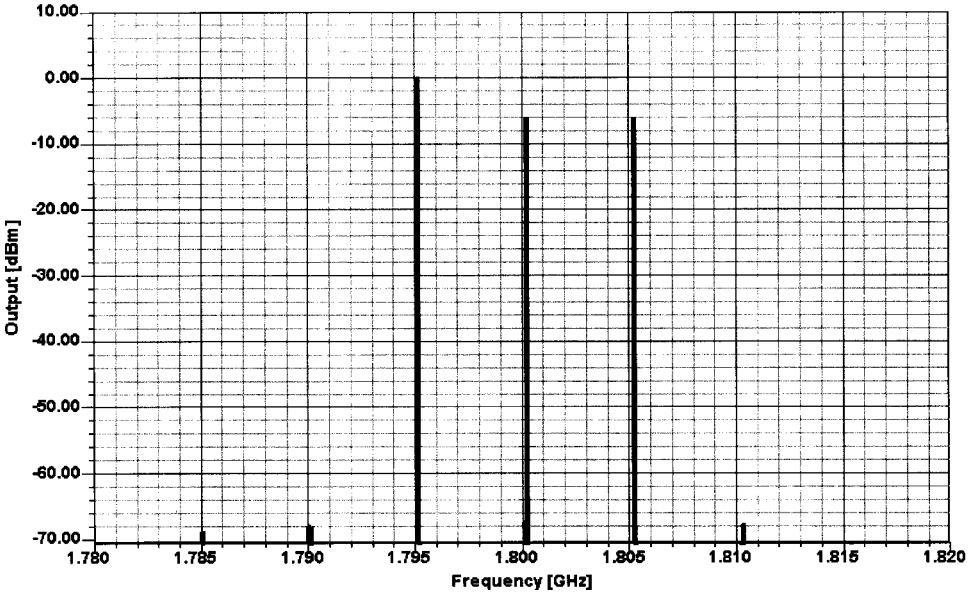


Figure 3-90 Simulated three-tone analysis of the GaAsFET amplifier with $I_D = 33$ mA.

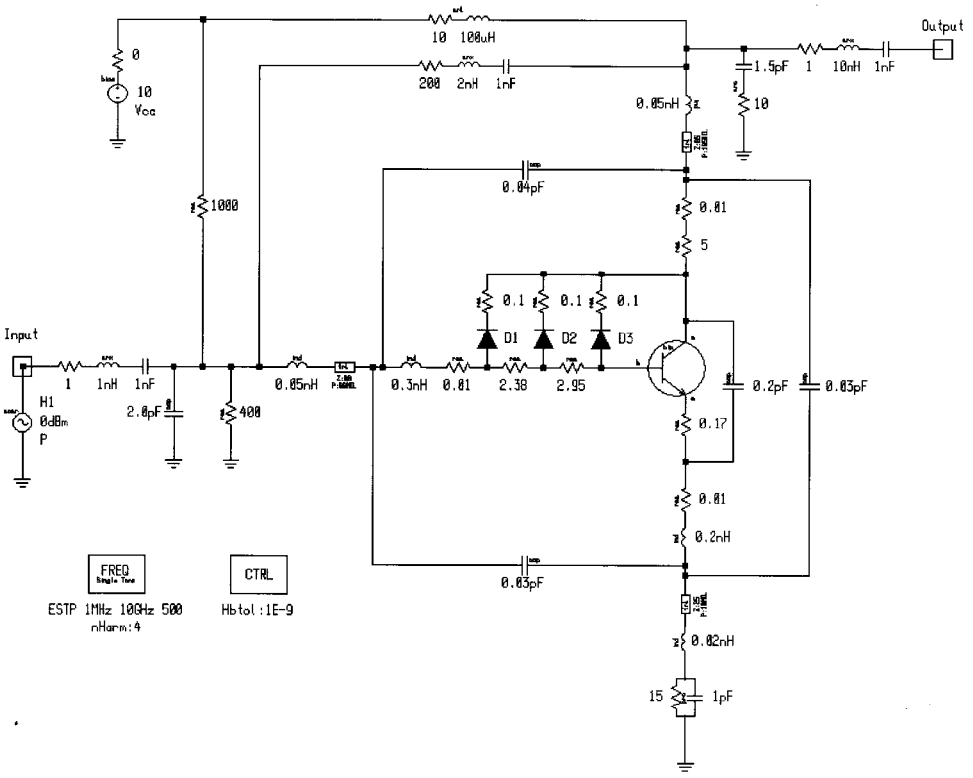


Figure 3-91 A high-linearity single-transistor amplifier modeled in great detail.

difficult to design, but there are monolithic circuits available where the higher gain–bandwidth product is being used to advantage.

Figure 3-92 shows the predicted frequency-dependent gain, matching, and noise figure at a bias point of 108 mA at 10 V. Using the same DIN three-tone test procedure, Figure 3-93 shows the expected good result.

Since this test requires two equal signals and one larger signal, we need to expect one IMD product larger than the rest. We have set our window for a dynamic range of 70 dB, and yes, in the lower left corner, we can see the product between the first two signals, while the other products are below the –50 dB level.

As a little task, Figure 3-94 shows a simple amplifier, based on a modern Siemens transistor, which is a feedback amplifier with fewer modeling details surrounding the transistor. Its associated frequency-dependent performance is shown in Figure 3-95. Interested readers are invited to improve its performance by using their skills and intuition.

Permitting higher currents and actually combining two transistors leads to our first MMIC. The following simulation is based on the MGA64135. Needless to say, there are some compensation networks inside the MMIC and distributed elements that affect the frequency response. However, with this arrangement, it is possible to obtain an appreciable gain at our wireless frequencies. The noise figure, based on the feedback, however, it not too impressive. This is a problem that occurs with most feedback amplifiers, which try to minimize the distortion and yet would like to keep the noise figure down. Figure 3-96 shows the schematic

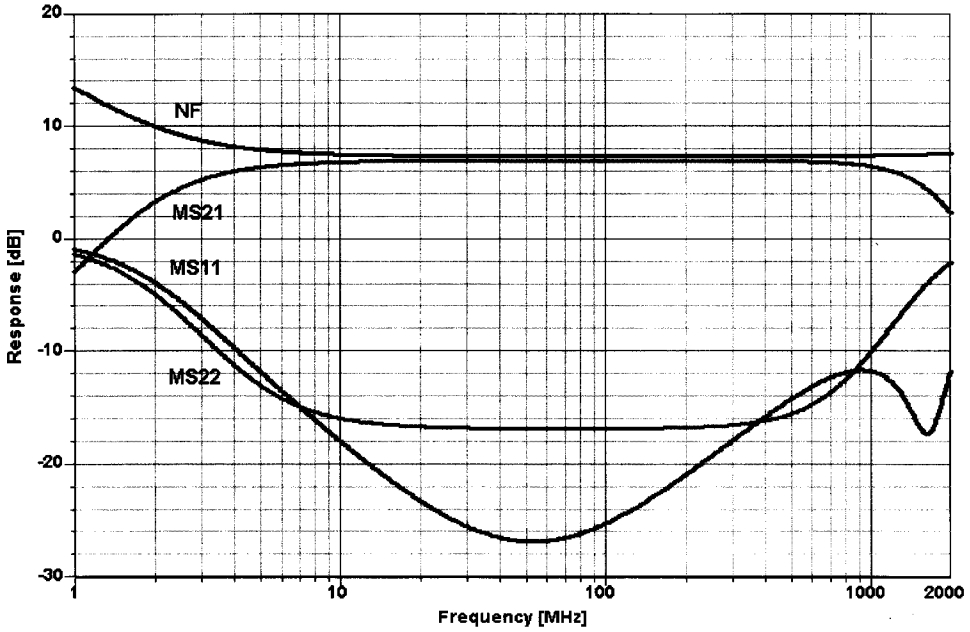


Figure 3-92 Simulated frequency-dependent gain, matching, and noise figure performance of the circuit in Figure 3-91.

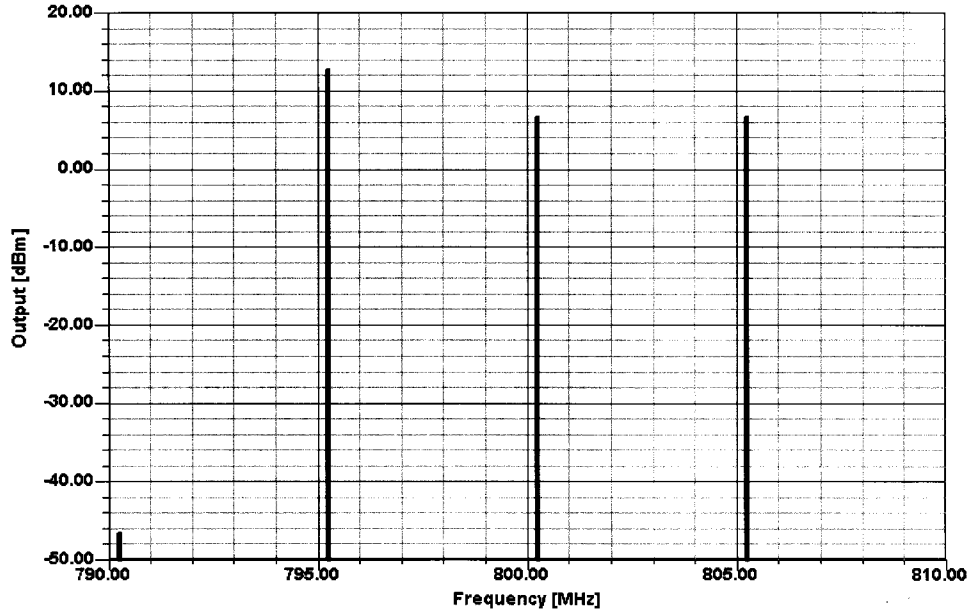


Figure 3-93 Simulated three-tone output spectrum of the amplifier in Figure 3-91.

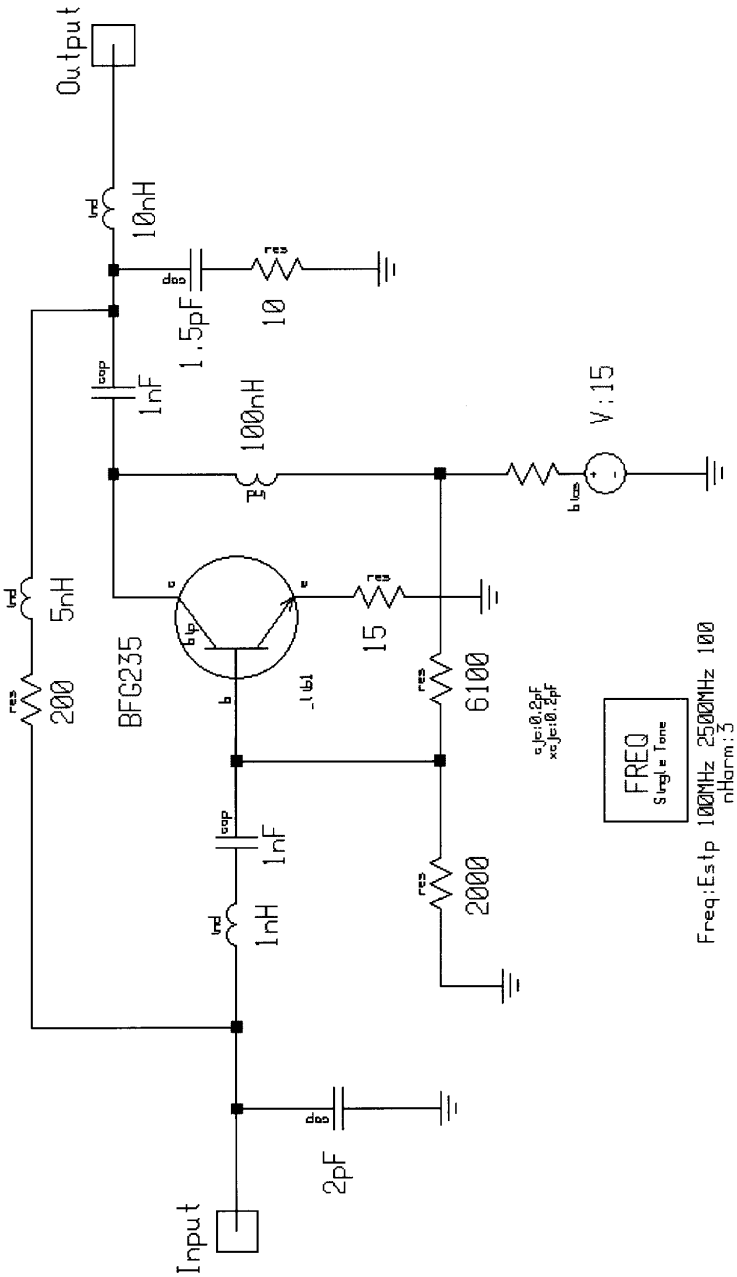


Figure 3-94 A high-linearity amplifier using the Siemens BFG235.

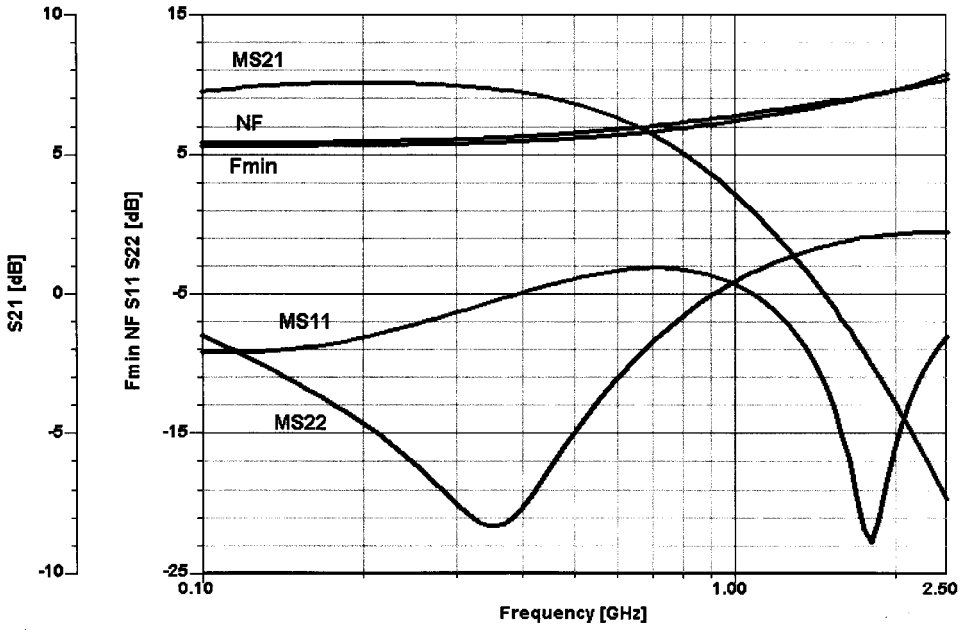


Figure 3-95 Frequency-dependent gain, matching, and noise performance of the amplifier in Figure 3-94.

of the GaAs MMIC; Figure 3-97 shows its frequency-dependent gain, matching, and noise performance; Figure 3-98 shows its two-tone response, which is used to obtain the intercept point ($IP_{3,in}$ = 8.5 dBm, $IP_{3,out}$ = 22 dBm); and, finally, Figure 3-99 shows the three-tone test results at an input RF level of –21/–27 dBm. While the gain–bandwidth product is considerably larger than the single-stage amplifier we examined before, the intercept point and three-tone performance do not reach the same values. As a general rule, the dynamic

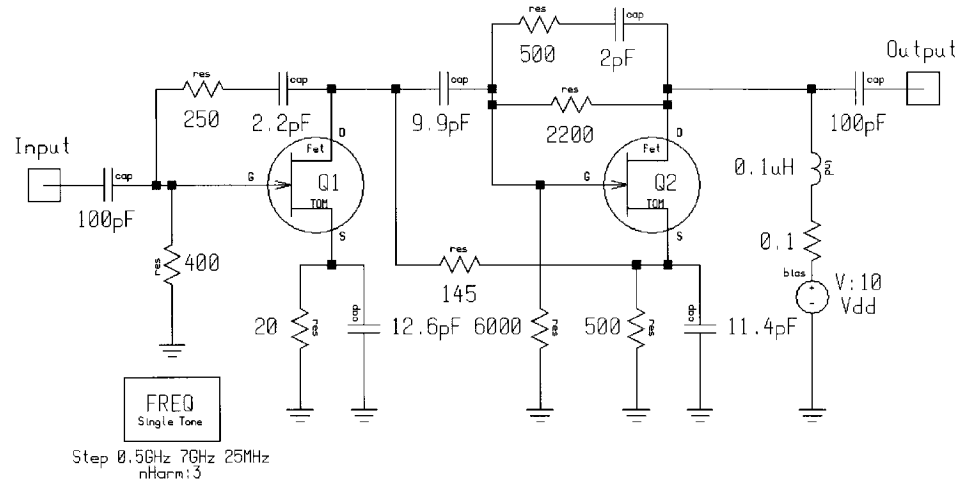


Figure 3-96 Simulation schematic for the MGA64135 MMIC amplifier. Q1 runs at 17.4 mA; Q2 at 32.2 mA.

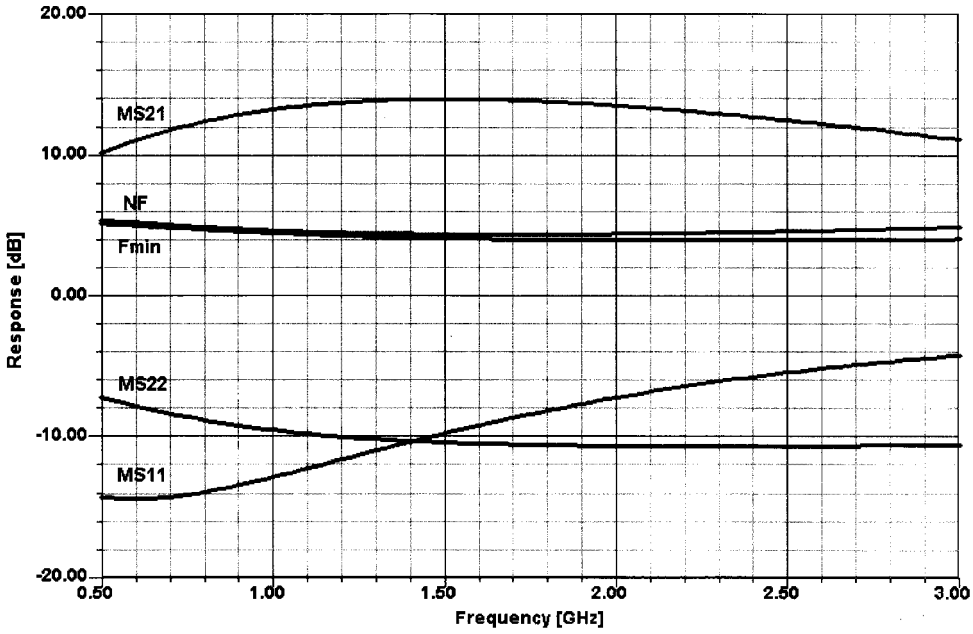


Figure 3-97 Frequency-dependent gain, matching, and noise figure performance for the MGA64135 MMIC amplifier.

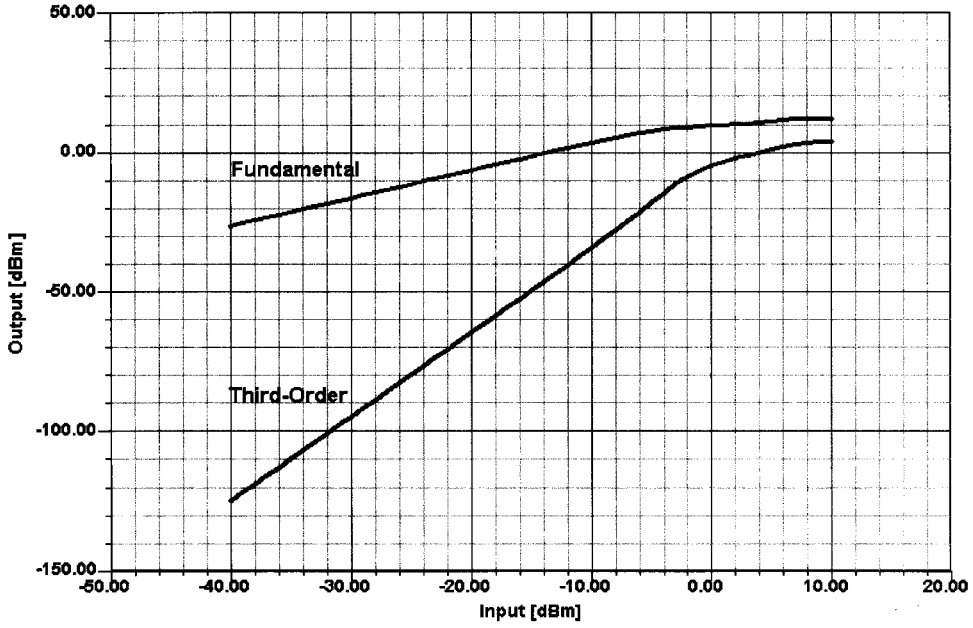


Figure 3-98 Simulated two-tone testing of the MGA64135 MMIC amplifier. $IP_{3,in}$ = 8.5 dBm and $IP_{3,out}$ = 22 dBm.

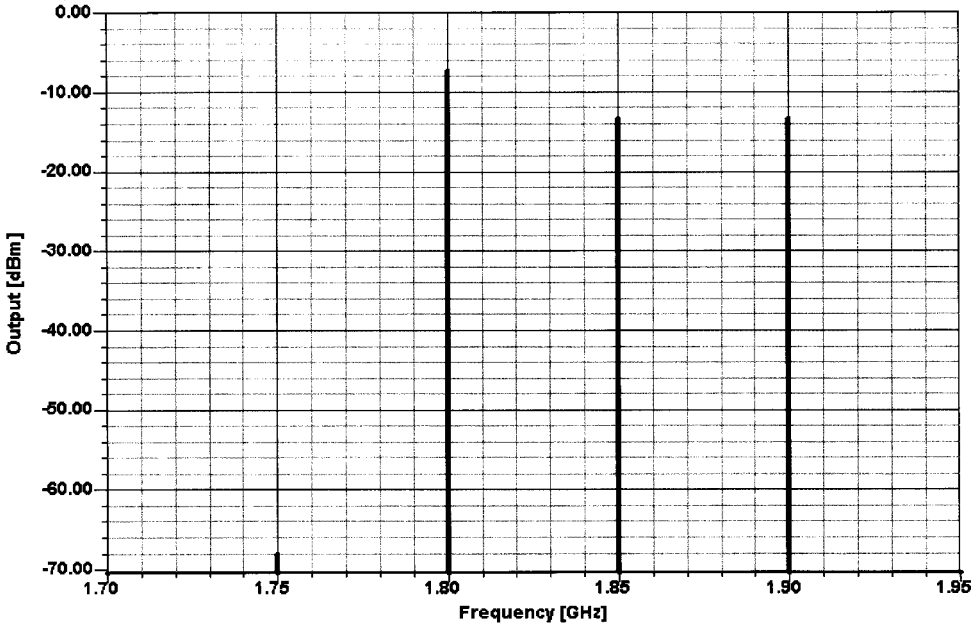


Figure 3-99 Simulated three-tone test of the MGA64135 MMIC. The test tones were 1.80 GHz (–21 dBm), 1.85 GHz (–27 dBm), and 1.90 GHz (–27 dBm). The 1.75-GHz IMD product is down 60.5 dB relative to the 1.80-GHz signal.

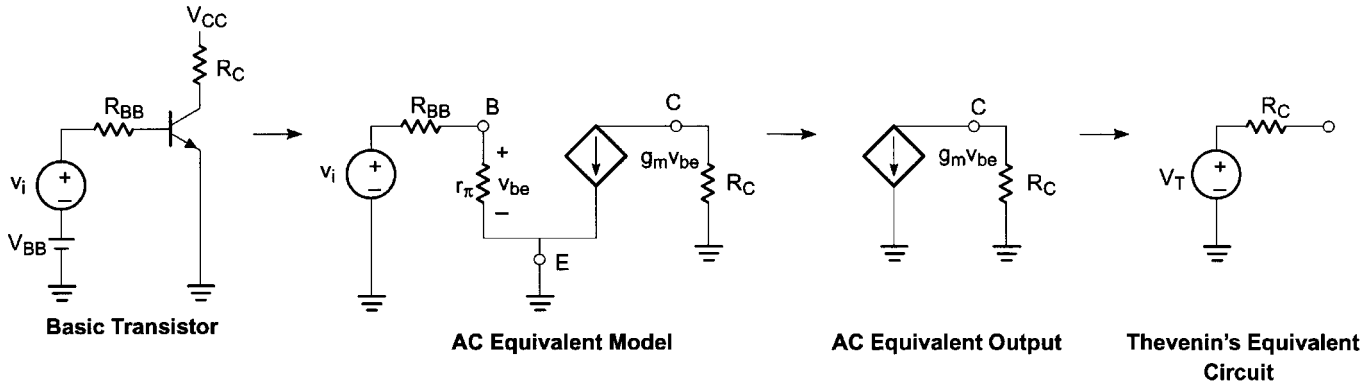
range is directly proportional to the power dissipated in the devices: specifically, the current. As explained earlier several times, the first-order nonlinearity comes from the voltage-to-current transfer characteristic, such as the transconductance itself, and can only be improved by a combination of resistive feedback in the emitter/source and by high currents. At these dc levels, the base–emitter (gate–source) junction gets linearized because the differential values are getting smaller compared to the parasitic loss resistances, including external dc/RF feedback. Once again, the voltage feedback reduces the voltage clipping but is initially more responsible for the input and output matching.

3-2-4 Low-Voltage Open-Collector Design [22–25]*

Why R_C Acts Like a Source Resistor. An open-collector output allows a designer the flexibility to choose the value of the R_C resistor. Choice of this resistor value not only sets the dc bias point of the device but also defines the source impedance value. Figure 3-100 shows the ac model of the transistor. Converting the output structure by applying a Norton and Thevenin transformation, one can conclude that R_C becomes the source resistance. Thus, by choosing R_C to be equal to the load, maximum power transfer will then occur.

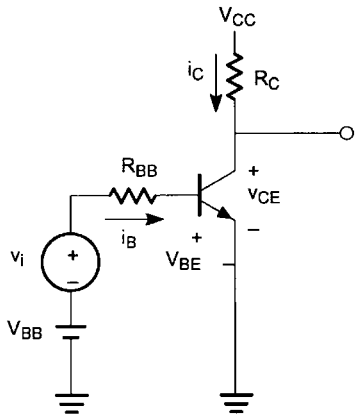
Figure 3-101 shows an active transistor with collector and base resistors. From basic transistor theory, Eq. A in Figure 3-101 is generated and has the same form as the general equation for a straight line ($y = mx + b$).

*Based on portions of the Philips Semiconductors/Signetics RF Communications Products Application Note AN1777, "Low-Voltage Front End Circuits: SA601, SA620," August 20, 1997. Used with permission.



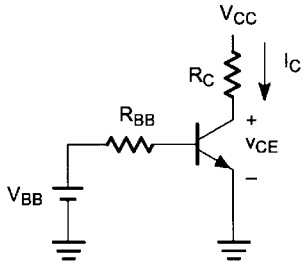
Symbol Convention:
DC: represented by uppercase symbols with uppercase subscripts, i.e., I_C
AC: represented by lowercase symbols with lowercase subscripts, i.e., i_C
Combined AC & DC: represented by lowercase symbol with uppercase subscript, i.e., i_C

Figure 3-100 *Rc* as source resistor.



DC Model

- Kill AC sources
- Open capacitors
- Short inductors

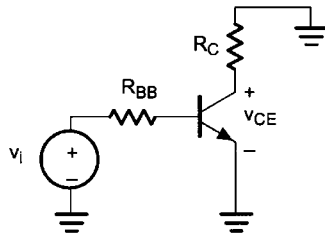


$$I_C = \left(\frac{-1}{R_C}\right) V_{CE} + \frac{V_{CC}}{R_C}$$

Eq. A

AC Model

- Kill DC sources
- Short capacitors
- Open inductors



$$i_C = \frac{-1}{R_C} (v_{CE})$$

Eq. B

Figure 3-101 Basic transistor analysis.

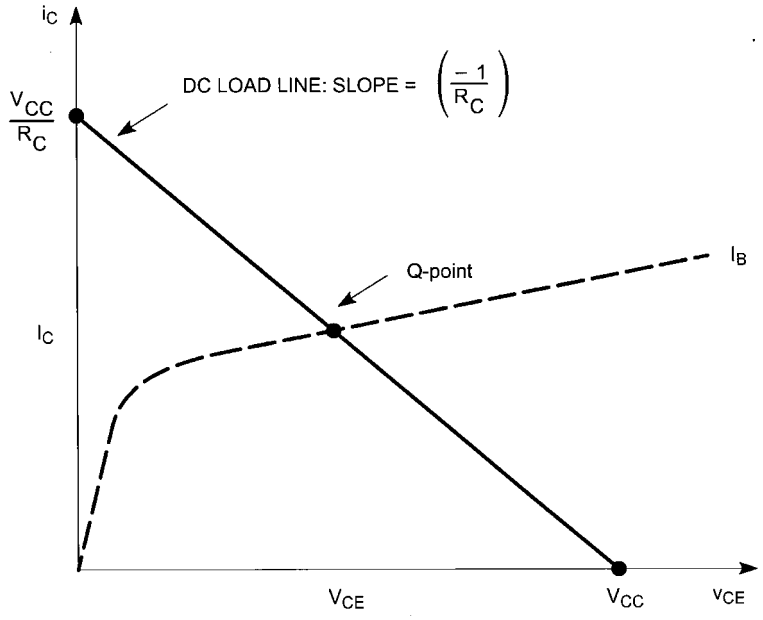


Figure 3-102 Load line and Q-point graph.

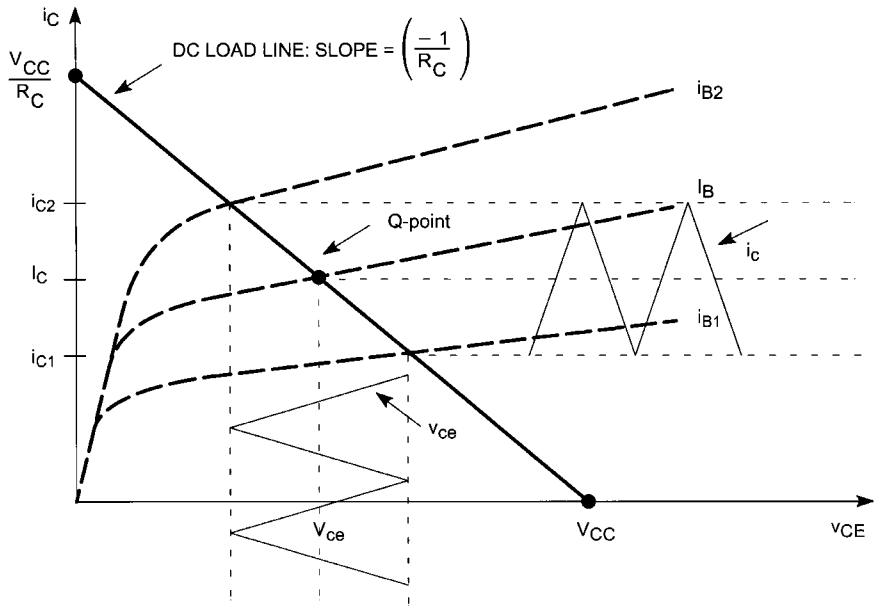


Figure 3-103 Graphical analysis for the circuitry in Figure 3-101.

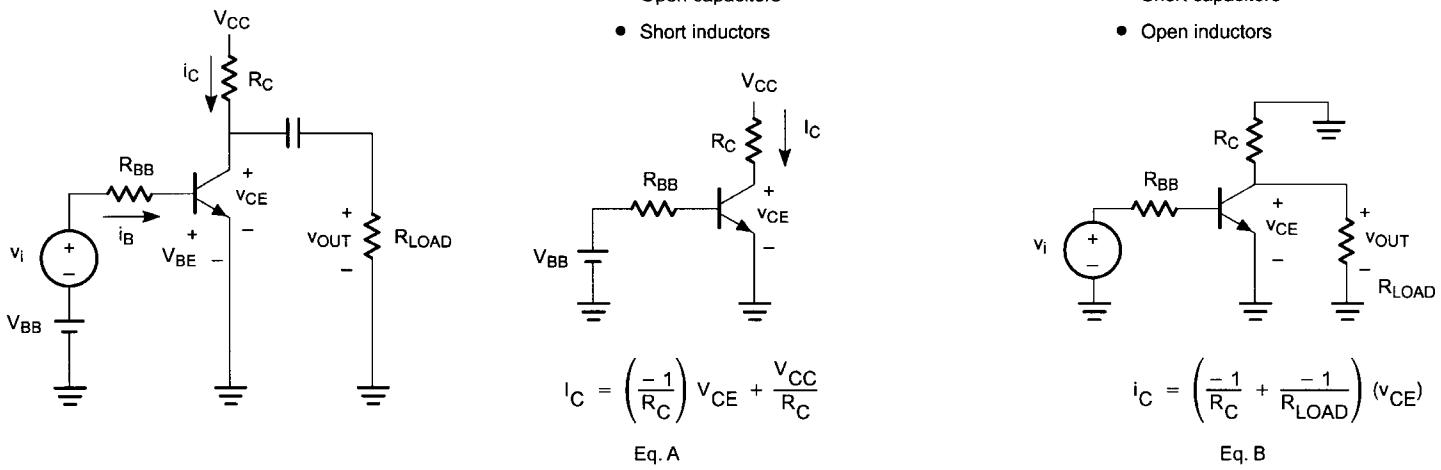


Figure 3-104 Basic transistor analysis with R_{LOAD}.

The slope of the dc load line is generated by the value of the collector resistor ($m = (1/R_C)$ and is shown in Figure 3-102. For a given small-signal base current, the collector current is shown by the dotted curve. The intersection of the dotted curve and the dc load line is called the *quiescent point* (Q-point) and dc bias determinant. The location of the Q-point is important because it determines whether the transistor is operated in the cutoff, active, or saturation region(s). In most cases, the Q-point should be in the active region because this is where the transistor acts like an amplifier.

Figure 3-103 shows the ac collector–emitter voltage (V_{CE}) output swing with respect to an ac collector current (I_C). Collector current is determined by the ac voltage presented to the input transistor's base (v_i) because it affects the base current (i_b), which then affects i_c. This is how v_i is amplified and seen at the output. Recall that this is with no external load (R_{LOAD}) present at the collector. Since no load is present, the slope of the ac load line is identical to that of the dc load line as seen in Eqs. A and B of Figure 3-101 ($m = -1/R_C$).

Open Collector with R_{LOAD}. A filter with some known input impedance is a typical load for the output of the transistor. For simplicity, we will assume a resistive load (R_{LOAD}) and neglect any reactance. Since a resistive load is used (see Figure 3-104), the ac output swing is measured at V_{OUT} or V_{CE}.

A dc blocking capacitor is used between the R_{LOAD} and the V_{CE} output to assure that the Q-point is not influenced by R_{LOAD}. It is also necessary to avoid passing dc to the load in applications where the load is a SAW filter. However, R_{LOAD} will affect the ac load line, which is seen in Eq. B in Figure 3-104. Note that the V_{CE} voltage swing is reduced; thus, the V_{OUT} signal is reduced (see Figure 3-105).

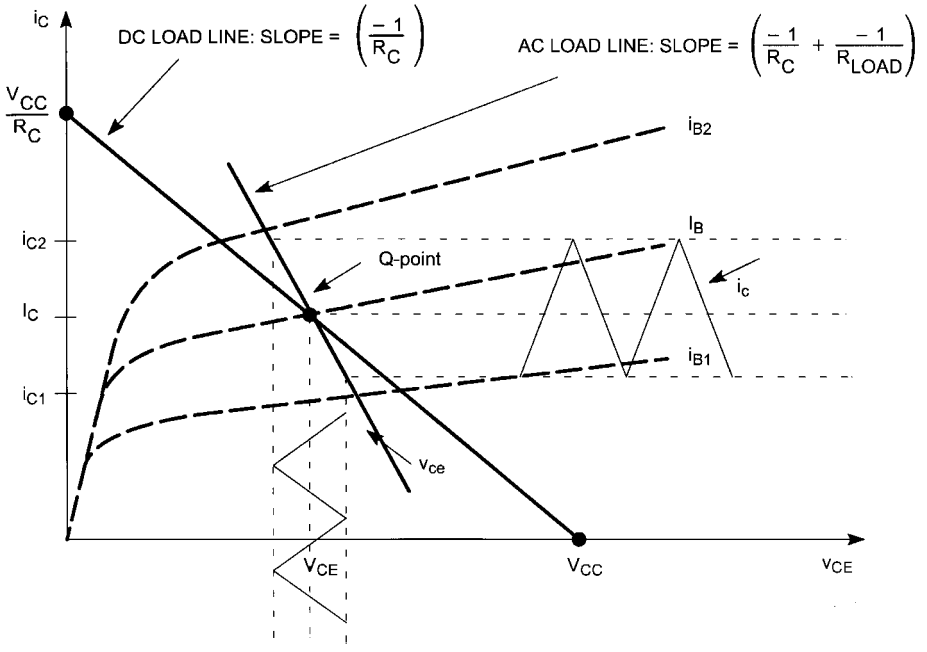


Figure 3-105 Graphical analysis for the circuitry in Figure 3-104. The ac load line intersects the Q point, and the slope is steeper than the dc load line due to R_{LOAD}, which makes V_{CE} smaller.

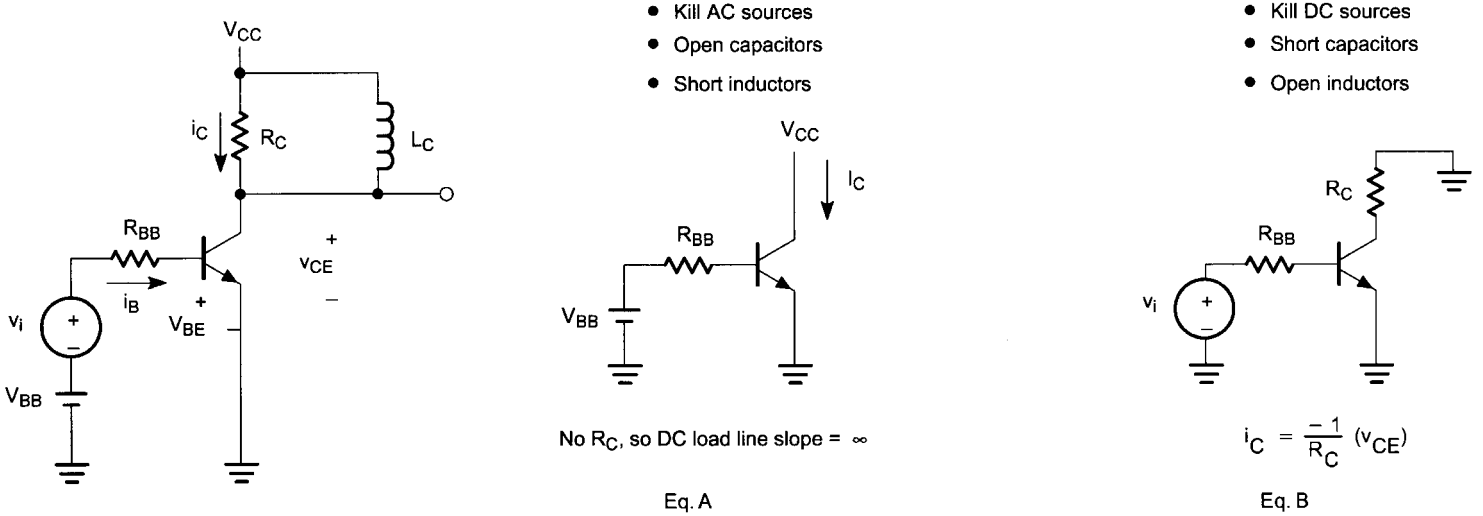


Figure 3-106 Basic transistor analysis with inductor added to collector.

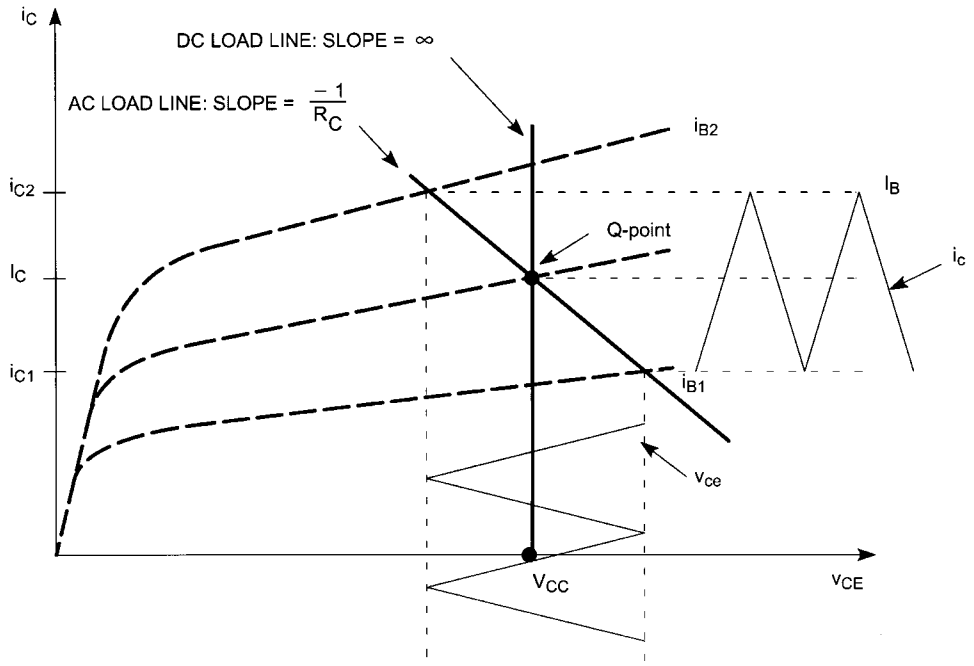


Figure 3-107 Graphical analysis for the circuitry in Figure 3-106.

Since the values of R_C and R_{LOAD} affect the ac load line slope, the value chosen is important. The higher the impedance of R_{LOAD} and R_C, the greater the ac output swing will be at the output, which means more conversion gain in a mixer. This is due to the slope getting flatter, thus allowing for more output swing.

Open Collector with Inductor (L_C). Adding an inductor in parallel with R_C can increase the ac output signal V_{CE}. Figure 3-106 shows the dc and ac analyses of this circuit configuration. In Eq. A in Figure 3-106, there is no R_C influence because the inductor acts like a short in the dc condition. This means the slope of the dc load is infinite and causes the Q-point to be centered around $V_{CC,}$ thus moving it to the right of the curve. The ac load line slope is set only by R_C because no load is present. Note that it has the same ac load line slope as the first condition in Eq. B of Figure 3-101.

Referring to Figure 3-107, one might notice that the base current (ac and dc) curves spread open as V_{CE} increases. This is caused by a noninfinite Early voltage (see Figure 3-108), which causes the collector current to be dependent on V_{CE}. Taking advantage of this nonideal condition, the peak-to-peak ac output swing V_{CE} can thereby be increased by moving the dc Q-point to the right due to the wider spreading between the curves corresponding to different base currents. Figure 3-109 combines Figures 3-103 and 3-107 to show the different ac output signals with different Q-points.

Looking at the ac output level, one might ask how the V_{CE} peak voltage can exceed the supply voltage V_{CC}. Recall that the inductor is an energy-storing device ($v = L\ di/dt$). Therefore, total instantaneous voltage is V_{CC} plus the voltage contribution of the inductor.

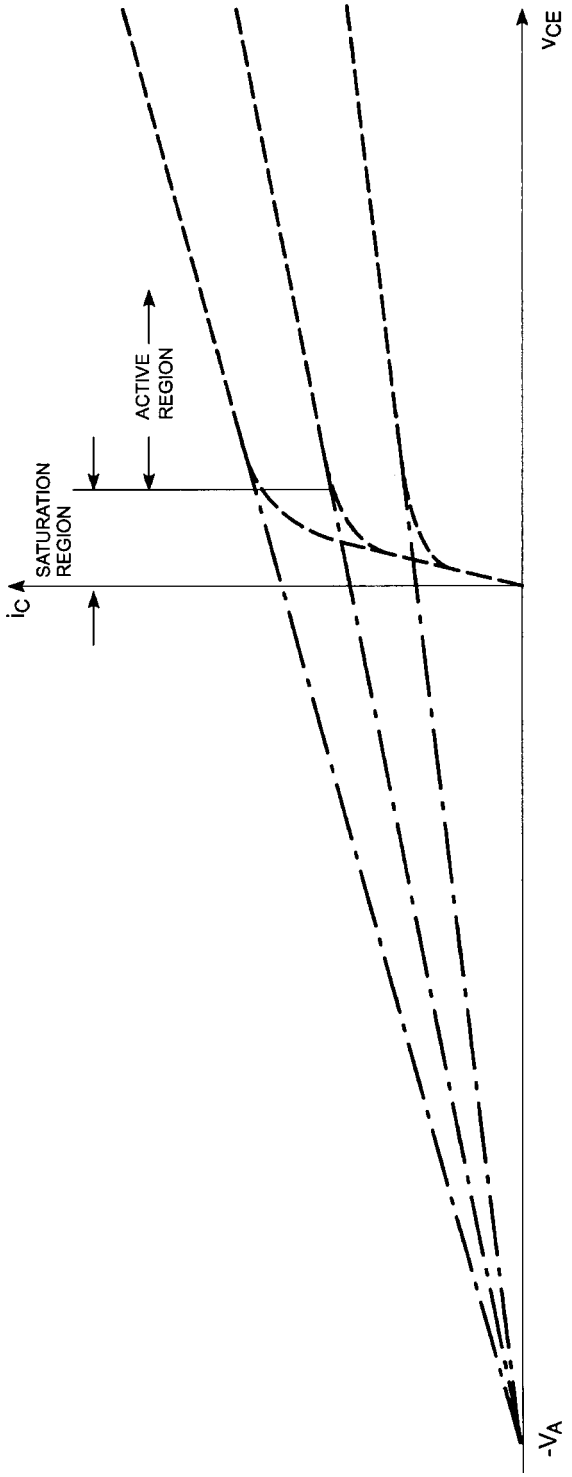


Figure 3-108 Graphical representation of Early voltage effect.

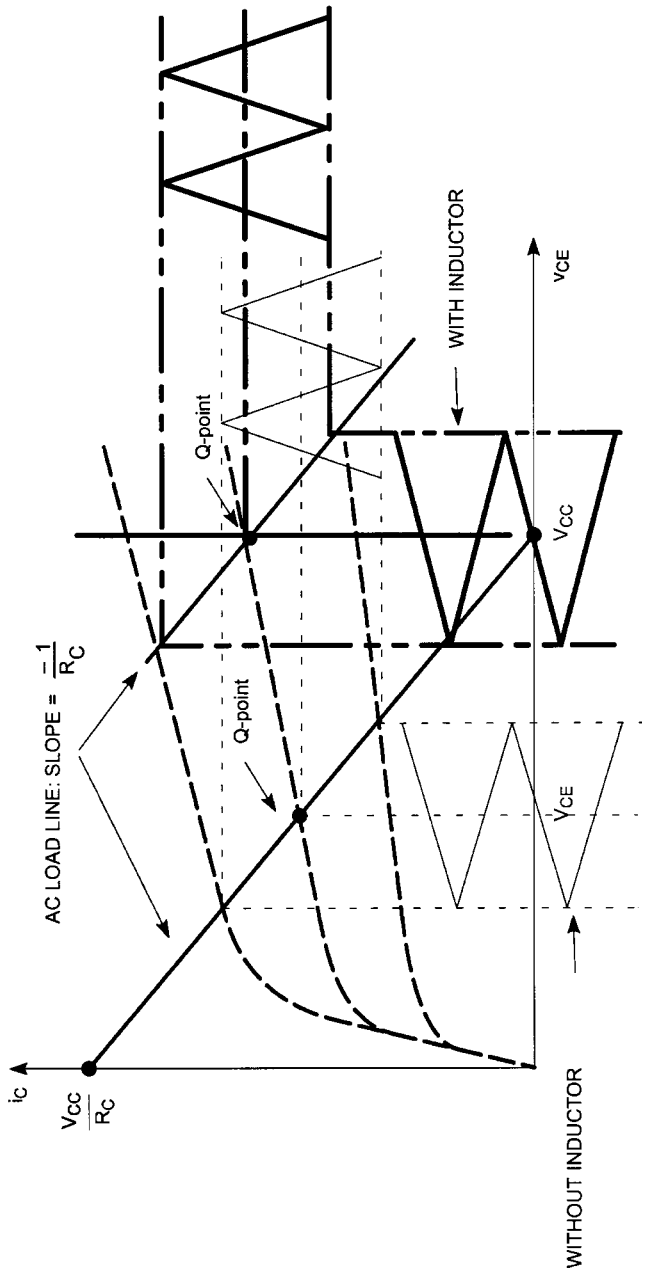


Figure 3-109 Comparison of open-collector circuit with inductor versus without inductor.

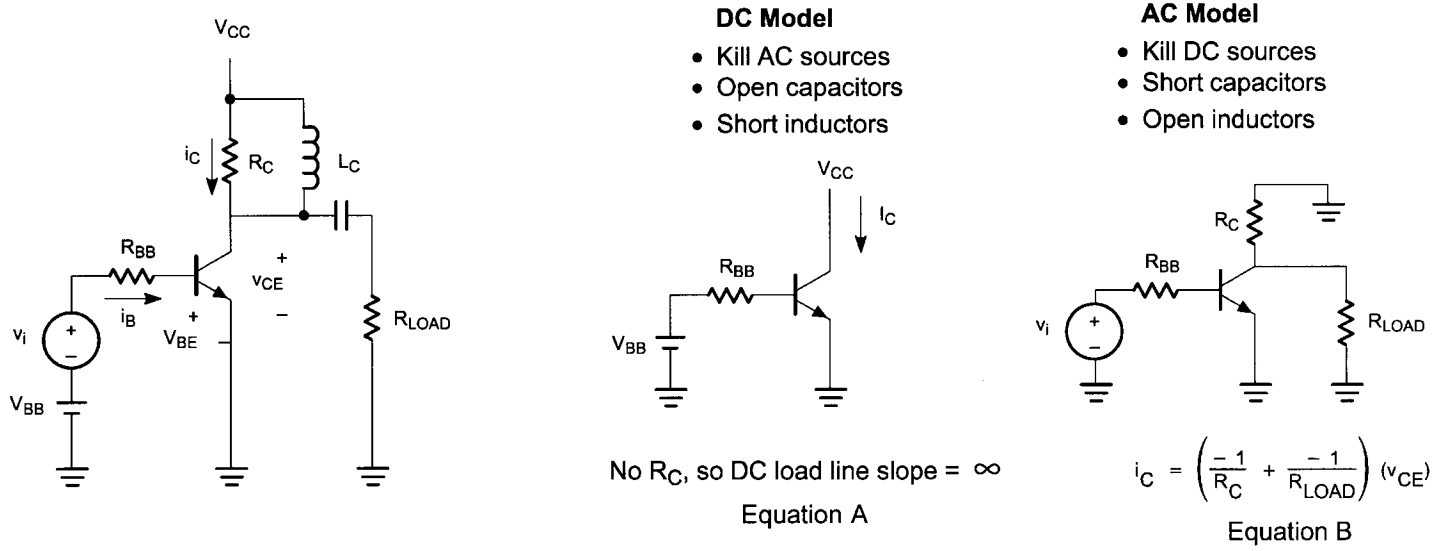


Figure 3-110 Basic transistor analysis with inductor and R_{LOAD}.

Open Collector with Inductor (L_C) and R_{LOAD}. Figure 3-110 shows the dc and ac analyses with the inductor and load resistor. Again, from the dc analysis, the inductor causes R_C to be nonexistent so the dc load line is vertical. In the ac analysis, the ac load line slope is influenced by both R_C and R_{LOAD} (see Eq. B in Figure 3-110). The ac load line slope is the same as the example of the open collector without the inductor. Figure 3-111 shows the response for the open collector with L_C and R_{LOAD}. Figure 3-112 combines Figures 3-105 and 3-111 showing the increase in ac output swing.

In conclusion, for the load line, R_C plays a role in setting up the bias-determined Q-point as well as the ac source impedance. However, when an inductor is placed in parallel with R_C, a different Q-point is set and the ac source impedance is altered. Moving the Q-point takes advantage of the transistor's nonideal i_c dependence on V_{CE} to get more signal output without having to change the base current. Since R_C is in parallel with R_{LOAD} in the ac condition, it influences the ac load line slope.

Flexible Matching Circuit [25]. A useful variation of the open-collector matching concepts previously outlined provided the capability of delivering equal power to two unequal resistive loads. This allows the power delivered to the load to be measured indirectly at another test point in the circuit where the impedance can be defined arbitrarily. If this impedance is defined to be 50 Ω, a spectrum analyzer can easily be placed directly into the circuit. This is an excellent troubleshooting technique and a valuable option to have available in high production environments.

Figure 3-113 shows the schematic for this flexible matching circuit. In this circuit, C_B functions only as a dc blocking capacitor and presents a negligible impedance at the frequency of interest. Recall from the previous open-collector matching discussions that,

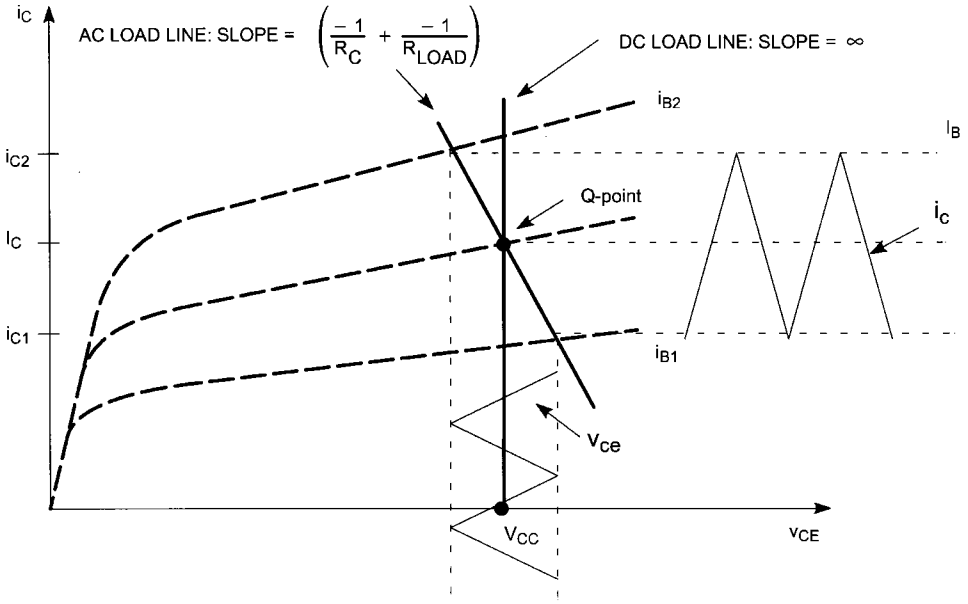


Figure 3-111 Graphical analysis for the circuitry in Figure 3-110.

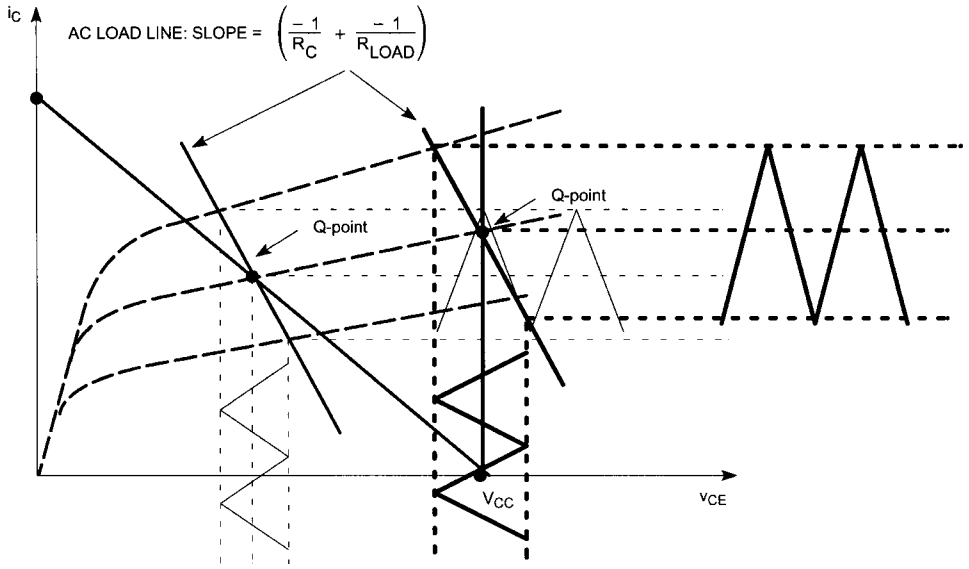


Figure 3-112 Comparison of open-collector R_{LOAD} circuit with inductor versus without inductor.

when R_C is placed in parallel with an inductor, it has no effect on the Q-point but does influence the slope of the ac load line as

$$\text{Slope} = \frac{-1}{R_C} + \frac{-1}{R_{\text{LOAD}}} \tag{3-166}$$

The capacitor C_S not only functions as a dc blocking capacitor but is also chosen such that the impedance presented by the combination of L, R_C, and C_S is equal to R_{LOAD} for optimum power transfer. The analysis is done in the following manner. First, note that inductor L is connected to V_{CC}, which is an affective ac ground. So L can be redrawn to ground. Next, R_C and C_S are converted to their parallel equivalent values as shown in Figure 3-114.

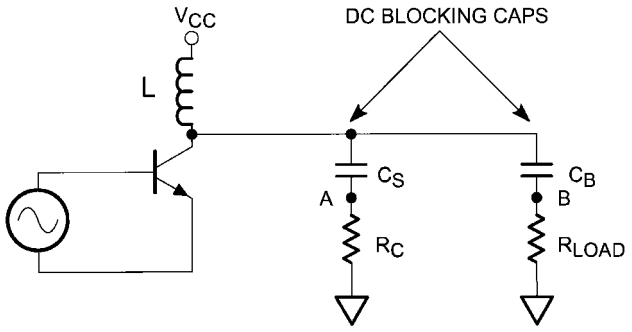


Figure 3-113 Flexible matching circuit.

Figure 3-114 Converting from series to parallel configuration.

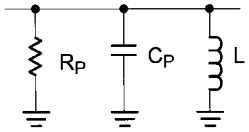


Figure 3-115 Converting from series to parallel configuration. Goal: Find L_C and C_p such that R_p looks like the R_{LOAD} value at the resonant frequency.

The resulting parallel *LCR* circuit is shown in Figure 3-115. At resonance, the parallel L–C_P combination will be an effective open circuit leaving only R_P. R_P is then simply chosen to be equal to R_{LOAD}.

Starting with the most simple transistor configuration and using *RC* types of dc-coupled amplifier circuits, we next present an overview of possible configurations for this application.

3-3 SINGLE-STAGE FEEDBACK AMPLIFIERS

This amplifier (Figure 3-116) is a standard minimal configuration amplifier with reduction of current distortion because of the emitter feedback. The voltage gain is

$$\text{VG} = \frac{R_L}{(26\text{mV})/I_C + R_e} \tag{3-167}$$

where R_L is the total load seen at the collector.

The feedback amplifier in Figure 3-117 uses voltage feedback to "correct" the input and output impedance to be close to 50 Ω but does not improve the current distortion. The voltage gain of this circuit is

$$\text{VG} = \frac{R_F}{R_g} \tag{3-168}$$

$$R_{\text{IN}} = R_{\text{OUT}}$$

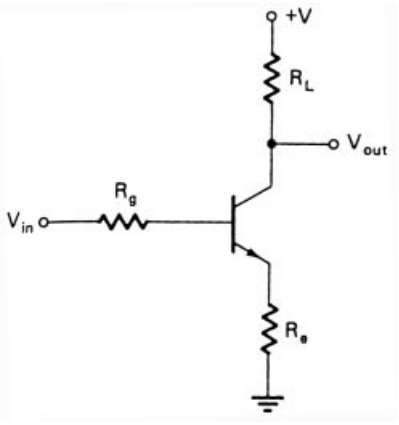


Figure 3-116 Single-stage amplifier with current feedback (provided by R_e).

$$= R_F R_D \quad \text{with} \quad R_D = \frac{26\text{mV}}{I_C} \tag{3-169}$$

Example. For $I_C = 10$ mA $\rightarrow R_D = 2.6\ \Omega$ and assuming $R_F = 200\ \Omega$, the input resistance $R_{\text{IN}} = 22.63\ \Omega$. The gain is calculated by getting the inverse of $R_D = g_m = 0.39$ A/V. Once this is multiplied with the output impedance of 22.63 Ω, the resulting gain becomes 10 log (8.6) or 9.3 dB. The gain can be increased by either using a larger collector current or a larger feedback resistor. It should be reminded that this circuit does not improve the distortion caused by the base–emitter junction.

To reduce the current distortion as well, we use the circuit shown in Figure 3-118, which contains both types of feedback. In reality, the collector–base resistor is frequently split into an unbypassed section responsible for the feedback and a larger resistor responsible for the appropriate biasing. The example circuit in Figure 3-119 shows this.

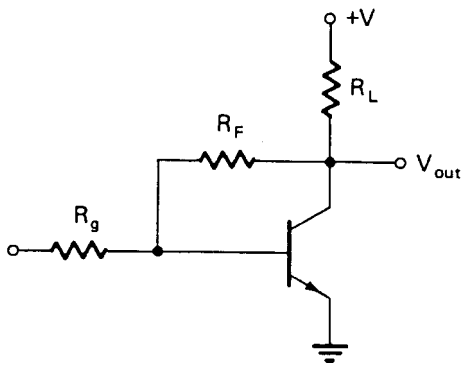


Figure 3-117 Single-stage amplifier with voltage feedback (provided by R_F).

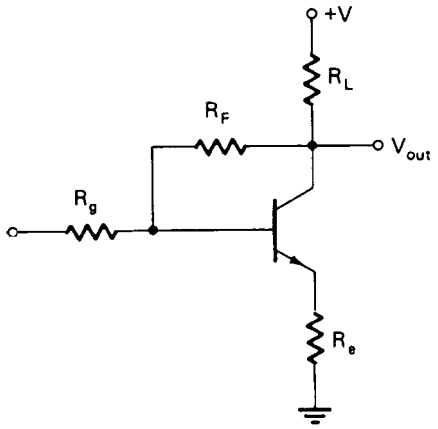


Figure 3-118 Single-stage amplifier with voltage and current feedback.

By using two of the fairly recent BFP420 transistors in parallel, and balancing the transistors by using the appropriate emitter resistors (which in total will result in 10 Ω) plus the bypassed value, we obtain the frequency-dependent gain, matching, and noise responses shown in Figure 3-120. To determine the appropriate feedback resistor R_F, we recommend the formula

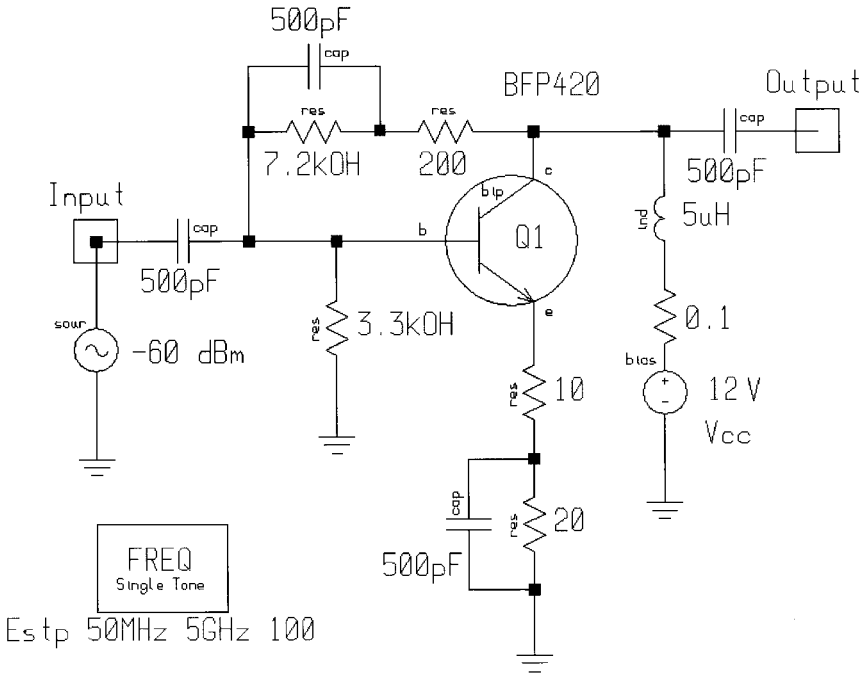


Figure 3-119 Practical amplifier circuit showing voltage and current feedback, and illustrating how reactances (in this case, 500-pF bypass capacitors) and multiple resistors can be used to provide different ac and dc resistances for the paths from collector to base and emitter to ground.

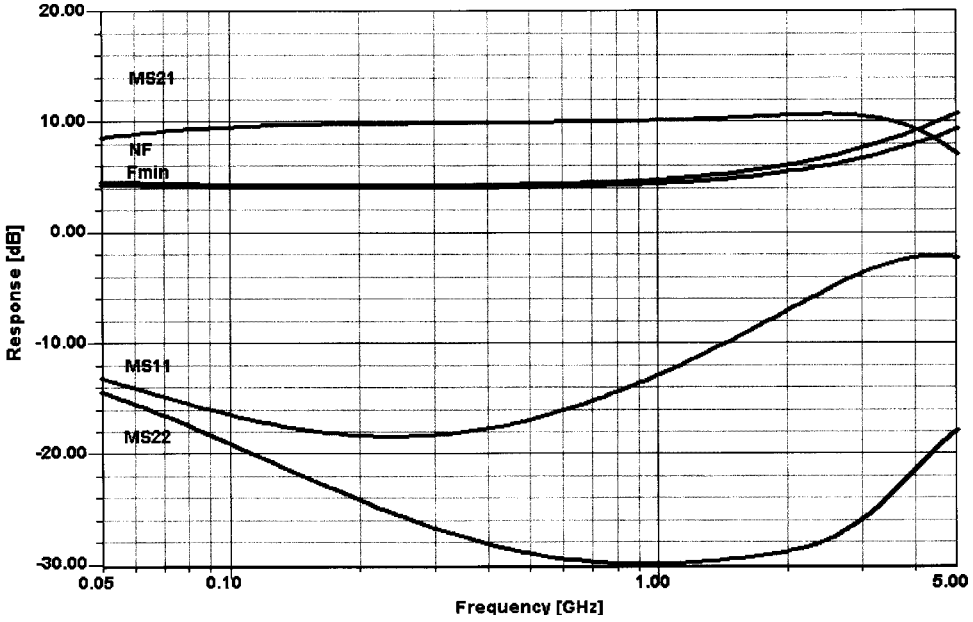


Figure 3-120 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-119.

$$R_F = \frac{(Z_{\text{in}}\, Z_{\text{out}})}{R_E} \tag{3-170}$$

This equation requires some additional comments. The combined shunt and series feedback only work as long as a sufficient transconductance is provided. In other words, there must be enough open-loop gain that will be reduced to obtain more bandwidth. A necessary condition for this approach, which really comes from the frequency range of less than 30 MHz, is that

$$Z_{\text{in}} = \frac{26\ \text{mV}}{I_C}\,\beta >> R_F \tag{3-171}$$

This requires a minimum transconductance of

$$g_m = \frac{1 + S_{21}}{Z_{\text{in}}} \tag{3-172}$$

where S_{21} must be a negative number that reflects the 180° phase shift. Of course, life becomes interesting when S_{21} goes through 1, but it (hopefully) never becomes negative and larger than 1 at the same time.

Example. The BFP420 has been used here frequently. A 10-dB gain would require an S_{21} of 3.16 ($\sqrt{10}$); therefore, we need to inspect the datasheet, which shows that the magnitude

of S_{21} is 3.16 around 4.5 GHz, yet the phase shift has changed from 180° to around 20°. In order to design an amplifier, we now set

$$R_F = Z_{\text{in}}(1 + |S_{21}|) \tag{3-173}$$

which is valid until S_{21} becomes too small. For a given 10 dB or 3.16 for S_{21}, the resulting RF is 4.16×50, or 208 Ω. The transconductance under these conditions becomes

$$g_m = \frac{1 + |S_{21}|}{Z_0} \tag{3-174}$$

where $Z_0 = 50$ Ω (the desired input impedance). g_m now becomes 83.2 mS. We now test the input and find

$$\frac{\beta}{g_m} = Z_{\text{in}} = \frac{50}{0.0832} = 600\ \Omega \tag{3-175}$$

Therefore, the emitter feedback resistor is not necessary to meet the previously mentioned condition whereby the input impedance needs to be higher than RF. Assuming the transconductance now is 0.00832, or 10% of the previous value, the input impedance would be 60 Ω, no longer meeting the prescribed requirement. Therefore, we need an unbypassed emitter resistor that fulfills the equation

$$R_E = \frac{Z_0^2}{R_F} - \frac{1}{g_{m_{\text{dc}}}} \tag{3-176}$$

where

$$g_{m_{\text{dc}}} = \frac{1 - S_{21_{\text{max}}}}{Z_0} \tag{3-177}$$

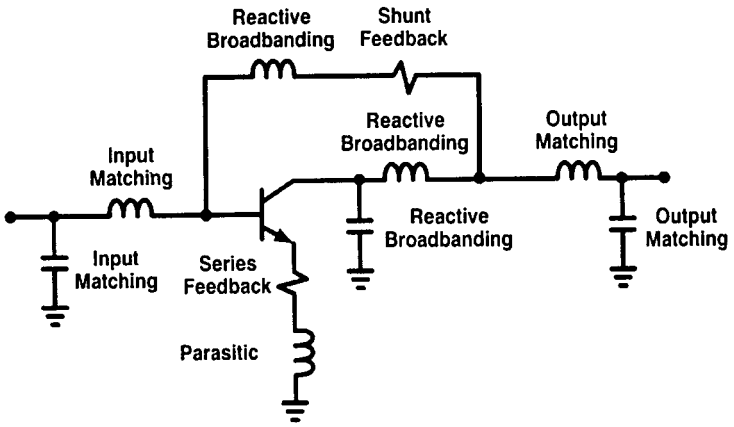


Figure 3-121 Elements of a feedback amplifier.

Again, the reader should be cautioned that in the common-emitter circuit, there is a 180° phase shift, and therefore $S_{21_{\max}}$ is a negative quantity, resulting in reality in

$$g_{m_{\text{dc}}} = \frac{1 + S_{21_{\max}}}{Z_0}$$

In our case,

$$g_{m_{\text{dc}}} = \frac{-S_{21}}{2Z_0} = \frac{38}{2 \times 50} = \frac{38}{100} = 0.38$$

so

$$R_E = \frac{50^2}{208} - \frac{1}{0.38} = 12 - 2.6 = 9.4\ \Omega \tag{3-178}$$

In practice, the designer would use a 10-Ω resistor.

Since the gain of the transistor continues to fall as a function of frequency, one can compensate to a degree for the cutoff frequency by putting an inductance in series that follows the equation

$$L_F = \frac{R_F}{2\pi f_{3\text{dB}}} \tag{3-179}$$

Assuming $f_{3\text{dB}}$ = 2 GHz, the resulting series inductance L_F = 16.6 nH. Basically, the same applies to circuits where the second transistor is used as an emitter follower to feed the signal back from the collector to the base of the first transistor.

In summary (see Figure 3-121):

- Resistive negative feedback at lower frequencies reduces open-loop gain and simultaneously matches input and output.
- Reactive elements set the gain and matching at high frequencies. Feedback is reduced at high frequencies.
- Emitter bypass capacitors are not used, as they would destroy the impedance match.

This is equally valid for FETs.

The next level of complexity can be reached by using two transistors with dc feedback.

3-3-1 Lossless or Noiseless Feedback

All the amplifiers shown above are based on a combination of resistive and reactive feedback. The very moment resistive feedback is used, the noise figure increases, as does the intercept point. As long as the intercept point increases overproportionally compared to the noise figure, this is an acceptable method. However, there are many cases in which an extremely low noise figure and a higher than normally achievable intercept point are required. To avoid this degradation of performance, Norton of Anzac developed a negative feedback structure based on the capabilities of the ferrite transformer. This patented technique (U.S. Patent Nos. 3,624,536 and 3,891,934, since expired), known as *lossless feedback*, provides lower noise figure and higher linear output than can be achieved using resistive feedback with the same transistor. Transformer feedback is used in which the transformer in the feedback network acts as a directional coupler. The coupling ratio between the input and output of the coupler

determines the magnitude of the feedback, and hence the gain. Across the frequency band over which the transformer operates as intended, the feedback is negative due to the in-phase coupling and the inversion of the active device.

This technique was used in amplifiers ranging from 5 to 500 MHz, 20 to 1000 MHz, and 300 to 1800 MHz. The difficulty associated with this type of amplifier is the fact that it is based on transformers that use twisted pairs of wires as transmission lines. In some implementations used at higher frequencies, actual directional couplers were used at the expense of bandwidth; in these circuits, the coupling was determined by the spacing of the coupler lines rather than transformer turns ratio. These microstrip couplers have much narrower bandwidths than transformers and are fairly large.

Another headache with the lossless feedback design is the fact that many of these transistor configurations are used in either grounded-base or grounded-emitter with transistors that have f_T values of 25 GHz and higher. Such devices are very hot, and achieving freedom from oscillation over a wide frequency range is a continuing problem, even for the circuits we have shown so far. A similar clever technique has been implemented in the termination-insensitive mixer, which we will examine in Chapter 4.

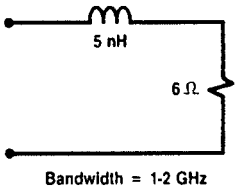
3-3-2 Broadband Matching

The circuitry above has enabled us to extend the frequency range of the amplifier and avoid the normal sharp cutoff. According to Fano [26], there is a limit to the bandwidth of matching. The concept is that impedance plus bandwidth determines the ability to match, since there is a Q of the circuit and there is a Q of the device. The following equations are valid:

$$|\rho_{\text{Bode}}| = e^{-(\pi Q_T / Q_{BP})} \tag{3-180}$$

$$Q_T = \frac{\sqrt{f_1 f_2}}{f_2 - f_1} \tag{3-181}$$

$$Q_{BP} = \frac{\text{Im}(Z)}{\text{Re}(Z)} = \frac{X_s}{R_s} \tag{3-182}$$



$$Q_{BP} = \frac{2\pi\sqrt{(2\text{ GHz})(1\text{ GHz})(5\text{ nH})}}{6\,\Omega} = 7.40$$

$$Q_T = \frac{\sqrt{(2\text{ GHz})(1\text{ GHz})}}{(2\text{ GHz} - 1\text{ GHz})} = 1.414$$

$$|\rho_{\text{Bode}}| = e^{-\left(\frac{\pi(1.414)}{7.40}\right)} = .549$$

$$\text{VSWR}_{\text{Bode}} = \frac{1 + .549}{1 - .549} = 3.43 : 1$$

Figure 3-122 Broadband matching example showing the Bode limit.

Figure 3-122 shows an example. This combination results in a VSWR of 3.43, which for practical purposes is already too high. The workaround is special circuitry (such as exponentially staggered impedance jumps in the matching network like 50 Ω > 20 Ω > 7 Ω > 3 Ω). This means we first transform the 50-Ω source of load impedance down to 3 Ω, and then look for a matching network that matches the transistor output/input, which is within a few ohms of 3 Ω. A detailed discussion of such matching can be found in the "Numerical Design" portion of Section 3-12-2 under the title "Broadband Matching Using Bandpass-Filter Networks—High-Q Case."

3-4 TWO-STAGE AMPLIFIERS

There are two popular configurations that have been used mostly for low-frequency applications. The first one (Figure 3-123, implemented in the circuit shown in Figure 3-124) shows a nice high-frequency performance assuming the emitter and collector currents are essentially the same. As for distortion products, this circuit behaves equivalently to the shunt impedance example from Figures 3-118 and 3-119, with the additional gain of the second transistor. Its frequency matching and noise performance are also shown, in Figure 3-125.

The next two-stage circuit (Figure 3-126, implemented in Figure 3-127) uses the collector current of the second stage as the feedback to the first. Proper biasing of this circuit is quite difficult and interactive. Figure 3-128 shows the frequency-dependent gain, matching, and noise figure. We believe that this circuit can be optimized further, but we leave such investigations to the reader.

A *cascode* amplifier is a special form of two-stage circuit. Although a cascaded pair acts like a single transistor, its merits are in its reduced feedback and resulting higher gain–bandwidth product. Figure 3-129 shows a simple cascode application that assumes the presence

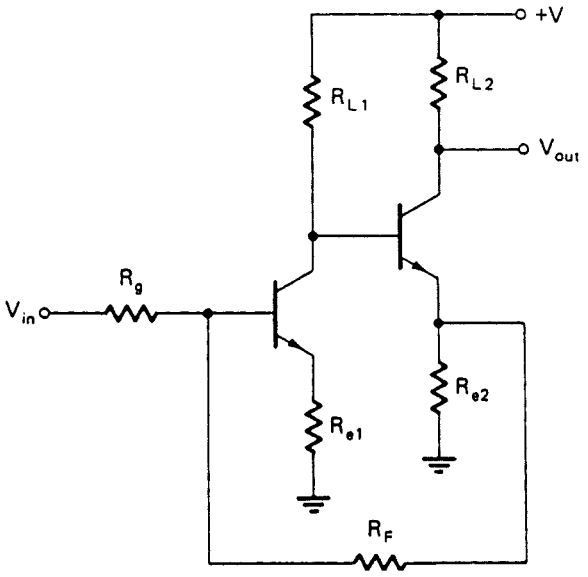


Figure 3-123 Two-stage amplifier with voltage feedback.

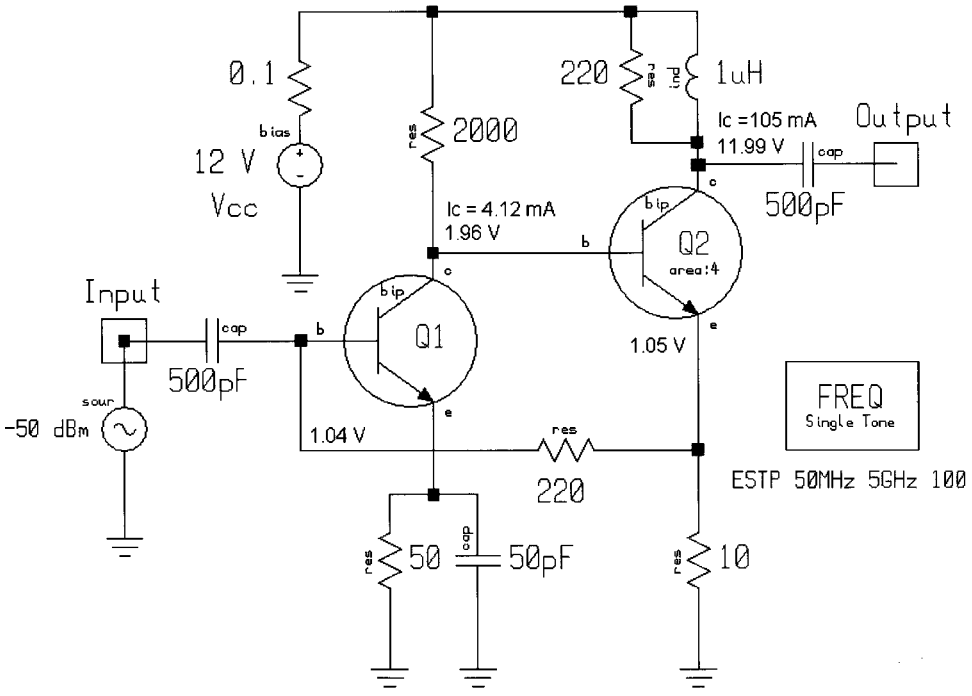


Figure 3-124 Implementation of the circuit in Figure 3-123.

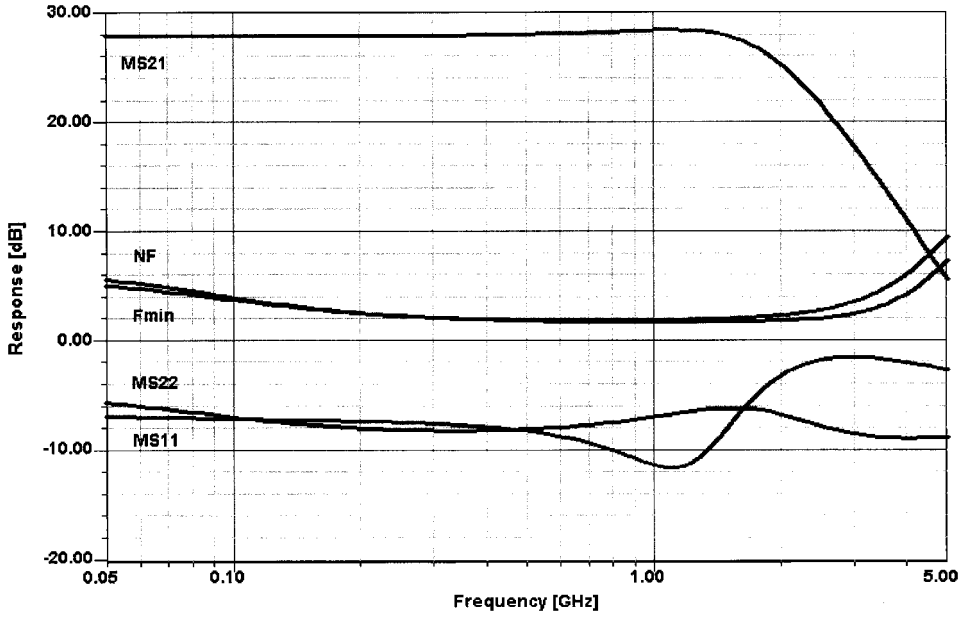


Figure 3-125 Frequency-dependent gain, matching, and noise performance of the circuit in Figure 3-124.

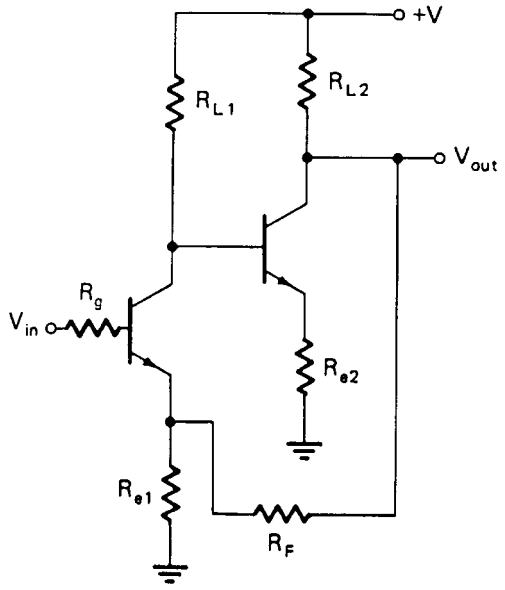


Figure 3-126 Two-stage amplifier with voltage feedback from Stage 2 to Stage 1, and current feedback in both stages.

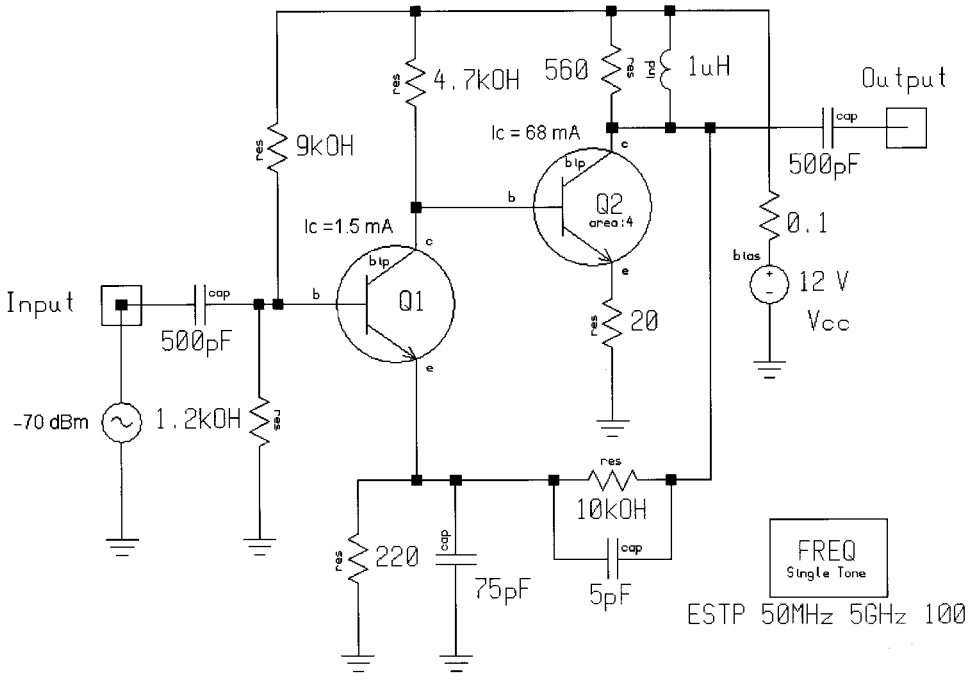


Figure 3-127 Simulated two-stage amplifier with feedback as shown in Figure 3-126.

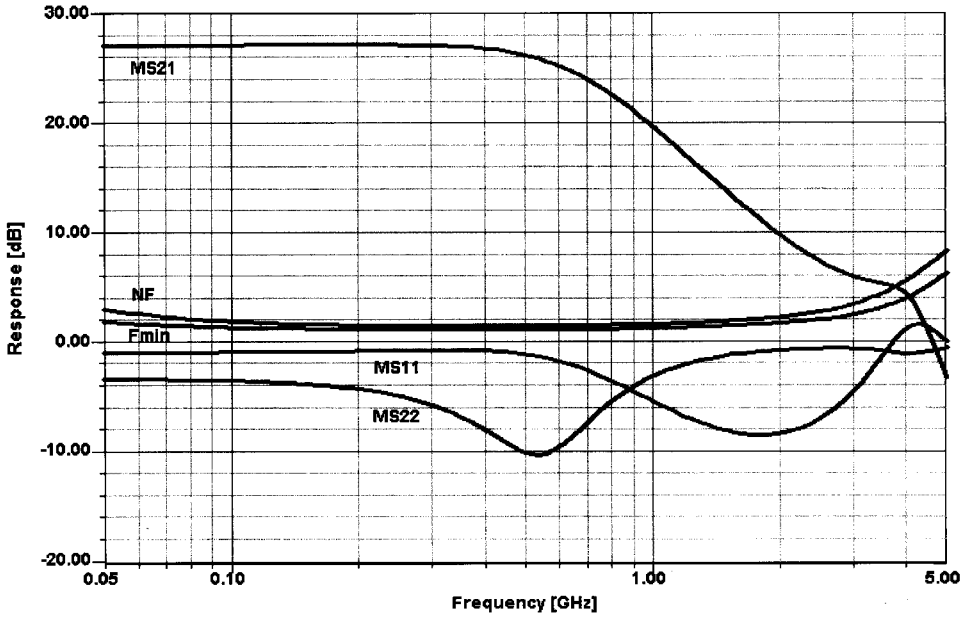


Figure 3-128 Gain, matching, and noise performance of the amplifier in Figure 3-127.

of an output transformer that absorbs the output capacitance. Figure 3-130 shows this circuit's frequency-dependent gain, matching, and noise performance.

It has been obvious to us that most published wireless application circuits are evaluated on a no-input-selectivity basis; that is, their reported performance does not take into account the filtering that is usually present in practical implementations. Evaluated in this way, a low-noise design may achieve a noise figure of less than 2 dB—1.2 dB or so is typical—but this is misleading because filters and connectors introduce insertion loss, resulting in a higher noise figure.

The cascode is a nice preamplifier combination. It combines (depending on the transistor) a low noise figure with essentially no feedback. Since the output of a grounded-base transistor provides a high-impedance source, this configuration is ideal to work into the SAW filters frequently used in wireless applications at the operating frequency. Special transistor pairs are available for use in cascode; as can be seen from this example, not every transistor combination is ideal. This one gives a higher noise figure than wanted and at frequencies above 2 GHz it tends to show possible instabilities. Both Siemens and Motorola are now offering transistors appropriate for these cascodes. Mostly, one would resort to transistors of less-exotic cutoff frequencies to guarantee stability. Typically, grounded-base/gate transistors tend to become oscillators, specifically because this configuration, like the emitter follower, allows a higher operating frequency while maintaining the same gain–bandwidth product; the actual gain for the last two circuit configurations is less. A grounded-base/gate stage has a high voltage gain and the emitter/source follower has power gain at low impedance levels. We invite our readers to experiment with the capacitor from the base of Q1 to ground in Figure 3-129 to get an interesting education about these transistors and their willingness to oscillate.

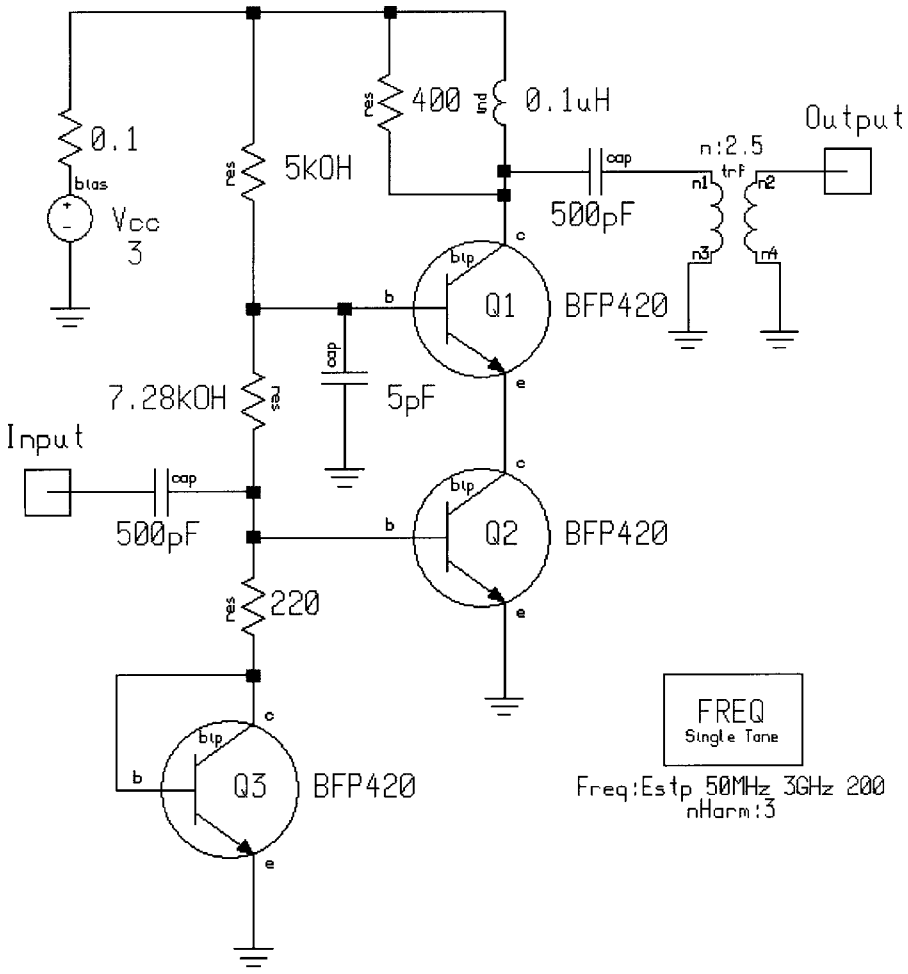


Figure 3-129 Schematic of the cascode amplifier.

The MC13144 is an example of a commercially available VHF to 2-GHz low-noise amplifier with programmable bias. As can be seen from Table 3-11, it is ideal for the 900-MHz range and its programmable gain/bias makes it very attractive. This design (Figure 3-131) is based on the cascode mentioned above and, because of high integration, uses many transistors for bias and level shifting.

The 900-MHz application circuit in Figure 3-132 shows the required external components for a successful design. Input and output are made tunable. The 5.6- and 39-nH inductances can be made from small air-wound coils. The output ceramic filter provides the necessary selectivity ahead of the mixer. With such a highly selective device in the input, the noise figure would deteriorate too much. The antenna is rarely in the vicinity of sufficiently strong transmitters to require such strong input filtering.

There are two more interesting applications for two-stage amplifiers: The one implemented in the Hewlett-Packard MSA-0735 MMIC has become a quite common one, and its usefulness up to higher frequencies depends mostly on the cutoff frequency of the transistors,

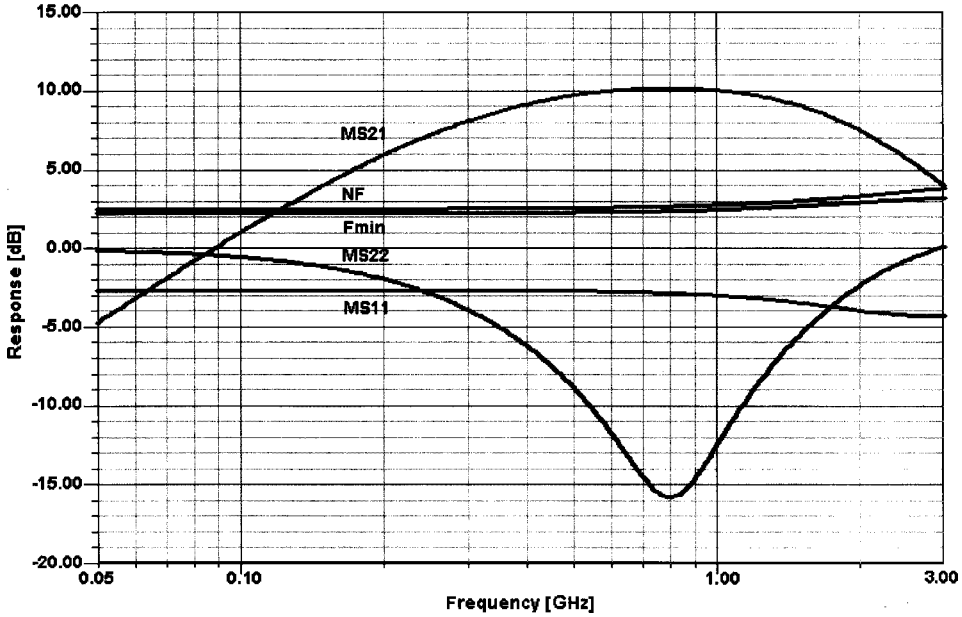


Figure 3-130 Frequency-dependent gain, matching, and noise performance of the cascode amplifier.

their parasitics, and the phase shift of the gain. The typical configuration is shown in Figure 3-133, and its frequency-dependent gain, match, and noise figure in Figure 3-134. There is an increase of S_{22} at higher frequencies; this indicates unwanted phase shift in the circuit. Some manufacturers hide the compensation circuits to overcome such peaking; most of the schematics shown by the manufacturers are really intended only to show operating principles, with proprietary details omitted.

Thanks to the properties of GaAs transistors, we can extend the frequency range significantly. The amplifier we are about to look at goes back several years and is described in great detail in Goyal [27]. Essentially what we have, seen in Figure 3-135, is an FET version of Figure 3-127. The main goal of this circuit was to have an all-monolithic circuit, avoiding all ac coupling and accomplishing all matching and interaction with dc coupling. The transistor Q1 is the main gain stage, which has the transistor Q2 as an active load. The output impedance of Q2 is $1/g_m$—typically in the area of 50–200 Ω, depending on the biasing,

Table 3-11 MC13144 features

Gain at 900 MHz, 17 dB
Noise figure at 900 MHz, 1.4 dB
P_{-1dB}, −7.0 dBm
$IP_{3,in}$, −5.0 dBm
Low operating supply voltage (1.8–6.0 Vdc)
Programmable bias with Enable 1 and Enable 2. [Enable 1 and Enable 2 programmed high for optimal noise figure and gain associated with NF. Enable can be overridden for external bias programming (up to 15 mA.)]

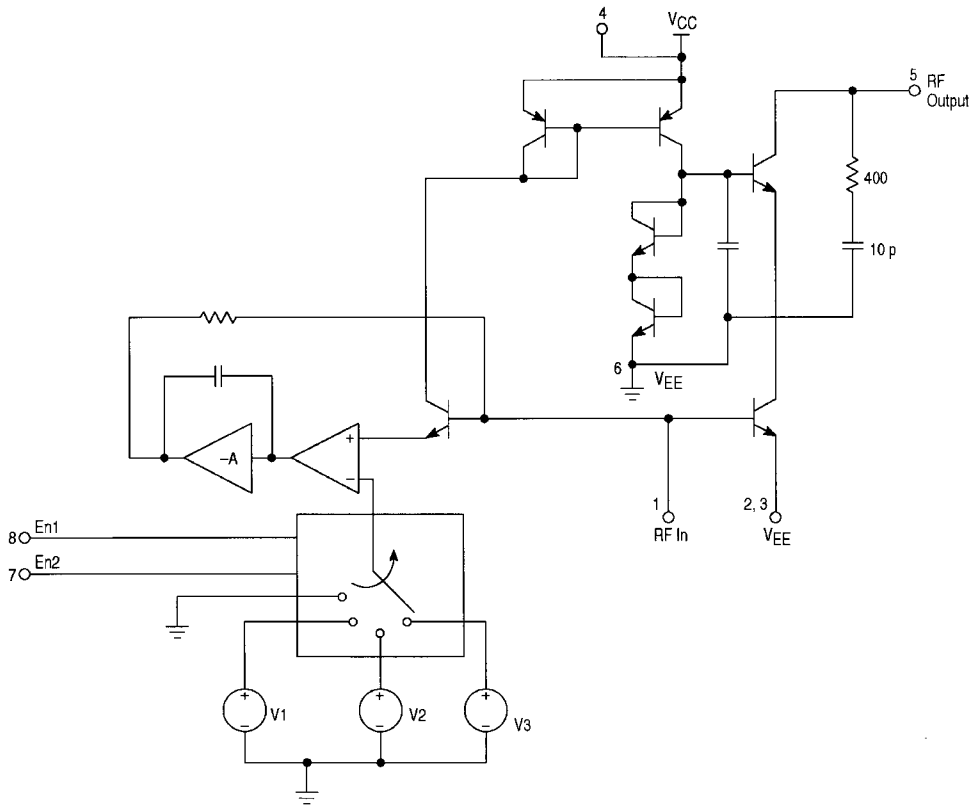


Figure 3-131 Schematic of the Motorola MC13144 cascode LNA.

which affects the transconductance. Transistor Q3 has the same function as an emitter/source follower; again, its dc bias determines its transconductance and therefore its output impedance. (Given this, a device that has a transconductance of 20 mS, or close to it, would establish a very good output match to 50 Ω.) The drains of Q1 and Q5 are at approximately 3.5 V. The diodes at the source of Q3 are level shifters that must shift the dc voltage at the gate of Q5 and at the drain of Q4 to approximately −1.5 V relative to ground—the necessary bias condition for Q1 and Q5. The standard shunt feedback circuit would show resistive feedback between the drain of Q1 (gate of Q2) and the gate of Q1; in the previous examples, we calculated the magic value of this part to be about 200 Ω. Adding in parallel to Q1 the transistor Q5 allows the designer to use a feedback scheme that isolates the feedback loop from the input. Dc-wise, the two transistors are tied together via a 10-kΩ resistor, which can be included in the actual packaged device. Insufficient information was published to allow us to duplicate the design, considering all the intricacies involved (we somewhat object to the fact that even in the *IEEE MTT-S Transactions*, so little information is given about the actual implementation of this design—using space constraints as an excuse—that its educational value is usually minimal); it shows a new principle without revealing the design steps necessary to obtain a practical circuit. Several variations of this type of circuit are available; the literature is full of discussions about them. The magnitude of the feedback is set by the ratio of the width of Q5 to that of Q1. Typical values for the Q5/Q1 ratio range

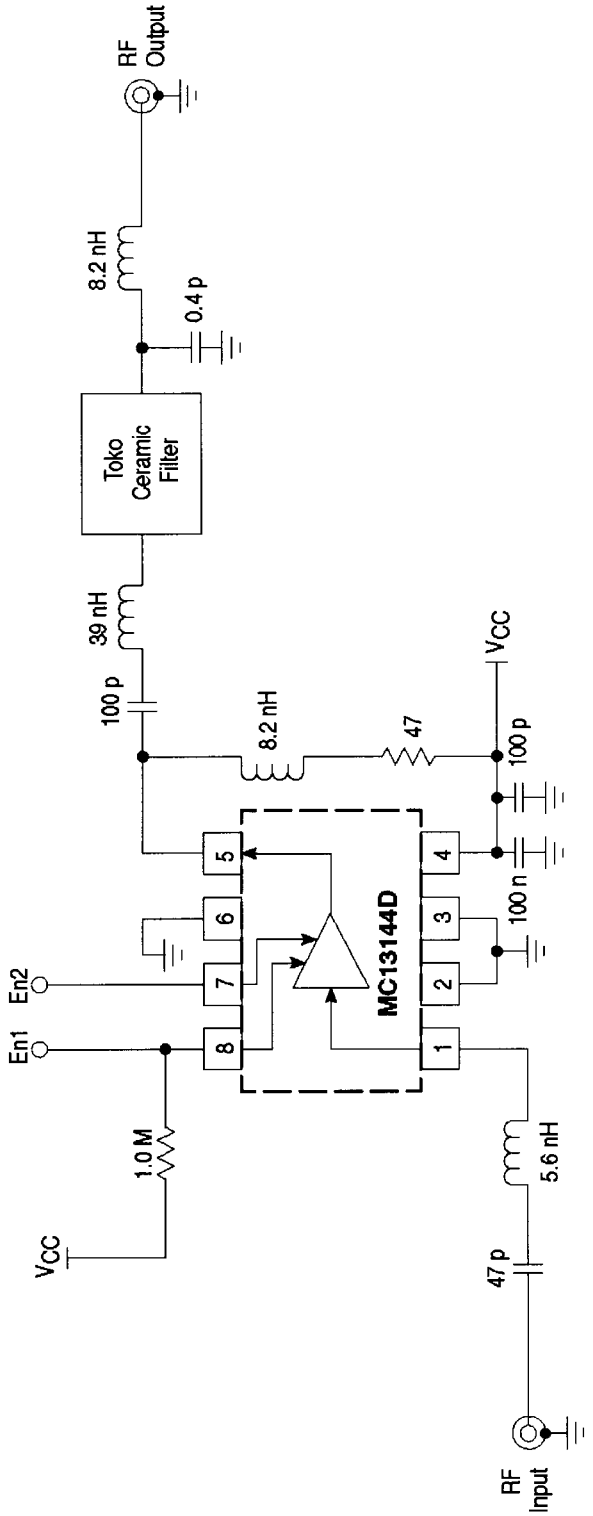


Figure 3-132 An MC13144 application circuit for 926.5 MHz.

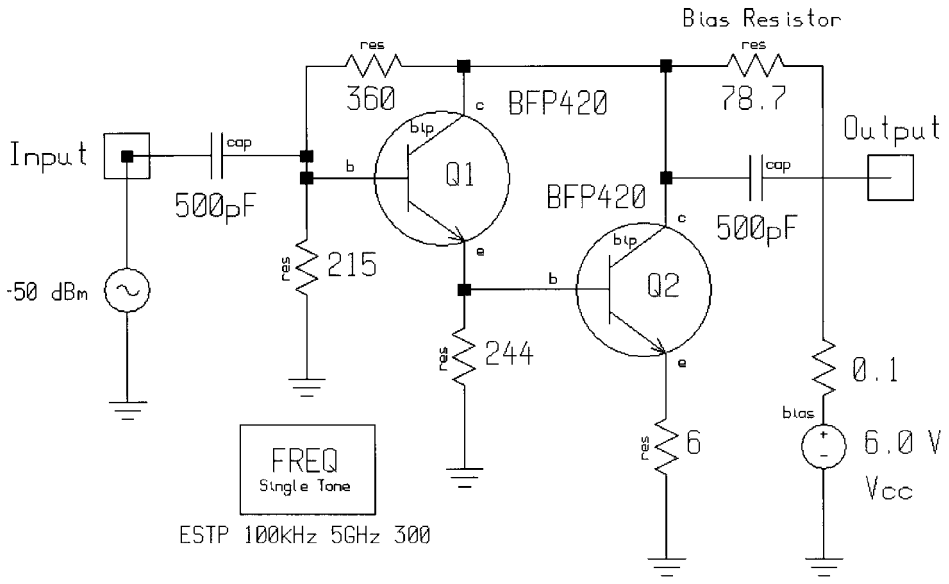


Figure 3-133 Schematic of the MSA-0375 MMIC amplifier.

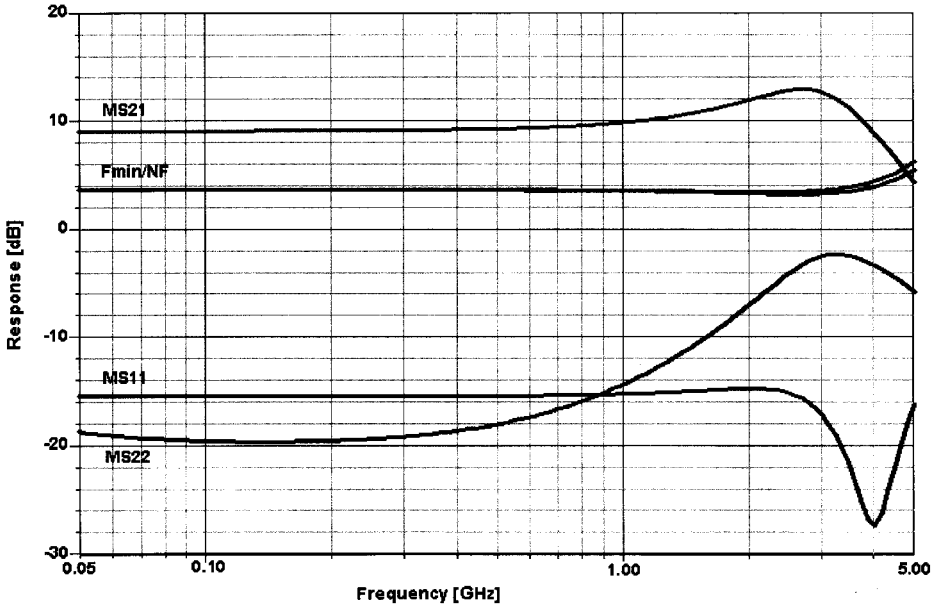


Figure 3-134 Frequency-dependent gain, matching, and noise performance of the MSA-0735 MMIC amplifier.

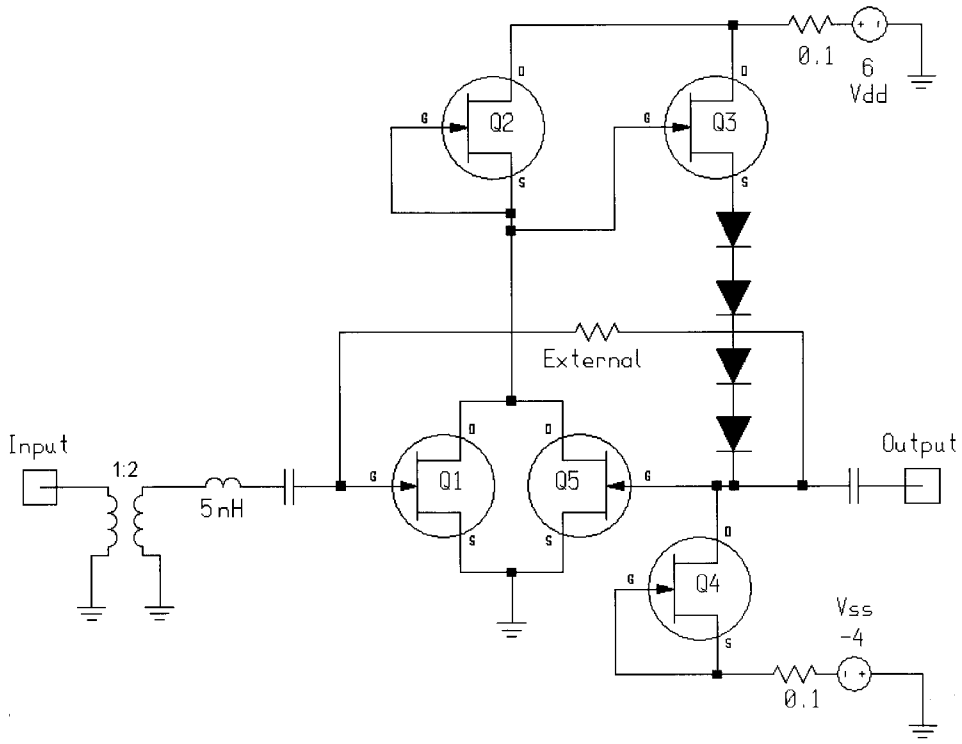


Figure 3-135 Schematic of the dc-coupled GaAsFET amplifier.

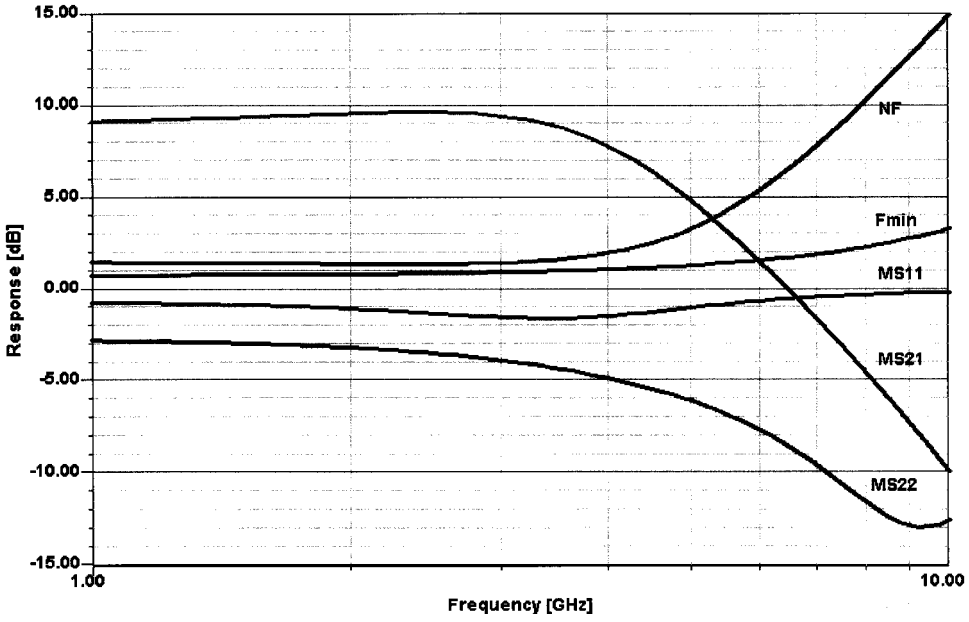


Figure 3-136 Frequency-dependent gain, matching, and noise performance of the dc-coupled GaAsFET amplifier.

from 0.15 to 0.30. In simulation, sizing the devices can best be accomplished by using the scaling factor for the FET model.

One of our questions was: "What would the perfect input termination for Q1 be?" By using an ideal transformer and playing with its turns ratio (needless to say, we did this with CAD), we determined experimentally that, relative to the gate, 200 Ω and a series inductor of 5–8 nH—the sort of details that designers often avoid telling their audiences—gave the ideal frequency response. Overall, we believe that the inventor of this circuit can be quite proud of the overall performance. The FETs we chose were taken from the Texas Instruments nonlinear foundry library; however, any similar device from another company would have worked equally well. The feedback loop used in this IC is an example of an approach frequently referred to as *active feedback*. The performance of this amplifier can be evaluated from Figure 3-136, which shows the frequency-dependent gain, matching, and noise figures (even including F_{min}).

3-5 AMPLIFIERS WITH THREE OR MORE STAGES

Most three-stage amplifiers are obtained by adding another buffer or power-gain stage to the previously described principal designs. As an example, we decided to evaluate the recent NEC UPC2749 IC, a silicon MMIC intended for application as a low-noise 1900-MHz amplifier operating at 3 V. Table 3-12 briefly summarizes its electrical specifications.

Table 3-12 NEC UPC2749 electrical characteristics

(T_A = 25 °C, $Z_L = Z_S$ = 50 Ω, V_{CC} = 3.0 V)			
	Part Number Package Outline		UPC2749T TO6
Symbols	Parameters and Conditions	Units	Typical
I_{CC}	Circuit current (no signal)	mA	6
G_S	Small-signal gain, f = 900 MHz dB	14.5	
	f = 1900 MHz	dB	16
f_{corner}	–3-dB Gain corner frequency	GHz	2.9
P_{1dB}	1 dB Compressed output power at 1900 MHz	dBm	–12.5
P_{SAT}	Saturated output power, f = 1900 MHz	dBm	–6
NF	Noise figure, f = 900 MHz	dB	3.2
	f = 1900 MHz	dB	4.0
RL_{IN}	Input return loss, f = 1900 MHz	dB	10
RL_{OUT}	Output return loss, f = 1900 MHz	dB	12.5
ISOL	Isolation, f = 1900 MHz	dB	30
OIP_3	SSB output third-order intercept, f_1 = 500 MHz, f_2 = 510 MHz	dBm	–3
	f_1 = 1000 MHz, f_2 = 1010 MHz	dBm	–3
	f_1 = 1900 MHz, f_2 = 1902 MHz	dBm	–3.5
	f_1 = 2000 MHz, f_2 = 2010 MHz	dBm	–4
$R_{TH(J-A)}$	Thermal resistance (junction to ambient)		
	Free air	°C/W	
	Mounted on a 50 × 50 × 1.6 mm epoxy glass PWB	°C/W	

In modeling the UPC2749's three transistors, we chose to use our favorite, the BFP420, which is not without its headaches at the higher frequencies. Figure 3-137 shows the circuit diagram, which closely resembles the one published by NEC; again, the only hard facts we had were the IC's basic circuit topology, gain, and dc current. Given those numbers, we staggered the current in the three devices by setting them at 0.64 mA for the first stage, 1.85 mA for the second stage, and 3 mA for the third stage. It should be noted that the two output transistors are not unlike the approach used for the Hewlett-Packard MSA-0735 (Figure 3-133). Taking (hopefully) all the right assumptions, we end up with a design that closely resembles the manufacturer's specifications for both power consumption, gain, and noise. This can be verified by looking at the results plotted in Figure 3-138.

The final example is a more complex derivative of the Philips NE/SA5204A amplifier, which was intended for standard 6-V operation and has a fairly "normal" circuit arrangement. This can be seen from Figure 3-139. A much more elaborate cousin of this is the low-voltage NEC UPC2710 amplifier. According to its specifications, it operates at 5 V but has significantly more gain than the NE/SA5204. Table 3-13 shows its electrical specifications and Figure 3-140 shows its measured gain as published by the manufacturer.

The actual interior of the circuit is shown in Figure 3-141. It consists of the standard preamplifier and then a Darlington configuration with the collectors tied together. The middle section of the circuit acts as a power supply responsible for the bias relative to the emitter of the first transistor influencing the rest of the circuit as well. Since NEC mentioned that they use their NESAT III process (f_T of 20 GHz at V_{ce} = 3) in producing the UPC2710T and we

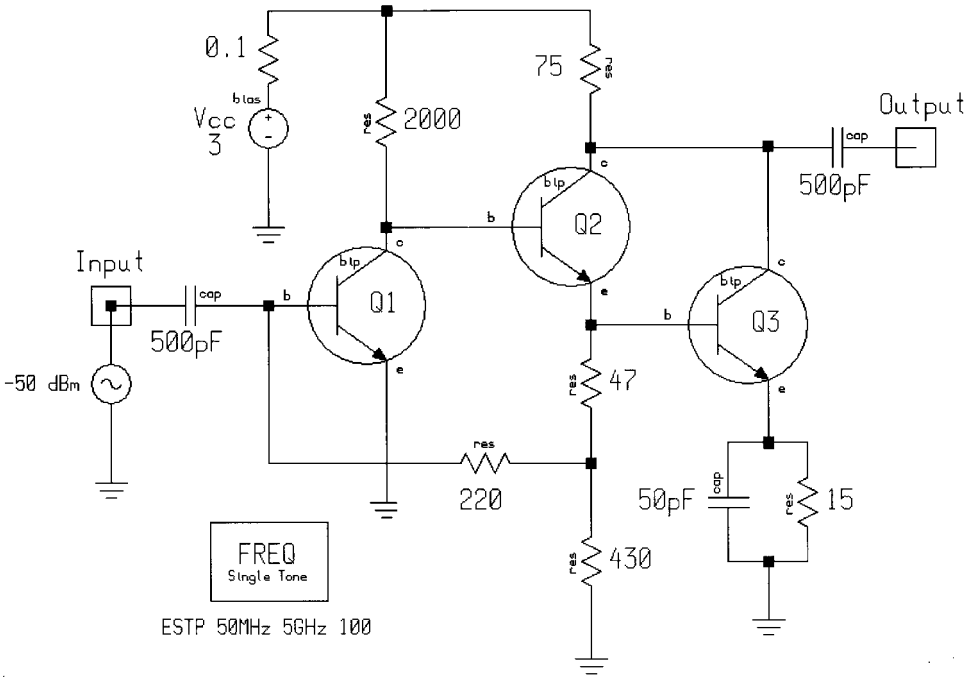


Figure 3-137 Circuit of the UPC2749 MMIC. The component values were not supplied by the manufacturer.

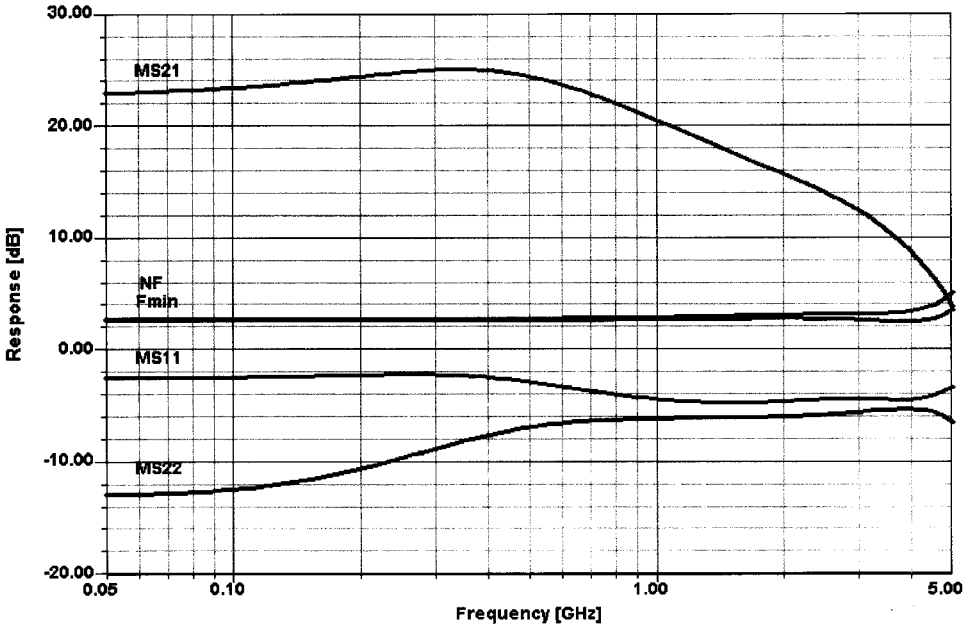


Figure 3-138 Simulated gain, matching, and noise performance of the NEC UP2749 MMIC.

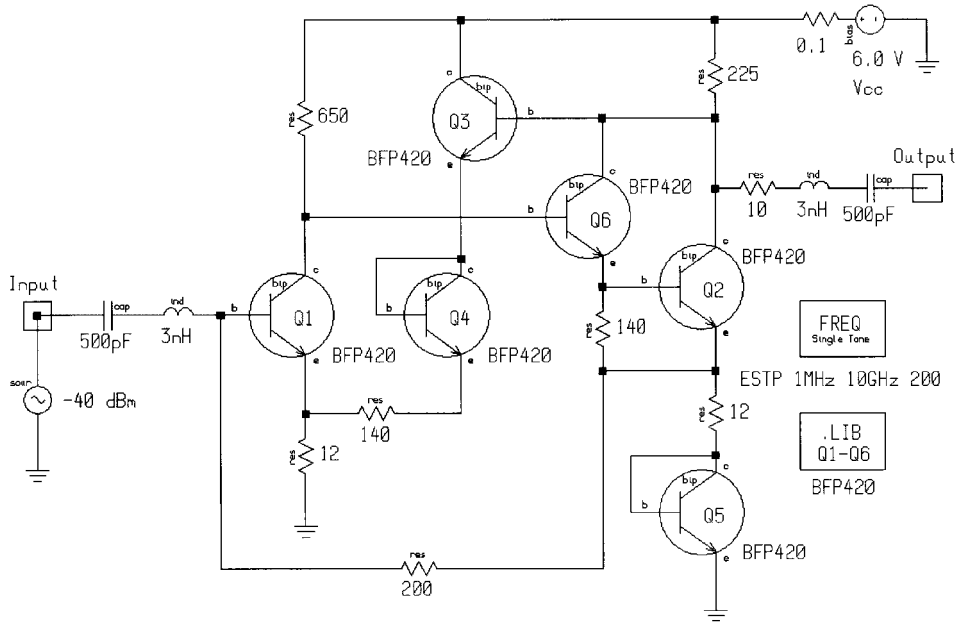


Figure 3-139 Schematic of the Philips NE/SA5204A amplifier IC entered for simulation.

Table 3-13 UPC2710 electrical specifications

	T_A = 25 °C, f = 0.5 GHz, V_{CC} = 5 V)		
	Part Number Package Outline		UPC2710T TO6
Symbols	Parameters and Conditions	Units	Typical
I_{CC}	Circuit current (no signal)	mA	22
G_S	Small-signal gain	dB	33
f_{corner}	–3-dB Gain corner frequency	GHz	1.0
ΔG_S	Gain flatness, f = 0.1–0.6 GHz f = 0.1–0.8 GHz	dB	±0.8
P_{SAT}	Saturated output power	dBm	13.5
P_{1dB}	Output power at 1dB compression point	dBm	7.5
NF	Noise figure	dB	3.5
RL_{IN}	Input return loss	dB	6
RL_{OUT}	Output return loss	dB	12
ISOL	Isolation	dB	39
ΔG_T	Gain–temperature coefficient	dB/°C	–0.006
R_{TH}	Thermal resistance (junction to ambient)	°C/W	

obtained some information regarding it, we decided to use our standard BFP420 (f_T of 25 GHz; see the datasheet in Chapter 2) for the purpose of simulation.

In our effort to simulate this circuit, we chose a combination of values that pretty much complies with the manufacturer's published specifications, but S_{22} is not close, as shown in Figure 3-142. We stopped at this point and invite our readers with CAD capabilities to provide the final touch in meeting all the published specifications for the circuit. We are particularly

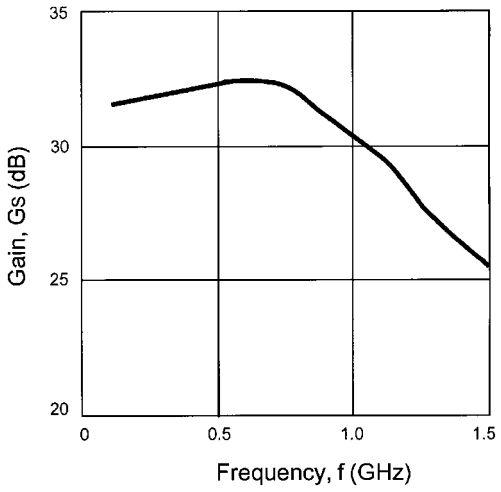


Figure 3-140 Manufacturer-supplied gain versus frequency for the NEC UPC2710T MMIC.

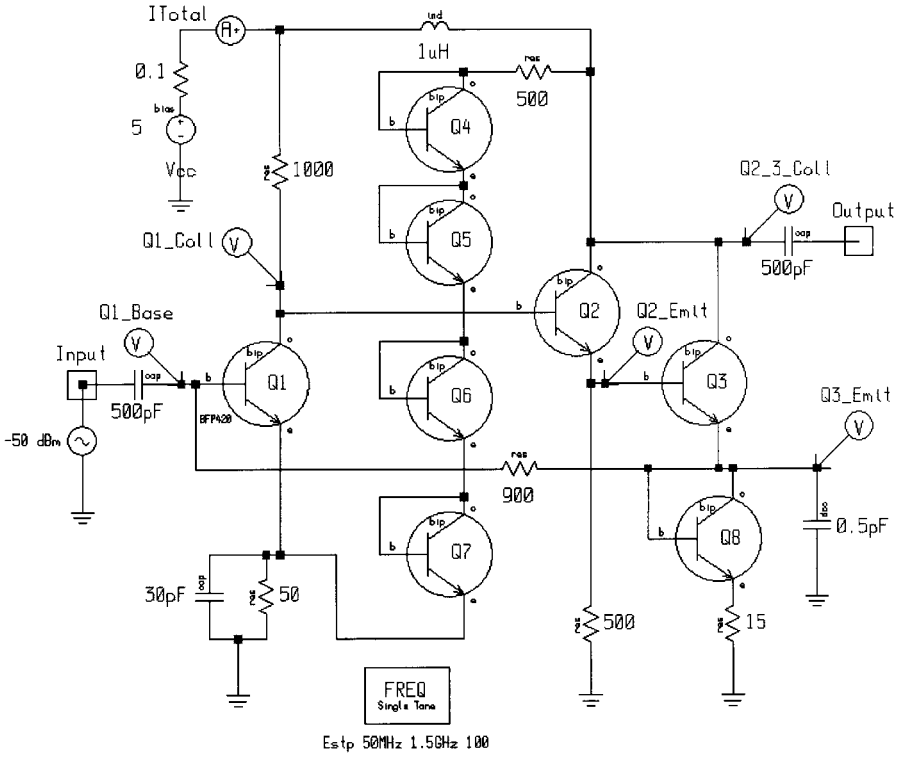


Figure 3-141 Schematic of the UPC2710T silicon MMIC entered for simulation.

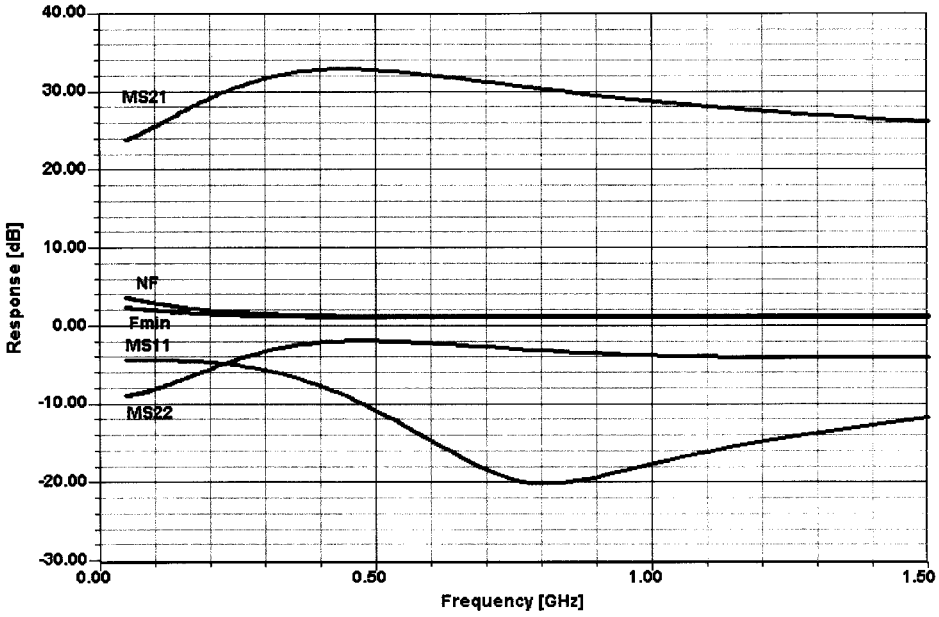


Figure 3-142 Simulated gain, match, and noise performance of the UPC2710T.

recommending such an effort because it shows very nicely the interaction of all the circuit elements in a fairly drastic way. Because of this, we admire the IC design even more, since the manufacturer has to sell a pretty consistent amplifier as a function of different production runs.

3-5-1 Stability of Multistage Amplifiers

Stability, needless to say, is a big issue, and the delay of the various stages adds to the phase shift, resulting in possible instabilities. The fact that S_{11} and/or S_{22} can become larger than 0 (in dB) means there is the potential of oscillating, depending on the reactances available. If a cable of inappropriate length is added, hell will break loose. Figure 3-143 shows the dangerous peaking of our simulation of the Philips NE5204A, in which the internal compensation was intentionally not assumed correctly. We will all agree that the high-end performance of this amplifier can generate high anxiety. As the production process of the aging device becomes more modern—meaning that transistors with much higher f_T will be the basis for the production—it will become quite difficult to maintain a previously achieved stability. This, by the way, applies also to discrete transistors, which for reasons of economics and multiple manufacturers will be improved over time. Such evolution results in generally smaller base-spreading resistances and capacitances, and other parameter changes, which, overall, cause headaches in manufacturing. We will leave the topic of three-stage amplifiers on this note.

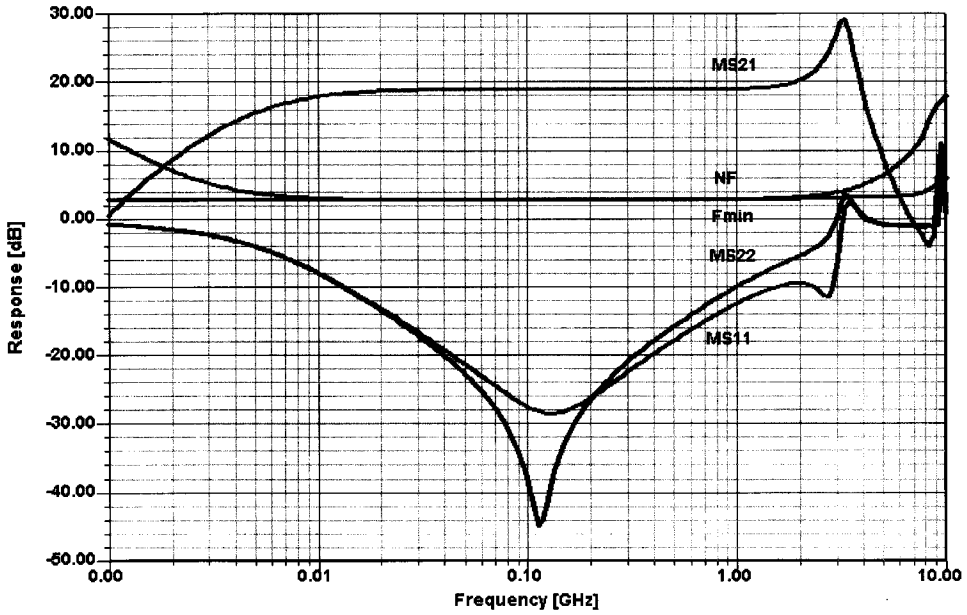


Figure 3-143 Simulated gain, match, and noise performance of the NE5204A IC.

3-6 A NOVEL APPROACH TO VOLTAGE-CONTROLLED TUNED FILTERS INCLUDING CAD VALIDATION [28]

Modern receivers control input stages as well as the oscillator band and frequency by electrical rather than mechanical means. Tuning is accomplished by voltage-sensitive capacitors (varactor diodes), and band switching by diodes with low forward conductance. Since the wireless band (essentially 400 MHz to 2.4 GHz) is so full of strong signals, the use of a tracking filter is desired as a solution to improve the performance and prevent second-order IMD products or other undesired overload effects. The dc control voltage needed for the filter can easily be derived from the VCO control voltage. There may be a small dc offset, depending on the IF used.

3-6-1 Diode Performance

The capacitance versus voltage curves of a varactor diode depend on the variation of the impurity density with the distance from the junction. When the distribution is constant, there is an "abrupt junction" and capacitance follows the law

$$C = \frac{K}{(V_d + V)^{1/2}} \tag{3-183}$$

where V_d is the contact potential of the diode and V is applied voltage.

Such a junction is well approximated by an alloyed junction diode. Other impurity distribution profiles give rise to other variations, and the above equation is usually modified to

$$C = \frac{K}{(V_d + V)^n} \tag{3-184}$$

where n depends on the diffusion profile and $C_0 = K/V_d^n$.

A so-called graded junction, having a linear decrease in impurity density with the distance from the junction, has a value of n. This is approximated in a diffused junction.

In all cases these are theoretical equations, and limitations on the control of the impurity pattern can result in a curve that does not have such a simple expression. In this case, the coefficient n is thought of as varying with voltage. If the impurity density increases away from the junction, a value of n higher than 0.5 can be obtained. Such junctions are called hyperabrupt. A typical n value for a hyperabrupt junction is about 0.75. Such capacitors are used primarily to achieve a large tuning range for a given voltage change. Figure 3-144 shows the capacitance–voltage variation for the abrupt and graded junctions as well as for a particular hyperabrupt-junction diode. Varactor diodes are available from a number of manufacturers, such as Motorola, Siemens, and Philips. Maximum values range from a few to several hundred picofarads, and useful capacitance ratios range from about 5 to 15.

Figure 3-145 shows three typical circuits that are used with varactor tuning diodes. In all cases, the voltage is applied through a large resistor R_e or, better, an RF choke in series with a small resistor. The resistance is shunted across the lower diode and may be converted to a shunt load resistor across the inductance to estimate Q. The diode also has losses that may result in lowering the circuit Q at high capacitance, when the frequency is sufficiently high. This must be considered in the circuit design.

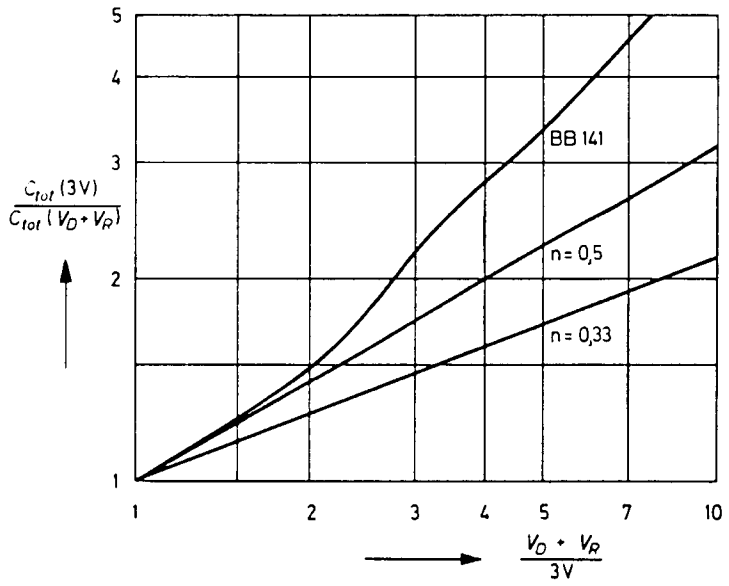


Figure 3-144 Voltage-dependent change of capacitance of different types of diodes. The BB141 is a hyperabrupt-junction diode with $n = 0.75$.

The frequency-dependent performance is not determined solely by applying the dc tuning voltage to Eq. (3-184). If the RF voltage is sufficient to drive the diode into conduction on peaks, an average current will flow in the circuits of Figure 3-145, which will increase the bias voltage. The current is impulsive, giving rise to various harmonics of the circuit. Even in the absence of conduction, Eq. (3-184) deals only with the small-signal capacitance. When the RF voltage varies over a relatively large range, the capacitance changes. In this case, Eq. (3-184) must be changed to

$$\frac{dQ}{dV} = \frac{K}{(V + V_d)^n} \tag{3-185}$$

Here Q is the charge on the capacitor. When this relation is substituted in the circuit differential equation, it produces a nonlinear differential equation, dependent on the parameter n. Thus, the varactor my generate direct current and harmonics of the fundamental frequency. Unless the diodes are driven into conduction at some point in the cycle, the direct current must remain zero.

The current of Figure 3-145c can be shown to eliminate the even harmonics, and it permits a substantially larger RF voltage without conduction than either circuit in Figure 3-145a or 3-145b. When $n = 0.5$, only the second harmonic is generated by the capacitor, and this can be eliminated by the back-to-back connection of the diode pair. Integrating Eq. (3-185), we find

$$Q + Q_A = \frac{K}{1 - n}(V + V_d)^{1-n}$$

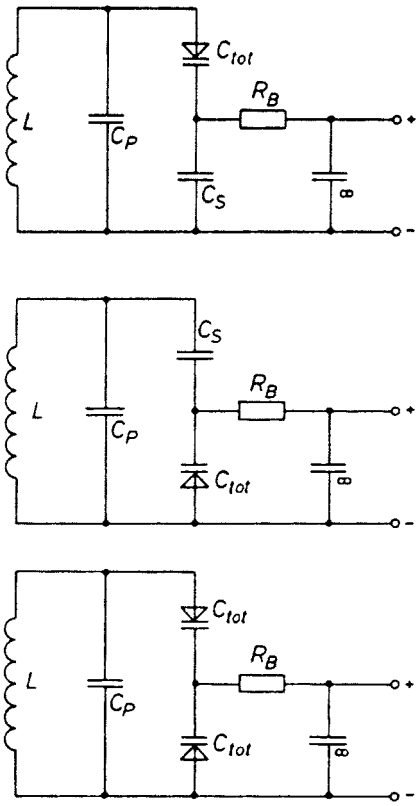


Figure 3-145 Various configurations to apply tuning diodes in a tuned circuit. The bottom version shows the lowest distortion.

$$= \frac{C_v}{1-n}(V + V_d) \tag{3-186}$$

C_v is the value of Eq. (3-184) for applied voltage V, and Q_A is a constant of integration. By letting $V = V_1 + v$ and $Q = Q_1 + q$, where the lowercase letters represent the varying RF and the uppercase letters indicate the values of bias when RF is absent, it follows that

$$q + Q_1 + Q_A = \frac{K}{1-n}[v + (V_1 + V_d)]^{1-n} \tag{3-187}$$

$$1 + \frac{v}{V'} = \left(1 + \frac{q}{Q'}\right)^{1/(1-n)} \tag{3-188}$$

where $V' = V_1 + V_d$ and $Q' = Q_1 + Q_A$. For the back-to-back connection of identical diodes, $K_{11} = K_{12} = K_1$, $V_1' = V_2' = V'$, $Q_1' = Q_2' = Q'$, $q = q_1 = -q_2$, and $v = v_1 - v_2$. Here the new subscripts 1 and 2 refer to the top and bottom diodes, respectively; v is the RF voltage across

the pair in series, and q is the charge transferred through the pair in series. This notation results in

$$\frac{v}{V'} \equiv \frac{v_1 - v_2}{V'} = \left(1 - \frac{q}{Q'}\right)^{1/(1-n)} - \left(1 - \frac{q}{Q'}\right)^{1/(1-n)} \tag{3-189}$$

For all n, this eliminates the even powers of q, and hence the even harmonics. This can be shown by expanding Eq. (3-189) in a series expansion and performing term by term combination of the equal powers of q. In the particular case $n = \frac{1}{2}$, $v/V' = 4q/Q'$, and the circuit becomes linear.

The equations hold as long as the absolute value of v_1/V' is less than unity, so that there is no conduction. At the point of conduction, the total value of v/V' may be calculated by noting that when $v_1/V' = 1$, $q/Q' = -1$, so $q_2/Q' = 1$, $v_2/V = 3$, and $v/V = -4$. The single-diode circuits conduct at $v/V' = -1$, so the peak RF voltage should not exceed this. The back-to-back configuration can provide a fourfold increase in RF voltage handling over the single diode. For all values of n, the back-to-back configuration allows an increase in the peak-to-peak voltage without conduction. For some hyperabrupt values of n, such that $1/(1 - n)$ is an integer, many of the higher-order odd harmonics are eliminated, although only $n = \frac{1}{2}$ provides elimination of the third harmonic. For example, $n = \frac{2}{3}$ results in $1/(1 - n) = 3$. The fifth harmonic and higher odd harmonics are eliminated, and the peak-to-peak RF without conduction is increased eightfold; for $n = \frac{3}{4}$ the harmonics 7 and above are eliminated, and the RF peak is increased 16 times. It must be noted in these cases that the RF peak at the fundamental may not increase so much, since the RF voltage includes the harmonic voltages.

Since the equations are only approximate, not all harmonics are eliminated, and the RF voltage at conduction, for the back-to-back circuit, may be different from that predicted. For example, abrupt-junction diodes tend to have n of about 0.46–0.48 rather than exactly 0.5. Hyperabrupt junctions tend to have substantial changes in n with voltage. The diode illustrated in Figure 3-144 shows a variation from about 0.6 at low bias to about 0.9 at higher voltages, with wiggles from 0.67 to 1.1 in the midrange. The value of V_d for varactor diodes tends to be in the vicinity of 0.7 V.

3-6-2 A VHF Example

The application of tuning diodes in double-tuned circuits in TV tuners has been common practice for many years. Figure 3-146 shows the circuit diagram. The input impedance of 50 Ω gets transformed up to 10 kΩ. The tuned circuits consist of the 0.3-μH inductor and two sets of antiparallel diodes. By dividing the RF current in the tuned circuit and using several diodes instead of just one pair, intermodulation distortion is reduced.

The coupling between the two tuned circuits is tuned via the 6-nH inductor that is common to both circuits. This type of inductance is usually printed on the circuit board. The diode parameters used for this application were equivalent to the Siemens BB515 diode. The frequency response of this circuit is shown in Figure 3-147.

The coupling is less than critical. This results in an insertion loss of about 2 dB and relatively steep passband sides. Once the circuit's large-signal performance (Figure 3-148) is seen, a third-order intercept point of about –2 dBm is not so unexpected. The reason for this poor performance is the high impedance (high L/C ratio), which provides a large RF voltage swing across the diodes. A better approach appears to be using even more diodes and at the same time changing the impedance ratio (L/C ratio).

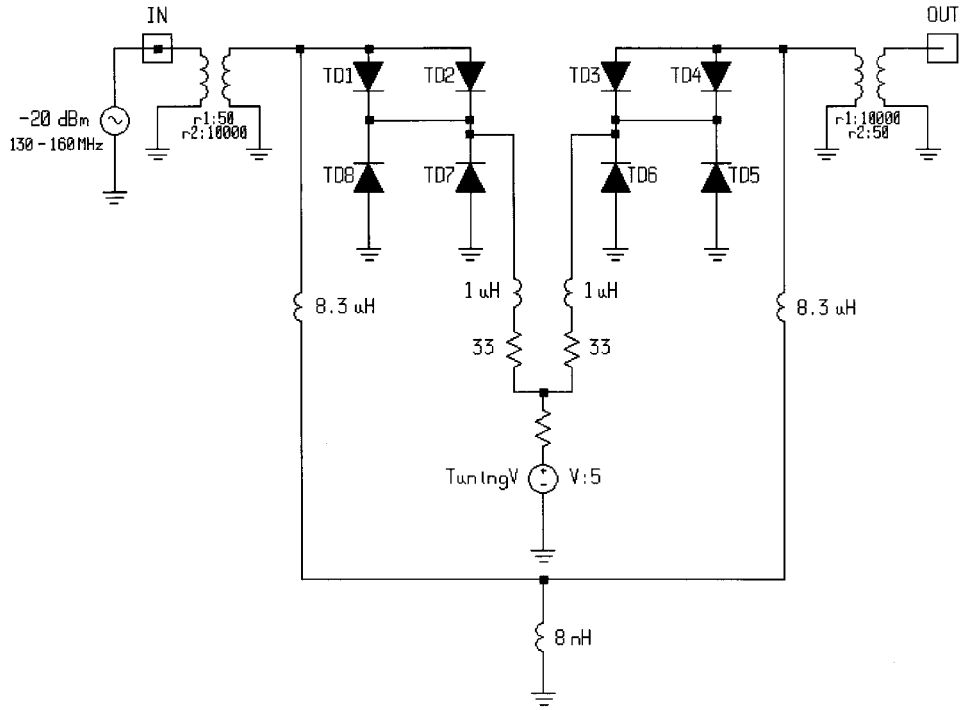


Figure 3-146 Double-tuned filter at 161 MHz using hyperabrupt-junction tuning diodes. By using several parallel diodes, the IMD performance improves.

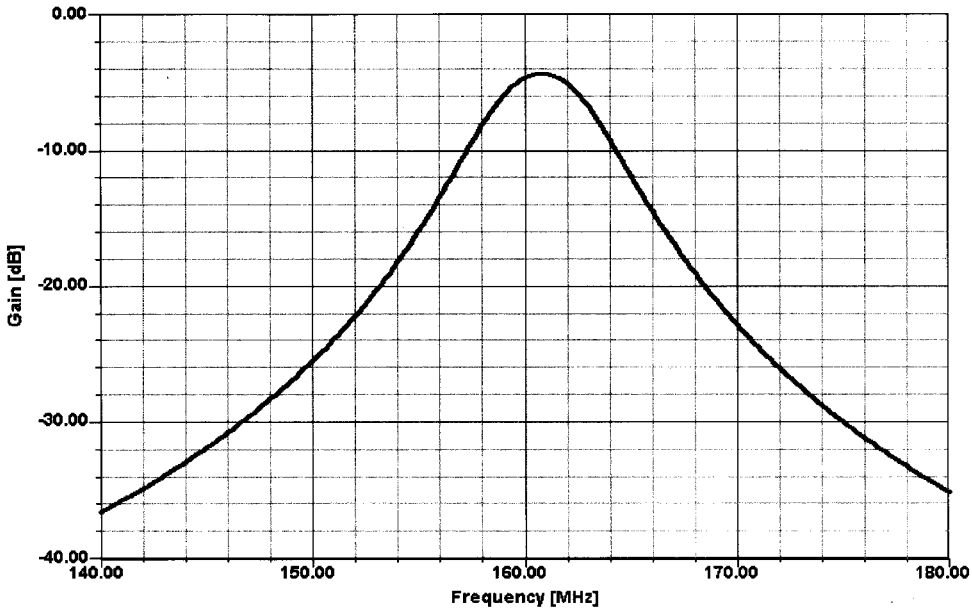


Figure 3-147 Frequency response of the tuned filter shown in Figure 3-146. The circuit is undercoupled (less than transitional coupling); $Q \times k < 1$.

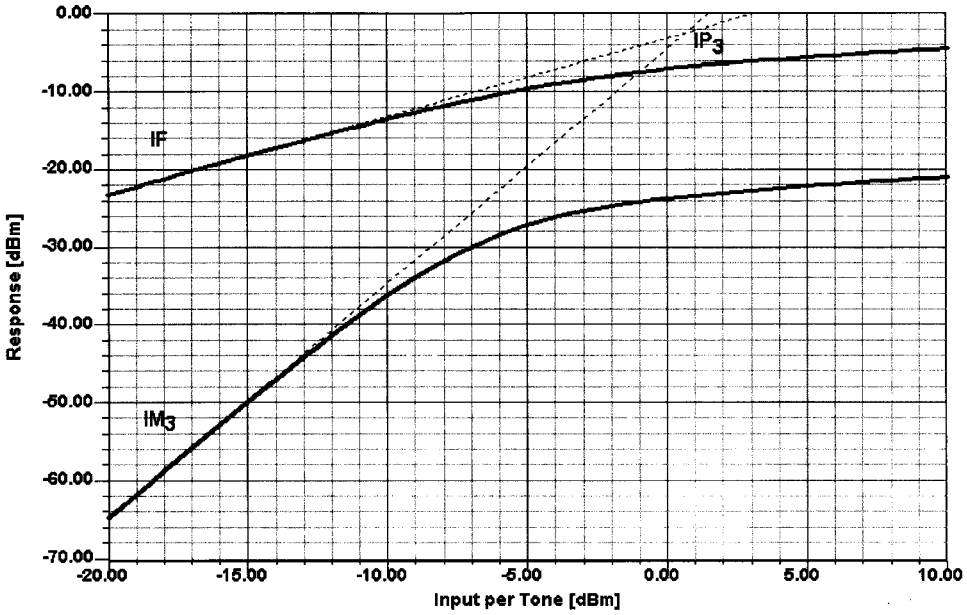


Figure 3-148 Prediction of intercept point of the double-tuned circuit shown in Figure 3-146. Note the compression of both the input signal and the IMD product.

3-6-3 An HF/VHF Voltage-Controlled Filter

The above example used a step-up procedure typically done by a tap at the input and output inductance. This method allows an impedance transformation; however, if one desires to change it into a series-tuned arrangement, it has to be done with a transformer. The large-signal conditions in the frequency range from 10 to 30 MHz on a medium- to large-sized antenna are equivalent to, if not worse than, the conditions for VHF operation. The only exception would be line of sight into the transmitter such as a tower in the middle of the city. Examples of such hostile conditions would be any large city, such as Munich, New York, Miami, Chicago, and San Francisco, where the authors had significant experience with intermodulation distortion problems.

By translating the circuit into a low-impedance-drive arrangement, building it symmetrically, and reducing the tuning range somewhat by capacitors and many series diode pairs, the filter's large-signal performance was significantly improved. The tuning diodes had about 125 pF at 1 Vdc per unit. Initial tests trying to use a high-capacitance diode, such as the BB112, resulted in much higher IMD products. The B112 has a 1-V capacitance of 470 pF down to about 20 pF at 8 V. Standard diodes show about 30 pF at 1 V, while the diodes used in the experiment had a 125-pF capacitance at 1 V. These experimental diodes were made available from a well-known manufacturer for this evaluation. Figure 3-149 shows the circuit diagram.

A step-down transformer drives the tuned circuit with a source impedance of 12.5 Ω. The circuit is symmetrical. There are two 0.8-μH inductors and 2×5 diodes in the loop. The output of the circuit is transformed back up to 50 Ω. The tuning voltage is supplied via a heavily filtered arrangement. This circuit has actually been implemented in Rohde & Schwarz field strength meter equipment. They guarantee an intercept point of about +20 dBm. The selectivity of these circuits is quite reasonable and mostly intended to reduce

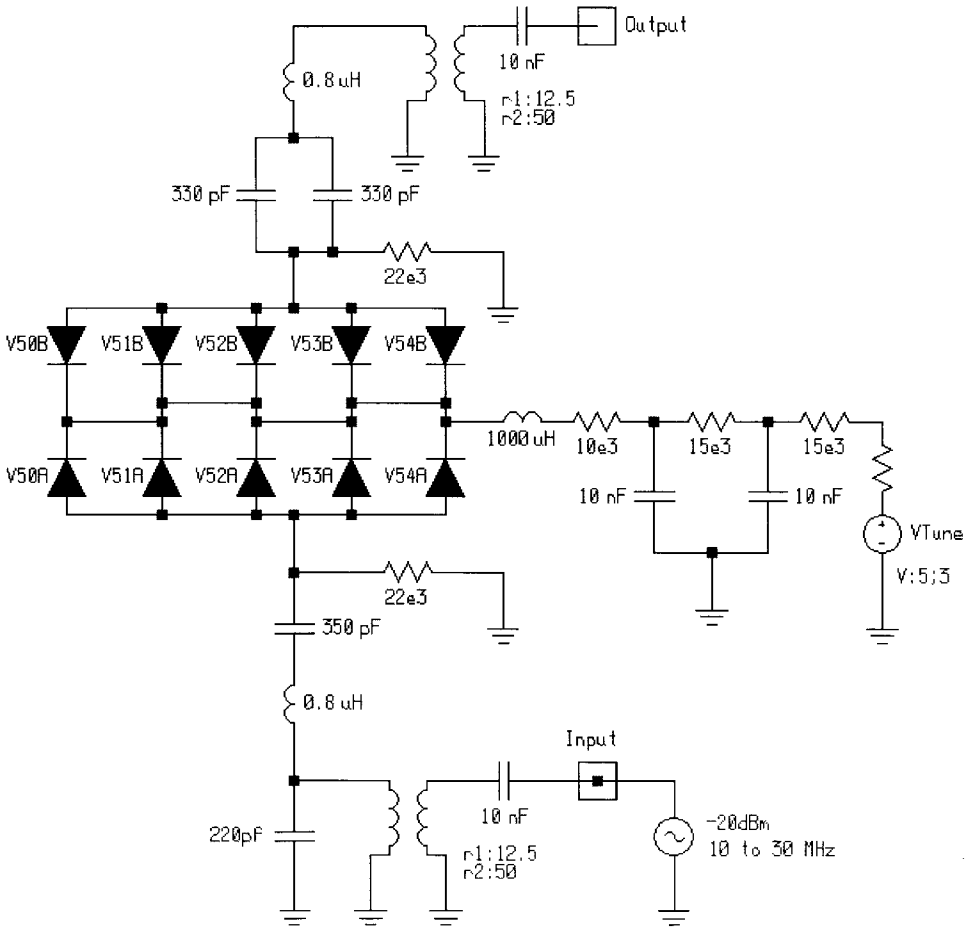


Figure 3-149 High dynamic range of a 10–20-MHz filter.

second-order intermodulation distortion products by about 10 dB, at the same time having a high third-order intercept point. Figure 3-150 shows the selectivity curves at two different tuning voltages.

The most interesting number, however, is the third-order intercept point (already determined as about 20 dBm), which still had to be simulated. This sometimes sounds like a contradiction: Once the measured values are available and are acceptable, why would one want—or, rather, need—to do a simulation besides the necessity to develop a high-input intercept filter? It was desirable to validate the nonlinear models and prove that the above-mentioned equations will hold true. This type of simulation is now more difficult because the number of nonlinear elements went from four to ten, and some numerical problems, such as convergence difficulties, can be expected, and the effect of harmonic frequency cancellation (compensation) can also be seen. By using diode combinations that result in $1(1 - n)$, n being an integer number, these IMD products can drastically be reduced. This implies that each of the five diode pairs has a selected value for n to meet this condition. Later, a final attempt will be made to improve the first VHF filter with the proper diode combinations.

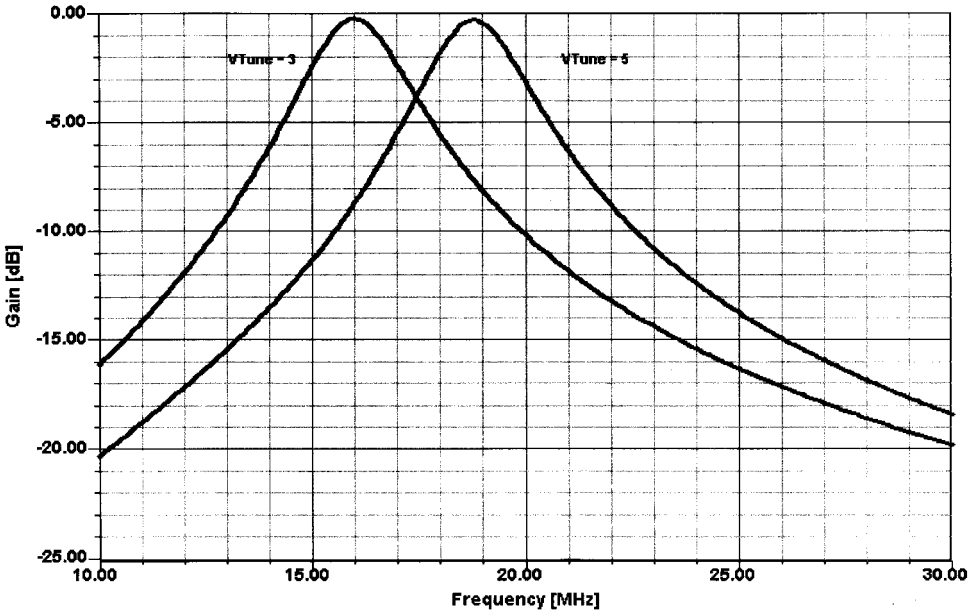


Figure 3-150 Frequency response of the filter shown in Figure 3-149. Filters of this type are intended to reduce second-order IMD by providing about 20-dB suppression at half the center frequency.

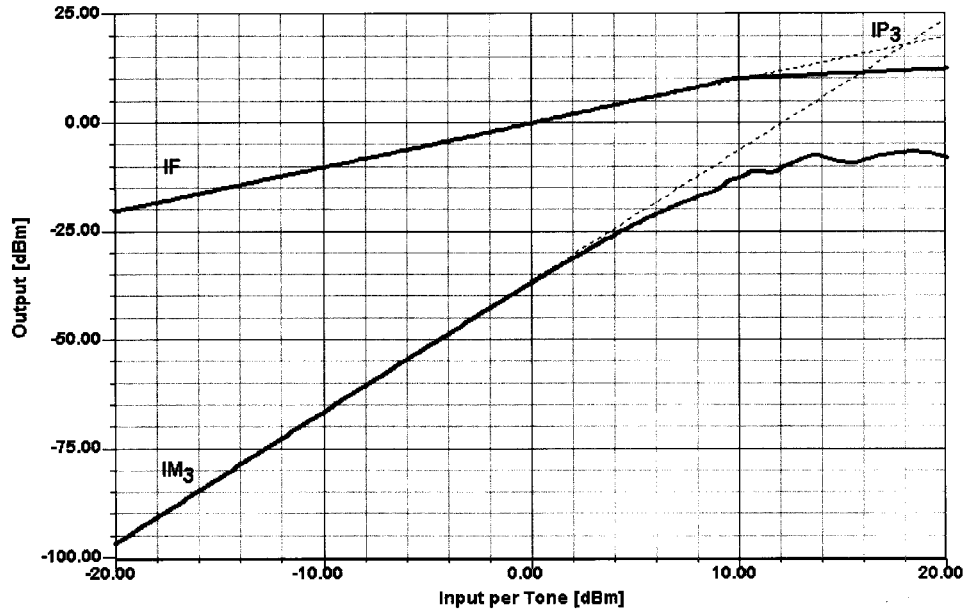


Figure 3-151 Prediction of third-order intercept point of the filter shown in Figure 3-149. The reason for the curves on the IMD plot lies in the interaction between the ten nonlinear devices.

Figure 3-151 shows the calculation of third-order intercept point for the arrangement shown in Figure 3-149. This number has now increased to +18 dBm. The difference between selection and measurement is approximately 2 dB. The reason for this more pessimistic value compared to the measured value probably has to do with slight variations of the exponent of the diodes' voltage-dependent capacitance. Therefore, the authors consider both the circuit performance and the simulation accuracy to be extremely good. This type of circuit, as mentioned earlier, has wide application in oscillator circuits. Danzeisen of Rohde & Schwarz was probably the first to have used this type of circuit by paralleling many diodes for improved performance.

3-6-4 Improving the VHF Filter

By selecting the appropriate diode combinations and circuit modifications as shown in Figure 3-152, a significant IMD improvement of the circuit shown in Figure 3-146 is obtained. A special shunt arrangement of several diodes with different exponents has been developed, which allows its value to be "adjusted." This circuit, for which a patent has been applied, showed an improvement from –2 dBm to 32 dBm. The same high operating Q was maintained. Figure 3-153 shows the frequency response after the modification. Note that a coupling slightly greater than critical ($Q \times k = 1.1$) has been selected.

3-6-5 Conclusion

After explaining some of the nonlinearities in mathematical terms, we have given some examples of voltage-tuned circuits and discussed their large-signal performance. This section has also shown that modern CAD tools can accurately predict the performance of such circuits.

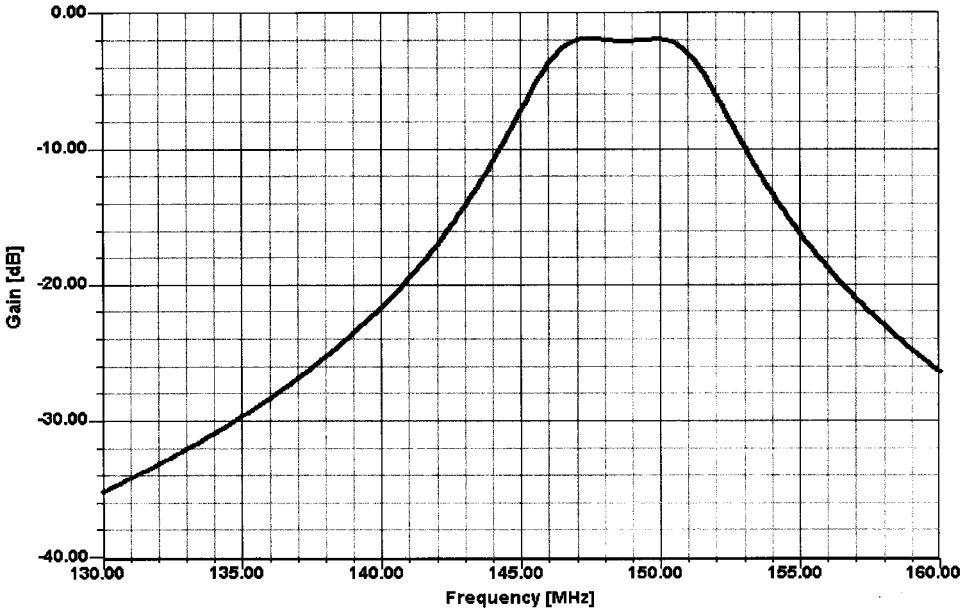


Figure 3-152 The frequency response of the improved circuit derived from Figure 3-146.

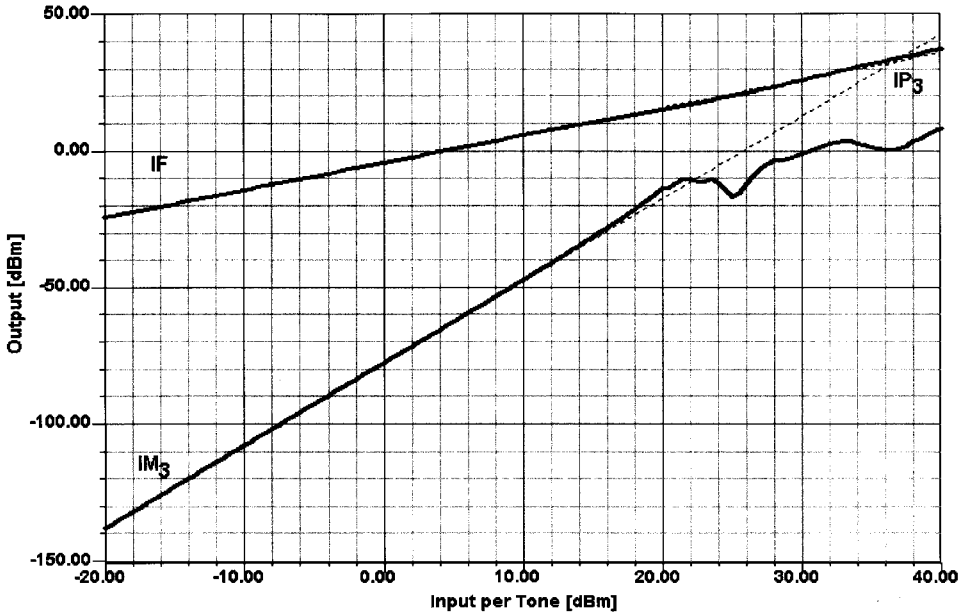


Figure 3-153 Predicted IMD performance of the improved version of the circuit shown in Figure 3-146. This type of improvement is significant for all applications. A patent is in the process of being obtained for this; therefore, the circuit cannot be disclosed at present.

3-7 DIFFERENTIAL AMPLIFIERS

The differential amplifier (Figure 3-154) goes back many years and also was the first step from TTL to ECL. The differential amplifier consists of an emitter follower (Q1) and a grounded-base stage (Q2). The constant-current generator (Q3) between ground and the emitters of Q1 and Q2 determines the current for these stages; the dc operating currents of Q1 and Q1 are the same. Most modern integrated circuits use this differential amplifier in one way or another.

If the circuit is operated about its point of symmetry, the dc component of the current through either Q1 or Q2 remains half of the emitter constant current through Q3 for all symmetrical driving signals. In addition, for input signals that are aperiodic, no even harmonics are generated. This assumes that the stage is driven symmetrically at both inputs. In our case, the differential amplifier has its second stage grounded, meaning that there is a base-to-ground capacitor, and the stage is asymmetrically fed. As the input voltage is increased, the differential stage becomes a limiter, which limits the current between plus and minus I_S, with I_S being the source current. The resulting output will become a square wave. This stage can also be used as a line receiver, translating a sine wave to an output logic suitable for driving emitter-coupled logic (ECL) stages.

One can assume that the differential amplifier has a transconductance of

$$g_m = \frac{\alpha g_{\text{in}}}{2} \tag{3-190}$$

where α is the ratio of collector to emitter current. Within the limits of the output being a sine wave, this stage has very low distortion. The output current now is determined by

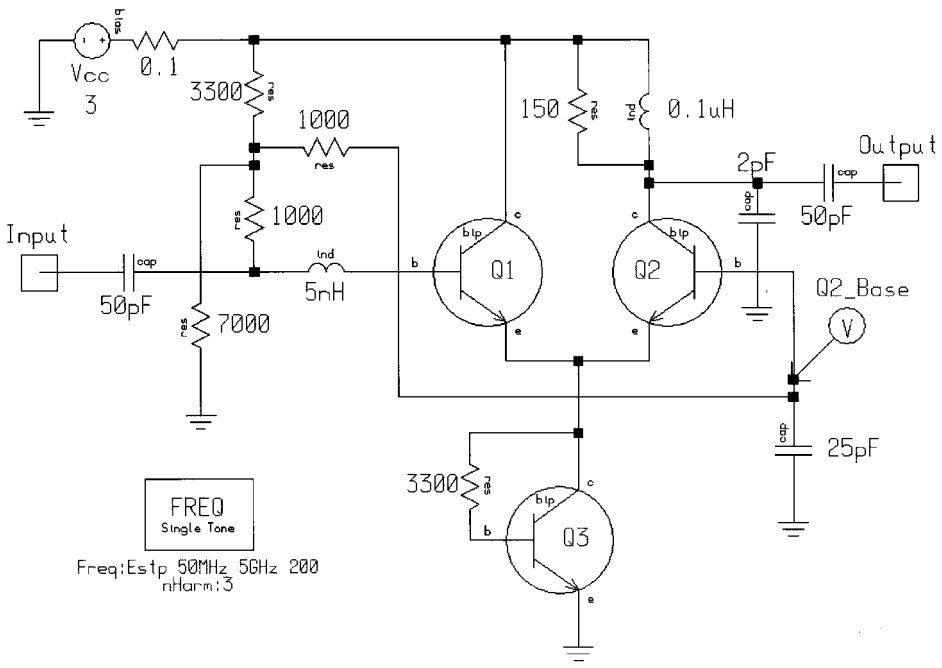


Figure 3-154 The differential amplifier schematic. The transistors are Siemens BFP420s. Figure 3-155 shows the circuit's frequency-dependent gain, matching, and noise performance.

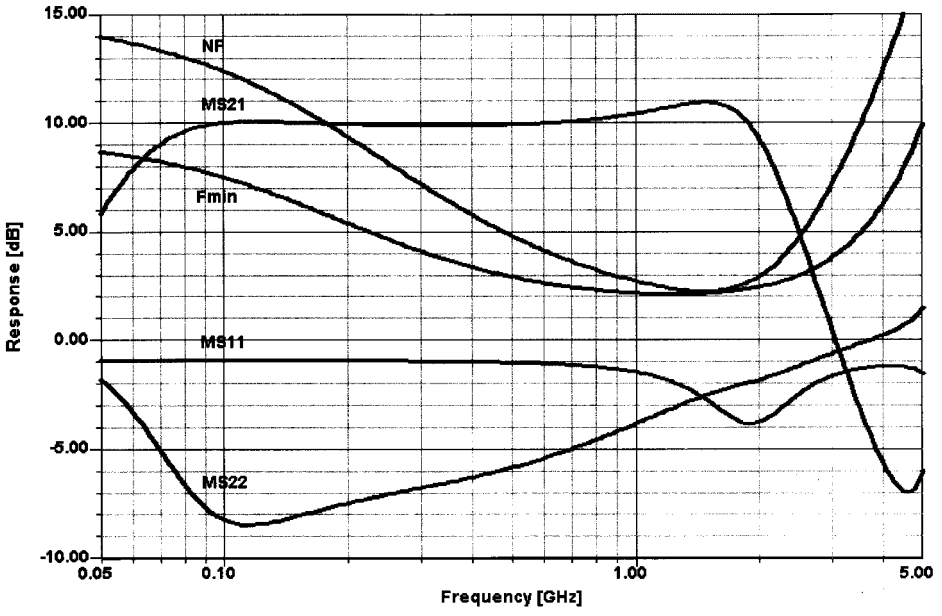


Figure 3-155 Frequency-dependent gain, matching, and noise performance of the differential amplifier.

$$i = \frac{I_{\text{in}}}{2} \tanh\left(\frac{x}{2} \cos \omega t\right) \tag{3-191}$$

where ωt is the input frequency and I_{in} is the current forced by Q3. The reason why the transconductance has to be divided by 2 comes from the fact that the current is split into two equal components for Q1 and Q2. Multiplying the new transconductance with the output load gives the voltage gain of this amplifier; its power gain can be obtained by multiplying this with $\sqrt{\text{output resistance} \div \text{load resistance}}$.

Having done the linear analysis and having learned that because of the hyperbolic tangent functions this amplifier can also be used as a line receiver, and converting sinusoidal waves to square waves, it may be useful to take a closer look at these issues. If we take a CAD oscilloscope, we see the resulting output waveform being a nonequal duty cycle output voltage, which results in high harmonic content. The amount of harmonic content can be controlled by the drive power and by the actual bias. Figure 3-156 shows the resulting output waveform.

If we now switch to a CAD spectrum analyzer, we can evaluate the harmonic content (Figure 3-157). Figure 3-158 shows the output transistor's dc I–V curves and ac load line under these conditions. Based on the previously defined conducting angle of the differential amplifier, one can optimize for harmonic content. Therefore, it is logical at this point to look at frequency multipliers, as they are sometimes needed. Finally, we are curious to see the so-called load line of the output transistor, which is, based on the interaction of all the stages, not looking the same way as the expected one, which we know from previous amplifier examples. Part of the reason for this is that for a certain voltage range, the collector voltage is in the negative region, and all kinds of saturation effects result in this surprising waveform. By reducing the input power, one would obtain a more "expected" load line. We invite readers with CAD capabilities to play with this example, as it gives a lot of insight into the operation of ECL stages.

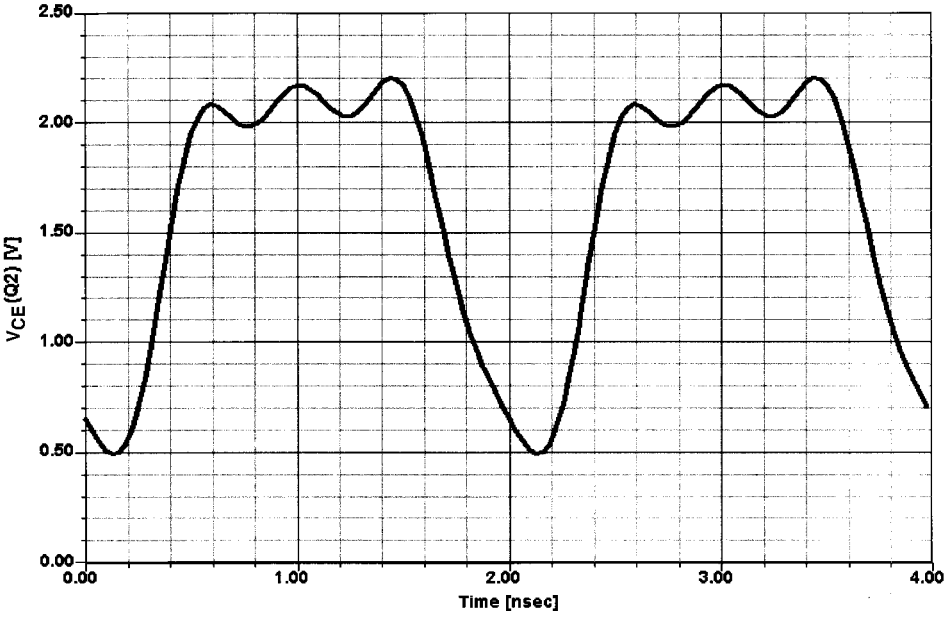


Figure 3-156 V_{CE} versus time for output transistor Q2 with a drive of 10 dBm.

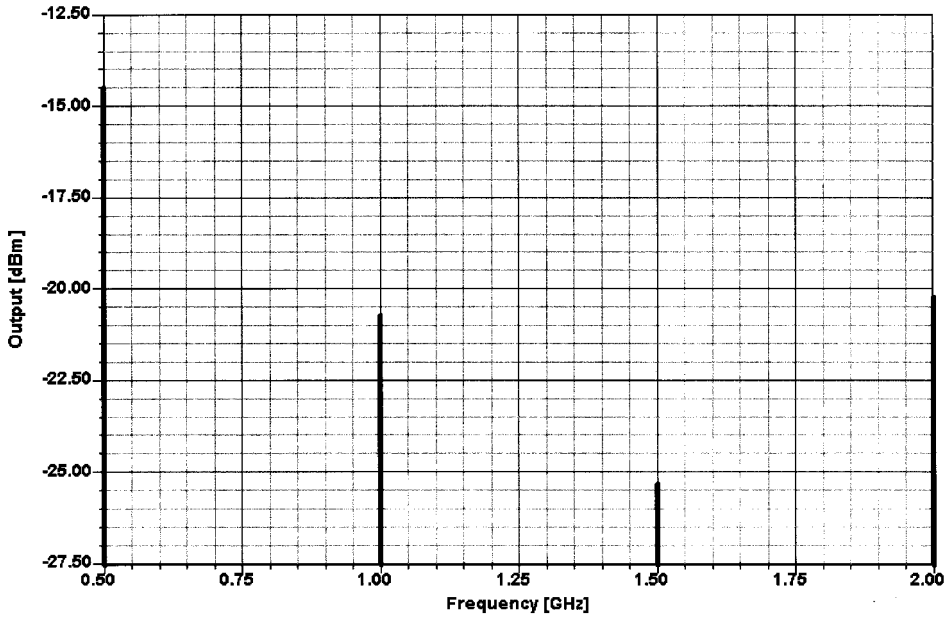


Figure 3-157 Output spectrum of the differential amplifier with a drive of 10 dBm. The biasing is not optimal for frequency multiplication.

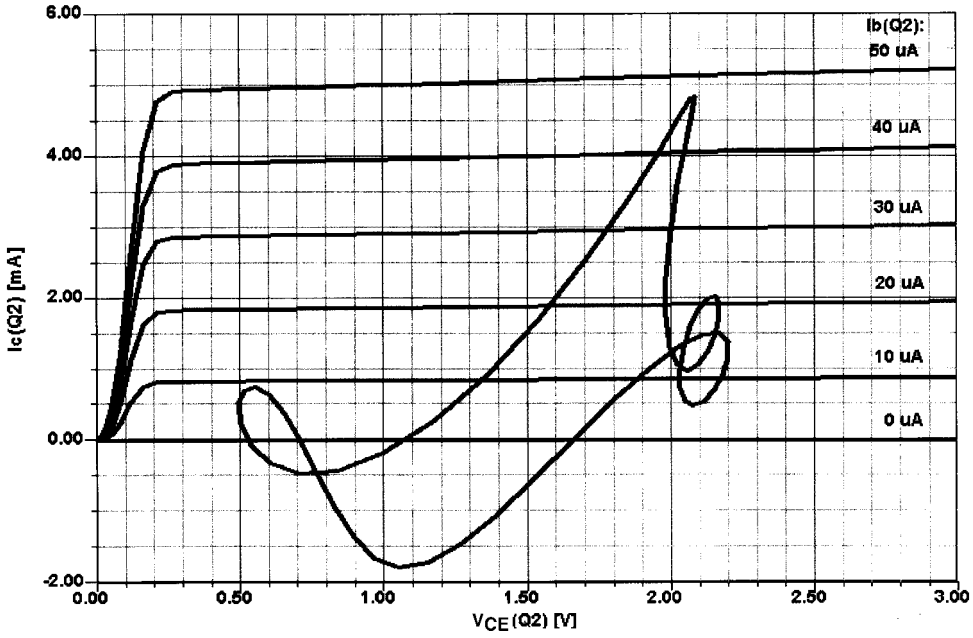


Figure 3-158 The dc *I–V* curves and ac load line for output transistor Q2 with a drive of 10 dBm.

3-8 FREQUENCY DOUBLERS

Having seen that the harmonic content invites us to bias a stage from Class A into Class B or C if our goal is a harmonically rich current, we will now evaluate a single stage that will be biased close to cutoff and driven by a relatively low frequency (500 MHz). To obtain the best results, we will use a parallel notch filter in the output, followed by a high-pass matching circuit to "catch" the desired harmonic. The input frequency is heavily suppressed by the notch filter. Based on its finite Q, there is a limited range over which this frequency can be varied with constant subharmonic attenuation. Figure 3-159 will be our test circuit.

This circuit consists of a bias transistor, Q2. This method has been discussed before and solves the temperature-dependent bias emitter thermal runaway problems through "temperature compensation." The purists among our readers may point out that the currents drawn by Q1 and Q2 are not quite identical (which would give 100% cancellation), but the approach shown is sufficient to do its job. Also, we have chosen a Class AB operation point, which is defined by Q1's 1-mA dc standing current. The conducting angle has not been optimized for the second harmonic—a task that the interested reader can do by using the conducting angle relationship shown at the beginning of this chapter. To give readers with access to nonlinear CAD some homework, we have not set this at the optimum point.

The second transistor, here called Q1 (because of its importance at RF), has a matching network at the input. This is recommended for optimum energy transfer. The matching condition can be validated by examining Figure 3-160. If we look at the higher frequency response, the input shows a possible trend of oscillation (negative loss); this could actually be eliminated by using a low-pass filter at the input instead of the high-pass filter used now. Findings of this type always make us nervous, and we hope that the nonlinear models are sufficiently accurate at higher frequencies to predict such behavior correctly. In the case of the FET doubler mentioned in Chapter 2, we found that the device modeling done by Compact (now Ansoft) gave the best results, the university answers were fairly close, and the device manufacturer data were worst. We have since become highly skeptical about device modeling unless we can validate it ourselves. Synergy Microwave has a complete set of test equipment in the laboratory, including network and spectrum analyzers with special software created by Rohde & Schwarz that can be used in addition to the Scout program to do in-house validation. This was achieved by the use of nonlinear optimization at a drive level of 0 dBm. One could say that this is similar to using a load-pull approach at the input, and then resorting to the large-signal S parameters obtained. Another note of caution: Capacitors at these low values tend to have a very large percentage error; for instance, a 1-pF capacitor, if not selected carefully, can vary between 0.7 and 1.3 pF—a 50% error. Needless to say, this is not acceptable for useful production.

The output consists of a high-pass filter and a notch filter with the purpose of reducing the subharmonic significantly. We have again done this on purpose, leaving room for the reader to come up with a more inventive output circuit like a Cauer-type (same as elliptical) low-pass filter with discrete notches at higher harmonics of the input frequency, such as 1.5 GHz, 2 GHz, and so on. In addition, we have shown three ways how to implement the inductance for the notch filter. This is equally applicable for other inductors in this value range. As frequency increases, this becomes more interesting but one needs to remember that many materials like silicon and gallium arsenide are extremely lossy. The same inductor approach has been used in Figure 5-61 of Chapter 5. For silicon, even the manufacturer becomes an issue. Figure 3-161 shows a top view of a "rectangular" inductor that actually approaches the spiral inductor; its implementation can be seen in Figure 3-162. Modeling this was a

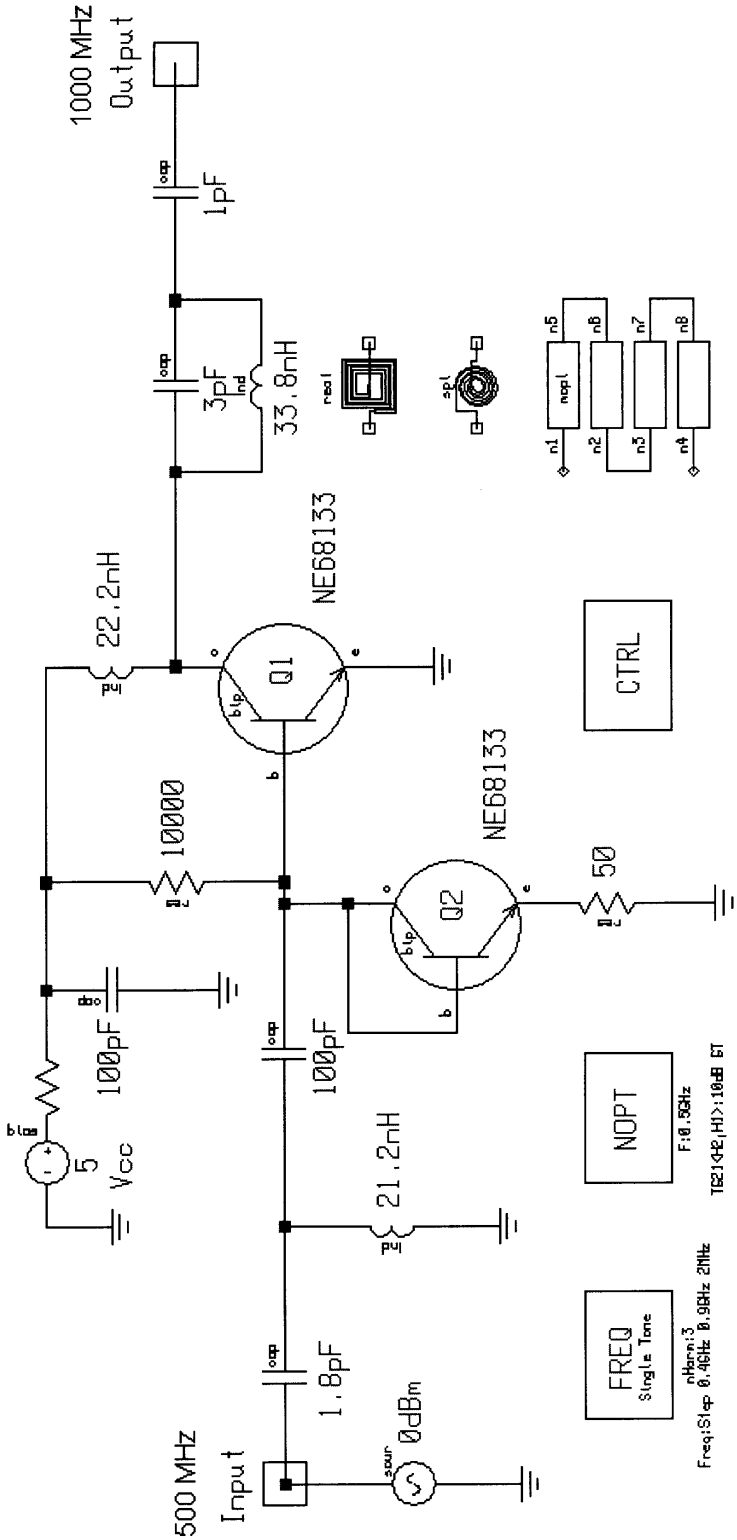


Figure 3-159 BJT frequency doubler schematic. A parallel-tuned trap in Q1's collector attenuates fundamental feedthrough.

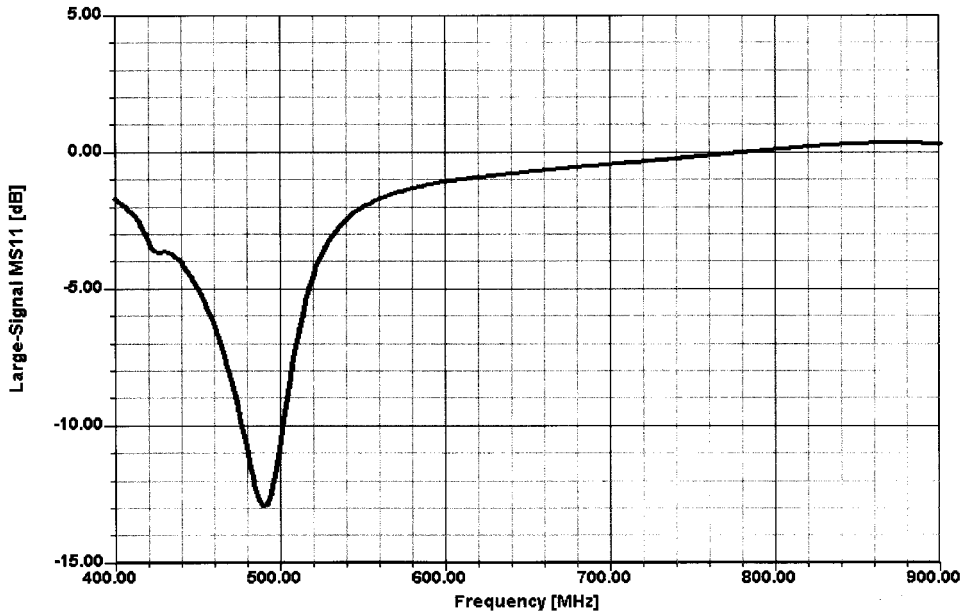


Figure 3-160 Large-signal MS_{11} for the BJT doubler. Note the positive region above 800 MHz.

nightmare, and during our Compact days, we actually had to modify the electromagnetic simulator to accommodate this and get results where measurements and simulation agreed. Figure 3-163 shows the measured and simulated Q as a function of frequency of a silicon-based inductor. In the case of other materials, such as PC board or GaAs, the models in the CAD tools are sufficient to predict the losses of the inductors. Following this initial exercise, Motorola then developed an improved silicon spiral inductor; its frequency-dependent Q is shown in Figure 3-164. Note that this Q was measured and not simulated.

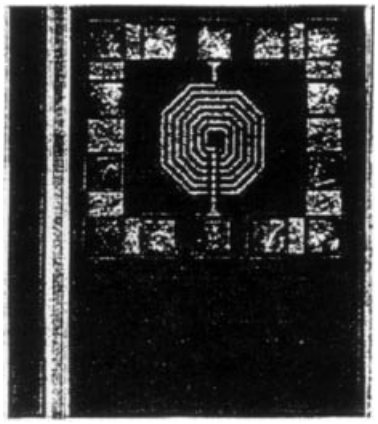


Figure 3-161 A 6.5-turn spiral inductor in silicon.

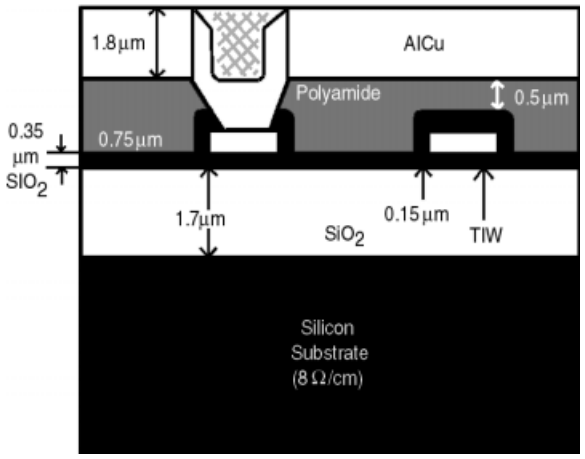


Figure 3-162 Cross section of the silicon inductor.

One of the reasons why we use a transistor multiplier is that we expect gain (because of the square-law characteristic, an FET would have been better tailored for a doubler); therefore, we have "measured" the gain of this circuit while driving it with 0 dBm. Another important factor is that, because of the way the CAD tool displays it, the gain is referenced to the 500-MHz input. Mentioning this is important; otherwise, the reader may look for the 1-GHz gain, which does not exist. Figure 3-165 shows the doubler gain.

By now we are curious to see the doubler in action, and by reviewing Figure 3-166 we can see the output spectrum. The output spectrum validates the statement that the output circuit, while doing a marvelous job of removing the 500-MHz input frequency, does not cure the problem of unwanted multiples of the input frequency. Again, a Cauer low-pass filter, while introducing some possible additional loss, should be considered for this circuit to have a bandpass response. On the other hand, many oscillators have very little harmonic suppression, and the same type of postfiltering needs to be considered. There are two-transistor solutions available also, such as push–push for frequency doubling (push–pull would

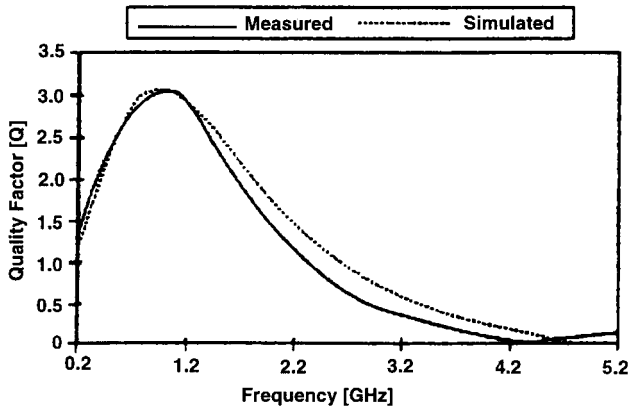


Figure 3-163 Measured and simulated *Q* versus frequency for the simulated inductor.

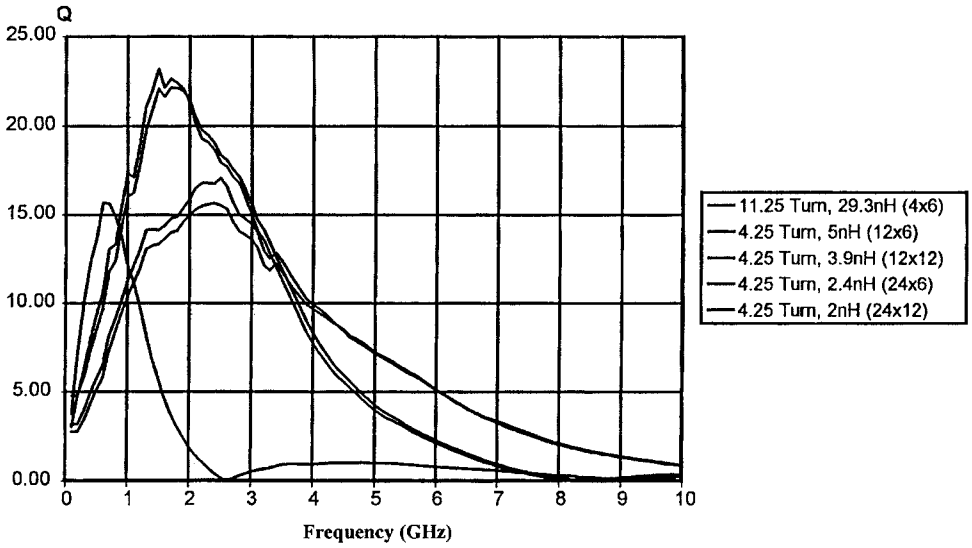


Figure 3-164 *Q* versus frequency for Motorola improved inductor.

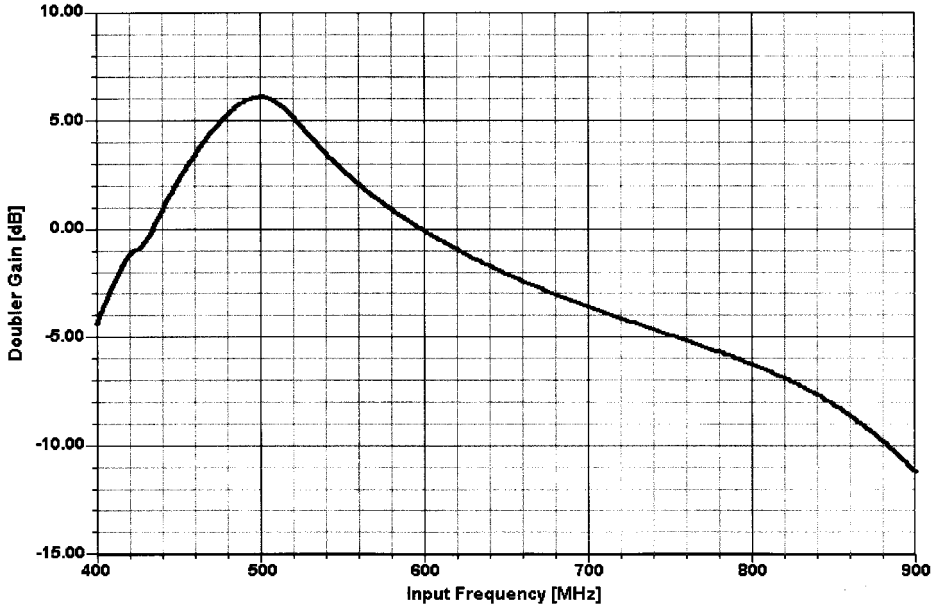


Figure 3-165 Frequency-dependent gain of the BJT doubler.

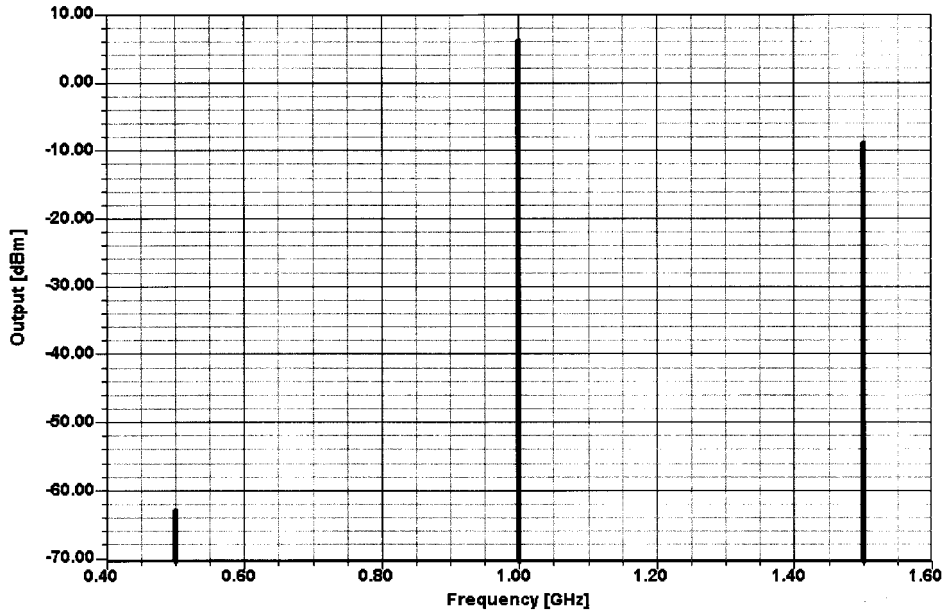


Figure 3-166 Output spectrum of the BJT doubler.

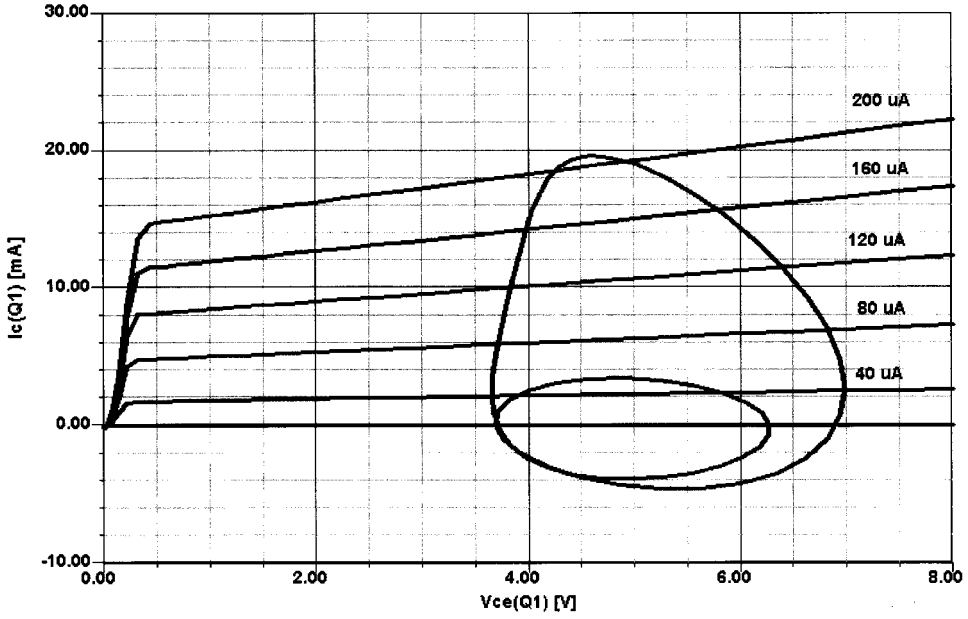


Figure 3-167 The dc *I–V* curves and ac load line for RF transistor Q1 in the BJT frequency doubler.

eliminate the desired second harmonic quite well). Actually, two diodes could also have been used for doubling, but since this requires transformers and a postamplifier, we have shown the preferred application. The only advantage of the diode application is that it is wideband, while this BJT circuit, as we have just learned, must be made narrowband for reasons of spectral purity.

Finally, we are curious to see the actual operating state of this amplifier, and this can be done by analyzing the load line (Figure 3-167). This is as complicated as the analysis for the differential amplifier. Because of the stored energy in the various tuned circuits, the load line looks quite impressive and would probably be difficult to predict without a nonlinear CAD tool.

3-9 MULTISTAGE AMPLIFIERS WITH AUTOMATIC GAIN CONTROL (AGC)

Having discussed the basic differential amplifier, we will now evaluate the performance of the MC1350/1490 video amplifier chip from Motorola. Most of the circuitry in the chip is self-explanatory. Figure 3-168 shows its inner circuit, and Figure 3-169 show a recommended application. By grounding one side of the input, we effectively get the same performance as the differential amplifier just discussed, with one difference being that the input is now fed to the constant-current generator. To minimize the distortion at the two lower transistors, they are decoupled by a 66-Ω resistor. Figure 3-170 shows the frequency-dependent gain, matching, and noise figure. Since this particular device has been specifically designed for AGC purposes, we also show in Figure 3-171 its gain reduction versus applied AGC voltage.

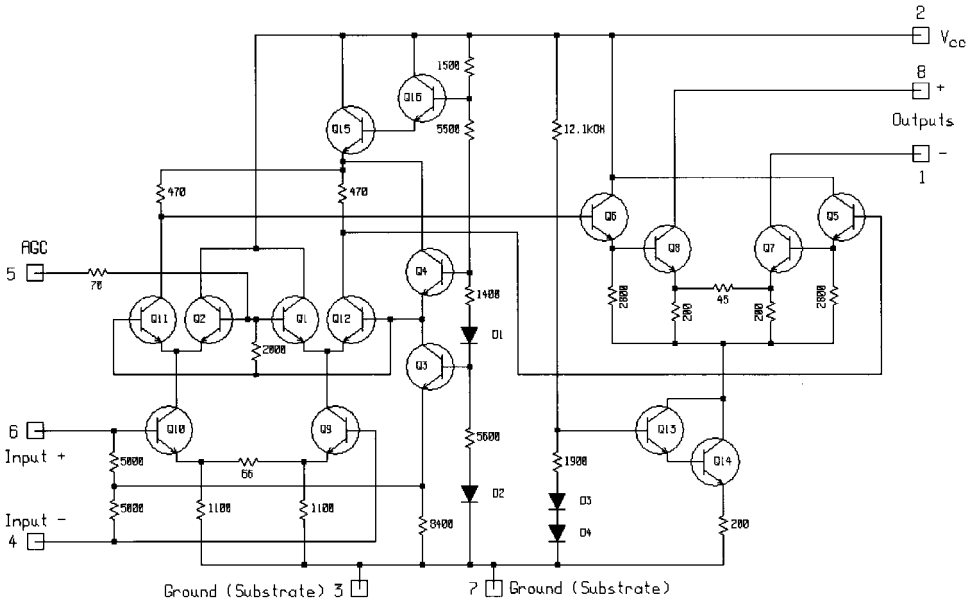


Figure 3-168 Schematic of the Motorola MC1350/1490 entered for simulation.

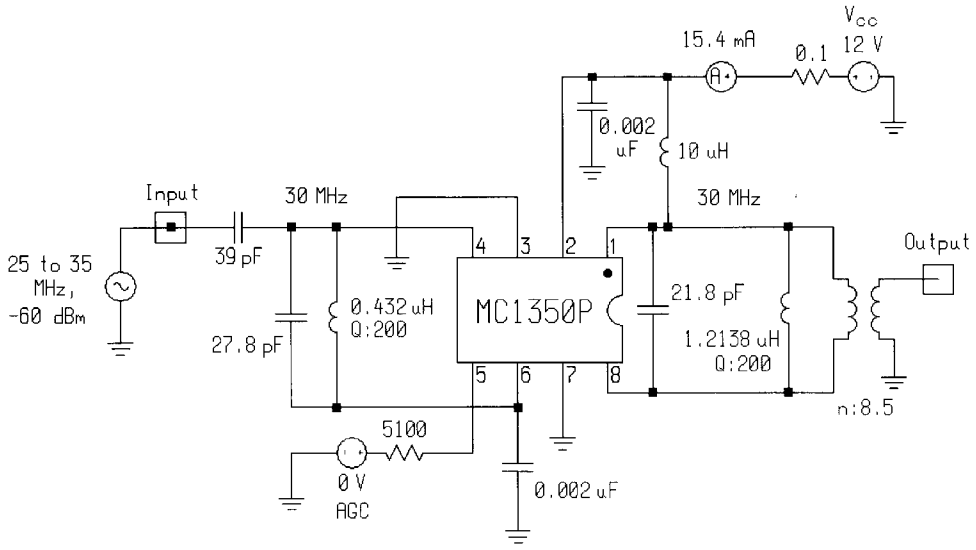


Figure 3-169 MC1350/1490 30-MHz application circuit.

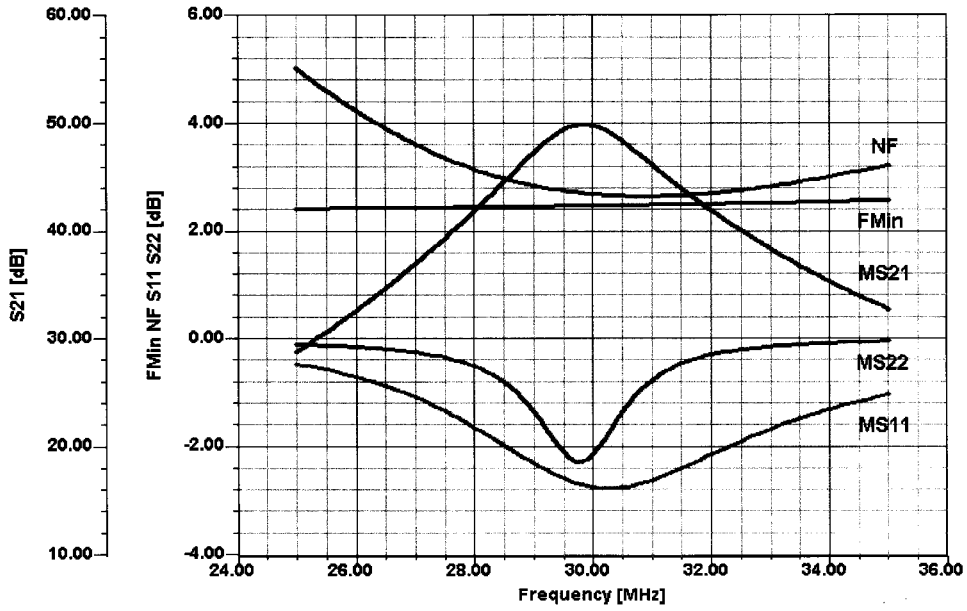


Figure 3-170 Frequency-dependent gain, matching, and noise performance of the MC1350/1490 in the circuit of Figure 3-169.

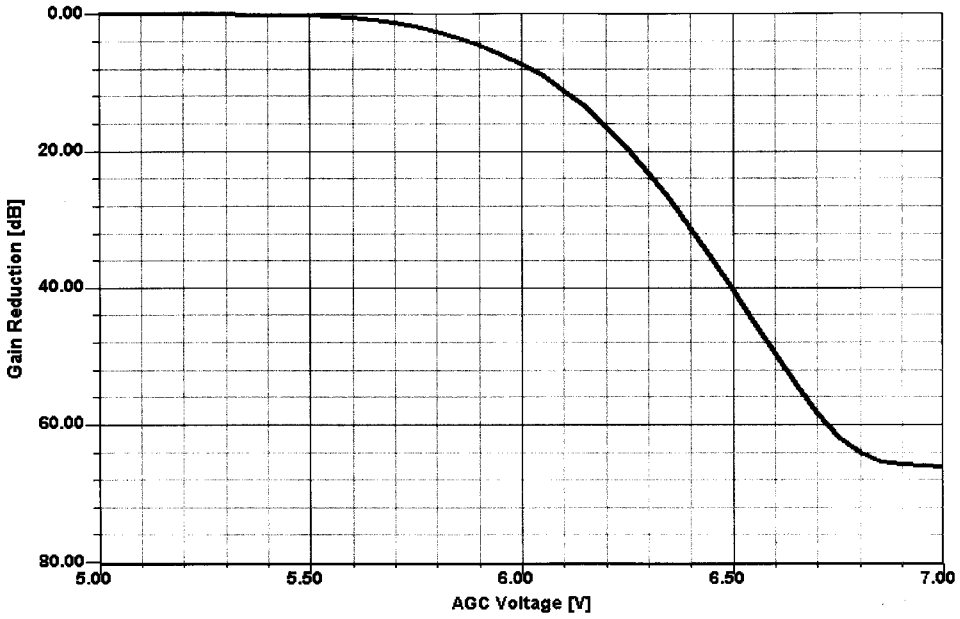


Figure 3-171 Gain reduction versus applied AGC voltage for the MC1350/1490 IC.

3-10 BIASING

A lot has been written about biasing, and since the purpose of this book is not to focus too much on dc, the following examples including RF consequences have been selected and considered sufficient. Early application reports by Motorola and others have provided a wealth of data. As a side effect of the introduction of the wireless era, we find that the engineers are so pressed for time that the number of complete application reports has approached zero. This is one of the reasons why we have tried to minimize the mathematical basis of this book (some is still needed), while still providing practical assistance. Figure 3-172 shows a textbook type of approach for RF.

In order to decouple the transistor from the biasing, we use RF chokes and dc decoupling (bypass) capacitors. This is the technique used in previous examples, so one might ask, "What's new?" As frequency increases, these inductors are either not manufacturable or have such a low Q that their use becomes questionable. This is the point where one may introduce the so-called distributed elements.

Figure 3-173 shows the very same circuit but resorting to distributed rather than lumped elements. The newly introduced elements (part of any good, up-to-date CAD tool; see element library for how to use their physical element descriptions) are described next.

- *Transmission Line.* Any printed connection between two points on a circuit board is a transmission line (Figure 3-174). Its characteristic electrical impedance is a function of the square root of the dielectric constant (ε_r), the width, metallization, thickness and height above substrate of the line, and the loss tangent of the substrate. Since lines frequently have to be laid out in the form of curved connections or have a bend in their direction, we have to add elements capable of describing the high-frequency conse-

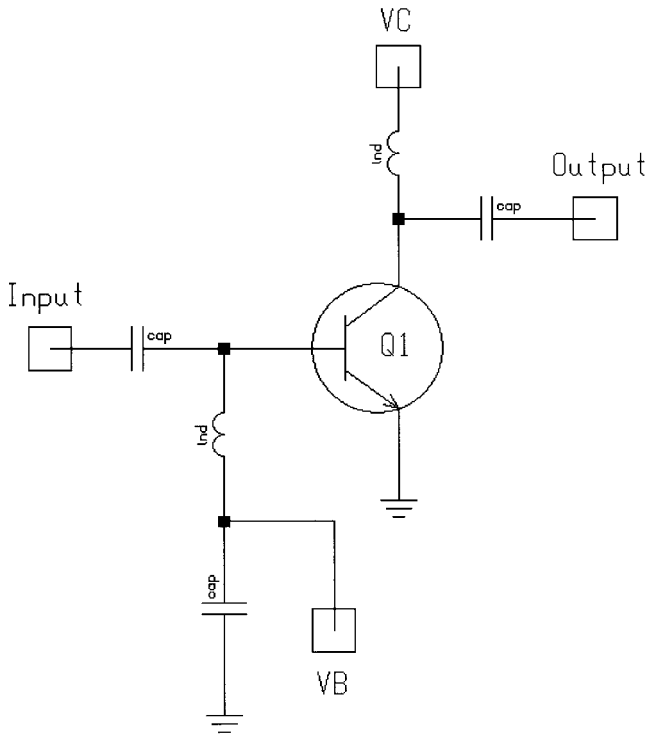


Figure 3-172 Simple BJT RF amplifier with lumped elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

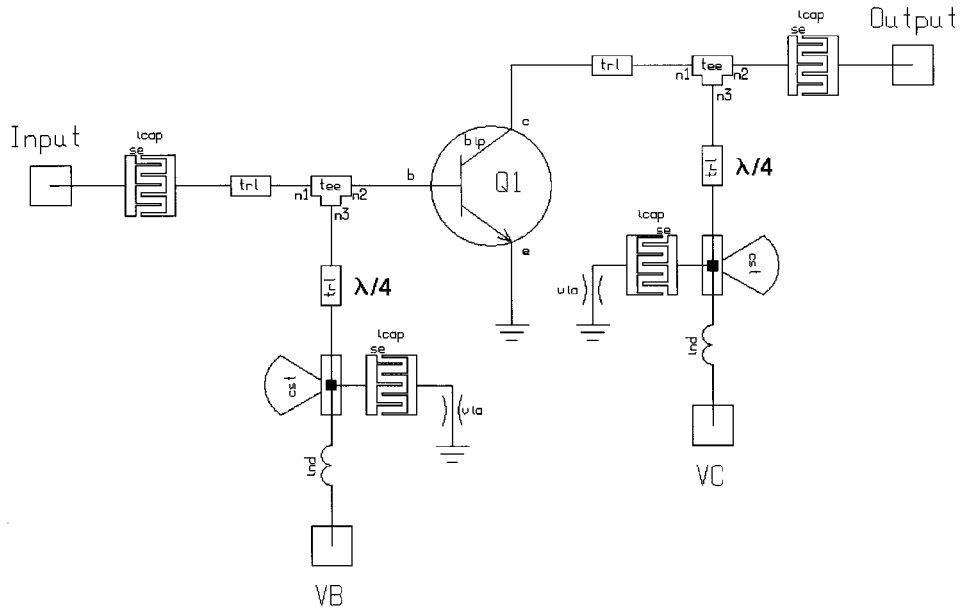


Figure 3-173 Simple BJT RF amplifier with distributed elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

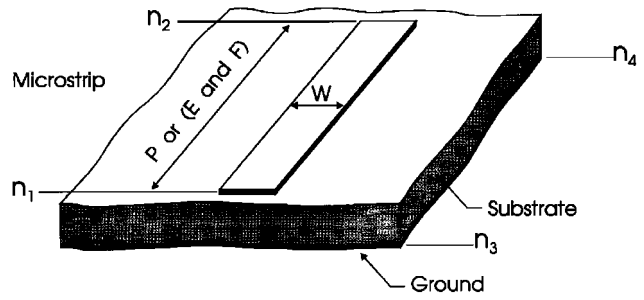


Figure 3-174 Transmission line in microstrip.

quences of such connections. Figures 3-175 and 3-176 show mitered and radial bend elements that perform this function.

- T *Junction, Cross, and* Y *Junction.* By the time a point like a collector or base, or its FET equivalent, spreads out into connections with other elements, we need additional modeling capability to describe ⊤ connections, crossings, and Y junctions. Figures 3-177, 3-178, and 3-179 show the way in which these connections must be modeled.

If the need exists, the standard inductances must be replaced with a transmission line whose length is λ/8 at the operating frequency. At higher frequencies, these transmission lines, however, then go into λ/4 resonant mode and later become capacitive. This type of design makes it fairly narrowband. A way around this is the use of printed inductors, as shown in Figures 3-180 and 3-181.

These inductors have a self-resonant frequency similar to the transmission line mentioned above, but the safety margin is significantly higher.

Talking about printed inductors, a logical extension of this is the printed transmission-line-based transformer, as shown in Figure 3-182. One can consider this as two interlaced rectangular inductors and, based on the substrate material, they are useful over a wide frequency range. Besides being used as a transformer, they can also be used to transit from

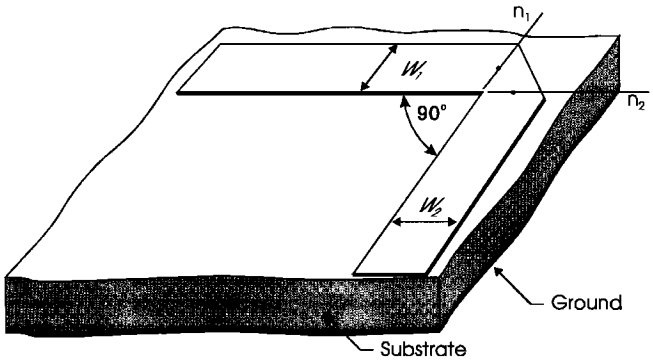


Figure 3-175 Mitered bend.

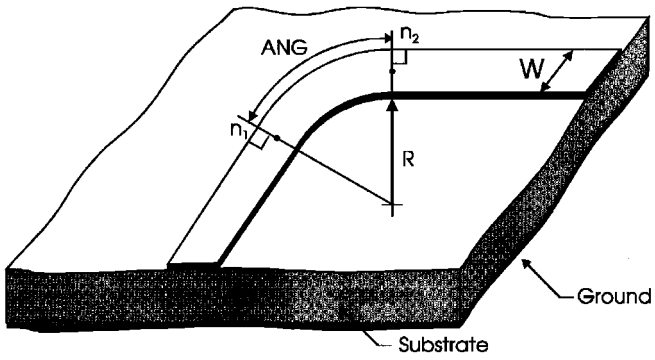


Figure 3-176 Radial bend.

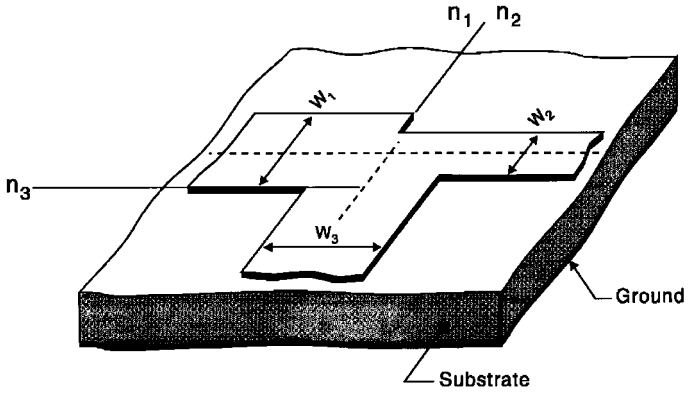


Figure 3-177 T junction.

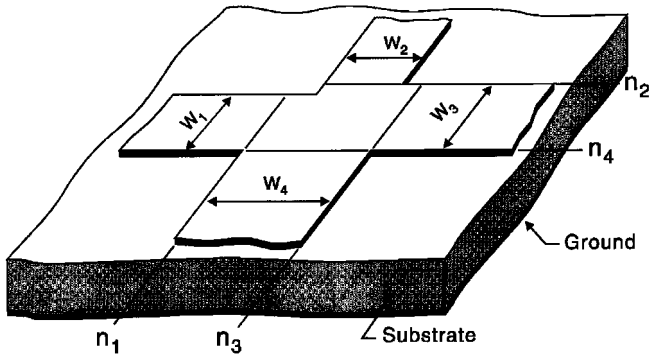


Figure 3-178 Cross.

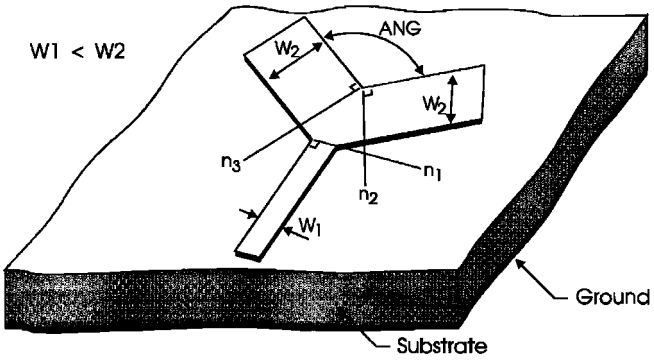


Figure 3-179 Y junction.

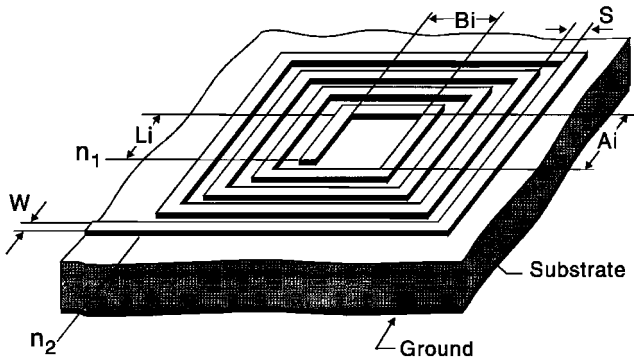


Figure 3-180 Rectangular inductor.

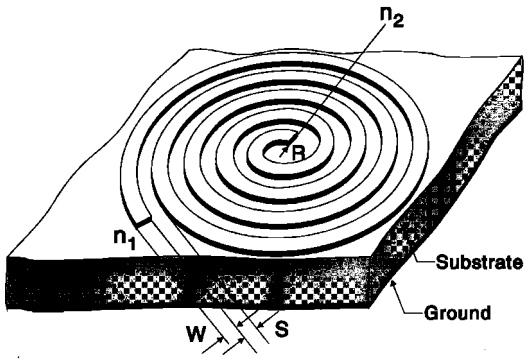


Figure 3-181 Spiral inductor.

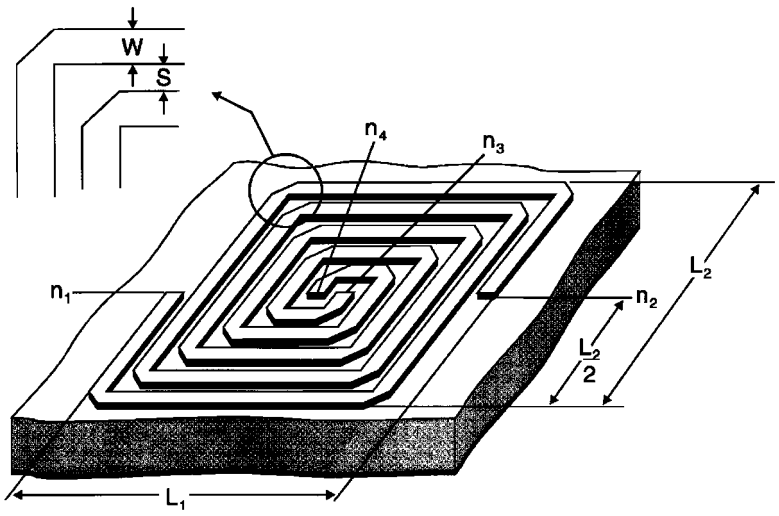


Figure 3-182 Transformer in microstrip.

unbalanced to balanced transmission provided that the difference in length from a connection point of view does not cause any problems (this is a layout issue).

A popular form of combining stages is the so-called Lange coupler (Figure 3-183) invented by the German Julius Lange. It is one of the major contributions in wideband applications.

Useful application of the Lange coupler probably starts at 4 GHz. It consists of parallel transmission lines with the appropriate connections as shown. Lange couplers are typically built with four, six, and eight fingers. The Ansoft Serenade product has a Lange coupler synthesis program that can be used to gain more insight into this coupler's application. We assume that other modern software has similar capabilities.

Where meander-type inductors are necessary, a neat way to implement and simulate them is to use the multiple coupled lines element (Figure 3-184) of the Serenade product, which both quickly and accurately calculates the behavior of the meander, including self-resonances and losses. We made use of this arrangement in our previous examples.

- *Interdigital Capacitors.* The issue of tolerances of small capacitors already has been brought up. The interdigital capacitor can be made on printed circuit board material as

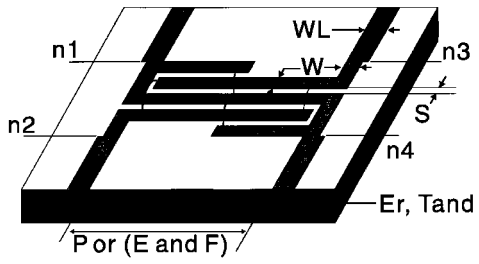


Figure 3-183 Four-strip Lange coupler.

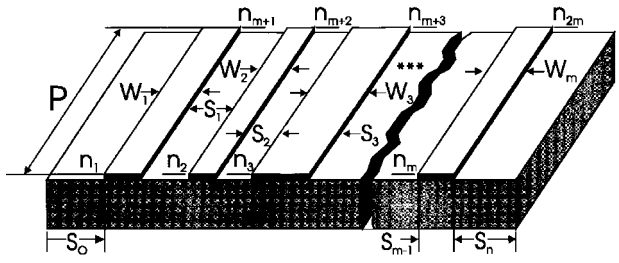


Figure 3-184 Multiple coupled lines element in microstrip.

well as gallium arsenide, and if its dimensions are continuous with the transmission-line width it does not cause any abrupt changes in impedance. This type of capacitor permits one to obtain very small values. An alternative to this is the use of transmission lines with a length of $\frac{3}{8}\lambda$. We learned earlier that a transmission line is inductive at frequencies below its fundamental resonance, is resistive at below resonance, and becomes capacitive above resonance. Again, bandwidth is also an issue. An interdigital capacitor consists of a number of parallel fingers as shown in Figure 3-185, and its capacitance can be varied by adjusting the number of fingers and their spacing. The advantage of the interdigital capacitor compared to discrete components is its low variation in value.

- *Radial Stubs.* The radial stub (Figure 3-186) is not much different from a $\lambda/4$ resonator, but its bandwidth is much greater than a simple transmission line. This is another way to ground the "cold" side of a transmission line or part of a circuit that needs to be grounded for RF. Of course, the interdigital capacitor comes in a version that is a combination of a capacitor and a via hole.
- *Via Holes.* The "cold" end of the transmission line, being considered either an inductor or a capacitor, needs to be connected to a sufficiently large copper backplane. One very efficient way to do this, especially if there is not enough copper left on the top of the board, is the use of via holes (Figure 3-187). One could theoretically generate

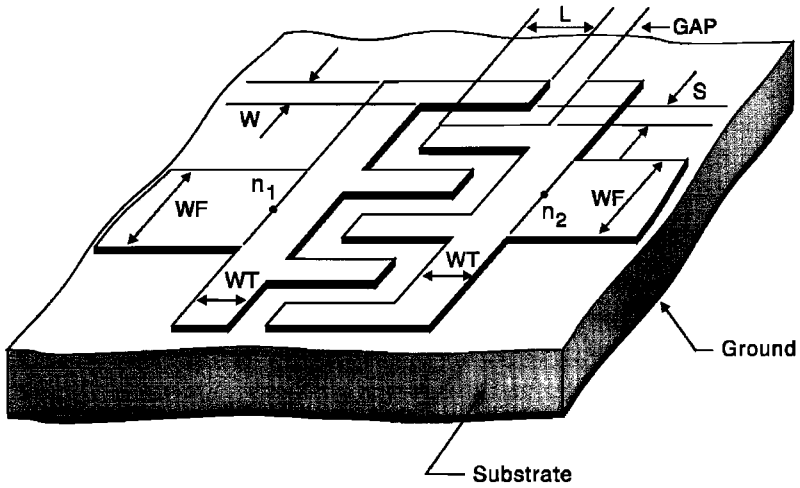


Figure 3-185 Interdigital capacitor.

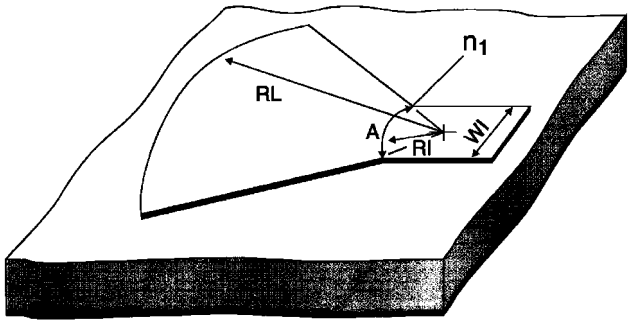


Figure 3-186 Radial stub.

a via hole with a rivet, but most manufacturing processes don't allow this; the normal solution is to use plated-through holes left open. In PC boards, via holes are typically cylindrical; on substrates like GaAs, they may be conical.

- *Correction Elements.* Although the behavior of an actual circuit proceeds regardless of our ability to measure and describe it, we do not enjoy this luxury in simulating circuit behavior in software. In a simulator, effects that are insufficiently described will be inaccurately simulated. For instance, segments of high-impedance transmission line (e.g., 120 Ω) are frequently used for dc feeds and RF chokes. In predicting the effect of a transition from 120-Ω line to 50-Ω line, a simulator must be alerted to the discontinuity so it can do the necessary mathematical corrections to account for the impedance jump. To do this, a specific circuit element, the STEP (Figure 3-188), must be inserted between the 120-Ω and 50-Ω line elements in the simulation circuit file. In addition to substrate data, we characterize a STEP by providing the widths of its input and output lines. The element itself has no physical length.

A similar correction is necessary if a transmission line is used as a resonator or just "left open" at one end. Such a transmission line tends to radiate and, because of its high-impedance properties, it reacts differently as far as its electrical length is concerned. A zero-length one-port element, the OPEN (Figure 3-189), must be added to such a line for mathematically correct calculation.

We end this excursion into distributed elements here. Interested readers may want to obtain a CD containing the element library of their favorite CAD tools that combine capabilities in the microwave area with a full set of distributed circuit elements. It is most important to keep

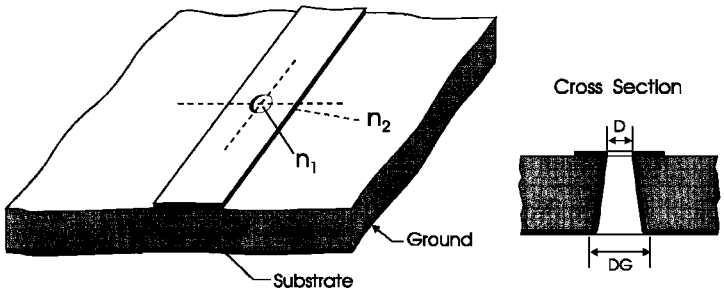


Figure 3-187 Via hole.

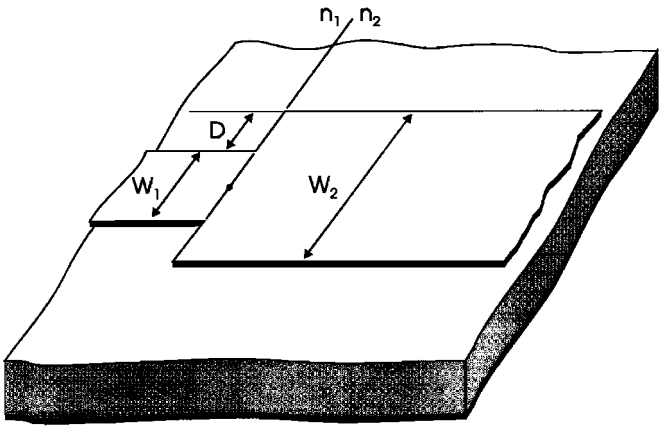


Figure 3-188 The STEP element tells the simulator to calculate the effects of joining transmission lines of differing characteristics.

in mind that, as frequency increases, we rapidly move into the area where we must consider all these distributed elements to achieve accurate simulations—even if doing so makes simulation a painful and time-consuming effort.

Finally, anyone who adventures into this area *must* obtain a foundry manual from the company that will build the integrated circuit or hybrid under design. There are basically two foundries, one applicable for MOS technology and one for GaAs. In the case of the bipolar transistor, the foundry service is not yet well-established. While we were able to interest one German company and one U.S. company in generating a custom bipolar IC within the activities of this book, the only "open" foundry we know of for bipolar (actually HBT) technology is the one operated by TRW. As of 1999, this seems to be the leading company for BJT-related products, and many currently available ICs have been developed with this foundry service.

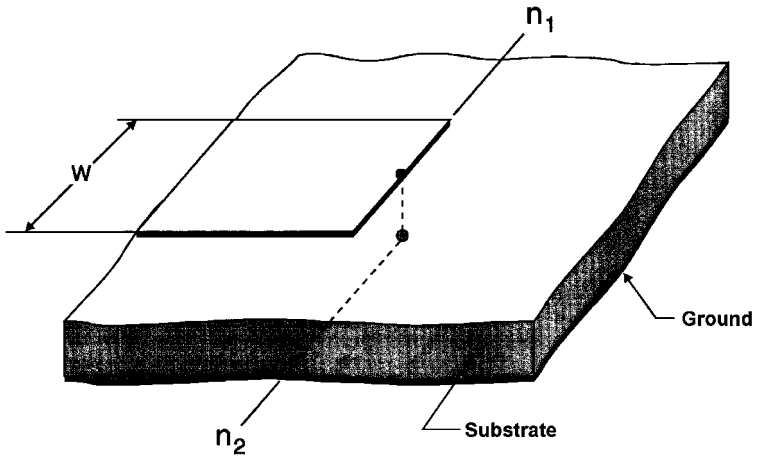


Figure 3-189 The OPEN element tells the simulator to calculate the effects of leaving the end of a transmission line unconnected.

We mention the issue of foundries here again because each foundry has its own proprietary approach to modeling discontinuities. The availability of a foundry service somewhat eases the requirement that a designer be fully up to speed on the nuances of discontinuities, because a foundry's designer service will help customers account for all relevant parasitics or discontinuities in their designs. In addition, there are tables of *S* parameters for standard cells of either capacitors, resistors, or inductors. The designer may then be forced to adjust the circuit so that it will work with a particular inductance value or value of another component within the resolution of the table that describes these elements. Information on the active parts, such as diodes and transistors, was given in Chapter 2.

3-10-1 RF Biasing

Applying the knowledge we have just acquired about distributed elements, Figure 3-190 shows a somewhat exaggerated case of using the various elements. While it is consistent with the simple RF case with which we started, modeling of this type is certainly necessary at higher frequencies. Many but not all of the elements we have just described are used, but we have not solved the question of the actual dc biasing.

3-10-2 dc Biasing

The following discussion is the recommended procedure for biasing a transistor. We will start with the collector current (in our example, Figure 3-191, 10 mA). To separate RF and dc, we have a combination of a load resistor (R_L) and an inductance (L_C). Information about this operation was given already in Section 3-2-6. The emitter resistor should have a voltage

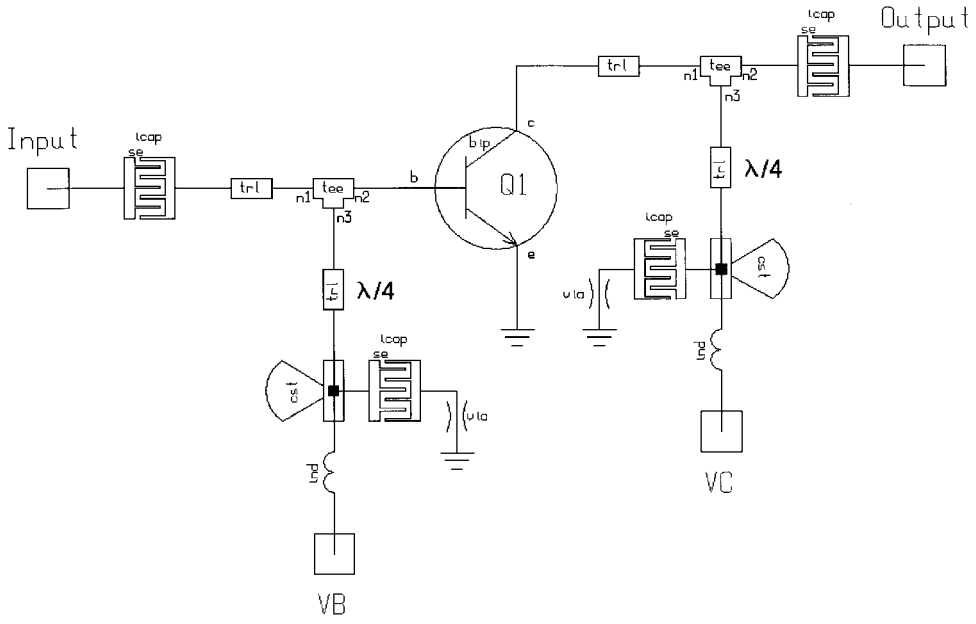


Figure 3-190 Simple BJT RF amplifier with distributed elements used for bypassing and dc blocking, and for base bias and collector supply feeds.

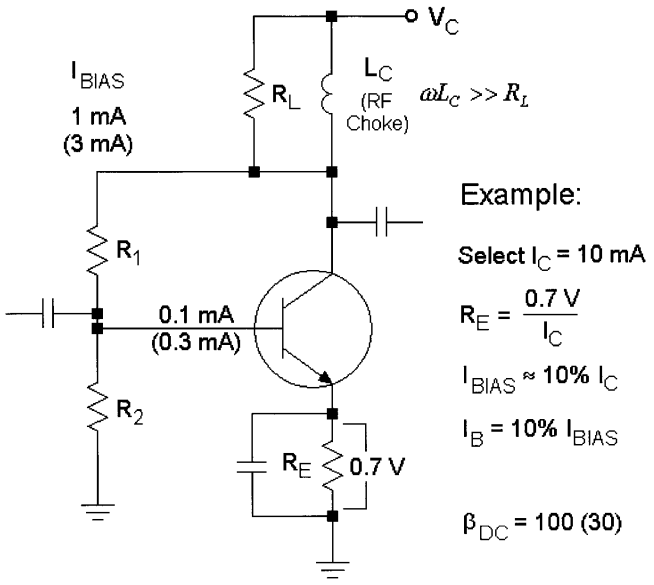


Figure 3-191 A dc biasing example.

drop of about 0.7 V, and therefore the emitter resistor is $0.7/I_C$. The next assumption is that the transistor has a dc current gain of 100, and therefore we decide to have the bias resistor chain (R_1 and R_2) draw 10% of I_C, or 1 mA. Assuming a collector voltage of 5 V, the value of $R_1 + R_2$ has to be 5 kΩ to result in the required 1 mA. Because the base current now is 10% of the divider current, the base will draw 0.1 mA. The voltage across R_2 has to be the voltage drop across R_E (0.7 V) + V_{BE}, which is also 0.7 V, or 1.4 V. Given the 1 mA, R_2 = 1.4/1E–3 = 1.4 kΩ. R_1 therefore is 5 kΩ – 1.4 kΩ = 3.6 kΩ. This concludes the calculation.

In the case of a dc beta of only 30, we need to raise I_{BIAS} from 1 mA to 3 mA since the base current now will be 0.3 mA. This changes the resistor values. $R_1 + R_2$ = 5 V/3 mA = (rounded) 1.6 kΩ. Following the same approach, we need 1.4/3E–3 = (rounded) 470 Ω. This makes R_1 = 1600 Ω – 470 Ω = (rounded) 1100 Ω. These roundings were necessary because resistors can only be bought in certain resistance steps.

However, for low battery voltages, the 0.7 V between the emitter and ground cannot be wasted. Therefore, a dc stabilization circuit, to which we have already referred, will be considered next. Figure 3-192 shows a combination of an RF transistor, Q1, and a dc stabilization transistor, Q2. The most important thing about this circuit is the fact that the difference between the supply voltage and the collector of Q1 is 0.3 V, which is approximately half of the voltage we had to "waste" across R_E in the above-mentioned circuit. Plus, the voltage drop is now in the collector, and given the fact that we feed Q1's collector through the inductor of a quasi-tuned circuit rather than through only a resistor, we get a much higher voltage swing. The base–collector diode temperature dependency of Q2 is compensated for with the silicon diode feeding the base. It is also fascinating to see that, because of the low currents, both the voltage drop of the diode and V_{BE} of Q2 are about the same, and yet Q1, operating at a reasonable current, shows the expected 0.7–0.8 V V_{BE}. We have supplied all the necessary dc voltages to give the reader an incentive to calculate the mechanism of this stabilization. The key equation to this is to determine the current in the voltage divider feeding

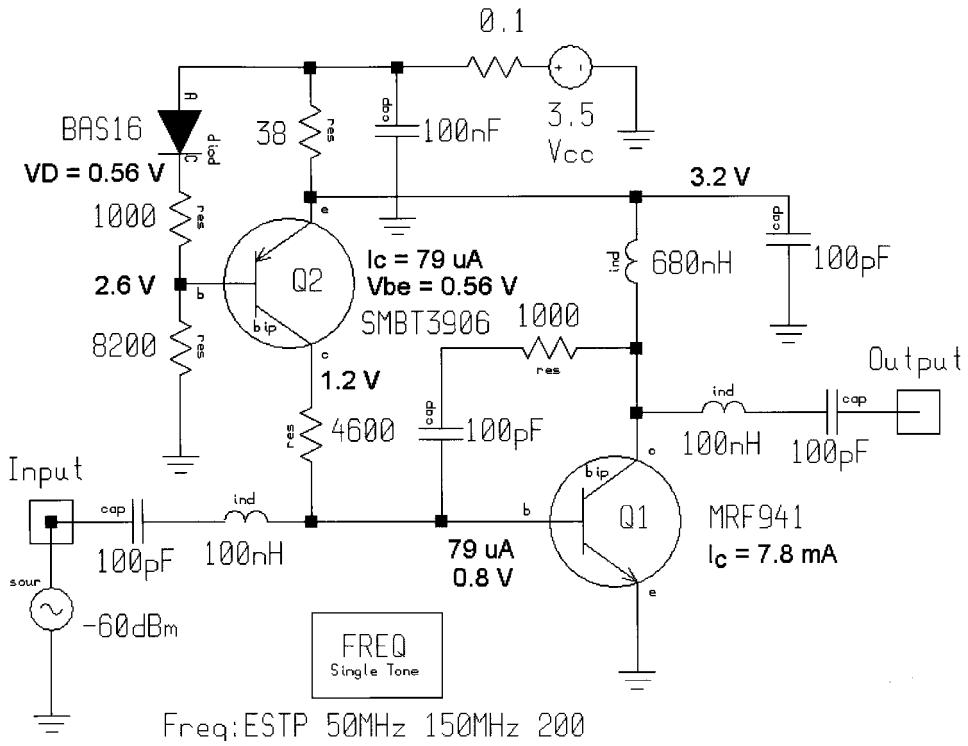


Figure 3-192 An RF amplifier with active biasing. (After the low-noise amplifier circuit in [29].)

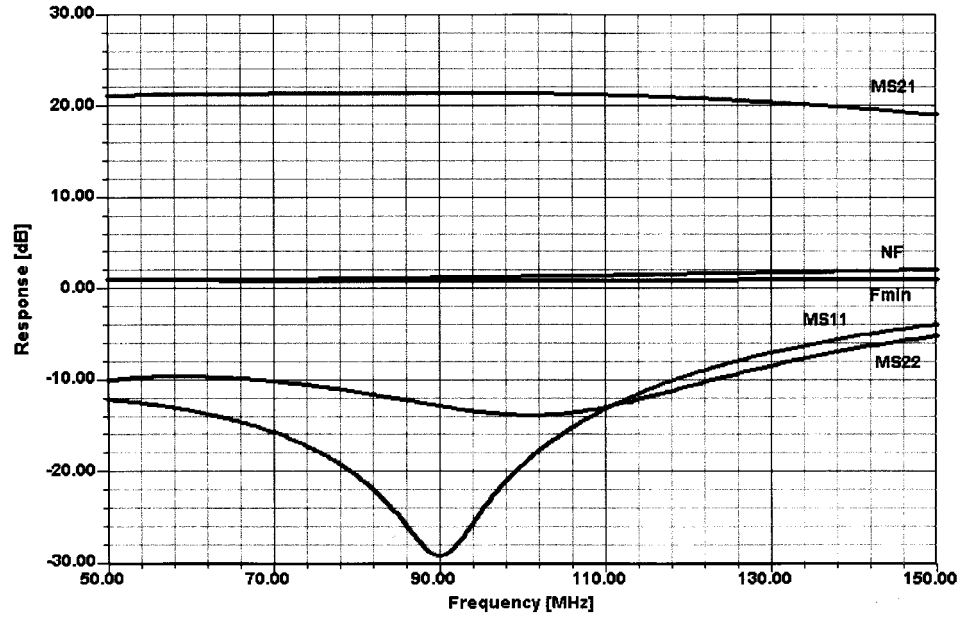


Figure 3-193 Frequency-dependent gain, matching, and noise responses for the amplifier in Figure 3-192.

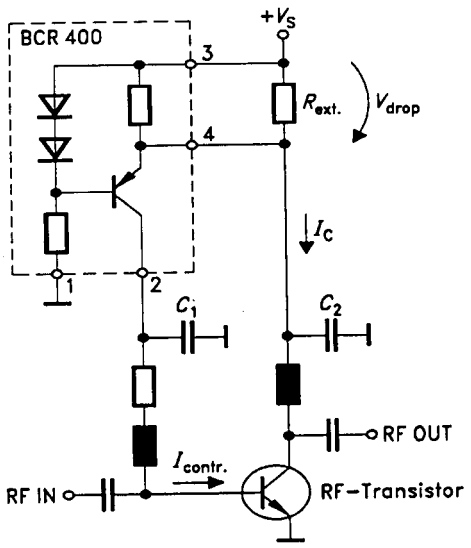


Figure 3-194 The Siemens BCR400 active bias controller can be applied to FETs as well as BJTs.

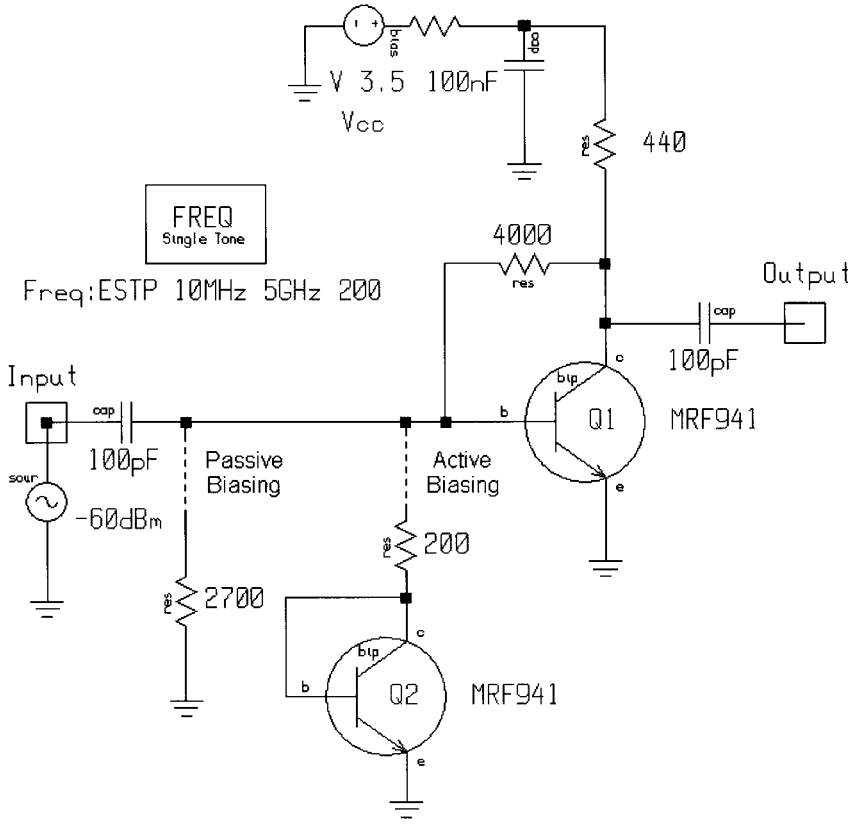


Figure 3-195 Amplifier stage (I_C = 3 mA) showing active and passive bias alternatives.

Q2, and proceeding from there. On the other hand, having a nice CAD tool that provides insight into the dc voltages and allows the addition of dc voltage probes makes life much easier. To prove that this circuit actually works, Figure 3-193 shows its frequency-dependent gain, matching, and noise figure performance. It is based on a Motorola application, which had excessive voltage drop for Q1 and which we have improved to be suitable for low-voltage operation [29].

Integrated active bias solutions are also available. Figure 3-194 shows the Siemens BCR400 active bias controller applied to a BJT; it can also be applied to FETs and TR switching diodes, as we saw in Figures 3-57 and 3-59, respectively.

3-10-3 dc Biasing of IC-Type Amplifiers

In integrated circuits, we do not have the unlimited flexibility described earlier. This results frequently in constant-current sources or transistors being used as diodes for biasing purposes. Our first suspicion is that this is going to cause a lot of noise in the IC; Figure 3-195 shows a test example.

As suspected, choosing the same bias point of 3 mA, the noise figure using active bias is quite a bit higher than the passive one, as shown in Figure 3-196.

3-11 PUSH–PULL/PARALLEL AMPLIFIERS

Since it is practically impossible to obtain $S_{11} = S_{22} = 0$, using a combiner at the input and output reduces the problem. Figure 3-197 shows the 3-dB hybrid approach. It has a practical

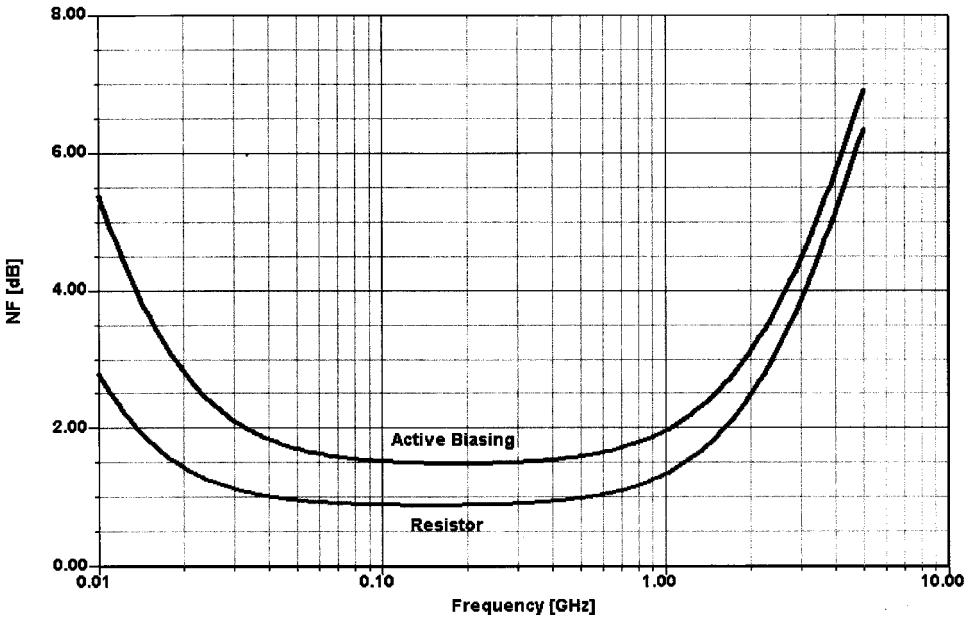


Figure 3-196 Frequency-dependent noise figure of the amplifier circuit shown in Figure 3-195. The active biasing curve is considerably noisier.

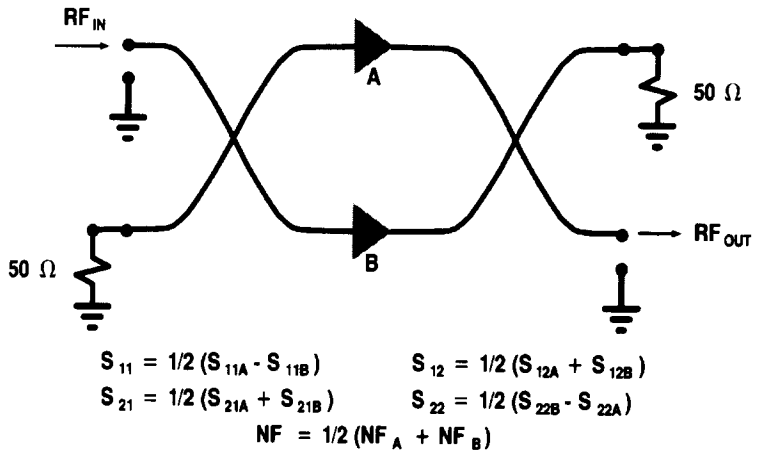


Figure 3-197 Combining power with 3-dB hybrid couplers.

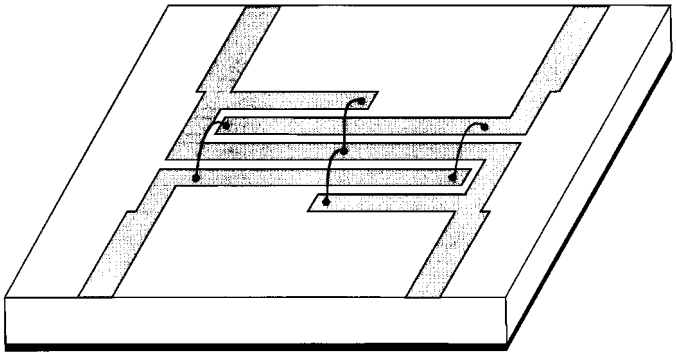


Figure 3-198 Four-strip version of the Lange coupler, a broadband 90° hybrid.

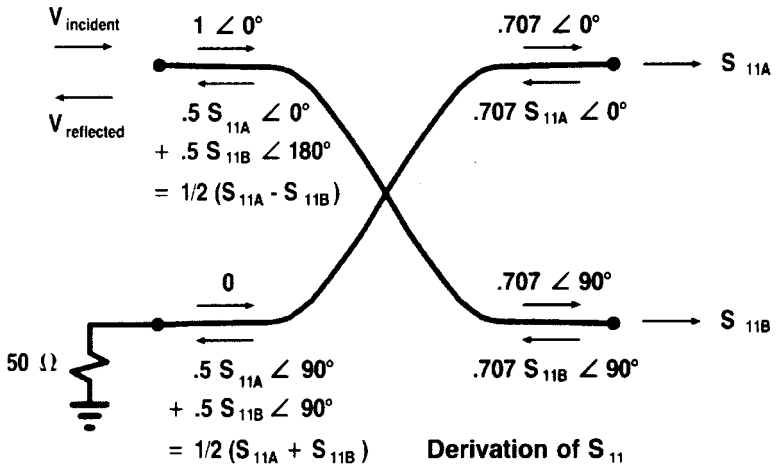


Figure 3-199 3-dB hybrids. This analysis method may be applied to any coupler.

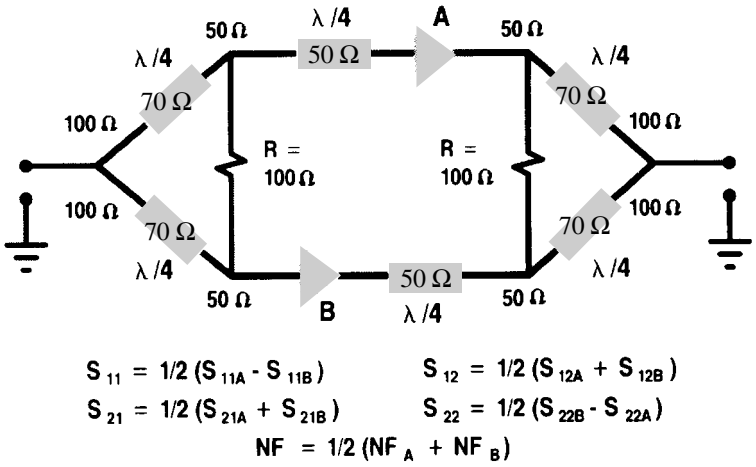


Figure 3-200 Wilkinson divider/combiners.

bandwidth limit of about 4:1; its advantage is that it is compact and ideal for cascading. If one stage opens, P_{OUT} and gain drop only 6 dB; however, it does not protect against load mismatch. A recommended distributed form is the Lange coupler (Figure 3-198). To analyze this approach, Figure 3-199 shows the resulting impedances and necessary phase shifts. If a Wilkinson coupler [30] is used instead of the 3-dB hybrid, the bandwidth is reduced. The Wilkinson coupler is larger in size than the 3-dB hybrid but more easily realized, leading to lower cost. Figure 3-200 shows the actual impedances necessary for matching, with A1 and A2 being the actual amplifiers combined.

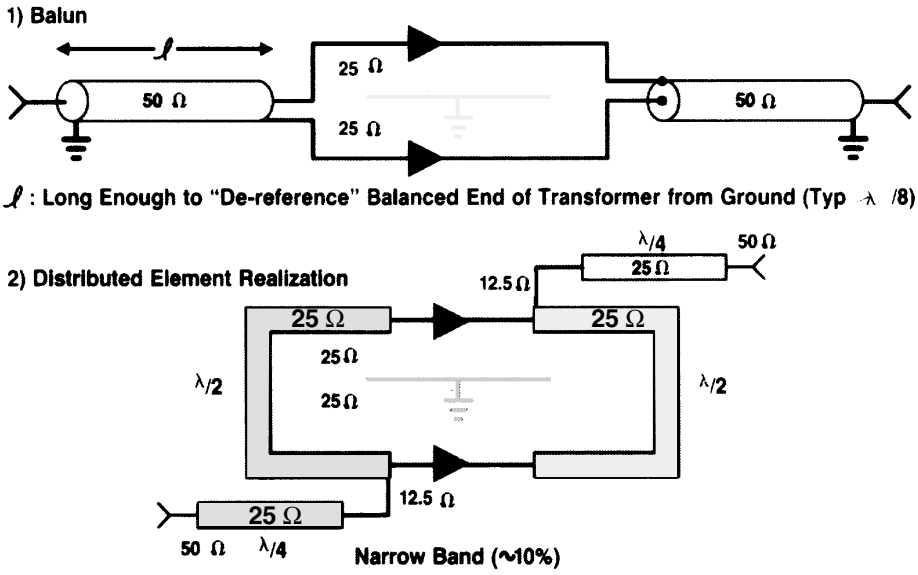


Figure 3-201 Push–pull amplifiers.

The solution for push–pull amplifiers resorts to a balun with the following results obtained: Bandwidths of greater than one decade are achievable; the balun structure gives 4:1 impedance advantage; there is no isolation; $S_{11} \neq 0$, $S_{22} \neq 0$; and cancellation of even-order products is provided.

Since most amplifiers have 50-Ω rather than 25-Ω impedances, Figure 3-201 shows a circuit that matches 25 Ω to 50 Ω on both sides. Since the standard amplifier operates at 50 Ω rather than 25 Ω, an additional network may be necessary to transform the impedance up to 50 Ω again. A similar problem occurs at the amplifier output.

3-12 POWER AMPLIFIERS

Power amplifiers are devices that are tasked to transform as much dc power into RF/microwave power as possible. Depending on the technology, devices are available that can produce at least 100 W up to 2 GHz. Since the transconductance of FETs at conducting angles of less than 180° becomes very small, the actual power gain sharply decreases compared to Class A operation while the power-added efficiency is optimized to values of 55% (Class B) or so. The dissipation of these devices is, of course, quite high, while in Class B operation this is even more of an issue. Power devices in Class C are dying to have big heat sinks to survive. One of the nice things about FETs is a much reduced thermal runaway effect compared to their bipolar brothers.

Power amplifier classes are defined in terms of the device conduction angle and/or the type of device operation:

Class A	"Linear" operation; 180° conduction angle
Class AB	"Linear" operation; conduction angle between 90° and 180°
Class B	"Linear" operation; 90° conduction angle
Class C	Fixed drive; less than 90° conduction angle
Class D	Switched operation, conduction range may vary with time from 0° to 180° or may be fixed.

There are also Classes E and F, which are currently under evaluation and showing some promising results.

Figure 3-202 shows Class C operation. The output of the amplifier is a current pulse train, which frequently leads to misunderstanding as to the calculation of its load. Given the fact that the amplifier is a current generator, it is silly to assume one has to do a conjugate matching looking backward into the amplifier. This is the condition for Class A operation. Unfortunately, this is still not understood by some engineers, but we hope we can clear this up today. The theoretical efficiency is 89.7% at a conducting angle of less than 180°; in practice, efficiencies more than 85% are not obtainable. (All of these efficiencies are those of the device itself; losses in matching networks are not taken into account in these efficiency values.) Considering optimal gain and optimal output power, the conducting angle actually settles somewhere between 100° and 140°. Again, the power gain in Class C operation is less than in Class A and Class B operation.

3-12-1 Example 1: 7-W Class C BJT Amplifier for 1.6 GHz

The following demonstrates how to design a 1.6-GHz, 7-W amplifier operating in Class C. We need to point out again that all modulation forms that contain the information in both

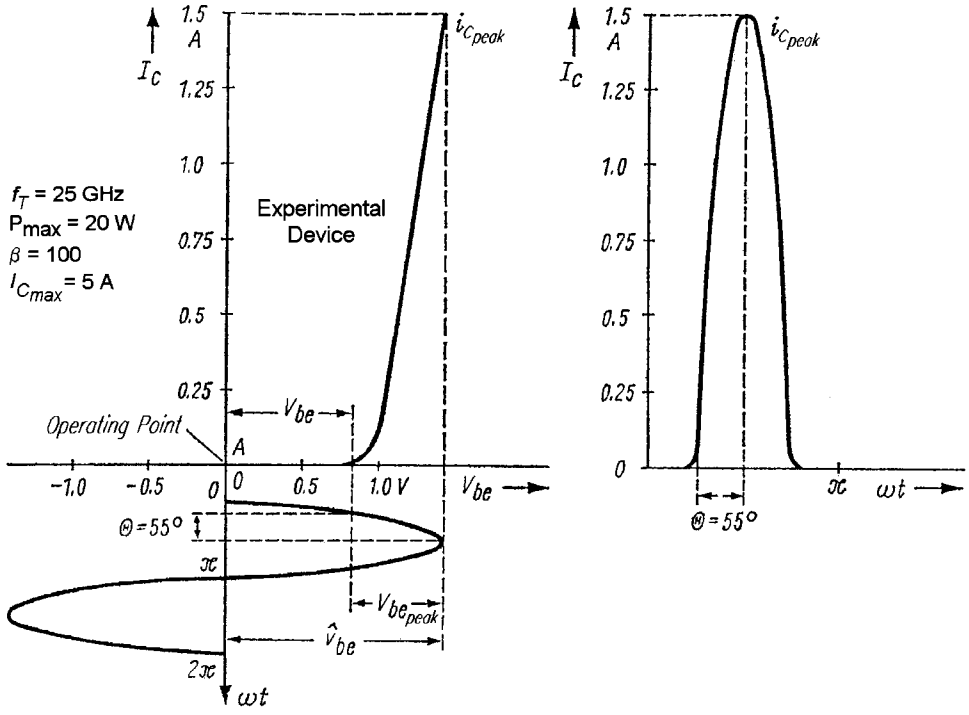


Figure 3-202 Output current as a function of drive voltage, taking into consideration the knee voltage of the base–emitter *pn* junction. The output is a series of current pulses with a duty cycle determined by the conduction angle.

amplitude and phase require more linearity than Class C affords. Needless to say, this affects the efficiency; therefore, at a later point we will also show a Class AB power amplifier. This is a tutorial, so we will present a UHF example and walk the reader through all the problems involved in the design. The selection of the transistor is done by first looking at the f_T, which in our case should be at least 4–5 GHz. The first order of business is to define the point of operation for the output amplifier. To do this, we need to determine the maximum voltage and current values supplied by the manufacturer. We have chosen a device that operates at 28 V and can sustain a peak collector current of 1.5 A. Figure 3-202 lets us determine the conducting angle Θ from the I_C/V_{BE} plane. This figure shows on the left side the I_C/V_{BE} plane and the necessary voltage to drive the amplifier.

Initially, we assume that the base–emitter junction is operated without any bias ($V_{BE} = 0$). The dc I–V curve shows that the collector current starts at about 650 mV. The amplifier itself, therefore, is already in Class C operation. For full output power, the drive voltage $V_{BE,\text{peak}}$ has to be 1.14 V. Therefore, the conducting angle can be determined by

$$\cos \Theta = \frac{V_{BE}}{\hat{\boldsymbol{v}}_{BE}} = \frac{0.65}{1.14} = 0.57$$

$$\cos 0.57 = 55^\circ$$

V_{BE} is the dc voltage at which the transistor starts conducting. From Figure 3-65 and Eqs. (3-160), (3-161), and (3-162), the fundamental current i_{c1} can be determined from

$$i_{c1} = i_{\text{peak}} \cdot f_1(\Theta)$$

or

$$i_{c1} = 1.5 \text{ A} \times 0.36 = 0.54 \text{ A}$$

The average collector dc current can be determined from

$$i_{c2} = i_{\text{peak}} \cdot f_2(\Theta)$$

or

$$i_{c2} = 1.5 \text{ A} \times 0.21 = 0.315 \text{ A}$$

The collector saturation voltage $V_{CE(\text{sat})}$, unfortunately, is around 2–3 V; for the purpose of the calculation, we assume that the saturation voltage is 3 V, and, therefore, we can assume that the maximum collector ac voltage $\hat{v}_{CE} \approx (V_B - 3) = 25$.

The appropriate collector load has to be

$$R_C = \frac{\hat{v}_{CE}}{i_{C1}} = \frac{25}{0.54} = 46.3 \ \Omega$$

The maximum output power is

$$P_{\text{out}} = \frac{(V_{CE} - V_{CE(\text{sat})})^2}{2R_C} = \frac{625}{92.6} = 6.75 \text{ W}$$

The only way to increase the power would be to have a smaller collector load at a higher current. In this family of Motorola transistors, there was no 10-W transistor; that's why we have chosen a 20-W experimental device, which for this purpose is overkill. The theoretical efficiency of the amplifier is defined as

$$\eta = \frac{P_{\text{out}}}{P_{\text{in}}} = \frac{6.75 \text{ W}}{0.315 \times 28} = 85\%$$

If the saturation voltage would be less, our efficiency would automatically increase, since for the given load, the actual output power would increase (6.75 W $\rightarrow$ 7.5 W). Assuming that this stage would operate at 1900 MHz with a gain of roughly 22 dB, the required drive would be significantly less than the output power, or roughly 40 mW. For the purpose of the following calculations, we will assume that the drive power is 0.7 W. The total dissipation of the device now adds up to be

$$P_T = P_{\text{in(dc)}} - P_{\text{out}} + P_{\text{drive}}$$

$$P_T = (28 \times 0.315) - 6.75 + 0.7 = 2.77$$

Power-added efficiency is derived from this. Knowing the power dissipation, one can calculate a thermal resistance R_{therm} of the heat sink and the transistor:

$$R_{\text{therm}} = \frac{T_J - T_A}{P_T}$$

Where T_J is the device junction temperature and T_A is the ambient temperature. For a maximum ambient temperature of 60 °C and a maximum junction temperature of 150 °C, the thermal resistance is

$$R_{\text{therm}} = \frac{120 - 60}{2.77} = 21.7\ ^\circ\text{C/W}$$

R_{therm} is the sum of thermal resistance of the transistor and its package, and the thermal resistance between the transistor package and the heat sink, and the thermal resistance between the heat sink and the ambient air. Therefore,

$$R_{\text{therm}} = R_{\text{th(case)}} + R_{\text{th(case/heat sink)}} + R_{\text{th(heat sink)}}$$

The required thermal resistance of the heat sink is calculated from

$$R_{\text{th(heat sink)}} = 21.7 - 10 - 1.4 = 15.6\ ^\circ\text{C/W}$$

Assuming an aluminum heat sink (a 2-mm-thick plate), the required heat-sink surface area A (cm^2) is

$$A = \frac{1}{\alpha R_{\text{th(heat sink)}}}; \quad \alpha \approx 1.5\ \frac{\text{mW}}{^\circ\text{Ccm}^2}$$

Therefore, the required area is

$$A = \frac{1}{1.5 \times 0.0156} = 42.7\ \text{cm}^2$$

The next task is to match both the input and output to the (most likely) 50-Ω source and termination. The following section will guide us through the procedure. The input impedance Z_{11} (after some lengthy search) was found to be $1 + j10\ \Omega$. This means that the input consists of the base-spreading resistance (approximately 1 Ω) and an input capacitance, which because of phase shift now turns out to be inductive. Since the popular method for input matching these days is a pi configuration, we are going to use a simple pi filter both for input and output matching. There are several equations in the literature that give the values for the two capacitors and inductors; most of them result in an incorrect resonance frequency. Since we have access to software tools, we decided to take three approximate values and use the optimizer to give us the appropriate matching. The only remaining variable in such an approach is the bandwidth. If we specify $S_{21} = 1$ and $S_{11} = 0$ (100% energy transfer and perfect return loss), we will realize that this is possible at only one frequency. We therefore need to identify a frequency band or bandwidth over which an acceptable match is needed. As the bandwidth increases, the match for such a simple circuit will get poorer; there are not enough elements available to do better. In this case, we took the center frequency ±100 MHz.

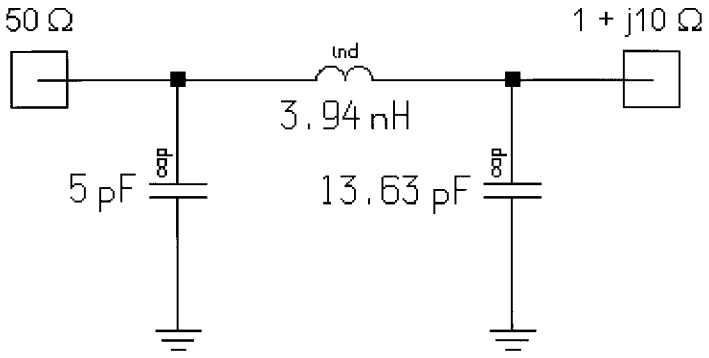


Figure 3-203 Input matching network for the amplifier used in this tutorial.

Figures 3-203 and 3-204 show the actual circuit with the element values and the resulting input match. Likewise, we did the same thing at the output, but the actual resulting dynamic output capacitance is not known *a priori*. For low frequencies, the integral equation gives a value of $2C_{CE}$; however, practical tests in frequencies above 500 MHz show that the actual increase is only 1.3 times the dc output capacitance. Therefore, we decided to design the filter with no output capacitance assumed, and the input capacitance of about 5 pF must be reduced for proper reactive matching. Mathematically, this could result in a negative number, specifically if the resulting output capacitance exceeds the 5-pF value. To compensate for this, one needs to use an RF choke that tunes out the resulting dynamic output capacitance and then one can revert to the original 5 pF as calculated.

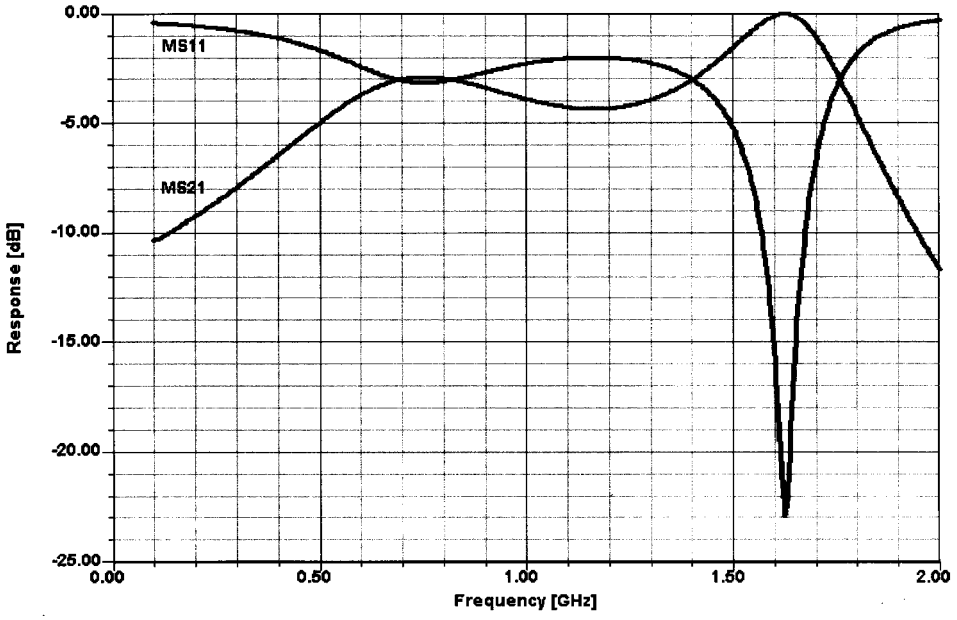


Figure 3-204 Frequency response of the input matching network.

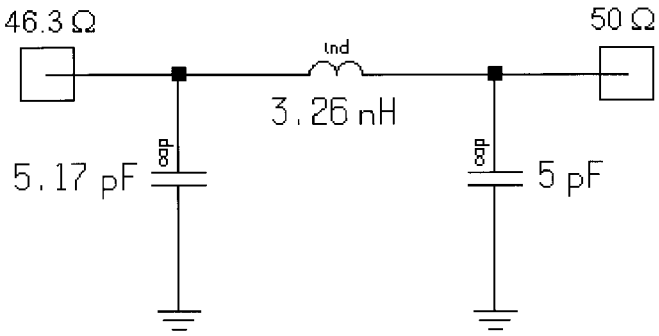


Figure 3-205 Output matching network for the amplifier.

The selection of the actual component is another issue, because these capacitors need to handle several amperes of RF current. It is not always easy to find capacitors with the appropriate value that can handle the current. The inductors can be printed. The paragraph following this section therefore will deal with the issue of how to translate the lumped values into distributed values, which will make the current handling much easier (we hope).

For this, we did the same optimizing approach by using a 50-Ω source and a 46.3-Ω load. Figures 3-205 and 3-206 show the circuit with the appropriate values and its simulated frequency-dependent responses.

The next step is to connect the input and output matching network to an "actual" transistor. This is a point where difficulties arise, since very few, if any, companies provide SPICE-type

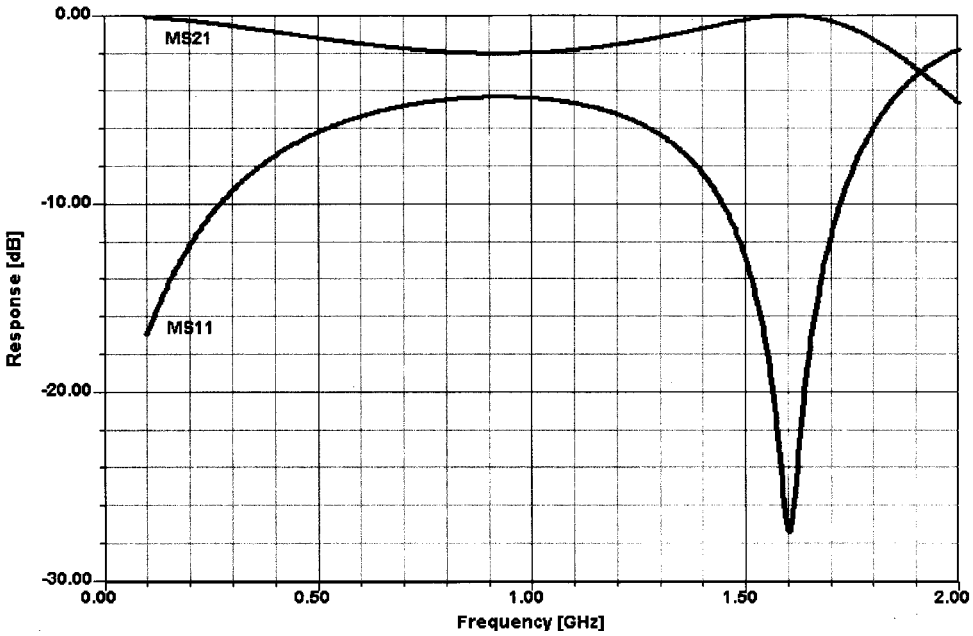


Figure 3-206 Frequency response of the output matching network.

parameters for transistors operating at currents of more than 200 mA. Figures 3-207 and 3-208 show the overall schematic and frequency response of the actual amplifier with all filters connected. Figure 3-209 details the amplifier's peak gain response. If more bandwidth is needed, the matching circuit must be increased in complexity, as will be shown in the following section.

Now is a good time to look at the final results of our simulation, which is based on the various assumptions as outlined above. Figure 3-210 shows the dc *I–V* curves with a straight load line and no reactances. This was accomplished by using a wideband transformer instead of an output matching circuit. Figure 3-211, however, shows the "load line" that results when the transformer is replaced with a real matching network. Based on the harmonics content and its improper termination, one can see how the output load line opens and shows ringing in the low-current area. If we expect to see the saturation voltage, we will be disappointed because the saturation voltage is frequency dependent (increases with frequency) and this figure gives an erroneous impression. Hence, we show dc *I–V* curves. To display the conduction angle, which we showed previously, we examine the output current as a function of time at a drive level of +16.5 dBm (Figure 3-212). During the negative half, the transistor is not conducting as predicted, so this validates our assumptions. Next, we are interested in determining the power gain. Therefore, we look at the output spectrum for a drive level of +16.5 dBm (Figure 3-213). It shows an output level of 38.6 dBm. If we subtract the drive level (16.5 dBm) from this, we see that the power gain is 22.1 dB at this drive level. Since we are used to thinking in watts, Figure 3-214 shows the absolute output power and the harmonics prior to filtering. Figure 3-215 shows the output spectrum after replacing the transformer with a simple pi output network.

During the conduction period, there will be a voltage across the transistor 180° out of phase, as shown in Figure 3-216. Finally, to highlight the nonlinearity of a Class C amplifier, Figure 3-217 shows the amplifier's output power versus drive, and Figure 3-218 shows the amplifier's gain versus drive. When the transistor turns on is clearly evident.

A particular difficulty arises in the case of FETs, where the charge energy makes for a difference between dc and pulsed measurements. Figure 3-219 shows this. Some of the more modern transistors, such as SiGe types, may need some pulsing also; as can be seen in Figure 3-219, even depending on the pulsing, one may get different dc *I–V* curves. This area is still under investigation, and even the "experts" do not agree on all aspects of this. In real life, it turns out that a lot of experimentation is the answer. Anyone who thinks that a pure CAD approach will give the right design answers immediately will encounter many surprises! This is not a fault of the CAD tools—assuming that they are properly engineered—but rather the availability of appropriately "tweaked" models for high-power applications, and the willingness of device suppliers to standardize on certain models and support them.

Because of the difficulties involved in cooling transistors under test and simulating RF pulse and CW conditions, depending on the application, manufacturers have so far stayed away from investing money in this area. The designer must therefore resort to application notes published by device manufacturers. Because even slight layer changes can have a large impact on the frequency response of a given device, application-note-based designs can run into difficulty if a device's fabrication has significantly evolved since the publication of its application notes.

Given the multitude of models outlined in Chapter 2, the CAD user is sometimes left hanging as to which is the best model to use. The latest indications are that the MEXTRAM (*m*ost *e*xquisite *t*ransistor *m*odel), developed by the University of Delft, Holland, with Philips of Holland, shows the most promising result, specifically in the area of IMD products. While

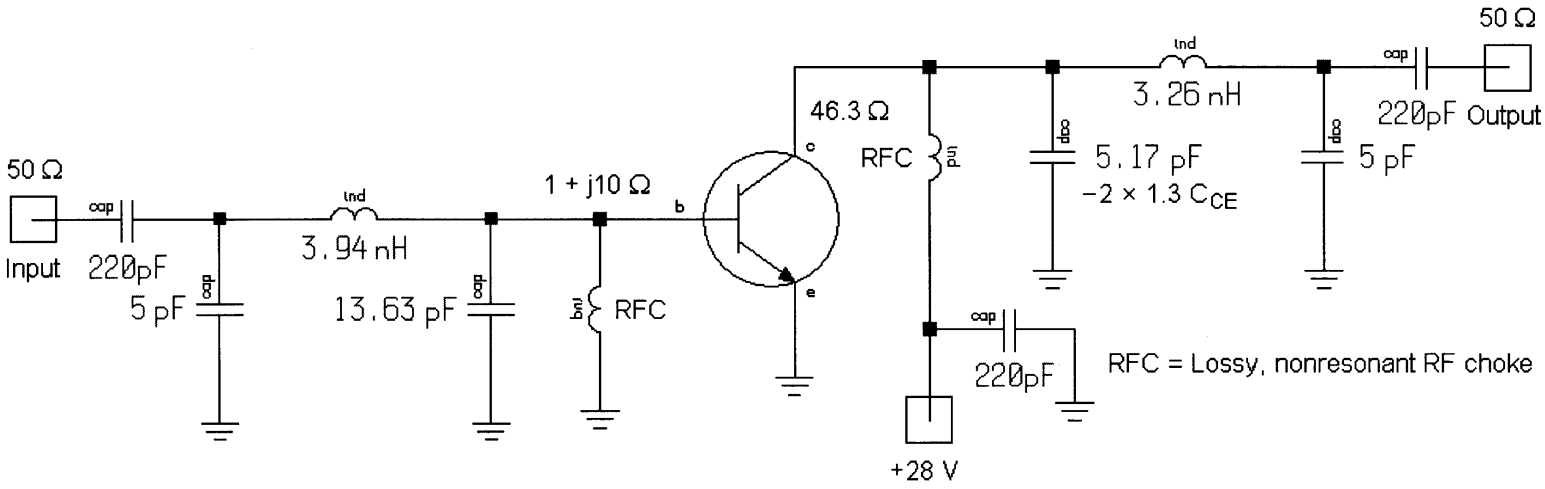


Figure 3-207 Actual schematic of our experimental 1.6-GHz amplifier. The RF chokes have enough RF losses not to be resonant and yet don't load the circuit. This is frequently accomplished with either ferrite beads or resistors in parallel with the chokes.

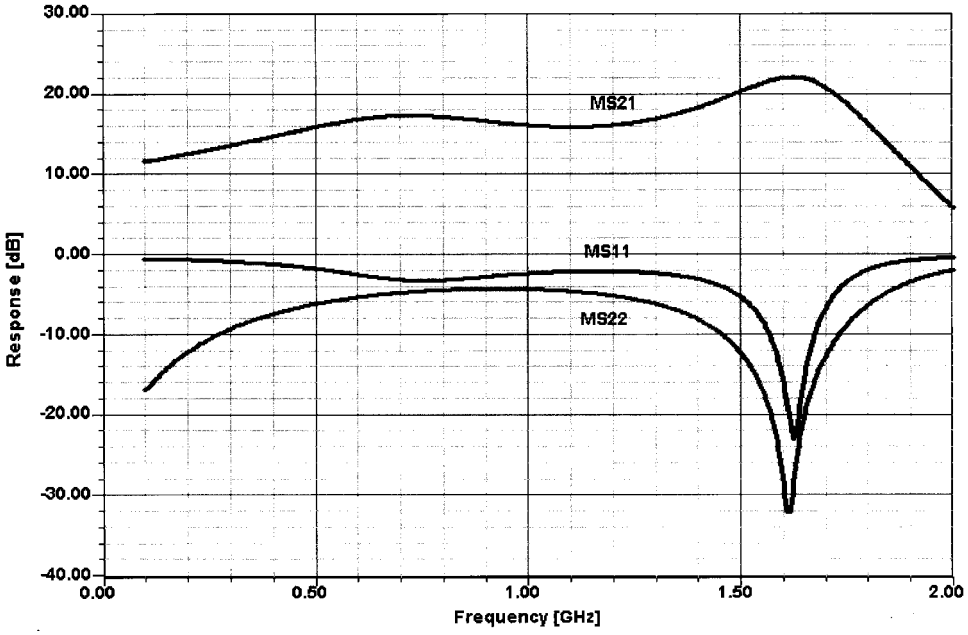


Figure 3-208 Overall frequency response of our experimental amplifier. Input and output matching, and gain, are shown as a function of frequency.

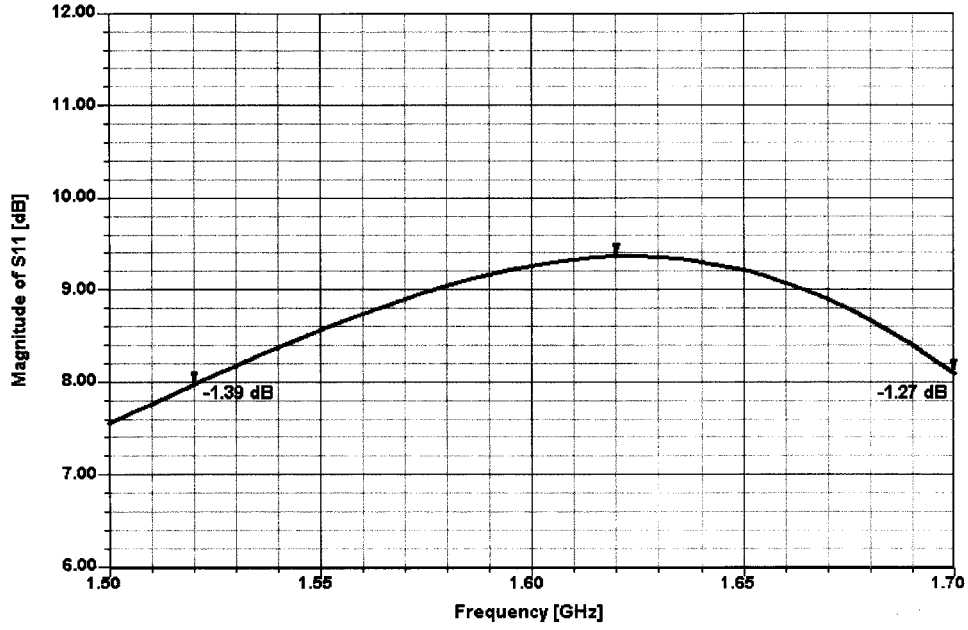


Figure 3-209 Close-in display of the frequency-dependent gain.

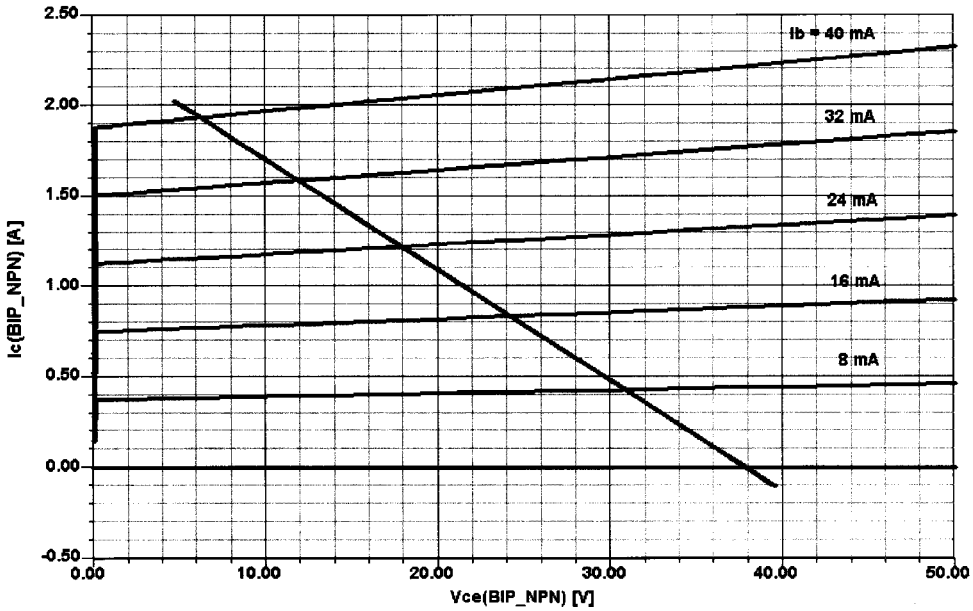


Figure 3-210 The dc *I–V* curves assuming a perfect resistive load. No RF saturation voltage or other RF saturation effects are evident.

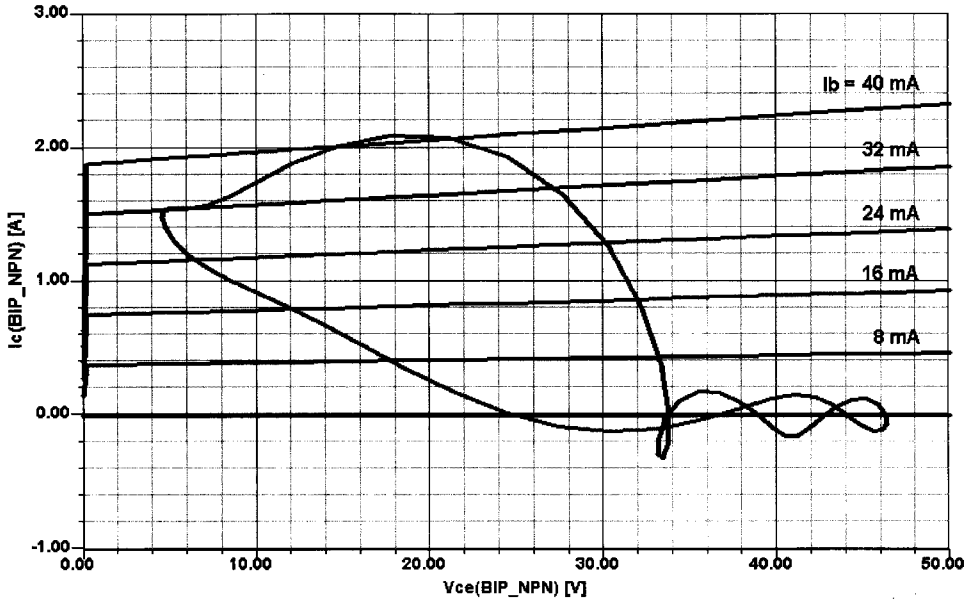


Figure 3-211 Load line with the output transformer replaced with a real matching circuit.

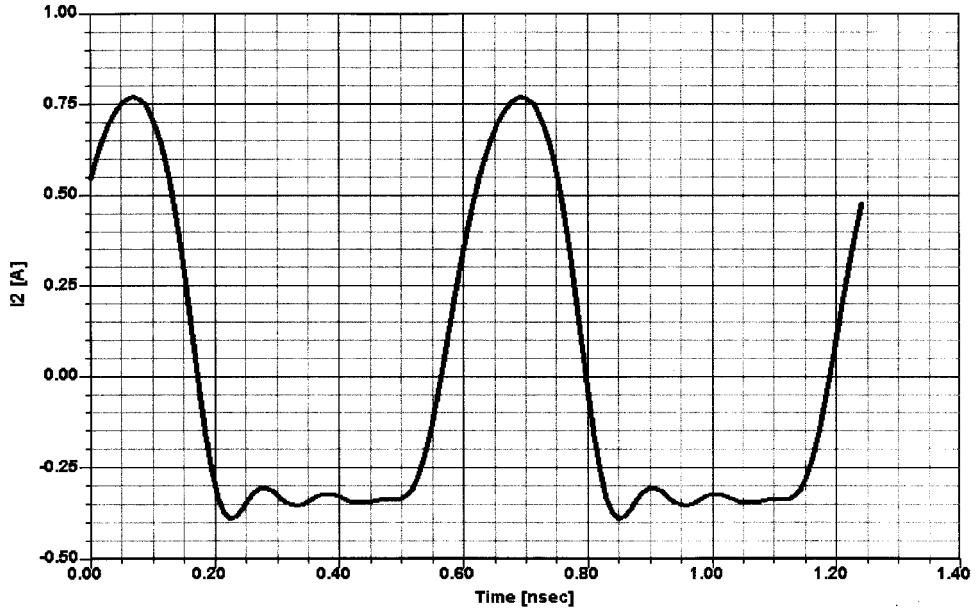


Figure 3-212 Output current as a function of time. The duty cycle is determined by the device conduction angle.

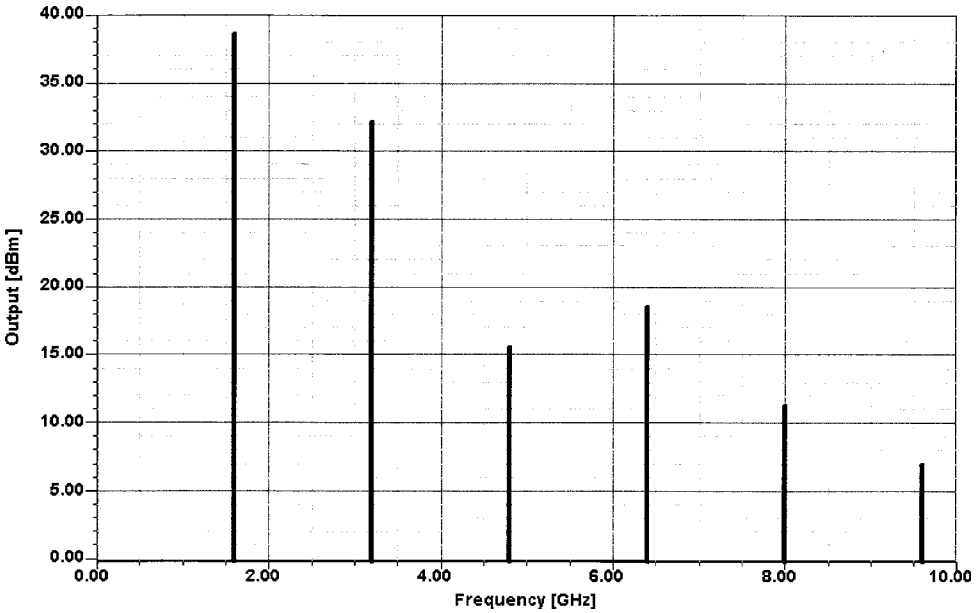


Figure 3-213 Power output at the collector prior to any filtering, assuming a real output termination.

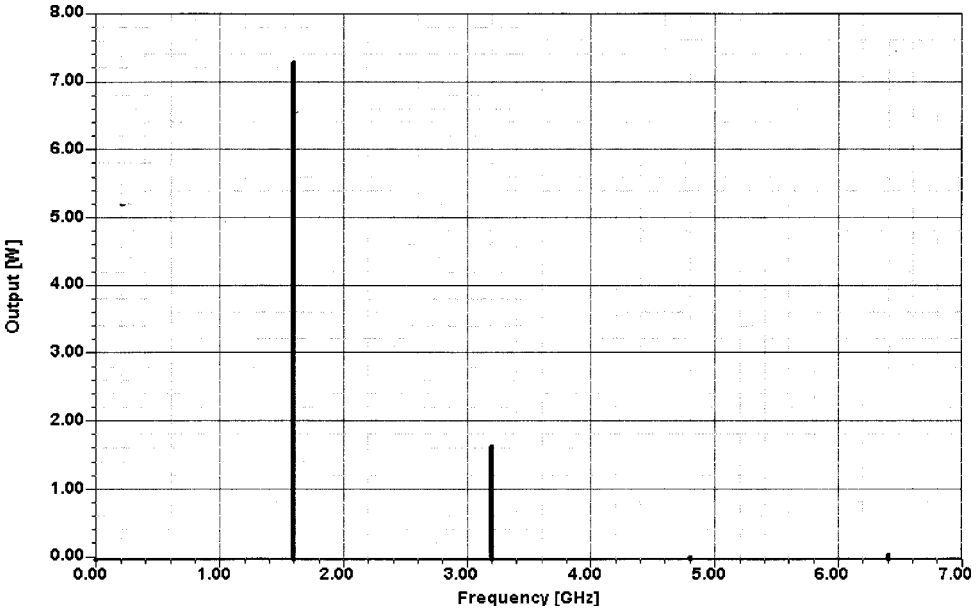


Figure 3-214 Available output power in watts, prior to any filtering, assuming a real output termination.

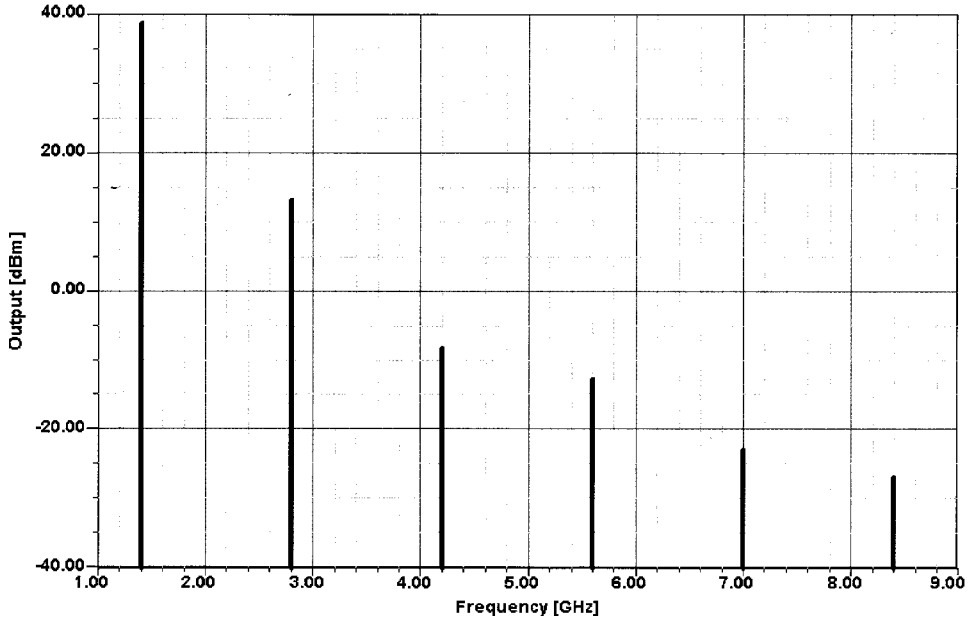


Figure 3-215 The output spectrum after replacing the transformer with a simple pi output network.

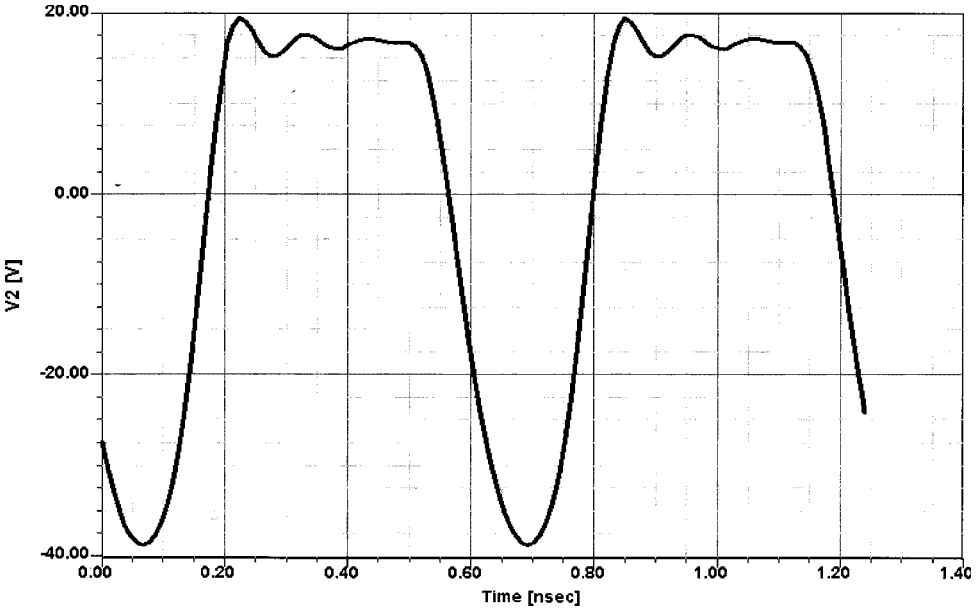


Figure 3-216 Output voltage at the collector. The voltage is asymmetrical compared to 0; this is important in the selection of a device capable of handling these voltages.

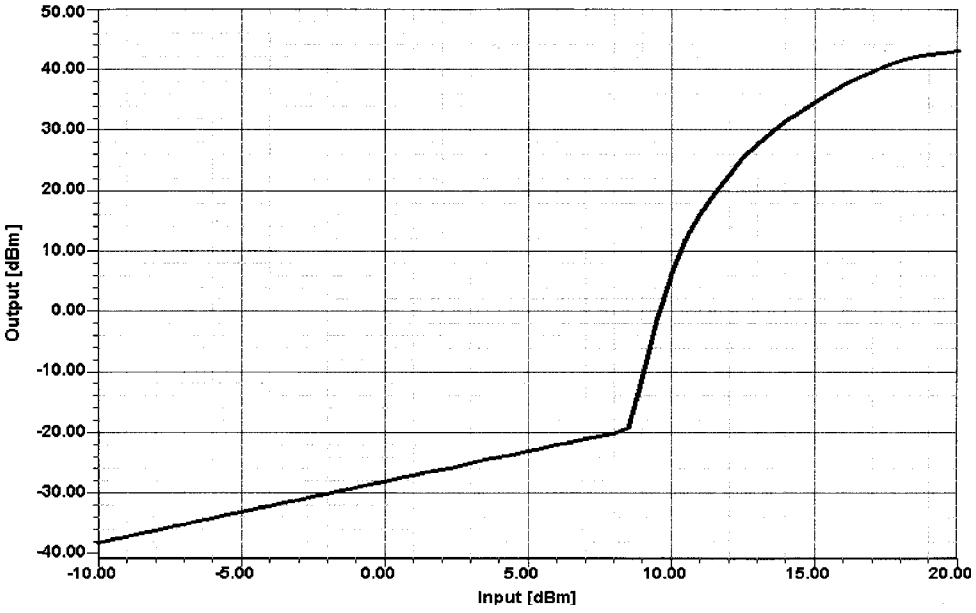


Figure 3-217 Output power versus input power for the Class C amplifier. Until the transistor turns on (at a drive level of about 8.5 dBm), only fedthrough power appears at the output.

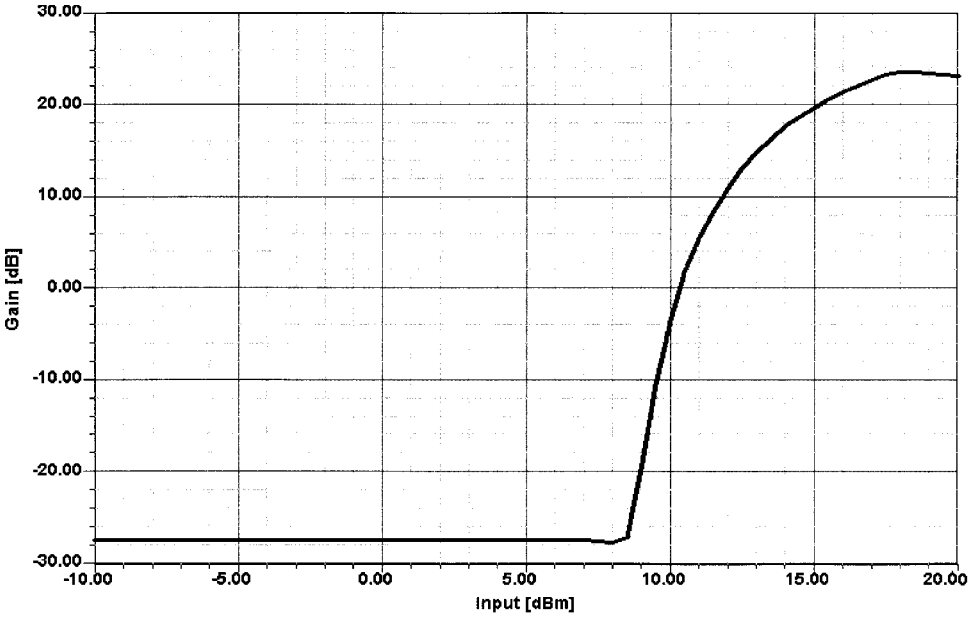


Figure 3-218 Amplifier gain as a function of drive. It is evident that one needs to cross the 0.65-V point for the transistor to become conductive and show gain. The initial amplification as a function of drive shows a very steep curve that then levels off and drops as the amplifier moves into saturation.

other later models have up to seven internal nodes, MEXTRAM fortunately uses only five. (The MEXTRAM model actually consists of three transistors, which makes modeling fun.) Power-added efficiency simulation is not available, since the MEXTRAM model was developed for small-signal applications, meaning circuits that operate at less than 100 mW of RF output power.

As if the absence of trustworthy high-power device models is not challenge enough, in developing this 1.6-GHz Class C amplifier example, we ran into another complication: Using a CAD optimizer to simultaneously match the circuit's input and output networks failed because of the device's input-impedance variation with drive. Every matching adjustment that improved the drive to the device resulted in an input-impedance change that reduced the drive to the device! In the end, the optimizer couldn't keep up. In actually building such a circuit, one would make the matching-network values adjustable on the initial breadboard and then, with the necessary input drive level applied, iteratively adjust the input and output matching for the desired output level. It appears that the CAD program could not handle such a degree of nonlinearity.

We then went to the approach of using large-signal *S* parameters at a drive level slightly less than 16 dBm, but this detuned the input filter, resulting in a poorer input match than the one initially predicted. Likewise, the output matching, which in a power stage must result in maximum output power rather than a conjugate match, was less impressive. The large-signal *S* parameters at the output for a power amplifier are somewhat dubious because Class C operation generates so many harmonics that the output termination will reflect a lot of energy and, depending on the output filter topology, different results can be expected. We firmly believe that there will be further mathematical improvement in the CAD tools, and we very much invite

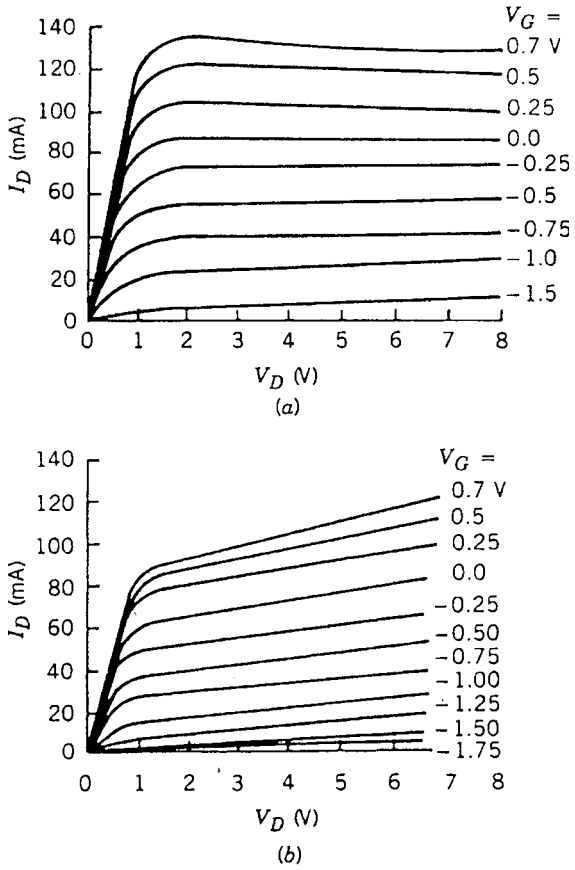


Figure 3-219 Comparison of FET drain current characteristics obtained from (a) pulsed *I–V* and (b) rectified sine-wave evaluation. The rectified sine-wave curves can be considered equivalent to curves obtained at dc.

our readers to use such a nonlinear example to test their tools, provided that a high-power model becomes available in this power range.

It has been rumored that the load-pull technique, rather than the application of large-signal *S* parameters, is a vehicle to solve such matching problems. However, the load-pull technique is really applied only to the fundamental frequency and does not deal with the issue of harmonics of the output current. More details about the load-pull technique and its capabilities are found in Appendix B.

Actual measurements of amplifiers of this type show that our simulated efficiency and drive level come quite close to reality, but at a frequency somewhat offset from that chosen for the design. As an example, the input and output filters tuned at 1.4 GHz instead of 1.6 GHz using the large-signal *S* parameters. Again, an iterative process, similar to manually trimming a prototype, will give a similar result. This is consistent with our earlier statement that most final designs are really hand-tweaked, and that CAD can only bring us close, but not all the way, to a good solution. For passive problems, the CAD does a significantly better job; the nonlinear cases are the ones that still cause headaches.

3-12-2 Impedance Matching Networks Applied to RF Power Transistors*

Introduction. Some graphical and numerical methods of impedance matching will be reviewed here. The examples given refer to high-frequency power amplifiers.

Although matching networks normally take the form of filters and therefore are also useful to provide frequency discrimination, this aspect will be considered only as a corollary of the matching circuit.

Matching is necessary for the best possible energy transfer from stage to stage. In RF power transistors, the input impedance is of low value, decreasing as the power increases or as the chip size becomes larger. This impedance must be matched either to a generator—of generally 50-Ω internal impedance—or to a preceding stage. Impedance matching has to be made between two complex impedances, which makes the design still more difficult, especially if matching must be accomplished over a wide frequency band.

Device Parameters

Input Impedance. The general shape of the input impedance of RF power transistors is as shown in Figure 3-220. It is a large-signal parameter, expressed here by the parallel combination of a resistance R_p and a reactance X_p [31].

The equivalent circuit shown in Figure 3-221 accounts for the behavior illustrated in Figure 3-220. With the presently used stripline or flange packaging, most of the power devices for VHF low band will have their R_p and X_p values below the series resonant point f_s. The input impedance will be essentially capacitive.

Most of the VHF high-band transistors will have the series-resonant frequency within their operating range; that is, they will be purely resistive at one single frequency f_s, which the parallel-resonant frequency f_p will be outside.

Parameters for 1- or 2-GHz transistors will be beyond f_s and approach f_p. They show a high value of R_p and X_p with inductive character.

A parameter that is very often used to judge on the broadband capabilities of a device is the input Q or Q_{IN}, defined simply as the ratio R_p/X_p. Practically, Q_{IN} ranges around 1 or less for VHF devices and around 5 or more for microwave transistors.

Q_{IN} is an important parameter to consider for broadband matching. Matching networks normally are low-pass or pseudo-low-pass filters. If Q_{IN} is high, it can be necessary to use bandpass filter-type matching networks and to allow insertion losses. But broadband matching is still possible. This will be discussed later.

Output Impedance. The output impedance of RF power transistors, as given by all manufacturers' datasheets, generally consists of only a capacitance C_{OUT}. The internal resistance of the transistor is supposed to be much higher than the load and is normally neglected. In the case of a relatively low internal resistance, the efficiency of the device would decrease by the factor

$$1 + R_L/R_T \tag{3-192}$$

*Based on Motorola application note AN-721 ("Impedance Matching Networks Applied to RF Power Transistors"). Copyright of Motorola; used by permission.

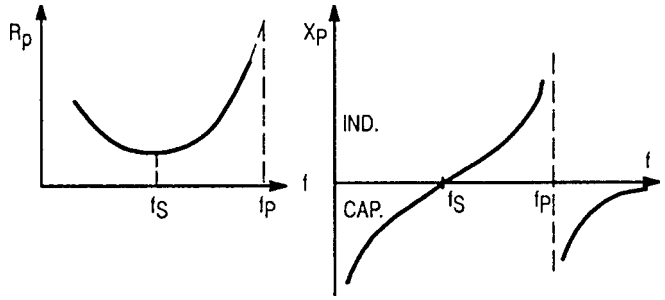


Figure 3-220 Input impedance of RF power transistors as a function of frequency.

where R_L is the load resistance, seen at the collector–emitter terminals, and R_T is the internal transistor resistance equal to

$$\frac{1}{\omega_T (C_{TC} + C_{DC})} \tag{3-193}$$

defined as a single parameter, where ω_T is the transit angular frequency and C_{TC} and C_{DC} are the transition and diffusion capacitances, respectively, at the collector junction.

The output capacitance, C_{OUT}, which is a large-signal parameter, is related to the small-signal parameter C_{CB}, the collector–base transition capacitance.

Since a junction capacitance varies with the applied voltage, C_{OUT} differs from C_{CB} in that it has to be averaged over the total voltage swing. For an abrupt junction and assuming certain simplifications, $C_{\text{OUT}} = 2C_{CB}$.

Figure 3-222 shows the variation of C_{OUT} with frequency. C_{OUT} decreases partly due to the presence of the collector lead inductance, but mainly because of the fact that the base–emitter diode does not shut off anymore when the operating frequency approaches the transit frequency, f_T.

Output Load. In the absence of a more precise indication, the output load R_L is taken as

$$R_L = \frac{[V_{CC} - V_{CE(\text{sat})}]^2}{2P_{\text{OUT}}} \tag{3-194}$$

with $V_{CE(\text{sat})}$ equal to 2 or 3 V, increasing with frequency.

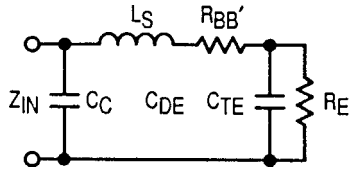


Figure 3-221 Equivalent circuit for the input impedance. R_E is the emitter diffusion resistance; C_{DE} and C_{TE} are the diffusion and transition capacitances, respectively, of the emitter junction; $R_{BB'}$ is the base-spreading resistance; C_C is the package capacitance; and L_S is the base lead inductance.

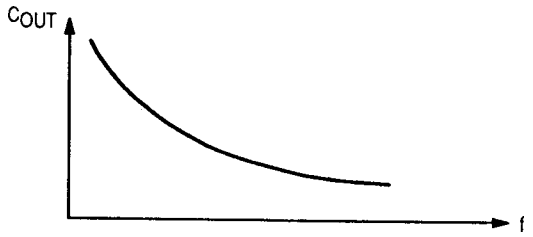


Figure 3-222 Output capacitance C_{OUT} as a function of frequency.

The above equation just expresses a well-known relation, but it also shows that the load, in first approximation, is not related to the device, except for $V_{CE(sat)}$. The load value is primarily dictated by the required output power and the peak voltage; it is not matched to the output impedance of the device.

At high frequencies this approximation becomes less exact, and for microwave devices the load that must be presented to the device is indicated on the datasheet.

Strictly speaking, impedance matching is accomplished only at the input. Interstage and load matching are more impedance transformations of the device input impedance and of the load into a value R_L (sometimes with additional reactive component) that depends essentially on the power demanded and the supply voltage.

Matching Networks. In the following, matching networks will be described by the order of complexity. These are ladder-type reactance networks. The different reactance values will be calculated and determined graphically. Increasing the number of reactances broadens the bandwidth. However, matching networks consisting of more than four reactances are rare. Above four reactances, the improvement is small.

Numerical Design

TWO-RESISTANCE NETWORKS. Resistance terminations will first be considered. Figure 3-223 shows the reactive L section and the terminations to be matched.

Matching or exact transformation from R_2 into R_1 occurs at a single frequency f_0. At f_0, X_1 and X_2 are equal to

$$X_1 = \pm R_1 \sqrt{\frac{R_2}{R_1 - R_2}} = R_1 \frac{1}{\sqrt{n-1}} \tag{3-195}$$

$$X_2 = \mp \sqrt{R_2(R_1 - R_{2)}} = R_1 \frac{\sqrt{n-1}}{1} \tag{3-196}$$

At f_0, $X_1X_2 = R_1R_2$. X_1 and X_2 must be of opposite sign. The shunt reactance is in parallel with the larger resistance.

The frequency response of the L section is shown in Figure 3-224, where the normalized current is plotted as a function of the normalized frequency.

If X_1 is capacitive and consequently X_2 is inductive, then

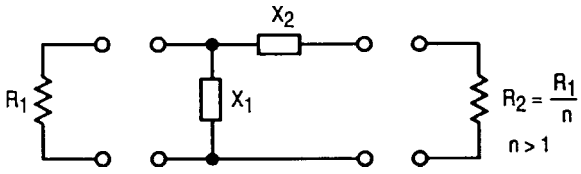


Figure 3-223 Two-reactance matching network.

$$X_1 = -\frac{f_0}{f} R_1 \sqrt{\frac{R_2}{R_1 - R_2}} = -\frac{f_0}{f} R_1 \frac{1}{\sqrt{n-1}} \tag{3-197}$$

and

$$X_2 = \frac{f}{f_0} \sqrt{R_2(R_1 - R_2)} = \frac{f}{f_0} R_1 \frac{\sqrt{n-1}}{1} \tag{3-198}$$

The normalized current absolute value is equal to

$$\left|\frac{I_2}{I_0}\right| = \frac{2\sqrt{n}}{\sqrt{(n-1)^2 \cdot (f/f_0)^4 - 2\,(f/f_0)^2 + (n+1)^2}} \tag{3-199}$$

where $I_0 = \sqrt{n}E/2R_1$ and is plotted in Figure 3-224 [32].

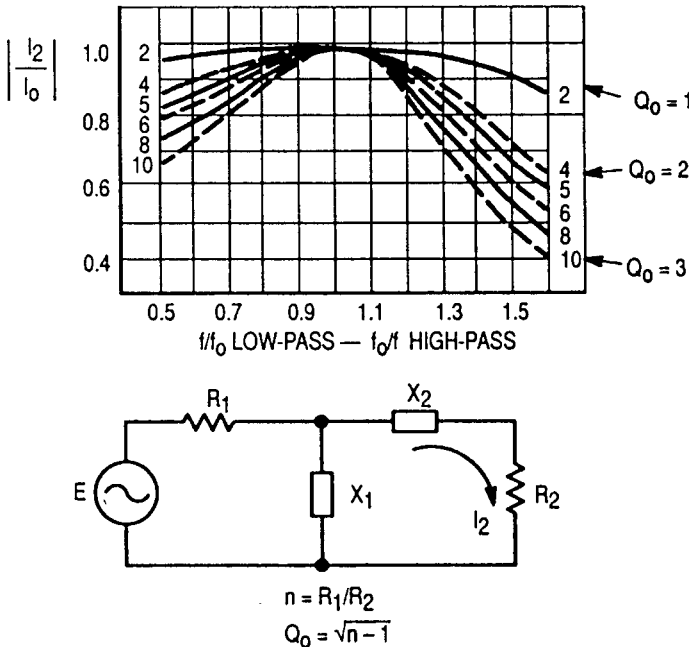


Figure 3-224 Normalized frequency response for the L section in low-pass or high-pass form.

If X_1 is inductive and consequently X_2 is capacitive, the only change required is a replacement of f by f_0 and vice versa. The L section has low-pass form in the first case and high-pass form in the second case.

The Q of the circuit at f_0 is equal to

$$Q_0 = \frac{X_2}{R_2} = \frac{R_1}{X_1} = \sqrt{n-1} \tag{3-200}$$

For a given transformation ratio n, there is only one possible value of Q. On the other hand, there are two symmetrical solutions for the network that can be either a low-pass filter or a high-pass filter.

The frequency f_0 does not need to be the center frequency, $(f_1 + f_2) / 2$, of the desired band limited by f_1 and f_2. In fact, as can be seen from the low-pass configuration of Figure 3-224, it may be interesting to shift f_0 toward the high-band edge frequency f_2 to obtain a larger bandwidth w, where

$$w = \frac{2(f_1 + f_2)}{f_2 - f_1} \tag{3-201}$$

This will, however, be at the expense of poor harmonic rejection.

Example. For a transformation ratio of $n = 4$, we can determine the following values from the above relations:

Bandwidth w	0.1	0.3
Maximum insertion losses	0.025	0.2
X_1/R_1	1.730	1.712

If the terminations R_1 and R_2 have a reactive component X, the latter may be taken as part of the external resistance as shown in Figure 3-225. This compensation is applicable as long as

$$Q_{int} = \frac{X_{\text{int}}}{R_2} \quad \text{or} \quad \frac{R_1}{X_{\text{int}}} < n - 1 \tag{3-202}$$

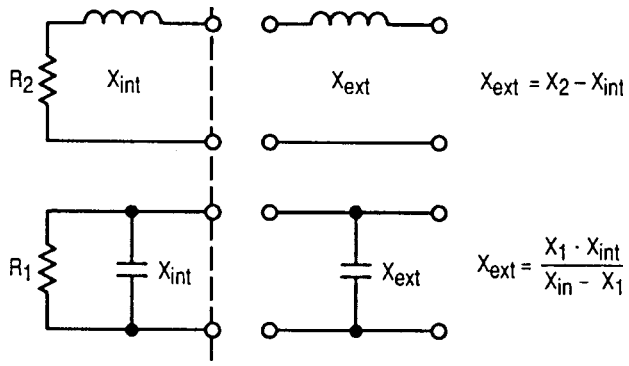


Figure 3-225 Termination reactance compensation.

Tables giving reactance values can be found in Matthaei et al. [33, 34].

USE OF TRANSMISSION LINES AND INDUCTORS. In the preceding section, the inductance was expected to be realized by a lumped element. A transmission line can be used instead (Figure 3-226).

As can be seen from the computed selectivity curves (Figure 3-227) for the two configurations, transmission lines result in a larger bandwidth. The gain is important for a transmission line having a length $L = \lambda/4$ ($\Theta = 90°$) and a characteristic impedance $Z_0 = \sqrt{R_1 R_2}$. It is not significant for lines short with respect to $\lambda/4$. One will notice that there is an infinity of solutions, one for each value of C, when using transmission lines.

THREE-REACTANCE MATCHING NETWORKS. The networks that will be investigated are shown in Figure 3-228. They are made of three reactances alternately connected in series and shunt.

A three-reactance configuration allows the designer to make the quality factor (Q) of the circuit and the transformation ratio $n = R_2/R_1$ independent of each other and consequently to choose the selectivity between certain limits.

For narrowband designs, one can use the following formulas (see [35], where tables are given):

Network (a)

$$X_{C1} = R_1/Q \quad Q \text{ must be selected first} \tag{3-203}$$

$$X_{C2} = R_2 \sqrt{\frac{R_1 R_2}{(Q^2 + 1) - R_1/R_2}} \tag{3-204}$$

$$X_L = \frac{QR_1 + (R_1 R_2/X_{C2})}{Q^2 + 1} \tag{3-205}$$

Network (b)

$$X_{L1} = R_1 Q \quad Q \text{ must be selected first} \tag{3-206}$$

$$X_{L2} = R_2 B \tag{3-207}$$

$$A = R_1(1 + Q^2) \tag{3-208}$$

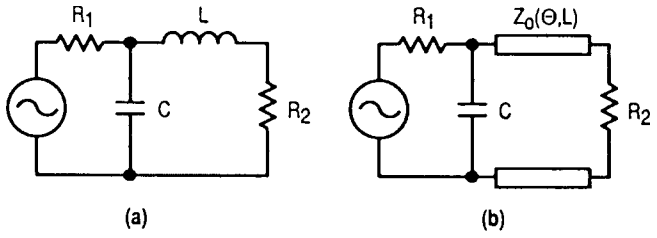


Figure 3-226 Use of a transmission line in the L section.

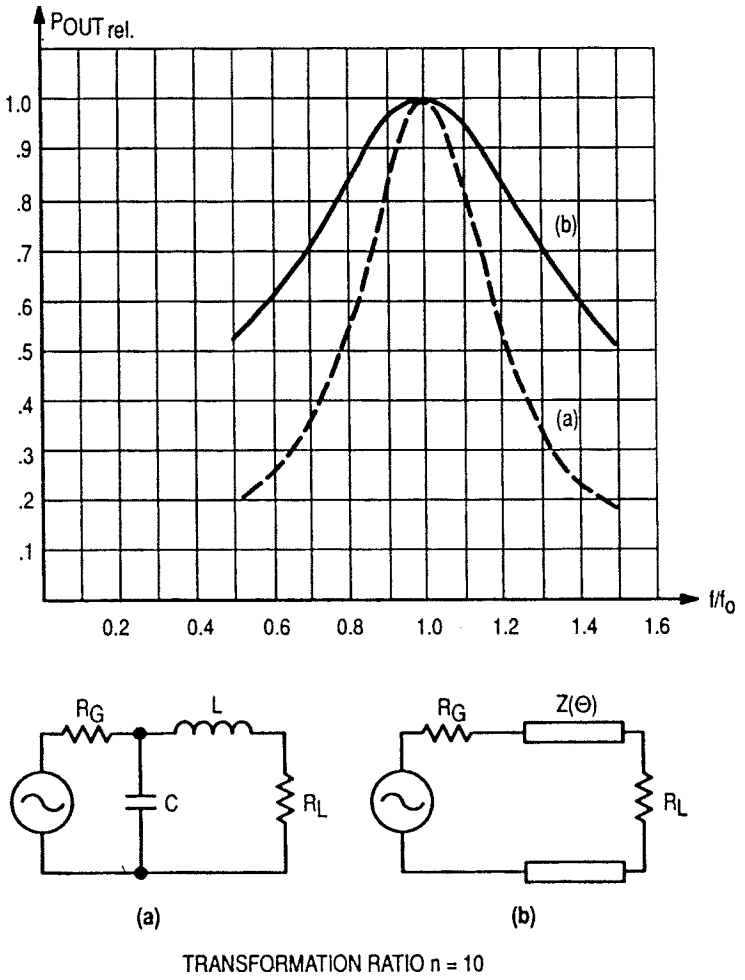


Figure 3-227 Bandwidth of the L section for $n = 10$ (a) with lumped constants and (b) with transmission line (λ/4).

$$X_{C1} = \frac{A}{Q + B} \tag{3-209}$$

$$B = \sqrt{\frac{A}{R_2} - 1} \tag{3-210}$$

Network (c)

$$X_{L1} = QR_1 \quad Q \text{ must be selected first} \tag{3-211}$$

$$X_{C2} = AR_2 \tag{3-212}$$

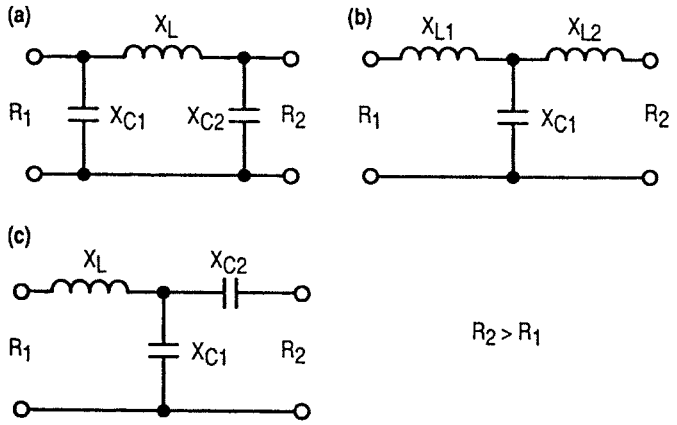


Figure 3-228 Three-reactance matching networks.

$$A = \sqrt{\frac{R_1(1+Q^2)}{R_2} - 1} \tag{3-213}$$

$$X_{C1} = \frac{B}{Q - A} \tag{3-214}$$

$$B = R_1(1+Q^2) \tag{3-215}$$

The network that yields the most-practical component values should be selected for a given application.

The three-reactance networks can be thought of as being formed of an L section (two reactances) and of a compensation reactance. The L section essentially performs the impedance transformation, while the additional reactance compensates for the reactive part of the transformed impedance over a certain frequency band.

Figure 3-229 shows a representation in the Z plane of the circuit of Figure 3-228a split into two parts: $R_1 - C_1 - L_1$ and $C_2 - R_2$. Exact transformation from R_1 into R_2 occurs at the points of intersection M and N. Impedances are then conjugate, or $Z' = R' + jX'$ and $Z'' = R'' + jX''$ with $R' = R''$ and $X' = -X''$.

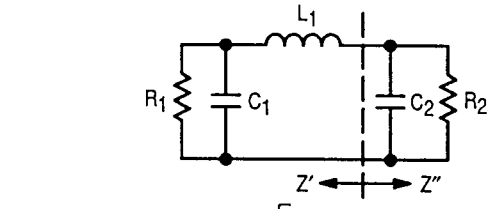
The only possible solution is obtained when X' and $-X''$ are tangential to each other. For the dashed curve, representing another value of L_1 or C_1, a wider frequency band could be expected at the expense of some ripple inside the band. However, this can only be reached with four reactances, as will be shown below.

With a three-reactance configuration, there are not enough degrees of freedom to permit $X' = -X''$ and simultaneously obtain the same variation of frequency on both curves from point M' to point N'. Exact transformation can, therefore, be obtained at only one frequency.

The values of the three reactances can be calculated by making

$$X' = -X'', \quad R' = R'', \quad \text{and} \frac{dX'}{dR'} = -\frac{dX''}{dR''} \tag{3-216}$$

The general solution of these equations leads to complicated calculations. Therefore, computed tables should be used.



$$Z' = \frac{R_1}{1 + \omega^2 R_1^2 C_1^2} + j\omega \left[L1 - \frac{R_1^2 C_1}{1 + \omega^2 R_1^2 C_1^2} \right] = R' + jX'$$

$$Z'' = \frac{R_2}{1 + \omega^2 C_2^2 R_2^2} - \frac{j\omega^2 C_2 R_2^2}{1 + \omega^2 C_2^2 R_2^2} = R'' + jX''$$

$$|x'| = \left(\frac{L_1}{C_1 R_1^2} - \frac{R'}{R_1} \right) \sqrt{\frac{R_1}{R'} - 1}$$

$$|x''| = \sqrt{R' R_2 - R'^2}$$

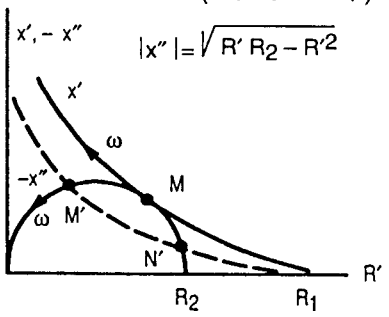


Figure 3-229 *Z*-plane representation of the circuit of Figure 3-228a.

One will note from Figure 3-229 that the compensation reactance contributes somewhat to the impedance transformation; that is, R' varies when going from M to R_2.

The circuit of Figure 3-228b is dual with respect to the first one and gives exactly the same results in a *Y*-plane representation. The circuit of Figure 3-228c is somewhat different since only one intersection M exists, as shown in Figure 3-230. Narrower frequency bands must be expected from this configuration. The widest band is obtained for $C_1 = \infty$.

Again, if one of the terminations has a reactive component, the latter can be taken as a part of the matching network, provided that it is not too large (see Figure 3-227).

FOUR-REACTANCE NETWORKS. Four-reactance networks are used essentially for broadband matching. The networks that will be considered in the following consist of two two-reactance sections in cascade. Some networks have pseudo-low-pass filter character; others, bandpass filter character. In principle, the former show narrower bandwidth since they extend the impedance transformation to very low frequencies unnecessarily, while the latter ensure good matching over a wide frequency band around the center frequency only (see Figure 3-233).

The two-reactance sections used in the networks in Figure 3-231 have either transformation properties or compensation properties. Impedance transformation is obtained with one series reactance and one shunt reactance. Compensation is made with both reactances in series or in shunt. If two cascaded transformation networks are used, transformation is accomplished partly by each one.

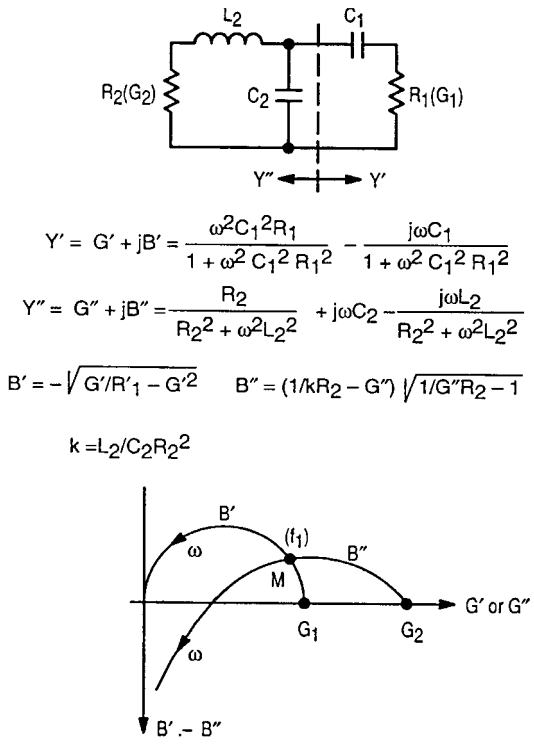


Figure 3-230 *Y*-plane representation of the circuit of Figure 3-228c.

With four-reactance networks there are two frequencies, f_1 and f_2, at which the transformation from R_1 into R_2 is exact. These frequencies may also coincide. For the network in Figure 3-231b, for instance, at point M, R_1 or R_2 is transformed into $\sqrt{R_1R_2}$ when both frequencies fall together. At all points (M), Z_1 and Z_2 are conjugate if the transformation is exact.

In the case of Figure 3-231b, the reactances are easily calculated for equal frequencies:

$$X_1 = \frac{R_1}{\sqrt{(n-1)^{1/2}}} \tag{3-217}$$

$$X_2 = R_1\sqrt{\frac{n^{1/2}-1}{n}} \tag{3-218}$$

$$X_1X_4 = R_1R_2 = X_2X_3 \tag{3-219}$$

$$X_3 = \frac{R_1}{\sqrt{n(n^{1/2}-1)}} \tag{3-220}$$

$$X_4 = \frac{R_1}{n}\sqrt{n^{1/2}-1} \tag{3-221}$$

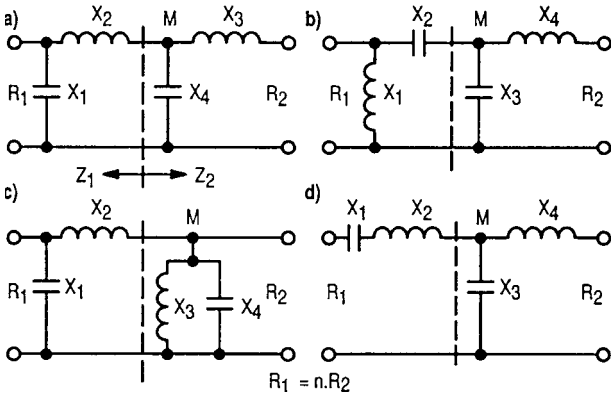


Figure 3-231 Four-reactance networks.

For the network in Figure 3-231a, normally at point M, Z_1 and Z_2 are complex. This pseudo-low-pass filter has been computed elsewhere [33]. Many tables can be found in the literature for networks of four and more reactances having Chebyshev character or maximally flat response [33, 34, 36].

Figure 3-232 shows the transformation path from R_1 to R_2 for networks (a) and (b) in Figure 3-231 on a Smith Chart.

Case (a) of Figure 3-232 has been calculated using tables in Matthaei [34]. Case (b) has been obtained from the relationship given above for $X_1, \ldots, X_4$. Both apply for a transformation ratio equal to 10 and for $R_1 = 1$. There is no simple relationship for $X_1', \ldots, X_4'$ of network (b) if f_1 is made different from f_2 for larger bandwidth. Figure 3-233 shows the respective bandwidths of networks (a) and (b) for the circuits shown in Figure 3-232.

If the terminations include a reactive component, the computed values for X_1 or X_4 may be adjusted to compensate for this.

For configuration (a), it can be seen from Figure 3-232 that, in the considered case, the Q values are equal to 1.6. For configuration (b), Q_1', which is equal to Q_2', is fixed for each transformation ratio

n	2	4	8	10	16
$Q'_1 = Q'_2$	0.65	1	1.35	1.46	1.73

$$Q' = \sqrt{n^{1/2} - 1} \qquad (3\text{-}222)$$

The maximum value of reactance that the terminations may have for use in this configuration can be determined from the above values of Q'.

If R_1 is the load resistance of a transistor, the internal transistor resistance may not be equal to R_1. In this case, the selectivity curve will be different from the curves given in Figure 3-233. Figure 3-234 shows the selectivity for networks (a) and (b) when the source resistance R_1 is infinite. From Figure 3-234 it can be seen that network (a) is more sensitive to changes in R_1 than network (b).

As mentioned earlier, the four-reactance network can also be thought of as two cascaded two-reactance sections: one used for transformation, the other for compensation. Figure 3-235 shows commonly used compensation networks, together with the associated L section.

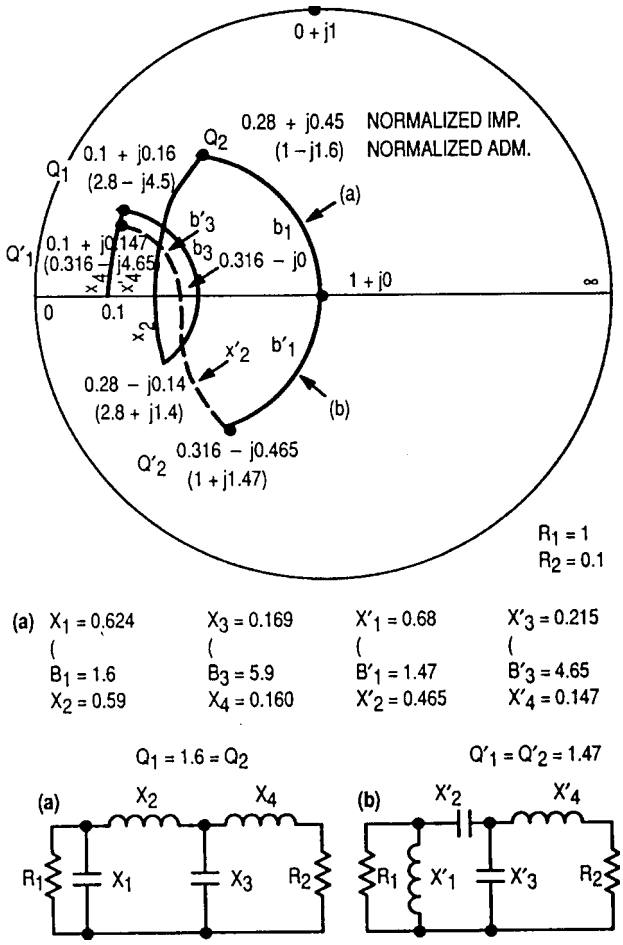


Figure 3-232 Transformation paths for networks (a) and (b).

The circuit of Figure 3-235a can be compared to the three-reactance network shown in Figure 3-228c. The difference is that capacitor C_2 of that circuit has been replaced by an *LC* circuit. The resulting improvement can be seen by comparing Figure 3-236 with Figure 3-230.

By adding one reactance, exact impedance transformation is achieved at two frequencies. It is now possible to choose component values such that the point of intersection M' occurs at the same frequency f_1 on both curves and simultaneously that N' occurs at the same frequency f_2 on both curves. Among the infinite number of possible intersections, only one allows the achievement of this.

When M' and N' coincide in M, the new $dX'/df = dX''/df$ condition can be added to the condition $X' = -X''$ (for three-reactance networks) and similarly $R' = R''$ and $dR'/df = dR''/df$.

Again, a general solution of the above equations leads to still more complicated calculations than in the case of three-reactance networks. Therefore, tables are preferable [33, 34, 36].

The circuit of Figure 3-235b is the dual of the circuit of Figure 3-233a and does not need to be treated separately. It gives exactly the same results in the *Z* plane. Figure 3-235c shows a higher-order compensation requiring six reactive elements.

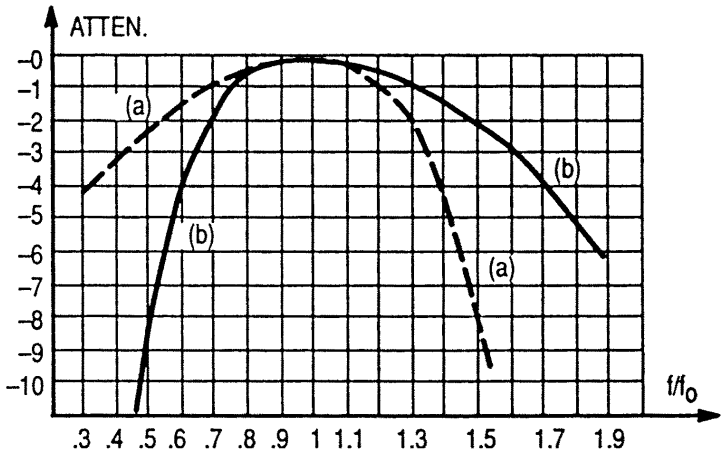


Figure 3-233 Selectivity curves for networks (a) and (b) of Figure 3-232.

The above-discussed matching networks employing compensation circuits result in narrower bandwidths than the former solutions (see p. 573) using two transformation sections. A matching arrangement with higher-order compensation, such as in Figure 3-235c, is not recommended. Better use can be made of the large number of reactive elements by using them all for transformation.

When the above configurations are realized using short portions of transmission lines, the equations or the usual tables no longer apply. The calculations must be carried out on a computer because of their complexity. However, a graphical method can be used (see the next section) that will consist essentially in tracing a transformation path on the Z–Y chart using the computed lumped element values and replacing it by the closest path obtained with distributed constants. The bandwidth change is not significant as long as short portions of lines are used.

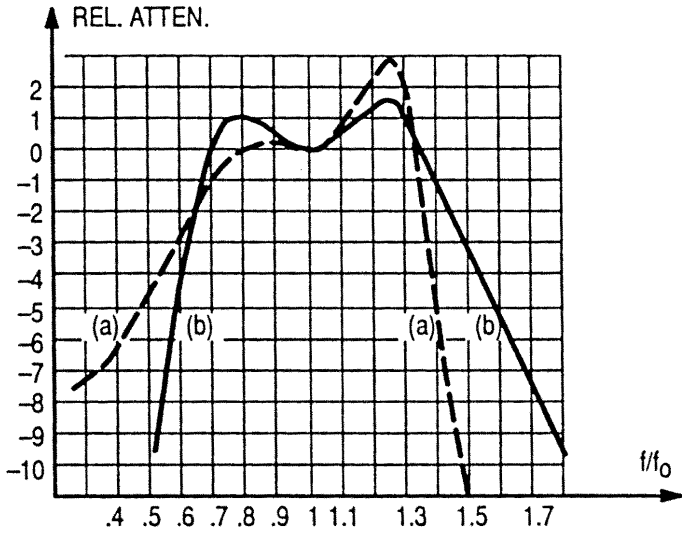


Figure 3-234 Selectivity curves for networks (a) and (b) of Figure 3-232 with infinite R_1.

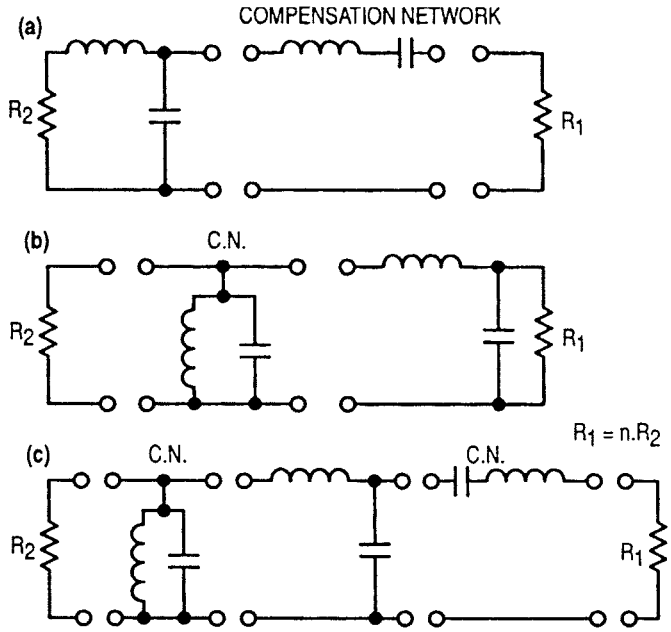


Figure 3-235 Compensation networks used with an L section.

MATCHING NETWORKS USING QUARTER-WAVE TRANSFORMERS. At sufficiently high frequencies, where λ/4-long lines of practical size can be realized, broadband transformations can easily be accomplished by the use of one or more λ/4 sections. Figure 3-237 summarizes the main relations for (a) one-section and (b) two-section transformations. A compensation network can be realized using a λ/2-long transmission line. Figures 3-238 and 3-239 show the selectivity curves for different transformation ratios and section numbers.

EXPONENTIAL LINES. Exponential lines have largely frequency-independent transformation properties. The characteristic impedance of such lines varies exponentially with their length l:

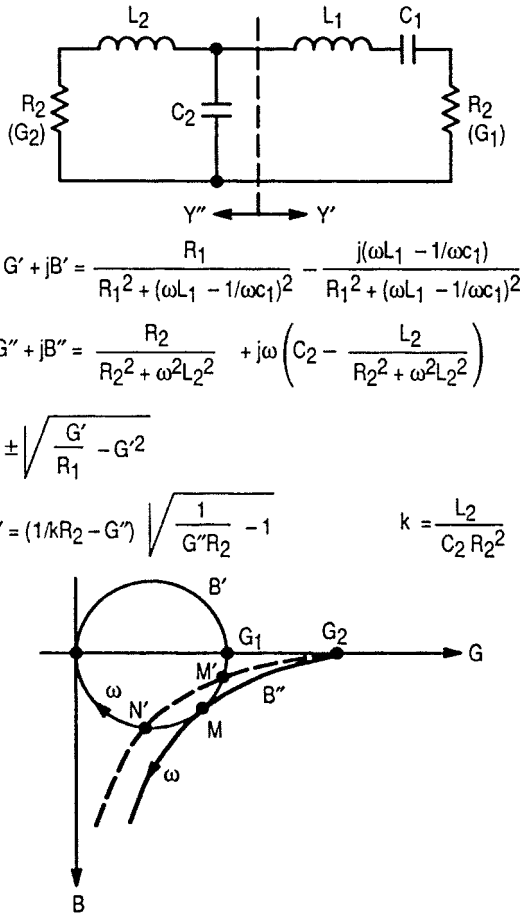
$$Z = Z_0 e^{kl} \tag{3-223}$$

where k is a constant, but these properties are preserved only if k is small.

BROADBAND MATCHING USING BANDPASS FILTER NETWORKS—HIGHQ CASE. The above circuits are applicable to devices having low input or output Q, if broadband matching is required. Generally, if the impedances to be matched can be represented, for instance, by a resistor R in series with an inductor L (sometimes a capacitor C) within the band of interest and if L is sufficiently low, the latter can be incorporated into the first inductor in the matching network. This is also valid if the representation consists of a shunt combination of a resistor and a reactance.

Practically, this is feasible for Q values around 1 or 2. For higher Q values or for input impedances consisting of a series or parallel resonant circuit (see Figure 3-221), as it appears to be for large bandwidths, a different treatment must be followed.

Let us first recall that, as shown by Bode [37] and Fano [38], limitations exist on the impedance matching of a complex load. In the example of Figure 3-240, the load to be matched consists of a capacitor C and a resistor R in shunt.



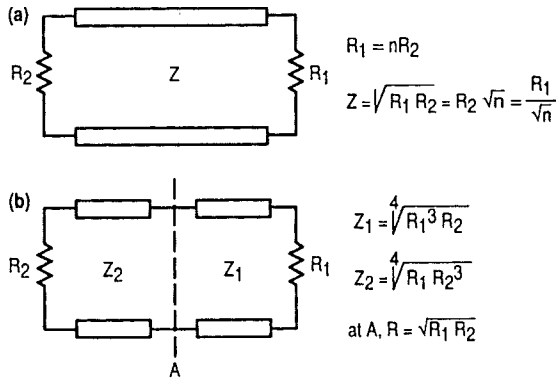
$$Y' = G' + jB' = \frac{R_1}{R_1^2 + (\omega L_1 - 1/\omega c_1)^2} - \frac{j(\omega L_1 - 1/\omega c_1)}{R_1^2 + (\omega L_1 - 1/\omega c_1)^2}$$

$$Y'' = G'' + jB'' = \frac{R_2}{R_2^2 + \omega^2 L_2^2} + j\omega\left(C_2 - \frac{L_2}{R_2^2 + \omega^2 L_2^2}\right)$$

$$B' = \pm\sqrt{\frac{G'}{R_1} - G'^2}$$

$$B'' = (1/kR_2 - G'')\sqrt{\frac{1}{G''R_2} - 1} \qquad k = \frac{L_2}{C_2 R_2^2}$$

Figure 3-236 *Y*-plane representation of the circuit of Figure 3-235a.



$$R_1 = nR_2$$

$$Z = \sqrt{R_1 R_2} = R_2\sqrt{n} = \frac{R_1}{\sqrt{n}}$$

$$Z_1 = \sqrt[4]{R_1^3 R_2}$$

$$Z_2 = \sqrt[4]{R_1 R_2^3}$$

$$\text{at A, } R = \sqrt{R_1 R_2}$$

Figure 3-237 Transformation networks using λ/4-long transmission lines.

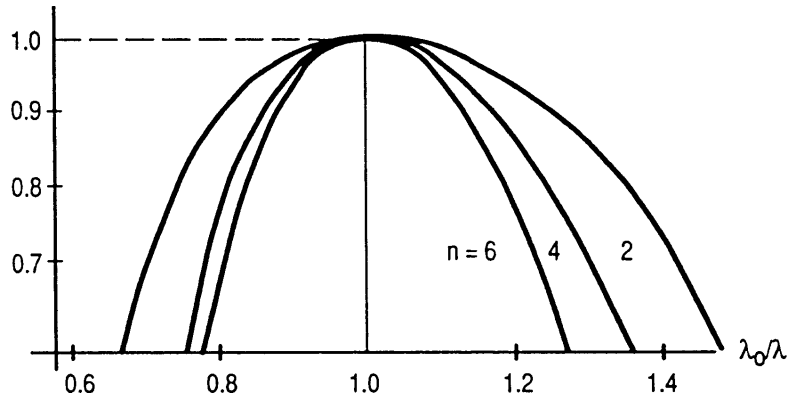


Figure 3-238 Selectivity curves for two λ/4-section networks at different transformation ratios.

The reflection coefficient between the transformed load and generator is equal to

$$\Gamma = \frac{Z_{\text{IN}} - R_G}{Z_{\text{IN}} + R_G} \tag{3-224}$$

When $\Gamma = 0$, there is perfect matching; when $\Gamma = 1$, there is total reflection. The ratio of reflected to incident power is

$$\frac{P_r}{P_i} = |\Gamma|^2 \tag{3-225}$$

The fundamental limitation on the matching takes the form

$$\int_{\omega=0}^{\infty} \ln\left(\frac{1}{|\Gamma|}\right) d\omega \le \frac{\pi}{RC} \qquad \text{Bode equation} \tag{3-226}$$

and is represented in Figure 3-241.

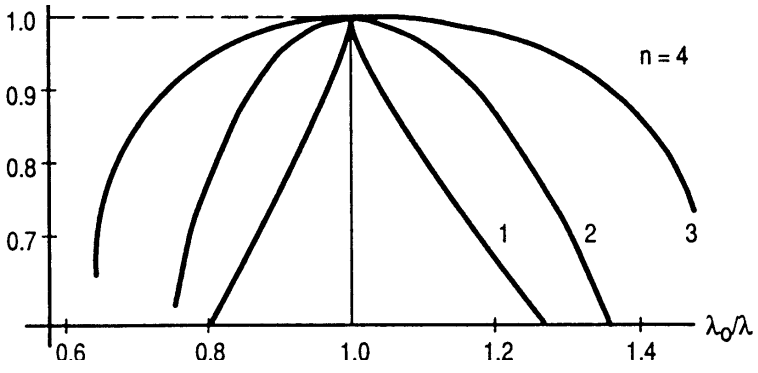


Figure 3-239 Selectivity curves for one, two, and three λ/4 sections.

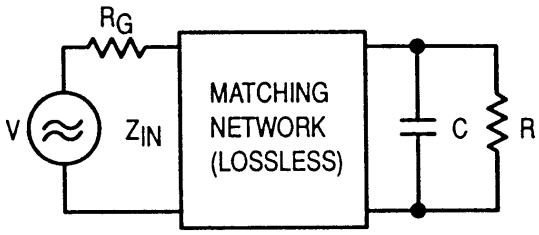


Figure 3-240 General matching conditions.

The meaning of the Bode equation is that the area S under the curve cannot be greater than π/RC and, therefore, if matching is required over a certain bandwidth, this can only be done at the expense of less power transfer within the band. Thus, power transfer and bandwidth appear as interchangeable quantities.

It is evident that the best utilization of the area S is obtained when $|\Gamma|$ is kept constant over the desired band ω_c and made equal to 1 over the rest of the spectrum. Then

$$|\Gamma| = e^{-\pi/\omega_c RC} \tag{3-227}$$

within the band and no power transfer happens outside. A network fulfilling this requirement cannot be obtained in practice because an infinite number of reactive elements would be necessary.

If the attenuation is plotted versus frequency for practical cases, one may expect to have curves like the ones shown in Figure 3-242 for a low-pass filter having Chebyshev character.

For a given complex load, an extension of the bandwidth from ω_1 to ω_2 is possible only with a simultaneous increase of the attenuation a. This is especially noticeable for Q values exceeding 1 or 2 (see Figure 3-243). Thus, devices having relatively high input Q values are usable for broadband operation, provided the consequent higher attenuation or reflection introduced is acceptable.

For average insertion losses or attenuation a (neglecting the ripple), the attenuation decreases if the number n of the network element increases. But above $n = 4$, the improvement is small. For a given attenuation a and bandwidth, the larger the n the smaller the ripple; for a given attenuation, the larger the n the larger the bandwidth.

Computations show that for $Q < 1$ and $n \leq 3$, the attenuation is below 0.1 dB (approximately). The impedance transformation ratio is not free here. The network is a

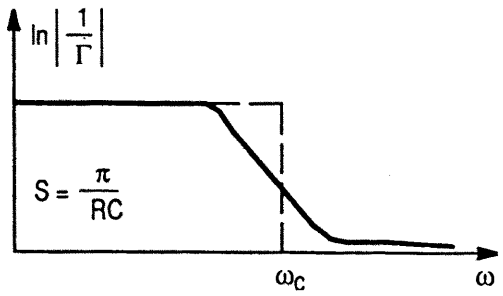


Figure 3-241 Representation of the Bode equation.

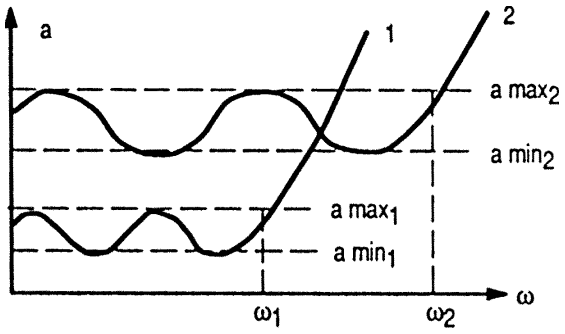


Figure 3-242 Attenuation versus angular frequency for different bandwidths with the same load.

true low-pass filter. For a given load, the optimum generator impedance will result from the computation.

Before impedance transformation is introduced, a conversion of the low-pass prototype into a bandpass filter network must be made. Figure 3-244 summarizes the main relations for this conversion. r is the conversion factor. For the bandpass filter, $Q_{\text{IN max}}$, or the maximum possible input Q of the device to be matched, has been increased by the factor r (from Figure 3-244, $Q'_{\text{IN max}} = rQ_{\text{IN max}}$).

Impedance inverters will be used for impedance transformation. These networks are suitable for insertion into a bandpass filter without affecting the transmission characteristics.

Figure 3-245 shows four impedance inverters. Note that one of the reactances is negative and must be combined in the bandpass network with a reactance of at least equal positive value. Insertion of the inverter can be made at any convenient place [32, 39].

When using the bandpass filter for matching the input impedance of a transistor, reactances $L'_1C'_1$ should be made to resonate at ω_0 by addition of a convenient series reactance.

As stated earlier, the series combination of R_0, L'_1 and C'_1 normally constitutes the equivalent input network of a transistor when considered over a large bandwidth. This is a good approximation up to about 500 MHz.

In practice, the normal procedure for using a bandpass filter matching network will be the following:

1. For a given bandwidth center frequency and input impedance of a device to be matched—for example, 50 Ω—first determine Q'_{IN} from the datasheet as

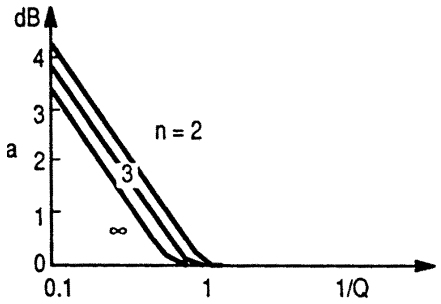


Figure 3-243 Insertion losses as a function of 1/*Q*.

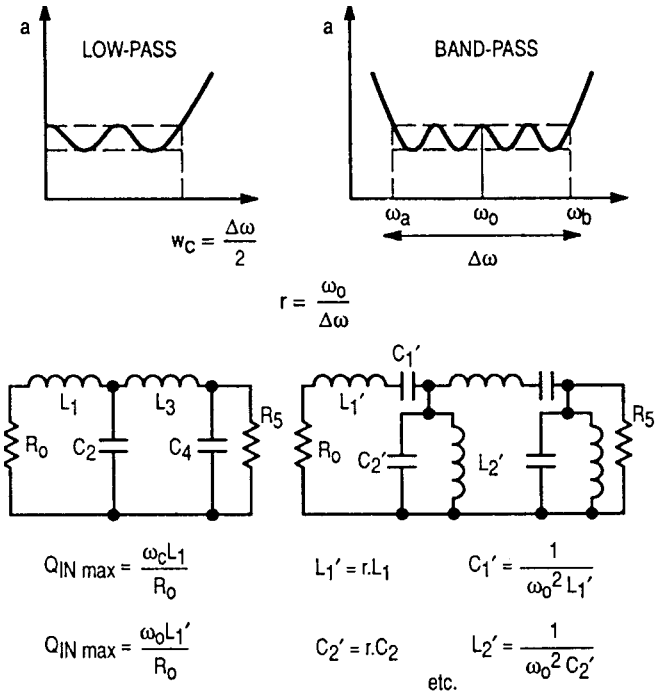


Figure 3-244 Conversion from low-pass filter into bandpass filter.

$$\frac{\omega_0 L_1'}{R_0} \tag{3-228}$$

after having eventually added a series reactor for centering.

2. Convert the equivalent circuit $R_0 L_1' C_1'$ into a low-pass prototype $R_0 L_1$ and calculate Q_{IN} using the formulas of Figure 3-244.
3. Determine the other reactance values from tables [33] for the desired bandwidth.
4. Convert the element values found by Step 3 into series- or parallel-resonant circuit parameters.
5. Insert the impedance inverter in any convenient place.

In the above discussions, the gain roll-off has not been taken into account. This is of normal use for moderate bandwidths (e.g., 30%). However, several methods can be employed to obtain a constant gain within the band despite the intrinsic gain decrease of a transistor with frequency. Tables have been computed elsewhere [40] for matching networks approximating 6 dB/octave attenuation versus frequency.

Another method consists of using the above-mentioned network and then adding a compensation circuit, as shown in Figure 3-246.

Resonance ω_b is placed at the high edge of the frequency band. Choosing Q correctly, roll-off can be made 6 dB/octave. The response of the circuit of Figure 3-246 is expressed by

$$\frac{1}{1 + Q^2(\omega/\omega_b - \omega_b/\omega)^2} \tag{3-229}$$

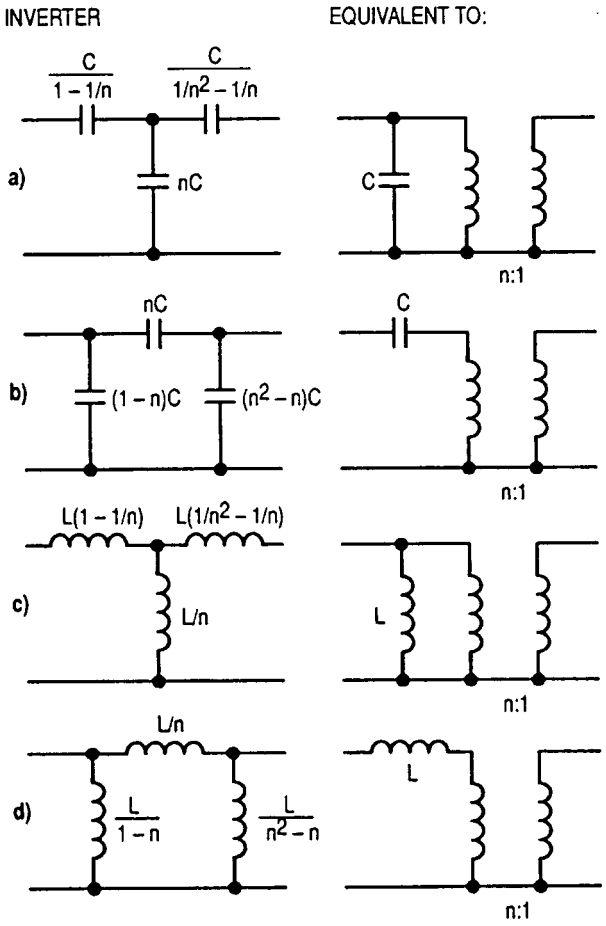


Figure 3-245 Impedance inverters.

where $\omega < \omega_b$. This must be equal to ω/ω_b for 6 dB/octave attenuation.

At the other band edge a, exact compensation can be obtained if

$$Q = \frac{(\omega_b/\omega_a)^2 - 1}{\omega_a/\omega_b - (\omega_b/\omega_a)^2} \tag{3-230}$$

There are several methods available for synthesis of matching networks. They should not be confused with filters because they can be either high-pass, low-pass, or bandpass, and their main purpose is to match one complex impedance to another. In many cases, one of the terminations happens to be 50 Ω and real, but this is just a "lucky" case of the general approach. An excellent paper on this [41] refers to the "real frequency" technique. Another equally important publication is the one by Cuthbert [42], since followed by his book [43]. Two more good references are Vendelin et al. [44] and Potter [45].

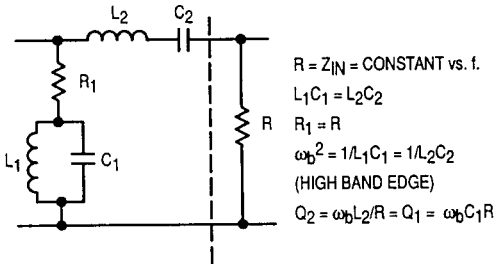


Figure 3-246 Roll-off compensation network.

3-12-3 Example 2: Low-Noise Amplifier Using Distributed Elements

In our previous examples, the amplifiers were built around lumped elements. Since a transmission line length of less than $\lambda/4$ is an inductor and $\frac{3}{4}\lambda$ is a capacitor, the whole amplifier can be built around microstrip technology.

If anyone wonders why we are going to show a low-noise amplifier in the middle of the power section, the reason is that we will select a low-power Siemens transistor, which, being operated in Class A, has an intercept point of +30 dBm and can easily achieve a noise figure of less than 2 dB. Such an amplifier is attractive when feeding an input stage from an antenna that sees many signals and curtails the bandwidth with an input filter of good selectivity that needs a good output termination, which in many cases works against the achievement of a good noise figure because when designing for best noise figure the input impedance of the transistor stage is far from 50 Ω.

First, we will start with the textbook approach and use the Smith diagram to obtain the best noise figure of which the device is capable. Figure 3-247 shows the actual approach to input and output matching.

We start our design by querying the Smith Tool about the minimum noise figure and the gain that comes with it. We mark the point of the lowest noise figure (labeled Γ_{opt}). The small

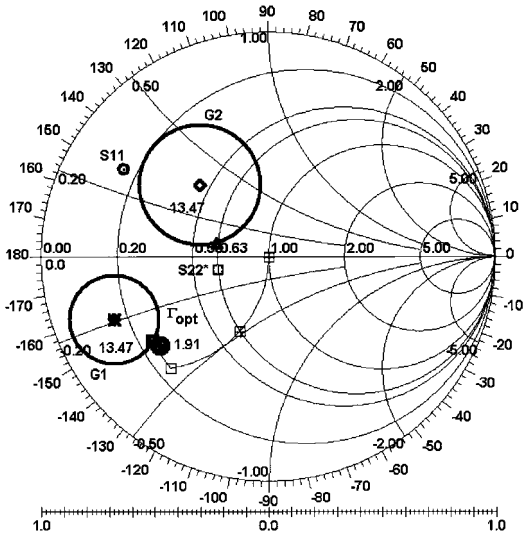


Figure 3-247 Input and output matches for 13.47 dB of gain and NF = 1.91 dB as displayed by the interactive Smith Tool.

noise circle yields the location for Γ_{opt} and tells us the minimum noise figure is going to be 1.91 dB. (By the way, this assumes no further losses, even in soldering the transistor to the PC board.) Then we construct the input gain circle (G1), which touches Γ_{opt} and turns out to represent 13.47 dB of available gain.

The next step is to find a matching network between Γ_{opt} and the 50-Ω source impedance. Immediately there is a conflict: If we look at S_{11}, which is the input reflection coefficient shown in the Smith diagram, it is on the other side of the complex plane and sufficiently far away, indicating that there is going to be quite a difference between noise matching and power-gain matching. We have two options for matching the input: Either we start from the origin and move toward the position of Γ_{opt}, or we start with Γ_{opt} and go to 50 Ω. Depending on the direction in which we do this, the result can be different topologies, such as a high-pass or low-pass filter. The high-pass filter is obtained by starting with Γ_{opt}; the low-pass topology is obtained by starting at the 50-Ω point. The Smith Tool then provides us with the corresponding values of capacitors and inductors after the transformation has been accomplished.

The final step is to do the same type of matching at the output. We start with S_{22}, and either select S_{22}^* (the complex conjugate for highest gain) or S_{22load}. As a rule, the second one is always a lower impedance (in terms of the real part) than the conjugate match. This graphic procedure is somewhat trial and error. As to the topology, a high-pass match should be preferred at the input and a low-pass match at the output. The reason for this is that the high-pass filter will tend to protect the stage from the enormous number of signals present below 1 GHz; using a low-pass filter at the output suppresses unwanted harmonics that the amplifier may generate. Figure 3-248 shows one of the approaches we like best simply because it gives a good input termination and we were trying for a simultaneous match for best noise figure and power gain.

The way we were able to achieve this was to take advantage of the existing feedback in the transistor and select a less-than-optimal output termination but gaining at the input. Taking advantage of the Miller effect, we rotate the phase angle of the feedback with the load to move the optimum noise and gain matching points closer. The danger in this is that the circuit is now highly sensitive to both input and output termination. The output *must* see 50 Ω to maintain this good input match, but since the second transistor stage can now be optimized for best input match, a two-stage combination can be configured to meet this

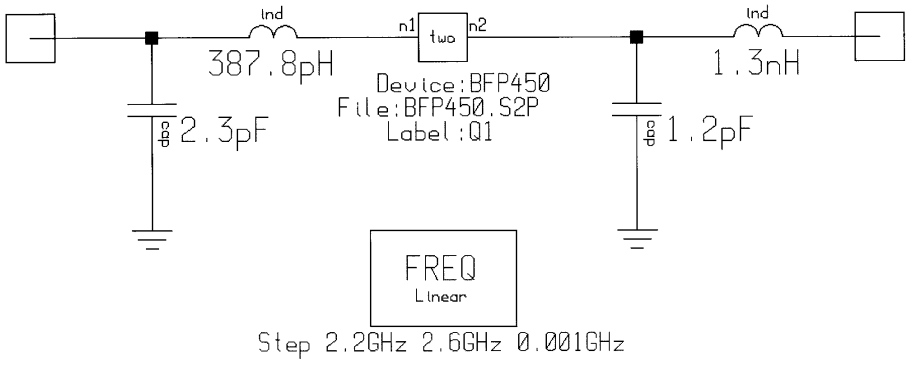


Figure 3-248 A 2.4-GHz amplifier using a BFP450 transistor and lumped-element input and output matching.

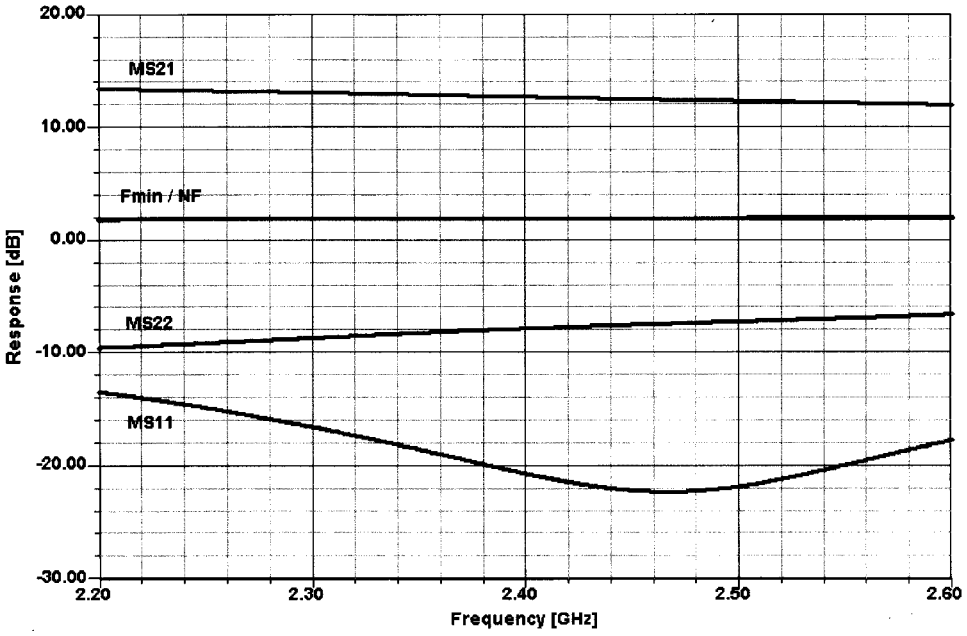


Figure 3-249 Frequency-dependent gain, matching, and noise characteristics of the matched high-dynamic-range BFP450 2.4-GHz amplifier.

requirement. Also, let us not forget that the device is running at 50 mA and as far as bipolar transistors are concerned, this is probably the best amplifier combination we have seen for getting the highest dynamic range and lowest noise figure. The resulting gain, matching, and noise figure characteristics are shown in Figure 3-249.

Now we turn to using distributed elements and at the same time use a shunt feedback at the output; to be specific, we will load it with 150 Ω. The reason for this is to be able to obtain a good compromise at the input that will depend less on the output and at the same

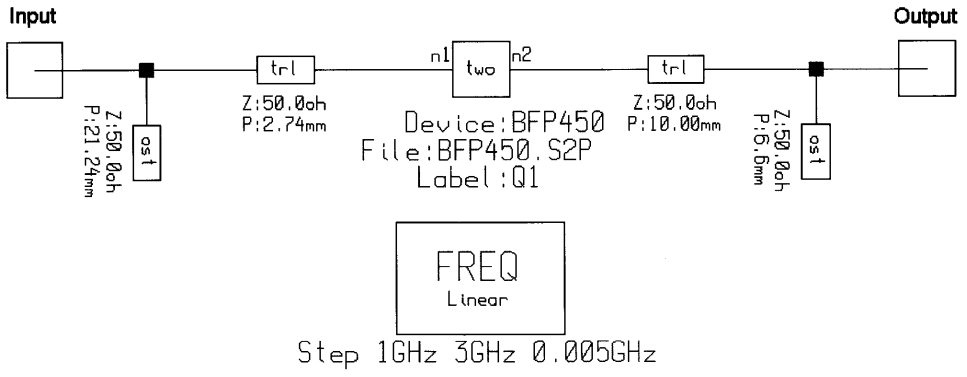


Figure 3-250 Distributed-element input and output matching for the 2.4-GHz BFP450 amplifier.

time improve the S_{22} matching at the output. The procedure for doing the matching with a transmission line offers only certain degrees of freedom, while doing the matching with lumped elements allows the choice of series or parallel inductors or capacitors. In the case of a transmission line, we have open stubs, shorted stubs, and transmission lines. This may result in somewhat more complex matching schemes. On the positive side, they can easily be trimmed, are reproducible, and certainly cost less than exotic capacitors, which need to have small tolerances and to be able to handle high RF currents.

The resulting input and output matching is shown in Figure 3-250, and its frequency response in Figure 3-251. We were able to maintain a good input and output match with an insignificant deterioration of the noise figure.

Since this is a high-dynamic-range amplifier, we also show its predicted two-tone response in Figure 3-252. The manufacturer specifies an $IP_{3,in}$ of 29 dB for a bias point of $V_{CE} = 3$ V and $I_C = 50$ mA with $Z_S = Z_{S,opt}$ and $Z_L = Z_{L,opt}$ at 1.8 GHz. Significantly, however, the manufacturer-supplied nonlinear BFP450 library we used exhibits considerably different characteristics at $V_{CE} = 4$ V and $I_C = 50$ mA than those reflected in the corresponding S parameters supplied by the same manufacturer. In our circuit, the S parameters resulted in about 13.5 dB of gain (based on a compromise between matching for optimum noise and maximum gain) versus about 9.5 dB of gain for the nonlinear library (match optimized for maximum gain). The difference between the 2.4-GHz F_{min} values returned for the two modeling approaches is similarly striking: 1.90 dB for the bare S parameter set and 7.76 dB for the bare nonlinear library biased to $V_{CE} = 4$ V and $I_C = 50$ mA. Clearly, the nonlinear modeling of the BFP450 needs work!

Having fully evaluated the amplifier at 2.4 GHz as a working example, we note that this frequency is used by radio amateurs for some applications. We will therefore supply

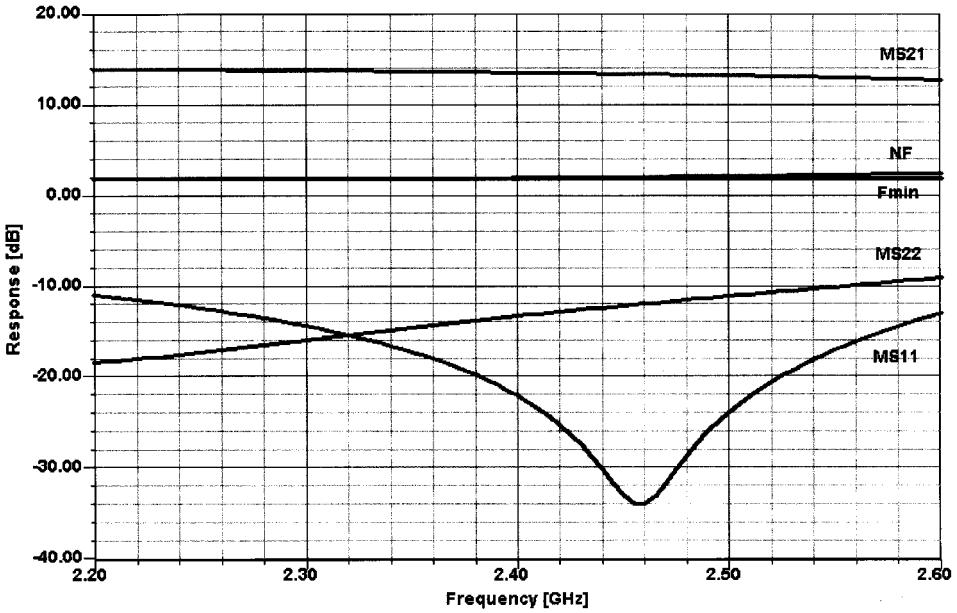


Figure 3-251 Predicted frequency-dependent gain, matching, and noise performance for the BFP450 amplifier with distributed-element matching.

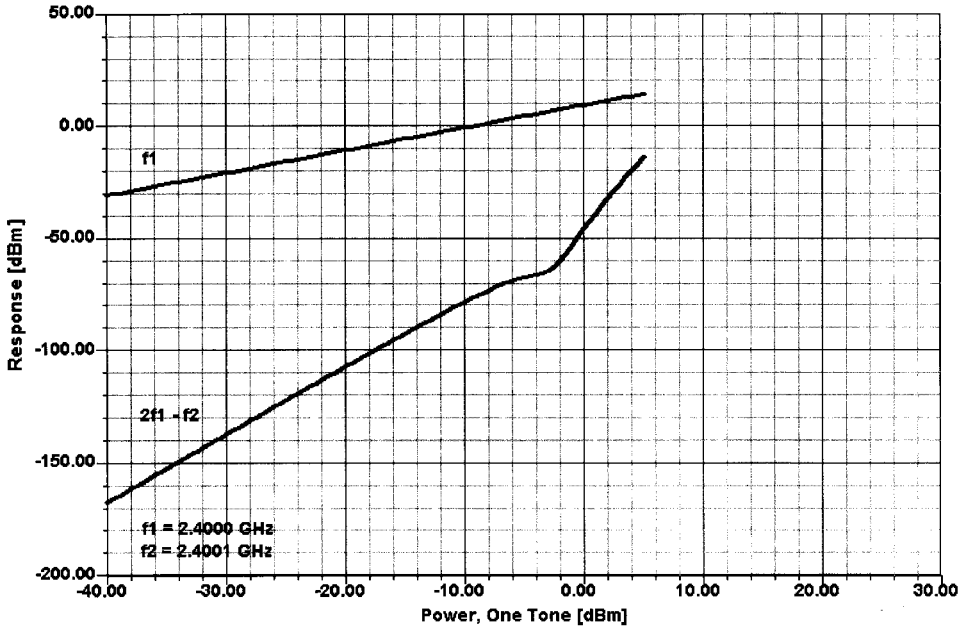


Figure 3-252 Predicted two-tone response of the BFP450 amplifier with the circuit adjusted for maximum gain (9.5 dB): $IP_{3,in} \approx 28.5$ dBm and $IP_{3,out} \approx 38$ dBm. That this version's gain is significantly lower than that of the *S*-parameter-based equivalent illustrates the need for better accuracy in parameter extraction.

examples for 432 and 1296 MHz, popular spot frequencies in the amateur 70- and 23-cm bands, respectively. Figures 3-253 and 3-254 show the 432-MHz version's schematic and frequency-dependent responses. Figures 3-255 and 3-256 show the schematic and frequency-dependent responses for the 1296-MHz version.

Using a bipolar transistor that operates at much less current, the noise figure can be improved at the expense of the dynamic range/intercept point. It is possible to reduce the noise figure by about 1 dB or so, but the reduction in intercept point will be much more dramatic. Another choice is the use of GaAsFETs for this application, but their matching in comparable circuits has been described as a nightmare simply because the devices have too much gain and very high input impedances (essentially 0.2 pF in series with 3 Ω) combined with reasonable output impedances in the vicinity of 240 Ω. Here is a wide field in which engineers can experiment and examine price/performance and performance/real estate issues. GaAsFETs also tend to require more than 30–40 mA to be alive, and even the low operating voltage of 4–5 V makes the device consume a lot of dc power.

3-12-4 Example 3: 1-W Amplifier Using the CLY15

The following is an application showing a fairly wideband amplifier with +30 dBm (1 W) output. We need to point out that in accordance with the various digital modulation modes used with this amplifier, this is not continuous operation, but a much less than 1:1 duty cycle operation. The device would probably not survive operating at full output under CW conditions. This is subject to a discussion with the manufacturer, Siemens, who has not

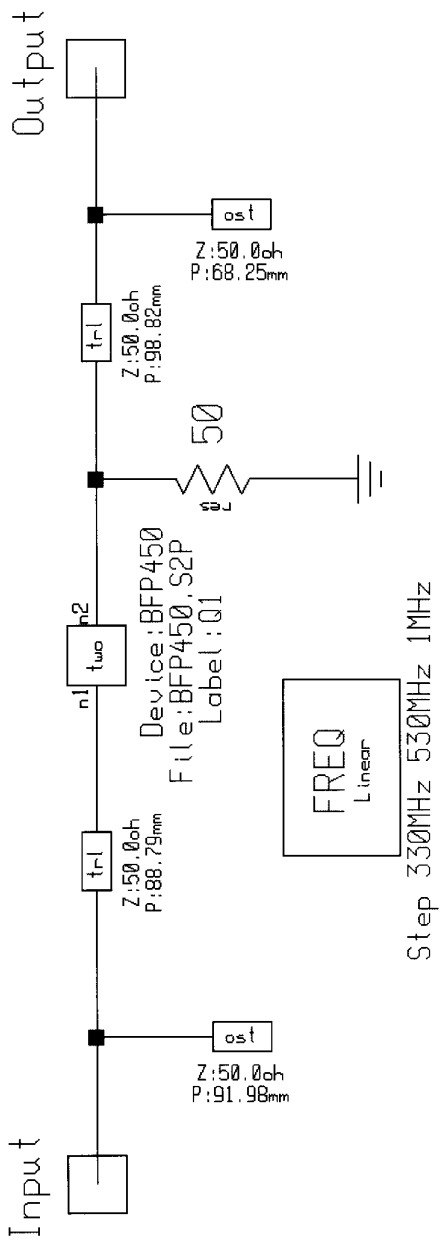


Figure 3-253 Schematic of BPF450 amplifier for 432 MHz. Because f_T = 25 GHz for the BFP450, at this lower frequency we must shunt the device output with a 50-Ω resistor to maintain stability.

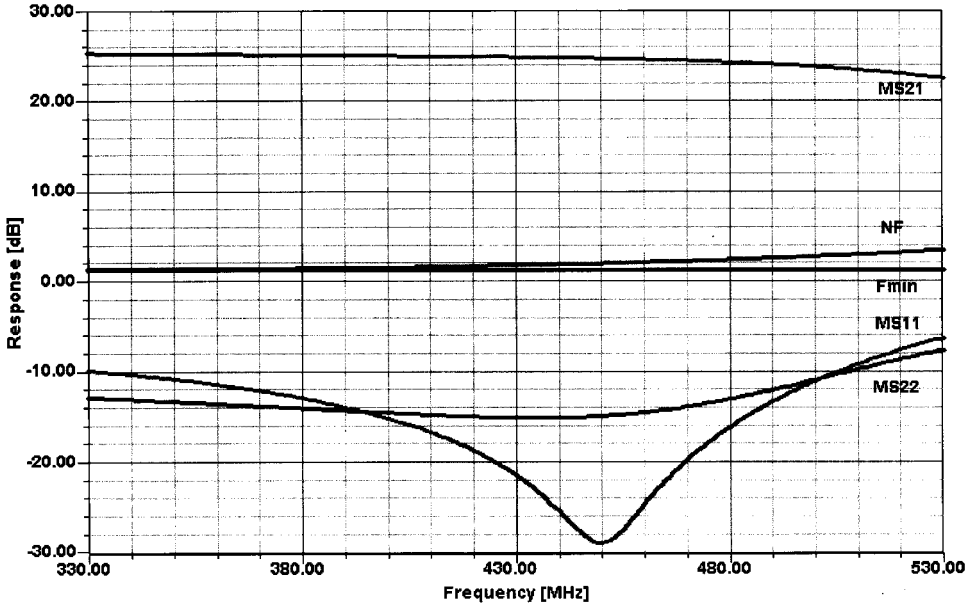


Figure 3-254 Frequency-dependent gain, matching, and noise performance for the BPF450 amplifier at 432 MHz. As with the circuit in Figure 3-253, the input and output network values have been chosen to achieve a low NF and good input matching at the expense of the output match.

released sufficient data for us to make a judgment as to the CLY15's continuous power capability. Since there is no useful large-signal model available at the time of this writing, we will design this Class A amplifier using large dc bias *S* parameters. In the actual design there will be some cut-and-try, but we find that in this type of application this is still the best way of handling the design.

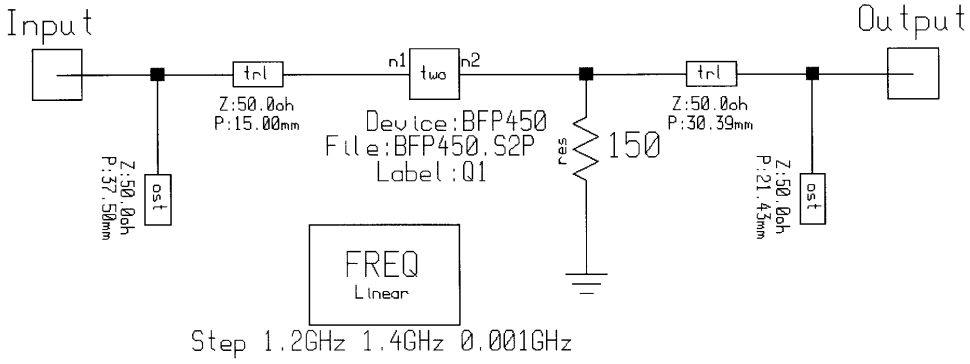


Figure 3-255 Schematic of the BPF450 amplifier for 1296 MHz. A loading resistor is still necessary to ensure stability, but because the transistor exhibits less gain at this frequency than at 432 MHz, we can increase the loading resistor to 150 Ω.

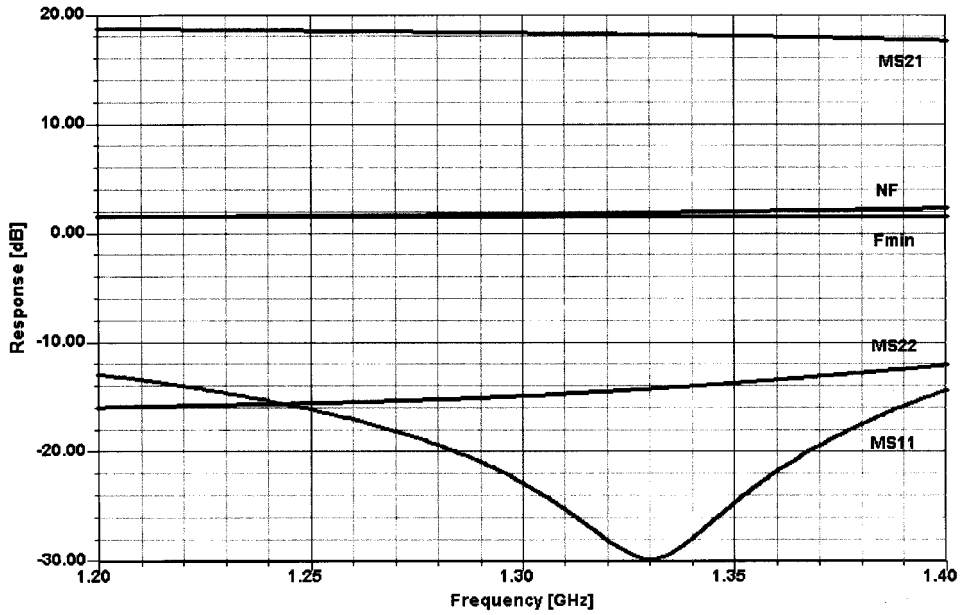


Figure 3-256 Frequency-dependent gain, matching, and noise responses of the BFP450 amplifier at 1296 MHz. As with the two previous versions, the input and output network values have been chosen to achieve a low NF and good input matching at the expense of the output match.

The CLY15 (gate width 16 mm) is the largest of a series of Siemens GaAsFETs; it was already mentioned in Chapter 2, where its datasheet is reproduced. Its device family also includes the CLY2 (gate width 2 mm), CLY5 (4 mm), and CLY10 (8 mm). It is possible that bigger devices will be offered in the future. One of the most critical parameters will be the saturation voltage; from the datasheet of the CLY15, we can see that at 1.5 A the saturation voltage is approximately 1 V. This drop of 1 V includes losses that result from other parasitic elements; the dc *I–V* curve (Figure 3-257a) indicates a saturation voltage of 0.75. Since we aim for 1-W output power, the resistive portion of the load will be $(5 - 1)^2 / (1 \times 2) = 8\ \Omega$. This assumes a supply voltage of 5 V.

The next step is to design the input and output matching network. Previewing the *S* parameters, it will be noticed that they are essentially inductive, which means that the input match, using lumped elements, will result in a combination of capacitors at both the input and the output. Figures 3-258 and 3-259 show the ways in which the input and output matches were done on the Smith diagram resulting in lumped matching elements. Inspecting the lumped input matching, we find the following. We start with S_{11}, which is in the lower part of the Smith diagram, and determine its conjugate match, S_{11}^*. We already pointed out that this device is inductive, which means that matching it from the conjugate point will require two capacitors. The first one is a series capacitor, as identified in the Smith diagram, and followed by a shunt capacitor to transform the impedance to 50 Ω. A close inspection of the Smith diagram shows that the resulting point is not quite on S_{11}; it is an iterative process to exactly find the crossover point, yet the input match of better than 20 dB is more than acceptable.

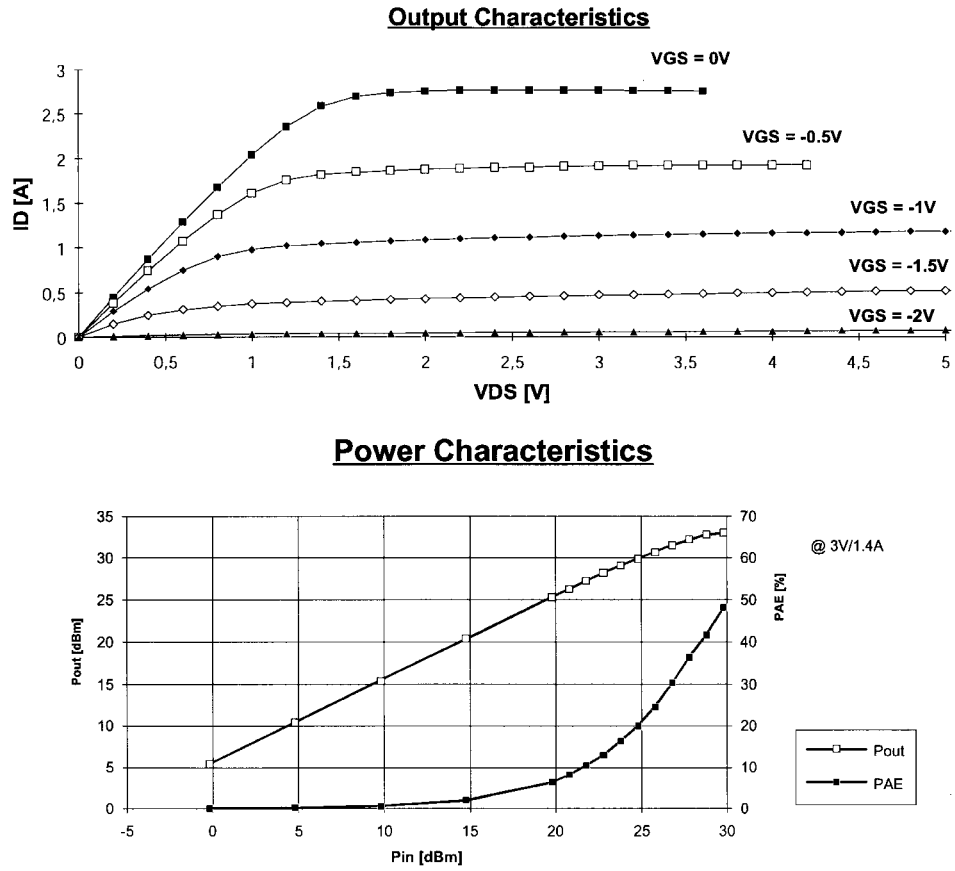


Figure 3-257 Output and power characteristics for the CLY15 medium-power GaAsFET.

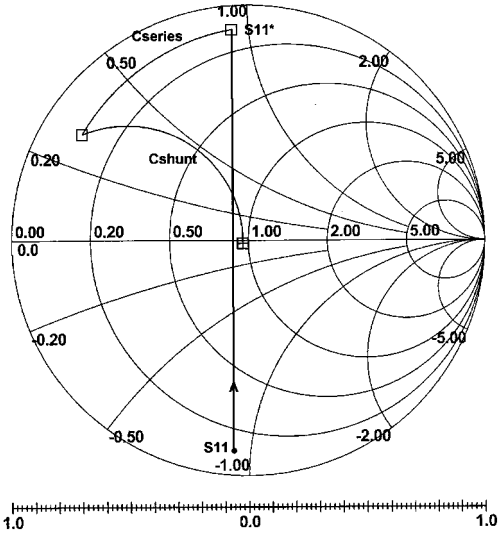


Figure 3-258 Input match for S_{11}^{*} using lumped elements.

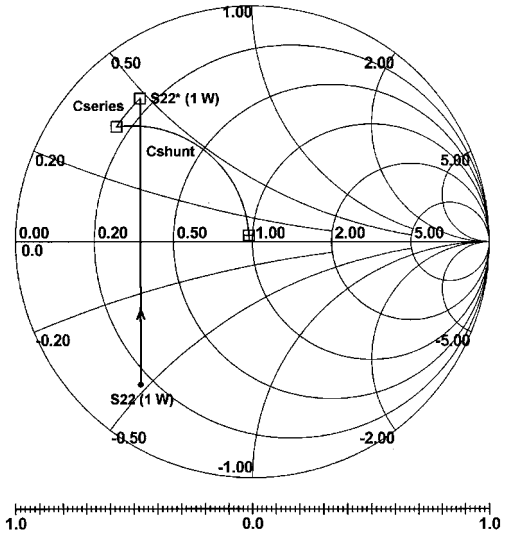


Figure 3-259 Output match for S_{22}^* (1 W) using ideal lumped elements.

As far as the output is concerned, we use the same technique by moving from S_{22} to S_{22}^*. S_{22} has been determined from the required output power; as we showed above, the real portion is 8 Ω. Using the lumped elements, the matching technique is identical; we first need to use a series capacitor and then a shunt capacitor for the transformation. The arc for the series capacitor this time is much longer, resulting in a series capacitance of 11.2 pF instead of the 2.1-pF capacitance used at the input. The actual schematic is shown Figure 3-260. The drawback of lumped elements is that it is not practical to buy components that can handle the heavy RF currents that flow in the input and output networks. Taking the values obtained

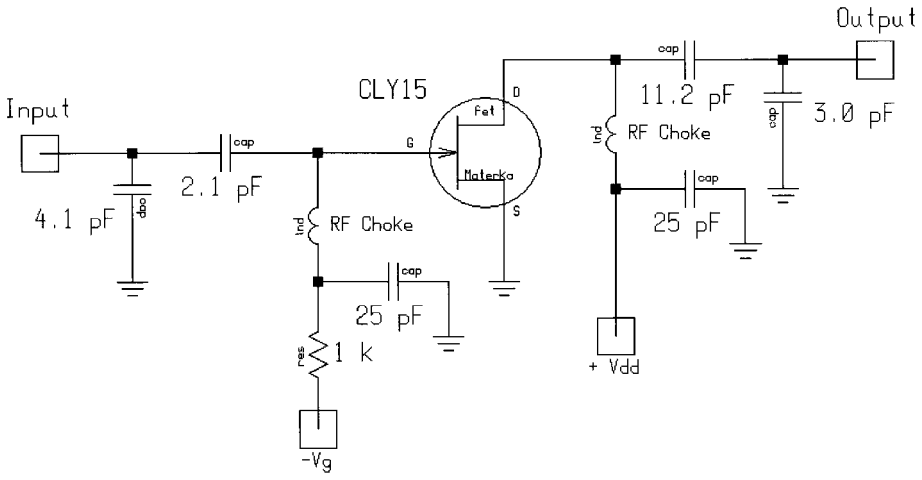


Figure 3-260 The complete amplifier using the CLY15 medium-power GaAsFET, with lumped-element input and output matching.

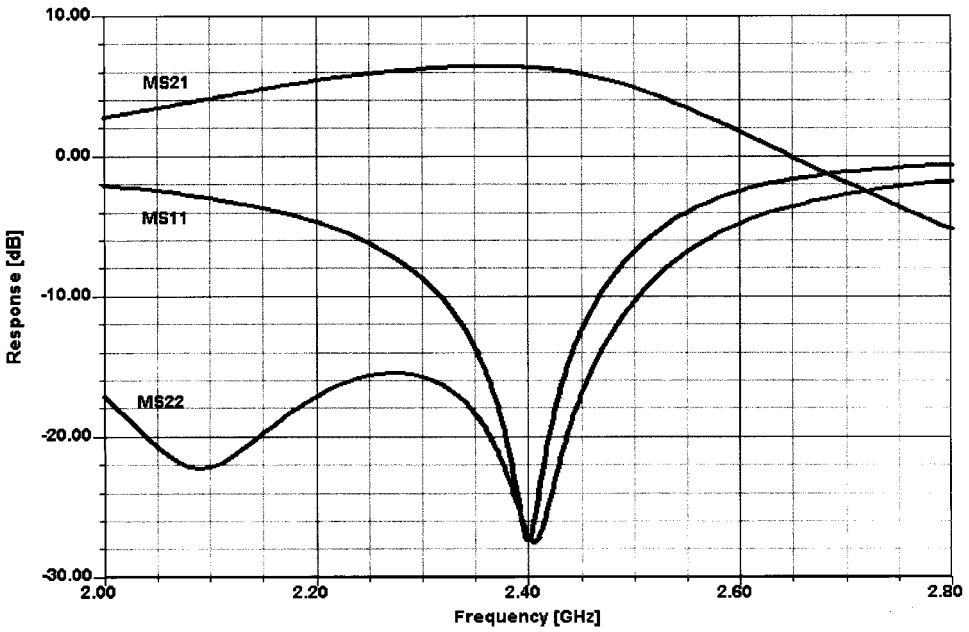


Figure 3-261 Frequency-dependent gain and matching responses of the CLY15 amplifier using ideal lumped elements.

from the lumped input and output matching approach, Figure 3-261 shows the resulting frequency response.

The next step is to provide the same matching condition using distributed elements, which of course have the advantage that they can be printed on the PC board. Instead of using ideal elements, we used a loss tangent of 2E–3, or a Q of 500. The manufacturer of the R4000 material (Rogers) provided this specification. Now we start with the same approach by moving from S_{11} to S_{11}^*. From there, we use an open stub with a transmission-line impedance of 50 Ω, followed by a transmission line to the point in the left lower corner, and, finally, a shorted transmission-line stub allows the 50-Ω matching. For this amplifier, we are not looking for noise matching but for input and output power matching. Figure 3-262 shows this match on the Smith Chart.

At the output, the procedure is again similar: We move from the S_{11} for 1 W to a conjugate value, then use an open stub that acts like a capacitor until we cross the real axis; finally, a λ/4 transmission line is used to match from approximately 75 Ω to 50 Ω. Figure 3-263 shows this match on the Smith Chart. This concludes the distributed matching design for the CLY15 amplifier.

Finally, the complete circuit using distributed elements is shown in Figure 3-264. From a dc point of view, we took advantage of the incoming shorted stub by using a transmission line of the appropriate length and a bypass capacitor to ground. This allows us to feed the gate with the required negative voltage without interfering with the frequency response of the amplifier. When evaluating the total frequency response (Figure 3-265), we noticed that this circuit seems to have a higher Q than the lumped equivalent. Again, to obtain a wider bandwidth more transformation stages will be needed. We leave this task to the interested reader.

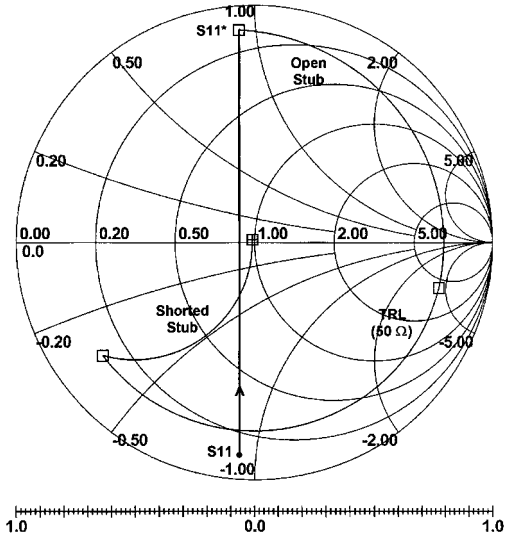


Figure 3-262 Input match for S_{11}^* (1 W) using distributed elements.

The 1-dB bandwidth of the amplifier using distributed, lossy elements is 120 MHz. If a wider bandwidth is required, matching networks with more elements are required. Because of the gain–bandwidth product, the peak gain for such a wider-bandwidth arrangement will be less than this narrowband example.

Having used a lumped to distributed transformation, we next present a power amplifier that uses coaxial lines, baluns, and other transformation steps capable of handling the increased power requirement.

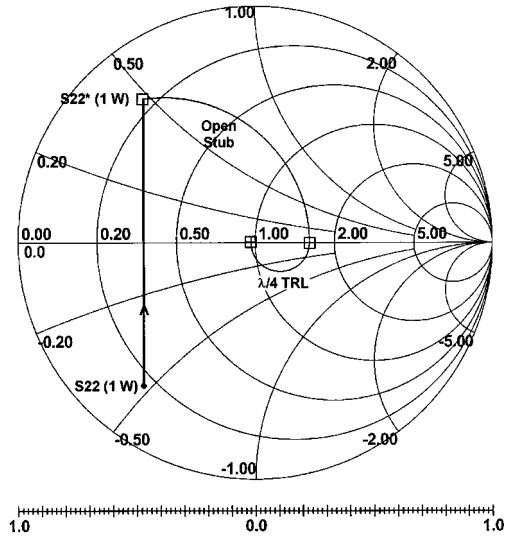


Figure 3-263 Output match for S_{22}^* (1 W) using distributed elements.

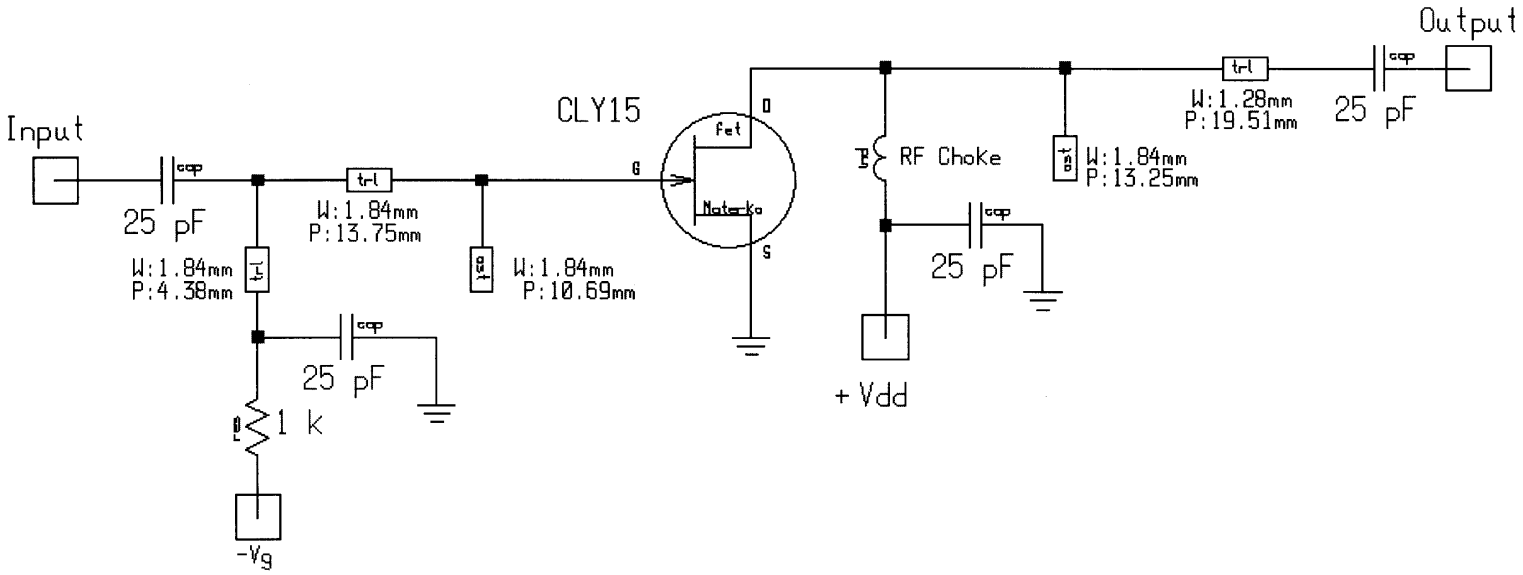


Figure 3-264 The complete amplifier using the CLY15 medium-power GaAsFET, with distributed-element input and output matching.

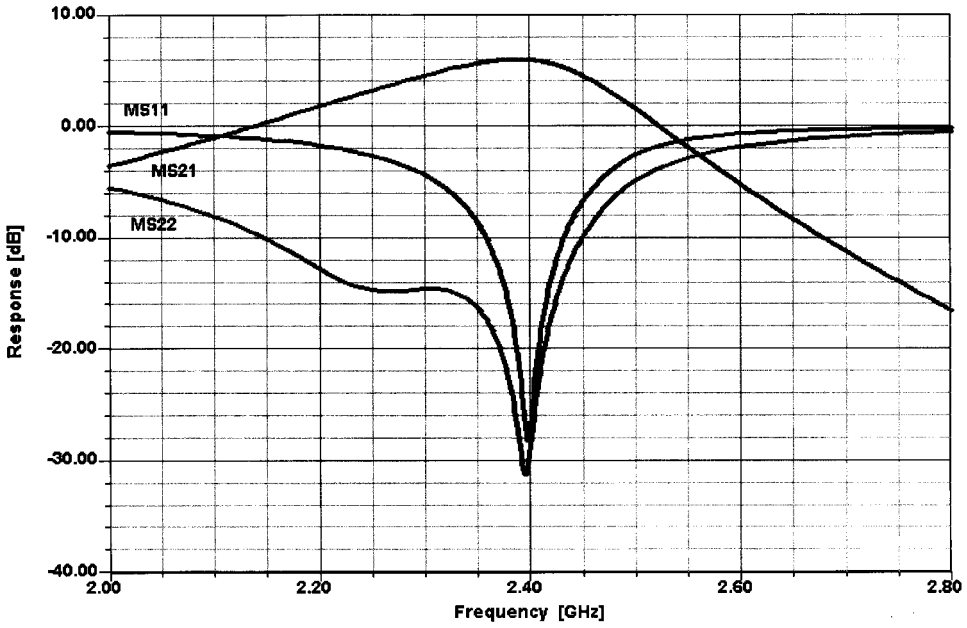


Figure 3-265 Frequency-dependent gain and matching using distributed, lossy elements.

3-12-5 Example 4: 90-W Push–Pull BJT Amplifier at 430 MHz

The last complete application we want to examine is a 90-W output, push–pull amplifier using two TRW transistors (that are no longer made—an all-too-frequent occurrence after a design is complete, resulting in a scramble for replacements). Figure 3-266 shows the circuit diagram of this UHF transistor amplifier, and Figure 3-267 shows its mechanical design. The amplifier requires 28 Vdc and will draw approximately 6 A, requiring 15 W of drive at 438 MHz. This amplifier operates in Class AB, which means it is not limited to amplifying FM and other angle-modulated signals, and can accommodate even single-sideband (SSB) operation.

The drive power coming from the source connected to the input N connector (15 W) is fed through a short piece of coaxial cable to the PC board sitting on top of the heat sink. To change from single ended to push–pull, a 50-Ω coaxial cable is coiled to provide sufficient electrical length to act as a balun that provides two symmetrical paths through the amplifier. The RC section of 10 Ω in parallel with a 220-pF capacitor stabilizes the amplifier at lower frequencies, while at higher frequencies the 10-Ω resistor has no contribution. There are two independent 4:1 transformers, one for each section; they are built using 25-Ω coaxial cable with a set of two binocular ferrite cores to force the transformation. The dc supply is also fed to the transistors at this point. The stabilization circuit consists of a temperature monitor using the *pn* junction of a transistor, side-mounted on the heat sink, and a combination of operational amplifiers and adjustments. This guarantees proper, temperature-independent biasing. The 4:1 input baluns result in 12.5 Ω impedance, or 3.125 Ω per transistor. Since the input impedance at this drive level is somewhere in the vicinity of $[(1 \ldots 2) + j(3 \ldots 5)]$ Ω; the final matching is accomplished by an *RC* section in which the inductance is actually printed

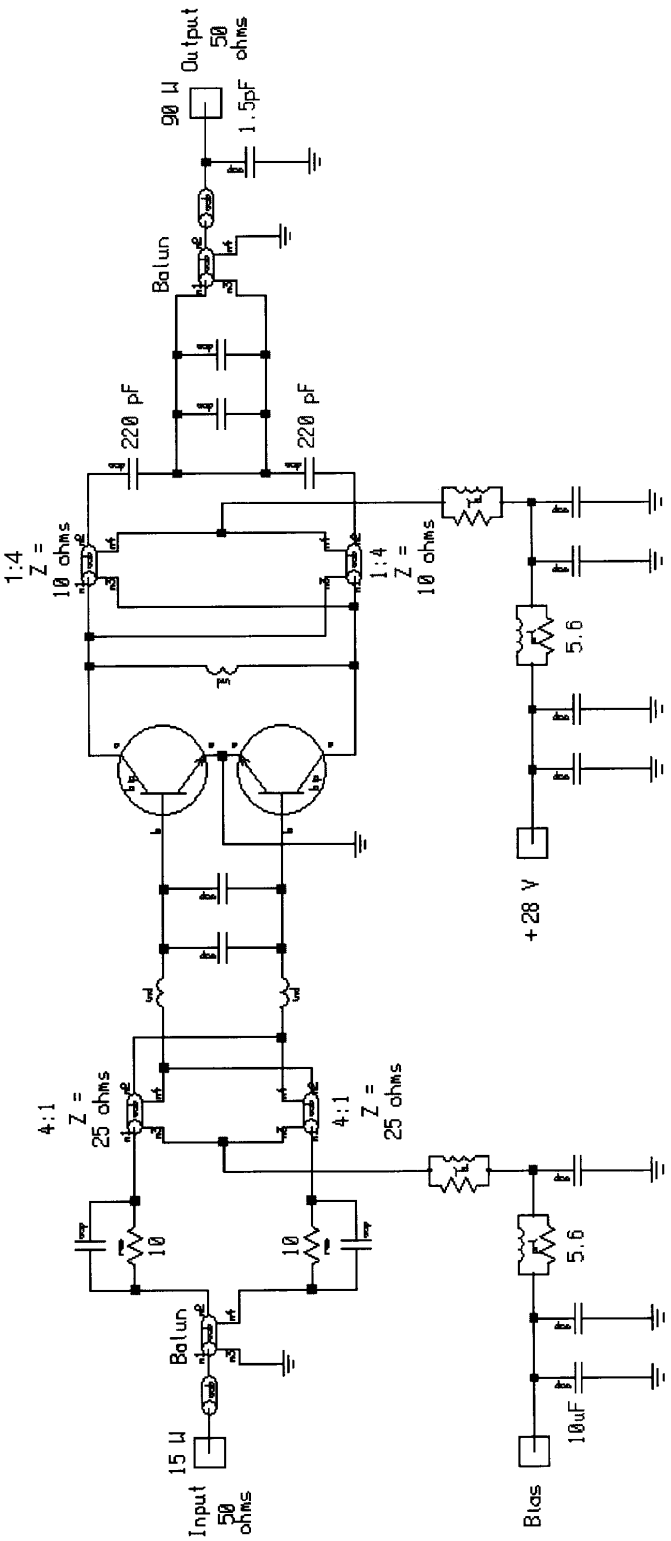


Figure 3-266 Circuit diagram of the 90-W, push–pull, 438-MHz linear amplifier.

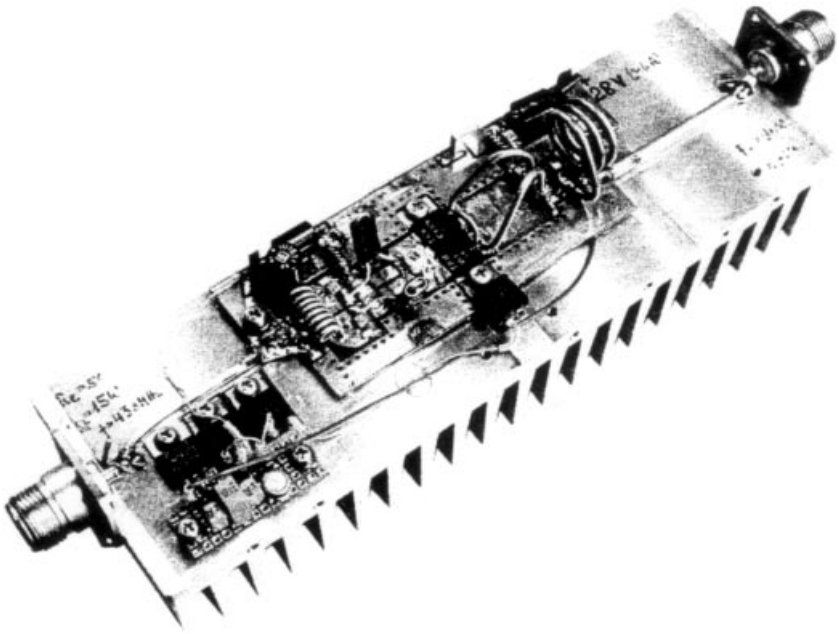


Figure 3-267 Mechanical assembly of the 90-W amplifier shown in Figure 3-266.

and the shunt capacitor across is soldered from base to base as close to the transistor bodies as possible. A ceramic-dielectric variable capacitor in parallel with this allows fine tuning.

The output consists of an inductor from collector to collector that tunes out the capacitive reactance and makes the transistor output impedance real. We then follow the same pattern by taking a 1:4 transformer, but this time built using 10-Ω rigid line instead of 25-Ω rigid line. Together with the output capacitance, prior to the output balun, this is a medium-bandwidth matching arrangement that transforms up to 50 Ω. The collector impedance per collector had to be $(28 - 2)^2/2 \times 90$, resulting in roughly 3.8 Ω—not that different from the input match at the balun point. This way, the matching networks at the input and output can be identical. At the output of the matching network, another coiled 50-Ω coaxial cable serves as balun; because it must operate at the 90-W level, larger-diameter line was used than in the input balun. Finally, another section of rigid line is used to connect the output balun on the PC board to the output N connector. The output capacitor again helps in the transformation.

Both the base and collector voltage supplies are heavily filtered using ferrite-bead-based RF chokes and heavy wire inductors wound on low-impedance resistors. A fairly large number of these amplifiers were actually built and are still in use.

3-12-6 Quasiparallel Transistors for Improved Linearity

Some applications require really low intermodulation distortion. One of the authors (Rohde), during his tenure at RCA Government Systems Division, and his team developed a parallel-device topology that is somewhat similar to the conventional parallel approach but shows vastly increased linearity. In creating higher-power devices, manufacturers parallel several transistors in one package, connecting them in a clever way so the overall drive, dissipation, and output power are equally divided between all the devices. Special nickel–chromium

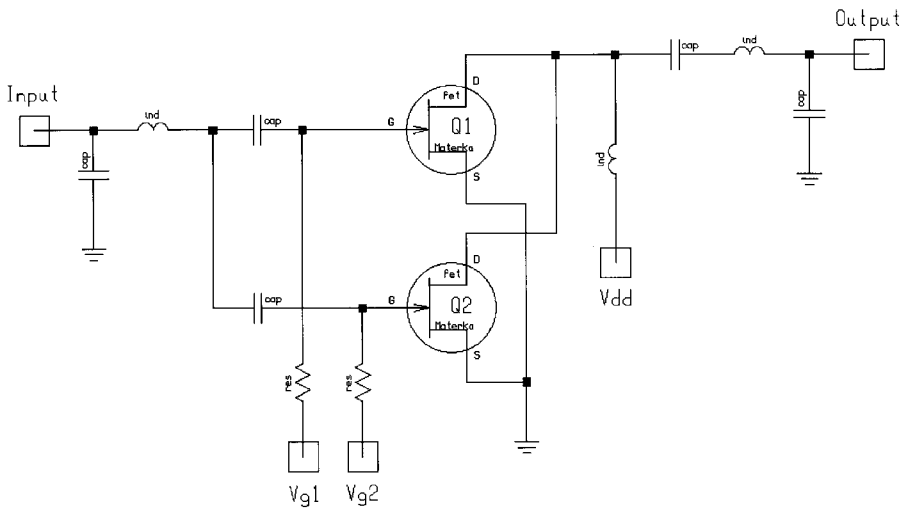


Figure 3-268 The quasiparallel amplifier topology looks like a standard parallel amplifier. Biasing the devices to different operating points nets an improvement in IMD performance.

ballasting resistors in the many fingers of the emitter connections of the multiple internal devices equalize any differences among the transistors while increasing the saturation voltage by only a small amount. In our high-linearity design, we used two transistors in parallel, one being biased somewhere between Class A and B, and the second one somewhere between Class B and C (Figure 3-268). The exact bias voltages and operating points had to be found through experiment. In practice, Q1, the "initially on" transistor, is driven to a point close to that at which its IMD products would increase. Q2 is biased such that at this same drive level it adds to the output power.

We frequently referred to this configuration as "turbocharged." The IMD products at full output power were down 40–45 dB, compared to 26–30 dB for a single transistor. While such an arrangement shows superior IMD performance, because of the somewhat abrupt kick-in of the second transistor, the transconductance response is semismooth, a characteristic that might make alert readers think that IMD will occur at this point. On the other hand, the transconductance for Q1, in effect, saturates or stays constant, so when the transconductance responses of both transistors are overlaid, the result is a smooth response overall.

This technique has been validated up to 100 W and up to at least 50 MHz. Its drawbacks include the need for two transistors with their associated costs, additional real estate, and higher input and output capacitances. Nonetheless, considering that MOS transistors capable of operation to 1 GHz were unavailable when this technique was developed, we believe that it should be revisited.

The Class D and higher operating modes require a large number of additional components and will make the efficiency higher (selective in frequency), but not necessarily reduce the IMD products. A further point of consideration is that it is incorrect to assume that reducing the input level to a given Class A will necessarily reduce its IMD products. As with tubes, the closer one gets to the knee, even in Class A operation, the worse the IMD. (This phenomenon was particularly troublesome in systems using 1-kW amplifiers up to 150 MHz;

underdrive with the intent of getting better IMD performance produced just the opposite result.) If a wide power range is considered, we recommend using a bias scheme that includes sufficient intelligence to find the best operating point.

Another hot amplifier issue is that of IMD as a function of load impedance. If an amplifier is used at output power considerably below its design value, its load impedance will be too low and, depending on the circuit configuration, the result may be additional IMD.

3-12-7 Distribution Amplifiers

It is not uncommon that a signal from a particular source like an antenna has to be distributed to several places. A condominium is a typical situation in which energy from a main signal source must be split for distribution to a larger number of users—there must be enough voltage or power to produce a noise-free picture. An amplifer capable of doing this is a *distribution* amplifier, not to be confused with a *distributed* (or traveling-wave) amplifier, such as that shown earlier in this chapter in Figure 3-6. An interesting means of achieving this is to use multiple Wilkinson couplers following an amplifier. Figure 3-269 shows an arrangement that makes heavy use of multiple power dividers to provide multiple outputs. Theoretically, the reverse is also possible: If we want to collect energy from a number of sources and combine them, as on a cable, we can use Wilkinson couplers at the *input* of a gain block. Note, however, that depending on the level and number of the combined signals, this may put enormous strain on the linearity of the amplifier. What we want to avoid is a lot of distortion, which, in the case of television, will show up as color confetti in the noise and signal.

3-12-8 Stability Analysis of a Power Amplifier

One of the hidden pitfalls in narrowband circuit design is that a circuit may be unstable outside the frequency range of interest. Circuits containing structures built from distributed elements, which exhibit multiple resonances, can be particularly susceptible to such instabilities. In this example, we analyze the stability of a simple FET power amplifier designed using small-signal concepts and the interactive Smith Tool utility in Ansoft's Serenade

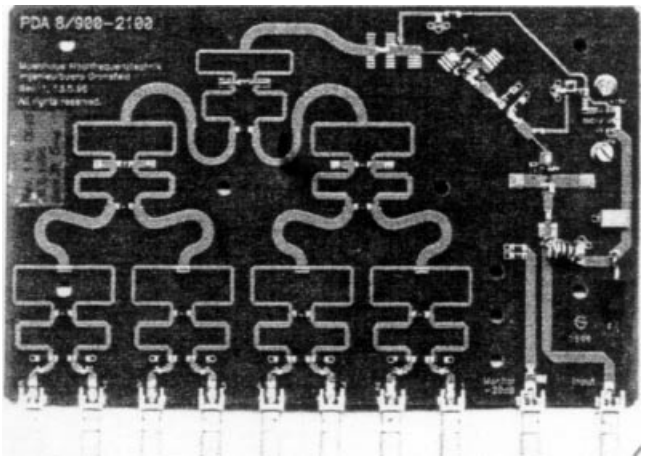


Figure 3-269 Distribution amplifier showing the use of multiple Wilkinson power dividers.

Design Environment. The evaluation shows how K-factor analysis alone is not sufficient to check for stability of a design.

Small-Signal ac Analysis. We begin by taking the Figure 3-270 amplifier, which was designed for about 16 dB of gain at 2.2 GHz. Due to the very low, capacitive input impedance of the FET and the proximity of the source-plane gain circles and stability circles to the edge of the Smith Chart, the matching network attempts to transform 50 Ω to an impedance of approximately $(2.5 + j15)$ Ω. The impedance lies close to the stability circles at 2 GHz but remains in a stable region. The bias point selected (–3.5 V) is chosen for Class AB operation. Higher linear gain can be achieved by reducing the gate voltage to –2.0 V; at this bias point, the amplifier gain will increase to about 19 dB.

First, we will do a small-signal ac analysis and examine the S parameters and K and B_1 factors [see Eqs. (3-154) and (3-157), respectively] from 1 GHz to 6 GHz—well beyond the bandwidth of the design. Figure 3-271 shows the circuit's frequency-dependent gain and matching characteristics.

If we examine B_1 and the stability factor K, we see that the amplifier is not unconditionally stable in the 2.2-GHz passband as shown in Figure 3-272. K dips slightly below 1, although B_1 remains above zero. In practice, a slight change in design would be required to bring K above 1 and remove this potential instability. However, for our example, this is of no consequence, as we shall see below.

Nyquist Stability Analysis. Now we will utilize a Nyquist stability analysis to check the stability of the amplifier. To fully realize the advantages of a Nyquist stability analysis, we must sweep frequency over a very wide range. We want to be able to pick up instabilities at *any* frequency, not just near the design frequency. In this example we sweep from 1 kHz to 20 GHz. A start frequency of 1 kHz is chosen so we can determine the starting point of the Nyquist plot and we ensure that any low-frequency resonances are picked up. The stop frequency is chosen past the f_{max} of the transistor (where we know it cannot oscillate).

The analysis results are shown in polar plot form in Figure 3-273 and in magnitude–angle form in Figure 3-274. The polar graph shows the real and imaginary parts of the system determinant as frequency is swept. The magnitude–angle graph shows the magnitude and cumulative angle (no cut at ±180°) of the system determinant as frequency is swept. The key to interpreting these graphs lies in the following statement drawn from the Nyquist criterion:

The circuit will be unstable if the Nyquist plot encircles the origin in a clockwise direction.

The term "unstable" refers to the existence of natural frequencies lying in the right-half of the complex frequency plane (RHP). These natural frequencies on the RHP will cause oscillations. To encircle the origin in a clockwise direction means that the Nyquist plot will travel in a clockwise path, cross the negative real axis, and continue to completely surround the origin. If the path appears to encircle the origin but then "unravels" itself in a counter-clockwise direction, the origin is not actually encircled. Figure 3-273 shows that the origin is not encircled and therefore this circuit is stable.

The magnitude–angle plot can assist in determining if a clockwise encirclement has actually been made with the help of the following:

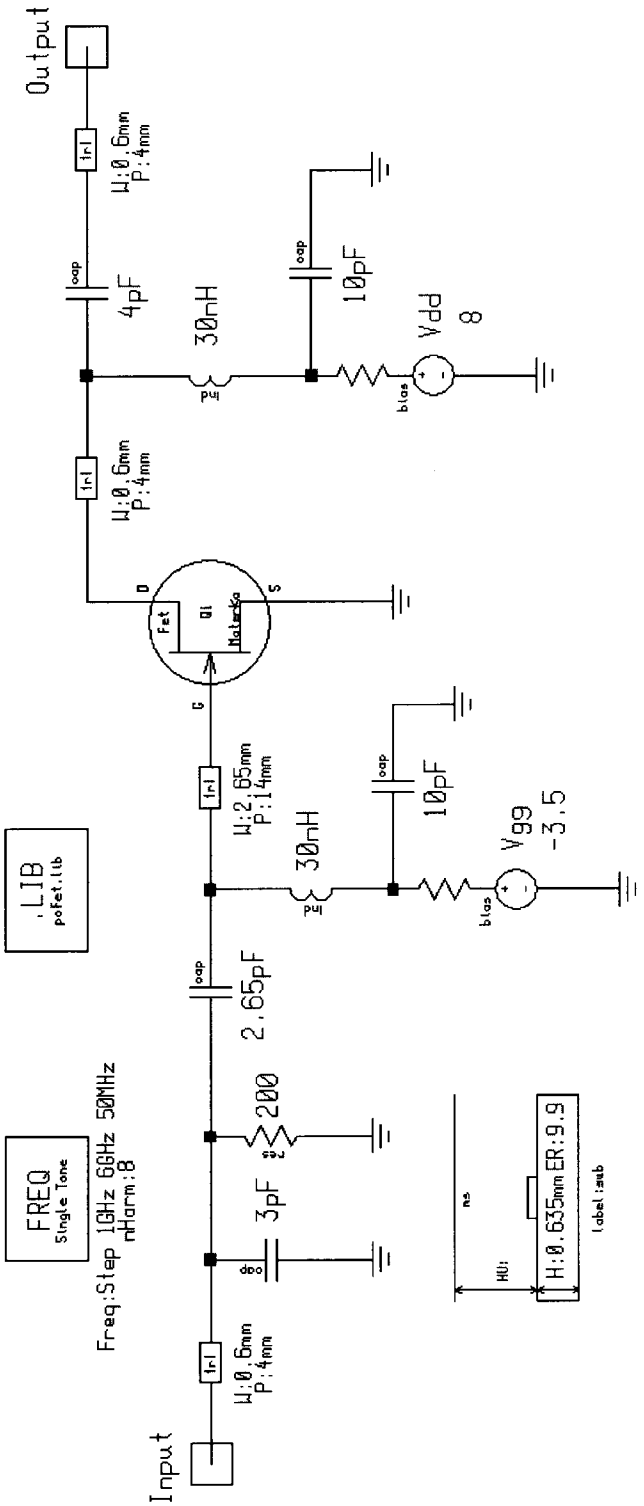


Figure 3-270 A 2.2-GHz amplifier design for stability analysis.

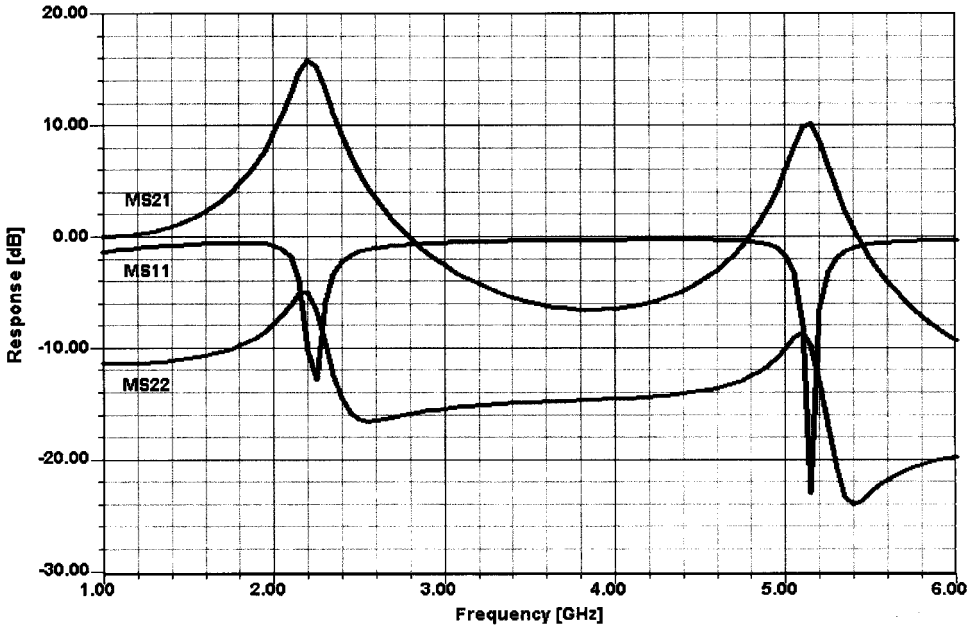


Figure 3-271 Frequency-dependent gain and matching characteristics of the amplifier. Although the circuit is designed for 2.2-GHz operation, the periodic responses of its distributed-element-based input and output matching networks result in a second gain peak near 5.1 GHz. As we shall see, however, instabilities may result at frequencies that cannot be intuited from this graph.

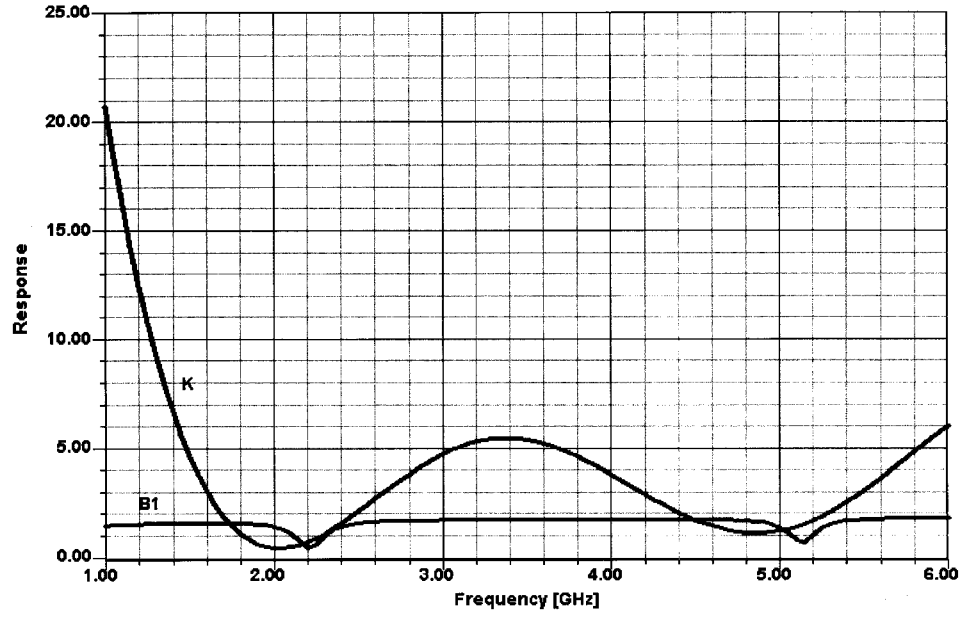


Figure 3-272 K and B_1 for the 2.2-GHz amplifier.

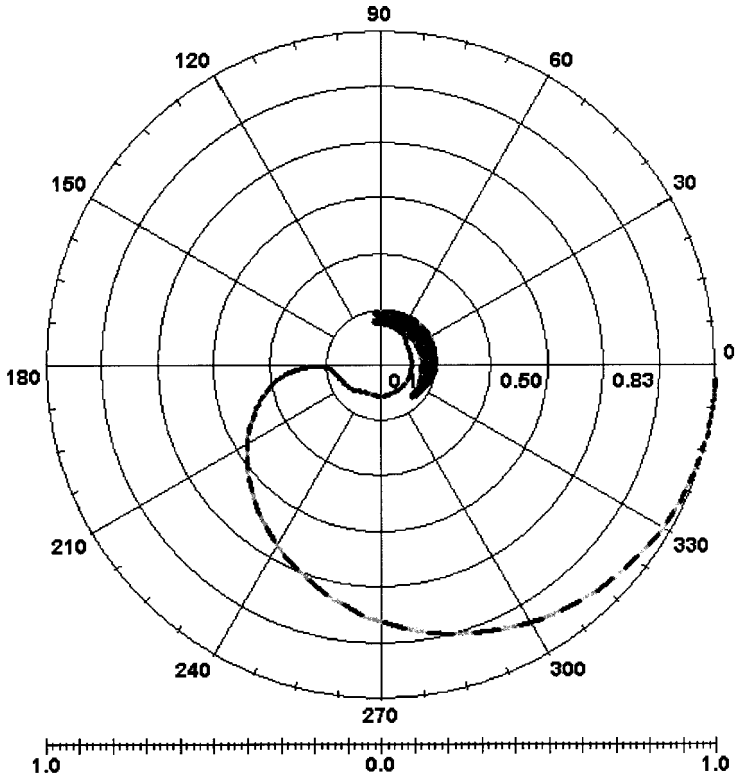


Figure 3-273 Polar Nyquist plot for the 2.2-GHz amplifier covering 1 kHz to 20 GHz.

- To travel in a clockwise direction on a polar chart means to decrease phase angle. On a magnitude–angle plot, the phase will become more negative as frequency is increased (although it may not be monotonic).
- If the path crosses the negative real axis on the polar chart, it will cross –180° on the magnitude–angle plot.
- Encirclement will occur if the phase angle continues to decrease to, say, –360°. If "unraveling" occurs, the phase angle will not stay below –360°. Figure 3-274 clearly shows that the phase angle decreases to about –180° but does not decrease any more. Rather, the angle returns to a positive value and oscillates between about –40° and +100°. Therefore an encirclement did not occur, and the circuit is stable.

An Unstable Case. Let us now change the gate bias to –2 V—a value that increases the amplifier gain enough to cause instability. Figures 3-275 and 3-276 show the results. From Figure 3-275 we can see that the origin is encircled in a clockwise direction. Figure 3-276 confirms this since the phase angle goes through –180° and continues without unraveling.

Approximate Frequency of Oscillation. The frequency where the Nyquist plot crosses –180° is often the approximate frequency of the unstable right-half-plane zeros and provides an estimate of the potential oscillation. For this example, there is a sharp crossing at 200 MHz that is characteristic of a circuit resonance. Note that the frequency is not related

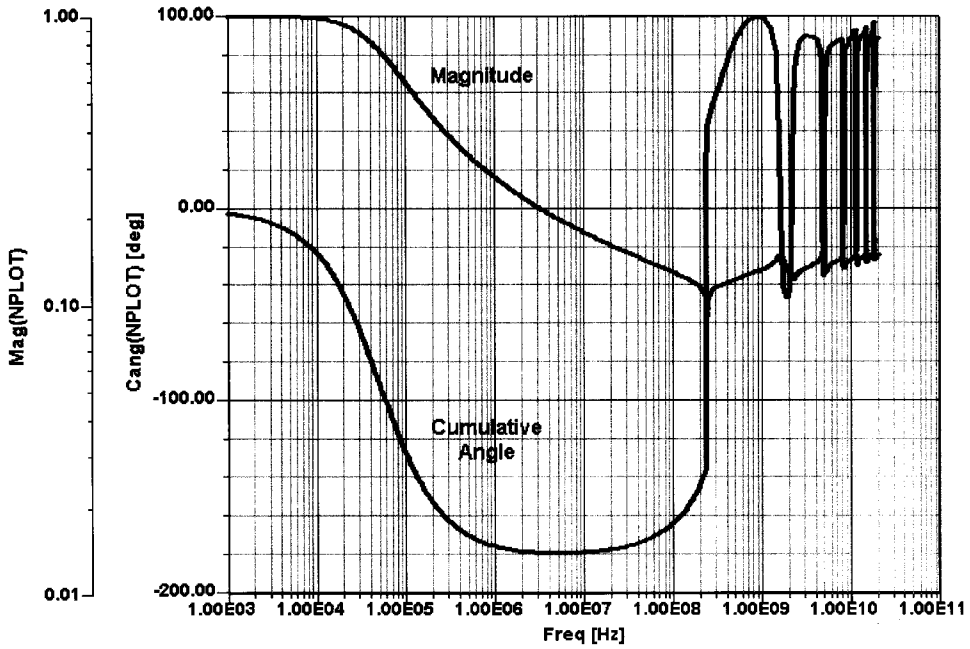


Figure 3-274 Magnitude and cumulative angle Nyquist plot for the 2.2-GHz amplifier covering 1 kHz to 20 GHz.

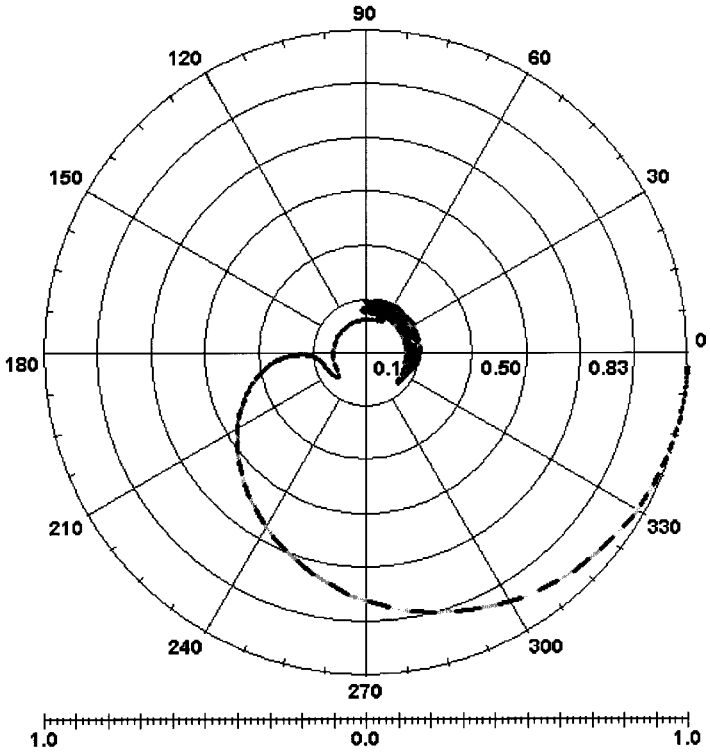


Figure 3-275 Polar Nyquist plot of the unstable amplifier.

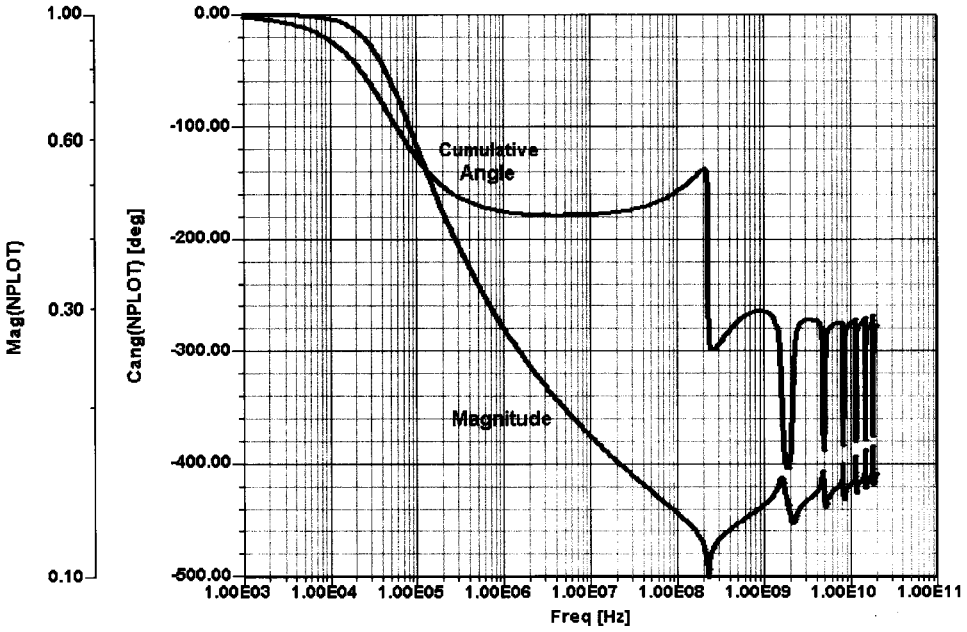


Figure 3-276 Magnitude–cumulative angle Nyquist plot of the unstable amplifier.

to a resonant frequency in a rigorous mathematical fashion, but it has been our experience that this frequency approximates the resonant frequency. To confirm the circuit's readiness for oscillation in the vicinity of this frequency, we evaluate it with Serenade's Oscillator Design Aid from 10 MHz to 1 GHz. The Design Aid finds an initial resonant frequency near 230.6 MHz—not too far from the –180° crossing frequency of 200 MHz of the Nyquist plot. Figure 3-277 shows this point.

Oscillator analysis finds a final oscillation frequency of 228.48 MHz. Figure 3-278 shows the resulting output spectrum at Port 2. At this point we decided to do a quick jump into Chapter 5 and take advantage of the CAD capability to predict the resulting phase noise, which is shown in Figure 3-279. As we compare this with other measured results, the approximately 228-MHz oscillator has a poor phase noise since the same phase noise is achievable at 1 GHz and higher. We remind the reader that this is an unwanted effect, but all of us who have built amplifiers will have experienced either low-frequency oscillations (motorboating) or very-high-frequency oscillation, sometimes caused by the components in the immediate vicinity of the transistor itself and found somewhere above 3 GHz. This case is really a demonstration case; a closer look reveals that the transconductance required for this condition is unrealistically high; but nonetheless it is a good example to demonstrate the stability issue.

Where Does Oscillation Begin? An interesting question to solve now is, "What gate bias voltage is needed for oscillation to begin?" Ansoft's Serenade Design Environment includes oscillator synchronous stability analysis, which uses a modification of the harmonic-balance technique to determine critical points, such as the voltage needed to "turn on" an oscillator. Mathematically, this is called a *Hopf bifurcation* because the behavior of

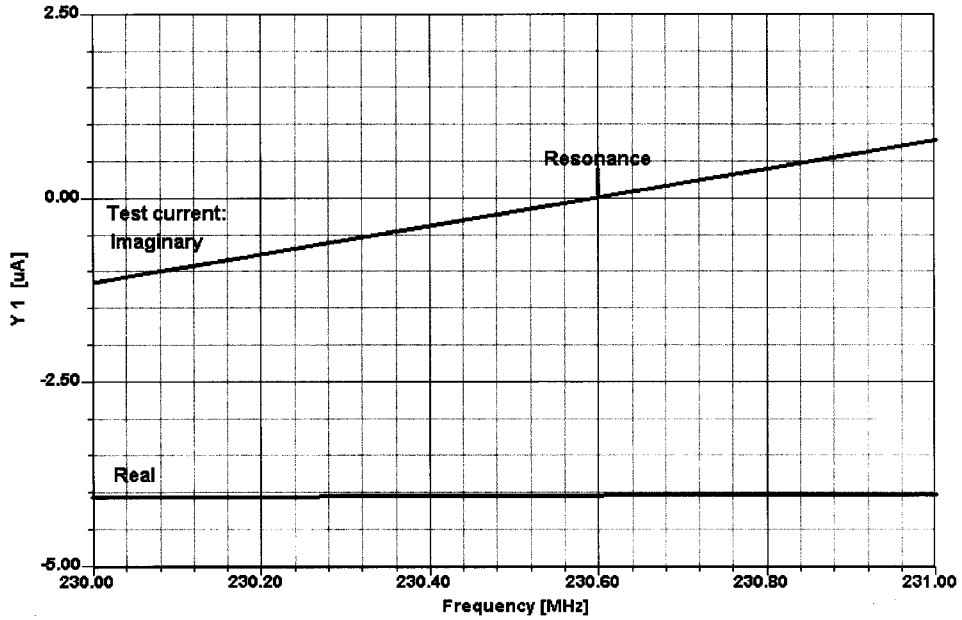


Figure 3-277 Graph from Serenade Oscillator Design Aid analysis of the amplifier. A resonant frequency is indicated where the imaginary part of the test current equals zero and the real part is negative.

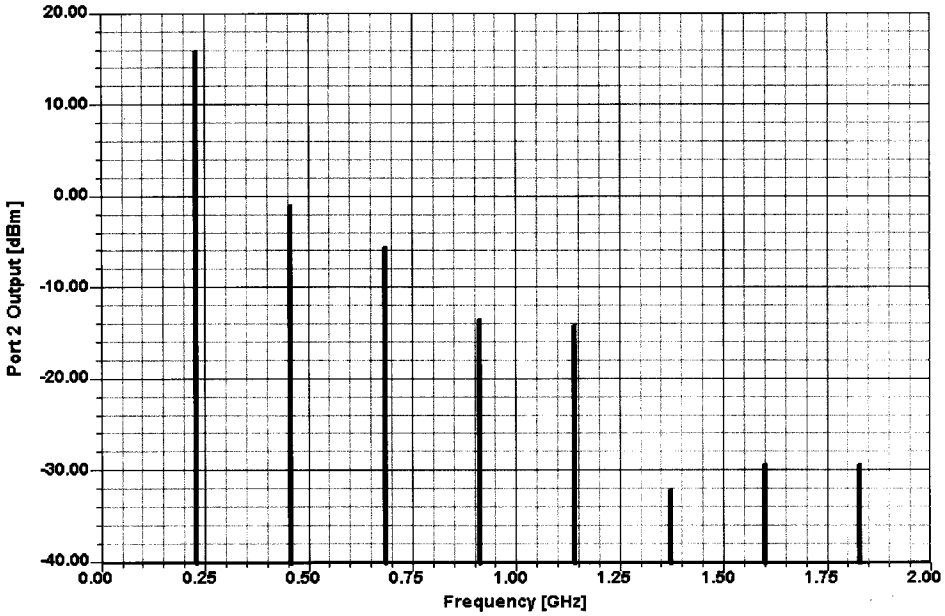


Figure 3-278 Predicted output spectrum of the unintended oscillator.

the circuit changes dramatically from a dc steady state to an oscillatory response. Figure 3-280 shows this effect by plotting output power versus gate bias. The turning point at –3.2 V (22 mW) is clearly visible, as is the bifurcation at –3.1 V (0 mW output).

Figure 3-280 can be interpreted as follows:

1. If the bias voltage is increased from, say, –4.0 V, the circuit will begin to oscillate at the bifurcation voltage and the output power will jump to 26 mW.
2. If the circuit is oscillating and the voltage is decreased from –2.0 V, the circuit will cease oscillating at the turning voltage of –3.2 V.
3. The branch between the turning point and the bifurcation point is unstable and cannot be realized in practice. This forms a hysteresis loop around points (–3.1 V, 0 mW), (–3.1 V, 26 mW), (–3.2 V, 22 mW), and (–3.2 V, 0 mW).

This amplifier stability analysis section is based on the stability analysis examples provided by Ansoft for the Serenade Design Environment.

An interesting example that combines both CAD and the feed-forward amplifier technique can be found in the Serenade Design Environment system simulator examples manual. This technique has been under evaluation for the last two years, but our feeling is that not enough measured data at frequencies above 500 MHz with reliable performance information have yet been made available in the literature. Along these lines, work done at Ansoft under a U.S. Air Force contract has provided very good insight into high-efficiency amplifiers at microwave frequencies [46].

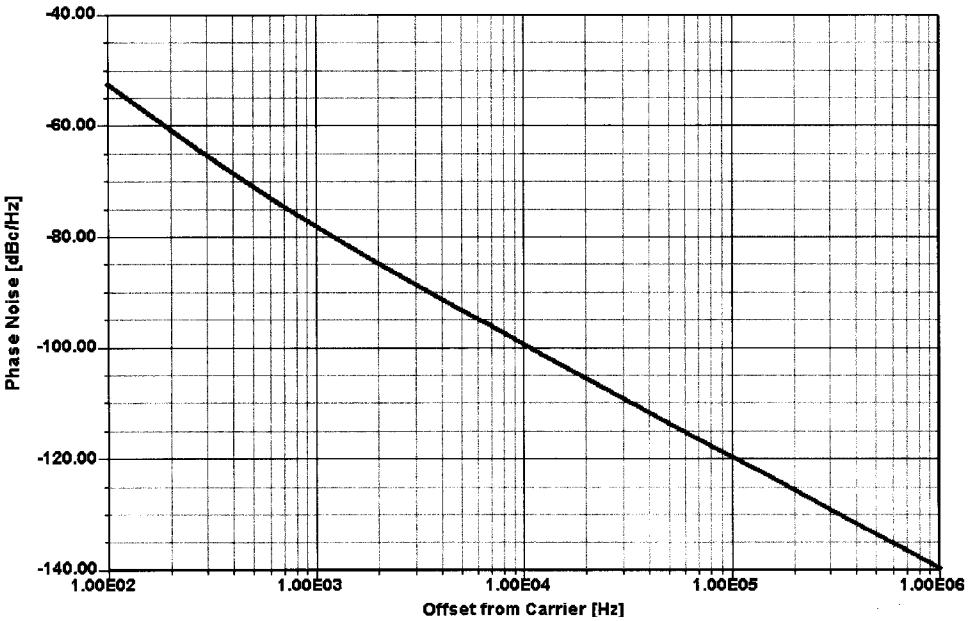


Figure 3-279 Predicted phase noise of the oscillating amplifier.

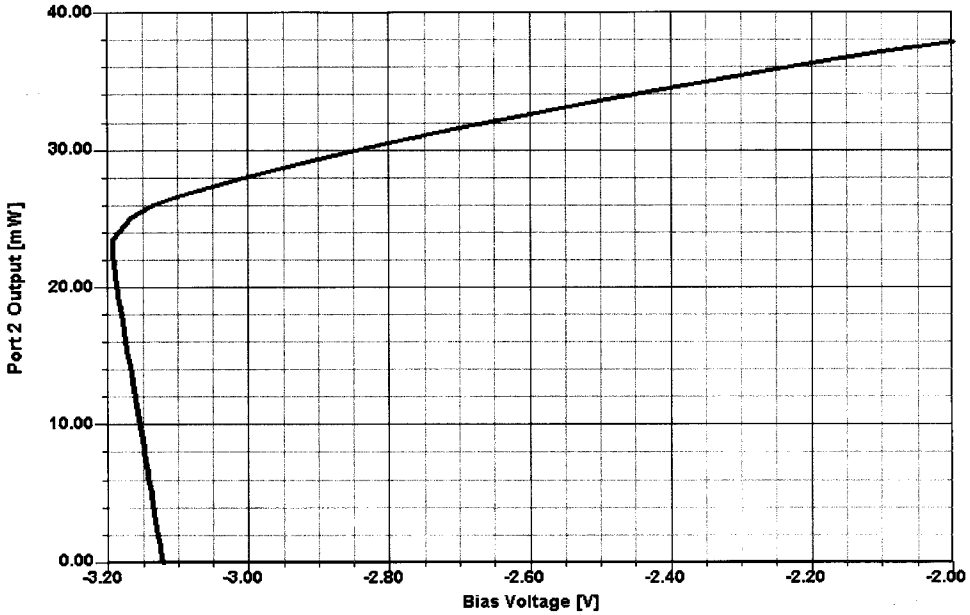


Figure 3-280 Output power versus gate bias for the unstable amplifier.

3-13 POWER AMPLIFIER DATASHEETS AND MANUFACTURER-RECOMMENDED APPLICATIONS

Having discussed the various aspects of BJTs and FETs, we have decided to use MOS examples from Ericsson and Motorola, and a bipolar example from Motorola, as good examples to show the technology required to build high-power transistor amplifiers in this frequency range. In the following pages, we reproduce, with permission, datasheets for Ericsson's PTF 10009 FET and Motorola's MRF186 FET, followed by Motorola's MRF899 BJT.

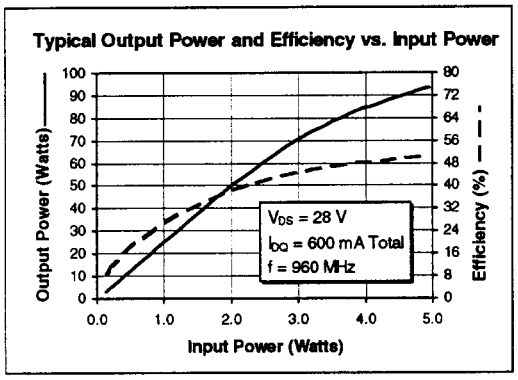
ERICSSON

PTF 10009
85 Watts, 1.0 GHz
LDMOS Field Effect Transistor

Description

The 10009 is a common source N-channel enhancement-mode lateral MOSFET intended for large signal amplifier applications to 1.0 GHz. It is rated at 85 watts minimum output power. Nitride surface passivation and gold metallization are used to ensure excellent device lifetime and reliability. 100% lot traceability is standard.

- Performance at 960 MHz, 28 Volts
  - Output Power = 85 Watts
  - Power Gain = 13.0 dB Typ
  - Efficiency = 50% Typ
- Gold Metallized
- Silicon Nitride Passivated
- Excellent Thermal Stability



Package 20230

Maximum Ratings

Parameter	Symbol	Value	Unit
Drain-Source Voltage	V_{DSS}	65	Vdc
Gate-Source Voltage	V_{GS}	±20	Vdc
Operating Junction Temperature	T_J	200	°C
Total Device Dissipation at T_{flange} = 25°C Above 25°C derate by	P_D	270 1.54	Watts W/°C
Storage Temperature Range	T_{STG}	-65 to 150	°C
Thermal Resistance (T_{flange} = 70°C)	$R_{\theta JC}$	0.65	°C/W

PTF 10009

ERICSSON

Electrical Characteristics (100% Tested)

Characteristic (per side)	Conditions	Symbol	Min	Typ	Max	Units
Drain-Source Breakdown Voltage	V_{GS} = 0 V, I_D = 25 mA	$V_{(BR)DSS}$	65	—	—	Volts
Drain-Source Leakage Current	V_{DS} = 28 V, V_{GS} = 0 V	I_{DSS}	—	—	1.0	mA
Gate Threshold Voltage	V_{DS} = 10 V, I_D = 75 mA	$V_{GS(th)}$	—	2.0	—	Volts
Forward Transconductance	V_{DS} = 10 V, I_D = 3 A	g_{fs}	—	2.8	—	Siemens

Dynamic Characteristics

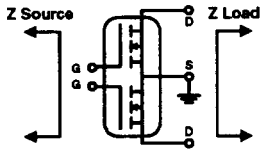
Characteristic (per side)	Symbol	Min	Typ	Max	Units
Input Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{iss}	—	90	—	pF
Output Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{oss}	—	36	—	pF
Reverse Transfer Capacitance (V_{DS} = 28 V, V_{GS} = 0 V, f = 1 MHz)	C_{rss}	—	1.9	—	pF

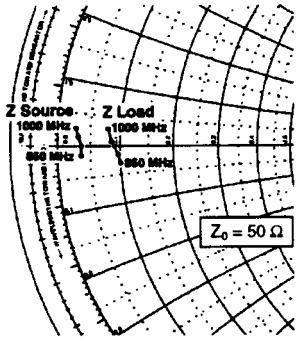
RF Specifications (100% Tested)

Characteristic	Symbol	Min	Typ	Max	Units
Gain (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz)	G_{ps}	12.0	13.0	—	dB
Drain Efficiency (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz)	η	47	50	—	%
Load Mismatch Tolerance (V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA, f = 960 MHz— all phase angles at frequency of test)	Ψ	—	—	5:1	—

Impedance Data (data shown for fixed-tuned broadband circuit)

(V_{DD} = 28 V, P_{out} = 85 W, I_{DQ} = 600 mA)



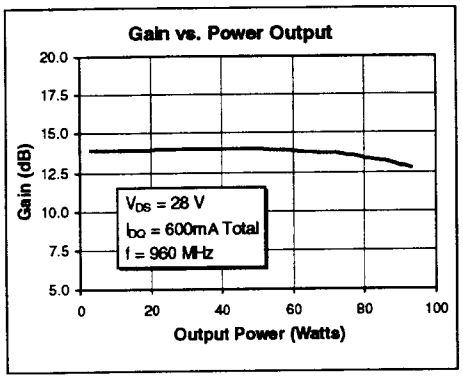


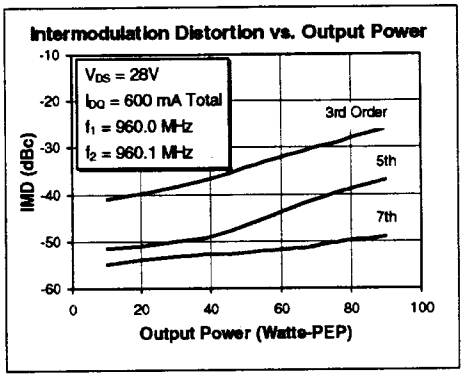
Frequency	Z Source		Z Load	
MHz	R	jX	R	jX
860	1.76	-0.78	5.00	-1.50
900	1.80	-0.05	4.80	-0.78
960	1.58	0.69	4.24	0.36
1000	1.39	1.35	3.95	1.41

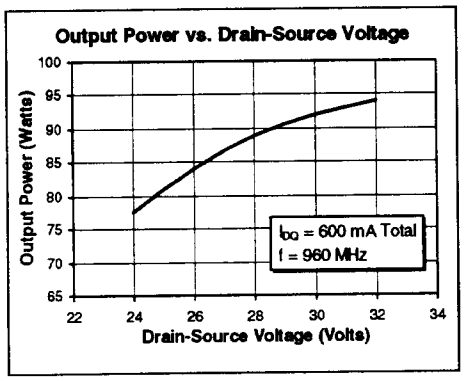
ERICSSON

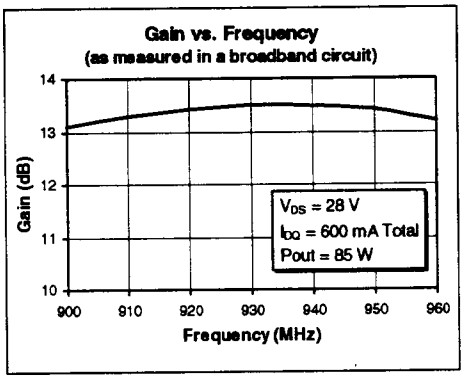
PTF 10009

Typical Performance





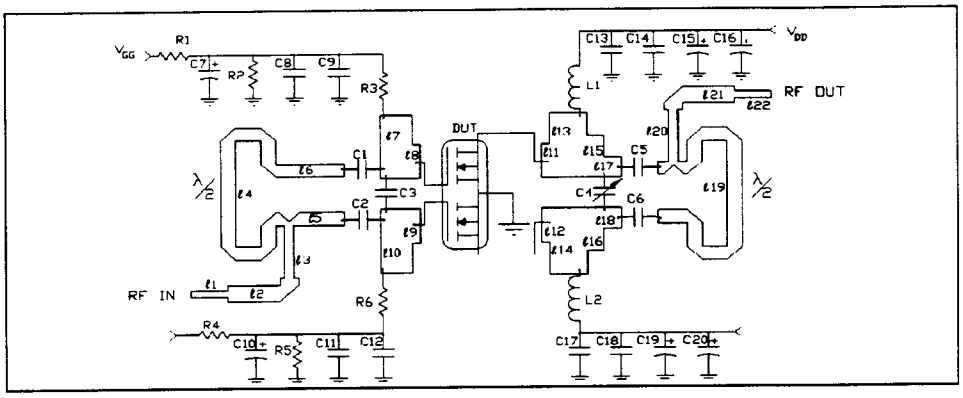




PTF 10009

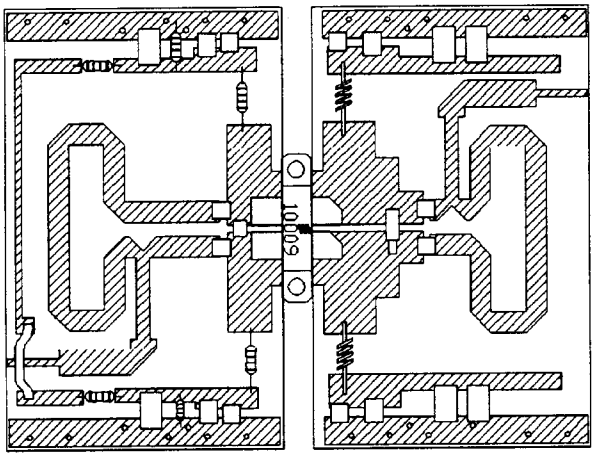
ERICSSON

Test Circuit



Schematic for f = 960 MHz

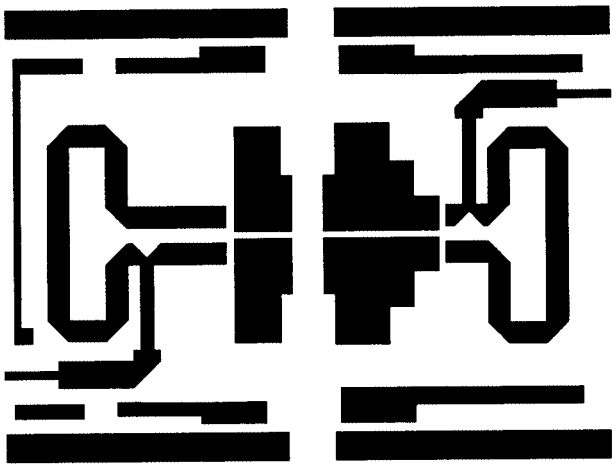
DUT	10009
C1-2, C5-6, C9, C12-13, C17	33 pF, Capacitor ATC 100 B
C3	11 pF, Capacitor ATC 100 B
C4	6.0 pF, Variable Capacitor, JMC 5701
C7, C10	10 µF, +10 V Electrolytic Capacitor
C8, C11, C14, C18	0.01 µF, Capacitor ATC 100 B
C15, C16, C19, C20	10 µF, +30 V Electrolytic Capacitor
L1. L2	4 Turn, #20 AWG, .120" I.D.
R1, R2, R4, R5	1.0 K, Ω Resistor
R3, R6	5.1 K, 1/4 Ω Resistor
ℓ1, ℓ22	50 Ω, .030 λ
ℓ2, ℓ21	20 Ω, .080 λ
ℓ3, ℓ20	32 Ω, .191 λ
ℓ4, ℓ19	25 Ω, .500 λ
ℓ5, ℓ6	25 Ω, .091 λ
ℓ7, ℓ10	7 Ω, .056 λ
ℓ8, ℓ9	13.0 Ω, .017 λ
ℓ11, ℓ12	13.0 Ω, .017 λ
ℓ13, ℓ14	7.0 Ω, .064 λ
ℓ15, ℓ16	10.0 Ω, .029 λ
ℓ17, ℓ18	19.0 Ω, .028 λ
Circuit Board	.031" Thick, ε_r = 4.0, AlliedSignal, G200



Parts Layout (not to scale)

ERICSSON

PTF 10009



Artwork (1 inch ——— *)*

Ericsson Components
RF Power Products
675 Jarvis Drive
Morgan Hill, CA 95037 USA
Telephone: 408-778-9434

1-877-GOLDMOS
(1-877-465-3667)
e-mail: rfpower@ericsson.com
www.ericsson.com/rfpower

Specifications subject to change
without notice.
L1
© 1997 Ericsson Inc.
EUS/KR 1301-PTF 10009 Uen Rev. B 09-29-98

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF186/D

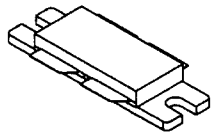
The RF MOSFET Line
RF Power Field-Effect Transistor
N–Channel Enhancement–Mode Lateral MOSFET

MRF186

120 W, 1.0 GHz
28 VOLTS
LATERAL N–CHANNEL
BROADBAND
RF POWER MOSFET

Designed for broadband commercial and industrial applications at frequencies from 800 MHz to 1.0 GHz. The high gain and broadband performance of this device makes it ideal for large–signal, common source amplifier applications in 28 volt base station equipment.

- Guaranteed Performance @ 960 MHz, 28 Volts
  Output Power — 120 Watts (PEP)
  Power Gain — 11 dB
  Efficiency — 30%
  Intermodulation Distortion — –28 dBc
- Excellent Thermal Stability
- 100% Tested for Load Mismatch Stress at all Phase Angles with 5:1 VSWR @ 28 Vdc, 960 MHz, 120 Watts CW



CASE 375B–02, STYLE 2

MAXIMUM RATINGS (2)

Rating	Symbol	Value	Unit
Drain–Source Voltage	V_{DSS}	65	Vdc
Drain–Gate Voltage (R_{GS} = 1 MΩ)	V_{DGR}	65	Vdc
Gate–Source Voltage	V_{GS}	±20	Vdc
Drain Current — Continuous	I_D	14	Adc
Total Device Dissipation @ T_C = 70°C Derate above 70°C	P_D	162.5 1.25	Watts W/°C
Storage Temperature Range	T_{stg}	– 65 to +150	°C
Operating Junction Temperature	T_J	200	°C

THERMAL CHARACTERISTICS (2)

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	0.8	°C/W

NOTE – **CAUTION** – MOS devices are susceptible to damage from electrostatic charge. Reasonable precautions in handling and packaging MOS devices should be observed.

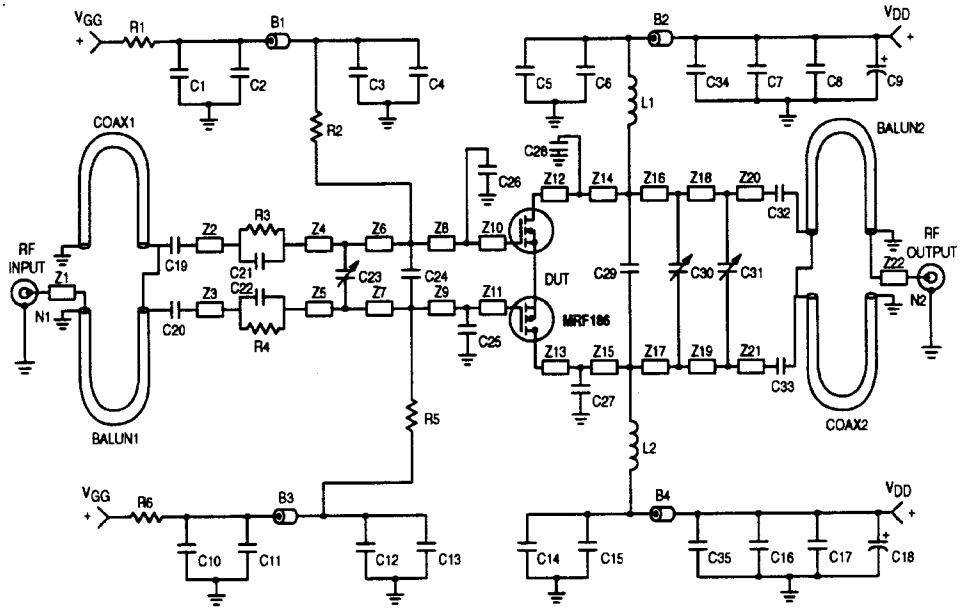
© Motorola, Inc. 1999

MOTOROLA

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted)

Characteristic	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS (1)					
Drain–Source Breakdown Voltage (V_{GS} = 0 Vdc, I_D = 50 μAdc)	$V_{(BR)DSS}$	65	—	—	Vdc
Zero Gate Voltage Drain Current (V_{DS} = 28 Vdc, V_{GS} = 0)	I_{DSS}	—	—	1	μAdc
Gate–Source Leakage Current (V_{GS} = 20 Vdc, V_{DS} = 0)	I_{GSS}	—	—	1	μAdc
ON CHARACTERISTICS (1)					
Gate Quiescent Voltage (V_{DS} = 26 Vdc, I_D = 300 μAdc Per Side)	$V_{GS(th)}$	2.5	3	4	Vdc
Gate Quiescent Voltage (V_{DS} = 26 Vdc, I_D = 300 mAdc Per Side)	$V_{GS(Q)}$	3.3	4.2	5	Vdc
Delta Gate Threshold Voltage (Side to Side) (V_{DS} = 28 V, I_D = 300 mA Per Side)	$\Delta V_{GS(Q)}$	—	—	0.3	Vdc
Drain–Source On–Voltage (V_{GS} = 10 Vdc, I_D = 3 Adc Per Side)	$V_{DS(on)}$	—	0.58	0.7	Vdc
Forward Transconductance (V_{DS} = 10 Vdc, I_D = 3 Adc Per Side)	g_{fs}	2.4	2.8	—	S
DYNAMIC CHARACTERISTICS (1)					
Input Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{iss}	—	177	—	pF
Output Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{oss}	—	45	—	pF
Reverse Transfer Capacitance (Per Side) (V_{DS} = 28 Vdc, V_{GS} = 0, f = 1 MHz)	C_{rss}	—	3.4	—	pF
FUNCTIONAL CHARACTERISTICS (In Motorola Test Fixture) (2)					
Two–Tone Common Source Amplifier Power Gain (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	G_{ps}	11	12.2	—	dB
Two–Tone Drain Efficiency (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	η	30	35	—	%
3rd Order Intermodulation Distortion (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	IMD	—	–32	–28	dBc
Input Return Loss (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 960.0 MHz, f2 = 960.1 MHz)	IRL	9	16	—	dB
Two–Tone Common Source Amplifier Power Gain (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	G_{ps}	—	12	—	dB
Two–Tone Drain Efficiency (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	η	—	33	—	%
3rd Order Intermodulation Distortion (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	IMD	—	–32	—	dBc
Input Return Loss (V_{DD} = 28 Vdc, P_{out} = 120 W PEP, I_{DQ} = 2 x 400 mA, f1 = 945.0 MHz, f2 = 945.1 MHz)	IRL	—	16	—	dB
Output Mismatch Stress (V_{DD} = 28 Vdc, P_{out} = 120 W CW, I_{DQ} = 2 x 400 mA, f = 960 MHz, VSWR = 5:1, All Phase Angles at Frequency of Tests)	Ψ	No Degradation In Output Power Before and After Test			

(1) Each side of device measured separately.
(2) Device measured in push–pull configuration.



B1 – B4	Fair Rite Products Short Ferrit Bead, 2743021446	C31	0.8 – 8.0 pF, Variable Capacitor, Johanson Gigatrim
C1, C7, C8, C10, C16, C17	10 µF, 50 V, Tantalum	L1, L2	3 Turns, #20 AWG, IDIA 0.126", 24.7 nH
C2, C11, C34, C35	0.1 µF, Chip Capacitor	N1, N2	Type N Connectors
C3, C6, C12, C15	330 pF, Chip Capacitor	R1, R6	1 kΩ, 1/4 W, Carbon Resistor
C4, C5, C13, C14, C19, C20, C32, C33	47 pF, Chip Capacitor	R2, R5	1.2 kΩ, 0.1 W, Chip Resistor
C9, C18	250 µF, 50 V, Electrolytic Capacitor	R3, R4	75 Ω, 0.1 W, Chip Resistor
C21, C22	12 pF, Chip Capacitor	Z1 – Z22	Microstrip (See Component Placement)
C23, C30	0.6 – 4.5 pF, Variable Capacitor, Johanson Gigatrim	Balun1, Balun2, Coax1, Coax2	2.20" 50 Ω, 0.086" OD Semi–Rigid Coax
C24, C25, C26	5.1 pF, Chip Capacitor	Board	1/32" Glass Teflon®, ε_r = 2.55
C27, C28	3.9 pF, Chip Capacitor		

Figure 1. 930 – 960 MHz Test Circuit Schematic

TYPICAL CHARACTERISTICS

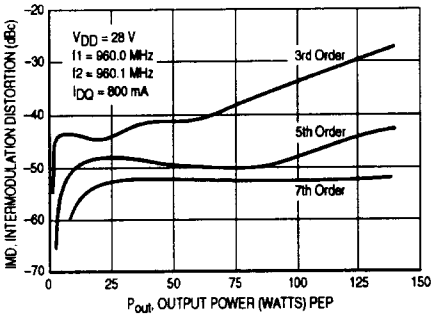


Figure 2. Intermodulation Distortion versus Output Power

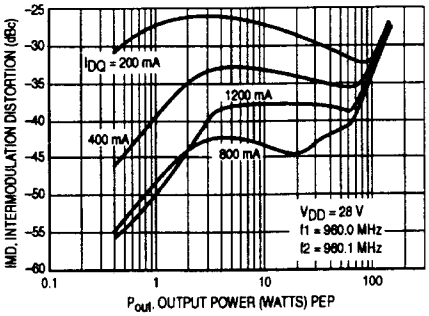


Figure 3. Intermodulation Distortion versus Output Power

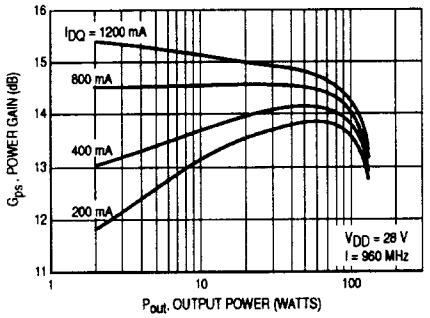


Figure 4. Power Gain versus Output Power

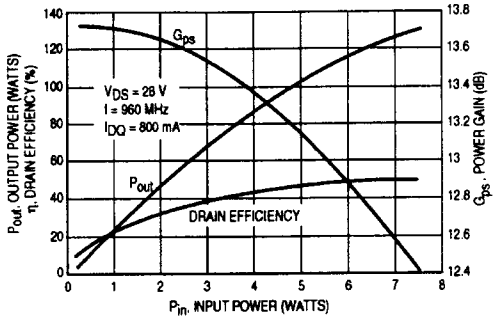


Figure 5. Output Power versus Input Power

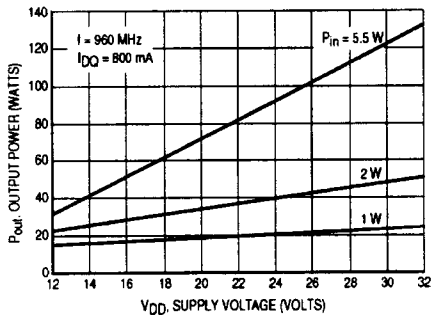


Figure 6. Output Power versus Supply Voltage

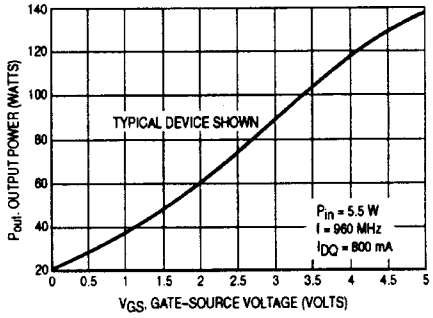


Figure 7. Output Power versus Gate Voltage

TYPICAL CHARACTERISTICS

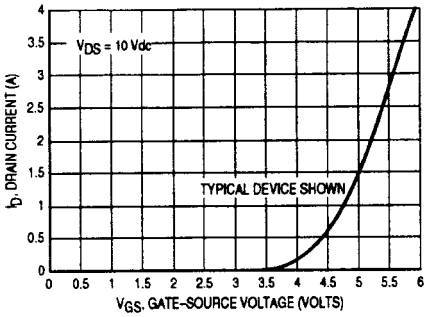


Figure 8. Drain Current versus Gate Voltage

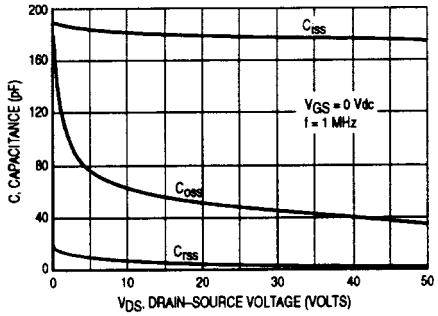


Figure 9. Capacitance versus Voltage

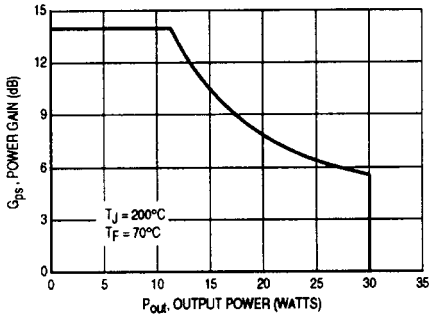


Figure 10. DC Safe Operating Area

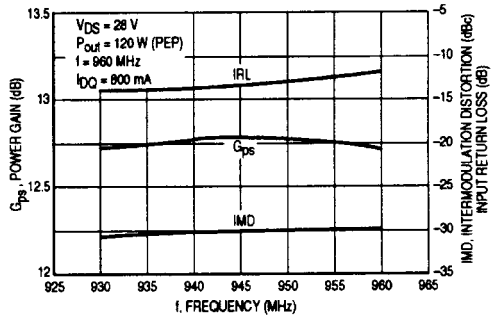
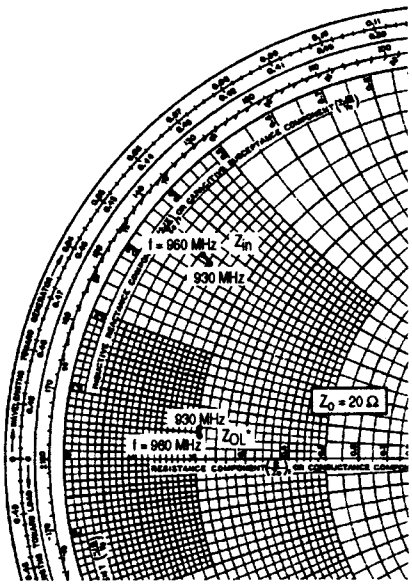


Figure 11. Broadband Circuit Performance



V_{CC} = 28 V, I_{DQ} = 2 × 400 mA, P_{out} = 120 Watts (PEP)

f MHz	Z_{in} Ω	Z_{OL}^* Ω
930	2.5 + j6.9	4.3 + j1.2
945	2.5 + j7.0	4.3 + j1.0
960	2.2 + j7.1	4.3 + j0.9

Z_{in} = Conjugate of impedance.

Z_{OL}^* = Conjugate of the optimum load impedance at a given output power, voltage, IMD, bias current, efficiency and frequency.

Note: Z_{OL}^* was chosen based on tradeoffs between gain, output power, drain efficiency and intermodulation performance. Impedances shown represent a single channel (1/2 of MRF186) impedance measurement.

Figure 12. Series Equivalent Input and Output Impedance

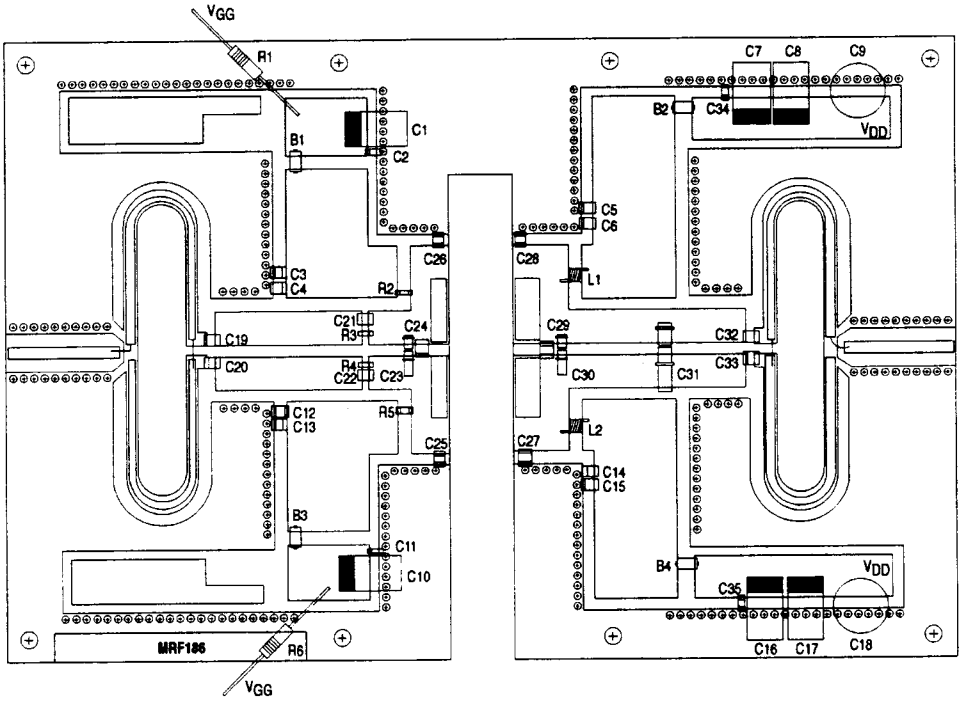
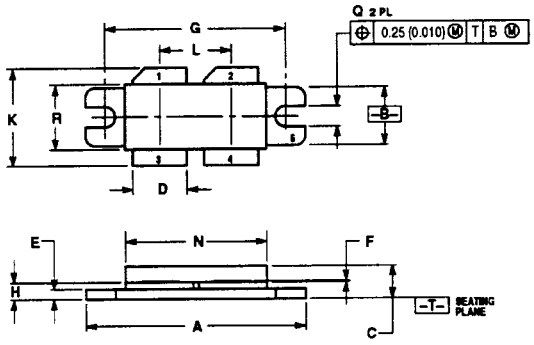


Figure 13. Component Placement Diagram of 930 – 960 MHz Broadband Test Fixture

PACKAGE DIMENSIONS



NOTES:
1. DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2. CONTROLLING DIMENSION: INCH.

DIM	INCHES MIN	INCHES MAX	MILLIMETERS MIN	MILLIMETERS MAX
A	1.330	1.350	33.78	34.29
B	0.375	0.395	9.52	10.03
C	0.180	0.205	4.57	5.21
D	0.320	0.340	8.13	8.64
E	0.060	0.070	1.52	1.77
F	0.004	0.006	0.11	0.15
G	1.100 BSC		27.94 BSC	
H	0.082	0.097	2.09	2.46
K	0.580	0.620	14.73	15.75
L	0.435 BSC		11.05 BSC	
N	0.845	0.875	21.46	22.23
Q	0.118	0.130	3.00	3.30
R	0.390	0.410	9.91	10.41

STYLE 1:
PIN 1. COLLECTOR
2. COLLECTOR
3. BASE
4. BASE
5. EMITTER

CASE 375A–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Motorola data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and Ⓜ are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

Mfax is a trademark of Motorola, Inc.

How to reach us:
USA/EUROPE/Locations Not Listed: Motorola Literature Distribution; P.O. Box 5405, Denver, Colorado 80217. 303–675–2140 or 1–800–441–2447

Mfax™: RMFAX0@email.sps.mot.com – TOUCHTONE 602–244–6609 – US & Canada ONLY 1–800–774–1848

INTERNET: http://motorola.com/sps

JAPAN: Nippon Motorola Ltd.: SPD, Strategic Planning Office, 4–32–1, Nishi–Gotanda, Shinagawa–ku, Tokyo 141, Japan. 81–3–5487–8488

ASIA/PACIFIC: Motorola Semiconductors H.K. Ltd.; 8B Tai Ping Industrial Park, 51 Ting Kok Road, Tai Po, N.T., Hong Kong. 852–26629298

MOTOROLA
SEMICONDUCTOR TECHNICAL DATA

Order this document
by MRF899/D

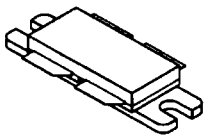
The RF Line
NPN Silicon RF Power Transistor

MRF899

150 W, 900 MHz
RF POWER
TRANSISTOR
NPN SILICON

Designed for 26 Volt UHF large–signal, common emitter, Class AB linear amplifier applications in industrial and commercial FM/AM equipment operating in the range 800–960 MHz.

- Specified 26 Volt, 900 MHz Characteristics
  Output Power = 150 Watts (PEP)
  Minimum Gain = 8.0 dB @ 900 MHz, Class AB
  Minimum Efficiency = 35% @ 900 MHz, 150 Watts (PEP)
  Maximum Intermodulation Distortion –28 dBc @ 150 Watts (PEP)
- Characterized with Series Equivalent Large–Signal Parameters from 800 to 960 MHz
- Silicon Nitride Passivated
- 100% Tested for Load Mismatch Stress at all Phase Angles with 5:1 VSWR @ 26 Vdc, and Rated Output Power
- Gold Metallized, Emitter Ballasted for Long Life and Resistance to Metal Migration
- Circuit board photomaster available upon request by contacting RF Tactical Marketing in Phoenix, AZ.



CASE 375A–01, STYLE 1

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Collector–Emitter Voltage	V_{CEO}	28	Vdc
Collector–Emitter Voltage	V_{CES}	60	Vdc
Emitter–Base Voltage	V_{EBO}	4.0	Vdc
Collector–Current — Continuous	I_C	25	Adc
Total Device Dissipation @ T_C = 25°C Derate above 25°C	P_D	230 1.33	Watts W/°C
Storage Temperature Range	T_{stg}	–65 to +150	°C

THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Case	$R_{\theta JC}$	0.75	°C/W

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
OFF CHARACTERISTICS					
Collector–Emitter Breakdown Voltage (I_C = 100 mAdc, I_B = 0)	$V_{(BR)CEO}$	28	37	—	Vdc
Collector–Emitter Breakdown Voltage (I_C = 50 mAdc, V_{BE} = 0)	$V_{(BR)CES}$	60	85	—	Vdc
Emitter–Base Breakdown Voltage (I_E = 10 mAdc, I_C = 0)	$V_{(BR)EBO}$	4.0	4.9	—	Vdc
Collector Cutoff Current (V_{CE} = 30 Vdc, V_{BE} = 0)	I_{CES}	—	—	10	mAdc
ON CHARACTERISTICS					
DC Current Gain (I_{CE} = 1.0 Adc, V_{CE} = 5.0 Vdc)	h_{FE}	30	75	120	—
DYNAMIC CHARACTERISTICS					
Output Capacitance (V_{CB} = 26 Vdc, I_E = 0, f = 1.0 MHz) (1)	C_{ob}	—	75	—	pF

(1) For information only. This part is collector matched.

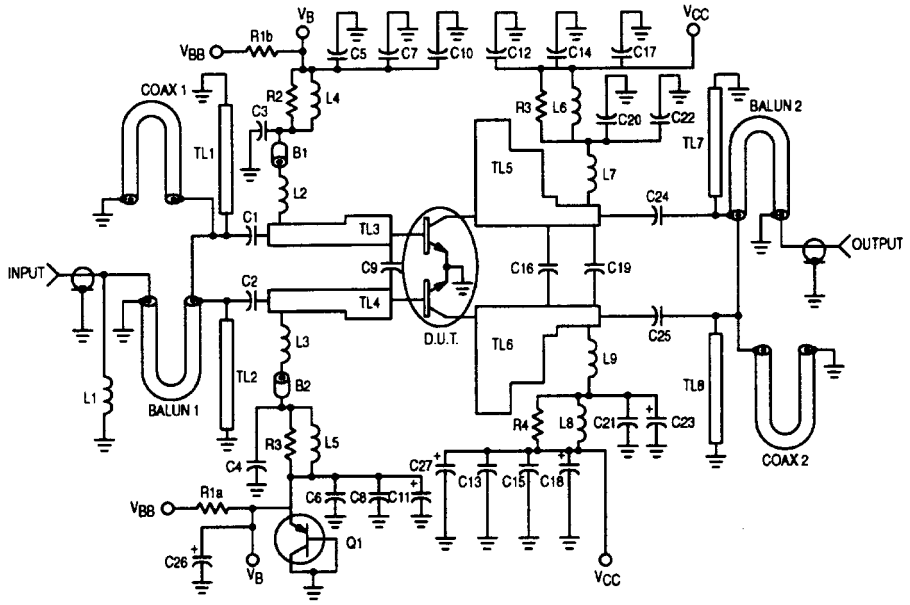
(continued)

© Motorola, Inc. 1997

MOTOROLA

ELECTRICAL CHARACTERISTICS — continued (T_C = 25°C unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
FUNCTIONAL CHARACTERISTICS					
Common–Emitter Amplifier Power Gain V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	G_{pe}	8.0	9.0	—	dB
Collector Efficiency V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	η	30	40	—	%
3rd Order Intermodulation Distortion V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz	IMD	—	–32	–28	dBc
Output Mismatch Stress V_{CC} = 26 Vdc, P_{out} = 150 Watts (PEP), I_{cq} = 300 mA, f_1 = 900 MHz, f_2 = 900.1 MHz, VSWR = 5:1 (all phase angles)	Ψ	No Degradation in Output Power Before and After Test			



B1, B2 — Ferrite Bead, Ferroxcube #56–590–65–3B
C1, C2, C24, C25 — 43 pF, B Case, ATC Chip Capacitor
C3, C4, C20, C21 — 100 pF, B Case, ATC Chip Capacitor
C5, C6, C12, C13 — 1000 pF, B Case, ATC Chip Capacitor
C7, C8, C14, C15 — 1800 pF, AVX Chip Capacitor
C9 — 9.1 pF, A Case, ATC Chip Capacitor
C10, C11, C17, C18, C22, C23 — 10 µF, Electrolytic Capacitor Panasonic
C16 — 3.9 pF, B Case, ATC Chip Capacitor
C19 — 0.8 pF, B Case, ATC Chip Capacitor
C26 — 200 µF, Electrolytic Capacitor Mallory Sprague
C27 — 500 µF Electrolytic Capacitor

L1 — 5 Turns 24 AWG IDIA 0.059″ Choke, 19.8 nH
L2, L3, L7, L9 — 4 Turns 20 AWG IDIA 0.163″ Choke
L4, L5, L6, L8 — 12 Turns 22 AWG IDIA 0.140″ Choke
N1, N2 — Type N Flange Mount, Omni Spectra
Q1 — Bias Transistor BD136 PNP
R2, R3, R4, R5 — 4.0 x 39 Ohm 1/8 W Chips in Parallel
R1a, R1b — 56 Ohm 1.0 W
TL1–TL8 — See Photomaster
Balun1, Balun2, Coax 1, Coax 2 — 2.20″ 50 Ohm 0.086″ o.d. Semi–rigid Coax, Micro Coax
Board — 1/32″ Glass Teflon, ε_r = 2.55″ Arlon (GX–0300–55–22)

Figure 1. 900 MHz Power Gain Test Circuit

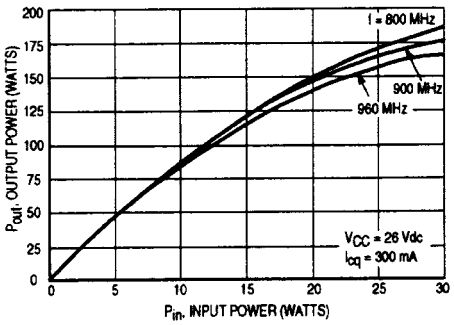


Figure 2. Output Power versus Input Power

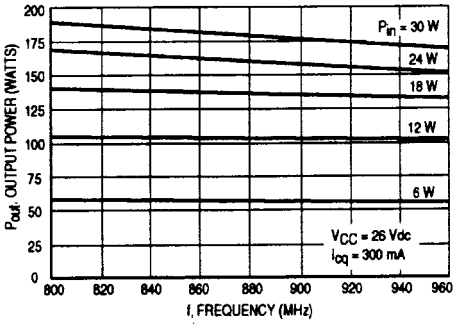


Figure 3. Output Power versus Frequency

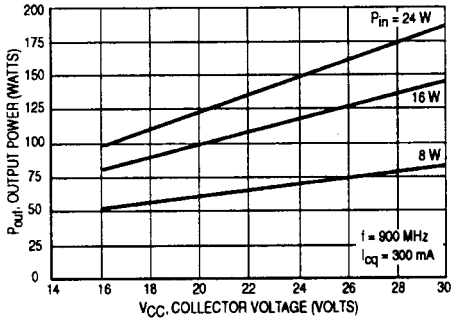


Figure 4. Output Power versus Supply Voltage

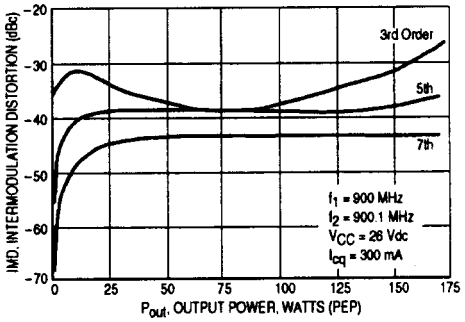


Figure 5. Intermodulation versus Output Power

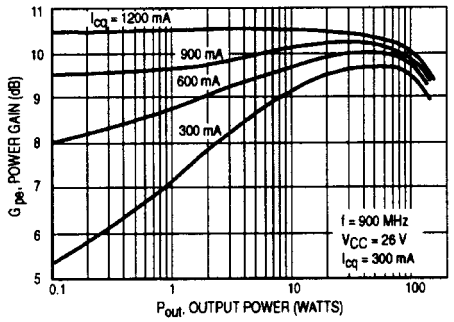


Figure 6. Power Gain versus Output Power

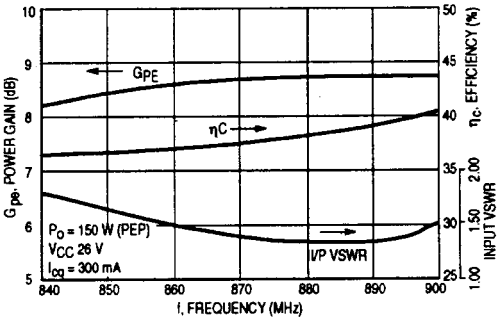
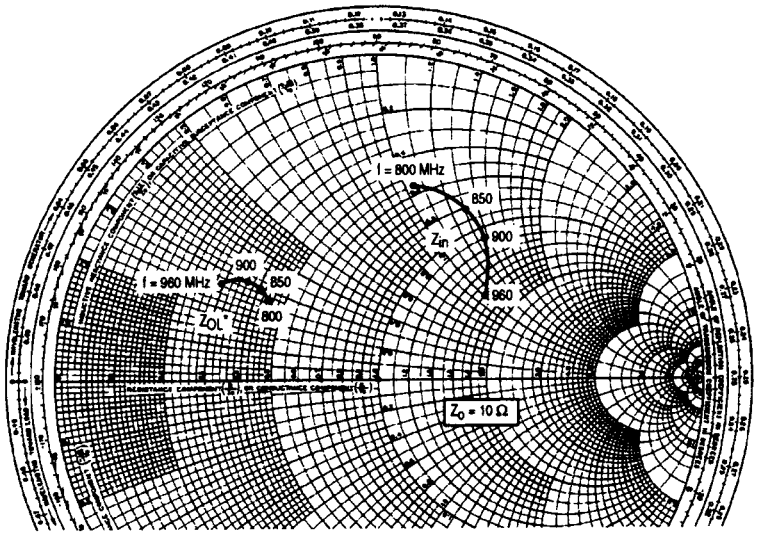


Figure 7. Broadband Test Fixture Performance



f MHz	Z_{in} Ohms	Z_{OL}^* Ohms
800	5.51 + j10.6	4.52 + j2.64
850	8.17 + j13.2	4.21 + j2.98
900	11.2 + j13.8	3.68 + j2.97
960	16.8 + j10.1	2.98 + j2.71

Z_{OL}^* = Conjugate of optimum load impedance into which the device operates at a given output power, voltage and frequency.

NOTE: Z_{in} & Z_{OL}^* are given from base–to–base and collector–to–collector respectively

Figure 8. Input and Output Impedances with Circuit Tuned for Maximum Gain @ P_O = 150 W (PEP), V_{CC} = 26 V

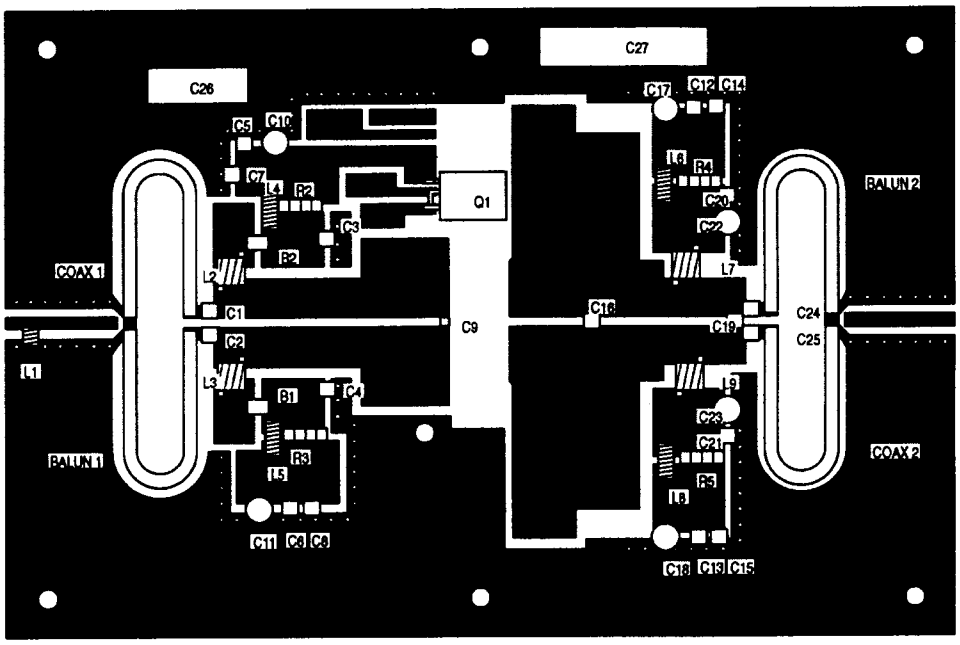
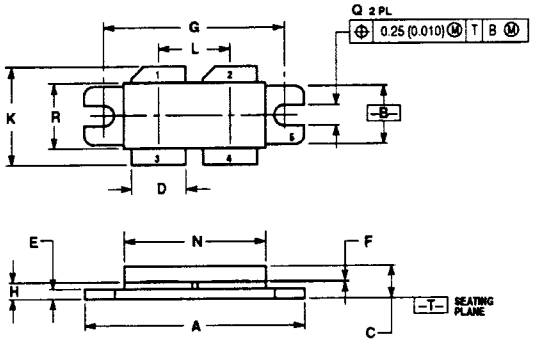


Figure 9. MRF899 Test Fixture Component Layout

PACKAGE DIMENSIONS



NOTES:
1. DIMENSIONING AND TOLERANCING PER ANSI Y14.5M, 1982.
2. CONTROLLING DIMENSION: INCH.

DIM	INCHES MIN	INCHES MAX	MILLIMETERS MIN	MILLIMETERS MAX
A	1.330	1.350	33.79	34.29
B	0.375	0.395	9.52	10.03
C	0.160	0.205	4.57	5.21
D	0.320	0.340	8.13	8.64
E	0.060	0.070	1.52	1.77
F	0.004	0.006	0.11	0.15
G	1.100 BSC		27.94 BSC	
H	0.062	0.097	2.08	2.46
K	0.580	0.620	14.73	15.75
L	0.435 BSC		11.05 BSC	
N	0.845	0.875	21.46	22.23
Q	0.116	0.130	3.00	3.30
R	0.390	0.410	9.91	10.41

STYLE 1:
PIN 1. COLLECTOR
2. COLLECTOR
3. BASE
4. BASE
5. EMITTER

CASE 375A–01
ISSUE O

Motorola reserves the right to make changes without further notice to any products herein. Motorola makes no warranty, representation or guarantee regarding the suitability of its products for any particular purpose, nor does Motorola assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters which may be provided in Motorola data sheets and/or specifications can and do vary in different applications and actual performance may vary over time. All operating parameters, including "Typicals" must be validated for each customer application by customer's technical experts. Motorola does not convey any license under its patent rights nor the rights of others. Motorola products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which the failure of the Motorola product could create a situation where personal injury or death may occur. Should Buyer purchase or use Motorola products for any such unintended or unauthorized application, Buyer shall indemnify and hold Motorola and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claim alleges that Motorola was negligent regarding the design or manufacture of the part. Motorola and M are registered trademarks of Motorola, Inc. Motorola, Inc. is an Equal Opportunity/Affirmative Action Employer.

Mfax is a trademark of Motorola, Inc.

How to reach us:
USA/EUROPE/Locations Not Listed: Motorola Literature Distribution;
P.O. Box 5405, Denver, Colorado 80217. 303–675–2140 or 1–800–441–2447

Mfax™: RMFAX0@email.sps.mot.com – TOUCHTONE 602–244–6609
– US & Canada ONLY 1–800–774–1848

INTERNET: http://motorola.com/sps

JAPAN: Nippon Motorola Ltd.: SPD, Strategic Planning Office, 4–32–1, Nishi–Gotanda, Shinagawa–ku, Tokyo 141, Japan. 81–3–5487–8488

ASIA/PACIFIC: Motorola Semiconductors H.K. Ltd.; 8B Tai Ping Industrial Park, 51 Ting Kok Road, Tai Po, N.T., Hong Kong. 852–26629298

REFERENCES

1. Ulrich L. Rohde, Anthony M. Pavio, and Robert A. Pucel, "Accurate Noise Simulation of Microwave Amplifiers Using CAD," *Microwave J.*, **31**(12), 130–141, December 1988.
2. H. T. Friis, "Noise Figures of Radio Receivers," *Proc. IRE*, **32**, 419–422, July 1944.
3. "IRE Standards on Methods of Measuring Noise in Linear Twoports," *Proc. IRE*, **48**, 60–68, January 1960.
4. H. Rothe and W. Dhalke, "Theory of Noisy Fourpoles," *Proc. IRE*, **44**, 811–818, June 1956.
5. "Representation of Noise in Linear Twoports," *Proc. IRE*, **48**, 69–74, January 1960.
6. H. A. Haus and R. B. Adler, *Circuit Theory of Linear Noisy Networks*, Cambridge, MA: MIT Press, 1959/New York: Wiley, 1959.
7. H. Fukui, "Noise Performance of Microwave Transistors," *IEEE Trans. Electron Devices*, **ED-13**, 329–341, March 1966.
8. S. Iversen, "The Effect of Feedback on Noise Figure," *Proc. IEEE*, **63**, 540–542, March 1975.
9. L. Besser, "Stability Considerations of Low-Noise Transistor Amplifiers with Simultaneous Noise and Power Match," *IEEE MTT-S International Microwave Symposium Digest*, 1975, pp. 327–329.
10. G. D. Vendelin, "Feedback Effects on the Noise Performance of GaAs Mesfets," *IEEE MTT-S International Microwave Symposium Digest*, 1975, pp. 324–326.
11. W. A. Suter, "Feedback and Parasitic Effects on Noise," *Microwave J.* **26**(2), 123–129, February 1983.
12. K. Hartmann and M. J. O. Strutt, "Changes in Four Noise Parameters Due to General Changes in Linear Two-Port Circuits," *IEEE Trans. Electron Devices*, **ED-20**(10), 874–877, October 1973.
13. H. Fukui, "Available Power Gain, Noise Figure, and Noise Measure of Two-Ports and Their Graphical Representation," *IEEE Trans. Circuit Theory*, **CT-13**, 137–142, June 1966.
14. C. R. Poole and D. K. Paul, "Optimum Noise Measure Terminations for Microwave Transistor Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**(11), 1254–1257, November 1985.
15. H. Hillbrand and P. H. Russer, "An Efficient Method for Computer Aided Noise Analysis of Linear Amplifier Networks," *IEEE Trans. Circuits Syst.*, **CAS-23**(4), 235–238, April 1976.
16. W. E. Pastori, *Topics on Noise*, Eaton Electronics Instrumentation Division (EID) Seminar Notes, Los Angeles, 1983.
17. W. E. Pastori, "Image and Second Stage Corrections Resolve Noise Figure Measurement Confusion," *Microwave Syst. News*, **13**(5), 67–86, May 1983.
18. R. Q. Lane, "The Determination of Device Noise Parameters," *Proc. IEEE*, **57**, 1461–1462, August 1969.
19. R. Q. Lane, "A Microwave Noise and Gain Parameter Test Set," *IEEE International Solid-State Circuits Conference*, 1978, pp. 172–173.
20. N. Kuhn, "Accurate and Automatic Noise Figure Measurements with Standard Equipment," *IEEE MTT-S International Microwave Symposium Digest*, 1980, pp. 425–427.
21. E. W. Strid, "Measurement of Losses in Noise-Matching Networks," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**, 247–252, March 1981.
22. Chris Bowick, *RF Circuit Design*. Indianapolis: Howard W. Sams & Co., 1982.
23. Adel S. Sedra and Kenneth C. Smith, *Microelectronic Circuits*. Chicago: Saunders College Publishing, 1991.
24. Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed. New York: Wiley, 1984.

25. S. H. Lee, A. K. Wong, and M. G. Wong, *A Current-Combiner Circuit for Better Mixer Conversion Gain*, Philips Semiconductors applications document.
26. F. M. Fano, "Theoretical Limitations in the Broadband Matching Arbitrary Impedances," *J. Franklin Inst.*, **249**, 57–85, 139–155, 1950.
27. Ravender Goyal, ed., *High-Frequency Analog Integrated-Circuit Design*. New York: Wiley, 1995, Chap. 6.
28. Stefan Georgi and Ulrich L. Rohde, "A Novel Approach to Voltage Controlled Tuned Filters Using CAD Validation," *Microwave Eng. Eur.*, 35–42, October 1997.
29. *Analog/Interface ICs Device Data*, Vol. II, Q4/96, DL128, Rev 6, Motorola, Inc., MC13158 Wideband FM IF Subsystem specifications, pp. 8-308 to 8-329.
30. E. J. Wilkinson, "An *N*-Way Power Divider," *IRE Trans. Microwave Theory Tech.*, **MTT-8**, 116, January 1960.
31. Motorola Application Note AN-282A, *Systemizing RF Power Amplifier Design*.
32. W. E. Everitt and G. E. Anner, *Communication Engineering*. New York: McGraw-Hill, 1956.
33. G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks and Coupling Structures*. New York: McGraw-Hill, 1980.
34. G. L. Matthaei, "Tables of Chebyshev Impedance-Transforming Networks of Low-Pass Filter Form," *Proc. IEEE*, August 1964.
35. Motorola Application Note AN-267, *Matching Network Designs with Computer Solutions*.
36. E. G. Cristal, "Tables of Maximally Flat Impedance-Transforming Networks of Low-Pass Filter Form," *IEEE Trans. Microwave Theory Tech.*, **MTT 3**(5), September 1965.
37. H. W. Bode, *Network Analysis and Feedback Amplifier Design*. New York: Van Nostrand, 1943.
38. R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," *J. Franklin Inst.*, 139–154, January-February 1950.
39. J. H. Horwitz, "Design Wideband UHF-Power Amplifiers," *Electron. Des.* **11**, x–x, May 24, 1969.
40. O. Pitzalis and R. A. Gilson, "Tables of Impedance Matching Networks Which Approximate Prescribed Attenuation Versus Frequency Slopes," *IEEE Trans. Microwave Theory Tech.*, **MTT-9**(4), April 1971.
41. Anthony N. Gerkis, "Broadband Impedance Matching Using the 'Real Frequency' Network Synthesis Technique," *Appl. Microwave & Wireless*, **10**(6), 26–36, July/August 1998.
42. T. Cuthbert, "Broadband Impedance Matching Methods," *RF Des.*, **17**(8), 64–71, August 1994.
43. T. Cuthbert, *Broadband Direct-Coupled and Matching RF Networks*. Greenwood, AR, T. R. Cuthbert, 1999.
44. George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990, pp. 279–283.
45. Andy Potter, "HP RF Compiler Automates Schematic Capture and Extends Capabilities of Circuit Synthesis." *Appl. Microwave & Wireless*, **11**(6), 106–118, June 1999.
46. Aron Kain, *Final Report for an Integrated CAD Tool for High-Power, High-Efficiency RF and Microwave Amplifiers*, STTR Phase II, Contract no. F33615-95-C-174, Dept. of the Air Force, Small Business Technology Transfer (STTR) Program.
47. Kenneth K. Clarke and Donald T. Hess, *Communication Circuits: Analysis and Design.* Reading, MA: Addison-Wesley, 1971.

FURTHER READING

P. L. D. Abrie, *The Design of Impedance-Matching Networks for Radio-Frequency and Microwave Amplifiers*. Norwood, MA: Artech House, 1985.

P. L. D. Abrie and P. Rademeyer, "A Method for Evaluating and the Evaluation of the Influence of the Reverse Transfer Gain on the Transducer Power Gain of Some Microwave Transistors," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**(8), 711–713, August 1985.

A. Ambrózy, *Electronic Noise*. New York: McGraw-Hill, 1982.

W. Baechtold and M. J. O. Strutt, "Noise in Microwave Transistors," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**(9), 578–585, September 1968.

A. Cappy, "Noise Modeling and Measurement Techniques," *IEEE Trans. Microwave Theory Tech.*, **MTT-36**, 1–10, January 1988.

R. S. Carson, *High Frequency Amplifiers*. New York: Wiley, 1975.

Cezar A. A. Carioca and Paulo H. P. de Carvalho, "Computer-Aided Design of Diode Frequency Multipliers," *Appl. Microwave & Wireless*, **11**(4), 38–46, April 1999.

H. J. Carlin, "A New Approach to Gain-Bandwidth Problems," *IEEE Trans. Circuits Syst.*, **CAS-23**, 170–175, April 1977.

H. J. Carlin and P. Amstutz, "On Optimum Broad-Band Matching," *IEEE Trans. Circuits Syst.*, **CAS-23**, 401–405, April 1977.

H. J. Carlin and P. Amstutz, "On Optimum Broadband Matching," *IEEE Trans. Circuits Syst.*, **CAS-28**(5), May 1981.

H. J. Carlin and J. J. Komiak, "A New Method of Broad-Band Equalization Applied to Microwave Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**, 93–99, February 1979.

H. J. Carlin and B. Yarman, "A Simplified Real Frequency Technique Applied to Broad-Band Multistage Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-30**(12), December 1982.

H. J. Carlin and B. S. Yarman, "The Double Matching Problem: Analytic and Real-Frequency Solutions," *IEEE Trans. Circuits Syst.*, **CAS-30**(1), 15–28, January 1983.

G. Caruso and M. Sannino, "Computer-Aided Determination of Microwave Two-Port Noise Parameters," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 639–642, September 1978.

Complex Math User's Guide, Compact Software, Paterson, NJ, 1988.

Steve C. Cripps, *RF Power Amplifiers for Wireless Communications*. Norwood, MA: Artech House, 1999.

T. Cuthbert, *Circuit Design Using Personal Computers*. New York: Wiley, 1983.

J. A. Eisenberg, "Designing Amplifiers for Optimum Noise Figure," *Microwaves*, **13**, 36–41, April 1974.

R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," *J. Franklin Inst.*, **249**, 57–83, January 1960/139–155, February 1960.

H. Fukui, *Low-Noise Microwave Transistors and Amplifiers*. New York: IEEE Press, 1981.

C. Gentili, *Microwave Amplifiers and Oscillators*. New York: McGraw-Hill, 1987.

Stefan Georgi and Ulrich L. Rohde, "A Novel Approach to Voltage-Controlled Tuned Filters Using CAD Validation," *Microwave Eng. Eur.*, 35–42, October 1997.

G. Gonzales, *Microwave Transistor Amplifier Analysis and Design*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

H. A. Haus and R. B. Adler, "Optimum Noise Performance of Linear Amplifiers," *Proc. IRE*, **46**, 1517–1533, August 1958.

Aron Kain, *Final Report for Bias Dependence Noise Modeling of Heterojunction Bipolar Transistors*, USAF SBIR Phase II (PIIN), F33615-95-C-1707, November 1997. Issued by USAF/AFMC/ASC, Wright Laboratory WL/AAKE BLD 7, 2530 C ST, Wright-Patterson AFB, OH 45433-7607.

M. J. Köppen, "A Single-Stage Power Amplifier for 17 watts Output in the 470-MHz Band Using the 266BLY, Operating from 12.5 volt Supply Voltage," C.A.B. Communications, Eindhoven, Holland, Electronic Application Laboratory Report, August 8, 1970.

William B. Kuhn and Naveen K. Yanduru, "Spiral Inductor Substrate Loss Modeling in SiliconRFICs," *Microwave J.*, **11**(3), 66–81, March 1999.

J. Lange, "Noise Characterization of Linear Two-Ports of Invariant Parameters," *IEEE J. Solid-State Circuits*, **SC-2**, 37–40, June 1967.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

David Lovelace, Natalino Camilleri, and George Kannell, "Silicon MMIC Inductor Modeling for High Volume, Low Cost Applications," *Microwave J.*, **39**(8), 60–71, August 1994.

Jarek Lucek and Robbin Damen, "Designing an LNA for a CDMA Front End," *RF Des.*, **22**(2), 20–30, February 1999.

Stephen A. Maas, "FET Models for Volterra-Series Analysis," *Microwave J.*, **42**(5), 260–270, May 1999.

Donald A. McClure, KB2Z, "Transmission-Line Transformers," *Commun. Q.*, 45–58, Summer 1997.

D. J. Mellor and J. G. Linvill, "Synthesis of Interstage Networks and Prescribed Gain Versus Frequency Slopes," *IEEE Trans. Microwave Theory Tech.*, **MTT-23**(12), 1013–1020, December 1975.

Sean Mercer, "Exact Simulation of LNAs Reduces Design Cycle Time," *Appl. Microwave & Wireless*, 74–82, January 1999.

R. P. Meys, "A Wave Approach to the Noise Properties of Linear Microwave Devices," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 34–37, January 1978.

C. K. S. Miller, W. C. Daywitt, and M. G. Arthur, "Noise Standards, Measurements and ReceiverNoise Definitions," *Proc. IEEE*, **55**, June 1967.

Brian J. Minnis, *Designing Microwave Circuits by Exact Synthesis* . Boston: Artech House, 1996.

R. Minton, *Design of Large Signal VHF Transistor Power Amplifiers* , RCA Application Note SMA-36, July 1964.

R. Minton and H. C. Lee, *Frequency Multiplication Using Overlay Transistors* , RCA Application Note SMA-40, September 1965.

M. Mitama and H. Katoh, "An Improved Computational Method for Noise Parameter Measurement," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**, 612–615, June 1979.

Motorola tutorial, "The Cascode Amplifier: A Useful RF Building Block," *Microwave Prod. Digest*, 39, 59, March 1999.

T. Mukaihata, "Applications and Analysis of Noise Generation in *N*-Cascaded Mismatched Two-Port Networks," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 699–708, September 1968.

W. W. Mumford and E. H. Scheibe, *Noise Performance Factors in Communications Systems* . Dedham, MA: Horizon House, 1968.

D. E. Norton, "High Dynamic Range Transistor Amplifier Using Lossless Feedback," *Microwave J.*, **19**(5), 53–57, May 1976.

D. E. Norton, "Strong Signal Survivability Design for Lossless Feedback Amplifiers," *Proceedings of the Eleventh Annual European Microwave Conference* , 1981.

H. S. Okean and P. O. Lombardo, "Noise Performance of M/W and MM-Wave Receivers," *Microwave J.*, **16**(1), January 1973.

T. Y. Otoshi, "The Effect of Mismatched Components on Microwave Noise-Temperature Calibrations," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 675–686, September 1968.

P. Penfeld, "Wave Representations of Amplifier Noise," *IRE Trans. Circuit Theory*, 84–86, March 1962.

Ulrich L. Rohde, Jerry C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*, 2nd ed. New York: McGraw-Hill, 1997, Chap. 3.

C. Rorabaugh, *Circuit Design and Analysis—Featuring C Routines* . New York: McGraw-Hill, 1992.

M. Sannino, "On the Determination of Device Noise and Gain Parameters," *Proc. IEEE*, **67**, 1364–1366, September 1979.

SYNMAT, matching synthesis program by Radom, Inc., developed by Douglas Mellor.

A. van der Ziel, *Noise in Solid State Devices and Circuits* . New York: Wiley, 1986.

Dave Vondran, "Noise Figure Measurement: Corrections Related to Match and Gain," *Microwave J.*, **42**(3), 22–38, March 1999.

D. F. Wait, "Thermal Noise from a Passive Linear Multiport," *IEEE Trans. Microwave Theory Tech.*, **MTT-16**, 687–691, September 1968.

B. S. Yarman and H. J. Carlin, "A Simplified 'Real Frequency' Technique Applied to Broad-Band Multistage Microwave Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-30**(12), 2216–2222, December 1982.

G. P. Young and S. O. Scanlan, "Matching Network Design Studies for Microwave Transistor Amplifiers," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**(10), 1027–1035, October 1981.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

4

MIXER DESIGN

4-1 INTRODUCTION

Radiocommunication requires that we shift a baseband information signal to a frequency or frequencies suitable for electromagnetic propagation to the desired destination. At the destination, we reverse this process, shifting the received radiofrequency (RF) signal back to baseband to allow the recovery of the information it contains. This frequency-shifting function is traditionally known as *mixing*; the stages that perform it are known as *mixers*. Any device that exhibits amplitude-nonlinear behavior can serve as a mixer, for, as we saw in Section 1-6-2, nonlinear distortion results in the production, from the signals present at the input of a device, of signals at new frequencies. Even a rusty screw or bolt on an antenna element can act as a mixer, producing unwanted IMD products that appear at the receiver input.

Although mixers are equally important in wireless transmission and reception, traditional mixer terminology favors the receiving case because mixing was first applied as such in receiving applications. Thus, the signal to be frequency shifted is applied to the mixer's RF port, and the frequency-shifting power or voltage [from a *local oscillator* (*LO*)] is applied to the mixer's LO port, resulting in two outputs at the mixer's *intermediate frequency* (*IF*) port. If the wanted IF is lower than the RF signal, the mixer is a *downconverter*; if the wanted IF is higher than the RF, the mixer is an *upconverter*. *Converter* may also be used as a term for a single stage that simultaneously acts as mixer and LO.

For a given RF signal, an ideal mixer with a perfect LO (i.e., an LO with no harmonics and no noise sidebands) would produce only two IF outputs: one at the frequency sum of the RF and LO, and another at the frequency difference between the RF and LO. Filtering can be used to select the desired IF product and reject the unwanted one, which is sometimes referred to as the *IF image*.

The simultaneous generation of LO + RF and LO – RF outputs results not from a departure of mixer performance from the ideal, but from the mathematics of mixing itself. Another unavoidable mixing artifact, the *RF image* response, also results from the mathematics of mixing rather than mixer nonideality. Just as a given RF/LO combination produces two IF outputs (LO + RF and LO – RF, the IF and IF image), the mixer will produce output at the

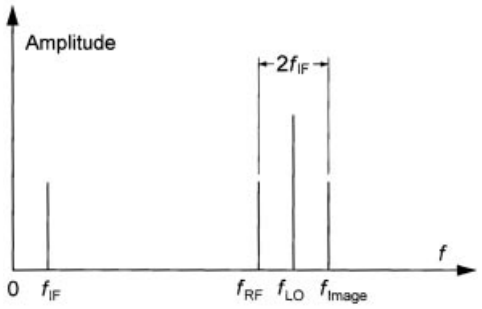


Figure 4-1 Relationship between a mixer's image and desired-signal responses. The image is $2f_{IF}$ away from the desired signal.

desired IF (LO + RF *or* LO – RF) in response to *two* possible RF inputs: one at LO + IF and another at LO (IF (Figure 4-1). The undesired response, the *RF image* (traditionally referred to merely as the *image*), is $2f_{IF}$ removed from the desired response. Even if no human-generated signals exist at the RF image frequency, reducing a mixer's RF image response can be important because noise at that frequency, including that produced by circuitry between the mixer and antenna, will still be mixed to the desired IF, degrading the signal-to-noise ratio. Filtering and phasing techniques can be used to reduce the RF or IF image responses: filtering if the image is sufficiently removed from the desired response that filtering will provide the necessary rejection; phasing if the desired and image responses are insufficiently spaced for filtering to work, as in the case of a double-conversion receiver in which signals at a high first IF (e.g., 50–70 MHz) must be converted to a very low first IF, such as 25 kHz.

The output of every real mixer includes a vast number of additional unwanted products, including noise, the fundamentals of the mixer's RF and LO signals and their harmonics, and the sums and differences of the RF and LO and their harmonics. Intermodulation distortion between multiple signals present at the RF port and IF output resulting from the mixing to IF of LO noise-sideband energy by strong adjacent signals (*reciprocal mixing*, Section 1-6-2) further complicate a mixer's output spectrum and may compromise system performance.

All mixers are *multipliers* in the sense that the various new outputs they produce can be described mathematically as the multiplicative products of their inputs. From an implementation standpoint, however, a given mixer circuit can be characterized as additive or multiplicative depending on how RF and LO signals are applied to it. Additive mixing occurs

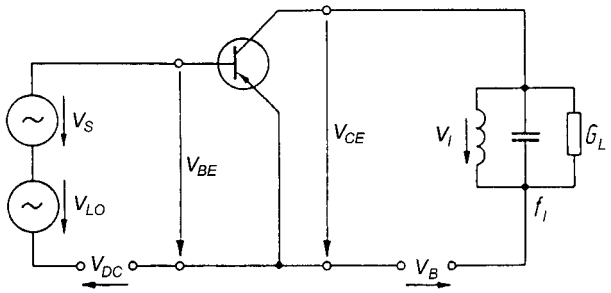


Figure 4-2 Additive mixing in a BJT [1].

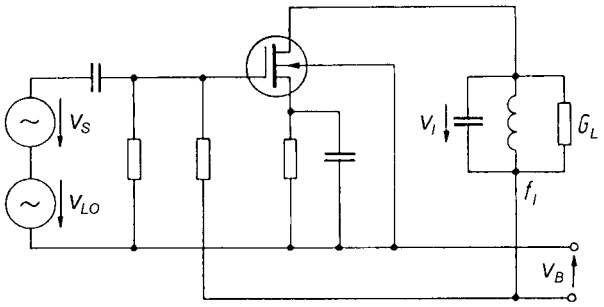


Figure 4-3 Additive mixing in a single-gate MOSFET [1].

when the RF and LO signals are applied to the same input port, as in Figures 4-2 and 4-3. Multiplicative mixing occurs when the RF and LO signals are applied to separate ports, as in Figure 4-4. As a rule, multiplicative mixers afford better isolation between their LO and RF ports than additive mixers, and this enhanced interport isolation is their principal merit. Multiplicative mixing does not in itself suppress unwanted products; the spurious response of a basic multiplicative mixer cell is poor unless it is used in a push–pull or quad configuration.

Let us now consider the basic theory of mixers. Mixing is achieved by the application of two signals to a nonlinear device. Depending on the particular device, the nonlinear characteristic may differ. However, it can generally be expressed in the form

$$I = K(V + v_1 + v_2)^n \tag{4-1}$$

The exponent n is not necessarily an integer, V may be a dc offset voltage, and the signal voltages v_1 and v_2 may be expressed as $v_1 = V_1 \sin(\omega_1 t)$ and $v_2 = V_2 \sin(\omega_2 t)$.

When $n = 2$, Eq. (4-1) may then be written as

$$I = K[V + V_1 \sin(\omega_1 t) + V_2 \sin(\omega_2 t)]^2 \tag{4-2}$$

This assumes the use of a device with a square-law characteristic. A different exponent will result in the generation of other mixing products, but this is not relevant for a basic understanding of the process. Expanding Eq. (4-2), we find

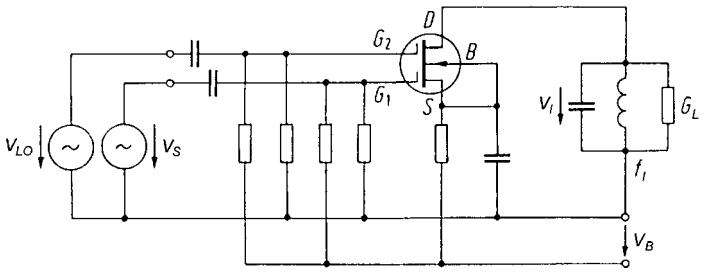


Figure 4-4 Multiplicative mixing in a dual-gate MOSFET. A dual-gate device is actually two single-gate devices in series [1].

$$I = K[V^2 + V_1^2 \sin^2(\omega_1 t) + V_2^2 \sin^2(\omega_2 t) + 2\,VV_1 \sin(\omega_1 t)$$

$$+ 2\,VV_2 \sin(\omega_2 t) + 2\,V_2V_1 \sin(\omega_2 t) \sin(\omega_1 t)] \tag{4-3}$$

The output comprises a direct current and a number of alternating current contributions. We are interested only in that portion of the current that generates the IF; so, if we neglect those terms that do not include both V_1 and V_2, we may write

$$I_{\text{IF}} = 2\,KV_1V_2 \sin(\omega_1 t) \sin(\omega_2 t)$$

$$I_{\text{IF}} = KV_2V_1\{\cos[(\omega_2 - \omega_1)t] - \cos[(\omega_2 + \omega_1)t]\} \tag{4-4}$$

This means that, at the output, we have the sum and difference signals available, and the one of interest may be selected by the IF filter.

4-2 PROPERTIES OF MIXERS

4-2-1 Conversion Gain/Loss

Even though a mixer works by means of the amplitude-nonlinear behavior in its device(s), we generally want (and expect) it to act as a linear frequency shifter. The degree to which the frequency-shifted signal is attenuated or amplified is an important mixer property. *Conversion gain* can be positive or negative; by convention, negative conversion gains are often stated as *conversion loss*.

In the case of a diode (passive) mixer , the insertion loss is calculated from the various loss components:

$$\text{Loss (dB)} = \text{Conversion loss} + \text{Transformer loss}$$

$$+ \text{Losses due to harmonic generation} + \text{Diode loss} \tag{4-5}$$

In the case of a double-balanced mixer (DBM), we must add the transformer losses (on both sides) and the diode losses as well as the mixer sideband conversion, which accounts, by definition, for 3 dB. Ideally, the mixer produces only one upper and one lower sideband, which results in the 3-dB loss compared to the input signal. Also, the input and output transformers add about 0.75 dB on each side, and of course there are the diode losses because of the series resistances of the diodes.

Figure 4-5 shows the equivalent circuit of a diode. It consists of a series (loss) resistor R_S and a time-variable electronic resistor, typically called the *diffusion resistance*, R_D, which equals 26 mV/I_D, and a capacitance C_D shunting R_D. C_D can be found from

$$C_D = \frac{W^2}{2D}\frac{I_D}{V_T} \tag{4-6}$$

where D is the diffusion constant, a material-dependent value, and W is the physical width. The average value for R_D is somewhere between the calculated value of 26 mV/I_D and some

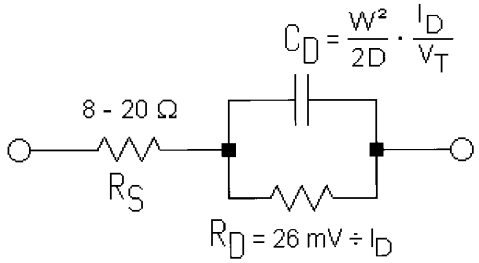


Figure 4-5 Equivalent circuit of a mixer diode.

leakage current, simply because it is generated by a rectification mechanism, which turns the LO power into an RF current and then into a combination of dc and RF currents.

We can calculate the diode loss according to

$$\text{Diode loss (db)} = \log 10\left(\frac{50 + (2 \times R_S)}{50}\right) \tag{4-7}$$

Assuming that $R_S = 8\ \Omega$, the diode loss for a diode-ring mixer is

$$\text{Diode loss (db)} = \log 10\left(\frac{50 + (2 \times 8)}{50}\right) = 0.5\ \text{dB} \tag{4-8}$$

From Eq. (4-5), the insertion loss for this mixer is therefore

$$\begin{aligned}\text{Loss (dB)} &= 3\ \text{dB (Conversion loss)} + 1.5\ \text{dB (Transformer loss)} \\ &\quad + 1\ \text{dB (Losses from harmonic generation)} + 0.5\ \text{dB (Diode loss)} \\ &= 6\ \text{dB}\end{aligned} \tag{4-9}$$

This assumes mixing at the fundamental frequency. Sometimes the diode loss resistor R_S is as high as 25 Ω per arm (as in a MOSFET switch), or 50 Ω total. This now results in

$$\text{Loss (dB)} = 3\ \text{dB} + 1.5\ \text{dB} + 1\ \text{dB} + 3\ \text{dB} = 8.5\ \text{dB} \quad \text{(Insertion loss)} \tag{4-10}$$

Since the value of R_S is partially determined by the threshold voltage of the diode and the diode diffusion resistance, R_D, a wide range of values can be noticed for different drive levels and mixer topologies. Figure 4-5 shows a shunt capacitance C_D, the so-called *diode diffusion capacitance*. When the diode is conducting, the influence of this nonlinearity is frequency dependent, which adds to the insertion loss. In this discussion we have not considered its frequency dependency. At wireless frequencies, modern Schottky diodes, also frequently called *hot-carrier* diodes, are operated far from their cutoff frequency, resulting in less than 1 dB of additional losses. There are also mixers with special circuitry to terminate the IF image. This is done with a diplexer circuit or equivalent circuitry. This makes insertion loss values as low as 4 dB possible; however, the large-signal condition or intercept point suffers.

4-2-2 Noise Figure

Like any network, a mixer contributes noise to the signals that it frequency-shifts. The degree to which a mixer's noise degrades the signal-to-noise ratio (see Section 1-7-1) of the signals it frequency-shifts is evaluated in terms of noise factor and noise figure as discussed in a subsection of Section 1-7-1 entitled "Noise Figure of Cascaded Networks."

For a long time, the literature has stated that the noise figure of a passive mixer, which is pretty much independent of its circuit arrangement, is equal to the mixer's insertion loss. But this neglects the influence of the white-noise contribution of the mixer's diode(s). This is ironic, considering that RF noise generators were long based on thermionic diodes operated in saturation (the 5722 noise diode was a popular type, as in the Rohde & Schwarz SKTU). Such a diode's noise power output can readily be determined from its saturation current. On the other hand, all Schottky diodes, while conducting, generate white noise that follows the same principle as above. This fact has been recognized by only a few companies that make modern noise-measurement equipment. Modern noise-measurement devices measure the noise figure of a system by a "hot and cold" technique, an approach based on knowledge of the absolute noise energy emitted under hot conditions (conductance). This method has the advantage that it can be used up to several tens of gigahertz, while the old vacuum-tube-based noise generators ran out of steam at around 1 GHz due to the inability to match the tube to the 50-Ω termination. This was typically accomplished by connecting a 50-Ω resistor between anode and ground (without dc connection), followed by a low-pass filter, which would match the tube capacitance and other parasitics to the required termination of 50 Ω, purely resistive.

In reality, we can take the loss calculation from above and add the Schottky noise generated by the diodes as they are driven by the local oscillator.

If a Schottky diode is used as a noise generator in the conductive mode, it generates a continuous frequency spectrum, possibly up to several gigahertz. There is a mathematical relationship between the noise power spectrum emitted by the diode and the time-averaged current of this diode, which generates the noise. If the noise source impedance is set (typically 50 Ω), the available noise power can be calculated according to

$$I_R = \sqrt{2e \times I_s \times \Delta f} \tag{4-11}$$

where $e = 1.6 \times 10^{-19}$ coulombs
I_s = saturation current of the diode
Δf = effective noise bandwidth

For $S_{11} = 0$ or proper termination of this circuit ($R_G = R_{term}$),

$$P_R = \left(\frac{I_R}{2}\right)^2 \times R_i$$

$$= \frac{e}{2} \times I_s \times \Delta f \times R_i \tag{4-12}$$

Calculated at a bandwidth of 1 Hz,

$$\frac{P_R}{\Delta f} = \frac{e}{2} \times I_s \times R_i \tag{4-13}$$

If

$$\frac{P_R}{\Delta f} = kT_0 \times F \tag{4-14}$$

the noise factor becomes

$$F = \frac{e \times I_s \times R_i}{2kT_0} \tag{4-15}$$

If the values for e and kT_0 are inserted,

$$F = 20 \times I_D \times \frac{26 \text{ mV}}{I_D} + \frac{R_S + R_G}{2R_G} \tag{4-16}$$

Example. Assume the passive mixer mentioned above with its 6-dB insertion loss is considered and a dc current of 15 mA results as a function of the LO drive. Since I_D gets canceled, the noise factor (F) of the diode portion equals

$$F = 20 \times I_D \times \frac{26 \text{ mV}}{I_d} + \frac{R_S + R_G}{2R_G}$$

$$= 0.52 + 0.58$$

$$= 1.1 \tag{4-17}$$

The noise figure, NF, is 10 log F, or 0.413 dB. Now this number and the insertion loss must be added. The resulting noise figure would be 6.413 dB. This is consistent with published measurement data.

Exact Mathematical Nonlinear Approach. The exact noise factor of a real mixer is computed by the formula

$$F = \frac{N_0(\omega_{\text{IF}}) + kT_0}{K_B T_0 G_{Tc}(\omega_{\text{RF}})} \tag{4-18}$$

where $N_0(\omega_{\text{IF}})$ = total noise power (per unit bandwidth) delivered to the IF load at intermediate frequency
K_B = Boltzmann's constant
T_0 = reference temperature (290 or 300 K is commonly used)
$G_{Tc}(\omega_{\text{RF}})$ = transducer conversion gain from ω_{RF} to ω_{IF}.

Let us now further elaborate on Eq. (4-18). We may write

$$N_0(\omega_{\text{IF}}) = N_S(\omega_{\text{IF}}) + N_{\text{INT}}(\omega_{\text{IF}}) + N_L(\omega_{\text{IF}}) + kT_0 \tag{4-19}$$

where N_S is noise generated by the RF source resistance and transferred to the IF load through frequency conversion, N_{INT} is noise generated internally to the mixer, and N_L is noise generated by the IF termination. If the source resistance is held at temperature T_0, N_S will basically originate from noise generated at the RF and image frequencies, which are transferred to the IF with approximately the same conversion gain, plus a relatively small contribution transferred from other sidebands with a smaller conversion gain. We may then write synthetically

$$N_S(\omega_{\text{IF}}) = 2aK_BT_0G_{Tc}(\omega_{\text{RF}}) \tag{4-20}$$

where a is a coefficient slightly larger than 1. $N_{\text{INT}}(\omega_{\text{RF}})$ is generated by transformer losses, by the diode Schottky noise, and by the diode resistive parasitics, and, in principle, it may take on any value; in particular, it may be zero if both the transformers and the diodes are ideal (i.e., if the latter are pure nonlinear resistors). As for $N_L(\omega_{\text{IF}})$, by Nyquist's theorem the IF load resistor R_L may be described as a noiseless resistor in series with a noise voltage source whose mean-square voltage (per unit bandwidth) is

$$|V_L|^2 = 4K_BT_LR_L \tag{4-21}$$

where T_L is the IF termination temperature. If the IF load is driven by a source with an output impedance $Z_{\text{out}}(\omega_{\text{IF}})$, the noise power actually delivered to the load will obviously be

$$N_{\text{out}} = \frac{4K_BT_LR_L^2}{|Z_{\text{out}}(\omega_{\text{IF}}) + R_L|^2} \tag{4-22}$$

In addition to N_{out}, the thermal noise originating from the IF termination delivered to the IF load at ω_{IF} will also include contributions from other sidebands that are backconverted by the mixer nonlinearities with a relatively small conversion gain. Thus, we may write

$$N_L(\omega_{\text{IF}}) = \frac{4bK_BT_LR_L^2}{|Z_{\text{out}}(\omega_{\text{IF}}) + R_L|^2} \tag{4-23}$$

where b is slightly larger than 1. If we now introduce the mixer conversion loss, namely,

$$L_C = \frac{1}{G_{Tc}(\omega_{\text{RF}})} \tag{4-24}$$

and combine Eq. (4-18) with Eqs. (4-19), (4-20), and (4-23), we finally get the noise factor expression:

$$F = 2a + \frac{N_{\text{INT}}(\omega_{\text{IF}}) + kT}{K_BT_0}L_C + \frac{4bT_LR_L^2}{T_0|Z_{\text{out}}(\omega_{\text{IF}}) + R_L|^2}L_C \tag{4-25}$$

In the normal region of operation of the mixer (sufficient LO drive) we may assume

Table 4-1 Noise figure and conversion gain versus LO power for a diode DBM

LO Power (dBm)	NF (dB)	Conversion Gain (dB)
–10.0	45.3486	–45.1993
–8.0	32.7714	–32.5264
–6.0	19.8529	–19.2862
–4.0	12.1154	–11.3228
–2.0	8.85188	–8.05585
0.0	7.26969	–6.51561
2.0	6.42344	–5.69211
4.0	5.85357	–5.15404
6.0	5.50914	–4.84439
8.0	5.31796	–4.66871
10.0	5.19081	–4.54960
12.0	5.08660	–4.45887
14.0	4.99530	–4.38806
16.0	4.91716	–4.33322
18.0	4.85920	–4.29407
20.0	4.82031	–4.26763

$$Z_{\text{out}}(\omega_{\text{IF}})R_L \tag{4-26}$$

so that Eq. (4-25) becomes

$$F = 2a + \frac{N_{\text{INT}}(\omega_{\text{IF}}) + kT}{K_B T_0} L_C + \frac{bT_L}{T_0} L_C \tag{4-27}$$

Multiplying $\log_{10} F$ by 10 gives us the exact mixer noise figure in dB.

The method just discussed is the basis for the far noise calculation for oscillators as mentioned in Chapter 5, in which it is referred to as conversion noise. This includes the AM-to-PM conversion noise.

Table 4-1 shows how the noise figure and conversion gain vary with LO power for a generic diode DBM (Figure 4-6). "Starving" a diode mixer by decreasing its LO drive rapidly degrades its performance in all respects.

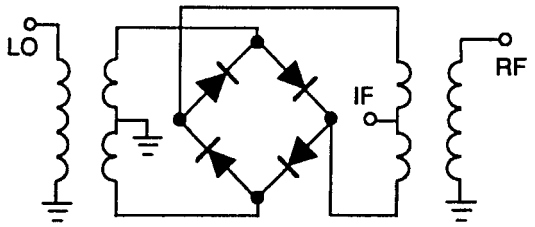


Figure 4-6 Generic diode DBM.

SSB Versus DSB Noise Figure. In radiocommunication, we typically use mixers as "single-sideband" devices; that is, we are interested in only one of the two possible RF inputs ($f_{LO} + f_{IF}$ or $f_{LO} - f_{IF}$) that produce output at the desired IF. The unused input is the *RF image* or simply *image*. A mixer's RF image response can be important even if no human-generated signals are present at f_{image} because noise there, including that produced by the mixer's RF-port termination, will be converted to the IF, possibly compromising the system's signal-to-noise ratio. As a result of this and the IEEE definition of noise figure [2], which, for a mixer, considers only the noise associated with a system's principal frequency transformation, some controversy exists about the definition and measurement of mixer noise figure. IEEE's NF definition for mixers assumes *no* noise contribution at f_{image}, not even from the mixer's RF-port termination at that frequency; yet, by convention (and out of necessity, since an absolutely noiseless $T = 0\,f_{image}$ termination is unavailable on the average test bench), the noise figures we measure always include *some* noise at f_{image}, even if, as a result of filtering, it arises only from the mixer's RF-port termination. The NF measurements and simulations presented in this book reflect this conventional SSB NF scenario. Stephen Maas [3] provides in-depth material on these issues, including the importance of double-sideband NF for some systems, and the limitations of NF as a figure of merit.

4-2-3 Linearity

The 1-dB Compression Point. Like other networks, a mixer is amplitude-nonlinear above a certain input level; above this point, the output level fails to track input-level changes proportionally. This figure of merit, P_{-1dB}, identifies the single-tone input-signal level at which the output of the mixer has fallen 1 dB below the expected output level. The 1-dB compression point in a conventional double-balanced diode mixer is approximately 6 dB below the LO power. For lower distortion mixers, it is usually 3 dB below the LO power.

The 1-dB Desensitization Point. This specification is another figure of merit similar to the 1-dB compression point. However, the 1-dB desensitization point refers to the level of an interfering (undesired) input signal that causes a 1-dB decrease in nominal conversion gain for the desired signal. For a diode-ring DBM, the 1-dB desensitization point is usually 2–3 dB below the 1-dB compression point.

Dynamic Range. The dynamic range of any RF/wireless system can be defined as the difference between the 1-dB compression point and the minimum discernible signal (MDS). These two points are specified in units of power (dBm), giving dynamic range in dB. When the RF input level approaches the 1-dB compression point, harmonic and intermodulation products begin to interfere with the system performance. High dynamic range is obviously desirable, but cost, power consumption, system complexity, and reliability must also be considered.

Harmonic Intermodulation Products (HIPs). These are spurious products that are harmonically related to the f_{LO} and f_{RF} input signals.

$$\text{HIP} = Mf_{LO} + Nf_{RF} \tag{4-28}$$

Table 4-2 Typical spurious responses of high-level double-balanced mixer (decibels below $f_{LO} \pm f_{RF}$ response)

RF Input Signal Harmonics		f_{LO}	$2f_{LO}$	$3f_{LO}$	$4f_{LO}$	$5f_{LO}$	$6f_{LO}$	$7f_{LO}$	$8f_{LO}$
$8f_{RF}$	100	100	100	100	100	100	100	100	100
$7f_{RF}$	100	97	102	95	100	100	100	90	100
$6f_{RF}$	100	92	97	95	100	100	95	100	100
$5f_{RF}$	90	84	86	72	92	70	95	70	92
$4f_{RF}$	90	84	97	86	97	90	100	90	92
$3f_{RF}$	75	63	66	72	72	58	86	58	80
$2f_{RF}$	70	72	72	70	82	62	75	75	100
f_{RF}	60	0	35	15	37	37	45	40	50
		60	60	70	72	72	62	70	70

Table 4-2 shows relative harmonic intermodulation product levels for a high-level diode DBM.

Intermodulation Distortion (IMD). Nonlinearities in the mixer devices give rise to intermodulation distortion products whenever two or more signals are applied to the mixer's RF port. Testing this behavior with two (usually closely spaced) input signals of equal magnitude can return several figures of merit depending on how the results are interpreted. A mixer's third-order output intercept point ($IP_{3,out}$) is defined as the output power level where the spurious signals generated by $(2f_{RF1} \pm f_{RF2}) \pm f_{LO}$ and $(f_{RF1} \pm 2f_{RF2}) \pm f_{LO}$ are equal in amplitude to the desired output signal as shown in Figure 4-7.

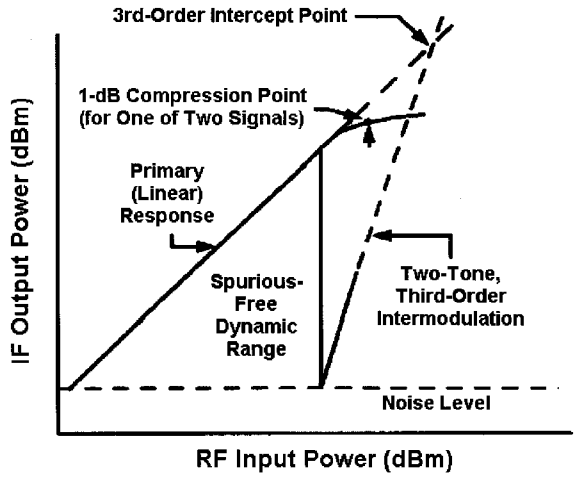


Figure 4-7 Mixer linearity evaluation, including compression and two-tone IMD dynamic range. P_{-1dB} for a single tone cannot be read directly from this graph because the values shown are the result of two-equal-tone drive.

The third-order input intercept point, $IP_{3,in}$—IP_3 referred to the input level—is of particularly useful value and is the most commonly used mixer IMD figure of merit. $IP_{3,in}$ can be calculated according to

$$IP_{n,in} = IMR \div (n - 1) + \text{Input power (dBm)} \tag{4-29}$$

where *IMR* is the intermodulation ratio (the difference in dB between the desired output and the spurious signal), and n is the IM order—in this case, 3. In a conventional diode double-balanced mixer, $IP_{3,in}$ is approximately 14 dB above the single-tone 1-dB compression point (P_{-1dB})—approximately 8 dB greater than the local oscillator power. As will be seen later, this does not apply to feedback-active mixers. They have their own agenda: Above a particular input level, their IMD products increase almost exponentially.

Although designers are usually more concerned with odd-order IM performance, second-order IM can be important in wideband systems (systems that operate over a 2:1 or greater bandwidth) as discussed in Section 1-7-2.

4-2-4 LO Drive Level

A mixer's specifications are usually guaranteed at a particular LO drive level, usually specified as a dBm value that may be qualified with a tolerance. Insufficient LO drive degrades mixer performance; excessive LO drive degrades performance and may damage mixer devices. Commercially available diode mixers are often classified by LO drive level; for example, a "Level 17" mixer requires 17 dBm of LO drive.

4-2-5 Interport Isolation

In a mixer, isolation is defined as the attenuation in dB between a signal input at any port and its level as measured at any other port. High isolation numbers are desirable. Figure 4-8 shows LO-to-IF and LO-to-RF isolation versus frequency for a triple-balanced diode DBM. Isolation is dependent mainly on transformer and physical symmetry and device balance, but the level of signals applied to the mixer also plays a role, as shown in Figure 4-9.

4-2-6 Port VSWR

The load presented by a mixer's ports to the outside world can be of critical importance to a designer. For example, high LO-port VSWR may result in inefficient use of available LO power, resulting in LO starvation (underdrive) that degrades the mixer's performance. Figure 4-10 shows LO-port VSWR versus frequency for a high-level diode DBM with two values of LO power. Like interport isolation, port VSWR can vary with the level of the signal applied.

4-2-7 dc Offset

Isolation between ports plays a major role in reducing dc offset in a mixer. Like isolation, dc offset is a measure of the imbalance of the mixer. In phase-detector and phase-modulator applications, dc offset is a critical parameter.

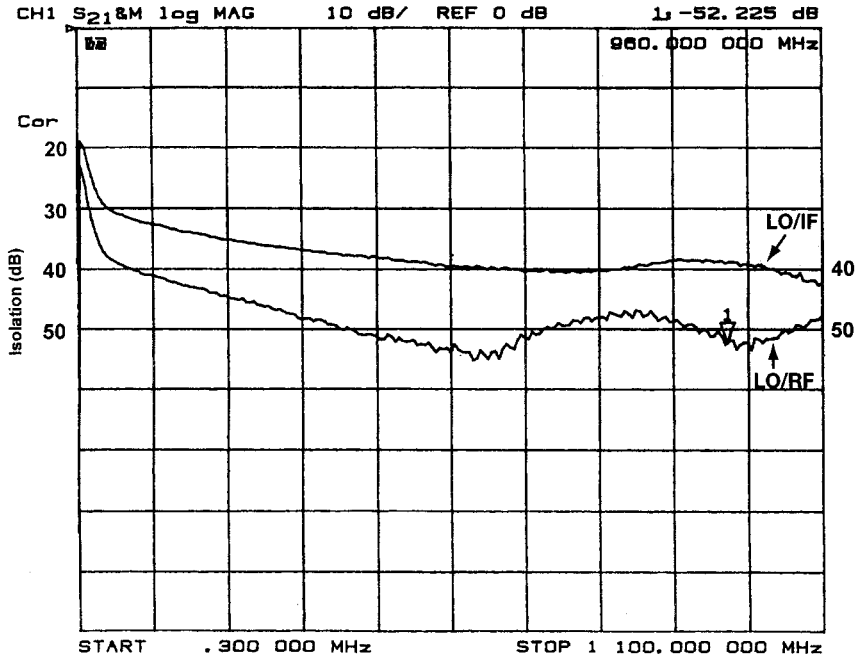


Figure 4-8 LO–IF and LO–RF isolation versus frequency for a high-level triple-balanced diode mixer. The periodic roughness of the traces is a measurement-system artifact, proving that even sophisticated FFT-based instruments are not perfect.

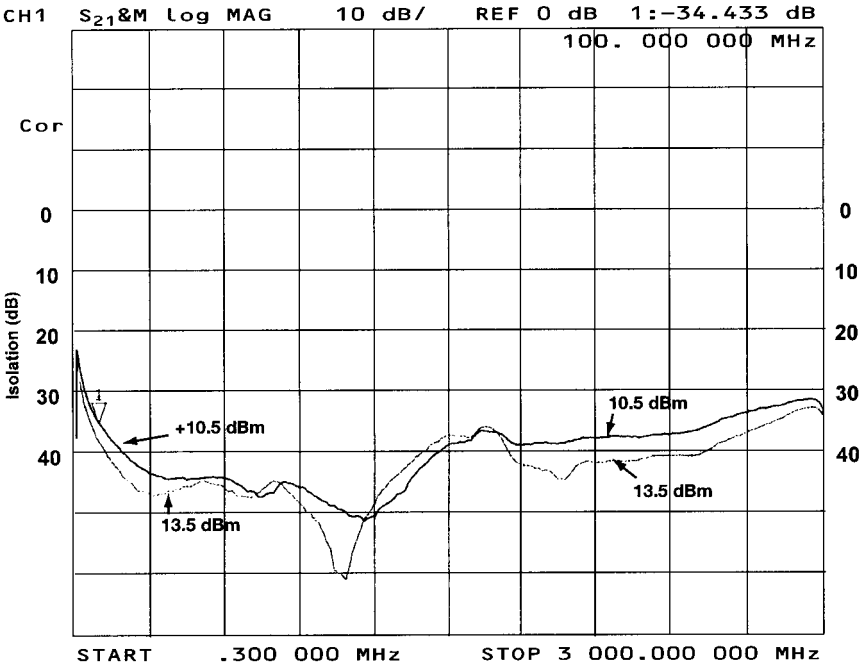


Figure 4-9 LO–IF isolation versus frequency and LO drive level for a high-level diode DBM.

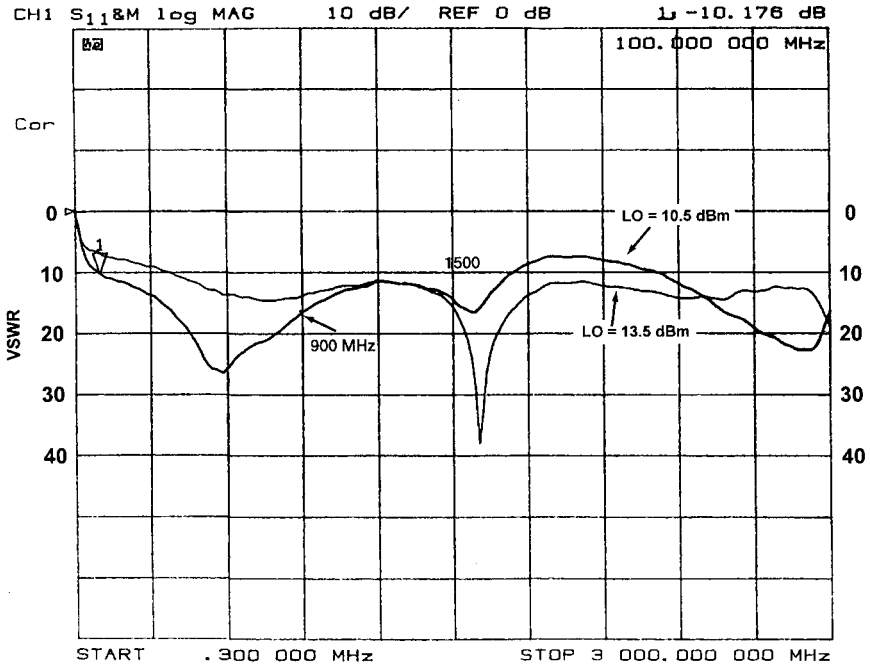


Figure 4-10 LO-port VSWR versus frequency for a high-level diode DBM.

4-2-8 dc Polarity

Unless otherwise specified, mixers with dc output are designed to have negative polarity when RF and LO signals are of equal phase.

4-2-9 Power Consumption

Circuit power consumption is always important, but in battery-powered wireless designs it is *critical.* Mixer choice may be significant in determining a system's power consumption, sometimes in ways that seem paradoxical at first glance. For instance, a passive mixer might seem to be a power-smart choice because it consumes *no* power—until we factor in the power consumption of the circuitry needed to provide the (often considerable) LO power a passive mixer requires. If a mixer requires a broadband resistive termination that will be provided by a postmixer amplifier operating at a high standing current, the power consumption of the amplifier stage must be considered as well. Evaluating the suitability of a given mixer type to a task therefore requires a grasp of its ecology as well as its specifications.

4-3 DIODE MIXERS

Passive mixers based on diode switches are common in base-station applications, where their high dynamic range and 50-Ω port impedances overcome objections to their inherent conversion loss, unsuitability to integration, and relatively high LO-power requirement. As we'll discuss later, FET-based passive mixers are overcoming some of these limitations in portable/mobile wireless applications.

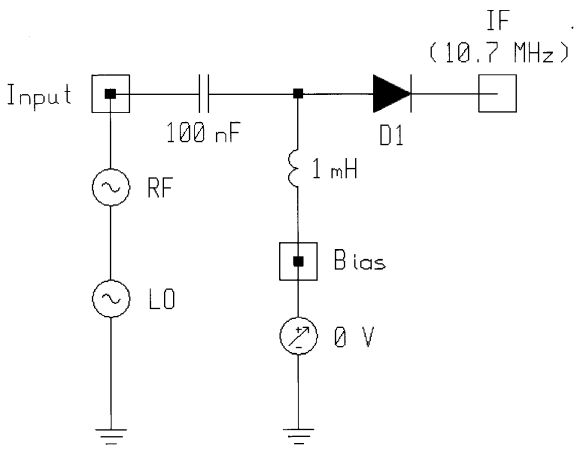


Figure 4-11 Schematic of the single-diode mixer. The circuit's 50-Ω output termination serves as the diode's dc return, as well as completing the circuit for RF, LO, and IF.

4-3-1 Single-Diode Mixer

Figure 4-11 shows the schematic of a simplistic single-diode mixer. The LO and RF signals are applied in series to the diode, with no effort made to match the sources to the diode or isolate the sources from each other. The LO and RF signals are both present in the diode simultaneously, so the mixing performed is additive. The LO signal switches the diode on and off, gating the RF signal to the circuit's output at the LO frequency. In this simple circuit, the output port's 50-Ω resistance serves as the diode's ac load and dc return. Because we want the diode to operate as a linear switch with respect to the RF signal, the diode should turn on hard and turn off as completely as possible, and the LO, not the RF source, should switch the diode on and off. The upper frequency limit of the circuit will depend on how rapidly the diode can switch between the on and off states. The LO can be only so strong before the diode's dissipation limits are exceeded; the RF source can be only so strong relative to the LO before it begins to play a role in switching the diode on and off. The LO must therefore be considerably stronger than the RF source—20 dB or more for low-distortion applications.

Figure 4-12 shows how the single-diode mixer's conversion gain and noise figure vary with applied LO power. Figure 4-13 shows the mixer's output spectrum.

Tables 4-3, 4-4, and 4-5 present sample data on diodes suitable for mixer service. The manufacturer's full datasheet on the Siemens BAT14-099 follows.

The simple single-diode-mixer circuit shown in Figure 4-11 is intended only as an illustration of the basic behavior of diode mixer behavior. A practical single-diode mixer would include filtering at its RF, LO, and IF ports—RF filtering for image rejection, reduced LO radiation and optimum matching of the RF source to the diode; LO filtering to keep RF out of the LO and optimally match the LO to the diode; and IF filtering to optimally match the diode to its IF (and IF image) load, preferably while providing some rejection of the mixer's unwanted outputs, the strongest (and most potentially troublesome) of which is the LO signal. Detail on the design of single-diode mixers can be found in Zinke and Brunswig [4].

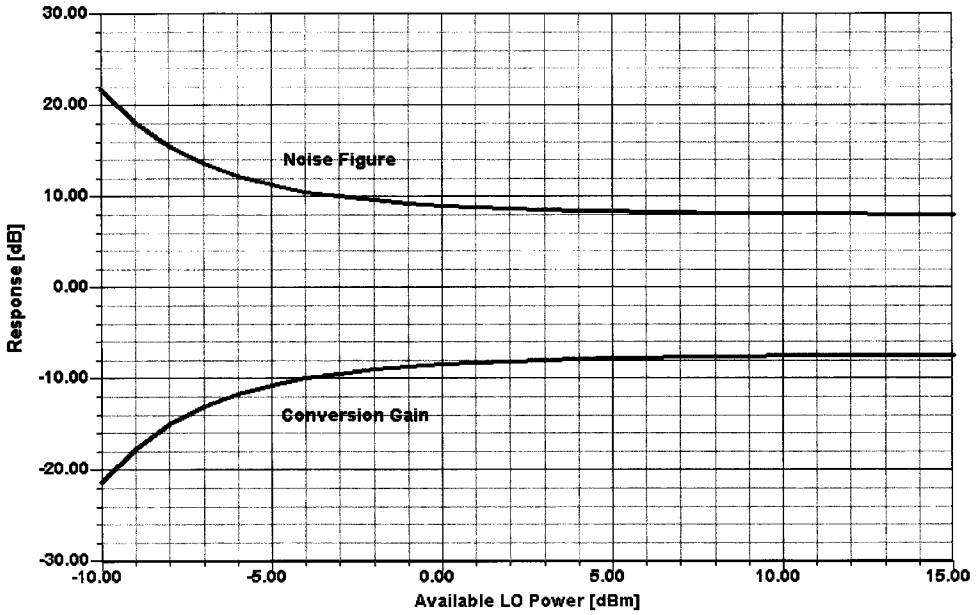


Figure 4-12 Conversion gain and noise figure versus LO power for the single-diode mixer. The values reported are worse than those predicted by theory because RF, LO, and IF matching has not been attempted. In this analysis, LO = 310.7 MHz (–10 to 25 dBm), RF = 300 MHz (–50 dBm), and IF = 10.7 MHz.

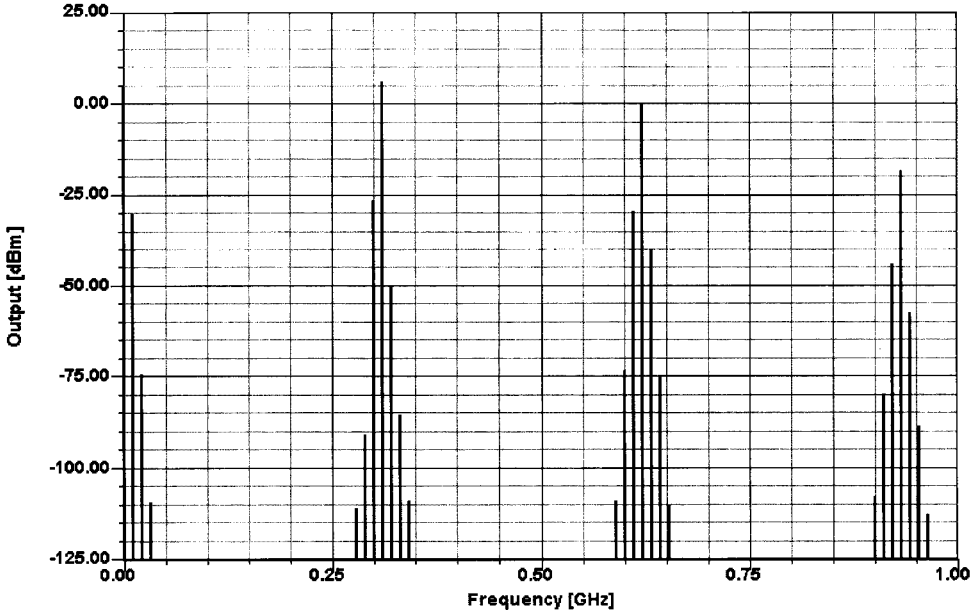


Figure 4-13 Output spectrum of the single-diode mixer. Three LO harmonics and 3 LO sidebands were used in this analysis. In this analysis, LO = 310.7 MHz (13 dBm), RF = 300 MHz (–20 dBm), and IF = 10.7 MHz.

Table 4-3 Example RF Schottky diodes (Siemens)

Type	Maximum Ratings		Characteristics (T_A = 25 °C)					Package
	V_R	I_F	C_T	V_F at I_F		V_F at I_F		
	(V)	(mA)	(pF)	(mV)	(mA)	(mV)	(mA)	
BAT14-03W	4	90	0.22	430	1	550	10	SOD-323
BAT14-099 (dual)	4	90	0.22	430	1	550	10	SOT-143
BAT14-099R (quad)	—	90	0.38	400	1	480	10	SOT-143
BAT15-03W	4	110	0.21	230	1	320	10	SOD-323
BAT15-099 (dual)	4	110	0.21	230	1	320	10	SOT-143
BAT15-099R (quad)	—	110	0.37	230	1	320	10	SOT-143
BAT17	4	130	0.55	340	1	425	10	SOT-23
BAT17-04 (dual)	4	130	0.55	340	1	425	10	SOT-23
BAT68	8	130	0.75	320	1	395	10	SOT-23
BAT68-04 (dual)	8	130	0.75	320	1	395	10	SOT-23
BAT68-05 (dual)	8	130	0.75	320	1	395	10	SOT-23
BAT68-06 (dual)	8	130	0.75	320	1	395	10	SOT-23
BAT68-07 (dual)	8	130	0.75	320	1	395	10	SOT-143
BAT68-03W	8	130	0.75	320	1	395	10	SOD-323
BAT114-099 (dual)	4	90	0.22	580	1	680	10	SOT-143
BAT114-099R (quad)	—	90	0.22	580	1	680	10	SOT-143

4-3-2 Single-Balanced Mixer

Figure 4-14 shows the schematic of a two-diode, *single-balanced* mixer (SBM). Unlike the single-diode mixer, it performs multiplicative mixing because its RF and LO signals are applied to different ports. In this more commonly seen two-diode mixer configuration, a balanced transformer drives the diodes out of phase for the LO and in phase for signals present at the RF port. Figure 4-15 shows how this mixer's conversion gain and noise figure vary with applied LO power. Figure 4-16 shows how the mixer's conversion gain and noise figure vary with frequency for a constant LO power.

Table 4-4 Example Schottky detector diodes (Siemens)

Type	Maximum Ratings		Characteristics (T_A = 25 °C)								Package
	V_R	I_F	C_T at V_R		V_F at I_V		R_O at V_F		I_R at V_R (μA)		
	(V)	(mA)	(pF)	(V)	(V)	(mA)	(kΩ)	(V)	(μA)	(V)	
BAT62	40	20	0.4	0	0.53	2	160	0	≤10	40	SOT-143
BAT62-03W	40	20	0.4	0	0.53	2	160	0	≤10	40	SOD-323
BAT63	3	100	0.65	0.2	0.19	1	30	0	≤10	3	SOT-143

Table 4-5 Example Schottky diodes and quads

Schottky Diodes for Mixer & Detector Applications

Barrier	V_B @ 10 μA (V)	C_T @ 0 V (pF)	V_F @ 1 mA (mV)	R_T* @ 10 mA (Ω)	SOD-323	SOT-23			SOT-143
	Min.			Max.	Single	Single	Series Pair	Reverse Series Pair	Unconnected Pair
Low	2	0.43–0.63	200–270	8	SMS1546-011	SMS1546-001	SMS1546-005		SMS1546-015
Low	2	0.2–0.26	260–320	18	SMS7621-011	SMS7621-001	SMS7621-005	SMS7621-006	SMS7621-015

*R_T is the slope resistance.

Schottky Quads

Barrier	V_B @ 10 μA (V)	C_J[1] @ 0 V (pF)	V_F @ 1 mA (mV)	ΔV_F @ 1 mA (mV)	R_T[2] @ 10 mA (Ω)	SOT-143			
	Min.			Max.	Max.	Ring Quad	Crossover Quad	Bridge Quad	Octo Quad
Low	2	0.3–0.5	200–270	10	8	SMS3926-022	SMS3926-023	SMS3929-021	SMS3938-026
Medium	3	0.3–0.5	310–370	10	8	SMS3927-022	SMS3927-023	SMS3930-021	SMS3939-026
High	4	0.3–0.5	520–580	10	8	SMS3928-022	SMS3928-023	SMS3931-021	SMS3940-026

1. C_J is the capacitance per junction.
2. R_T is the slope resistance.

The two-diode mixer is used mostly in the frequency range above 1 GHz in a manner akin to a phase discriminator, using step-recovery diodes in the LO feed for enhanced harmonic mixing. Such mixers are mainly used in medium-cost spectrum analyzers or microwave receivers up to several tens of gigahertz, with the necessary transformers and baluns printed on the circuit board. With perfectly matched diodes and perfect transformer and constructional symmetry, no LO energy arrives at the IF and RF ports, and there is only slight attenuation between the RF and IF ports. Both building and computer modeling such a mixer are impossible: building, because perfectly matched diodes and perfect transformer and constructional symmetry cannot be achieved in practice; computer modeling, because floating-point mathematics runs out of gas in handling the infinite amplitude spread involved in calculating the perfect cancellation of the LO signal as it travels to the RF and IF ports. That said, Figure 4-17 compares the mixer's *quasi*-ideal port-to-port isolation (matched diodes, a perfect transformer, no stray inductances and capacitances, and a 10-MΩ resistors

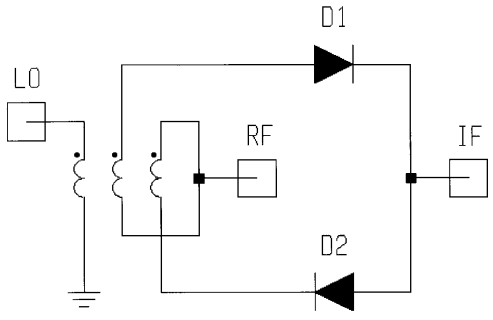


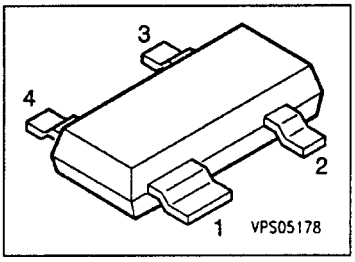
Figure 4-14 Schematic of the two-diode (also known as *single-balanced*) mixer.

SIEMENS

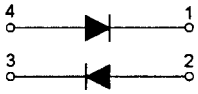
Silicon Dual Schottky Diode

BAT 14-099

- DBS mixer application to 12 GHz
- Low noise figure
- Medium barrier type



ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering Code (tape and reel)	Pin Configuration	Package[1]
BAT 14-099	S9	Q62702-A3461		SOT-143

Maximum Ratings per Diode

Parameter	Symbol	Values	Unit
Reverse voltage	V_R	4	V
Forward current	I_F	90	mA
Power dissipation, $T_S \leq 55$ °C	P_{tot}	100	mW
Storage temperature range	T_{stg}	– 55 ... + 150	°C
Operating temperature range	T_{op}	– 55 ... + 150	

Thermal Resistance

Junction – ambient[2]	$R_{th\,JA}$	≤ 1090	K/W
Junction – soldering point	$R_{th\,JS}$	≤ 930	

1) For detailed information see chapter Package Outlines.
2) Package mounted on alumina 15 mm × 16.7 mm to 0.7 mm.

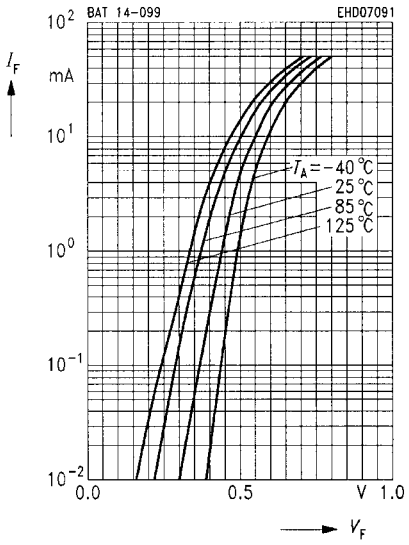
SIEMENS **BAT 14-099**

Electrical Characteristics per Diode
at T_A = 25 °C, unless otherwise specified.

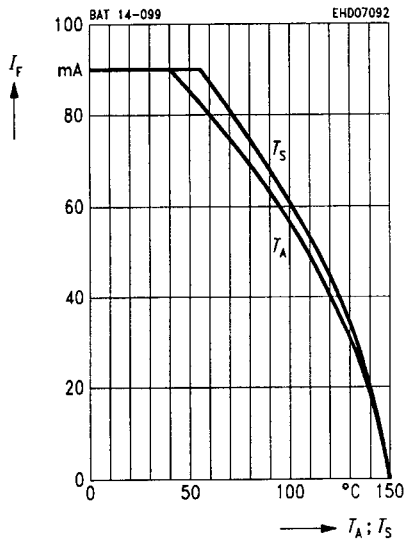
Parameter	Symbol	Values			Unit
		min.	typ.	max.	
Breakdown voltage I_R = 5 µA	V_{BR}	4	–	–	V
Forward voltage I_F = 1 mA I_F = 10 mA	V_F	 – –	 0.43 0.55	 – –	
Forward voltage matching I_F = 10 mA	ΔV_F	–	–	10	mV
Diode capacitance V_R = 0, f = 1 MHz	C_T	–	–	0.35	pF
Forward resistance I_F = 10 mA / 50 mA	R_F	–	5.5	–	Ω

SIEMENS **BAT 14-099**

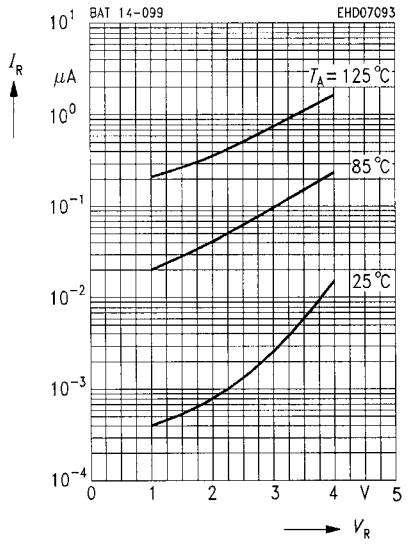
Forward current $I_F = f(V_F)$



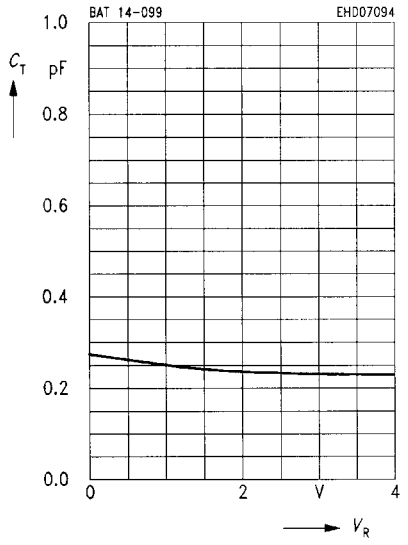
Forward current $I_F = f(T_S; T_A^*)$
*Package mounted on alumina



Reverse current $I_R = f(V_R)$



Diode capacitance $C_T = f(V_R)$
$f = 1$ MHz



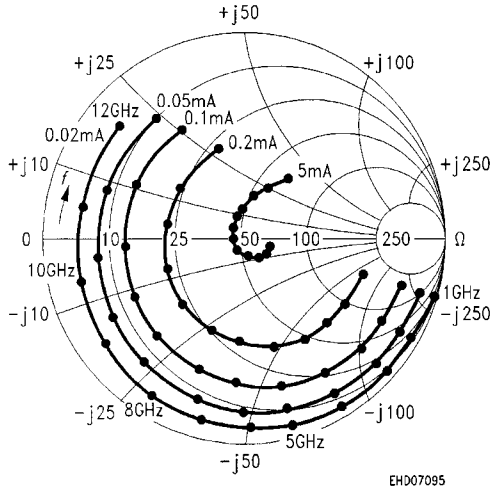
SIEMENS **BAT 14-099**

S_{11}-Parameters

Typical impedance characteristics (with external bias I and $Z_0 = \Omega$)

f	I = 0.02 mA		I = 0.05 mA		I = 0.1 mA		I = 0.2 mA		I = 0.5 mA	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
1	0.99	– 15.89	0.91	– 16.40	0.79	– 16.40	0.57	– 16.60	0.13	– 17.30
2	0.96	– 30.40	0.88	– 30.80	0.76	– 31.09	0.56	– 30.70	0.13	– 28.40
3	0.95	– 45.30	0.87	– 46.20	0.75	– 47.30	0.55	– 47.00	0.11	– 43.99
4	0.93	– 59.60	0.86	– 61.60	0.73	– 62.40	0.53	– 62.40	0.10	– 54.40
5	0.93	– 74.80	0.85	– 77.10	0.72	– 78.70	0.51	– 78.70	0.07	– 80.70
6	0.91	– 89.50	0.83	– 93.10	0.69	– 95.70	0.48	– 95.70	0.04	– 102.30
7	0.89	– 106.60	0.80	– 110.50	0.66	– 112.70	0.45	– 114.00	0.02	158.01
8	0.88	– 123.40	0.79	– 129.40	0.64	– 132.40	0.43	– 135.40	0.06	118.40
9	0.86	– 143.20	0.76	– 150.20	0.62	– 154.20	0.40	– 161.20	0.12	96.20
10	0.83	– 166.10	0.72	– 174.10	0.58	– 179.10	0.37	171.10	0.19	72.10
11	0.82	168.10	0.71	158.10	0.59	153.80	0.39	140.80	0.25	62.60
12	0.80	138.20	0.72	127.20	0.60	121.20	0.44	108.20	0.33	49.20

$S_{11} = f(f, I)$



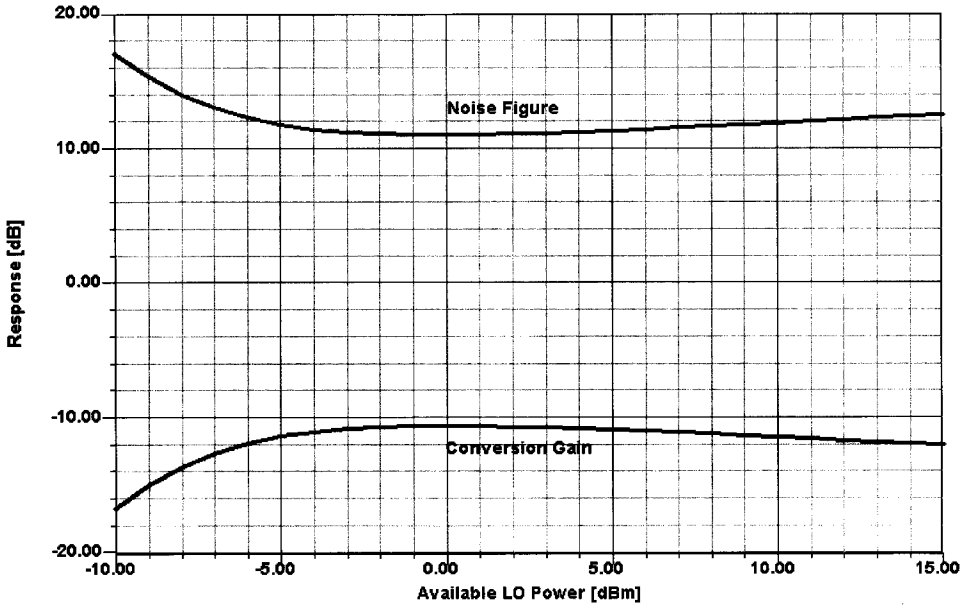


Figure 4-15 How the nonideal mixer's conversion gain and noise figure vary with available LO power. In this analysis, LO = 500 MHz (13 dBm), RF = 500.455 MHz (–20 dBm), and IF = 455 kHz.

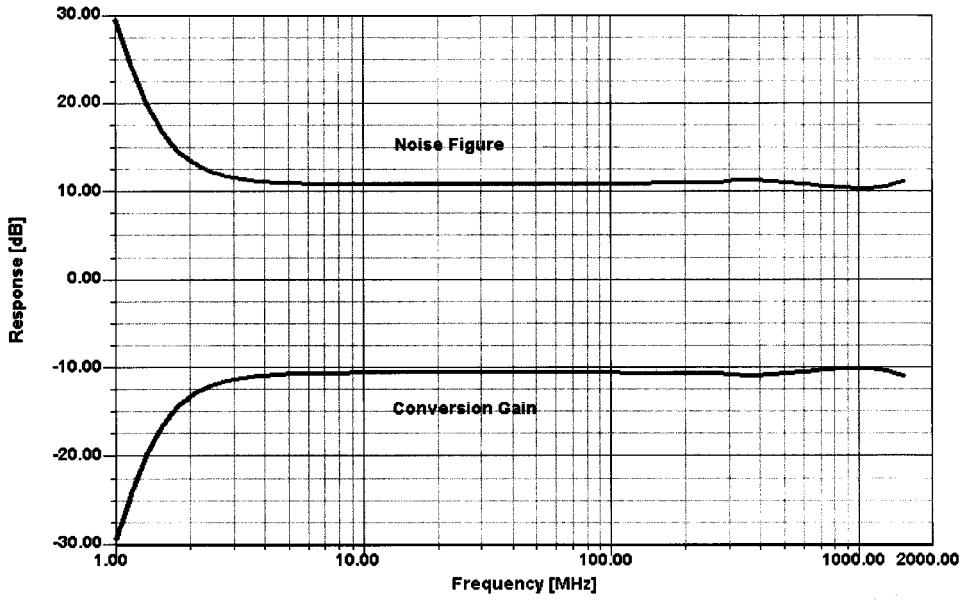


Figure 4-16 How the nonideal two-diode mixer's conversion gain and noise figure vary with frequency for a constant LO power. In this analysis, LO = 1–1500 MHz (2 dBm), RF = 1.455–1500.455 MHz (–40 dBm), and IF = 455 kHz.

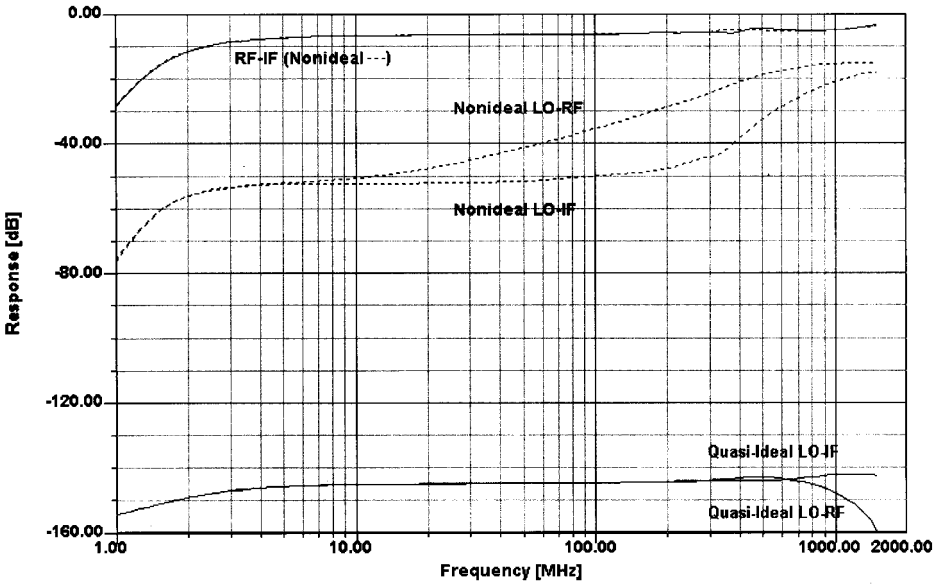


Figure 4-17 Port-to-port isolation of the quasi-ideal (identical diodes and no stray capacitance) and nonideal (slightly mismatched diodes and 0.5 pF of stray capacitance between the upper terminal of the middle transformer winding and ground) two-diode mixer. In this analysis, LO = 1–1500 MHz (2 dBm), RF = 1.455–1500.455 MHz (–40 dBm), and IF = 455 kHz.

connected from port to port) and nonideal port-to-port isolation (slightly mismatched diodes and 0.5 pF between the upper terminal of the middle winding and ground).

Figures 4-18a and 4-18b show the mixer's output spectrum for the quasi-ideal and nonideal cases, respectively.

Subharmonically Pumped Single-Balanced Mixer. Figure 4-19 shows a single-balanced mixer with a difference: antiparallel diode pairs take the place of single diodes, the RF and IF are buffered from each other only by filtering, and the LO is applied at one-half the frequency necessary to provide the desired frequency conversion. The RF-to-IF isolation is limited to that provided by the series input and output filtering, but the LO-to-IF isolation is higher at f_{LO} and much higher at $2f_{LO}$ than that achievable with a DBM with the LO signal at $2f_{LO}$(Figure 4-20). Although the example shown is for an up converting HF receiver, this technique finds application well into the microwave range as the basis for I/Q modulators, in which carrier leakage must be reduced to a level difficult to achieve with conventional DBMs [5–10].

4-3-3 Diode-Ring Mixer

Adding two more diodes and another transformer to the single-balanced mixer results in a double-balanced mixer (DBM) as shown in Figure 4-21. A DBM's frequency response is largely determined by the frequency response of its transformers, which act as transmission lines. The low-frequency limit is determined by the inductance of the transformer windings, the reactance of which, at the lowest frequency of interest, should be at least four times the impedance at which the transformer operates. The upper frequency limit is determined

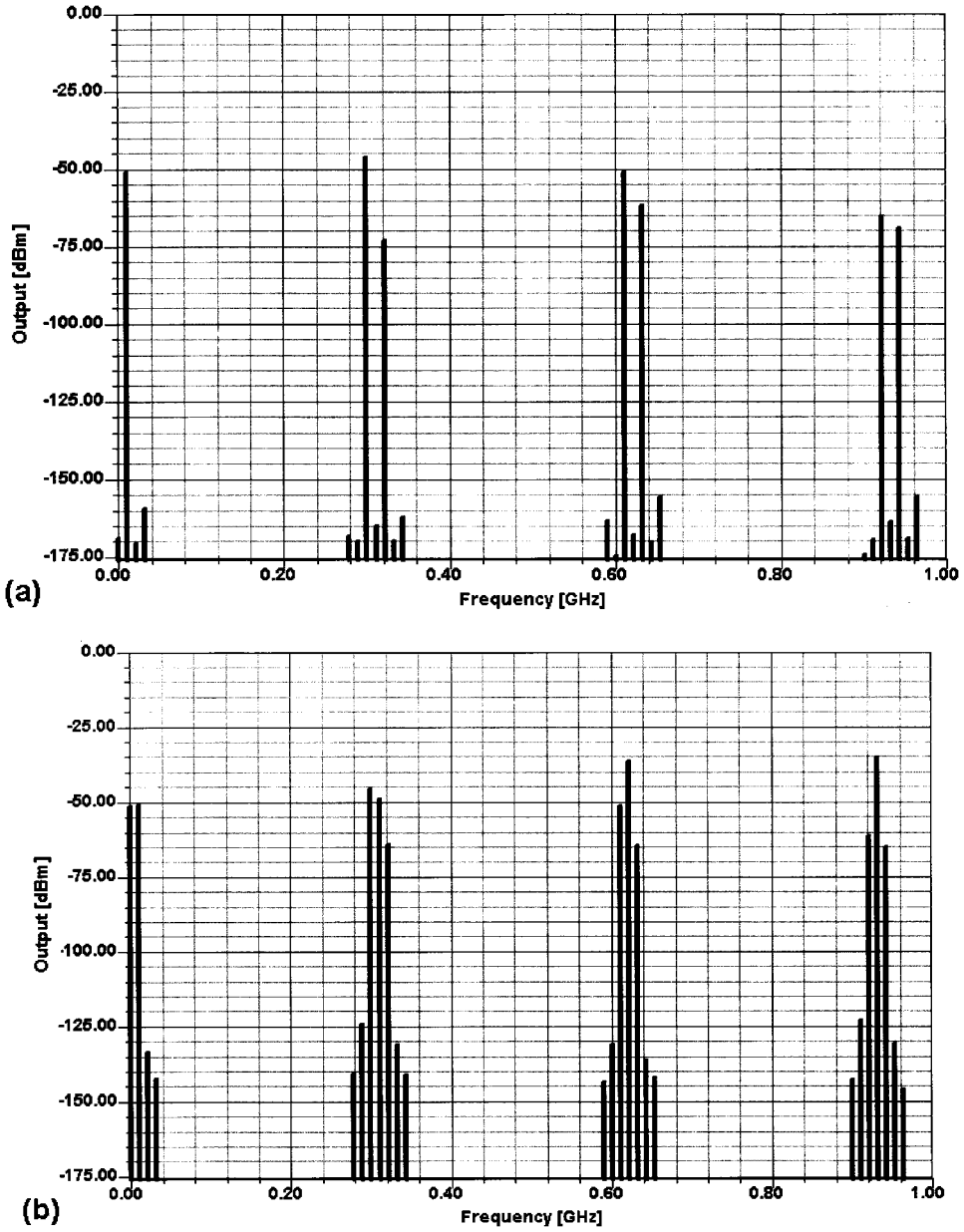


Figure 4-18 The nonideal mixer's output spectrum with (a) identical diodes and no stray capacitance and (b) slightly mismatched diodes and 0.5 pF of stray capacitance between the upper terminal of the middle transformer winding and ground. In these analyses, LO = 310.7 MHz (2 dBm), RF = 300.0 MHz (–40 dBm), and IF = 10.7 MHz. Four LO harmonics and 3 LO sidebands were used.

mainly by the degradation of the transformers' transmission-line behavior at higher frequencies, although the increasing importance of diode capacitance also plays a role.

A DBM's interport isolation is determined by the symmetry of its transformers, diodes, and physical construction. In practice, the effects of diode mismatch can be minimized by

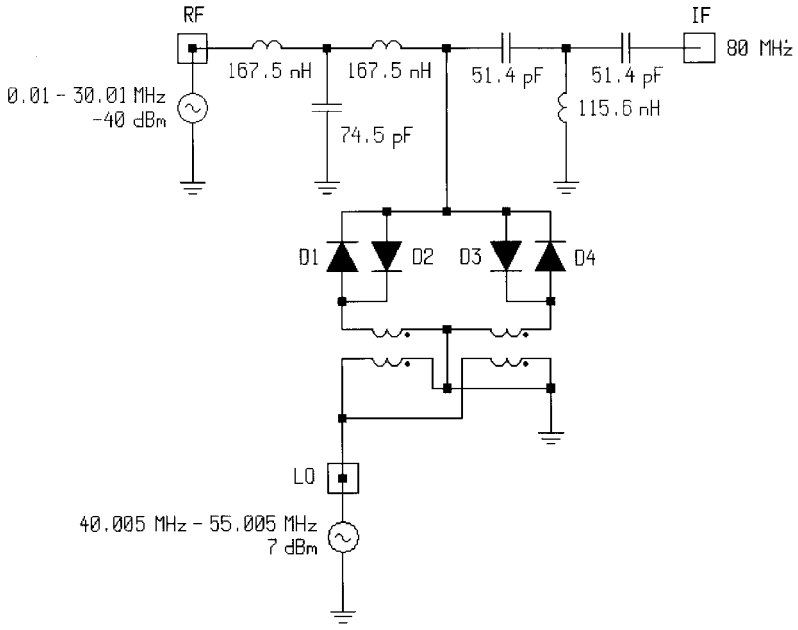


Figure 4-19 A subharmonically pumped single-balanced mixer using antiparallel diode pairs. The LO operates from 40.005 to 55.005 MHz to mix 0.01–30.01-MHz RF to an IF of 80 MHz.

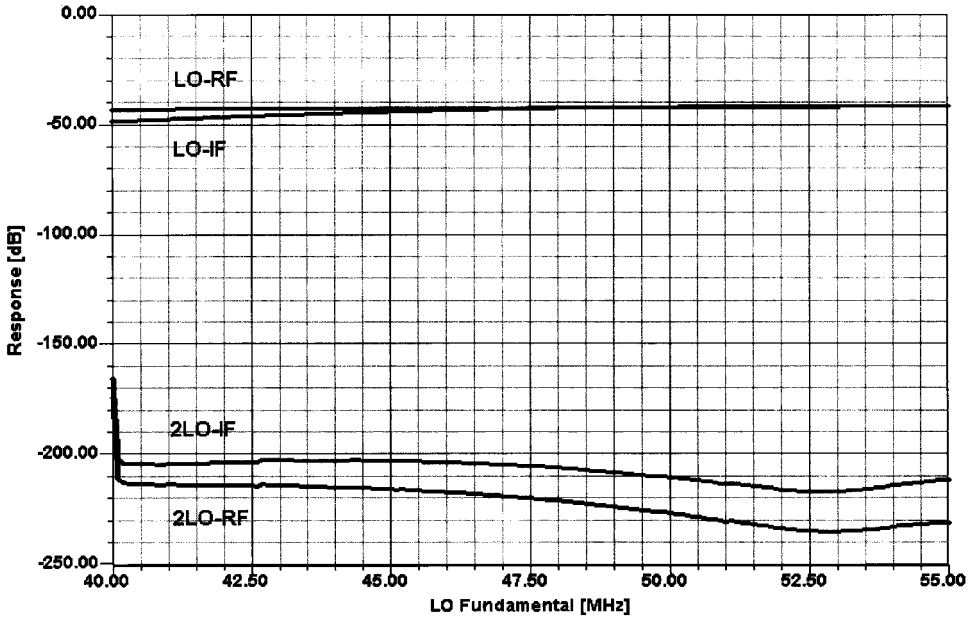


Figure 4-20 Simulated interport isolation of the subharmonic SBM. For realism, the diodes and transformers are slightly mismatched.

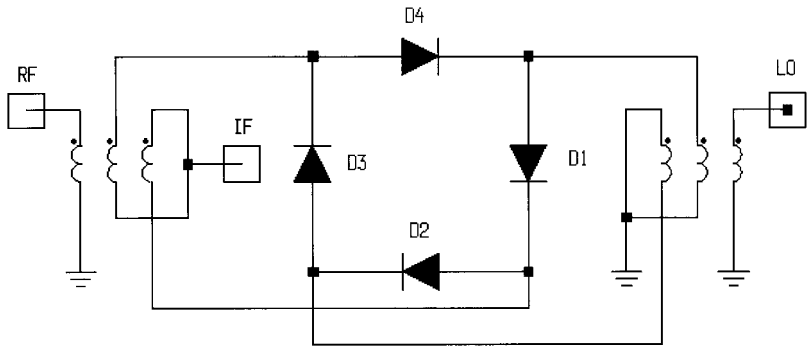


Figure 4-21 Schematic of the diode-ring double-balanced mixer.

using a dual diode, such as—in the case of the BAT15—the BAT15-099, the maximum V_F spread between the diodes being specified as 20 mV.

Figure 4-22 shows how the DBM's conversion gain and noise figure vary with applied LO power. Figure 4-23 shows how the DBM's conversion gain and noise figure vary with frequency for an LO power of 7 dBm, with quasi-ideal and nonideal balance. Figure 4-24 shows how the DBM's port-to-port isolation differs with quasi-ideal and nonideal balance for an LO power of 7 dBm. Figure 4-25 shows how the DBM's RF- and LO-port return losses vary with frequency; the sharp peak corresponds to a resonance caused by one of the stray capacitances added to simulate less-than-ideal balance in the modeled mixer.

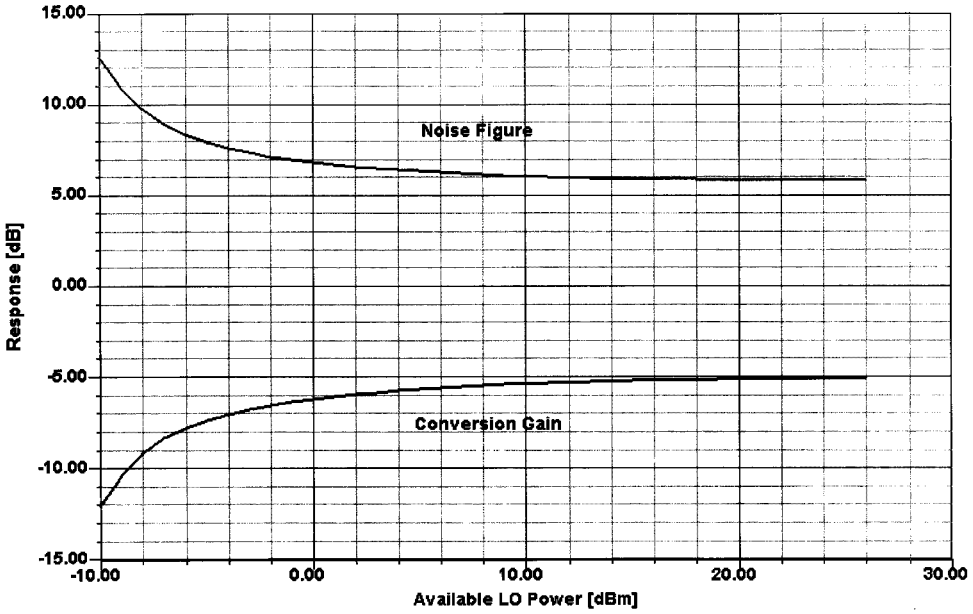


Figure 4-22 DBM conversion gain and noise figure versus LO power. In this analysis, LO = 310.7 MHz (–10 to 26 dBm), RF = 300 MHz (–40 dBm) and IF = 10.7 MHz.

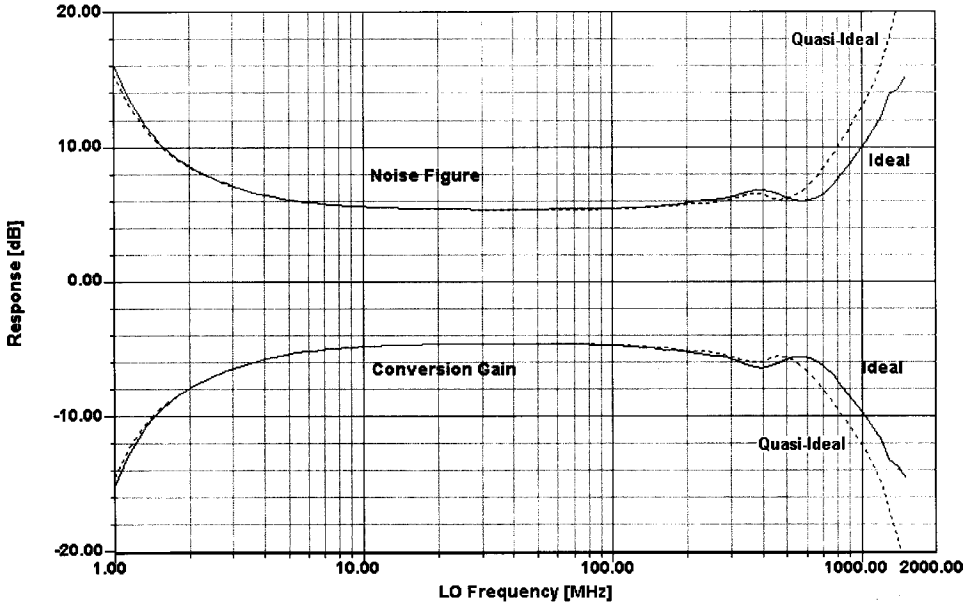


Figure 4-23 This plot of conversion gain (CG) and noise figure (NF) versus frequency for quasi-ideally and nonideally balanced versions of the same DBM reveals that balance plays a relatively minor role in the CG and NF performance achieved. In these analyses, the LO (–7 dBm) sweeps from 1 to 1500 MHz and the RF (–40 dBm) sweeps from 1.455 to 1500.455 MHz to produce an IF of 455 kHz.

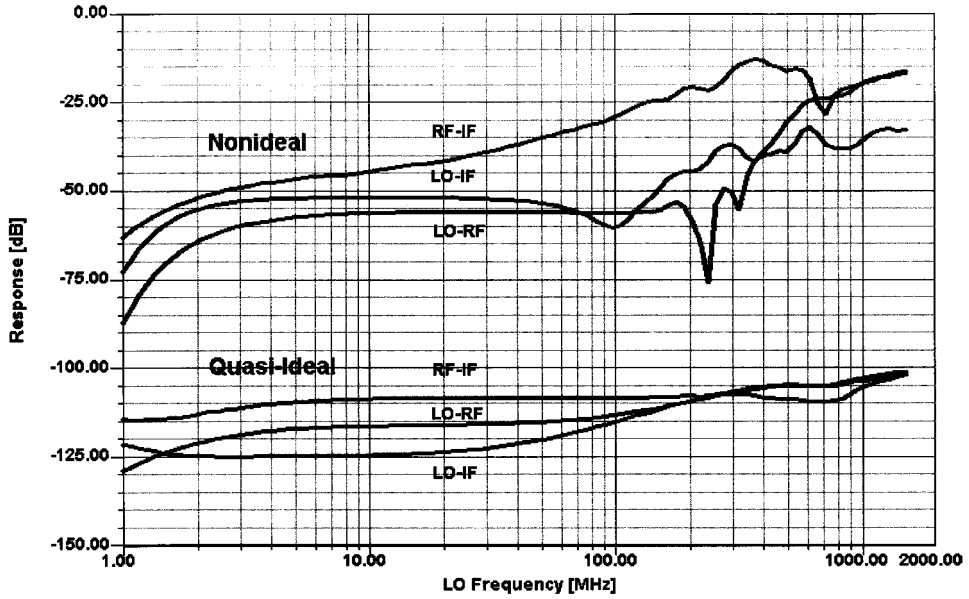


Figure 4-24 Interport isolation for DBMs with quasi-ideal and nonideal balance. In these analyses, the LO (–7 dBm) sweeps from 1 to 1500 MHz and the RF (–40 dBm) sweeps from 1.455 to 1500.455 MHz to produce an IF of 455 kHz.

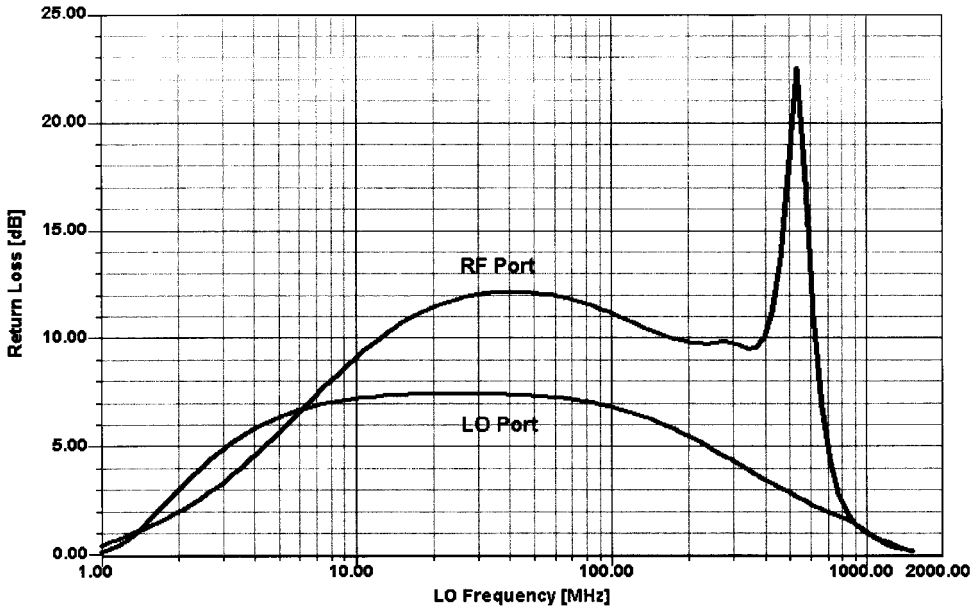


Figure 4-25 Return loss versus frequency for the DBM's RF and LO ports. The sharp peak results from a stray resonance.

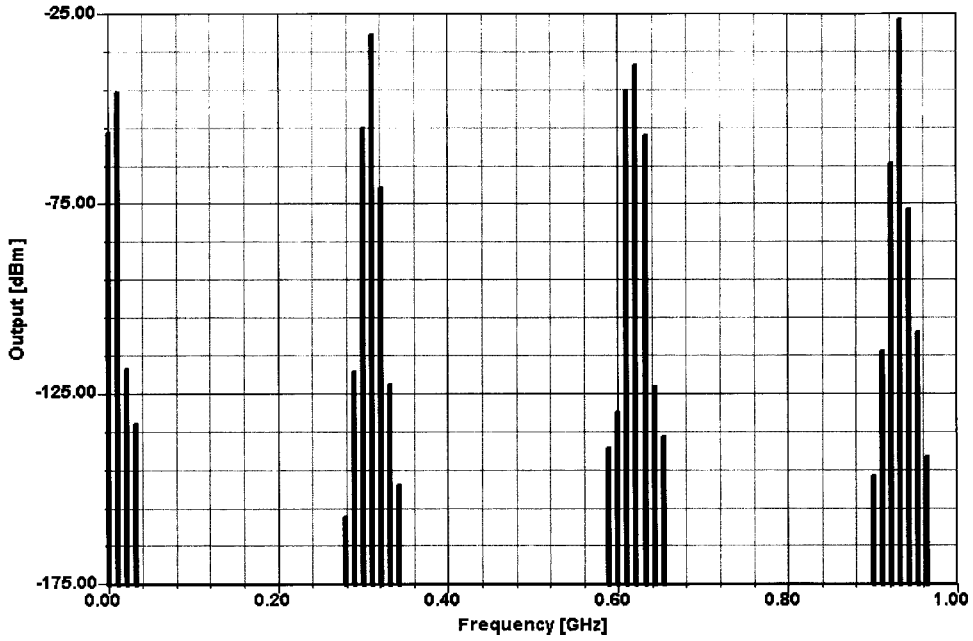


Figure 4-26 Output spectrum of a nonideally balanced DBM. In this analysis, LO = 310.7 MHz (–7 dBm) and RF = 300 MHz (–40 dBm) for an IF of 10.7 MHz. Four LO harmonics and three LO sidebands were used.

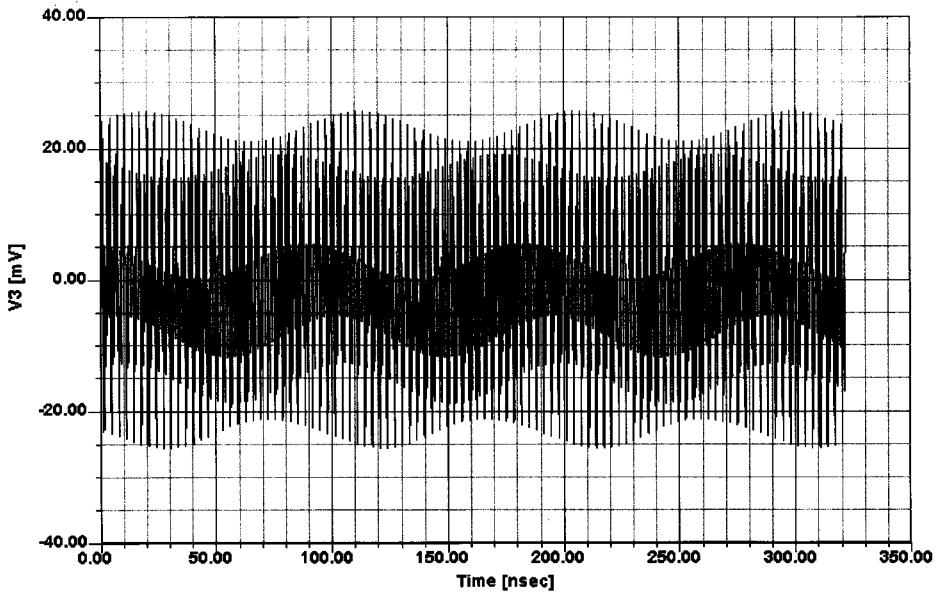


Figure 4-27 IF-port voltage waveform of the DBM over 100 cycles of the LO signal. The 310.7-MHz LO and 10.7-MHz IF components are clearly evident.

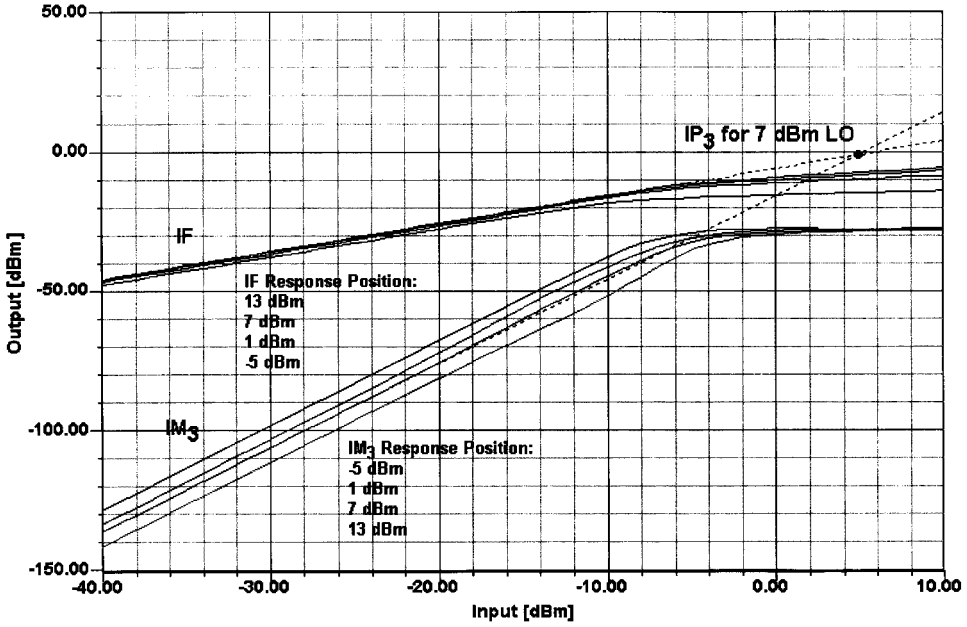


Figure 4-28 Diode DBM IF and IM_3 outputs versus RF power for four LO-drive levels. The responses for LO = 7 dBm have been extrapolated to show IP_3. Within limits, varying a mixer's LO drive affects its linear IF output relatively little while significantly affecting IMD. See Figure 4-29. For all four analyses, LO = 310.7 MHz, RF1 = 300.0 MHz (–40 to 10 dBm), and RF2 = 300.3 MHz (–40 to 10 dBm); four LO harmonics and three LO sidebands were used.

Figure 4-26 shows the DBM's output spectrum. Figure 4-27 shows the DBM's output waveform over 50 cycles of the LO signal.

Two-tone testing of the DBM allows us to characterize its IP_3 figures of merit. Figure 4-28 shows the nonideal DBM's IF and IM_3 responses for LO powers of –5, 1, 7, and 13 dBm. Figure 4-29 details how the DBM's IP_3 increases with LO drive, and Figure 4-30 shows the desired IF outputs and close third-order spurs near 10.7 MHz. Figure 4-31 shows the DBM's output voltage over 100 cycles of the LO signal, and Figure 4-32 shows the anode–cathode voltage of one of the ring's diodes, also over 100 LO cycles, both under two-tone IMD test conditions.

A double-balanced mixer, unless it is termination insensitive, is extremely sensitive to nonresistive termination. This is because the transmission-line transformers do not operate properly when they are not properly terminated, and the reflected power generates high voltage across the diodes. This effect results in much higher distortion levels than in a properly terminated transformer.

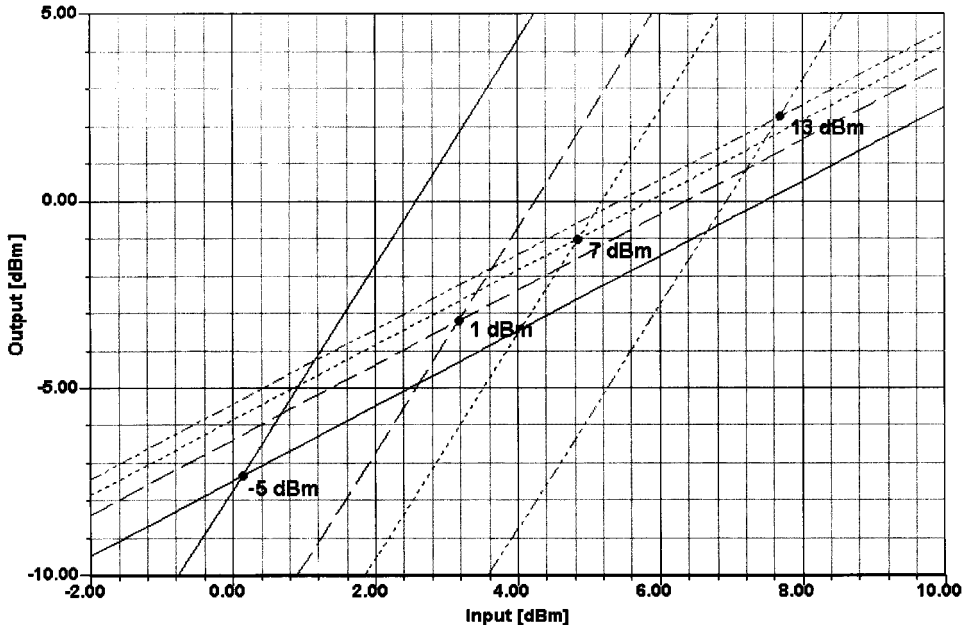


Figure 4-29 Extrapolating the responses for all four LO levels represented in Figure 4-28 shows how varying a diode DBM's LO drive shifts its third-order intercept point. Although these curves indicate that the simulated mixer's IP_3 generally increases with LO drive, the improvement in IP_3 is not as great as we might expect. The reason for this is that these four analyses, as well as the other diode-mixer analyses in this chapter, were done using diode models with a threshold voltage (V_J) of 0.23. If high-level diodes with a V_J of about 0.8 V had been used, $IP_{3,out}$ for the 13-dBm LO case shown here would increase to +13 dBm. $IP_{3,in}$ for the 13-dBm LO case would turn out to be 13 dBm + insertion loss = 13 dBm + 7 dB = 20 dBm. The issue of diode damage aside, attempting to increase IP_3 merely by driving a low- or medium-barrier diode harder eventually results in diminishing returns. High-barrier diodes are essential in getting the best IP_3 performance with high LO drive.

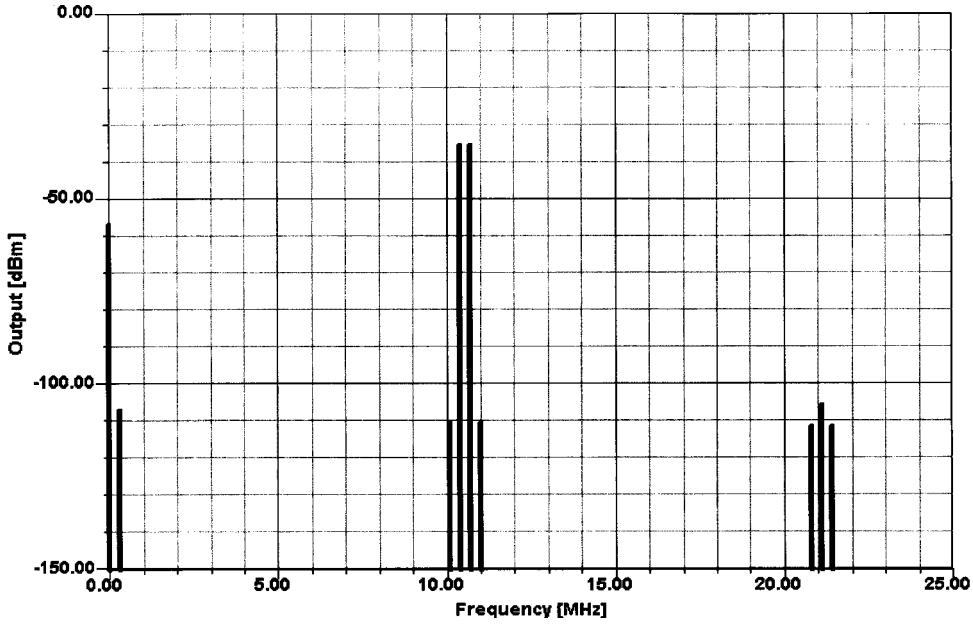


Figure 4-30 The DBM's output in the 11-MHz region during two-tone testing. The third-order products are clearly visible above and below the desired output signals. The test conditions for this analysis are those for Figure 4-28 with LO = 13 dBm.

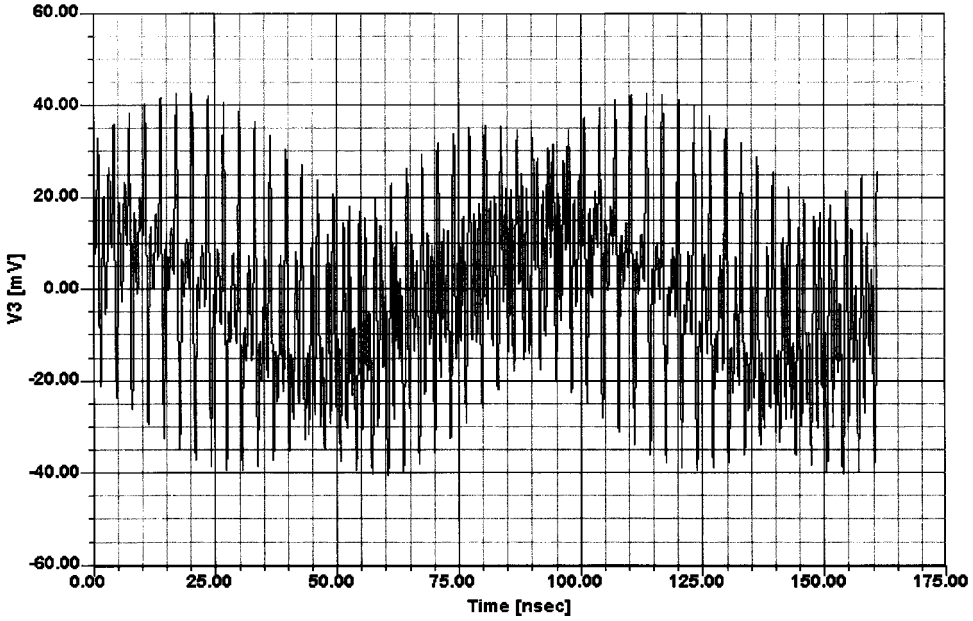


Figure 4-31 The DBM's IF-output voltage over 50 cycles of the LO signal. The test conditions for this analysis are those for Figure 4-28 with LO = 13 dBm.

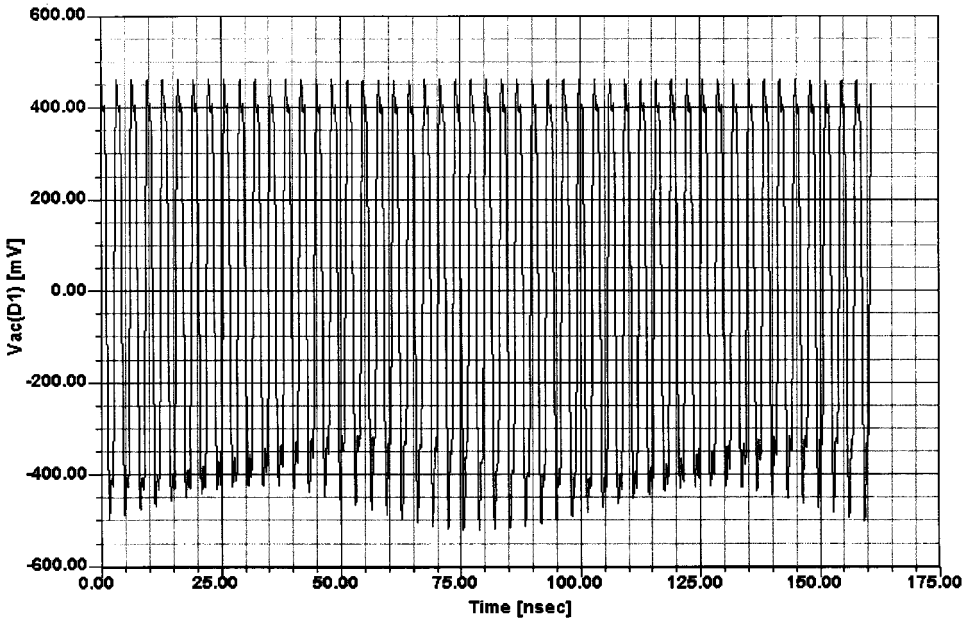


Figure 4-32 The anode-cathode voltage of one of the DBM's diodes, also over 50 LO cycles under two-tone IMD test conditions. The test conditions for this analysis are those for Figure 4-28 with LO = 13 dBm.

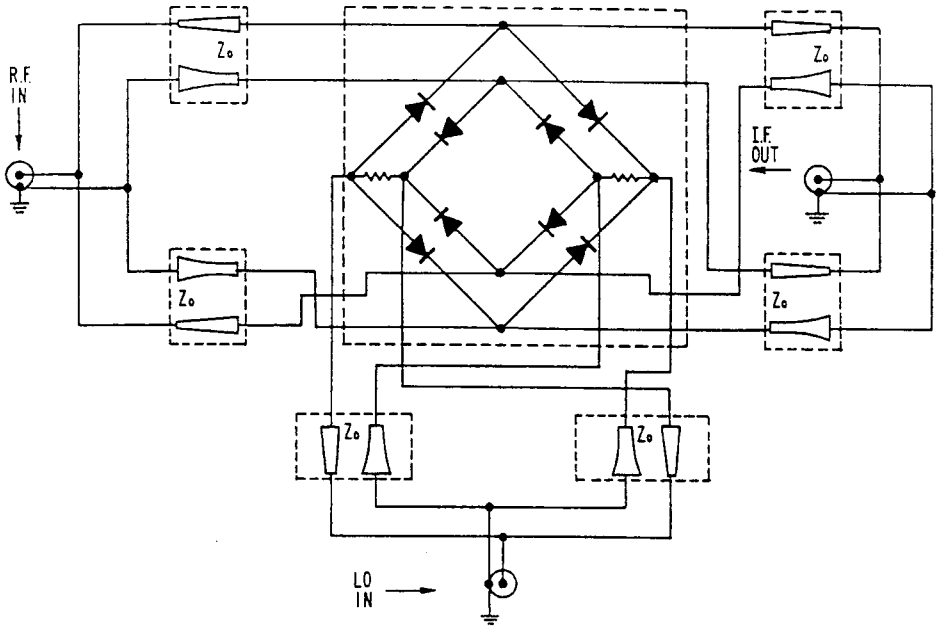


Figure 4-33 Example of a termination-insensitive mixer from U.S. Patent No. 4,224,572 (Adams-Russell Co., 1980) [11].

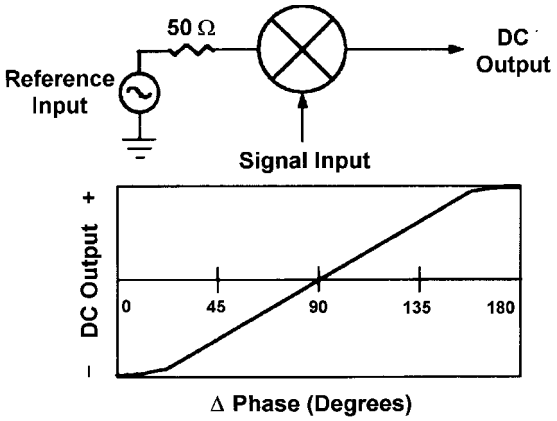


Figure 4-34 A mixer with a dc-coupled IF port can be used as a phase detector.

Termination-Insensitive Mixer. Figure 4-33 shows a mixer circuit that tolerates a fairly high VSWR at its output without significant degradation of its third-order IM performance [11].

Phase Detector. Theoretically, any mixer with a dc-coupled IF port can be used as a phase detector. When two signals of equal frequency are applied simultaneously to the reference and incoming signal ports, the phase detector produces a dc output at the IF port proportional to the cosine of the phase difference (Figure 4-34).

Binary Phase Shift Keying (BPSK) Modulator. Binary phase modulation occurs when a positive and negative signal current shifts the RF carrier between 0° and 180°. Figure 4-35 shows a double-balanced mixer operating as a BPSK modulator.

Quadrature Phase Shift Keying (QPSK) Modulator. A typical QPSK modulator consists of two biphase modulators, a 90° divider, and a 0° power combiner as shown. Data inputs at the control ports will cause the carrier to shift between 0°, 90°, 180°, and 270° as shown in Figure 4-36.

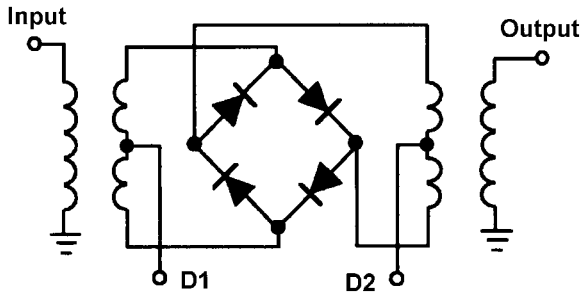


Figure 4-35 A diode-ring mixer as phase modulator.

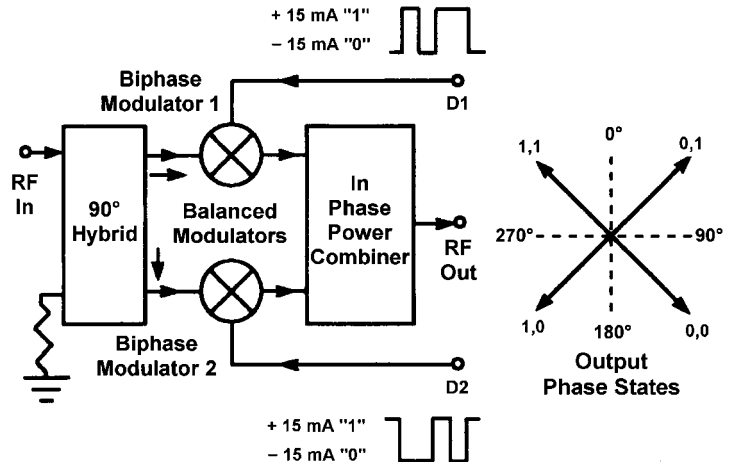


Figure 4-36 Two biphase modulators form the basis for a QPSK modulator.

Quadrature IF Mixer. A quadrature IF mixer produces two IF outputs in phase quadrature. Its basic structure consists of two double-balanced mixers, a 90° splitter, and a 0° splitter. The basic block diagram is shown in Figure 4-37.

Image-Reject Mixer. The image-reject mixer consists of a basic quadrature IF mixer with an additional 90° hybrid at the IF ports as shown in Figure 4-38. The primary function is to differentiate between the real signal and the image signal. This type of device is especially useful in applications where the desired RF signal and image are so close in frequency that rejecting the image with filtering is not practical.

Diode Attenuator/Switch. A ring of PIN diodes can be used as electronic attenuators by applying variable forward bias to the diodes (Figure 4-39). Maximum attenuation is achieved when the current at the control port is zero. The maximum attenuation is the isolation between the input and output ports. Minimum attenuation (insertion loss) is achieved when the IF-port current is 20-mA.

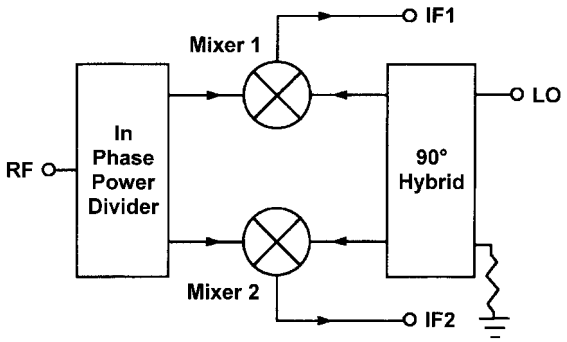


Figure 4-37 Quadrature mixer.

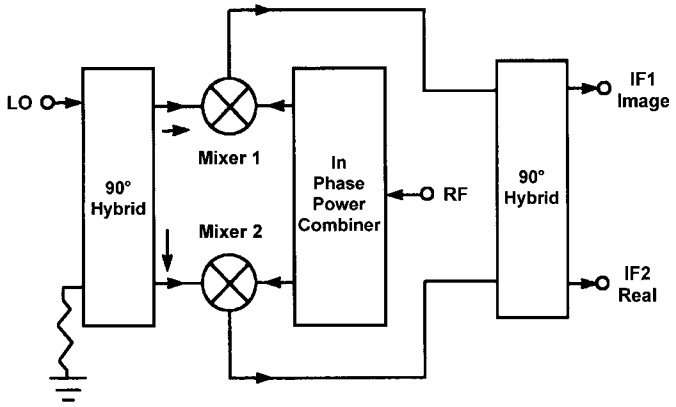


Figure 4-38 An image-reject mixer uses phasing to differentiate between its LO + IF and LO – RF IF outputs.

Single-Sideband (SSB) or In-Phase/Quadrature (I/Q) Modulator. SSB or I/Q modulators are useful in discriminating and removing the lower sideband (LSB) or upper sideband (USB) generated during frequency conversion, especially when the sidebands are very close in frequency and attenuation of one of the sidebands cannot be achieved with filtering. This is the case with audio and video modulation, where signals from dc to 10 MHz must be converted to a higher frequency that is appropriate for transmission. In such cases, both sidebands will be very close in frequency to the carrier frequency. With an I/Q modulator, one of the sidebands is easily canceled or attenuated along with its carrier.

Attenuation of the carrier has been the most troublesome aspect in the design of passive I/Q modulators. Isolation between the local oscillator (LO) port and the RF port of the mixers, which is the main parameter in determining carrier rejection, is usually insufficient at frequencies above 200 MHz.

I/Q modulator designs basically consist of two double-balanced mixers (Figure 4-40). The mixers are fed at the LO ports by a carrier phase-shifted through a 90° hybrid. Thus, the carrier signal's relative phase is 0° to one mixer and 90° to the other mixer. Modulation signals are fed externally in phase quadrature to the two mixers' IF ports. The mixers' modulated output signals are combined through a two-way, in-phase power divider/combiner.

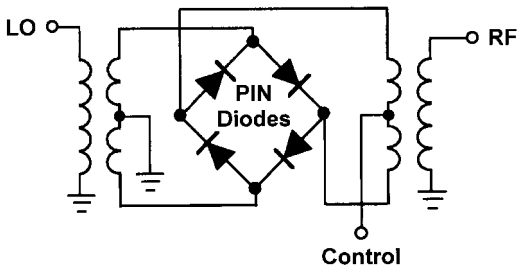


Figure 4-39 A diode DBM can be used as a dc-controlled attenuator if PIN diodes are used instead of Schottky devices in its ring.

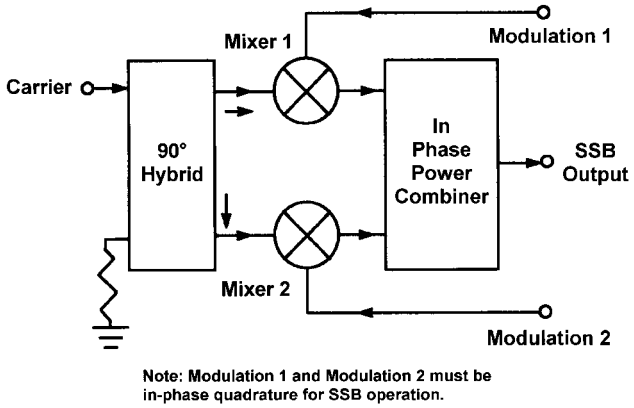


Figure 4-40 An SSB modulator matches two high-frequency mixers, a 90° hybrid, and an in-phase power combiner to produce an SSB output signal.

The circuit forms a phase-cancellation network to one of the sidebands and a phase-addition network to the other sideband. The carrier is somewhat attenuated and is directly dependent on the inherent LO-to-RF isolation of the mixers and the modulating signal level. In industry-standard I/Q modulators, USB suppression results when the first modulation port (Modulation 1) is fed with a signal that is 90° in advance of the signal feeding the second modulation port (Modulation 2). Opposite phasing can be arranged by changing the internal phase polarity of the mixers or by interchanging the 90° hybrid output ports to the LO ports of the mixers.

The phase and amplitude imbalances between the various components used in the manufacturing of the I/Q modulators must be tightly maintained for optimum SSB rejection. Matching of the two mixers for conversion loss and insertion phase is extremely critical, since differences in these parameters will add to amplitude- and phase-imbalance errors. The 90° hybrid in the LO port must be in nearly perfect phase quadrature.

Phase- and amplitude-imbalance errors adversely affect sideband suppression (Figure 4-41). In most cases, a typical passive I/Q modulator operates with a carrier input level of +10 dBm, which is required to drive the diodes in the mixers to operate in the linear range. The dynamic range of these mixers can be improved significantly by using diodes with a higher barrier height. The LO signal in this case must be increased in order to drive these diodes into conduction in their linear range.

Carrier rejection is also a problem when designing an SSB modulator, since only a few decibels of suppression can be achieved in standard high-frequency models. In the past, the major contributor to carrier suppression was the inherent LO-to-RF isolation through the mixers. Unfortunately, this isolation is usually poor at cellular frequencies (800–1000 MHz), where at least 25 dB of carrier rejection is necessary. In some cases, designers feed a small amount of dc into the IF ports to control the carrier rejection, which complicates the driver circuitry and calls for temperature compensation when operating at different temperatures.

As an example, an SSB modulator is assumed to operate with +10-dBm LO drive with each modulating signal at –10 dBm and in phase quadrature to each other when applied to the modulating ports (MOD 1 and MOD 2). The result will be a modulated signal at –16 dBm, assuming 6-dB conversion loss. For 20-dB carrier rejection with respect to the desired

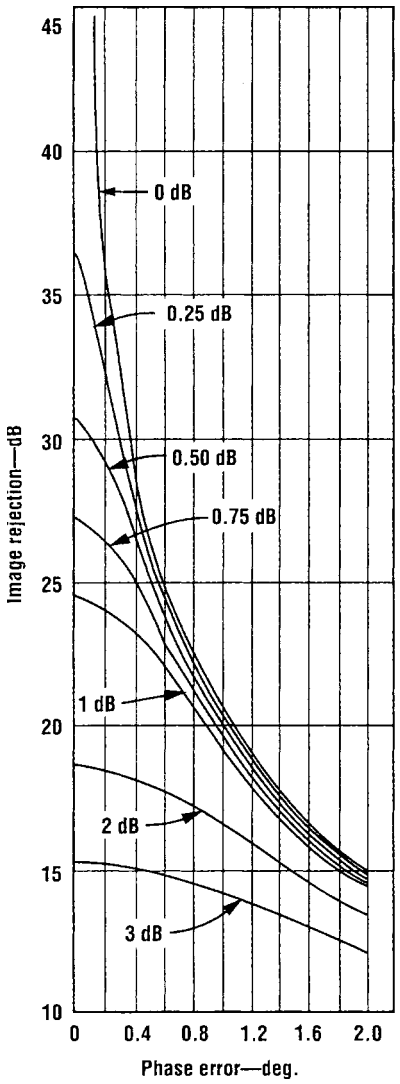


Figure 4-41 The level of SSB rejection improves as the phase- and amplitude-imbalance performance of an SSB modulator improves.

modulated signal, the carrier must be at –36 dBm, which translates to LO-to-RF isolation of 46 dB.

By employing a subharmonic approach, the performance of SSB modulators can be extended beyond the limits of conventional designs as reported by Joshi [7]. The approach is based on the use of subharmonic mixers in place of fundamental-frequency mixers and is applicable from about 140 to 3000 MHz. Subharmonic mixers use antiparallel diode pairs in their construction [12–14]. Matched antiparallel diode pairs used in single-ended or single-balanced mixer configurations cancel even-order intermodulation products (such as $2f_{LO} \times 2f_{RF}$, $3f_{LO} \times 3f_{RF}$, etc.) at all ports.

Single-ended mixers lack the port-to-port isolation needed for SSB modulator applications. Odd-order products of the RF and LO frequencies (even $f_{LO}\times$ odd f_{RF}) and (odd

$f_{LO} \times$ even f_{RF}) appear on all ports, requiring extensive filtering for satisfactory performance. For a single-balanced mixer, even harmonics of the LO combining with odd harmonics of the RF appear at the IF port, whereas odd harmonics of the LO combining with even harmonics of the RF appear at the RF and IF ports. This assumes that a balanced transformer is placed at the LO port, which is a logical choice due to the fact that the highest level signal appears at the LO port. Since the desired odd-order IF products appear at both the RF and IF ports, a need arises for a diplexing network to isolate the RF and IF signals.

The subharmonic modulator design provides a unique way to isolate the RF and IF signals. A single-balanced harmonic mixer offers good LO-to-RF and LO-to-IF isolation but poor RF-to-IF isolation. Fortunately, harmonically related signals are spaced well apart in the frequency spectrum, simplifying filtering of harmonically related signals.

Harmonic mixing also works well with low LO power levels, with somewhat lower 1-dB compression on the RF port than with fundamental-frequency mixing. The ability to operate with LO frequencies that are a fraction of the carrier frequency (1/2, 1/4, 1/6, etc.) significantly reduces the cost of an LO source, especially at higher frequencies. Also, using lower-frequency LO sources helps avoid the signal-leakage problems inherent with higher-frequency LO sources. Minimizing signal leakage, especially at higher frequencies, becomes expensive and bulky. Subharmonic mixing offers several advantages:

- The technique offers the ability to operate at LO frequencies that are 1/2, 1/4, or 1/6 of the carrier frequency. For example, for an IF of 100 MHz at an RF of 2 GHz, the LO can be $(2000 \pm 100) \div 2 = 950$ or 1050 MHz.
- The LO's even harmonics are strongly attenuated.
- The filtering requirements for fundamental frequency and odd harmonic signals of the LO are not critical.
- The cost of generating the LO is reduced due to the fact that the LO frequency need only be a fraction of the carrier frequency.

As an example of the performance improvements possible with the subharmonic mixers, units were evaluated at both cellular (935–960 MHz) and PCN/PCS (1.8–1.9 GHz) bands.

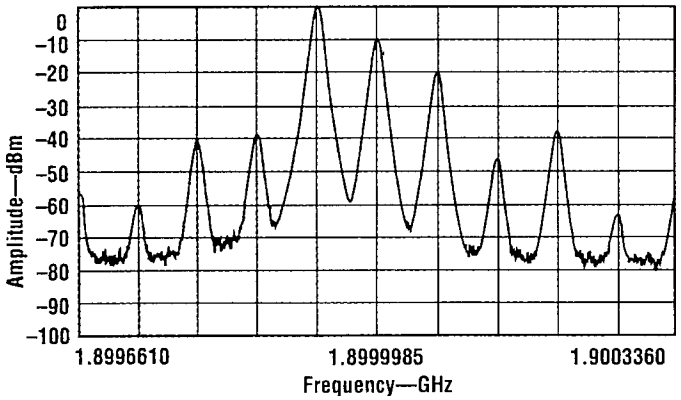


Figure 4-42 This plot of carrier and sideband rejection was measured for a conventional SSB modulator operating at 1.9 GHz.

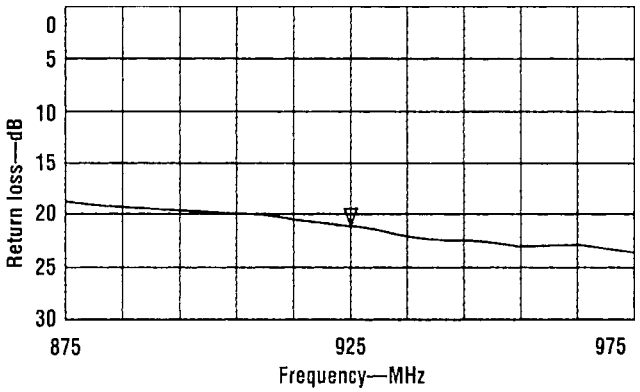


Figure 4-43 The SSB modulator's return loss as measured at the local oscillator (LO) port.

For a conventional SSB modulator at 1.9 GHz fed with +10-dBm modulation signals, carrier rejection is barely 10 dB (Figure 4-42).

Sideband rejection can be improved by tuning, but the carrier rejection is controlled by the LO-to-RF isolation of the double-balanced mixers. Conventional double-balanced mixers with high isolation at cellular and PCN bands are very expensive and large when special techniques are used to improve LO-to-RF isolation. In contrast, the subharmonic nature of the new approach allows the use of lower-frequency, less-expensive components in the modulators' construction.

The subharmonic modulators offer an improvement of more than 15 dB in carrier suppression compared to the conventional approach.

The measured VSWR (return loss) at the LO and RF ports is better than 1.50:1 (Figures 4-43 and 4-44). Measurements made on a cellular-band SSB modulator reveal carrier rejection on the order of 40 dB. Typical insertion loss is 7 dB while sideband rejection is 30 dB (Figure 4-45).

By virtue of harmonic mixing, even-order mixing products are attenuated by about 30 dB with respect to the desired modulated output signal. The fundamental-frequency feedthrough into the output port is approximately 5 dB lower than the desired modulated signal, whereas

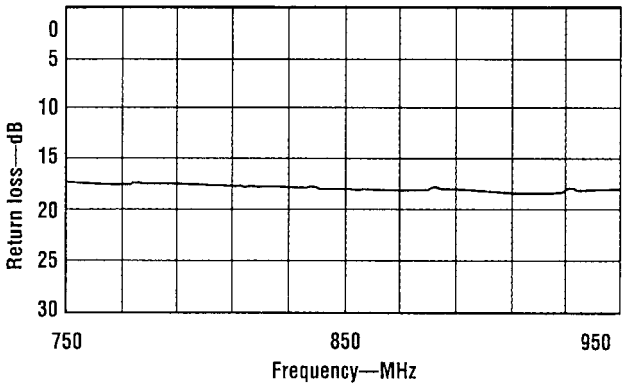


Figure 4-44 The novel harmonic SSB modulator's return loss as measured at the RF port.

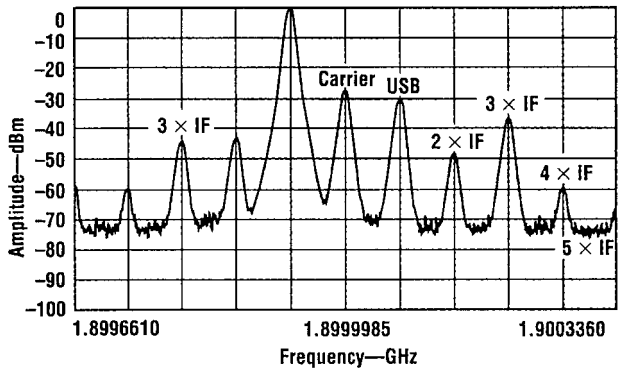


Figure 4-45 Plot of carrier and sideband rejection as measured for the novel harmonic SSB modulator operating at cellular frequencies.

the fourth harmonic mixing with the modulating signal is approximately 10 dB lower. Typical loss for fourth harmonic mixing is 17–19 dB while maintaining 30 dB of carrier rejection.

Since harmonically related products are well-spaced in frequency, filtering undesired signals is relatively inexpensive using standard octave-bandwidth filters. Low-cost commercial bandpass filters typically offer better than 40–50-dB attenuation of unwanted harmonic signals. Constant-impedance bandpass filters offering good impedance match at desired

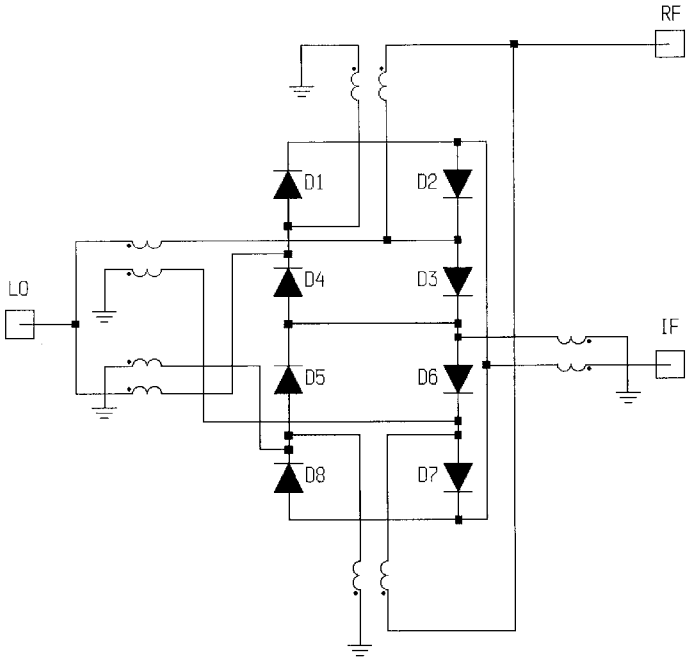


Figure 4-46 A triple-balanced diode mixer. A limitation of this configuration is that the internal dc common connections associated with its RF and LO transformers disallow usable IF response down to dc.

stopbands can also be used in cases where harmonically related products require impedance termination within a system.

The subharmonic modulator design is easily applied at custom frequencies. Conversion of an SSB modulator with output frequency corresponding to twice the LO frequency to one with output corresponding to four times the LO frequency requires only one component change, in the form of a signal-combining network at the modulator's output. Although the conversion loss of the fourth harmonic LO component mixing with the modulating signal is in the vicinity of 18 dB, the cost of generating the LO is drastically reduced with the subharmonic modulator. In spite of higher signal loss, the carrier rejection is still at least 30 dB at the fourth harmonic, and harmonically related products can be eliminated with an inexpensive filter.

Triple-Balanced Mixer. Two diode rings can be combined to form a double double-balanced mixer, or triple-balanced mixer (TBM), as shown in Figure 4-46. Triple-balanced mixers achieve higher dynamic range and interport isolation than double-balanced designs at the expense of LO power and increased complexity and size.

Figure 4-47 shows the interport isolation of the circuit in Figure 4-46 with the circuit configured in a less than ideally balanced form, with small variations in transformer-winding inductance and diode parameters introduced for more realistic modeling. Note that the mixer's interport isolation generally increases with frequency, rather than decreasing with frequency as with the DBM (Figure 4-24).

Rohde & Schwarz Subharmonically Pumped DBM. Figure 4-48 shows an example of a complete microwave diode mixer used as the input stage of a spectrum analyzer. Its LO is applied at 13.2 GHz, one-third the frequency necessary to mix the 40.1-GHz input

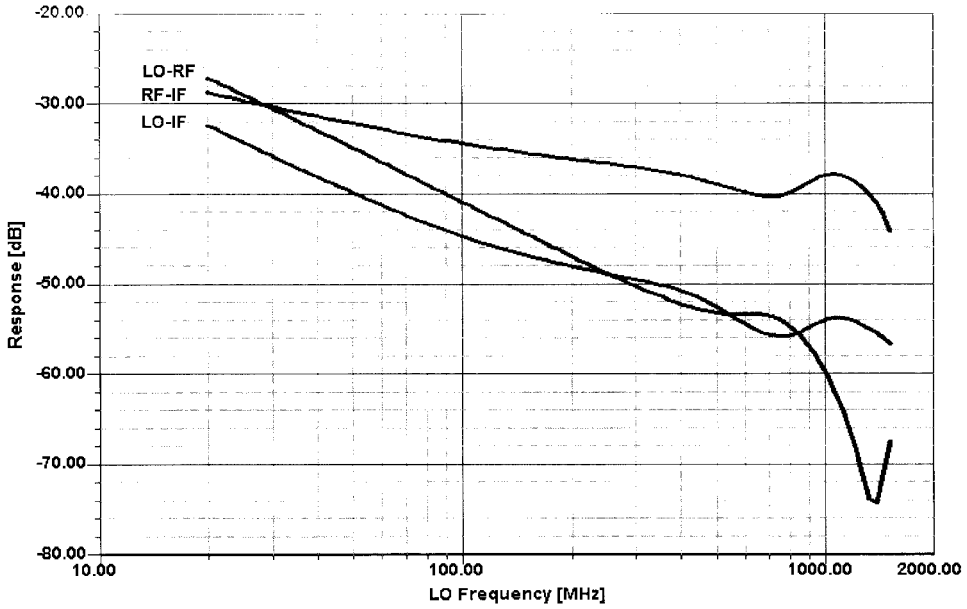


Figure 4-47 The triple-balanced mixer offers improved high-frequency isolation over a standard DBM.

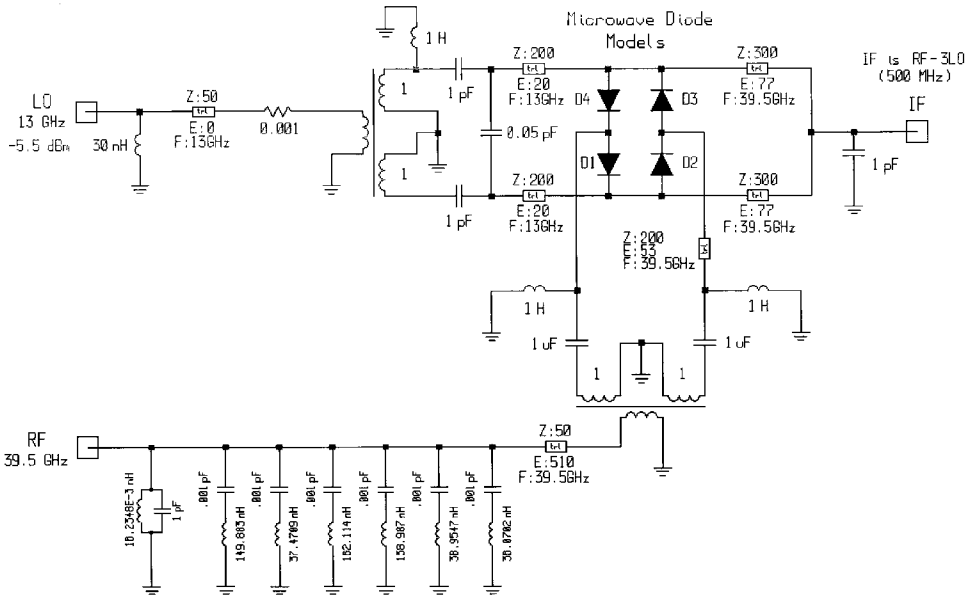


Figure 4-48 Schematic of the spectrum-analyzer mixer. The LO is applied at one-third the frequency necessary to mix the RF signal to the 500-MHz IF.

signal to the 500-MHz IF, with distortion in the diodes providing the 3× frequency multiplication. The circuit's conversion gain is –18.8 dB; its noise figure, 20.5 dB.

The subharmonic-drive technique exemplified by this mixer is important for another reason in addition to LO isolation. Depending on the application, directly generating a sufficiently phase-quiet LO signal at the higher wireless frequencies and on up into the microwave frequencies may not be feasible. In such cases, injecting the LO at a subharmonic of the desired LO frequency may provide better phase noise performance than a fundamental LO even though the phase noise of a frequency-multiplied source increases by $\log_{10} n$ dB, where n is the multiplication factor.

4-4 TRANSISTOR MIXERS

Diode mixers are lossy and termination-sensitive and require considerable LO power. Their attractiveness to designers of highly integrated wireless products is further reduced by their dependence on transformers for balance and port-to-port isolation. Transistor mixers are therefore used where high integration and reduced current drain are paramount—that is, in most non-base-station wireless applications.

Until ten years ago, single BJTs were commonly used in a simple additive-mixing arrangement, an example of which appears in Figure 4-49. Such a circuit behaves like a combination of a preamplifier and single-diode mixer. Single-BJT mixers were used in early AM–FM radios until 1980, and unfortunately in some handheld 2-meter and 70-cm ham equipment, to achieve the lowest possible power consumption. By definition, however, the presence of such a mixer also destroys any possibility of achieving a high intercept point. In

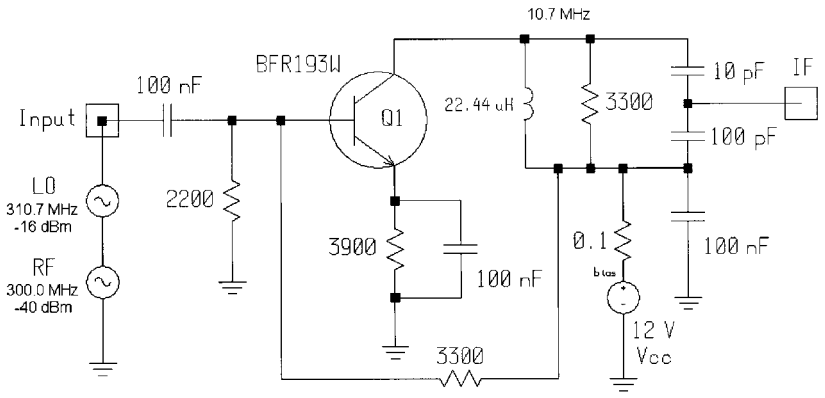


Figure 4-49 A single-BJT mixer.

addition to this, many combination frequencies occur despite elaborate input filtering. Although single-BJT mixers can exhibit considerable conversion gain (>10 dB), their dynamic range is restricted and their port-to-port isolation is poor. BJT mixers in today's competitive wireless designs use multiple transistors and are based almost exclusively on the Gilbert multiplier cell.

4-4-1 BJT Gilbert Cell

The classic active mixer (Figure 4-50), conceived by Barrie Gilbert in 1967, is the basis for most active mixers used in wireless products today. Figure 4-51 shows a Gilbert cell mixer implementation for analysis. Figure 4-52 shows its conversion gain and noise figure versus LO power. Figure 4-53 shows its IM_3 and IP_3 responses. As Gilbert wrote in a 1994 monograph [15]:

> This circuit is attractive because (a) it can be monolithically integrated with other signal-processing circuitry; (b) it can provide conversion gain, whereas a diode-ring mixer always has conversion loss; (c) it requires very low power to drive the LO port; (d) it provides excellent isolation between the signal ports; and (e) it is far less fussy about load-matching.

The two major advantages afforded by the Gilbert cell are port isolation (it should really be called a multiplicative mixer) and its significant reduction of even-order frequency combinations. As with the single-transistor mixer, the Gilbert cell does not particularly shine with high intercept point and low current consumption. Another major drawback is that it requires a high-impedance output, which gives the user headaches because most of the better filters are in 50-Ω technology, and only very simple monolithic filters offer input and output impedances around 1 kΩ. Impedances on this order can result in significant crosstalk in high-density circuit boards.

The reason the Gilbert cell can be so linear for small input signal levels is twofold: It is a differential amplifier and it makes use of the tanh function, which is linear over a wide range around the zero-crossing point [16]. Attempts have been made to increase the intercept point. The Plessey SL6440 IC was a special version of the Gilbert cell that achieved high $IP_{3,in}$ (+30 dBm) at the expense of dc current.

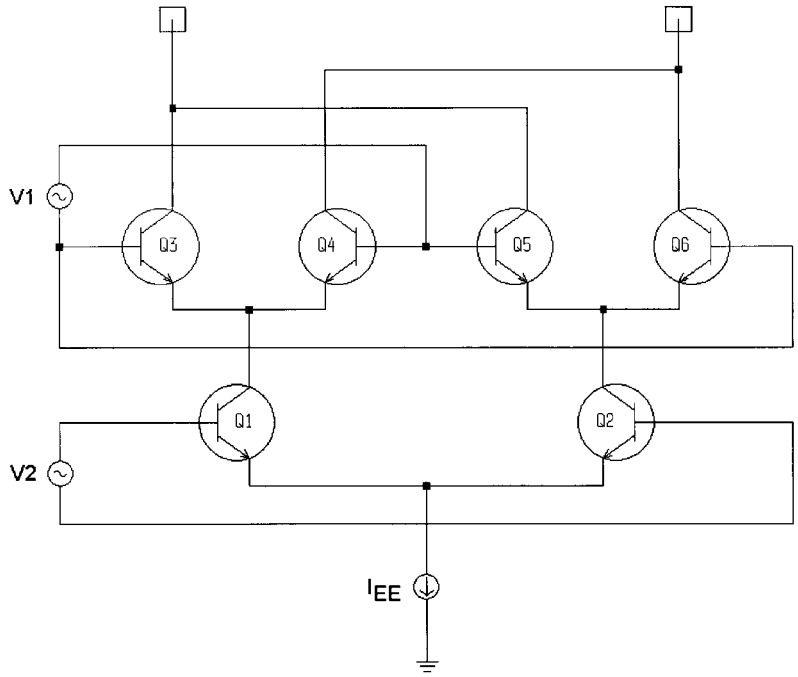


Figure 4-50 The basic Gilbert multiplier cell.

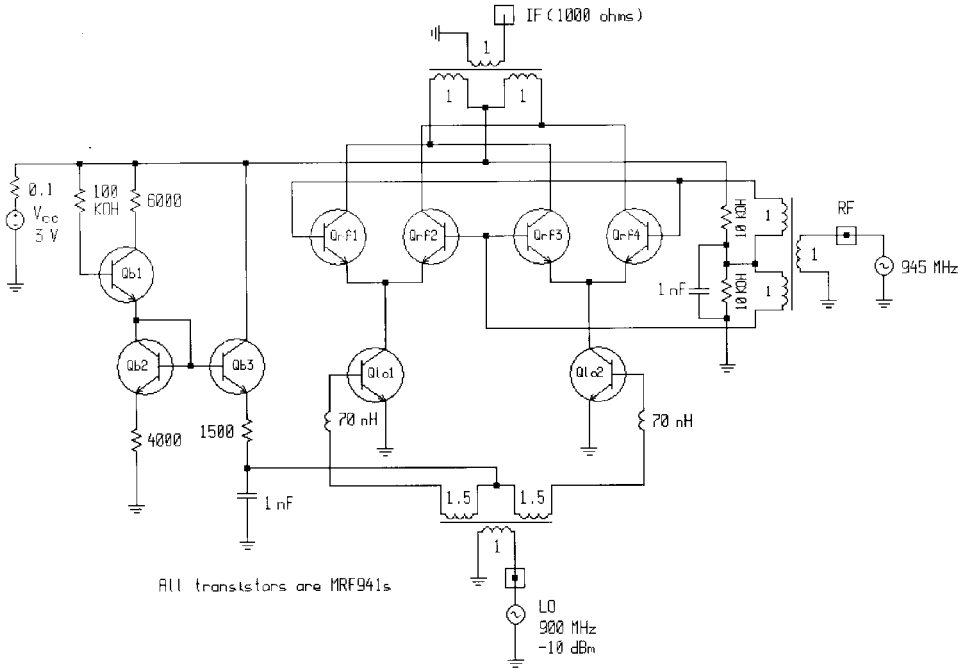


Figure 4-51 Gilbert cell mixer validation circuit.

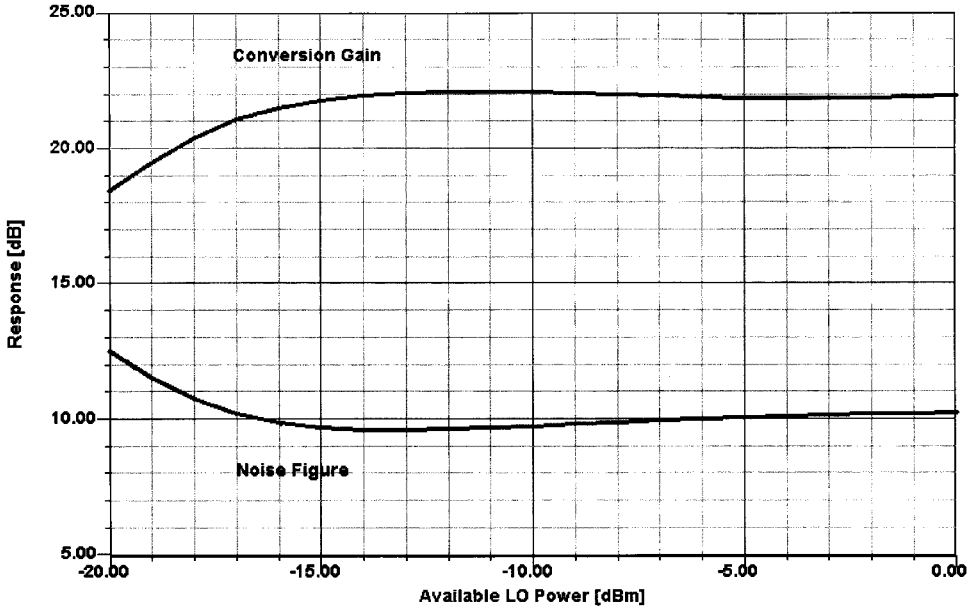


Figure 4-52 Conversion gain and noise figure versus LO power for the Gilbert cell mixer. In this analysis, LO = 900 MHz (–20 to 0 dBm) and RF = 945 MHz (–50 dBm) for an IF of 45 MHz.

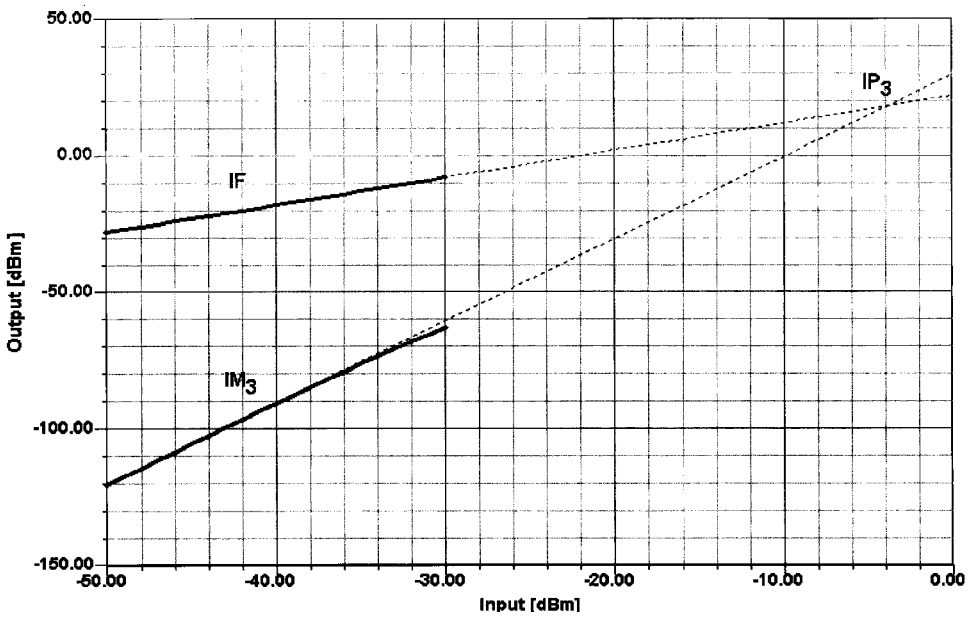


Figure 4-53 IM_3 and IP_3 responses for the BJT Gilbert cell mixer with RF signals at 945 and 946 MHz. In this analysis, LO = 900 MHz (–10 dBm), RF1 = 945 MHz (–50 to –30 dBm), and RF2 = 946 MHz (–50 to –30 dBm). Four LO harmonics and three LO sidebands were used.

Our suspicion, however, is that a combination of a modern double-balanced passive mixer using medium- to high-level diodes followed by a feedback FET amplifier, or even one of the late CATV transistors with f_T values of more than 25 GHz, would allow the achievement of a significantly higher intercept point. As pointed out in Chapter 3, CATV transistors are now available that combine a noise figure of less than 1 dB with a $IP_{3,in}$ of more than 30 dBm. Built-in push–pull structures using such transistors can achieve an $IP_{2,in}$ of 70 dBm, matching the numbers for the diode DBM. In Chapter 3 we learned that feedback always improves the dynamic range, mostly at the expense of gain, since the gain–bandwidth product gets reduced. Motorola has recently come out with a clever extension of its high-level preamplifier and modified it to become a mixer.

4-4-2 BJT Gilbert Cell with Feedback

Motorola's MC13143 low-power 2.4-GHz mixer IC uses a patented topology consisting of a Class-AB-biased Gilbert cell augmented by feedback. Its linearity can be programmed via

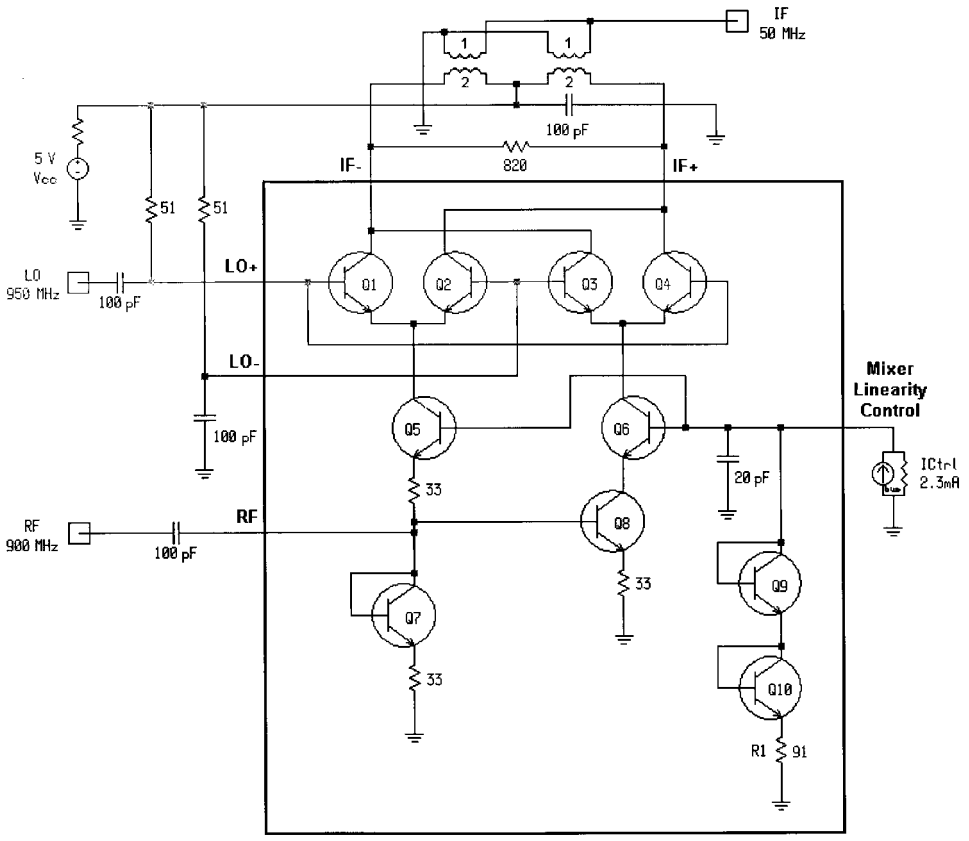


Figure 4-54 Motorola's MC13143 mixer IC uses dc feedback for improved linearity that can be programmed by applying a control current (0–2.3 mA) to its Mixer Linearity Control pin. Per Motorola, an $IP_{3,in}$ of 20 dBm may be achieved with a control current of 2.3 mA, at the expense of approximately 7 mA of additional supply current. In this validation circuit, R_1 (the resistor in Q10's emitter) has been adjusted for a total current drain (supply + linearity control) of 8 mA.

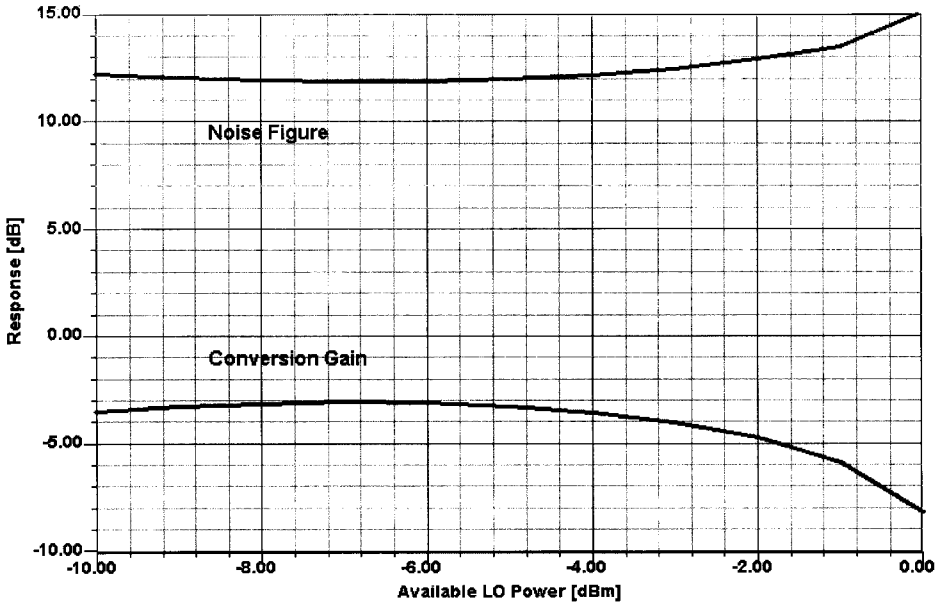


Figure 4-55 Calculated conversion gain and noise figure versus LO drive for the Motorola MC13143 IC operating at $V_{CC} = 2$. Motorola's specifications for the typical values of these characteristics are –5.0 dB and 12 dB, respectively, at an LO drive level of 0 dBm. The RF signal is at 900 MHz and the LO is at 950 MHz, for an IF of 50 MHz. In this analysis, LO = 950 MHz (–10 to 5 dBm) and RF = 901 MHz (–50 dBm) for an IF of 49 MHz.

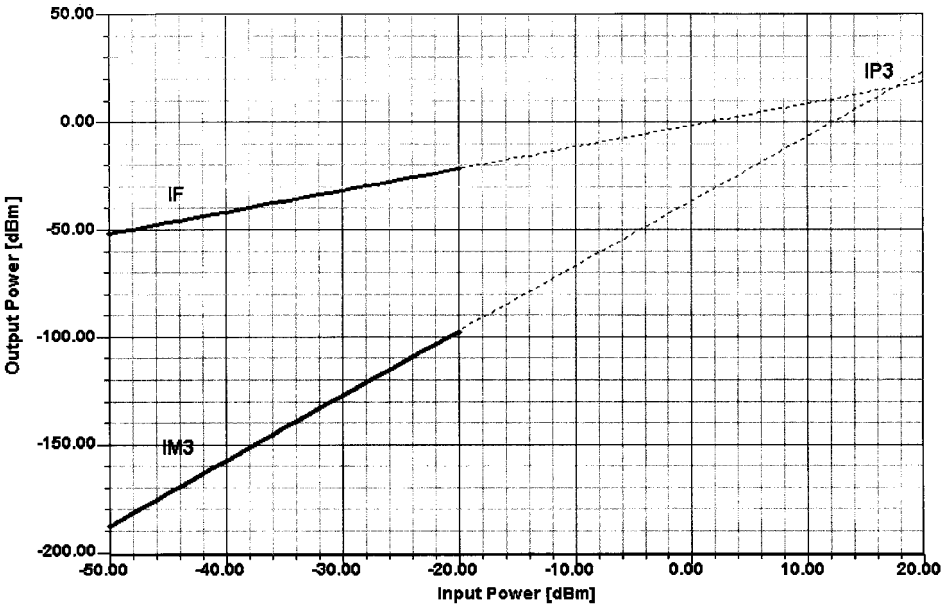


Figure 4-56 Calculated typical IF, IM_3, and IP_3 responses for the Motorola MC13143 mixer. In this analysis, LO = 950 MHz (–5 dBm), RF1 = 900 MHz (–50 to –20 dBm), and RF2 = 901 MHz (–50 to –20 dBm), and $V_{CC} = 5$. Four LO harmonics and three LO sidebands were used.

an external current source to achieve an $IP_{3,in}$ of 20 dBm at the expense of additional supply current.

The MC13143 contains 29 active transistors. Figure 4-54 shows the core of the MC13143 circuit reduced to its essentials and configured for computer analysis per Motorola's MC13143 test circuit. Figure 4-55 shows the circuit's conversion gain and noise figure versus LO drive, and Figure 4-56 shows the circuit's typical IF, IM_3, and IP_3 responses, both as calculated by Ansoft Corporation's Serenade 8.0 circuit simulator. A reproduction of the MC13143 datasheet, used by permission, follows.

In considering the approach used in the MC13143, one has to remember that feedback around many stages gives an IMD ripple, which means that the higher-order IPs are no longer a straight-line calculation. Above a certain level, the 3-dB/dB law fails, and as a result of this, the IMD characteristic can be as bad as 10 or 20 dB/dB—almost like a breakdown point.

High-frequency capability is an issue with silicon BJTs; the upper frequency limit of the technique is in the range of 3 GHz. Bipolar technology will be extended to 10 GHz by SiGe, but when these parts will be commercially available at competitive prices is currently unclear. There has not been enough published on the actual performance of Gilbert-cell-type mixers using SiGe transistors, but we assume that because of coupling on the silicon wafer, things are no longer as well behaved as we are used to at much lower frequencies.

4-4-3 FET Mixers

FETs, used in active and passive circuits, are still a popular solution for low-power integrated mixers up to 100 MHz (passive quad) and 1 GHz (dual-gate MOSFET). Cost-sensitive applications may require the use of a single-gate FET. These types of single-gate microwave mixers are frequently used to validate the quality of software but the user keeps forgetting this is a question of the quality of the model and not of the simulator. A good example of bad modeling is Figure 7.88 in Vendelin et al. [17]. Unfortunately, we have not seen the same circuit analyzed with modern tools, nor do the authors of the paper cited [18] provide sufficient details of the circuit and the transistor for us to do this ourselves.

Figure 4-57 shows a basic, active FET mixer. Like the single-diode and single-BJT mixers, it is additive, although it's common to lessen RF–LO interaction by injecting one signal at the gate and the other at the source as shown. Single-gate GaAs and LDMOS FETs are also used in this arrangement with gate biasing as appropriate to the device type. Figure 4-58 shows a dual-gate FET in an additive mixing configuration.

The linearity of FETs is based on the fact that a FET follows a square law and therefore the first derivative, its transconductance, is supposed to be constant. This is valid within a wide amplitude range. FET mixers have a noise figure similar to those found in bipolar mixers: typically around 6–9 dB, depending on the configuration. A FET mixer's intercept point is subject to load-impedance variations, with a purely resistive termination providing the best case. Terminating the mixer with a filter is problematic in that a filter looks purely resistive only within its 3-dB passband; in the transition band and beyond, the filter impedance rises rapidly and the mixer intercept point goes down. The best means devised so far to minimize this effect is to configure the FET output as a high-pass filter, using capacitive coupling from the output tuned circuit to a 50-(bandpass filter, as shown in Figure 4-59. Outside its passband, this impedance inverter acts more like a short circuit and maintains IMD products at a reasonable level. The alternative to this is the popular diplexer, which requires more components and has more insertion loss. The high-pass configuration has barely been mentioned in the literature.

Order this document by MC13143/D

MOTOROLA

Ultra Low Power DC - 2.4 GHz Linear Mixer

The MC13143 is a high compression linear mixer with single–ended RF input, differential IF output and differential LO inputs which consumes as little as 1.8 mW. A new circuit topology is used to achieve a high third order intermodulation intercept point, high linearity and high 1.0 dB output compression point while maintaining a linear 50 Ω input impedance. It is designed for Up or Down conversion anywhere from dc to 2.4 GHz.

Ultra Low Power: 1.0 mA @ V_{CC} = 1.8 to 6.5 V

- Wide Input Bandwidth: DC–2.4 GHz
- Wide Output Bandwidth: DC–2.4 GHz
- Wide LO Bandwidth: DC–2.4 GHz
- High Mixer Linearity: $P_{i1.0\ dB}$ = 3.0 dBm

Linearity Adjustment of up to IP_{3in} = 20 dBm

- 50 Ω Mixer Input
- Single–Ended Mixer Input
- Double Balanced Mixer Operation
- Differential Open Collector Mixer Output

ORDERING INFORMATION

Device	Operating Temperature Range	Package
MC13143D	T_A = –40 to 85°C	SO–8

MAXIMUM RATINGS (T_A = 25°C, unless otherwise noted.)

Rating	Symbol	Value	Unit
Power Supply Voltage	V_{CC}(max)	7.0	Vdc
Junction Temperature	T_{Jmax}	150	°C
Storage Temperature Range	T_{stg}	–65 to 150	°C

NOTE: ESD data available upon request.

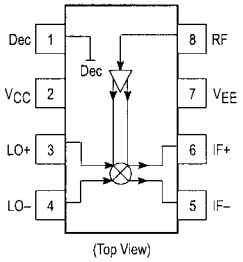
MC13143

ULTRA LOW POWER DC – 2.4 GHz LINEAR MIXER

SEMICONDUCTOR TECHNICAL DATA

D SUFFIX
PLASTIC PACKAGE
CASE 751
(SO–8)

PIN CONNECTIONS



(Top View)

This device contains 29 active transistors.

MC13143

RECOMMENDED OPERATING CONDITIONS

Rating	Symbol	Min	Typ	Max	Unit
Power Supply Voltage	V_{CC}	1.8	–	6.0	Vdc

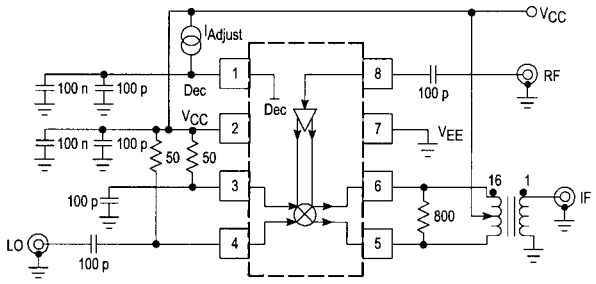
DC ELECTRICAL CHARACTERISTICS (T_A = 25°C, V_{CC} = 3.0 V, f_{RF} = 1.0 GHz, Pin = –25 dBm.)

Characteristic	Symbol	Min	Typ	Max	Unit
Supply Current (Lin Control Current = 0)	I_{CC1}	–	1.0	–	mA
Supply Current (Lin Control Current = 1.6 mA)	I_{CC2}	–	4.1	–	mA

AC ELECTRICAL CHARACTERISTICS (T_A = 25°C, V_{CC} = 3.0 V, f_{RF} = 1.0 GHz, Pin = –25 dBm.)

Characteristic	Symbol	Min	Typ	Max	Unit
Mixer Voltage Conversion Gain (R_P = R_L = 800 Ω)	VG_C	–	9.0	–	dB
Mixer Power Conversion Gain (R_P = R_L = 800 Ω)	PG_C	–3.5	–2.6	–1.5	dB
Mixer Input Return Loss	Γin_{mx}	–	–20	–	dB
Mixer SSB Noise Figure	NF_{SSB}	–	14	15	dB
Mixer 1.0 dB Compression Point (Mx Lin Control Current = 1.6 mA)	$Pin_{-1.0\ dB}$	–1	0	–	dBm
Mixer Input Third Order Intercept Point (d_f = 1.0 MHz, $I_{control}$ = 1.6 mA)	$IP3_{in}$	–	16	–	dBm
LO Drive Level	LO_{in}	–	–5.0	–	dBm
LO Leakage to Mixer IF Outputs	P_{LO-IF}	–	–33	–25	dB
Mixer Input Feedthrough Output	P_{RFm-IF}	–	–25	–	dB
LO Leakage to Mixer Input	P_{LO-RFm}	–	–40	–25	dB
Mixer Input Leakage to LO	P_{RFm-LO}	–	–35	–	dB

Figure 1. Test Circuit



MC13143

TYPICAL PERFORMANCE CURVES

Figure 2. Power Conversion Gain and Supply Current versus Supply Voltage

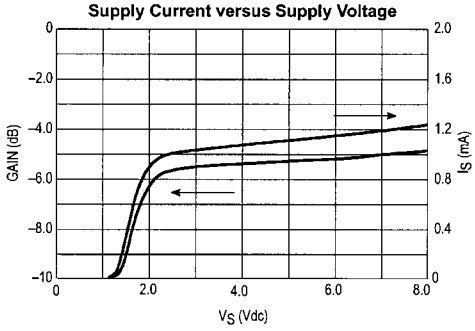


Figure 3. Noise Figure and Gain versus LO Power

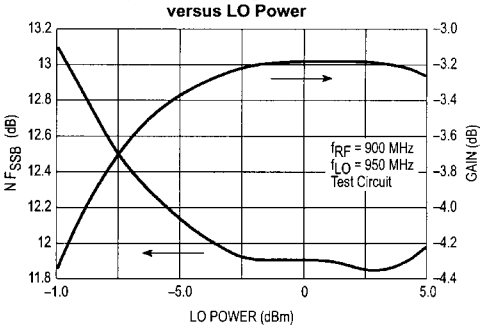


Figure 4. Mixer Input Return Loss versus RF Input Frequency

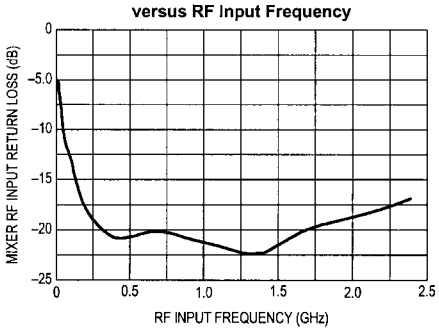


Figure 5. Power Conversion Gain and Supply Current versus RF Input Power

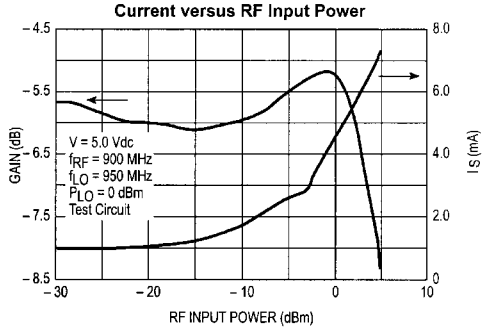


Figure 6. Noise Figure and Gain versus RF Frequency

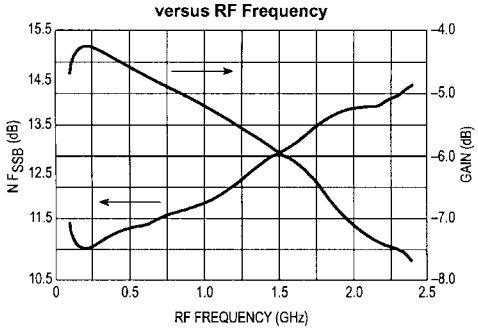
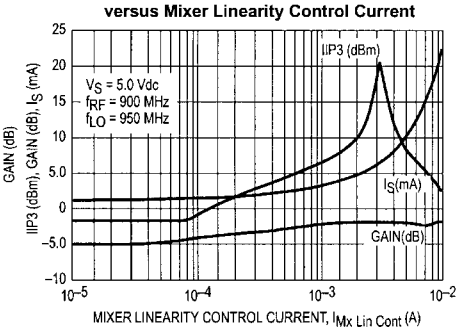


Figure 7. IIP3, Gain, Supply Current versus Mixer Linearity Control Current



MC13143

CIRCUIT DESCRIPTION

General

The MC13143 is a double–balanced Mixer. This device is designated for use as the frontend section in analog and digital FM systems such as Wireless Local Area Network (LAN), Digital European Cordless Telephone (DECT), PHS, PCS, GPS, Cellular, UHF and 800 MHz Special Mobile Radio (SMR), UHF Family Radio Services and 902 to 928 MHz cordless telephones. It features a mixer linearity control to preset or auto program the mixer dynamic range, an enable function and a wideband IF so the IC may be used either as a down converter or an up converter.

Current Regulation

Temperature compensating voltage independent current regulators provide typical supply current at 1.0 mA with no mixer linearity control current.

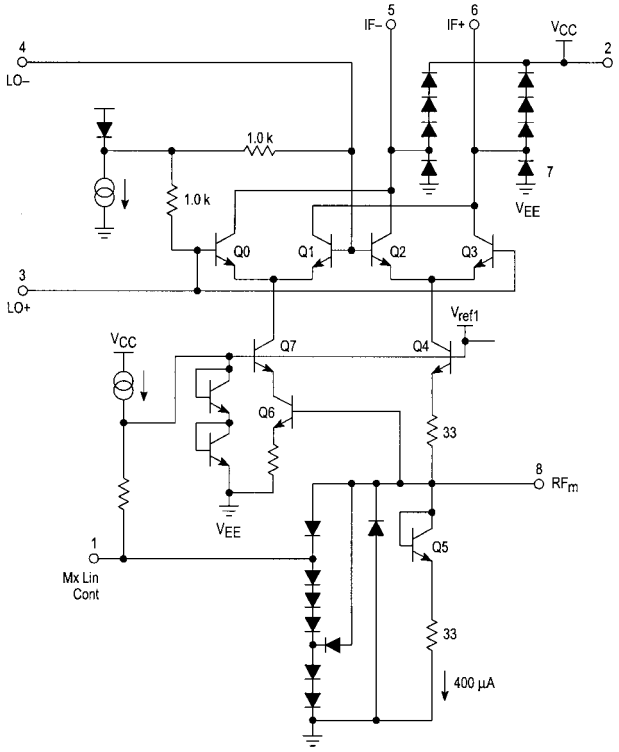
Mixer

The mixer is a unique and patented double–balanced four quadrant multiplier biased class AB allowing for programmable linearity control via an external current source. An input third order intercept point of 20 dBm may be achieved. All 3 ports of the mixer are designed to work up to 2.4 GHz. The mixer has a 50 Ω single–ended RF input and open collector differential IF outputs (see Internal Circuit Schematic for details). The linear gain of the mixer is approximately –5.0 dB with a SSB noise figure of 12 dB.

Local Oscillator

The local oscillator has differential input configuration that requires typically –10 dBm input from an external source to achieve the optimal mixer gain.

Figure 8. MC13143 Internal Circuit*



NOTE: * The MC13143 uses a unique and patented circuit topology.

MC13143

APPLICATIONS INFORMATION

Evaluation PC Board

The evaluation PCB is very versatile and is intended to be used across the entire useful frequency range of this device. The PC board is laid out to accommodate all SMT components on the circuit side (see Circuit Side Component Placement View).

Component Selection

The evaluation PC board is designed to accommodate specific components, while also being versatile enough to use components from various manufacturers. The circuit side placement view is illustrated for the components specified in the application circuit. The Component Placement View specifies particular components that were used to achieve the results shown in the typical curves and tables.

Mixer Input

The mixer input impedance is broadband 50 Ω for applications up to 2.4 GHz. It easily interfaces with a RF ceramic filter as shown in the application schematic.

Mixer Linearity Control

The mixer linearity control circuit accepts approximately 0 to 2.3 mA control current. An Input Third Order Intercept Point, IIP3 of 20 dBm may be achieved at 2.3 mA of control current (approximately 7.0 mA of additional supply current).

Local Oscillator Inputs

The differential LO inputs are internally biased at $V_{CC} - 1.0\ V_{BE}$; this is suitable for high voltage and high gain operation.

For low voltage operation, the inputs are taken to V_{CC} through 51 Ω.

IF Output

The IF is a differential open collector configuration which is designed to use over a wide frequency range for up conversion as well as down conversion.

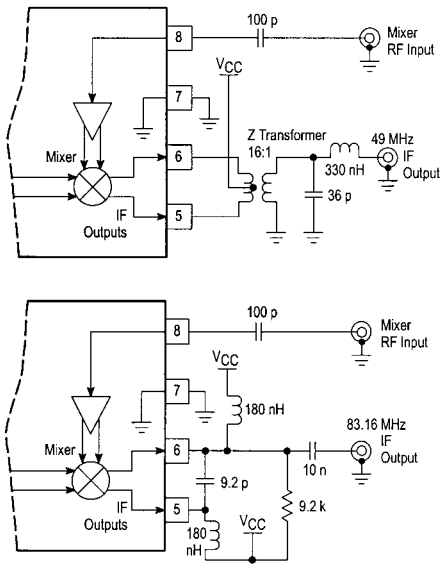
Input/Output Matching

It is desirable to use a RF ceramic or SAW filter before the mixer to provide image frequency rejection. The filter is selected based on cost, size and performance tradeoffs. Typical RF filters have 3.0 to 5.0 dB insertion loss. The PC board layout accommodates both ceramic and SAW RF filters which are offered by various suppliers such as Siemens, Toko and Murata.

Interface matching between the RF input, RF filter and the mixer will be required. The interface matching networks shown in the application circuit are designed for 50 Ω interfaces.

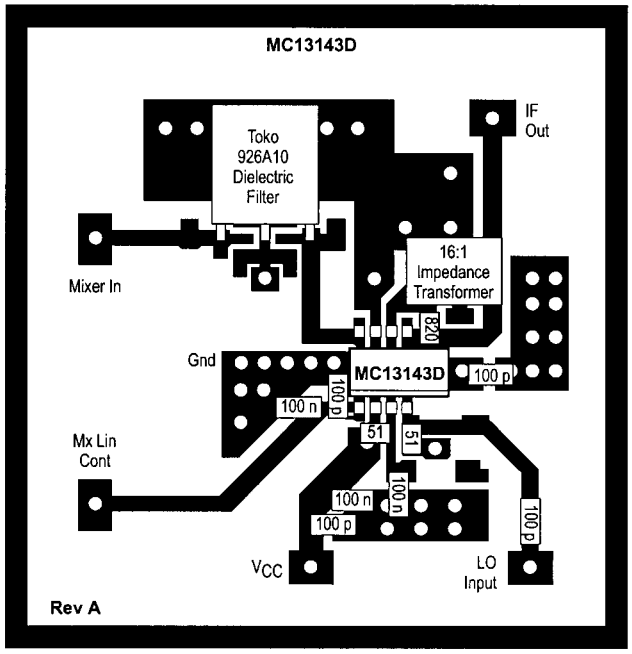
Differential to single–ended circuit configuration is shown in the test circuit. 6.0 dB of additional mixer gain can be achieved by conjugately matching the output of the MiniCircuits transformer to 50 Ω at the desired IF frequency. With narrowband IF output matching the mixer performance is 3.0 dB gain and 12 dB noise figure (see Narrowband 49 and 83 MHz IF Output Matching Options). Typical insertion loss of the Toko ceramic filter is 3.0 dB. Thus, the overall gain of the circuit is 0 dB with a 15 dB noise figure.

Figure 9. Narrowband IF Output Matching with 16:1 Z Transformer and LC Network



MC13143

Figure 10. Circuit Side Component Placement View



NOTES: 926.5 MHz preselect dielectric filter is Toko part # 4DFA–926A10; the 4DFA (2 and 3 pole SMD type) filters are available for applications in cellular and GSM, GPS, DECT, PHS, PCS and ISM bands at 902–928 MHz, 1.8–1.9 GHz at 2.4–2.5 GHz.

The PCB also accommodates a surface mount RF SAW filter in an eight or six pin ceramic package for the cellular base and handset frequencies. Recommended manufacturers are Siemens and Murata.

The PCB may also be used without a preselector filter; AC coupled to the mixer as shown in the test circuit schematic. All other external circuit components shown in the PCB layout above are the same as used in the test circuit schematic.

16:1 broadband impedance transformer is mini circuits part #TX16–R3T; it is in the leadless surface mount "TX" package. For a more selective narrowband match, a lowpass filter may be used after the transformer. The PCB is designed to accommodate lump inductors and capacitors in more selective narrowband matching of the mixer differential outputs to a single–ended output at a given IF frequency.

The local oscillator may also be driven in a differential configuration using a coaxial transformer. Recommended sources are the Toko Balun transformers type B4F, B5FL and B5F (SMD component).

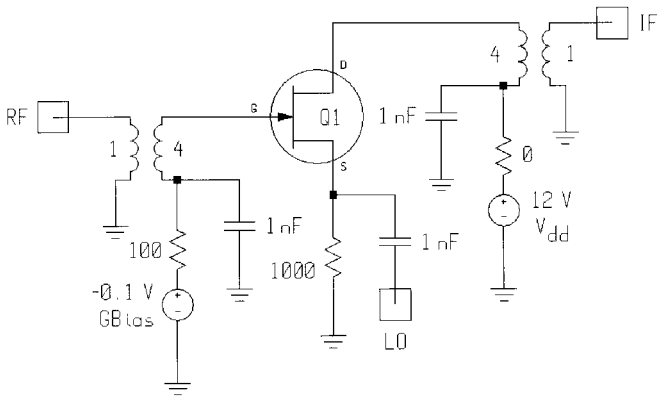


Figure 4-57 An additive JFET mixer cell using gate RF injection and source LO injection.

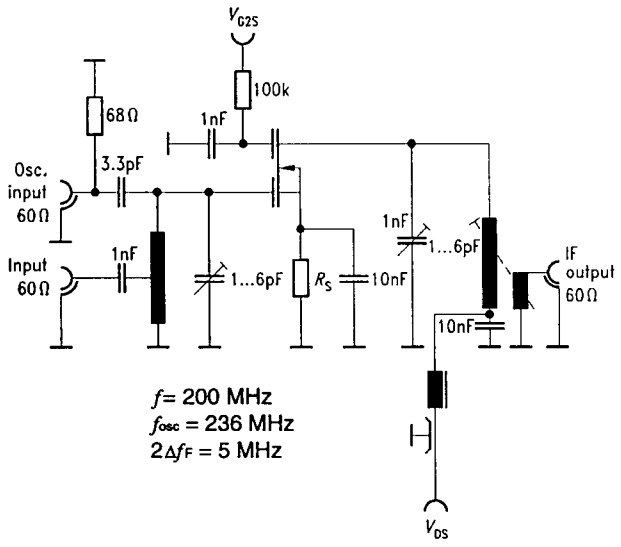


Figure 4-58 Test circuit for additive mixing using a dual-gate MOSFET.

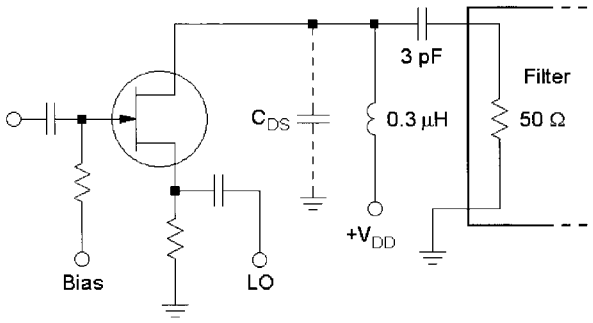


Figure 4-59 FET mixer output coupling that minimizes the effects of filter reactance outside the filter passband.

The active FET mixer achieves gain at the expense of intercept point; the difference can be as much as 20 dB. As an example, the passive MOSFET quad mixer, which is to be referenced, achieves an intercept point of greater than 40 dBm while similar active silicon FETs barely make 30 dBm. This 30-dBm performance is possible only at fairly low frequencies and requires careful selection and adjustment of the dc bias point of all transistors of the circuit. This performance also requires that the FETs be operated grounded-gate, which limits their gain. The other drawback of these active arrangements is that higher impedances at the output reduce the third-order intercept point overproportionally; as an example, a factor of two in impedance change can cost as much as 6 dB in the intercept point. In these discussions we assume that the IMD products are generated solely by the active device, and not by any passive devices, such as transformers. Actually, there are also some capacitors whose linearity varies with RF current—another unpleasant effect that is frequently overlooked.

On the other hand, one can use any FET as a passive device similar to a diode mixer, in which the source–drain channel gets switched on and off. This impedance modulation is somewhat similar to a diode mixer, but the gate electrode is isolated from both source and drain. It nonetheless falls into the category of additive mixers because there is sufficient interaction between gate and source, although the impedance at the gate changes significantly less than in an additive diode mixer. Implementation is a challenge in that building a high-performance passive FET mixer requires a pair or a quad of mixer cells that are sufficiently matched to suppress even-order IMD products. Intercept points, depending on the LO drive, vary between +20 and +45 dBm. The lower number is more applicable for microwave and RF frequencies, while the +45-dBm level is easily obtainable between and 5 and 30 MHz. It is necessary to remember that the FET is a voltage-driven switch, and LO matching becomes an issue: How do you generate 30 V peak-to-peak across a few picofarads over a wide frequency range? Physical layout can also be critical: Symmetry is not only important for the active devices, but also for the input, output, and LO circuit. We have seen cases where fractions of an inch in different lengths leading to the gate electrodes have cost up to 10 dB of IP_3—just because of the resulting lack of symmetry.

Figure 4-60 shows the circuit of Figure 4-57 modified to work with an LDMOS FET, the harmonic-balance version of the SPICE level 3 model. Figure 4-61 shows the LDMOS circuit's conversion gain and noise figure versus frequency. The "noisy" traces reflect the model's numerical problems. This model, which works up to about 0.8 μm and has been validated against the Motorola data and SPICE results, has already been made mathematically continuous but still shows the inability to properly model the real RF transistor. This is why more research needs to be done on RF and microwave applications for MOS in the modeling area. As can be seen from the references, several attempts have been made to improve the model, but the major headache remains with parameter extraction, and as the models get more complex, the modeling quality doesn't seem to improve as a linear function. The reader should be reminded of the load-pull measurement as outlined earlier.

Dual-Gate MOS/GaAs Mixers. Significantly better LO–RF isolation can be obtained by applying the RF to LO signals at separate gates of a dual-gate FET (Figure 4-62). The result is multiplicative mixing with fewer constraints on LO and RF filtering.

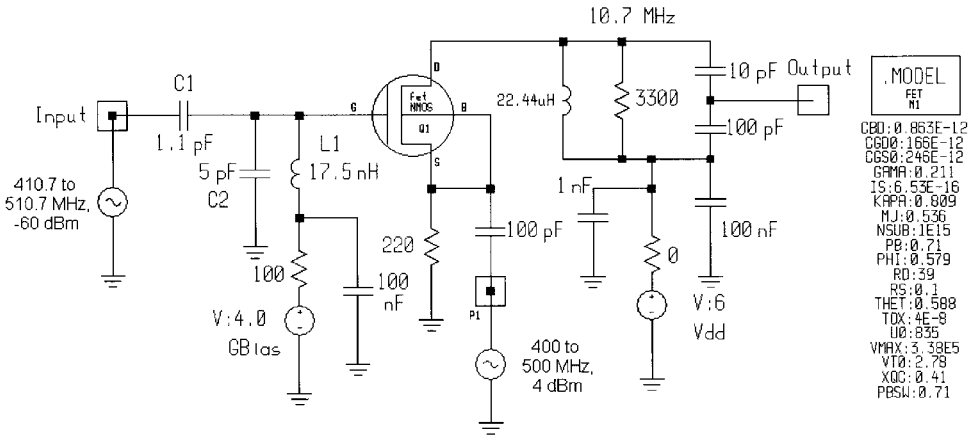


Figure 4-60 The N-JFET circuit of Figure 4-57 modified to work with an LDMOS FET.

4-4-4 MOSFET Gilbert Cell

The trend to go to smaller voltages has created some CMOS implementations of the Gilbert cell. While this may allow us to integrate reasonable mixers on the same chip, nobody should expect any miracles from these mixers. They are frequently starved in operating voltage and current and rarely fare better than a single diode mixer with proper drive level. On the other hand, the symmetry reduces some of the unwanted spurious frequencies and because of the high impedance, the actual RF power level is much less than the diode would need. There are so

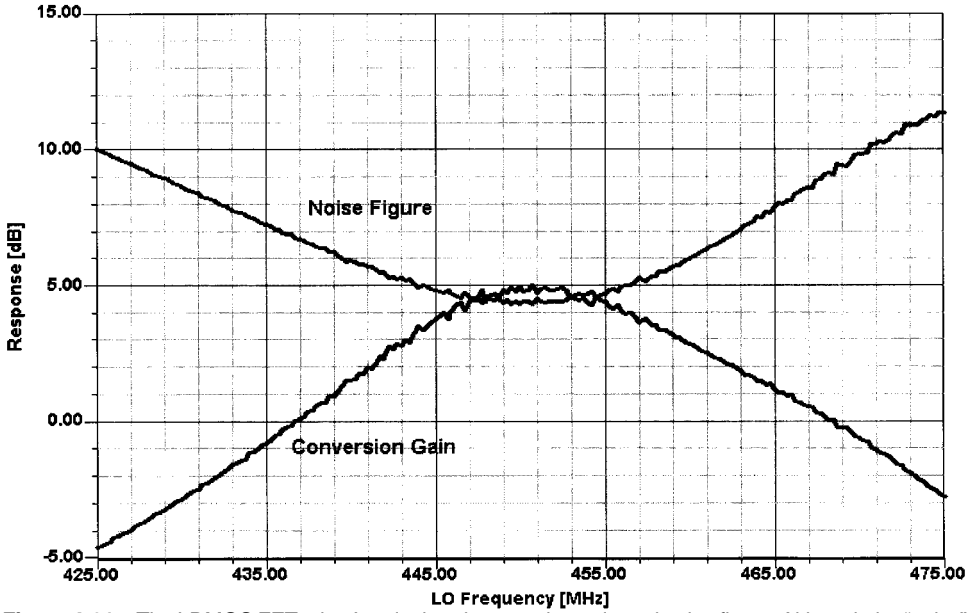


Figure 4-61 The LDMOS FET mixer's calculated conversion gain and noise figure. Although the "noisy" traces reflect numerical problems in the device model, the results can be put to good use in circuit design.

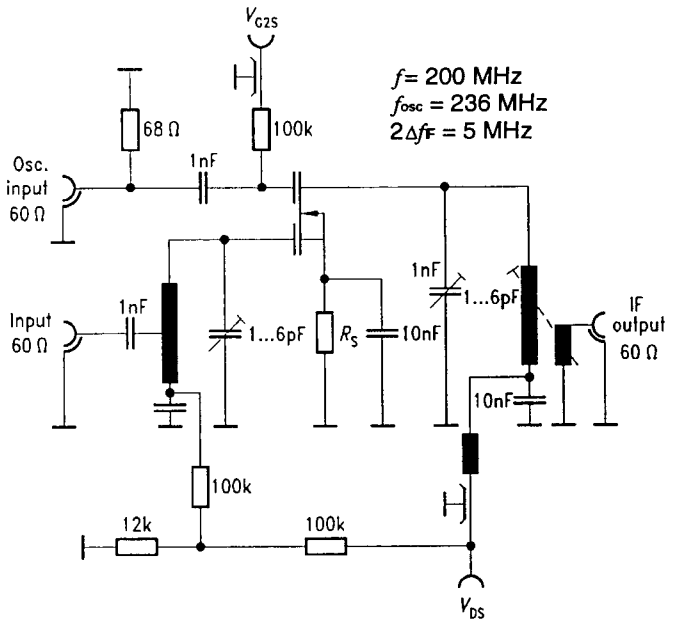


Figure 4-62 Test circuit for multiplicative mixing using a dual-gate MOSFET.

many CMOS transistors available in different processes that we cannot provide information about the general behavior, but this "low-cost" solution, as pointed out, is really only attractive because one can stay in the same technology while building the amplifier and other stages. Because of the flicker corner frequency of MOS, the noise figure of the mixer will be high compared to other devices, but not as bad as its GaAsFET brothers. In any case, to combine performance and simplicity, one needs to have a preamplifier that reduces the noise figure as well as the third-order intercept point of the mixer by the amount of preamplification.

4-4-5 GaAsFET Single-Gate Switch

In general, GaAsFET technology is the most expensive one. While GaAsFET competes with SiGe technology in frequency range at the time of this writing, it is a mature technology and well understood. The gallium arsenide wafers, however, are still unable to compete on a price basis with other technologies. There is a hybrid technology, silicon on sapphire, that also tries to go after the GaAsFET domain. The cutoff frequency of GaAsFET transistors is significantly higher than any other FETs, but the corner frequency f_c, also frequently referred to as flicker frequency, is somewhere between 10 and 100 MHz and therefore results in poor mixer performance from a noise figure point of view. The same applies, by the way, to oscillators. Another headache with GaAs is the low Q of the material, but this low Q gives transformers on GaAs a wider bandwidth. One of the pioneers of ICs for the wireless market is TriQuint, and we have been working with Wes Hayward of TriQuint in circuit analysis and validation. In doing so, we used some commercial and experimental circuits. In Chapter 3 we have already pointed out one amplifier that shows good agreement between measured and predicted data. By the time one has realized the entire circuit, including its associated biasing problems, even these circuits tend to grow. Figure 4-63 shows the circuit, which consists of an oscillator/amplifier, switching mixer, and a differential IF postamplifier.

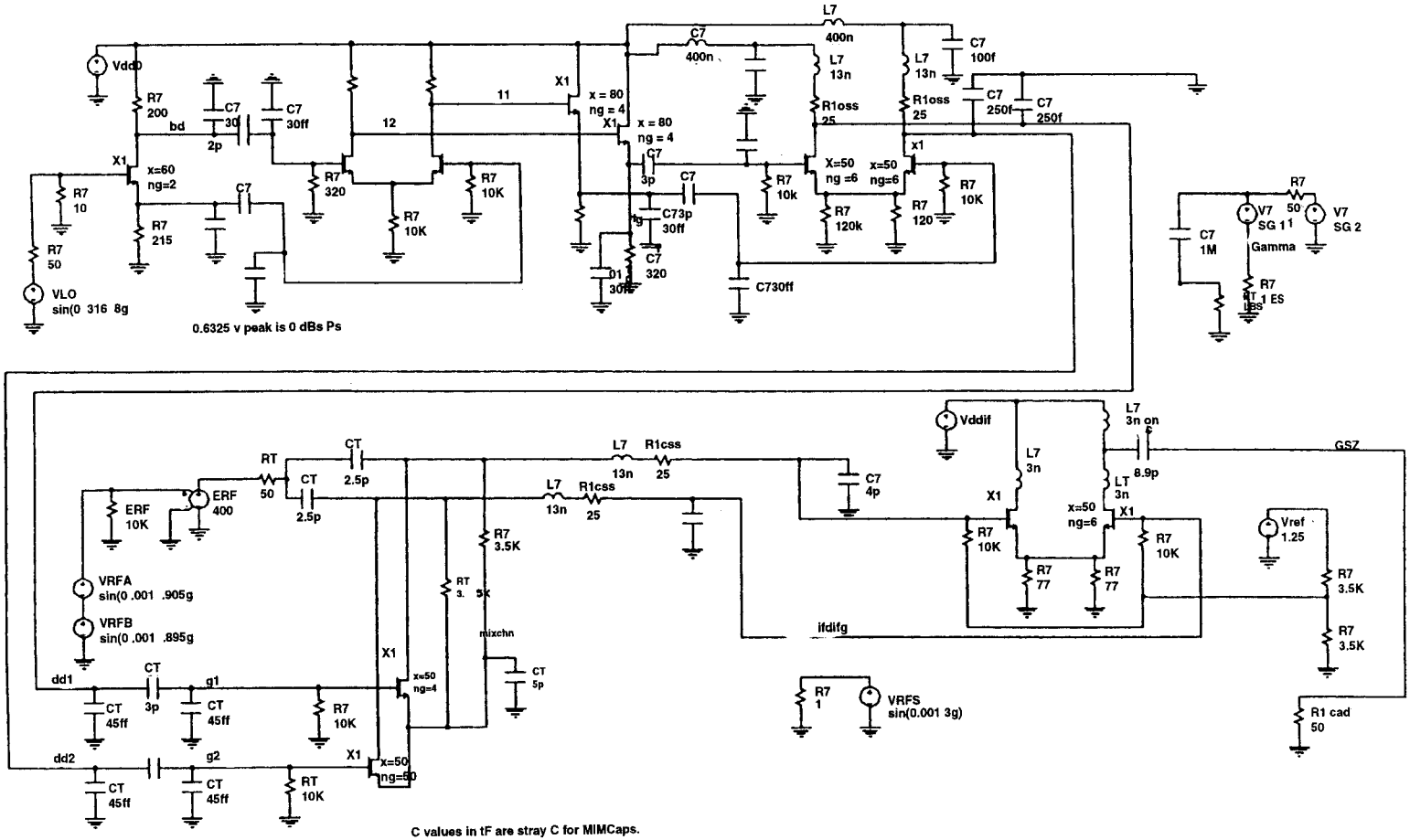
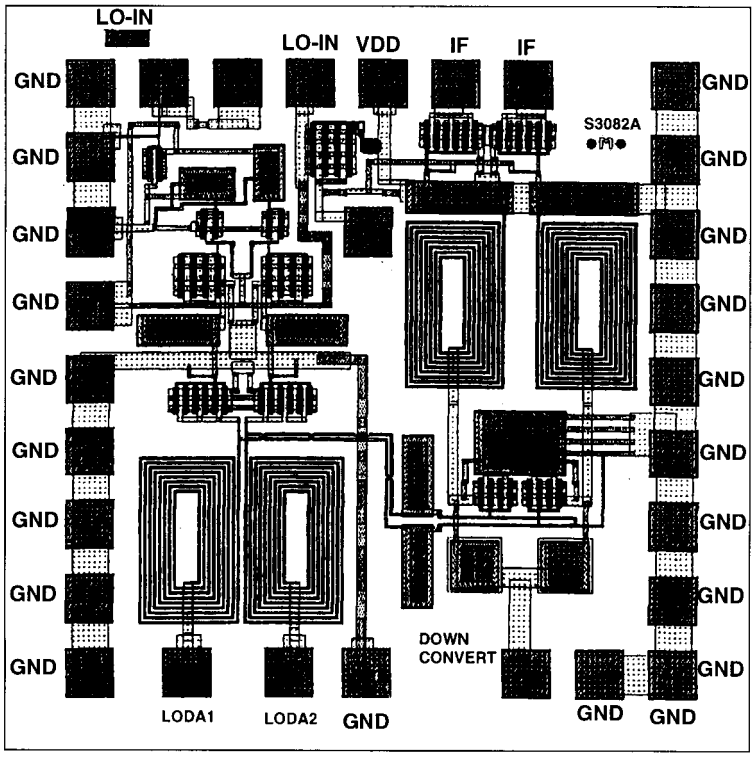


Figure 4-63 The complete switching mixer circuit.

Since the highest accuracy is needed to model these types of circuit, we found that only very recent simulators, such as Serenade 8 from Ansoft, can handle such a large number of FETs without barking at the user. Initially, we had all kinds of difficulties with these circuits because of the TriQuint TOM model, which had been developed for SPICE, and its IF/ELSE statements that made it mathematically discontinuous. The type of GaAsFETs used by TriQuint were both regular FETs and enhancement FETs, which require a positive voltage at the gate. Further details of this should be looked up in the foundry manual by TriQuint. In Chapter 2 we have shown some comparisons between different models and their prediction accuracy. Such converters need to resort to inductors in GaAs technology, and compared to other parts they become quite big (Figure 4-64).

In order to reduce some of the complexity in modeling and yet deal with the problems of mixers, a simpler mixer is needed that is a combination of a set of switching transistors and a postamplifier for the IF (Figure 4-65). The reader may notice that the small inductances of 15 nH have an extremely high loss modeled by a series resistor of 18 Ω. The first to recommend this type of switching mixers was probably Stephen Maas. The circuit diagram itself (Figure 4-65) is self-explanatory, but again we need to point out the resulting complexity.

The GaAsFET mixer giving the best performance on the market today is the Siemens model CMY210 (Figure 4-66). A reproduction of its datasheet, used by permission, follows the schematic.


LO-IN
LO-IN
VDD
IF
IF
S3082A
GND
GND
GND

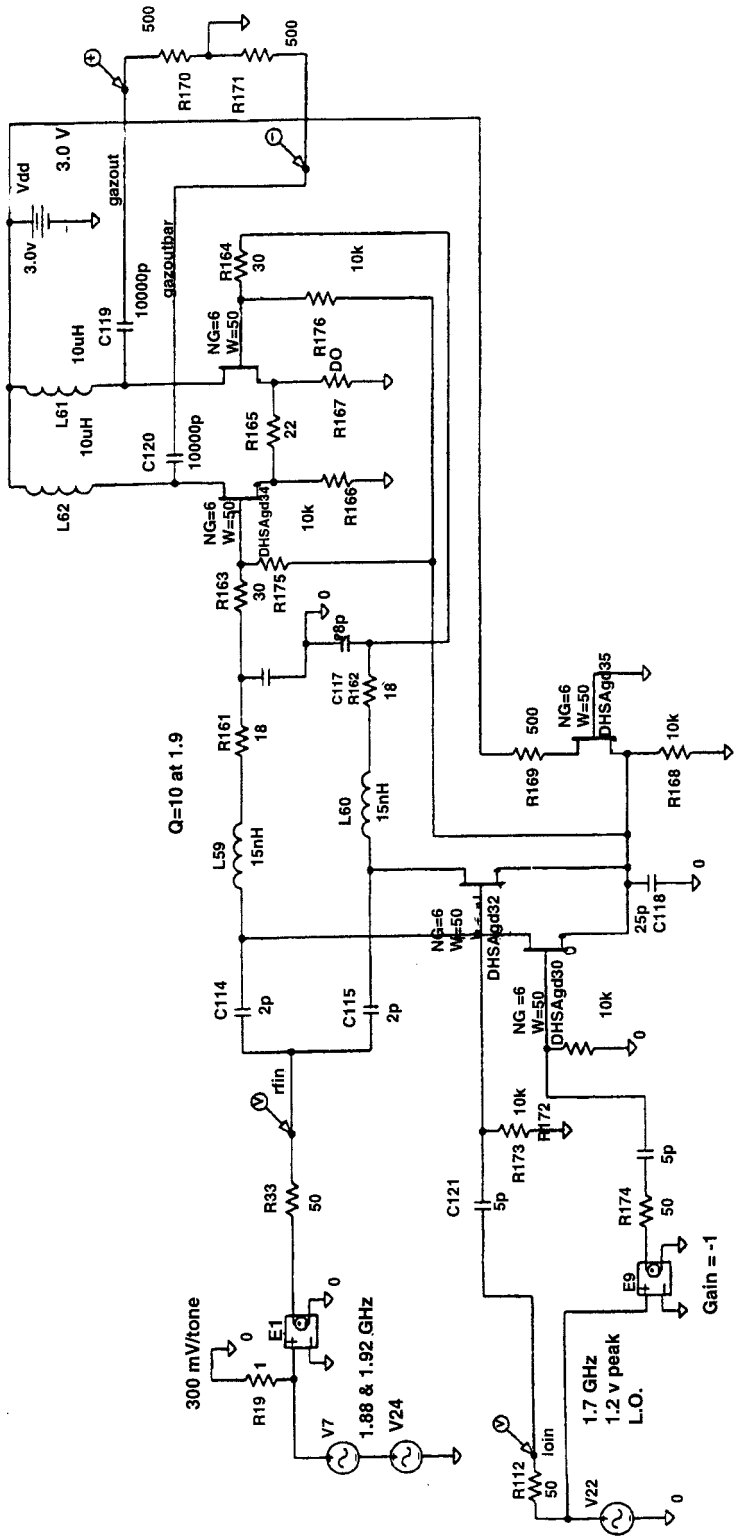


Figure 4-65 Simplified switching FET mixer.

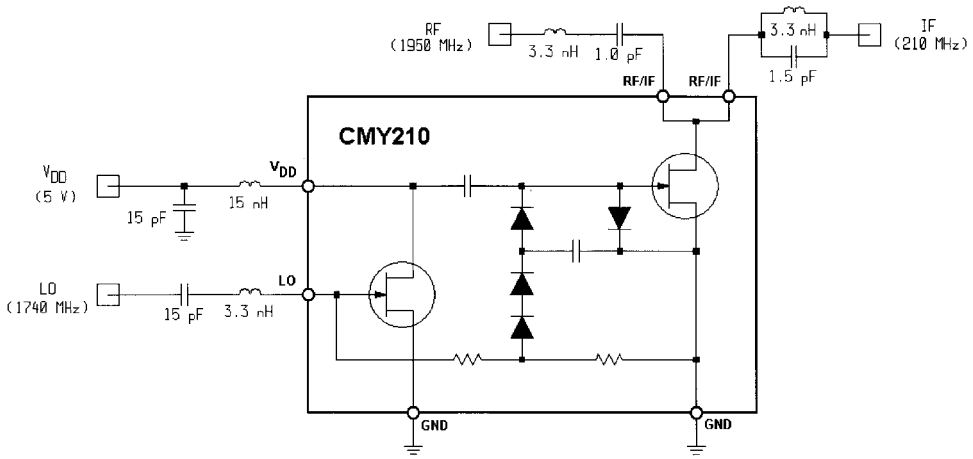


Figure 4-66 The Siemens CMY210 IC includes a shunt switch mixer and AGC-equipped LO amplifier for more constant LO drive. Despite its simplicity, this mixer achieves a typical $IP_{3,in}$ of 25 dBm.

The CMY210 circuit is a clever combination of a switching type of mixer with an automatic level control arrangement that provides the mixer with a reasonably constant drive level following the LO amplifier. Similar to cases we have shown, there is a clear relationship between the local oscillator power (invested dc power), and large-signal performance. This mixer externally has a series-tuned circuit at the input, which allows the UHF input signal

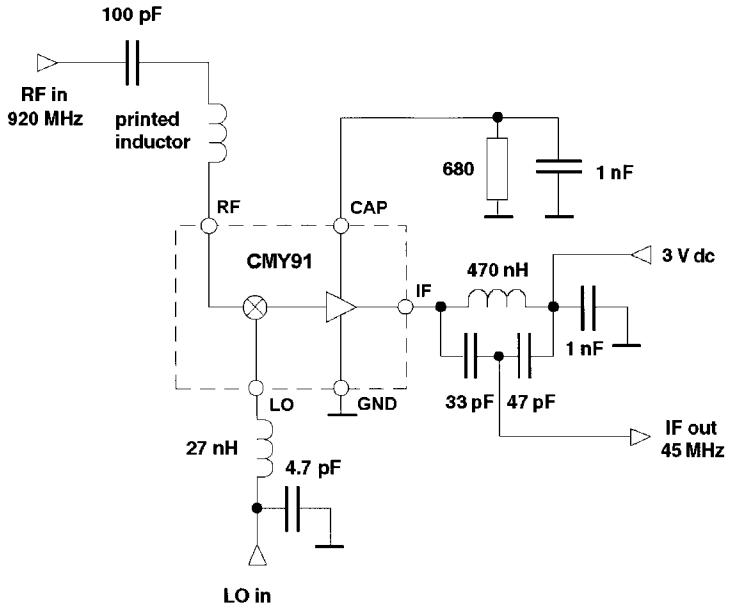


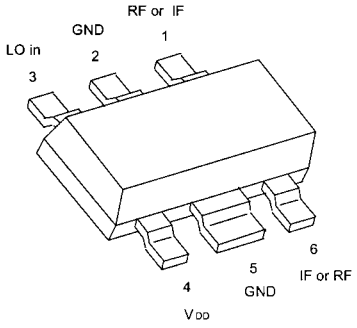
Figure 4-67 The CMY91 IC includes a postmixer amplifier instead of an AGC LO amplifier. Its $IP_{3,in}$ is typically –2 dBm (drain supply, 3 V at 1 mA) with no connection made to its CAP pin. In the configuration shown, its typical $IP_{3,in}$ is 0 dBm (drain supply, 3 V at 2.5 mA).

SIEMENS GaAs MMIC CMY 210

Preliminary Data

* **Ultralinear Mixer** with integrated LO-Buffer
* Very high **Input-IP3 of typical 25dBm**
* Very low **LO-Power demand of typ. 0dBm**
* Suited for Up- and Down-Conversion
* Wide LO-Frequency Range
  <500MHz to >2,5GHz
* Wide LO-Level Range
* Single ended Ports
* RF- and IF-Port Impedance 50 Ohm
* Operating Voltage Range: < 3 to 6V
* Very low Current Consumption of typical 7mA
* All Gold Metallisation

ESD: **E**lectro**s**tatic **d**ischarge sensitive device
Observe handling Precautions!



Type	Marking	Ordering code (tape and reel)	Package [1]
CMY210	**M3**	**Q62702 M 0016**	**MW-6**

Maximum Ratings	Port	Symbol	Value		Unit
			min	max	
Supply Voltage	4	V_{DD}	0	6	V
DC-Voltage at LO Input	3	V_3	-3	0,5	V
DC-Voltage at.RF-IF Ports [2]	1, 6	$V_{1,6}$	- 0,5	+ 0,5	V
Power into RF-IF Ports	1, 6	$P_{in,RF}$		10	dBm
Power into LO Input	3	$P_{in,LO}$		10	dBm
Channel Temperature		T_{Ch}		150	°C
Storage Temperature		T_{stg}	-55	150	°C
Thermal Resistance					
Channel to Soldering Point (GND)		R_{thChS}	≤100		K/W

1) For detailed dimensions see chapter Package Outlines
2) For DC test purposes only, no DC voltages at pins 1, 6 in application

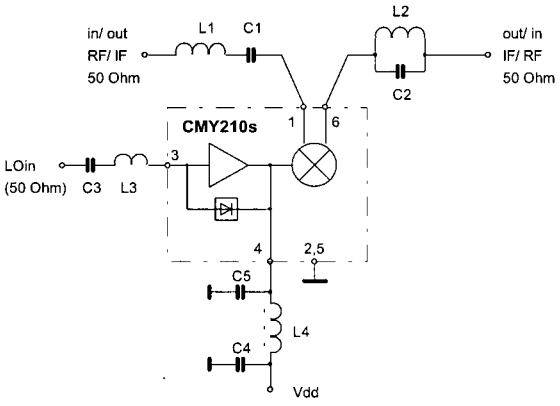
SIEMENS GaAs MMIC **CMY 210**

Electrical Characteristics

Test conditions: T_a = 25°C; V_{DD}= 3V, see test circuit; f_{RF} = 1620MHz; f_{LO} = 1500MHz; P_{LO} = 0dBm; f_{IF}= 120MHz, unless otherwise specified:

Parameter, Test Conditions	Symbol	min	typ	max	Unit
Operating Current	I_{op}	-	7	-	mA
Conversion Loss	L_c	-	5,5	-	dB
SSB Noise Figure	F_{ssb}	-	5,5	-	dB
2 Tone 3rd Order IMD $P_{RF1} = P_{RF2}$ = -3dBm f_{RF1} = 1615MHz; f_{RF2} = 1620MHz; f_{LO}=1500MHz	d_{IM3}	-	56	-	dBc
3rd Order Input Intercept Point	$IP3_{in}$	-	25	-	dBm
P_{-1dB} Input Power	P_{-1dB}	-	20	-	dBm
LO Leakage at RF/IF-Port (1,6)	$P_{LO\ 1,6}$	-	-6	-	dBm

Test circuit / application example



Notes for external elements:

L1, C1: Filter for upper frequency;
C2, L2: Filter for lower frequency;
each filter is a throughpath for the desired frequency (RF or IF) and isolates the other frequency (IF or RF) and its harmonics.
These two filters must be connected to pin 1 and pin 6 directly.
Parasitic capacitances at the ports 1 and 6 must be as small as possible.
L4 and C5 are optimized by indicating lowest I_{op} at used LO-frequency; same procedure for L3.
The ports 1, 3 and 6 must be DC open.

f *LO*	L1	C1	L2	C2	L3	C3	L4	C4	C5
MHz	nH	pF	nH	pF	nH	pF	nH	pF	pF
500	11	8.2	19	4.7	6	47	19	47	3.3
1000	7	4.7	9	3.3	6	33	16	33	0
1500	4	2.2	7,5	2.2	6	18	8.5	18	0
2000	*)	2.2	*)	1.5	4	15	*)	15	0
2500	*)	2.2	*)	1.5	3	15	*)	15	0
notes next pg	1)	1)	1)	1)	3)	3)	2)	2)	2)

SIEMENS GaAs MMIC **CMY 210**

Approximate values of used elements for down- (or up-) conversion with a lower frequency of 10MHz to 120MHz (IF or RF; in or out); parasitics will cause deviations; therefore exact values will be defined by application, especially for *).

General description and notes:

The CMY 210 is an all port single ended general purpose Up- and Down-Converter.
It combines small conversion losses and excellent intermodulation characteristics with a low demand of LO- and DC-power.
The internal level controlled LO-Buffer enables a good performance over a wide LO level range.
The internal mixers principle with one port RF and IF requires a frequency separation at pin 1 and 6 respectively.

Note 1:
Best performance with lowest conversion loss is achieved when each circuit or device for the frequency separation meets the following requirements:

Input Filter: Throughpass for the signal to be mixed; reflection of the mixed signal and the harmonics of both.

Output Filter: Throughpass for the mixed signal and reflection of the signal to be mixed and the harmonics of both.

The impedance for the reflecting frequency range of each filter toward the ports 1 and 6 should be as high as possible.
In the simplest case a series- and a parallel- resonator circuit will meet these requirements but also others as appropriate drop in filters or micro stripline elements can be used.
The two branches with filters should meet immediately at the package leads of the port 1 and 6.
Parasitic capacitances at these ports must be kept as small as possible.
The mixer also can be driven with a source- and a load impedance different to 50Ω, but performance will degrade at larger deviations.

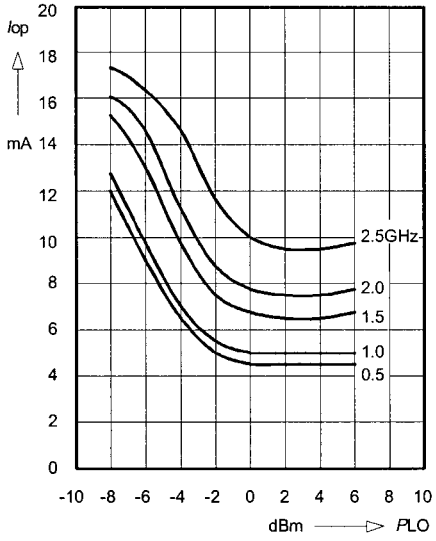
Note 2:
The LO-Buffer needs an external inductor L4 at port 4; the value of inductance depends on the LO frequency. It is tuned for minimum I_{op} consumption into port 4.
At lower LO frequencies it can be reduced by an additonal capacitor C5.

Note 3:
The LO Input impedance at Port 3 can be matched with a series inductor. It also can be tuned for a minimum current I_{op} into port 4. C3 is a DC blocking capacitor.
Since the input impedance of port 3 can be slightly negative at lower frequencies, the source reflection coefficient should be kept below 0.8 ($Z_0 = 50\ \Omega$) within this frequency range.

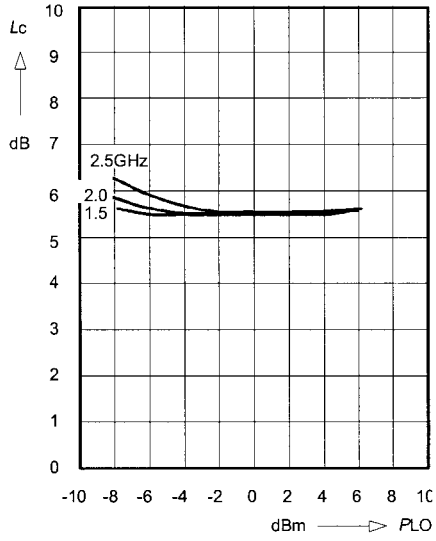
The Conversion Noise Figure Fssb is corresponding with the value of Conversion Loss L_c. The LO signal must be clean of noise and spurious at the frequencies $f_{LO} \pm f_{IF}$.

SIEMENS GaAs MMIC **CMY 210**

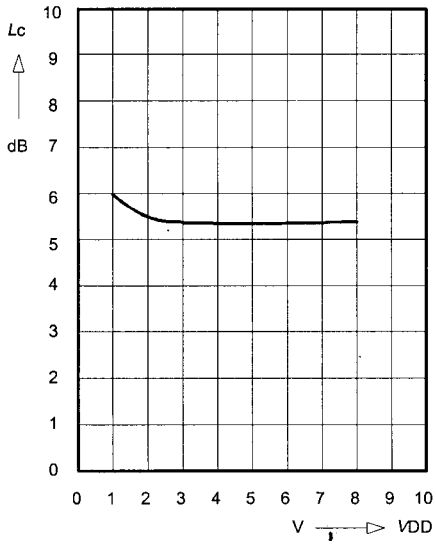
Operating Current $I_{op} = f(P_{LO})$
$V_{DD} = 3V$
f_{LO} = Parameter



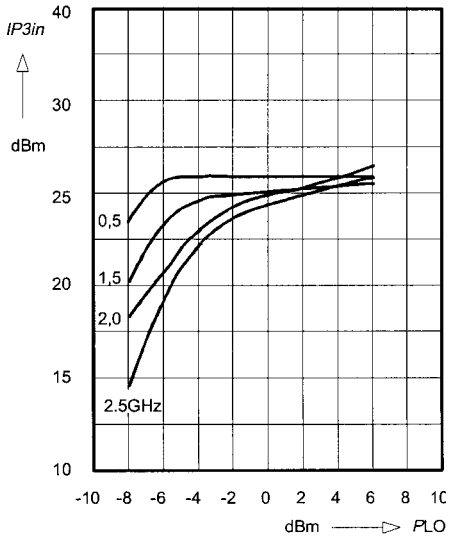
Conversion Loss $L_C = f(P_{LO})$
$V_{DD} = 3V$; $f_{IF} = 120MHz$
f_{LO} = Parameter



Conversion Loss $L_C = f(V_{DD})$
$P_{LO} = 0dBm$
$f_{LO} = 1500MHz$; $f_{IF} = 120MHz$

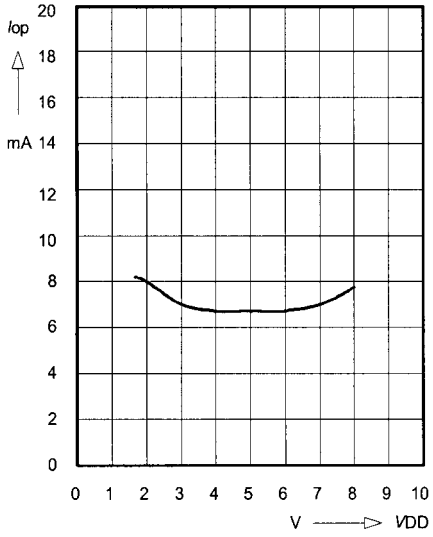


Third Order IP3 $IP3in = f(P_{LO})$
$P_{in} = 2$ x -3dBm; $f_{IF} = 40/45MHz$
$V_{DD} = 3V$; f_{LO} = Parameter;

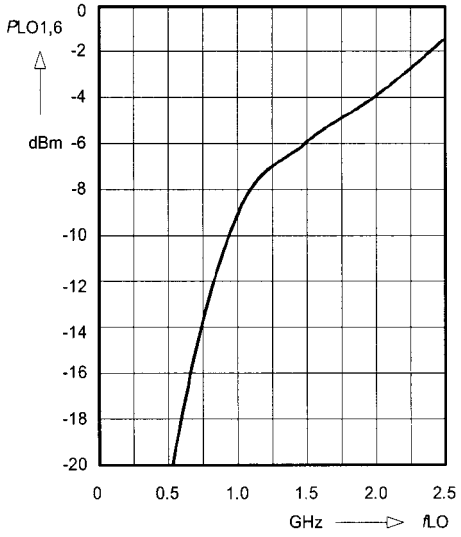


SIEMENS GaAs MMIC **CMY 210**

Operating Current $I_{op} = f(V_{DD})$
P_{LO} = 0dBm
f_{LO} = 1500MHz



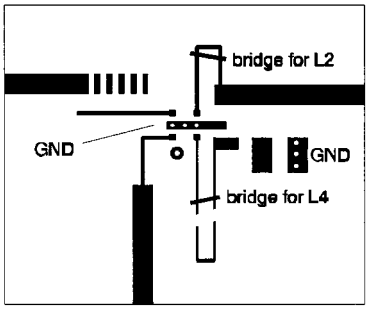
LO-Leakage at Port 1, 6 $P_{LO1,6} = f(f_{LO})$
P_{LO} = 0dBm
V_{DD} = 3V



SIEMENS GaAs MMIC **CMY 210**

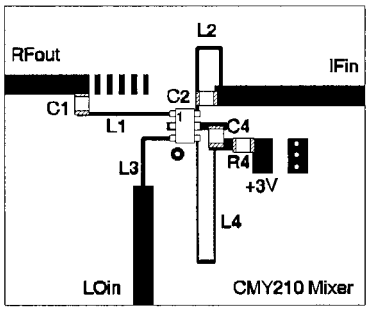
Additional informations and an example of a general purpose mixer pcb:

This general purpuse mixer demonstration board is used to show the performance of the CMY 210. The hints below will be helpful to achieve good intermodulation behaviour.



Material: Epoxy; size 24mm x 30mm; thickness 1mm

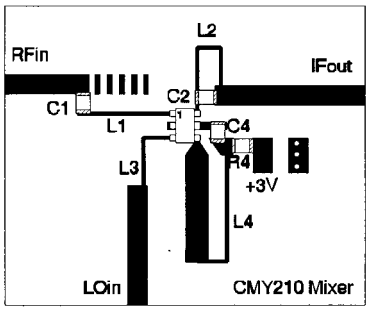
A good ground connection for CMY210 is necessary; here it is realized by 3 vias **under** the CMY 210 device. Best IM3 performance is obtained, when the capacitor C4 is grounded at the upper CMY 210 ground line without any additional vias (so the RF-Signal is encoupled from the LO-buffer best possible). Frequency tuning is done by selecting suitable capacitors C1 and C2, positioning C1 along the L1-line and by reducing the inductance of L2 and L4 by a bridge; unnecessary lines should be disconnected.



This example showes an up-converter.

The IM3 performance of upconverters mostly can be improved by tuning L4 slightly smaller then required for a minimum current consumption.

In other words, the resonant frequency of the buffer circuit at PIN4 (internal capacitor and L4) is tuned by L4 to a frequency slightly above the L.O.-frequency.



This example showes a down-converter.

Here an improved IM3-performance mostly can be obtained by a more capacitive load at CMY 210 port pin 4.
Here it is realized by a broader part of the L4 inductor line toward to port 4 and tuning L4 to a value sligthly smaller then for a minimum current consumption into port 4.

In other words, the resonant frequency of the buffer circuit at PIN4 (internal capacitor and L4) is tuned by L4 to a frequency slightly above the L.O.-frequency.

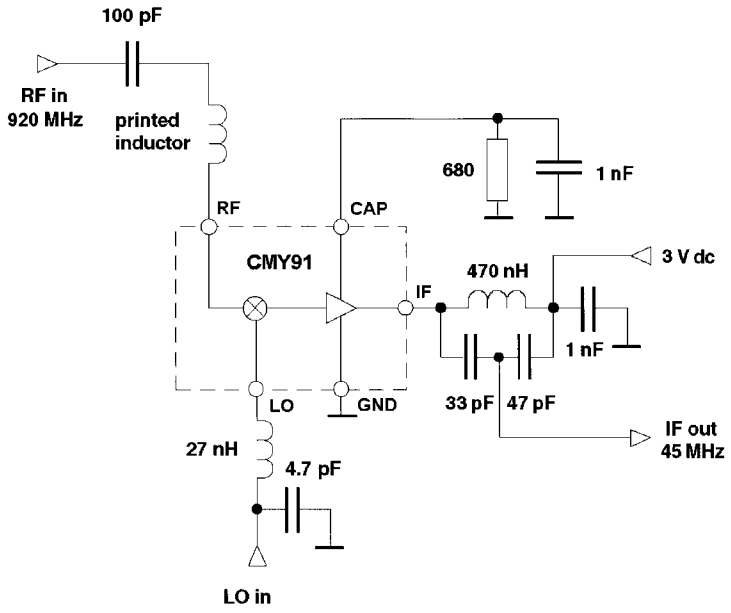


Figure 4-67 The CMY91 IC includes a postmixer amplifier instead of an AGC LO amplifier. Its $IP_{3,in}$ is typically –2 dBm (drain supply, 3 V at 1 mA) with no connection made to its CAP pin. In the configuration shown, its typical $IP_{3,in}$ is 0 dBm (drain supply, 3 V at 2.5 mA).

Table 4-6 Partial CMY91 specifications

Parameters		Symbol	Minimum	Typical	Maximum	Unit
Drain–source breakdown voltage						
I_{IF} = 500 µA	$V_{LO\text{-}GND}$ = 0 V					
$V_{RF\text{-}GND}$ = 4 V	CAP pin not connected	$V_{(BR)IF\text{-}GND}$	8	—	—	V
Drain current						
$V_{RF\text{-}GND}$ = 0 V	$V_{LO\text{-}GND}$ = 0 V					
$V_{IF\text{-}GND}$ = 3 V	CAP pin not connected	I_D	0.8	1	1.4	mA
Conversion gain						
f_{RF} = 920 MHz	f_{LO} = 965 MHz					
f_{IF} = 45 MHz	P_{LO} = –3 dBm	G_c	—	5.5	—	dB
Single-sideband noise figure						
f_{RF} = 920 MHz	f_{LO} = 965 MHz					
f_{IF} = 45 MHz	P_{LO} = –3 dBm	F_{SSB}	—	9	—	dB
Third-order intermodulation						
f_{RF} = 920 MHz	f_{LO} = 965 MHz					
f_{IF} = 45 MHz	P_{LO} = –3 dBm	IP_3	—	–2	—	dBm
LO/RF isolation						
f = 965 MHz		$Iso_{LO/RF}$	—	11	—	dB

T_A = 25 °C/V_D = 3 V; CAP pin connected to ground by 680-Ω resistor.

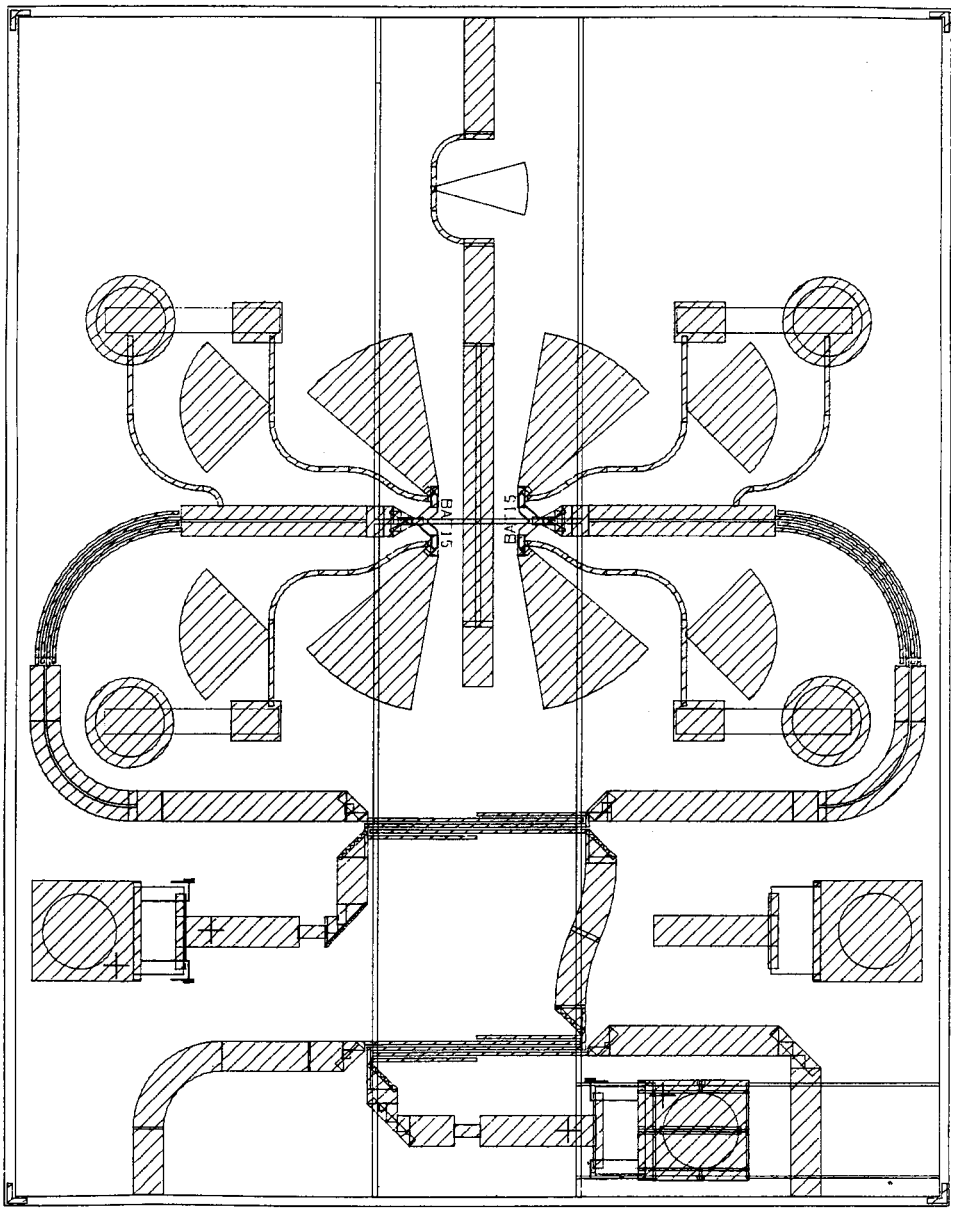


Figure 4-68 Layout of the 20-GHz diode mixer.

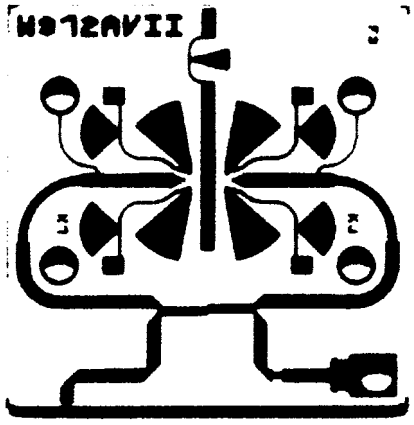


Figure 4-69 The 20-GHz diode mixer circuit board.

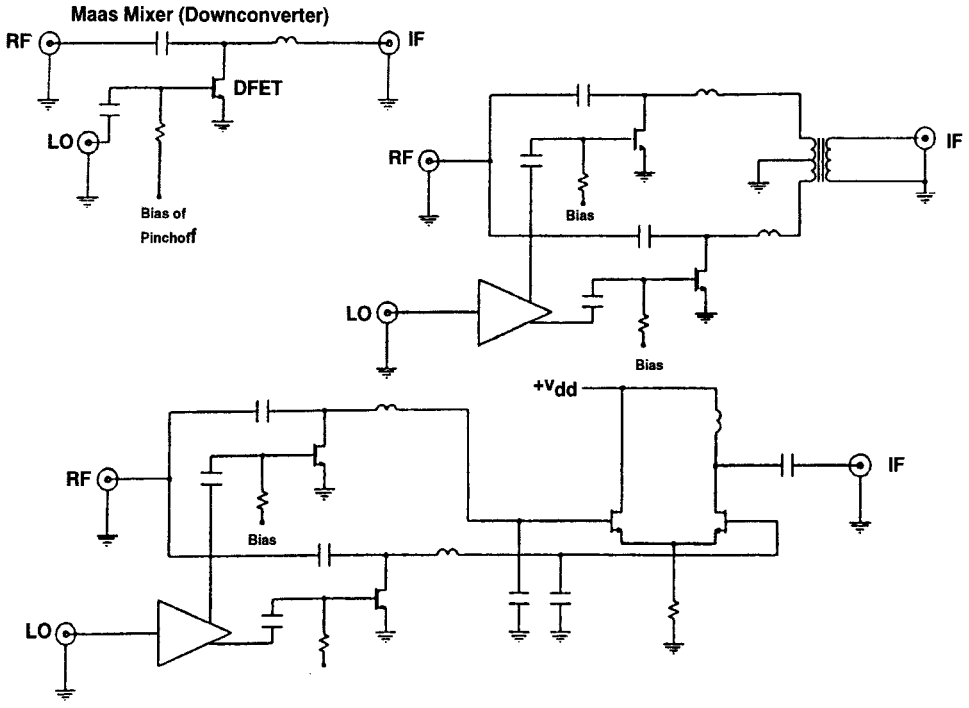


Figure 4-70a

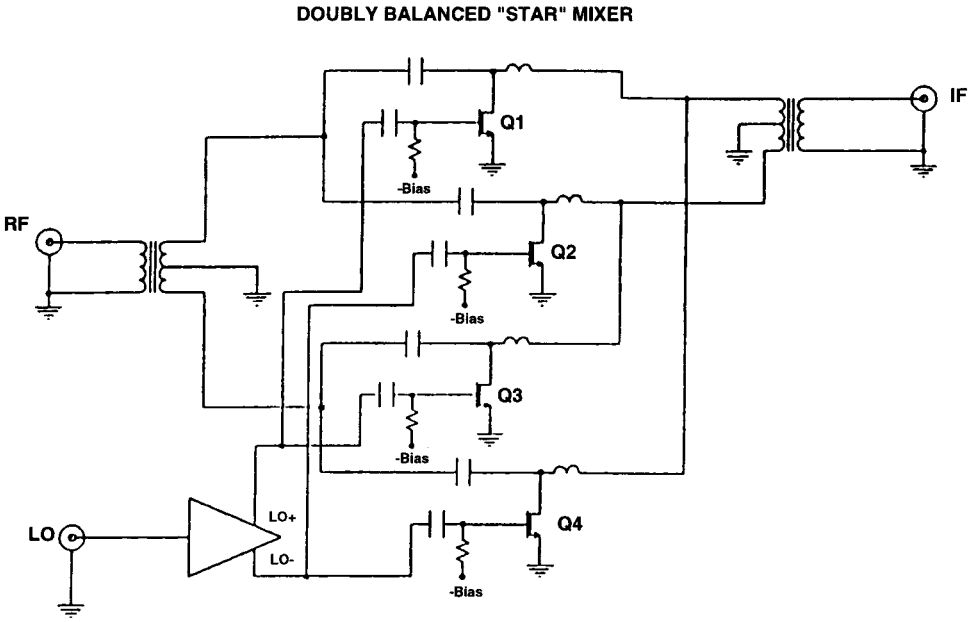


Figure 4-70b

to pass and has a parallel tuned circuit that keeps the input signal from getting into the IF stage. The resulting conversion loss of 5.5 dB is consistent with the example we presented before of 6 dB, and while the noise figure also is stated as 5.5 dB, that value is deceiving because it was determined using a 120-MHz IF. If an IF of 10 MHz or less was used, the noise figure would be tremendously worse. This is due to the frequently mentioned flicker-noise contribution. As a comparison, Figure 4-67 shows the predecessor to the CMY210, the CMY91, which includes a built-in IF amplifier. Table 4-6 shows its partial specifications. Since it has 5.5-dB gain, one can assume that the IF amplifier has something like 10 dB of gain, and the mixer portion, 6 dB losses. The resulting noise figure is the sum of the mixer noise figure and the IF amplifier noise figure. The IF transistor is probably not "noise matched" with the mixer, but the resulting 9-dB noise figure is not uncommon in this arrangement. Judging by previous examples, we would say that the noise figure of the mixer is somewhere around 6 dB, and therefore the mismatched (open at the input) IF amplifier has a noise figure of 3 dB, resulting in this total noise figure of 9 dB as stated. The third-order IP of –2 dBm is no match to the CMY210's $IP_{3,in}$ performance, which is typically 25 dBm.

For readers who are interested in what mixers look like when pushed well into the microwave region, we present the layout for a 20-GHz diode mixer. Figure 4-68 presents its layout, which includes radial stubs, circular stubs, and a Lange coupler. Figure 4-69 shows the layout implemented on a ceramic substrate.

As further food for thought and experimentation, Figure 4-70 presents mixer and related circuit notes as presented by Wes Hayward in TriQuint Semiconductor's GaAs Design Class.

HA Mixer Family Measurements

The mixers are all identical except for FET width. Mixers used the HA process and are half micron gate length. Bias was adjusted to best gain at +6 LO. The LO baluns were half wave lines of semi-rigid coax. F–LO = 1.65, RF = 1.9, and IF = 0.25 GHz. Input Power is –13 or –10 dBm RF. All measurements show downconversion results.

FET Width	P-av RF	P-LO dBm	P-out at IF	IMDR dB	Conv Gain dB	IP3-out dBm	IP3in dBm
50	–13	6	–31	44.7	–11.7	–2.35	9.35
50	–13	10	–30.5	53.5	–11.2	2.55	13.75
50	–10	10	–27.8	49	–11.5	3	14.5
100	–10	6	–25.7	41.5	–9.4	1.35	10.75
100	–10	10	–24.8	48	–8.5	5.5	14.0
100	–10	14	–24.5	68.8	–8.2	16.2	24.4
200	–10	6	–24.3	40.5	–8	2.25	10.25
200	–10	10	–23.7	45	–7.4	5.1	12.5
200	–10	14	–23.5	60.8	–7.2	13.2	20.4
200	–10	17	–24	60.7	–7.7	12.65	20.3
400	–10	6	–24	42.8	–7.7	3.7	11.4
400	–10	10	–23.5	52	–7.2	8.8	16.0
400	–10	14	–23.3	67.7	–7	16.85	23.85
400	–10	17	–23.2	57.5	–6.9	11.85	18.75
50	–10	6	–30.8	56	–14.5	3.5	18.0
50	–10	6	–28.6	42.5	–12.3	–1.05	11.25
50	–10	10	–28	49.2	–11.7	2.9	14.6
50	–10	14	–28.2	60.8	–11.9	8.5	20.4
50	–10	17	–28.3	63.5	–12	9.75	21.75
50	–10	17	–28	62.8	–11.7	9.7	21.4
100	–10	6	–29.5	42.5	–13.2	–1.95	11.25
100	–10	10	–25	47.2	–8.7	4.9	13.6
100	–10	14	–24.7	63.7	–8.4	13.45	21.85
100	–10	16	–25.5	62	–9.2	11.8	21.0

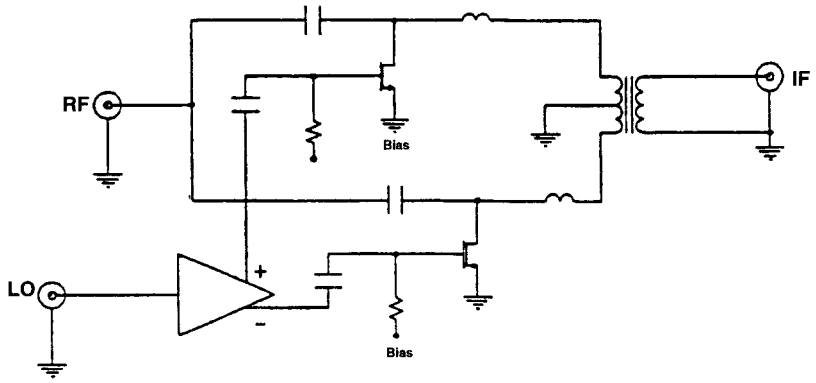
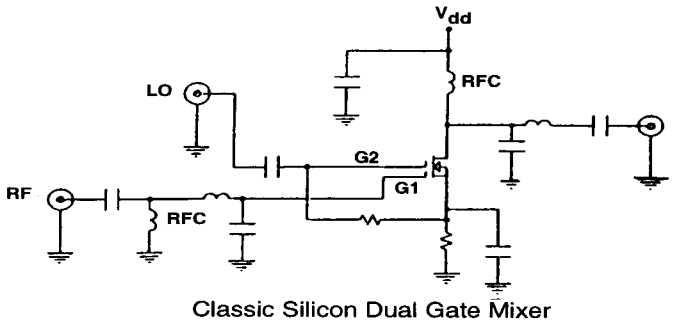


Figure 4-70c



Classic Silicon Dual Gate Mixer

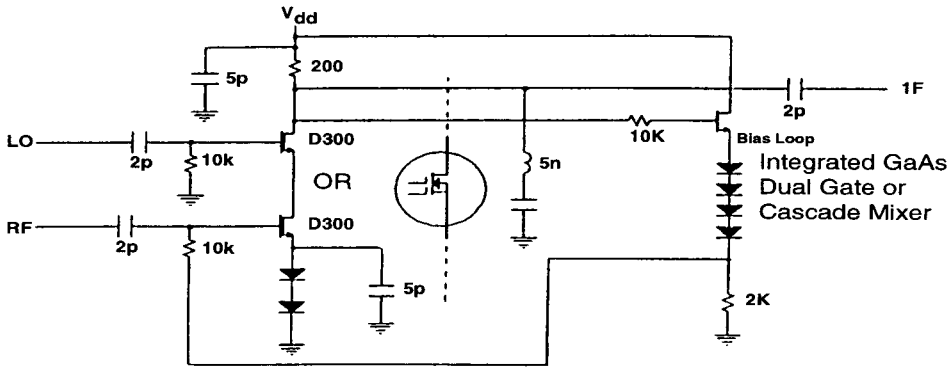


Figure 4-70d

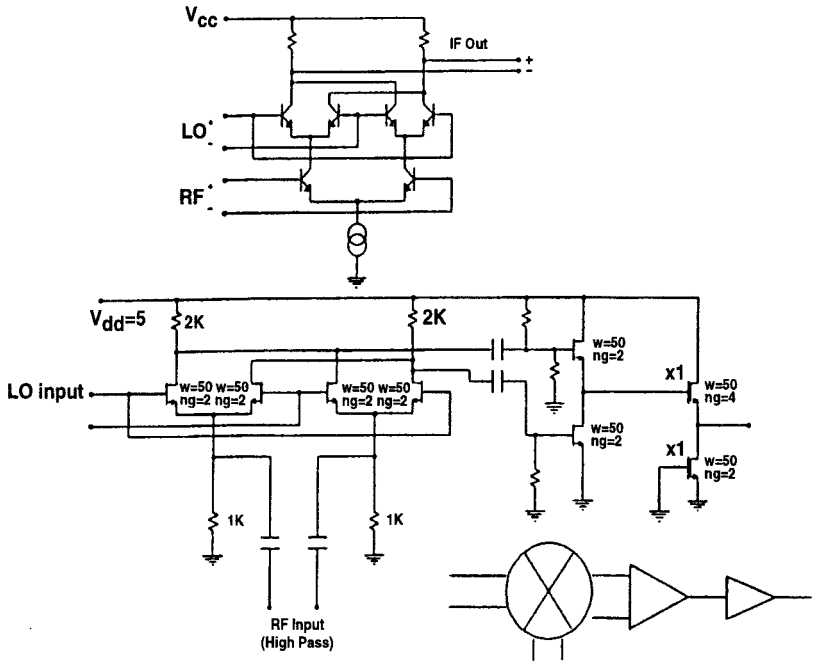


Figure 4-70e

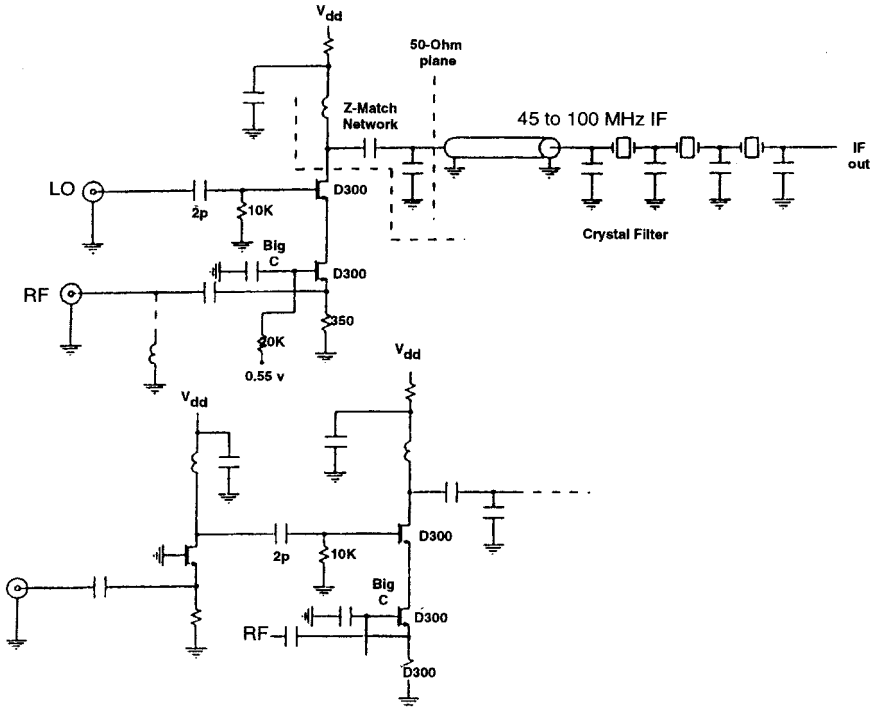


Figure 4-70f

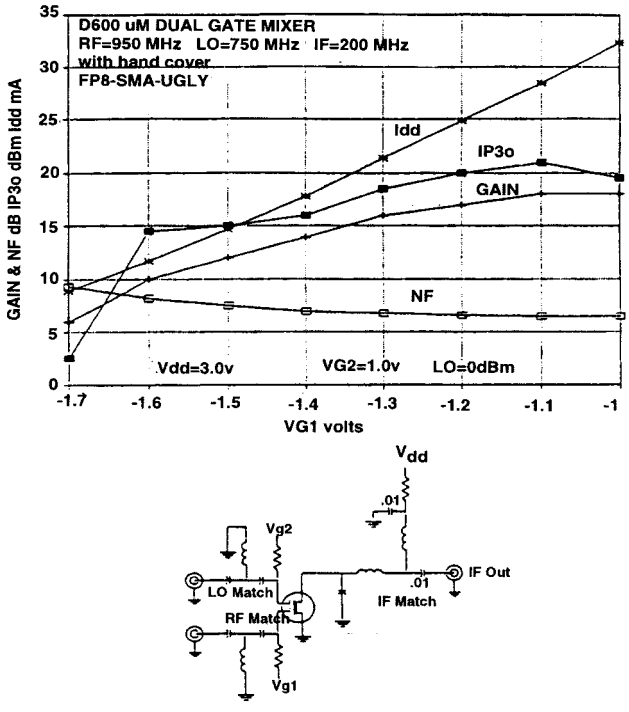


Figure 4-70g

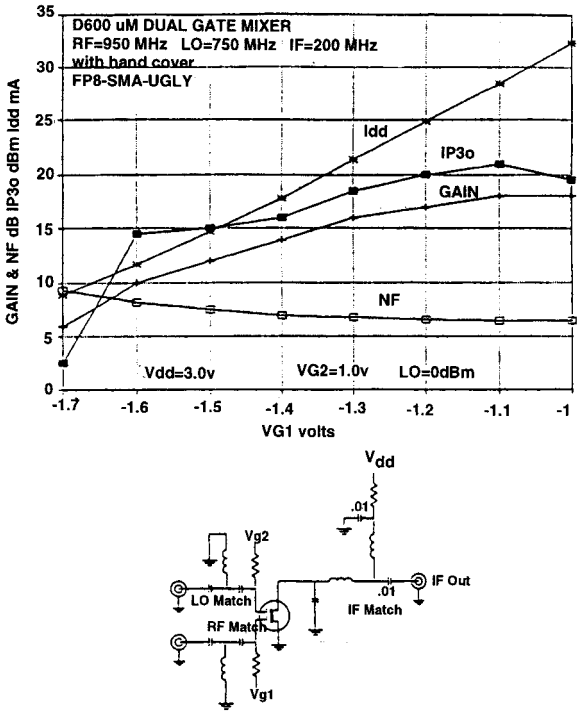


Figure 4-70h

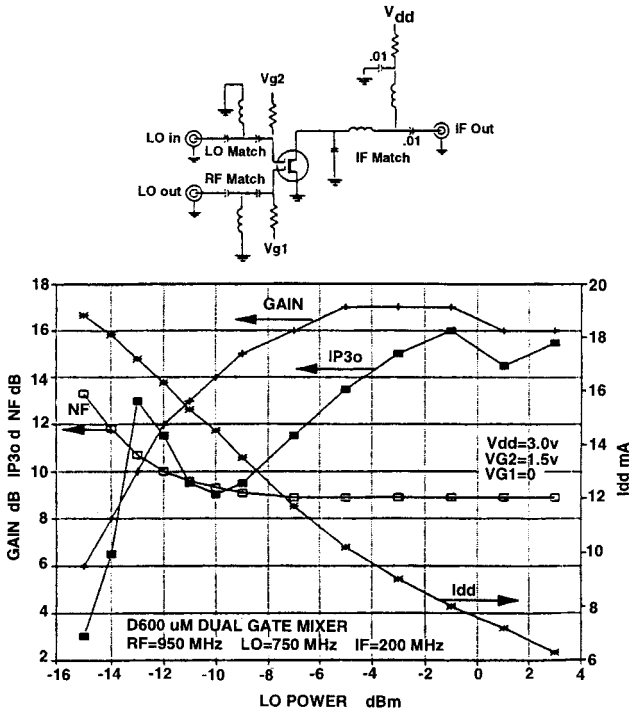


Figure 4-70i

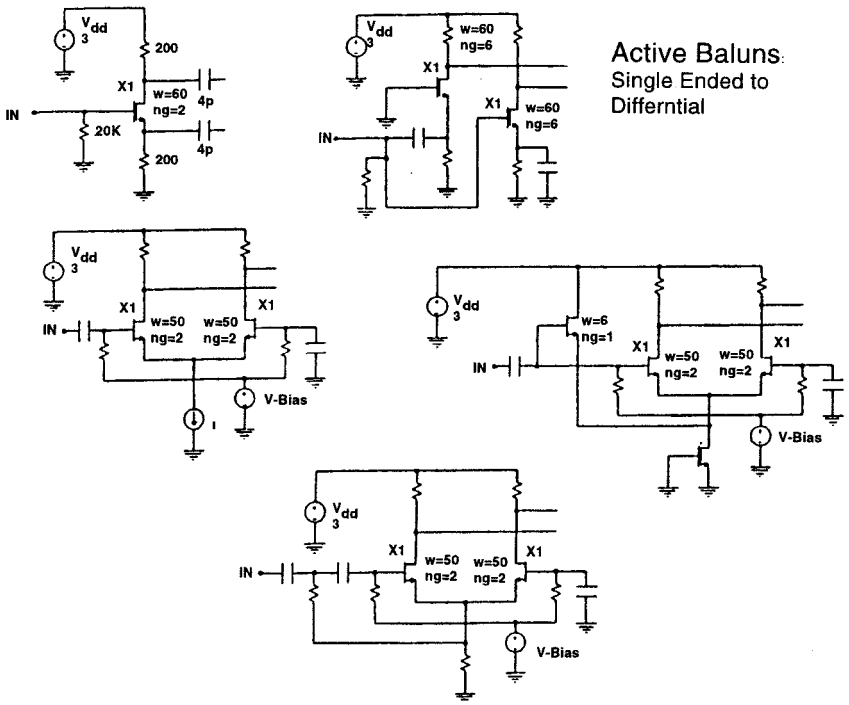


Figure 4-70j

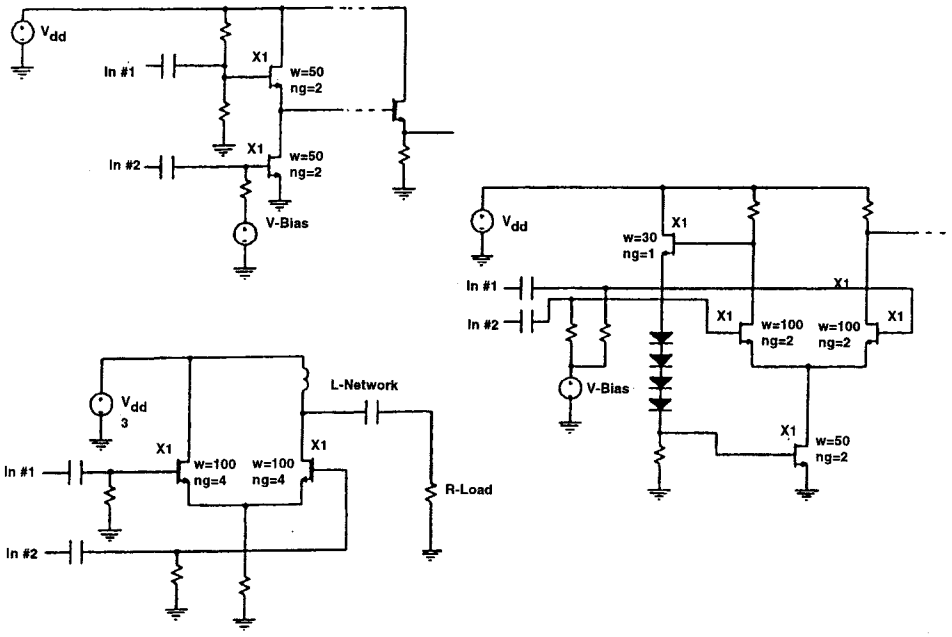


Figure 4-70k

REFERENCES

1. Otto Zinke and Hans Ludwig Hartnagel, *Lehrbuch der Hochfrequenztechnik*, *Band 2: Elektronik und Signalverarbeitung.* New York: Springer Verlag, 1987, pp. 429–430.
2. IRE Subcommittee 7.9 on Noise, "Description of Noise Performance in Receiving Systems," *Proc. IRE*, **51**, 436, 1963.
3. Stephen A. Maas, *Microwave Mixers*, 2nd ed. Boston: Artech House, 1993, pp. 153–171.
4. O. Zinke and H. Brunswig, *Lehrbuch der Hochfrequenztechnik*, Vol. 2, 3rd ed. New York: Springer Verlag, 1987, p. 435.
5. Cheng Qian, "Subharmonic Mixers Simplify Digital-Modulation Systems," *Microwaves & RF*, **37**(9), 62–70, September 1998.
6. Kenji Itoh et al., "Even-Harmonic Quadrature Modulator with Low Vector Modulation Error and Low Distortion for Microwave Digital Radio," *Microwave Theory and Techniques Society Symposium Digest*, 1996, pp. 967–970.
7. Shankar R. Joshi, "Novel I/Q Modulators Mix Cellular Signals," *Microwaves & RF*, **33**(11), 119–123, November 1994. Also see Shankar R. Joshi, "Modulator with Harmonic Mixers," Synergy Microwave Corporation, Paterson, NJ, U.S. Patent No. 5,416,449, May 16, 1995.
8. M. V. Schneider, R. Trambarulo, A. Gnauck, and M. J. Gans, "High-Speed Millimeter-Wave Modulator/Demodulator," *Microwave Theory and Techniques Society Symposium Digest*, 1990, pp. 611–614.
9. Jerzy Charamiec, "Subharmonically Pumped Schottky Diode Single-Sideband Modulator," *IEEE Trans. Microwave Theory Tech.*, **26**(9), 635–638, September 1978.
10. Ian Doyle, "A Simplified Subharmonic I/Q Modulator," *Appl. Microwave & Wireless*, **10**(8), 34–40, October 1998.
11. Peter Will, "Broadband Doubly Balanced Mixer Having Improved Termination Insensitivity Characteristics," Adams-Russell Co., Inc., Waltham, MA, U.S. Patent No. 4,224,572, September 23, 1980.
12. Joseph T. Lee, "Balanced Subharmonic Mixers," *Microwave J.*, **26**(8), August 1983.
13. Don Neuf, "Fundamental Versus Harmonic Mixing," *Microwave J.*, 1984.
14. Bert Henderson, "Full-Range Orthogonal Circuit Mixers Reach 2 to 26 GHz," *Microwave Syst. News*, January 1982.
15. Barry Gilbert, *Demystifying the Mixer* (self-published monograph), 1994, p. 27.
16. Kenneth K. Clarke and Donald T. Hess, *Communication Circuits: Analysis and Design*. Reading, MA: Addison-Wesley, 1971, p. 115.
17. George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990.
18. G. K. Tie and C. S. Aitchison, "Noise Figure and Associated Conversion Gain of a Microwave MESFET Gate Mixer," *Conference Proceedings of the 13th European Microwave Conference*, 1983.

FURTHER READING

Darrell L. Ash, "High-Performance TV Receiver," Texas Instruments Inc.

Sharief Babiker, Asen Asenov, Nigel Camerson, Stevfen P. Beaumont, and John R. Barker, "Complete Monte Carlo RF Analysis of 'Real' Short-Channel Compound FET's," *IEEE Trans. Electron Devices*, **Ed-45**(8), 1644–1652, August 1998.

Marc Borremans, Michel Steyaert, and Takashi Yoshitomi, "A 1.5 V, Wide Band 3 GHz, CMOS Quadrature Direct Up-Converter for Multi-Mode Wireless Communications," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference* , pp. 79–82.

A. Brunel, E. Douglas, K. Gallagher, J. Gillis, M. Goff, S. LeSage, E. Lewis, R. Pengelly, C. Ruhe, and J. Sano, "A Downconverter for Use in a Dual-Mode AMPS/CDMA Chip Set," *Microwave J.*, 20–42, February 1996.

P. Y. Chan, A. Rofougaran, K. A. Ahmed, and A. A. Abidi, "A Highly Linear 1-GHz CMOS Downconversion Mixer," *Proc. ESSCIRC*, Sevilla, September 1993, pp. 210–213.

Jan Crols and Michel S. J. Stayaert, "A 1.5-GHz Highly Linear CMOS Downconversion Mixer," *IEEE J. Solid-State Circuits*, **SC-30**(7), 736–742, July 1995.

Suman Datta, Shen Shi, Kenneth P. Roenker, Marc M. Chay, and William E. Stanchina, "Simulation and Design of InAlAs/InGaAs *pnp* Heterojunction Bipolar Transistors," *IEEE Trans. Electron Devices*, **Ed-46**(8), 1634–1641, August 1998.

Maria Luisa de la Fuente, Juan Pablo Pascual, and Eduardo Artal, "Analyze and Design a Cascode MESFET Mixer," *Microwaves & RF*, **37**(5) 129–138, May 1998.

Ravendar Goyal, *High-Frequency Analog Integrated Circuit Design* . New York: Wiley, 1995.

W. Hayward and S. Taylor, "Compensation Method and Apparatus for Enhancing Single Ended to Differential Conversion," TriQuint Semiconductor, Inc., Beaverton, OR, U.S. Patent No. 5,068,621, November 26, 1991.

Hewlett-Packard, Solutions from HP EEsof, *Low-Power Mixing Design Example Using HP Advanced Design System*, 1995.

K. Kanazawa et al., "A GaAs Double-Balanced Dual-Gate FET Mixer IC for UHF Receiver Front-End Applications," *IEEE MTT International Microwave Symposium Digest* , 1985, pp. 60–62.

K. Kimura, "Bipolar Four-Quadrant Analog Quarter-Square Multiplier," *IEEE J. Solid-State Circuits*, **29**(1), 46–55, January 1994.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

R. G. Meyer, "Intermodulation in High-Frequency Bipolar Transistor Integrated-Circuit Mixers," *IEEE J. Solid-State Circuits*, **SC-21**(4), 534–537, August 1986.

Ram Mohan Narayanan, "Multioctave, Double-Balanced Mixers Designed Using MIC Technology," *MSN & CT,* 108–115, January 1987.

D. Rabaey and J. Sevenhans, "The Challenges for Analog Circuit Design in Mobile Radio VLSI Chips," *Proceedings of the AACD Workshop* , Leuven, March 1993, Vol. 2, pp. 225–236.

J. Sevenhans, A. Vanwelsenaers, J. Wenin, and J. Baro, "An Integrated Si Bipolar Transceiver for a Zero IF 900 MHz GSM Digital Mobile Radio Front-End of a Hand Portable Phone," *Proc. CICC*, May 1991, pp. 7.7.1–7.7.4.

J. Sevenhans et al., "An Analog Radio Front-End Chip Set for a 1.9 GHz Mobile Radio Telephone Application," *Proc. ISSCC*, San Francisco, February 1994, pp. 44–45.

W. Titus, L. Holway, R. A. Pucel, P. Bernkopf, and A. Palevsky, "Broadband Control Components—FET Mixers," Final Technical Report, October 1988, Contract No. F33615-85-C-1725, PAY/RD/S-4270, Raytheon Company Research Division, Lexington, MA.

TQ9203 Mixer Noise Figure Measurement Note, TriQuint Semiconductor, Inc.

TriQuint Semicondutor GaAs Design Class Notes, Section 5 (Wes Hayward, "RFIC Lore and Philosophy" and "Non-Linear Designs") and Section 7 (Bob Hickey, "Design Example: Gilbert Cell Mixer with Active Baluns").

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

4

MIXER DESIGN

4-1 INTRODUCTION

Radiocommunication requires that we shift a baseband information signal to a frequency or frequencies suitable for electromagnetic propagation to the desired destination. At the destination, we reverse this process, shifting the received radiofrequency (RF) signal back to baseband to allow the recovery of the information it contains. This frequency-shifting function is traditionally known as *mixing*; the stages that perform it are known as *mixers*. Any device that exhibits amplitude-nonlinear behavior can serve as a mixer, for, as we saw in Section 1-6-2, nonlinear distortion results in the production, from the signals present at the input of a device, of signals at new frequencies. Even a rusty screw or bolt on an antenna element can act as a mixer, producing unwanted IMD products that appear at the receiver input.

Although mixers are equally important in wireless transmission and reception, traditional mixer terminology favors the receiving case because mixing was first applied as such in receiving applications. Thus, the signal to be frequency shifted is applied to the mixer's RF port, and the frequency-shifting power or voltage [from a *local oscillator* (*LO*)] is applied to the mixer's LO port, resulting in two outputs at the mixer's *intermediate frequency* (*IF*) port. If the wanted IF is lower than the RF signal, the mixer is a *downconverter*; if the wanted IF is higher than the RF, the mixer is an *upconverter*. *Converter* may also be used as a term for a single stage that simultaneously acts as mixer and LO.

For a given RF signal, an ideal mixer with a perfect LO (i.e., an LO with no harmonics and no noise sidebands) would produce only two IF outputs: one at the frequency sum of the RF and LO, and another at the frequency difference between the RF and LO. Filtering can be used to select the desired IF product and reject the unwanted one, which is sometimes referred to as the *IF image*.

The simultaneous generation of LO + RF and LO – RF outputs results not from a departure of mixer performance from the ideal, but from the mathematics of mixing itself. Another unavoidable mixing artifact, the *RF image* response, also results from the mathematics of mixing rather than mixer nonideality. Just as a given RF/LO combination produces two IF outputs (LO + RF and LO – RF, the IF and IF image), the mixer will produce output at the

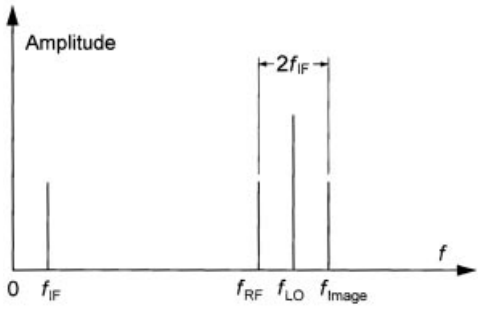


Figure 4-1 Relationship between a mixer's image and desired-signal responses. The image is $2f_{IF}$ away from the desired signal.

desired IF (LO + RF *or* LO – RF) in response to *two* possible RF inputs: one at LO + IF and another at LO (IF (Figure 4-1). The undesired response, the *RF image* (traditionally referred to merely as the *image*), is $2f_{IF}$ removed from the desired response. Even if no human-generated signals exist at the RF image frequency, reducing a mixer's RF image response can be important because noise at that frequency, including that produced by circuitry between the mixer and antenna, will still be mixed to the desired IF, degrading the signal-to-noise ratio. Filtering and phasing techniques can be used to reduce the RF or IF image responses: filtering if the image is sufficiently removed from the desired response that filtering will provide the necessary rejection; phasing if the desired and image responses are insufficiently spaced for filtering to work, as in the case of a double-conversion receiver in which signals at a high first IF (e.g., 50–70 MHz) must be converted to a very low first IF, such as 25 kHz.

The output of every real mixer includes a vast number of additional unwanted products, including noise, the fundamentals of the mixer's RF and LO signals and their harmonics, and the sums and differences of the RF and LO and their harmonics. Intermodulation distortion between multiple signals present at the RF port and IF output resulting from the mixing to IF of LO noise-sideband energy by strong adjacent signals (*reciprocal mixing*, Section 1-6-2) further complicate a mixer's output spectrum and may compromise system performance.

All mixers are *multipliers* in the sense that the various new outputs they produce can be described mathematically as the multiplicative products of their inputs. From an implementation standpoint, however, a given mixer circuit can be characterized as additive or multiplicative depending on how RF and LO signals are applied to it. Additive mixing occurs

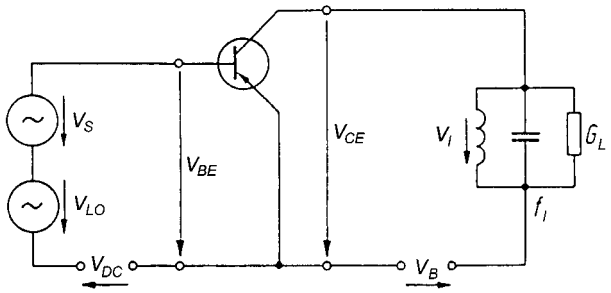


Figure 4-2 Additive mixing in a BJT [1].

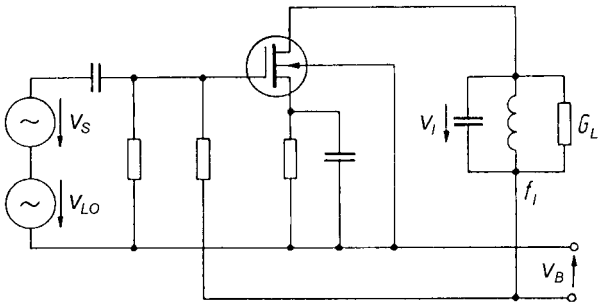


Figure 4-3 Additive mixing in a single-gate MOSFET [1].

when the RF and LO signals are applied to the same input port, as in Figures 4-2 and 4-3. Multiplicative mixing occurs when the RF and LO signals are applied to separate ports, as in Figure 4-4. As a rule, multiplicative mixers afford better isolation between their LO and RF ports than additive mixers, and this enhanced interport isolation is their principal merit. Multiplicative mixing does not in itself suppress unwanted products; the spurious response of a basic multiplicative mixer cell is poor unless it is used in a push–pull or quad configuration.

Let us now consider the basic theory of mixers. Mixing is achieved by the application of two signals to a nonlinear device. Depending on the particular device, the nonlinear characteristic may differ. However, it can generally be expressed in the form

$$I = K(V + v_1 + v_2)^n \tag{4-1}$$

The exponent n is not necessarily an integer, V may be a dc offset voltage, and the signal voltages v_1 and v_2 may be expressed as $v_1 = V_1 \sin(\omega_1 t)$ and $v_2 = V_2 \sin(\omega_2 t)$.

When $n = 2$, Eq. (4-1) may then be written as

$$I = K[V + V_1 \sin(\omega_1 t) + V_2 \sin(\omega_2 t)]^2 \tag{4-2}$$

This assumes the use of a device with a square-law characteristic. A different exponent will result in the generation of other mixing products, but this is not relevant for a basic understanding of the process. Expanding Eq. (4-2), we find

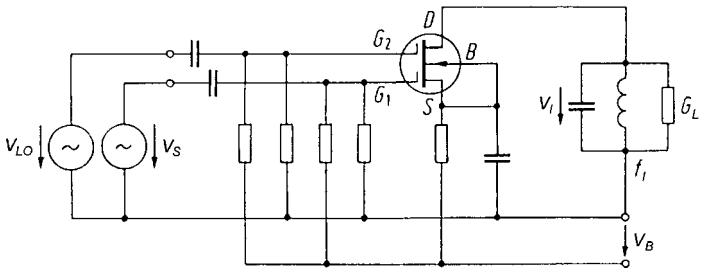


Figure 4-4 Multiplicative mixing in a dual-gate MOSFET. A dual-gate device is actually two single-gate devices in series [1].

$$I = K[V^2 + V_1^2 \sin^2(\omega_1 t) + V_2^2 \sin^2(\omega_2 t) + 2\,VV_1 \sin(\omega_1 t)$$

$$+ 2\,VV_2 \sin(\omega_2 t) + 2\,V_2V_1 \sin(\omega_2 t) \sin(\omega_1 t)] \tag{4-3}$$

The output comprises a direct current and a number of alternating current contributions. We are interested only in that portion of the current that generates the IF; so, if we neglect those terms that do not include both V_1 and V_2, we may write

$$I_{\text{IF}} = 2\,KV_1V_2 \sin(\omega_1 t) \sin(\omega_2 t)$$

$$I_{\text{IF}} = KV_2V_1\{\cos[(\omega_2 - \omega_1)t] - \cos[(\omega_2 + \omega_1)t]\} \tag{4-4}$$

This means that, at the output, we have the sum and difference signals available, and the one of interest may be selected by the IF filter.

4-2 PROPERTIES OF MIXERS

4-2-1 Conversion Gain/Loss

Even though a mixer works by means of the amplitude-nonlinear behavior in its device(s), we generally want (and expect) it to act as a linear frequency shifter. The degree to which the frequency-shifted signal is attenuated or amplified is an important mixer property. *Conversion gain* can be positive or negative; by convention, negative conversion gains are often stated as *conversion loss*.

In the case of a diode (passive) mixer , the insertion loss is calculated from the various loss components:

$$\text{Loss (dB)} = \text{Conversion loss} + \text{Transformer loss}$$

$$+ \text{Losses due to harmonic generation} + \text{Diode loss} \tag{4-5}$$

In the case of a double-balanced mixer (DBM), we must add the transformer losses (on both sides) and the diode losses as well as the mixer sideband conversion, which accounts, by definition, for 3 dB. Ideally, the mixer produces only one upper and one lower sideband, which results in the 3-dB loss compared to the input signal. Also, the input and output transformers add about 0.75 dB on each side, and of course there are the diode losses because of the series resistances of the diodes.

Figure 4-5 shows the equivalent circuit of a diode. It consists of a series (loss) resistor R_S and a time-variable electronic resistor, typically called the *diffusion resistance*, R_D, which equals 26 mV/I_D, and a capacitance C_D shunting R_D. C_D can be found from

$$C_D = \frac{W^2}{2D}\frac{I_D}{V_T} \tag{4-6}$$

where D is the diffusion constant, a material-dependent value, and W is the physical width. The average value for R_D is somewhere between the calculated value of 26 mV/I_D and some

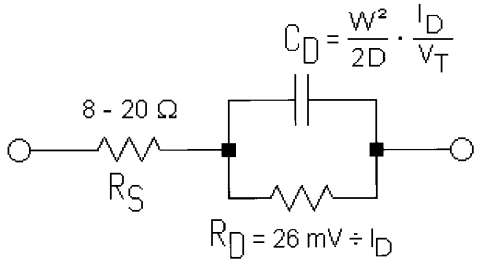


Figure 4-5 Equivalent circuit of a mixer diode.

leakage current, simply because it is generated by a rectification mechanism, which turns the LO power into an RF current and then into a combination of dc and RF currents.

We can calculate the diode loss according to

$$\text{Diode loss (db)} = \log 10\left(\frac{50 + (2 \times R_S)}{50}\right) \tag{4-7}$$

Assuming that $R_S = 8\ \Omega$, the diode loss for a diode-ring mixer is

$$\text{Diode loss (db)} = \log 10\left(\frac{50 + (2 \times 8)}{50}\right) = 0.5\ \text{dB} \tag{4-8}$$

From Eq. (4-5), the insertion loss for this mixer is therefore

$$\begin{aligned}\text{Loss (dB)} &= 3\ \text{dB (Conversion loss)} + 1.5\ \text{dB (Transformer loss)} \\ &\quad + 1\ \text{dB (Losses from harmonic generation)} + 0.5\ \text{dB (Diode loss)} \\ &= 6\ \text{dB}\end{aligned} \tag{4-9}$$

This assumes mixing at the fundamental frequency. Sometimes the diode loss resistor R_S is as high as 25 Ω per arm (as in a MOSFET switch), or 50 Ω total. This now results in

$$\text{Loss (dB)} = 3\ \text{dB} + 1.5\ \text{dB} + 1\ \text{dB} + 3\ \text{dB} = 8.5\ \text{dB} \quad \text{(Insertion loss)} \tag{4-10}$$

Since the value of R_S is partially determined by the threshold voltage of the diode and the diode diffusion resistance, R_D, a wide range of values can be noticed for different drive levels and mixer topologies. Figure 4-5 shows a shunt capacitance C_D, the so-called *diode diffusion capacitance*. When the diode is conducting, the influence of this nonlinearity is frequency dependent, which adds to the insertion loss. In this discussion we have not considered its frequency dependency. At wireless frequencies, modern Schottky diodes, also frequently called *hot-carrier* diodes, are operated far from their cutoff frequency, resulting in less than 1 dB of additional losses. There are also mixers with special circuitry to terminate the IF image. This is done with a diplexer circuit or equivalent circuitry. This makes insertion loss values as low as 4 dB possible; however, the large-signal condition or intercept point suffers.

4-2-2 Noise Figure

Like any network, a mixer contributes noise to the signals that it frequency-shifts. The degree to which a mixer's noise degrades the signal-to-noise ratio (see Section 1-7-1) of the signals it frequency-shifts is evaluated in terms of noise factor and noise figure as discussed in a subsection of Section 1-7-1 entitled "Noise Figure of Cascaded Networks."

For a long time, the literature has stated that the noise figure of a passive mixer, which is pretty much independent of its circuit arrangement, is equal to the mixer's insertion loss. But this neglects the influence of the white-noise contribution of the mixer's diode(s). This is ironic, considering that RF noise generators were long based on thermionic diodes operated in saturation (the 5722 noise diode was a popular type, as in the Rohde & Schwarz SKTU). Such a diode's noise power output can readily be determined from its saturation current. On the other hand, all Schottky diodes, while conducting, generate white noise that follows the same principle as above. This fact has been recognized by only a few companies that make modern noise-measurement equipment. Modern noise-measurement devices measure the noise figure of a system by a "hot and cold" technique, an approach based on knowledge of the absolute noise energy emitted under hot conditions (conductance). This method has the advantage that it can be used up to several tens of gigahertz, while the old vacuum-tube-based noise generators ran out of steam at around 1 GHz due to the inability to match the tube to the 50-Ω termination. This was typically accomplished by connecting a 50-Ω resistor between anode and ground (without dc connection), followed by a low-pass filter, which would match the tube capacitance and other parasitics to the required termination of 50 Ω, purely resistive.

In reality, we can take the loss calculation from above and add the Schottky noise generated by the diodes as they are driven by the local oscillator.

If a Schottky diode is used as a noise generator in the conductive mode, it generates a continuous frequency spectrum, possibly up to several gigahertz. There is a mathematical relationship between the noise power spectrum emitted by the diode and the time-averaged current of this diode, which generates the noise. If the noise source impedance is set (typically 50 Ω), the available noise power can be calculated according to

$$I_R = \sqrt{2e \times I_s \times \Delta f} \tag{4-11}$$

where $e = 1.6 \times 10^{-19}$ coulombs
I_s = saturation current of the diode
Δf = effective noise bandwidth

For $S_{11} = 0$ or proper termination of this circuit ($R_G = R_{term}$),

$$P_R = \left(\frac{I_R}{2}\right)^2 \times R_i$$

$$= \frac{e}{2} \times I_s \times \Delta f \times R_i \tag{4-12}$$

Calculated at a bandwidth of 1 Hz,

$$\frac{P_R}{\Delta f} = \frac{e}{2} \times I_s \times R_i \tag{4-13}$$

If

$$\frac{P_R}{\Delta f} = kT_0 \times F \tag{4-14}$$

the noise factor becomes

$$F = \frac{e \times I_s \times R_i}{2kT_0} \tag{4-15}$$

If the values for e and kT_0 are inserted,

$$F = 20 \times I_D \times \frac{26 \text{ mV}}{I_D} + \frac{R_S + R_G}{2R_G} \tag{4-16}$$

Example. Assume the passive mixer mentioned above with its 6-dB insertion loss is considered and a dc current of 15 mA results as a function of the LO drive. Since I_D gets canceled, the noise factor (F) of the diode portion equals

$$F = 20 \times I_D \times \frac{26 \text{ mV}}{I_d} + \frac{R_S + R_G}{2R_G}$$

$$= 0.52 + 0.58$$

$$= 1.1 \tag{4-17}$$

The noise figure, NF, is 10 log F, or 0.413 dB. Now this number and the insertion loss must be added. The resulting noise figure would be 6.413 dB. This is consistent with published measurement data.

Exact Mathematical Nonlinear Approach. The exact noise factor of a real mixer is computed by the formula

$$F = \frac{N_0(\omega_{\text{IF}}) + kT_0}{K_B T_0 G_{Tc}(\omega_{\text{RF}})} \tag{4-18}$$

where $N_0(\omega_{\text{IF}})$ = total noise power (per unit bandwidth) delivered to the IF load at intermediate frequency
K_B = Boltzmann's constant
T_0 = reference temperature (290 or 300 K is commonly used)
$G_{Tc}(\omega_{\text{RF}})$ = transducer conversion gain from ω_{RF} to ω_{IF}.

Let us now further elaborate on Eq. (4-18). We may write

$$N_0(\omega_{\text{IF}}) = N_S(\omega_{\text{IF}}) + N_{\text{INT}}(\omega_{\text{IF}}) + N_L(\omega_{\text{IF}}) + kT_0 \tag{4-19}$$

where N_S is noise generated by the RF source resistance and transferred to the IF load through frequency conversion, N_{INT} is noise generated internally to the mixer, and N_L is noise generated by the IF termination. If the source resistance is held at temperature T_0, N_S will basically originate from noise generated at the RF and image frequencies, which are transferred to the IF with approximately the same conversion gain, plus a relatively small contribution transferred from other sidebands with a smaller conversion gain. We may then write synthetically

$$N_S(\omega_{\text{IF}}) = 2aK_BT_0G_{Tc}(\omega_{\text{RF}}) \tag{4-20}$$

where a is a coefficient slightly larger than 1. $N_{\text{INT}}(\omega_{\text{RF}})$ is generated by transformer losses, by the diode Schottky noise, and by the diode resistive parasitics, and, in principle, it may take on any value; in particular, it may be zero if both the transformers and the diodes are ideal (i.e., if the latter are pure nonlinear resistors). As for $N_L(\omega_{\text{IF}})$, by Nyquist's theorem the IF load resistor R_L may be described as a noiseless resistor in series with a noise voltage source whose mean-square voltage (per unit bandwidth) is

$$|V_L|^2 = 4K_BT_LR_L \tag{4-21}$$

where T_L is the IF termination temperature. If the IF load is driven by a source with an output impedance $Z_{\text{out}}(\omega_{\text{IF}})$, the noise power actually delivered to the load will obviously be

$$N_{\text{out}} = \frac{4K_BT_LR_L^2}{|Z_{\text{out}}(\omega_{\text{IF}}) + R_L|^2} \tag{4-22}$$

In addition to N_{out}, the thermal noise originating from the IF termination delivered to the IF load at ω_{IF} will also include contributions from other sidebands that are backconverted by the mixer nonlinearities with a relatively small conversion gain. Thus, we may write

$$N_L(\omega_{\text{IF}}) = \frac{4bK_BT_LR_L^2}{|Z_{\text{out}}(\omega_{\text{IF}}) + R_L|^2} \tag{4-23}$$

where b is slightly larger than 1. If we now introduce the mixer conversion loss, namely,

$$L_C = \frac{1}{G_{Tc}(\omega_{\text{RF}})} \tag{4-24}$$

and combine Eq. (4-18) with Eqs. (4-19), (4-20), and (4-23), we finally get the noise factor expression:

$$F = 2a + \frac{N_{\text{INT}}(\omega_{\text{IF}}) + kT}{K_BT_0}L_C + \frac{4bT_LR_L^2}{T_0|Z_{\text{out}}(\omega_{\text{IF}}) + R_L|^2}L_C \tag{4-25}$$

In the normal region of operation of the mixer (sufficient LO drive) we may assume

Table 4-1 Noise figure and conversion gain versus LO power for a diode DBM

LO Power (dBm)	NF (dB)	Conversion Gain (dB)
–10.0	45.3486	–45.1993
–8.0	32.7714	–32.5264
–6.0	19.8529	–19.2862
–4.0	12.1154	–11.3228
–2.0	8.85188	–8.05585
0.0	7.26969	–6.51561
2.0	6.42344	–5.69211
4.0	5.85357	–5.15404
6.0	5.50914	–4.84439
8.0	5.31796	–4.66871
10.0	5.19081	–4.54960
12.0	5.08660	–4.45887
14.0	4.99530	–4.38806
16.0	4.91716	–4.33322
18.0	4.85920	–4.29407
20.0	4.82031	–4.26763

$$Z_{\text{out}}(\omega_{\text{IF}})R_L \tag{4-26}$$

so that Eq. (4-25) becomes

$$F = 2a + \frac{N_{\text{INT}}(\omega_{\text{IF}}) + kT}{K_B T_0} L_C + \frac{bT_L}{T_0} L_C \tag{4-27}$$

Multiplying $\log_{10} F$ by 10 gives us the exact mixer noise figure in dB.

The method just discussed is the basis for the far noise calculation for oscillators as mentioned in Chapter 5, in which it is referred to as conversion noise. This includes the AM-to-PM conversion noise.

Table 4-1 shows how the noise figure and conversion gain vary with LO power for a generic diode DBM (Figure 4-6). "Starving" a diode mixer by decreasing its LO drive rapidly degrades its performance in all respects.

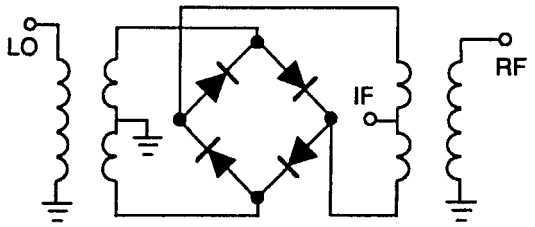


Figure 4-6 Generic diode DBM.

SSB Versus DSB Noise Figure. In radiocommunication, we typically use mixers as "single-sideband" devices; that is, we are interested in only one of the two possible RF inputs ($f_{LO} + f_{IF}$ or $f_{LO} - f_{IF}$) that produce output at the desired IF. The unused input is the *RF image* or simply *image*. A mixer's RF image response can be important even if no human-generated signals are present at f_{image} because noise there, including that produced by the mixer's RF-port termination, will be converted to the IF, possibly compromising the system's signal-to-noise ratio. As a result of this and the IEEE definition of noise figure [2], which, for a mixer, considers only the noise associated with a system's principal frequency transformation, some controversy exists about the definition and measurement of mixer noise figure. IEEE's NF definition for mixers assumes *no* noise contribution at f_{image}, not even from the mixer's RF-port termination at that frequency; yet, by convention (and out of necessity, since an absolutely noiseless $T = 0\,f_{image}$ termination is unavailable on the average test bench), the noise figures we measure always include *some* noise at f_{image}, even if, as a result of filtering, it arises only from the mixer's RF-port termination. The NF measurements and simulations presented in this book reflect this conventional SSB NF scenario. Stephen Maas [3] provides in-depth material on these issues, including the importance of double-sideband NF for some systems, and the limitations of NF as a figure of merit.

4-2-3 Linearity

The 1-dB Compression Point. Like other networks, a mixer is amplitude-nonlinear above a certain input level; above this point, the output level fails to track input-level changes proportionally. This figure of merit, P_{-1dB}, identifies the single-tone input-signal level at which the output of the mixer has fallen 1 dB below the expected output level. The 1-dB compression point in a conventional double-balanced diode mixer is approximately 6 dB below the LO power. For lower distortion mixers, it is usually 3 dB below the LO power.

The 1-dB Desensitization Point. This specification is another figure of merit similar to the 1-dB compression point. However, the 1-dB desensitization point refers to the level of an interfering (undesired) input signal that causes a 1-dB decrease in nominal conversion gain for the desired signal. For a diode-ring DBM, the 1-dB desensitization point is usually 2–3 dB below the 1-dB compression point.

Dynamic Range. The dynamic range of any RF/wireless system can be defined as the difference between the 1-dB compression point and the minimum discernible signal (MDS). These two points are specified in units of power (dBm), giving dynamic range in dB. When the RF input level approaches the 1-dB compression point, harmonic and intermodulation products begin to interfere with the system performance. High dynamic range is obviously desirable, but cost, power consumption, system complexity, and reliability must also be considered.

Harmonic Intermodulation Products (HIPs). These are spurious products that are harmonically related to the f_{LO} and f_{RF} input signals.

$$\text{HIP} = Mf_{LO} + Nf_{RF} \tag{4-28}$$

Table 4-2 Typical spurious responses of high-level double-balanced mixer (decibels below $f_{LO} \pm f_{RF}$ response)

RF Input Signal Harmonics		f_{LO}	$2f_{LO}$	$3f_{LO}$	$4f_{LO}$	$5f_{LO}$	$6f_{LO}$	$7f_{LO}$	$8f_{LO}$
$8f_{RF}$	100	100	100	100	100	100	100	100	100
$7f_{RF}$	100	97	102	95	100	100	100	90	100
$6f_{RF}$	100	92	97	95	100	100	95	100	100
$5f_{RF}$	90	84	86	72	92	70	95	70	92
$4f_{RF}$	90	84	97	86	97	90	100	90	92
$3f_{RF}$	75	63	66	72	72	58	86	58	80
$2f_{RF}$	70	72	72	70	82	62	75	75	100
f_{RF}	60	0	35	15	37	37	45	40	50
		60	60	70	72	72	62	70	70

Table 4-2 shows relative harmonic intermodulation product levels for a high-level diode DBM.

Intermodulation Distortion (IMD). Nonlinearities in the mixer devices give rise to intermodulation distortion products whenever two or more signals are applied to the mixer's RF port. Testing this behavior with two (usually closely spaced) input signals of equal magnitude can return several figures of merit depending on how the results are interpreted. A mixer's third-order output intercept point ($IP_{3,out}$) is defined as the output power level where the spurious signals generated by $(2f_{RF1} \pm f_{RF2}) \pm f_{LO}$ and $(f_{RF1} \pm 2f_{RF2}) \pm f_{LO}$ are equal in amplitude to the desired output signal as shown in Figure 4-7.

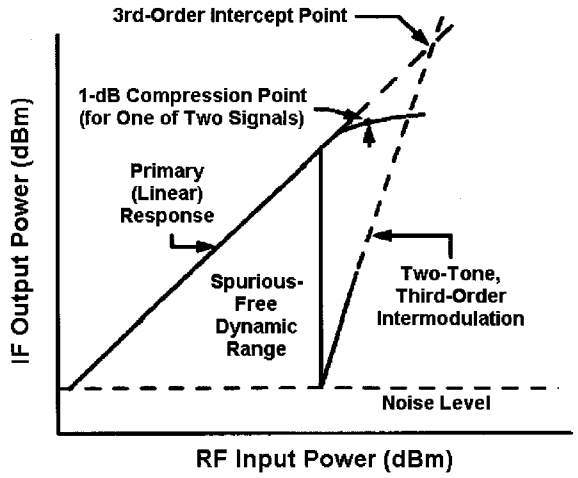


Figure 4-7 Mixer linearity evaluation, including compression and two-tone IMD dynamic range. P_{-1dB} for a single tone cannot be read directly from this graph because the values shown are the result of two-equal-tone drive.

The third-order input intercept point, $IP_{3,in}$—IP_3 referred to the input level—is of particularly useful value and is the most commonly used mixer IMD figure of merit. $IP_{3,in}$ can be calculated according to

$$IP_{n,in} = IMR \div (n - 1) + \text{Input power (dBm)} \tag{4-29}$$

where *IMR* is the intermodulation ratio (the difference in dB between the desired output and the spurious signal), and n is the IM order—in this case, 3. In a conventional diode double-balanced mixer, $IP_{3,in}$ is approximately 14 dB above the single-tone 1-dB compression point (P_{-1dB})—approximately 8 dB greater than the local oscillator power. As will be seen later, this does not apply to feedback-active mixers. They have their own agenda: Above a particular input level, their IMD products increase almost exponentially.

Although designers are usually more concerned with odd-order IM performance, second-order IM can be important in wideband systems (systems that operate over a 2:1 or greater bandwidth) as discussed in Section 1-7-2.

4-2-4 LO Drive Level

A mixer's specifications are usually guaranteed at a particular LO drive level, usually specified as a dBm value that may be qualified with a tolerance. Insufficient LO drive degrades mixer performance; excessive LO drive degrades performance and may damage mixer devices. Commercially available diode mixers are often classified by LO drive level; for example, a "Level 17" mixer requires 17 dBm of LO drive.

4-2-5 Interport Isolation

In a mixer, isolation is defined as the attenuation in dB between a signal input at any port and its level as measured at any other port. High isolation numbers are desirable. Figure 4-8 shows LO-to-IF and LO-to-RF isolation versus frequency for a triple-balanced diode DBM. Isolation is dependent mainly on transformer and physical symmetry and device balance, but the level of signals applied to the mixer also plays a role, as shown in Figure 4-9.

4-2-6 Port VSWR

The load presented by a mixer's ports to the outside world can be of critical importance to a designer. For example, high LO-port VSWR may result in inefficient use of available LO power, resulting in LO starvation (underdrive) that degrades the mixer's performance. Figure 4-10 shows LO-port VSWR versus frequency for a high-level diode DBM with two values of LO power. Like interport isolation, port VSWR can vary with the level of the signal applied.

4-2-7 dc Offset

Isolation between ports plays a major role in reducing dc offset in a mixer. Like isolation, dc offset is a measure of the imbalance of the mixer. In phase-detector and phase-modulator applications, dc offset is a critical parameter.

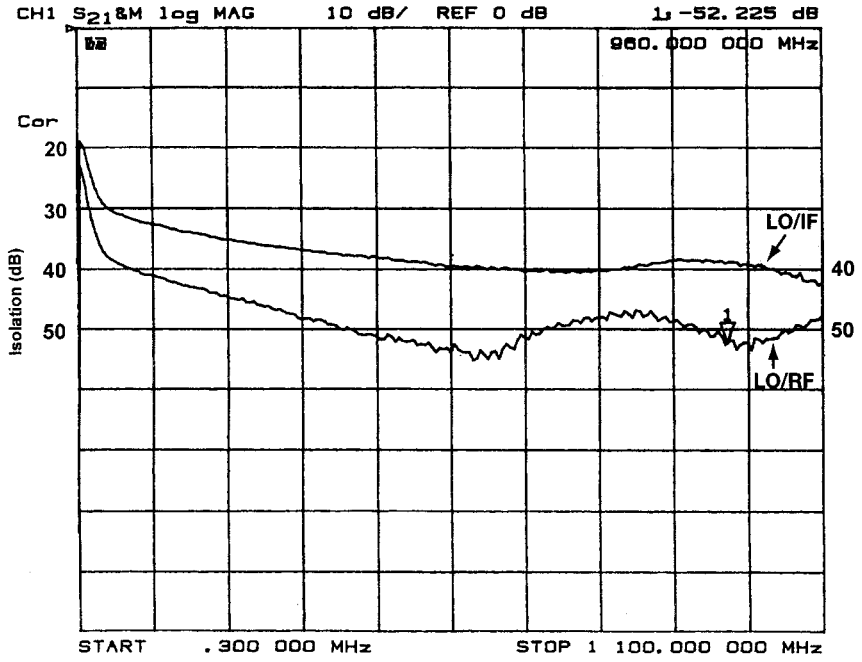


Figure 4-8 LO–IF and LO–RF isolation versus frequency for a high-level triple-balanced diode mixer. The periodic roughness of the traces is a measurement-system artifact, proving that even sophisticated FFT-based instruments are not perfect.

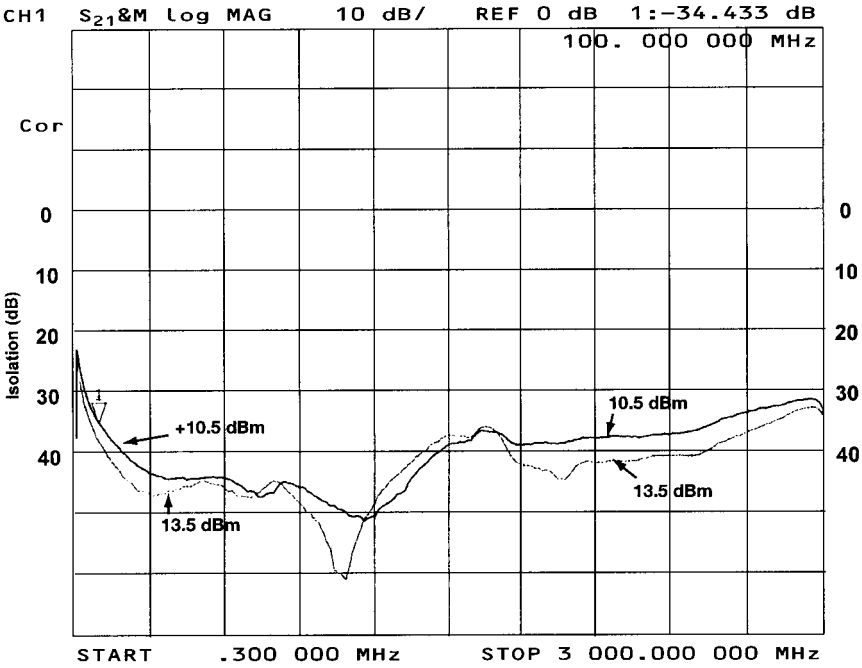


Figure 4-9 LO–IF isolation versus frequency and LO drive level for a high-level diode DBM.

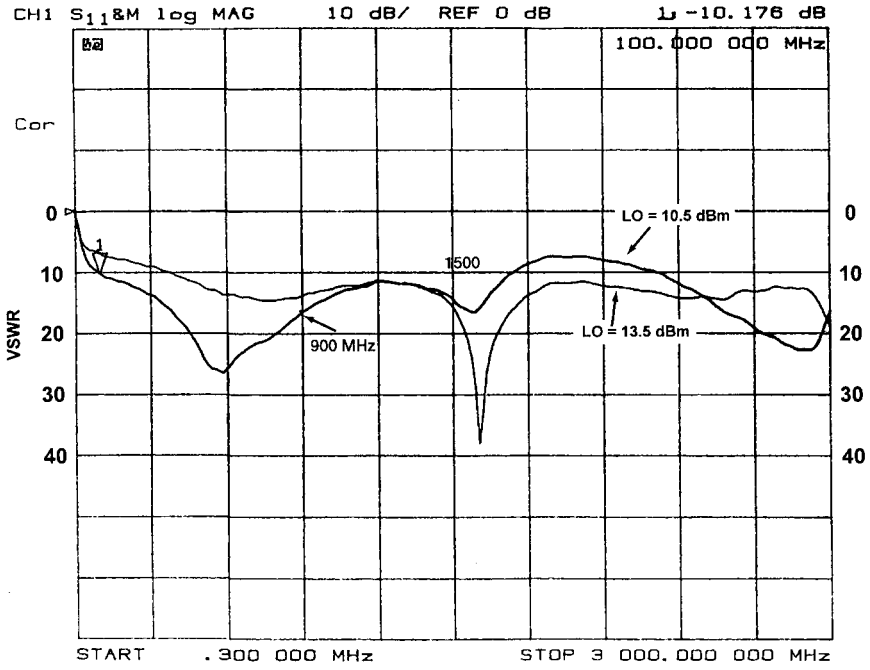


Figure 4-10 LO-port VSWR versus frequency for a high-level diode DBM.

4-2-8 dc Polarity

Unless otherwise specified, mixers with dc output are designed to have negative polarity when RF and LO signals are of equal phase.

4-2-9 Power Consumption

Circuit power consumption is always important, but in battery-powered wireless designs it is *critical.* Mixer choice may be significant in determining a system's power consumption, sometimes in ways that seem paradoxical at first glance. For instance, a passive mixer might seem to be a power-smart choice because it consumes *no* power—until we factor in the power consumption of the circuitry needed to provide the (often considerable) LO power a passive mixer requires. If a mixer requires a broadband resistive termination that will be provided by a postmixer amplifier operating at a high standing current, the power consumption of the amplifier stage must be considered as well. Evaluating the suitability of a given mixer type to a task therefore requires a grasp of its ecology as well as its specifications.

4-3 DIODE MIXERS

Passive mixers based on diode switches are common in base-station applications, where their high dynamic range and 50-Ω port impedances overcome objections to their inherent conversion loss, unsuitability to integration, and relatively high LO-power requirement. As we'll discuss later, FET-based passive mixers are overcoming some of these limitations in portable/mobile wireless applications.

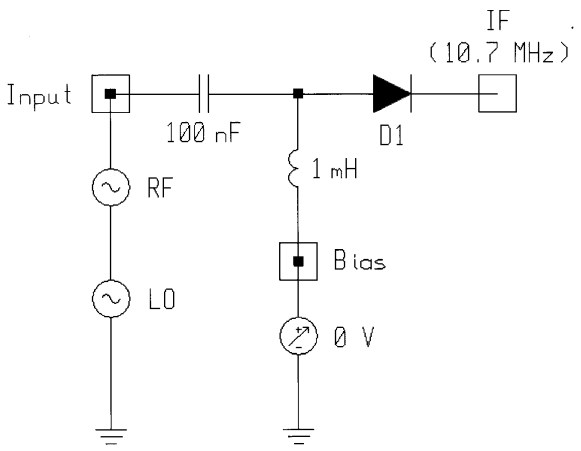


Figure 4-11 Schematic of the single-diode mixer. The circuit's 50-Ω output termination serves as the diode's dc return, as well as completing the circuit for RF, LO, and IF.

4-3-1 Single-Diode Mixer

Figure 4-11 shows the schematic of a simplistic single-diode mixer. The LO and RF signals are applied in series to the diode, with no effort made to match the sources to the diode or isolate the sources from each other. The LO and RF signals are both present in the diode simultaneously, so the mixing performed is additive. The LO signal switches the diode on and off, gating the RF signal to the circuit's output at the LO frequency. In this simple circuit, the output port's 50-Ω resistance serves as the diode's ac load and dc return. Because we want the diode to operate as a linear switch with respect to the RF signal, the diode should turn on hard and turn off as completely as possible, and the LO, not the RF source, should switch the diode on and off. The upper frequency limit of the circuit will depend on how rapidly the diode can switch between the on and off states. The LO can be only so strong before the diode's dissipation limits are exceeded; the RF source can be only so strong relative to the LO before it begins to play a role in switching the diode on and off. The LO must therefore be considerably stronger than the RF source—20 dB or more for low-distortion applications.

Figure 4-12 shows how the single-diode mixer's conversion gain and noise figure vary with applied LO power. Figure 4-13 shows the mixer's output spectrum.

Tables 4-3, 4-4, and 4-5 present sample data on diodes suitable for mixer service. The manufacturer's full datasheet on the Siemens BAT14-099 follows.

The simple single-diode-mixer circuit shown in Figure 4-11 is intended only as an illustration of the basic behavior of diode mixer behavior. A practical single-diode mixer would include filtering at its RF, LO, and IF ports—RF filtering for image rejection, reduced LO radiation and optimum matching of the RF source to the diode; LO filtering to keep RF out of the LO and optimally match the LO to the diode; and IF filtering to optimally match the diode to its IF (and IF image) load, preferably while providing some rejection of the mixer's unwanted outputs, the strongest (and most potentially troublesome) of which is the LO signal. Detail on the design of single-diode mixers can be found in Zinke and Brunswig [4].

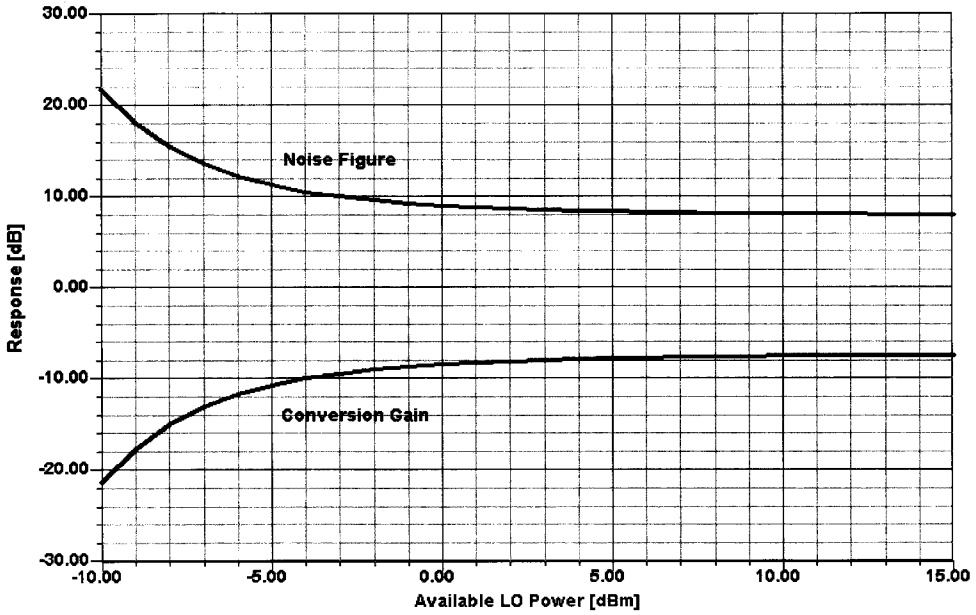


Figure 4-12 Conversion gain and noise figure versus LO power for the single-diode mixer. The values reported are worse than those predicted by theory because RF, LO, and IF matching has not been attempted. In this analysis, LO = 310.7 MHz (–10 to 25 dBm), RF = 300 MHz (–50 dBm), and IF = 10.7 MHz.

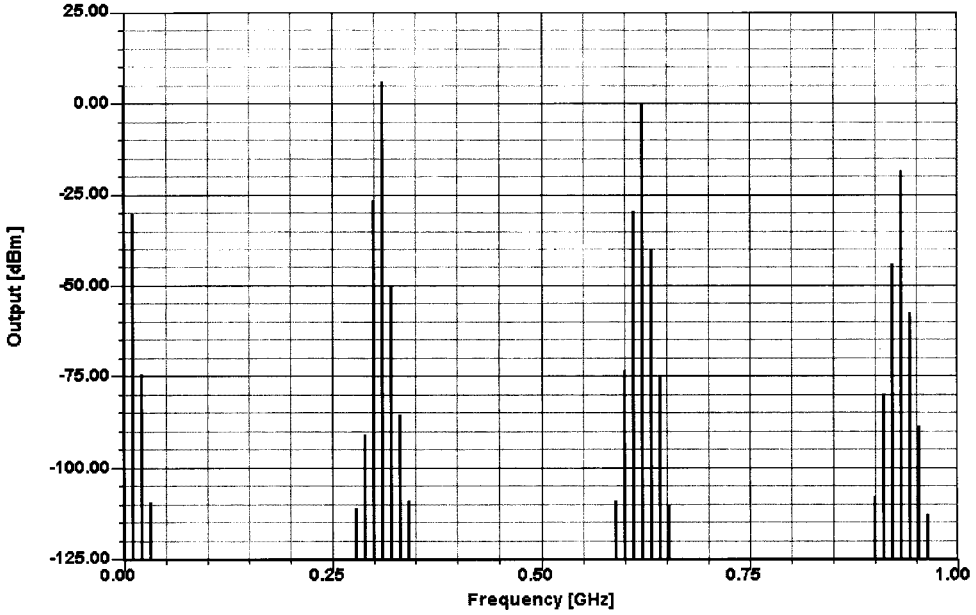


Figure 4-13 Output spectrum of the single-diode mixer. Three LO harmonics and 3 LO sidebands were used in this analysis. In this analysis, LO = 310.7 MHz (13 dBm), RF = 300 MHz (–20 dBm), and IF = 10.7 MHz.

Table 4-3 Example RF Schottky diodes (Siemens)

Type	Maximum Ratings		Characteristics (T_A = 25 °C)					Package
	V_R	I_F	C_T	V_F at I_F		V_F at I_F		
	(V)	(mA)	(pF)	(mV)	(mA)	(mV)	(mA)	
BAT14-03W	4	90	0.22	430	1	550	10	SOD-323
BAT14-099 (dual)	4	90	0.22	430	1	550	10	SOT-143
BAT14-099R (quad)	—	90	0.38	400	1	480	10	SOT-143
BAT15-03W	4	110	0.21	230	1	320	10	SOD-323
BAT15-099 (dual)	4	110	0.21	230	1	320	10	SOT-143
BAT15-099R (quad)	—	110	0.37	230	1	320	10	SOT-143
BAT17	4	130	0.55	340	1	425	10	SOT-23
BAT17-04 (dual)	4	130	0.55	340	1	425	10	SOT-23
BAT68	8	130	0.75	320	1	395	10	SOT-23
BAT68-04 (dual)	8	130	0.75	320	1	395	10	SOT-23
BAT68-05 (dual)	8	130	0.75	320	1	395	10	SOT-23
BAT68-06 (dual)	8	130	0.75	320	1	395	10	SOT-23
BAT68-07 (dual)	8	130	0.75	320	1	395	10	SOT-143
BAT68-03W	8	130	0.75	320	1	395	10	SOD-323
BAT114-099 (dual)	4	90	0.22	580	1	680	10	SOT-143
BAT114-099R (quad)	—	90	0.22	580	1	680	10	SOT-143

4-3-2 Single-Balanced Mixer

Figure 4-14 shows the schematic of a two-diode, *single-balanced* mixer (SBM). Unlike the single-diode mixer, it performs multiplicative mixing because its RF and LO signals are applied to different ports. In this more commonly seen two-diode mixer configuration, a balanced transformer drives the diodes out of phase for the LO and in phase for signals present at the RF port. Figure 4-15 shows how this mixer's conversion gain and noise figure vary with applied LO power. Figure 4-16 shows how the mixer's conversion gain and noise figure vary with frequency for a constant LO power.

Table 4-4 Example Schottky detector diodes (Siemens)

Type	Maximum Ratings		Characteristics (T_A = 25 °C)								Package
	V_R	I_F	C_T at V_R		V_F at I_V		R_O at V_F		I_R at V_R (μA)		
	(V)	(mA)	(pF)	(V)	(V)	(mA)	(kΩ)	(V)	(μA)	(V)	
BAT62	40	20	0.4	0	0.53	2	160	0	≤10	40	SOT-143
BAT62-03W	40	20	0.4	0	0.53	2	160	0	≤10	40	SOD-323
BAT63	3	100	0.65	0.2	0.19	1	30	0	≤10	3	SOT-143

Table 4-5 Example Schottky diodes and quads

Schottky Diodes for Mixer & Detector Applications

Barrier	V_B @ 10 μA (V)	C_T @ 0 V (pF)	V_F @ 1 mA (mV)	R_T* @ 10 mA (Ω)	SOD-323	SOT-23			SOT-143
	Min.			Max.	Single	Single	Series Pair	Reverse Series Pair	Unconnected Pair
Low	2	0.43–0.63	200–270	8	SMS1546-011	SMS1546-001	SMS1546-005		SMS1546-015
Low	2	0.2–0.26	260–320	18	SMS7621-011	SMS7621-001	SMS7621-005	SMS7621-006	SMS7621-015

*R_T is the slope resistance.

Schottky Quads

Barrier	V_B @ 10 μA (V)	C_J[1] @ 0 V (pF)	V_F @ 1 mA (mV)	ΔV_F @ 1 mA (mV)	R_T[2] @ 10 mA (Ω)	SOT-143			
	Min.			Max.	Max.	Ring Quad	Crossover Quad	Bridge Quad	Octo Quad
Low	2	0.3–0.5	200–270	10	8	SMS3926-022	SMS3926-023	SMS3929-021	SMS3938-026
Medium	3	0.3–0.5	310–370	10	8	SMS3927-022	SMS3927-023	SMS3930-021	SMS3939-026
High	4	0.3–0.5	520–580	10	8	SMS3928-022	SMS3928-023	SMS3931-021	SMS3940-026

1. C_J is the capacitance per junction.
2. R_T is the slope resistance.

The two-diode mixer is used mostly in the frequency range above 1 GHz in a manner akin to a phase discriminator, using step-recovery diodes in the LO feed for enhanced harmonic mixing. Such mixers are mainly used in medium-cost spectrum analyzers or microwave receivers up to several tens of gigahertz, with the necessary transformers and baluns printed on the circuit board. With perfectly matched diodes and perfect transformer and constructional symmetry, no LO energy arrives at the IF and RF ports, and there is only slight attenuation between the RF and IF ports. Both building and computer modeling such a mixer are impossible: building, because perfectly matched diodes and perfect transformer and constructional symmetry cannot be achieved in practice; computer modeling, because floating-point mathematics runs out of gas in handling the infinite amplitude spread involved in calculating the perfect cancellation of the LO signal as it travels to the RF and IF ports. That said, Figure 4-17 compares the mixer's *quasi*-ideal port-to-port isolation (matched diodes, a perfect transformer, no stray inductances and capacitances, and a 10-MΩ resistors

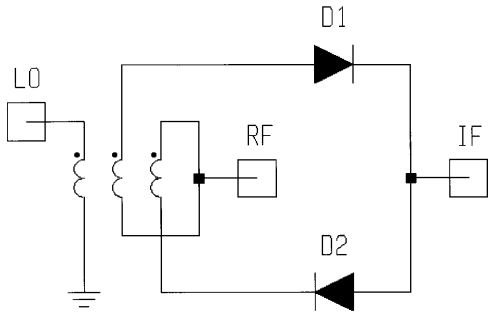


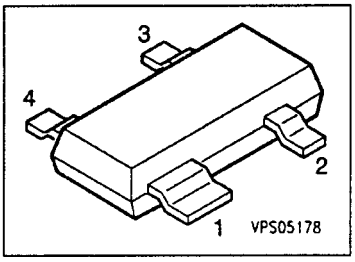
Figure 4-14 Schematic of the two-diode (also known as *single-balanced*) mixer.

SIEMENS

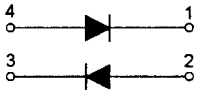
Silicon Dual Schottky Diode

BAT 14-099

- DBS mixer application to 12 GHz
- Low noise figure
- Medium barrier type



ESD: Electrostatic discharge sensitive device, observe handling precautions!

Type	Marking	Ordering Code (tape and reel)	Pin Configuration	Package[1]
BAT 14-099	S9	Q62702-A3461		SOT-143

Maximum Ratings per Diode

Parameter	Symbol	Values	Unit
Reverse voltage	V_R	4	V
Forward current	I_F	90	mA
Power dissipation, $T_S \leq 55$ °C	P_{tot}	100	mW
Storage temperature range	T_{stg}	– 55 ... + 150	°C
Operating temperature range	T_{op}	– 55 ... + 150	

Thermal Resistance

Junction – ambient[2]	$R_{th\,JA}$	≤ 1090	K/W
Junction – soldering point	$R_{th\,JS}$	≤ 930	

1) For detailed information see chapter Package Outlines.
2) Package mounted on alumina 15 mm × 16.7 mm to 0.7 mm.

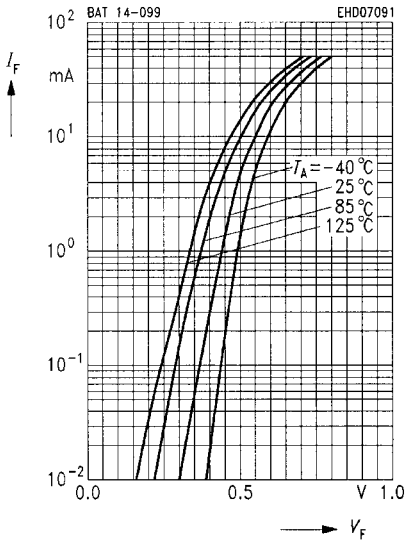
SIEMENS **BAT 14-099**

Electrical Characteristics per Diode
at T_A = 25 °C, unless otherwise specified.

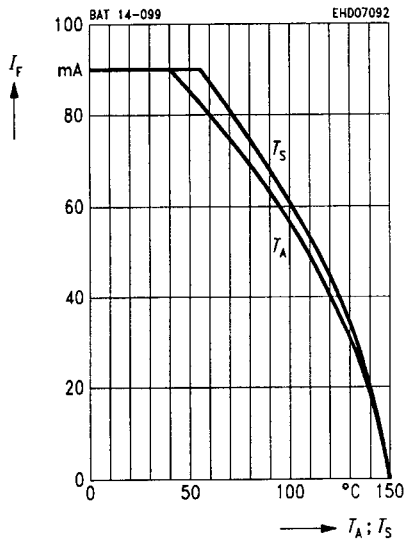
Parameter	Symbol	Values			Unit
		min.	typ.	max.	
Breakdown voltage I_R = 5 µA	V_{BR}	4	–	–	V
Forward voltage I_F = 1 mA I_F = 10 mA	V_F	 – –	 0.43 0.55	 – –	
Forward voltage matching I_F = 10 mA	ΔV_F	–	–	10	mV
Diode capacitance V_R = 0, f = 1 MHz	C_T	–	–	0.35	pF
Forward resistance I_F = 10 mA / 50 mA	R_F	–	5.5	–	Ω

SIEMENS **BAT 14-099**

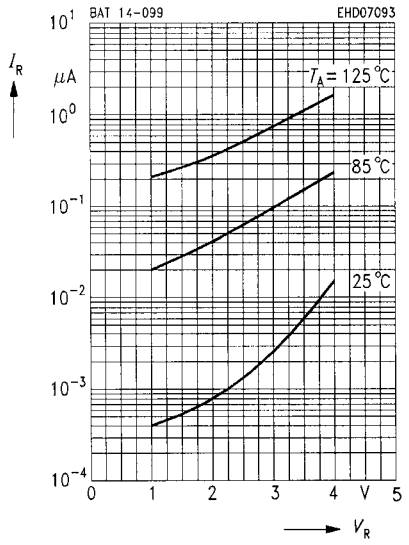
Forward current $I_F = f(V_F)$



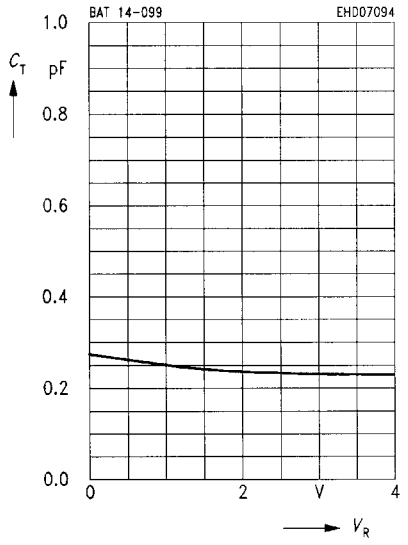
Forward current $I_F = f(T_S; T_A^*)$
*Package mounted on alumina



Reverse current $I_R = f(V_R)$



Diode capacitance $C_T = f(V_R)$
$f = 1$ MHz



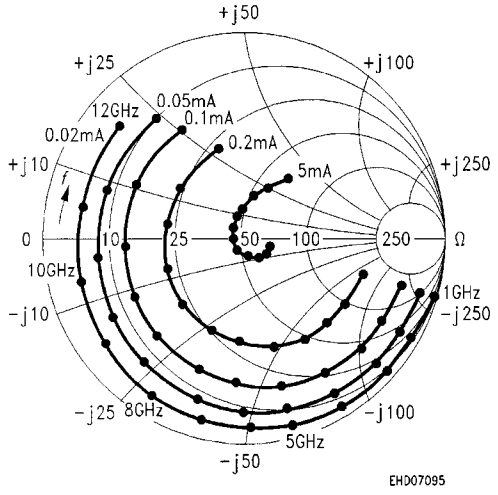
SIEMENS **BAT 14-099**

S_{11}-Parameters

Typical impedance characteristics (with external bias I and $Z_0 = \Omega$)

f	I = 0.02 mA		I = 0.05 mA		I = 0.1 mA		I = 0.2 mA		I = 0.5 mA	
GHz	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG	MAG	ANG
1	0.99	– 15.89	0.91	– 16.40	0.79	– 16.40	0.57	– 16.60	0.13	– 17.30
2	0.96	– 30.40	0.88	– 30.80	0.76	– 31.09	0.56	– 30.70	0.13	– 28.40
3	0.95	– 45.30	0.87	– 46.20	0.75	– 47.30	0.55	– 47.00	0.11	– 43.99
4	0.93	– 59.60	0.86	– 61.60	0.73	– 62.40	0.53	– 62.40	0.10	– 54.40
5	0.93	– 74.80	0.85	– 77.10	0.72	– 78.70	0.51	– 78.70	0.07	– 80.70
6	0.91	– 89.50	0.83	– 93.10	0.69	– 95.70	0.48	– 95.70	0.04	– 102.30
7	0.89	– 106.60	0.80	– 110.50	0.66	– 112.70	0.45	– 114.00	0.02	158.01
8	0.88	– 123.40	0.79	– 129.40	0.64	– 132.40	0.43	– 135.40	0.06	118.40
9	0.86	– 143.20	0.76	– 150.20	0.62	– 154.20	0.40	– 161.20	0.12	96.20
10	0.83	– 166.10	0.72	– 174.10	0.58	– 179.10	0.37	171.10	0.19	72.10
11	0.82	168.10	0.71	158.10	0.59	153.80	0.39	140.80	0.25	62.60
12	0.80	138.20	0.72	127.20	0.60	121.20	0.44	108.20	0.33	49.20

$S_{11} = f(f, I)$



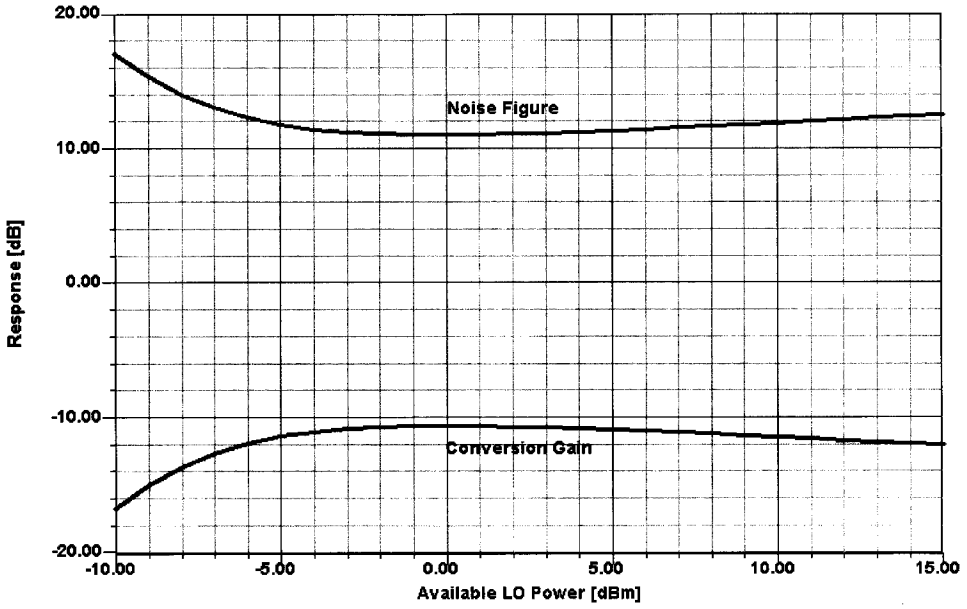


Figure 4-15 How the nonideal mixer's conversion gain and noise figure vary with available LO power. In this analysis, LO = 500 MHz (13 dBm), RF = 500.455 MHz (–20 dBm), and IF = 455 kHz.

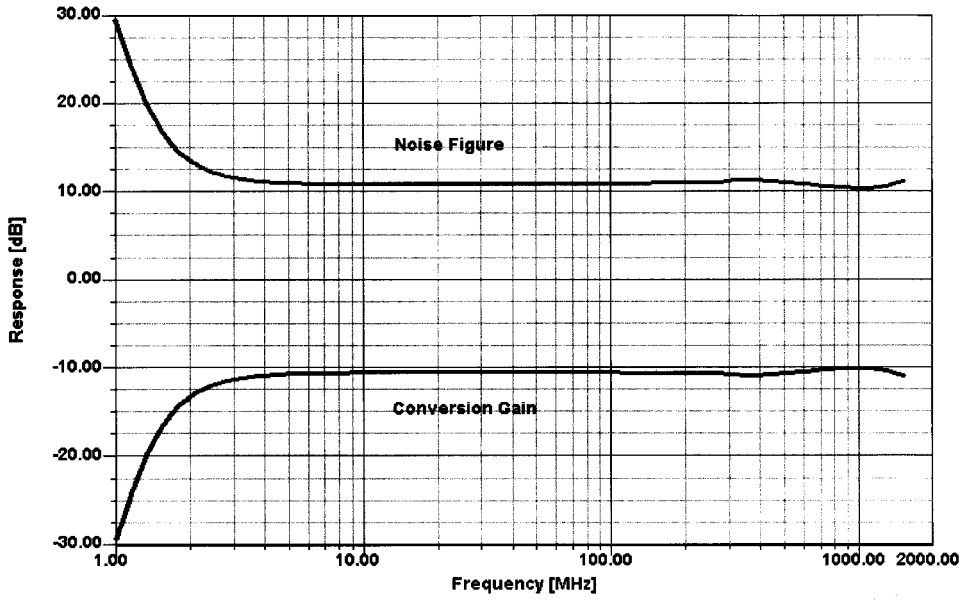


Figure 4-16 How the nonideal two-diode mixer's conversion gain and noise figure vary with frequency for a constant LO power. In this analysis, LO = 1–1500 MHz (2 dBm), RF = 1.455–1500.455 MHz (–40 dBm), and IF = 455 kHz.

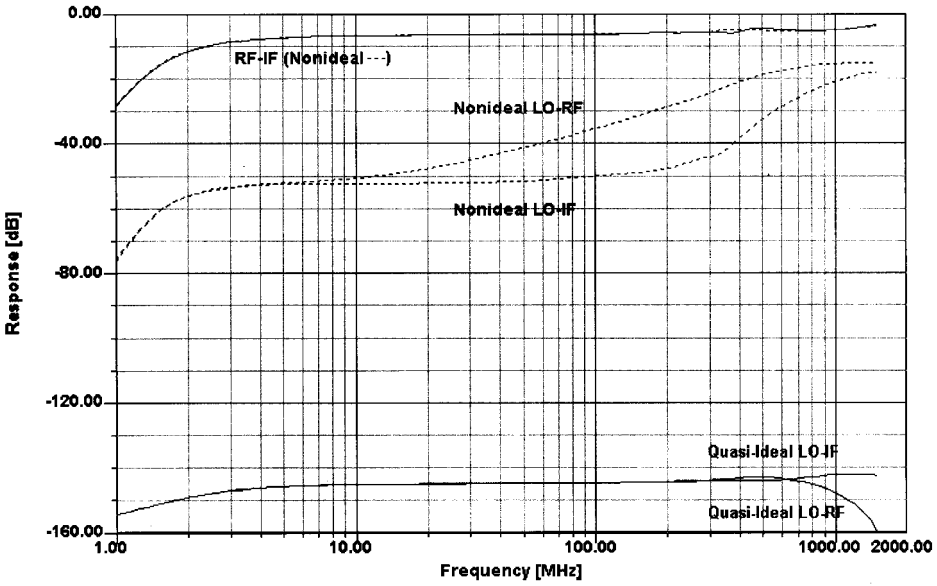


Figure 4-17 Port-to-port isolation of the quasi-ideal (identical diodes and no stray capacitance) and nonideal (slightly mismatched diodes and 0.5 pF of stray capacitance between the upper terminal of the middle transformer winding and ground) two-diode mixer. In this analysis, LO = 1–1500 MHz (2 dBm), RF = 1.455–1500.455 MHz (–40 dBm), and IF = 455 kHz.

connected from port to port) and nonideal port-to-port isolation (slightly mismatched diodes and 0.5 pF between the upper terminal of the middle winding and ground).

Figures 4-18a and 4-18b show the mixer's output spectrum for the quasi-ideal and nonideal cases, respectively.

Subharmonically Pumped Single-Balanced Mixer. Figure 4-19 shows a single-balanced mixer with a difference: antiparallel diode pairs take the place of single diodes, the RF and IF are buffered from each other only by filtering, and the LO is applied at one-half the frequency necessary to provide the desired frequency conversion. The RF-to-IF isolation is limited to that provided by the series input and output filtering, but the LO-to-IF isolation is higher at f_{LO} and much higher at $2f_{LO}$ than that achievable with a DBM with the LO signal at $2f_{LO}$(Figure 4-20). Although the example shown is for an up converting HF receiver, this technique finds application well into the microwave range as the basis for I/Q modulators, in which carrier leakage must be reduced to a level difficult to achieve with conventional DBMs [5–10].

4-3-3 Diode-Ring Mixer

Adding two more diodes and another transformer to the single-balanced mixer results in a double-balanced mixer (DBM) as shown in Figure 4-21. A DBM's frequency response is largely determined by the frequency response of its transformers, which act as transmission lines. The low-frequency limit is determined by the inductance of the transformer windings, the reactance of which, at the lowest frequency of interest, should be at least four times the impedance at which the transformer operates. The upper frequency limit is determined

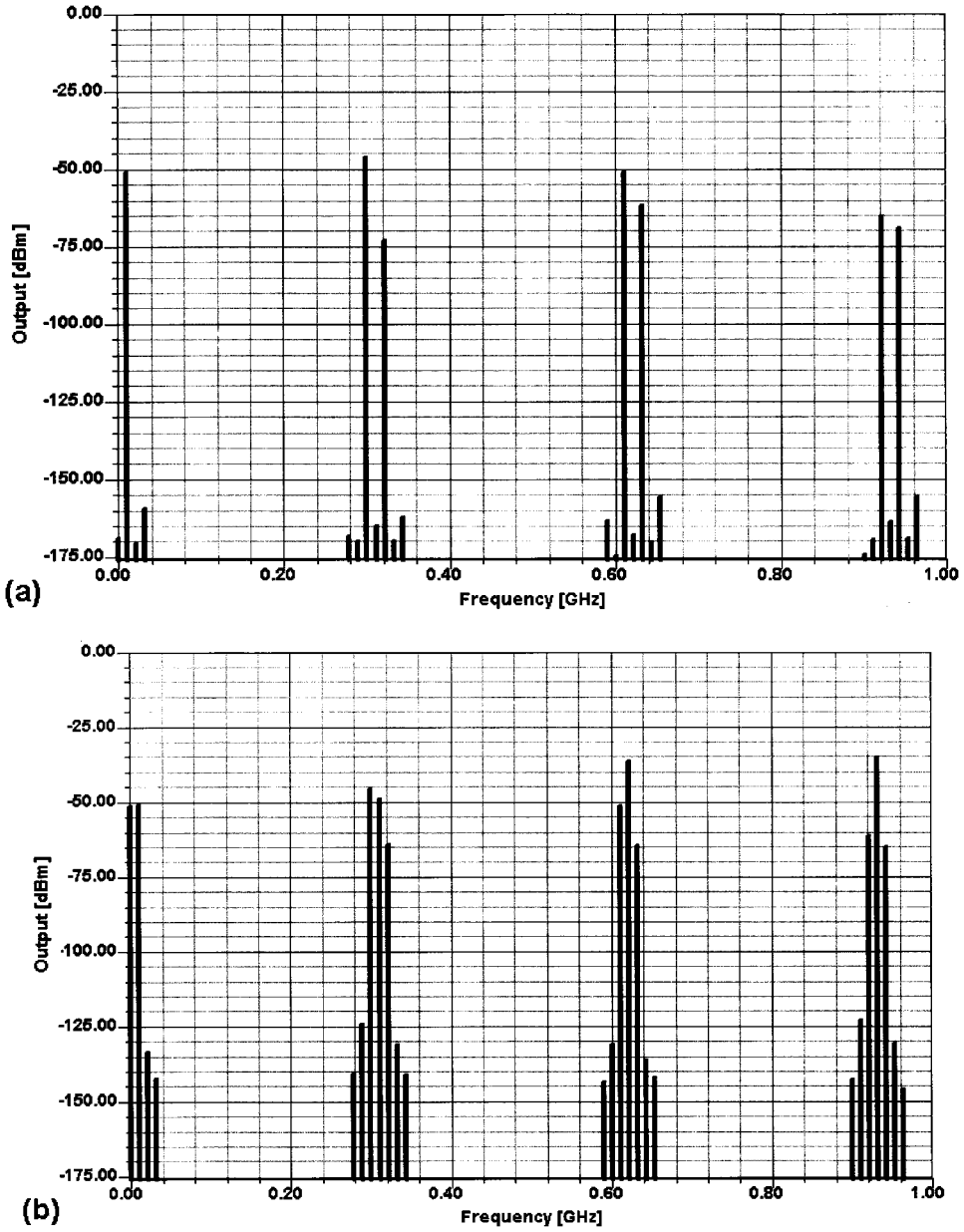


Figure 4-18 The nonideal mixer's output spectrum with (a) identical diodes and no stray capacitance and (b) slightly mismatched diodes and 0.5 pF of stray capacitance between the upper terminal of the middle transformer winding and ground. In these analyses, LO = 310.7 MHz (2 dBm), RF = 300.0 MHz (–40 dBm), and IF = 10.7 MHz. Four LO harmonics and 3 LO sidebands were used.

mainly by the degradation of the transformers' transmission-line behavior at higher frequencies, although the increasing importance of diode capacitance also plays a role.

A DBM's interport isolation is determined by the symmetry of its transformers, diodes, and physical construction. In practice, the effects of diode mismatch can be minimized by

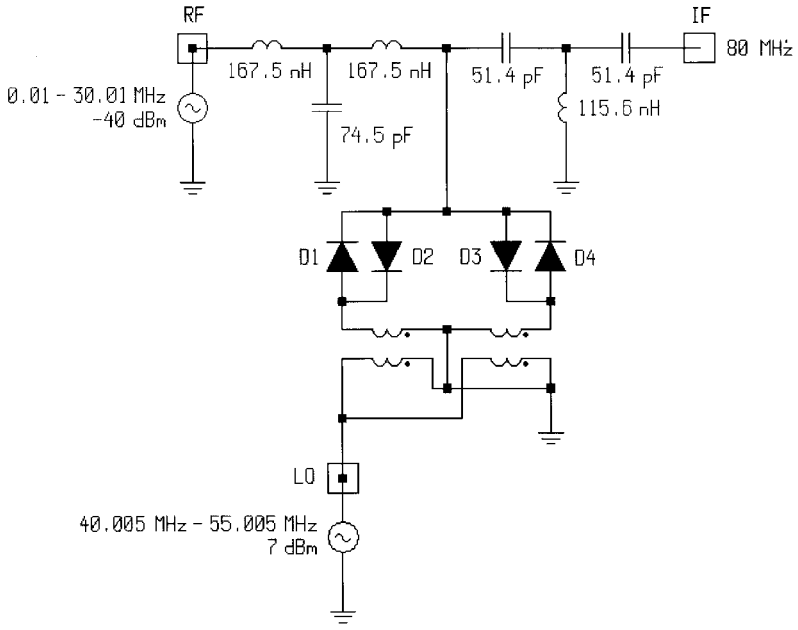


Figure 4-19 A subharmonically pumped single-balanced mixer using antiparallel diode pairs. The LO operates from 40.005 to 55.005 MHz to mix 0.01–30.01-MHz RF to an IF of 80 MHz.

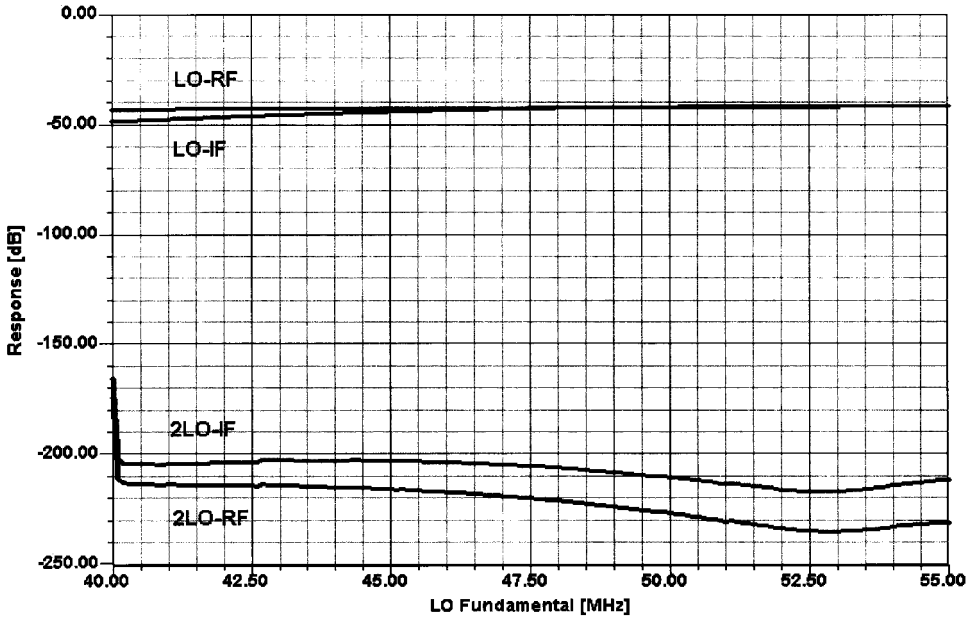


Figure 4-20 Simulated interport isolation of the subharmonic SBM. For realism, the diodes and transformers are slightly mismatched.

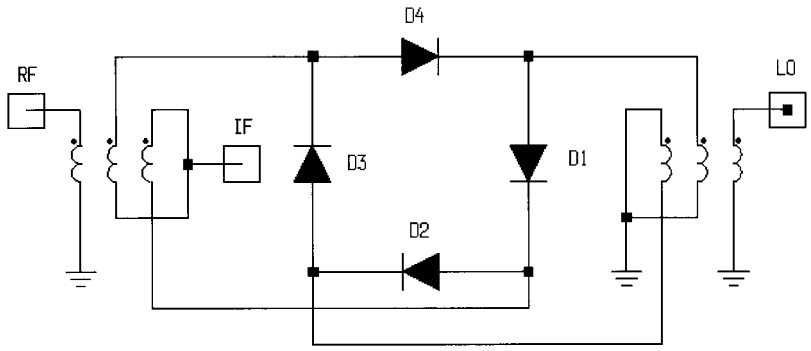


Figure 4-21 Schematic of the diode-ring double-balanced mixer.

using a dual diode, such as—in the case of the BAT15—the BAT15-099, the maximum V_F spread between the diodes being specified as 20 mV.

Figure 4-22 shows how the DBM's conversion gain and noise figure vary with applied LO power. Figure 4-23 shows how the DBM's conversion gain and noise figure vary with frequency for an LO power of 7 dBm, with quasi-ideal and nonideal balance. Figure 4-24 shows how the DBM's port-to-port isolation differs with quasi-ideal and nonideal balance for an LO power of 7 dBm. Figure 4-25 shows how the DBM's RF- and LO-port return losses vary with frequency; the sharp peak corresponds to a resonance caused by one of the stray capacitances added to simulate less-than-ideal balance in the modeled mixer.

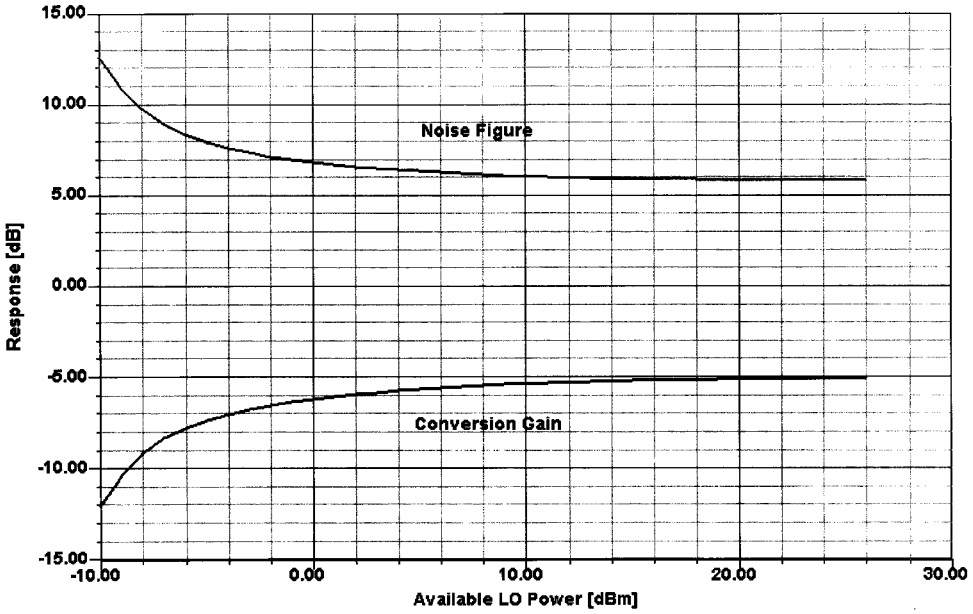


Figure 4-22 DBM conversion gain and noise figure versus LO power. In this analysis, LO = 310.7 MHz (–10 to 26 dBm), RF = 300 MHz (–40 dBm) and IF = 10.7 MHz.

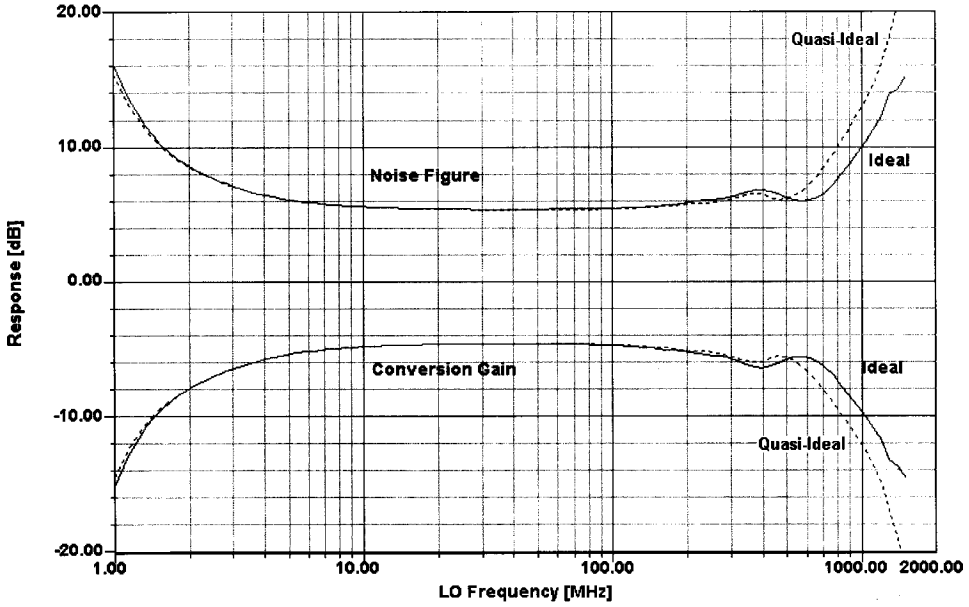


Figure 4-23 This plot of conversion gain (CG) and noise figure (NF) versus frequency for quasi-ideally and nonideally balanced versions of the same DBM reveals that balance plays a relatively minor role in the CG and NF performance achieved. In these analyses, the LO (–7 dBm) sweeps from 1 to 1500 MHz and the RF (–40 dBm) sweeps from 1.455 to 1500.455 MHz to produce an IF of 455 kHz.

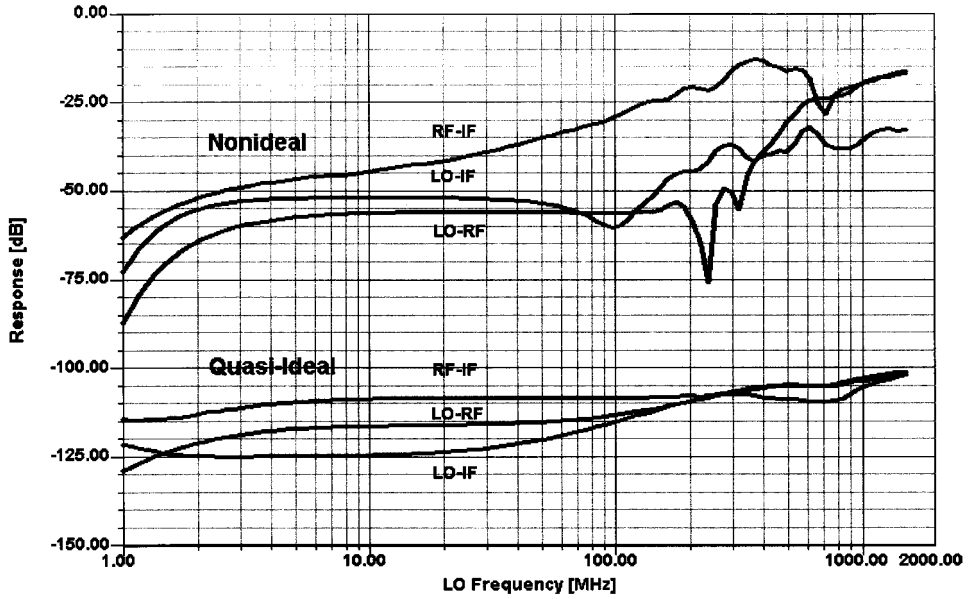


Figure 4-24 Interport isolation for DBMs with quasi-ideal and nonideal balance. In these analyses, the LO (–7 dBm) sweeps from 1 to 1500 MHz and the RF (–40 dBm) sweeps from 1.455 to 1500.455 MHz to produce an IF of 455 kHz.

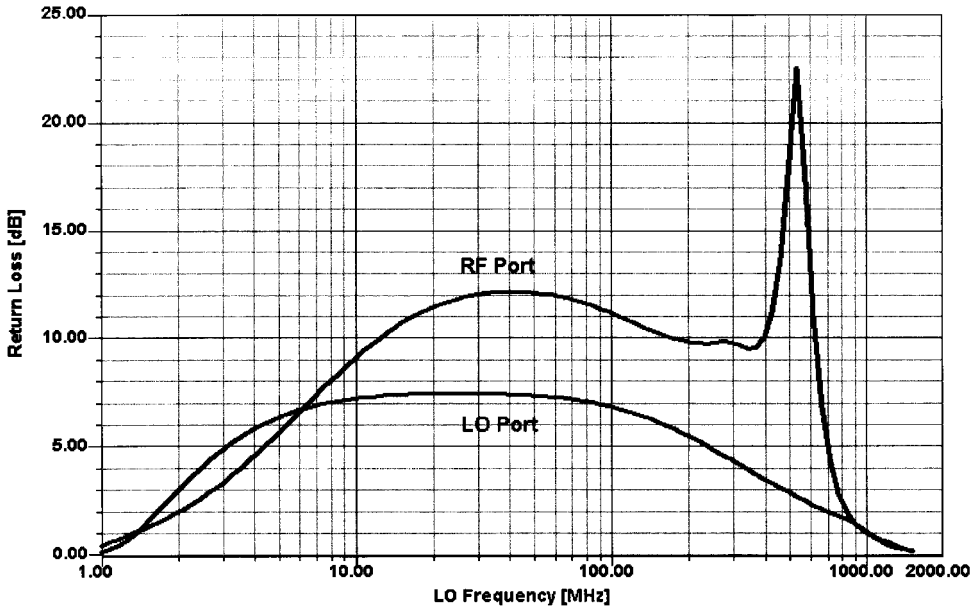


Figure 4-25 Return loss versus frequency for the DBM's RF and LO ports. The sharp peak results from a stray resonance.

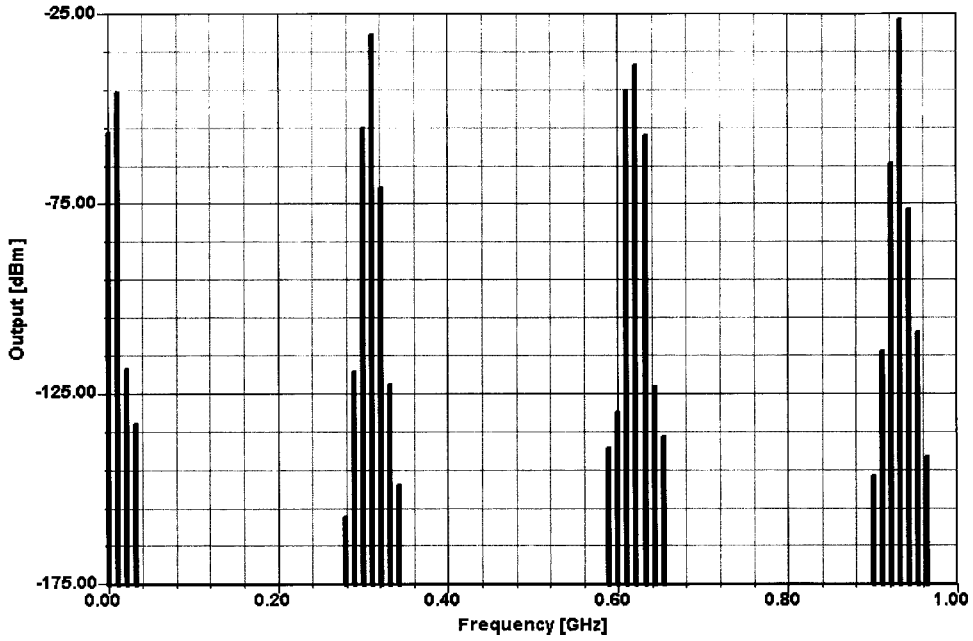


Figure 4-26 Output spectrum of a nonideally balanced DBM. In this analysis, LO = 310.7 MHz (–7 dBm) and RF = 300 MHz (–40 dBm) for an IF of 10.7 MHz. Four LO harmonics and three LO sidebands were used.

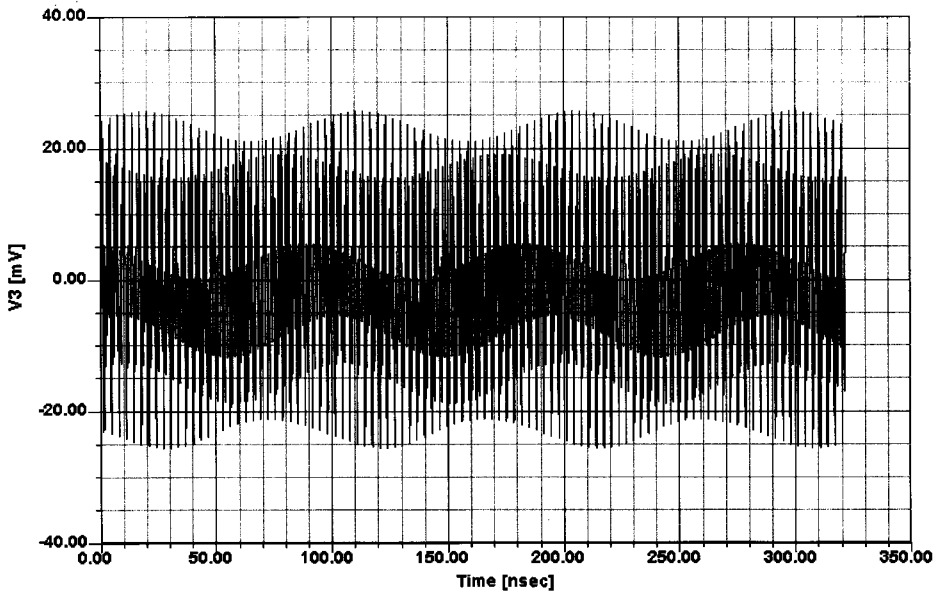


Figure 4-27 IF-port voltage waveform of the DBM over 100 cycles of the LO signal. The 310.7-MHz LO and 10.7-MHz IF components are clearly evident.

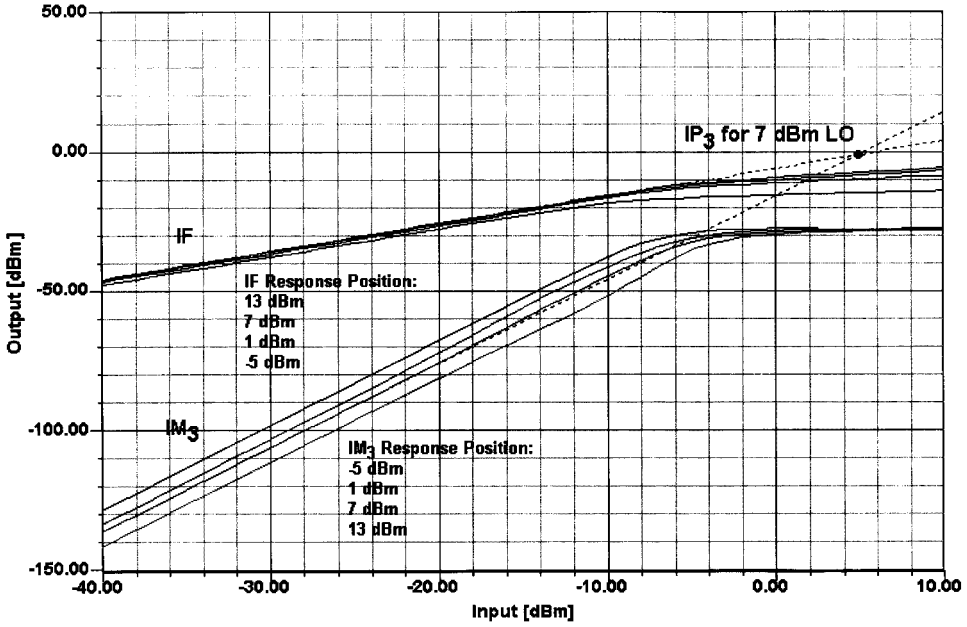


Figure 4-28 Diode DBM IF and IM_3 outputs versus RF power for four LO-drive levels. The responses for LO = 7 dBm have been extrapolated to show IP_3. Within limits, varying a mixer's LO drive affects its linear IF output relatively little while significantly affecting IMD. See Figure 4-29. For all four analyses, LO = 310.7 MHz, RF1 = 300.0 MHz (–40 to 10 dBm), and RF2 = 300.3 MHz (–40 to 10 dBm); four LO harmonics and three LO sidebands were used.

Figure 4-26 shows the DBM's output spectrum. Figure 4-27 shows the DBM's output waveform over 50 cycles of the LO signal.

Two-tone testing of the DBM allows us to characterize its IP_3 figures of merit. Figure 4-28 shows the nonideal DBM's IF and IM_3 responses for LO powers of –5, 1, 7, and 13 dBm. Figure 4-29 details how the DBM's IP_3 increases with LO drive, and Figure 4-30 shows the desired IF outputs and close third-order spurs near 10.7 MHz. Figure 4-31 shows the DBM's output voltage over 100 cycles of the LO signal, and Figure 4-32 shows the anode–cathode voltage of one of the ring's diodes, also over 100 LO cycles, both under two-tone IMD test conditions.

A double-balanced mixer, unless it is termination insensitive, is extremely sensitive to nonresistive termination. This is because the transmission-line transformers do not operate properly when they are not properly terminated, and the reflected power generates high voltage across the diodes. This effect results in much higher distortion levels than in a properly terminated transformer.

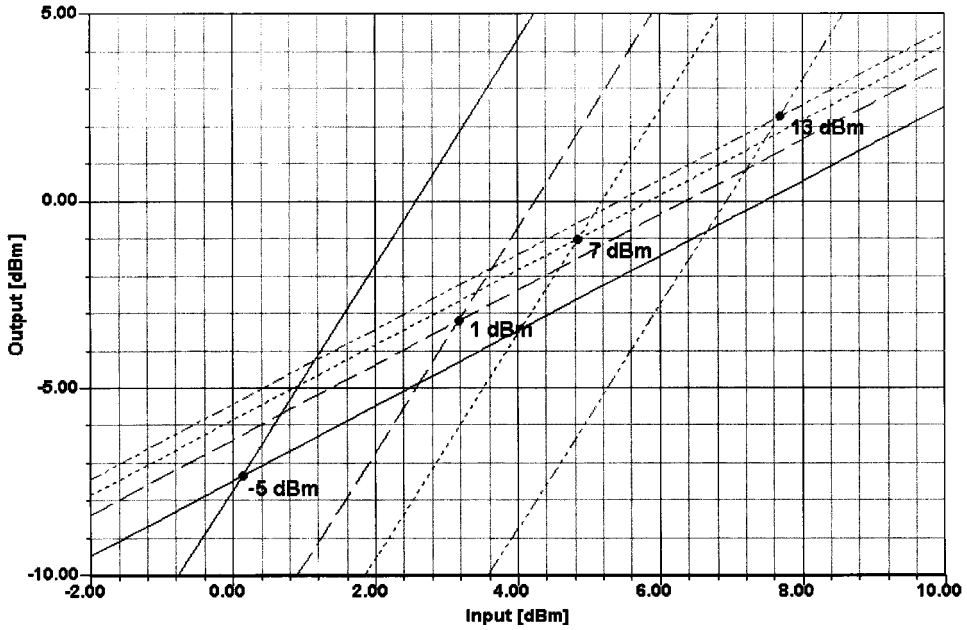


Figure 4-29 Extrapolating the responses for all four LO levels represented in Figure 4-28 shows how varying a diode DBM's LO drive shifts its third-order intercept point. Although these curves indicate that the simulated mixer's IP_3 generally increases with LO drive, the improvement in IP_3 is not as great as we might expect. The reason for this is that these four analyses, as well as the other diode-mixer analyses in this chapter, were done using diode models with a threshold voltage (V_J) of 0.23. If high-level diodes with a V_J of about 0.8 V had been used, $IP_{3,out}$ for the 13-dBm LO case shown here would increase to +13 dBm. $IP_{3,in}$ for the 13-dBm LO case would turn out to be 13 dBm + insertion loss = 13 dBm + 7 dB = 20 dBm. The issue of diode damage aside, attempting to increase IP_3 merely by driving a low- or medium-barrier diode harder eventually results in diminishing returns. High-barrier diodes are essential in getting the best IP_3 performance with high LO drive.

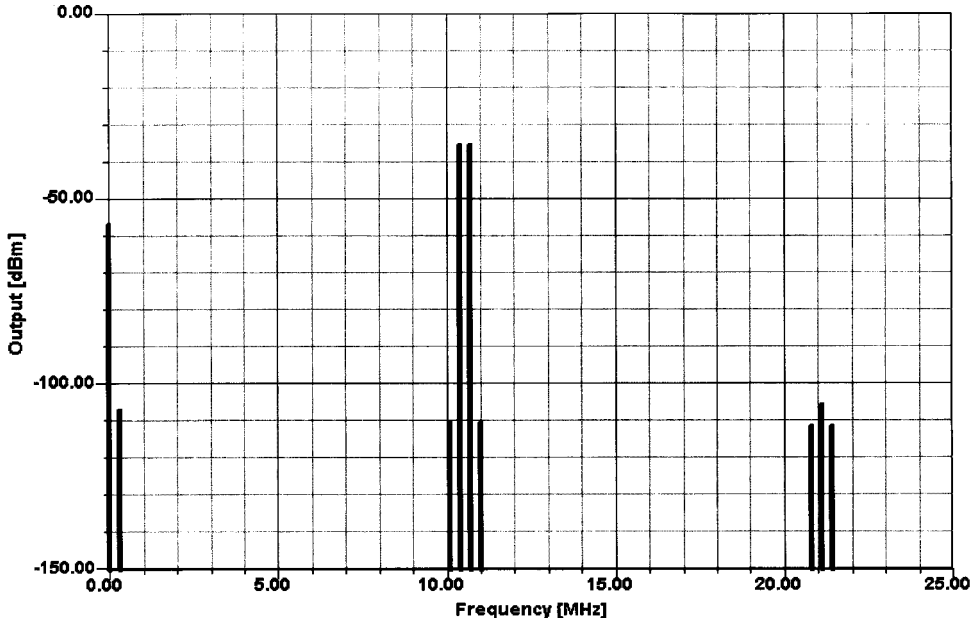


Figure 4-30 The DBM's output in the 11-MHz region during two-tone testing. The third-order products are clearly visible above and below the desired output signals. The test conditions for this analysis are those for Figure 4-28 with LO = 13 dBm.

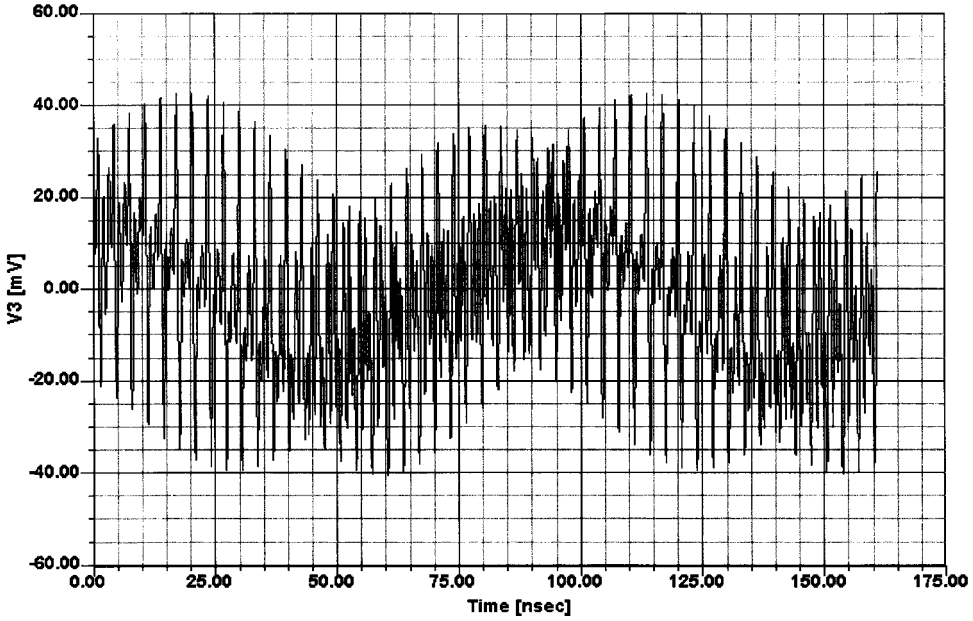


Figure 4-31 The DBM's IF-output voltage over 50 cycles of the LO signal. The test conditions for this analysis are those for Figure 4-28 with LO = 13 dBm.

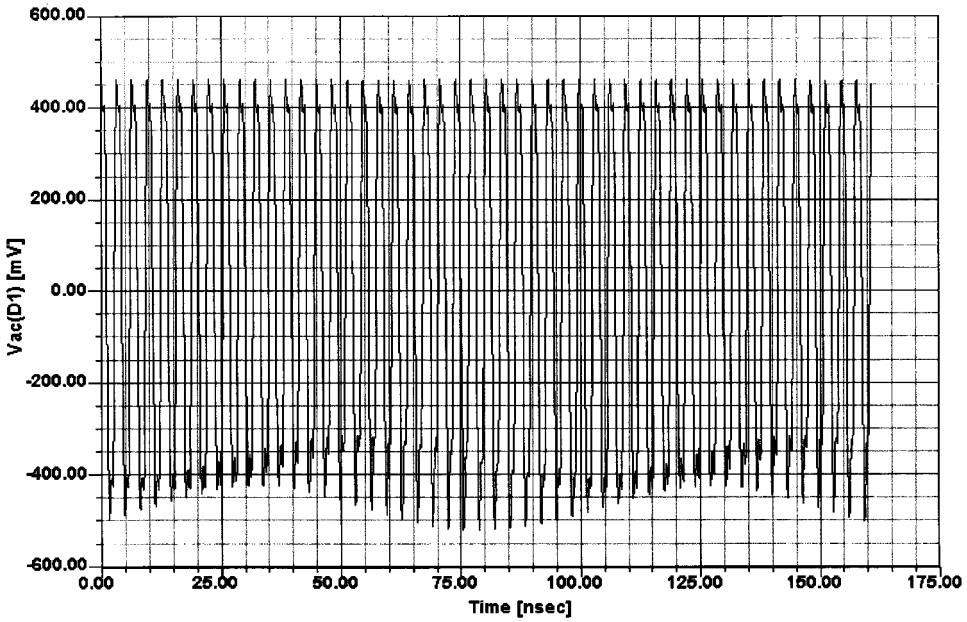


Figure 4-32 The anode-cathode voltage of one of the DBM's diodes, also over 50 LO cycles under two-tone IMD test conditions. The test conditions for this analysis are those for Figure 4-28 with LO = 13 dBm.

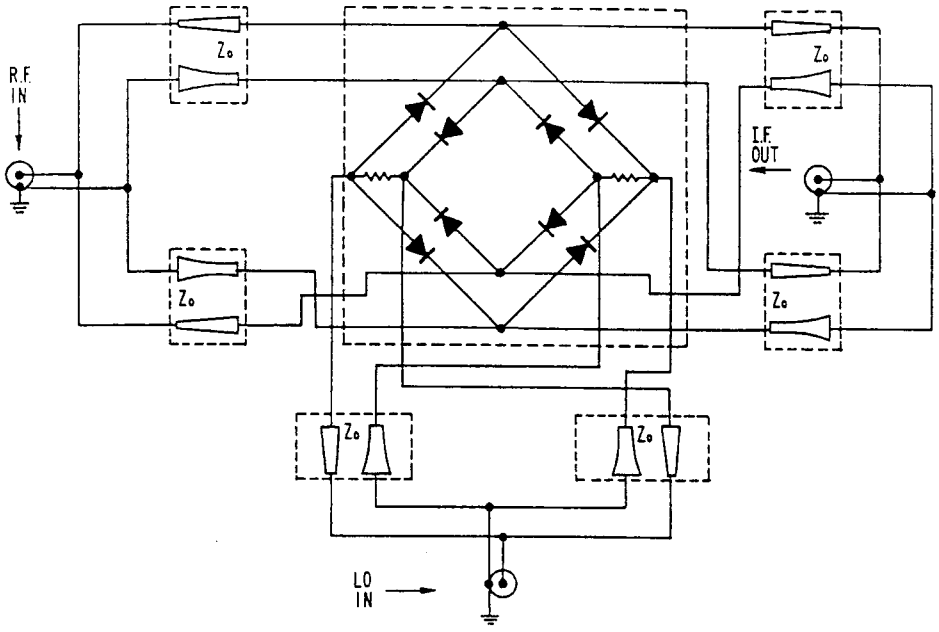


Figure 4-33 Example of a termination-insensitive mixer from U.S. Patent No. 4,224,572 (Adams-Russell Co., 1980) [11].

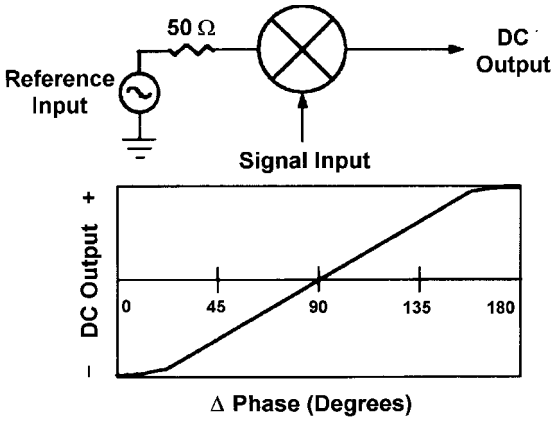


Figure 4-34 A mixer with a dc-coupled IF port can be used as a phase detector.

Termination-Insensitive Mixer. Figure 4-33 shows a mixer circuit that tolerates a fairly high VSWR at its output without significant degradation of its third-order IM performance [11].

Phase Detector. Theoretically, any mixer with a dc-coupled IF port can be used as a phase detector. When two signals of equal frequency are applied simultaneously to the reference and incoming signal ports, the phase detector produces a dc output at the IF port proportional to the cosine of the phase difference (Figure 4-34).

Binary Phase Shift Keying (BPSK) Modulator. Binary phase modulation occurs when a positive and negative signal current shifts the RF carrier between 0° and 180°. Figure 4-35 shows a double-balanced mixer operating as a BPSK modulator.

Quadrature Phase Shift Keying (QPSK) Modulator. A typical QPSK modulator consists of two biphase modulators, a 90° divider, and a 0° power combiner as shown. Data inputs at the control ports will cause the carrier to shift between 0°, 90°, 180°, and 270° as shown in Figure 4-36.

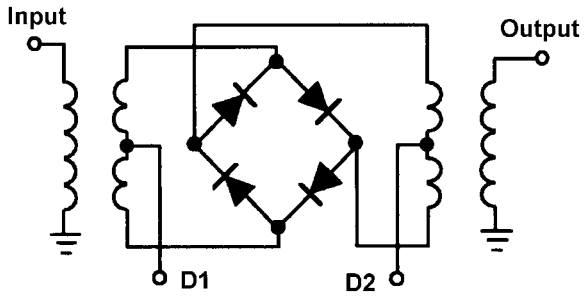


Figure 4-35 A diode-ring mixer as phase modulator.

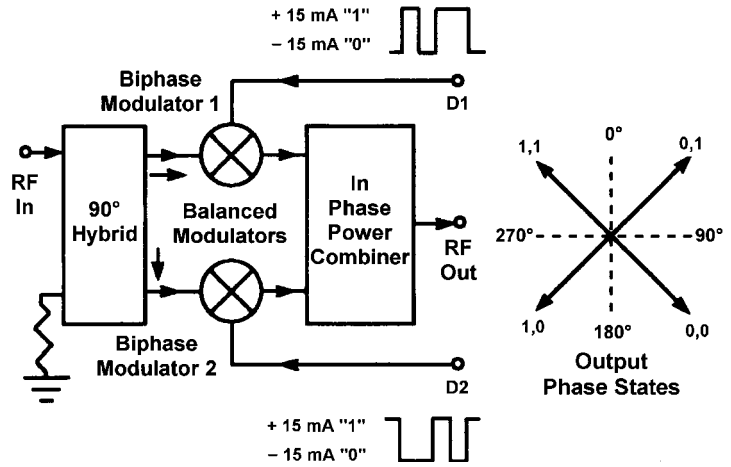


Figure 4-36 Two biphase modulators form the basis for a QPSK modulator.

Quadrature IF Mixer. A quadrature IF mixer produces two IF outputs in phase quadrature. Its basic structure consists of two double-balanced mixers, a 90° splitter, and a 0° splitter. The basic block diagram is shown in Figure 4-37.

Image-Reject Mixer. The image-reject mixer consists of a basic quadrature IF mixer with an additional 90° hybrid at the IF ports as shown in Figure 4-38. The primary function is to differentiate between the real signal and the image signal. This type of device is especially useful in applications where the desired RF signal and image are so close in frequency that rejecting the image with filtering is not practical.

Diode Attenuator/Switch. A ring of PIN diodes can be used as electronic attenuators by applying variable forward bias to the diodes (Figure 4-39). Maximum attenuation is achieved when the current at the control port is zero. The maximum attenuation is the isolation between the input and output ports. Minimum attenuation (insertion loss) is achieved when the IF-port current is 20-mA.

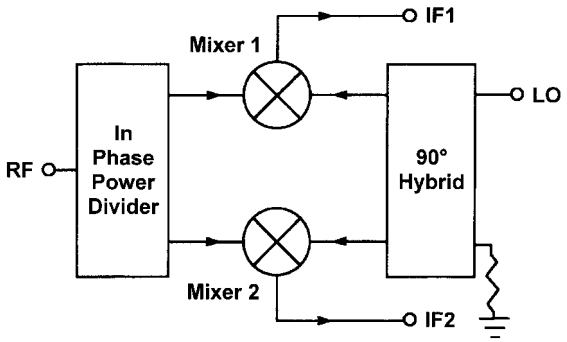


Figure 4-37 Quadrature mixer.

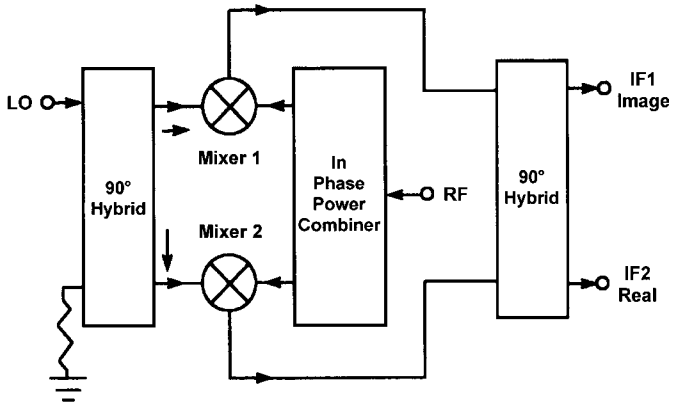


Figure 4-38 An image-reject mixer uses phasing to differentiate between its LO + IF and LO – RF IF outputs.

Single-Sideband (SSB) or In-Phase/Quadrature (I/Q) Modulator. SSB or I/Q modulators are useful in discriminating and removing the lower sideband (LSB) or upper sideband (USB) generated during frequency conversion, especially when the sidebands are very close in frequency and attenuation of one of the sidebands cannot be achieved with filtering. This is the case with audio and video modulation, where signals from dc to 10 MHz must be converted to a higher frequency that is appropriate for transmission. In such cases, both sidebands will be very close in frequency to the carrier frequency. With an I/Q modulator, one of the sidebands is easily canceled or attenuated along with its carrier.

Attenuation of the carrier has been the most troublesome aspect in the design of passive I/Q modulators. Isolation between the local oscillator (LO) port and the RF port of the mixers, which is the main parameter in determining carrier rejection, is usually insufficient at frequencies above 200 MHz.

I/Q modulator designs basically consist of two double-balanced mixers (Figure 4-40). The mixers are fed at the LO ports by a carrier phase-shifted through a 90° hybrid. Thus, the carrier signal's relative phase is 0° to one mixer and 90° to the other mixer. Modulation signals are fed externally in phase quadrature to the two mixers' IF ports. The mixers' modulated output signals are combined through a two-way, in-phase power divider/combiner.

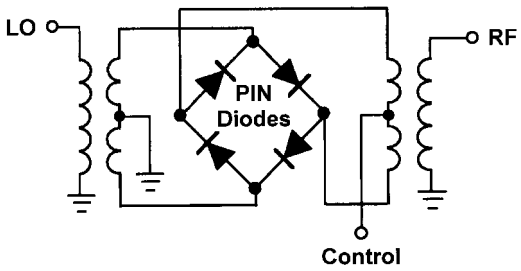


Figure 4-39 A diode DBM can be used as a dc-controlled attenuator if PIN diodes are used instead of Schottky devices in its ring.

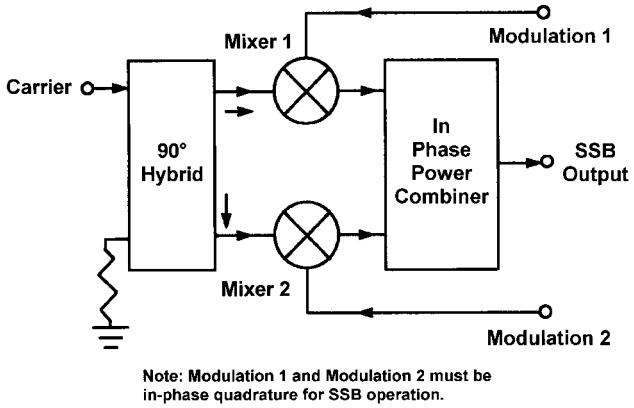


Figure 4-40 An SSB modulator matches two high-frequency mixers, a 90° hybrid, and an in-phase power combiner to produce an SSB output signal.

The circuit forms a phase-cancellation network to one of the sidebands and a phase-addition network to the other sideband. The carrier is somewhat attenuated and is directly dependent on the inherent LO-to-RF isolation of the mixers and the modulating signal level. In industry-standard I/Q modulators, USB suppression results when the first modulation port (Modulation 1) is fed with a signal that is 90° in advance of the signal feeding the second modulation port (Modulation 2). Opposite phasing can be arranged by changing the internal phase polarity of the mixers or by interchanging the 90° hybrid output ports to the LO ports of the mixers.

The phase and amplitude imbalances between the various components used in the manufacturing of the I/Q modulators must be tightly maintained for optimum SSB rejection. Matching of the two mixers for conversion loss and insertion phase is extremely critical, since differences in these parameters will add to amplitude- and phase-imbalance errors. The 90° hybrid in the LO port must be in nearly perfect phase quadrature.

Phase- and amplitude-imbalance errors adversely affect sideband suppression (Figure 4-41). In most cases, a typical passive I/Q modulator operates with a carrier input level of +10 dBm, which is required to drive the diodes in the mixers to operate in the linear range. The dynamic range of these mixers can be improved significantly by using diodes with a higher barrier height. The LO signal in this case must be increased in order to drive these diodes into conduction in their linear range.

Carrier rejection is also a problem when designing an SSB modulator, since only a few decibels of suppression can be achieved in standard high-frequency models. In the past, the major contributor to carrier suppression was the inherent LO-to-RF isolation through the mixers. Unfortunately, this isolation is usually poor at cellular frequencies (800–1000 MHz), where at least 25 dB of carrier rejection is necessary. In some cases, designers feed a small amount of dc into the IF ports to control the carrier rejection, which complicates the driver circuitry and calls for temperature compensation when operating at different temperatures.

As an example, an SSB modulator is assumed to operate with +10-dBm LO drive with each modulating signal at –10 dBm and in phase quadrature to each other when applied to the modulating ports (MOD 1 and MOD 2). The result will be a modulated signal at –16 dBm, assuming 6-dB conversion loss. For 20-dB carrier rejection with respect to the desired

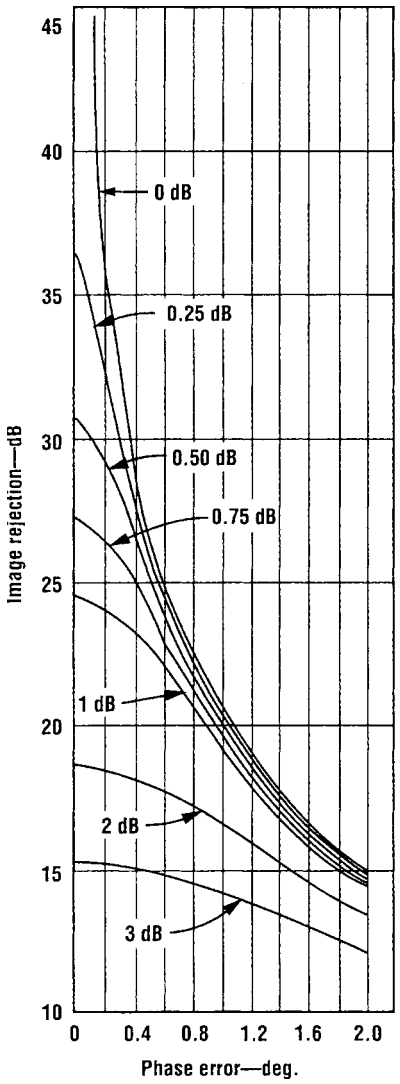


Figure 4-41 The level of SSB rejection improves as the phase- and amplitude-imbalance performance of an SSB modulator improves.

modulated signal, the carrier must be at –36 dBm, which translates to LO-to-RF isolation of 46 dB.

By employing a subharmonic approach, the performance of SSB modulators can be extended beyond the limits of conventional designs as reported by Joshi [7]. The approach is based on the use of subharmonic mixers in place of fundamental-frequency mixers and is applicable from about 140 to 3000 MHz. Subharmonic mixers use antiparallel diode pairs in their construction [12–14]. Matched antiparallel diode pairs used in single-ended or single-balanced mixer configurations cancel even-order intermodulation products (such as $2f_{LO} \times 2f_{RF}$, $3f_{LO} \times 3f_{RF}$, etc.) at all ports.

Single-ended mixers lack the port-to-port isolation needed for SSB modulator applications. Odd-order products of the RF and LO frequencies (even $f_{LO}\times$ odd f_{RF}) and (odd

$f_{LO} \times$ even f_{RF}) appear on all ports, requiring extensive filtering for satisfactory performance. For a single-balanced mixer, even harmonics of the LO combining with odd harmonics of the RF appear at the IF port, whereas odd harmonics of the LO combining with even harmonics of the RF appear at the RF and IF ports. This assumes that a balanced transformer is placed at the LO port, which is a logical choice due to the fact that the highest level signal appears at the LO port. Since the desired odd-order IF products appear at both the RF and IF ports, a need arises for a diplexing network to isolate the RF and IF signals.

The subharmonic modulator design provides a unique way to isolate the RF and IF signals. A single-balanced harmonic mixer offers good LO-to-RF and LO-to-IF isolation but poor RF-to-IF isolation. Fortunately, harmonically related signals are spaced well apart in the frequency spectrum, simplifying filtering of harmonically related signals.

Harmonic mixing also works well with low LO power levels, with somewhat lower 1-dB compression on the RF port than with fundamental-frequency mixing. The ability to operate with LO frequencies that are a fraction of the carrier frequency (1/2, 1/4, 1/6, etc.) significantly reduces the cost of an LO source, especially at higher frequencies. Also, using lower-frequency LO sources helps avoid the signal-leakage problems inherent with higher-frequency LO sources. Minimizing signal leakage, especially at higher frequencies, becomes expensive and bulky. Subharmonic mixing offers several advantages:

- The technique offers the ability to operate at LO frequencies that are 1/2, 1/4, or 1/6 of the carrier frequency. For example, for an IF of 100 MHz at an RF of 2 GHz, the LO can be $(2000 \pm 100) \div 2 = 950$ or 1050 MHz.
- The LO's even harmonics are strongly attenuated.
- The filtering requirements for fundamental frequency and odd harmonic signals of the LO are not critical.
- The cost of generating the LO is reduced due to the fact that the LO frequency need only be a fraction of the carrier frequency.

As an example of the performance improvements possible with the subharmonic mixers, units were evaluated at both cellular (935–960 MHz) and PCN/PCS (1.8–1.9 GHz) bands.

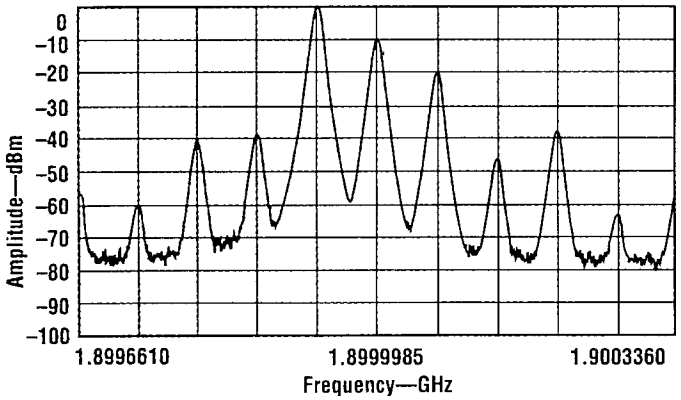


Figure 4-42 This plot of carrier and sideband rejection was measured for a conventional SSB modulator operating at 1.9 GHz.

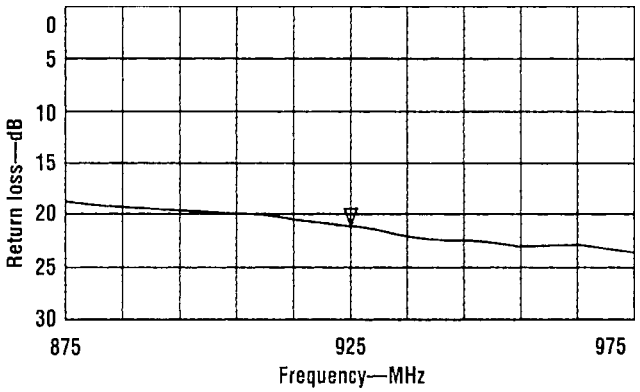


Figure 4-43 The SSB modulator's return loss as measured at the local oscillator (LO) port.

For a conventional SSB modulator at 1.9 GHz fed with +10-dBm modulation signals, carrier rejection is barely 10 dB (Figure 4-42).

Sideband rejection can be improved by tuning, but the carrier rejection is controlled by the LO-to-RF isolation of the double-balanced mixers. Conventional double-balanced mixers with high isolation at cellular and PCN bands are very expensive and large when special techniques are used to improve LO-to-RF isolation. In contrast, the subharmonic nature of the new approach allows the use of lower-frequency, less-expensive components in the modulators' construction.

The subharmonic modulators offer an improvement of more than 15 dB in carrier suppression compared to the conventional approach.

The measured VSWR (return loss) at the LO and RF ports is better than 1.50:1 (Figures 4-43 and 4-44). Measurements made on a cellular-band SSB modulator reveal carrier rejection on the order of 40 dB. Typical insertion loss is 7 dB while sideband rejection is 30 dB (Figure 4-45).

By virtue of harmonic mixing, even-order mixing products are attenuated by about 30 dB with respect to the desired modulated output signal. The fundamental-frequency feedthrough into the output port is approximately 5 dB lower than the desired modulated signal, whereas

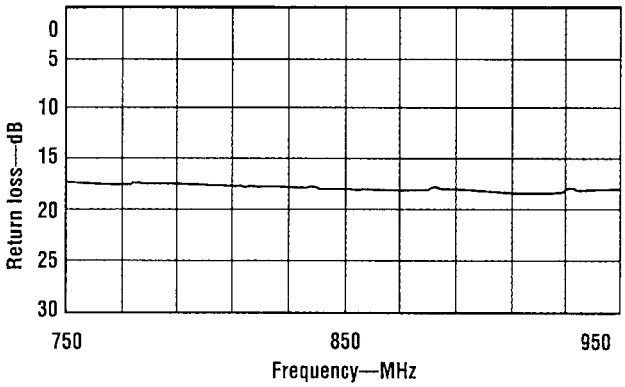


Figure 4-44 The novel harmonic SSB modulator's return loss as measured at the RF port.

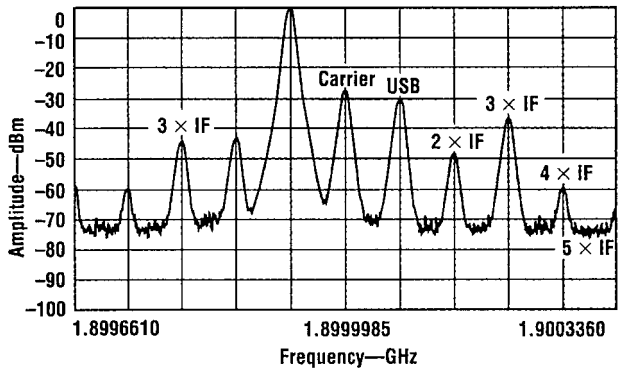


Figure 4-45 Plot of carrier and sideband rejection as measured for the novel harmonic SSB modulator operating at cellular frequencies.

the fourth harmonic mixing with the modulating signal is approximately 10 dB lower. Typical loss for fourth harmonic mixing is 17–19 dB while maintaining 30 dB of carrier rejection.

Since harmonically related products are well-spaced in frequency, filtering undesired signals is relatively inexpensive using standard octave-bandwidth filters. Low-cost commercial bandpass filters typically offer better than 40–50-dB attenuation of unwanted harmonic signals. Constant-impedance bandpass filters offering good impedance match at desired

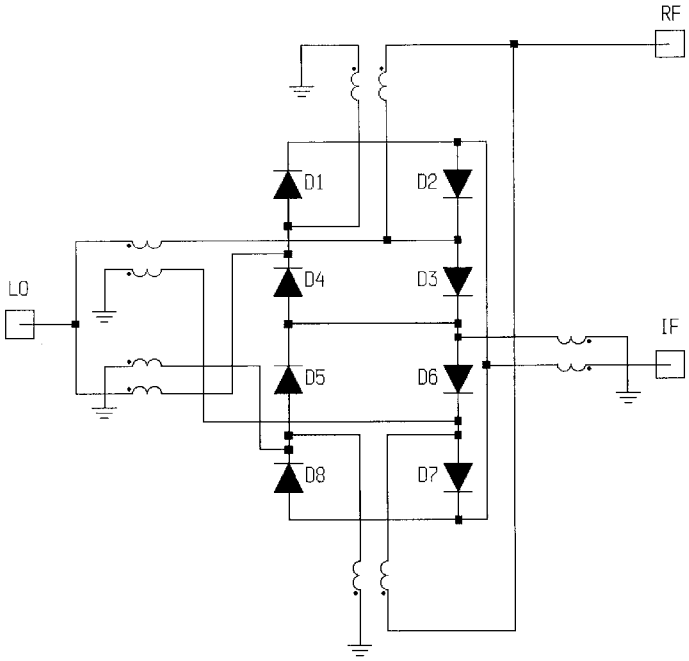


Figure 4-46 A triple-balanced diode mixer. A limitation of this configuration is that the internal dc common connections associated with its RF and LO transformers disallow usable IF response down to dc.

stopbands can also be used in cases where harmonically related products require impedance termination within a system.

The subharmonic modulator design is easily applied at custom frequencies. Conversion of an SSB modulator with output frequency corresponding to twice the LO frequency to one with output corresponding to four times the LO frequency requires only one component change, in the form of a signal-combining network at the modulator's output. Although the conversion loss of the fourth harmonic LO component mixing with the modulating signal is in the vicinity of 18 dB, the cost of generating the LO is drastically reduced with the subharmonic modulator. In spite of higher signal loss, the carrier rejection is still at least 30 dB at the fourth harmonic, and harmonically related products can be eliminated with an inexpensive filter.

Triple-Balanced Mixer. Two diode rings can be combined to form a double double-balanced mixer, or triple-balanced mixer (TBM), as shown in Figure 4-46. Triple-balanced mixers achieve higher dynamic range and interport isolation than double-balanced designs at the expense of LO power and increased complexity and size.

Figure 4-47 shows the interport isolation of the circuit in Figure 4-46 with the circuit configured in a less than ideally balanced form, with small variations in transformer-winding inductance and diode parameters introduced for more realistic modeling. Note that the mixer's interport isolation generally increases with frequency, rather than decreasing with frequency as with the DBM (Figure 4-24).

Rohde & Schwarz Subharmonically Pumped DBM. Figure 4-48 shows an example of a complete microwave diode mixer used as the input stage of a spectrum analyzer. Its LO is applied at 13.2 GHz, one-third the frequency necessary to mix the 40.1-GHz input

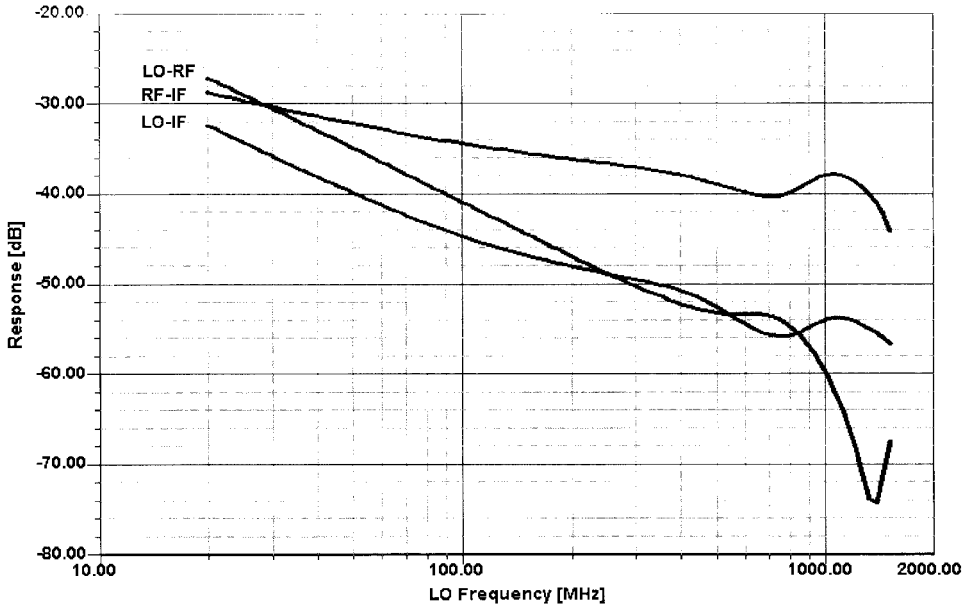


Figure 4-47 The triple-balanced mixer offers improved high-frequency isolation over a standard DBM.

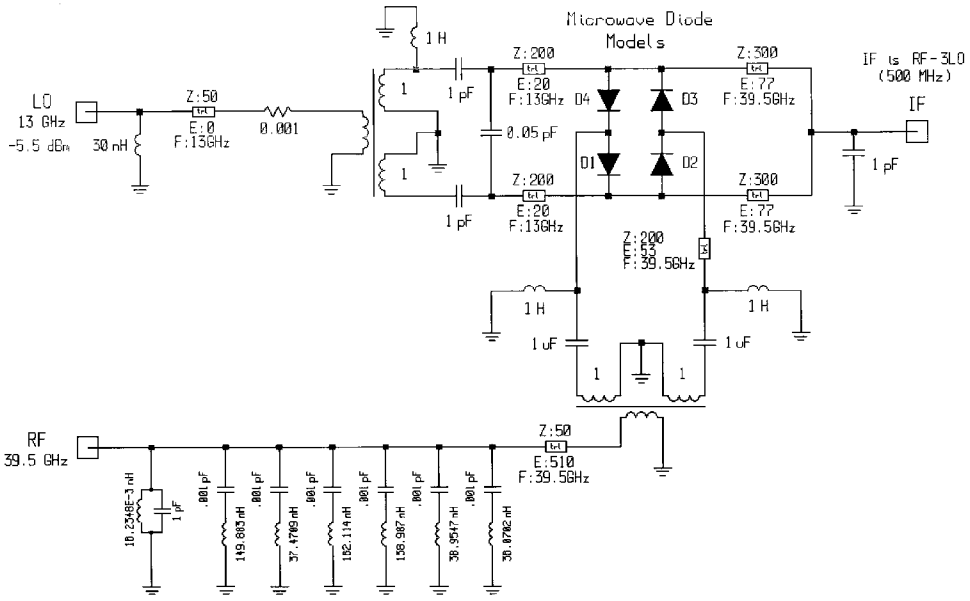


Figure 4-48 Schematic of the spectrum-analyzer mixer. The LO is applied at one-third the frequency necessary to mix the RF signal to the 500-MHz IF.

signal to the 500-MHz IF, with distortion in the diodes providing the 3× frequency multiplication. The circuit's conversion gain is –18.8 dB; its noise figure, 20.5 dB.

The subharmonic-drive technique exemplified by this mixer is important for another reason in addition to LO isolation. Depending on the application, directly generating a sufficiently phase-quiet LO signal at the higher wireless frequencies and on up into the microwave frequencies may not be feasible. In such cases, injecting the LO at a subharmonic of the desired LO frequency may provide better phase noise performance than a fundamental LO even though the phase noise of a frequency-multiplied source increases by $\log_{10} n$ dB, where n is the multiplication factor.

4-4 TRANSISTOR MIXERS

Diode mixers are lossy and termination-sensitive and require considerable LO power. Their attractiveness to designers of highly integrated wireless products is further reduced by their dependence on transformers for balance and port-to-port isolation. Transistor mixers are therefore used where high integration and reduced current drain are paramount—that is, in most non-base-station wireless applications.

Until ten years ago, single BJTs were commonly used in a simple additive-mixing arrangement, an example of which appears in Figure 4-49. Such a circuit behaves like a combination of a preamplifier and single-diode mixer. Single-BJT mixers were used in early AM–FM radios until 1980, and unfortunately in some handheld 2-meter and 70-cm ham equipment, to achieve the lowest possible power consumption. By definition, however, the presence of such a mixer also destroys any possibility of achieving a high intercept point. In

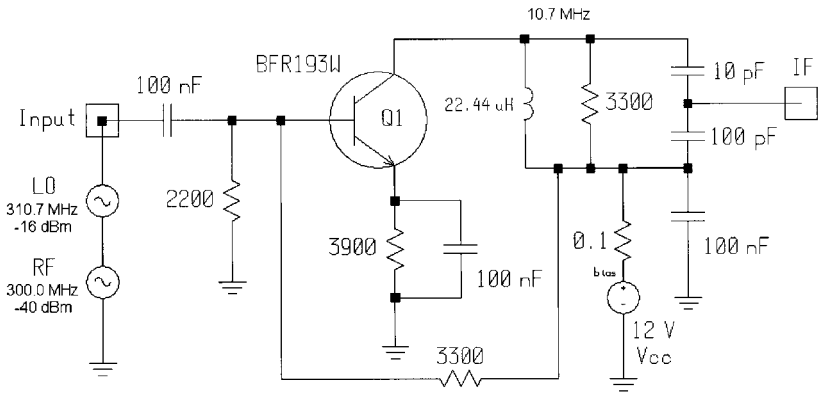


Figure 4-49 A single-BJT mixer.

addition to this, many combination frequencies occur despite elaborate input filtering. Although single-BJT mixers can exhibit considerable conversion gain (>10 dB), their dynamic range is restricted and their port-to-port isolation is poor. BJT mixers in today's competitive wireless designs use multiple transistors and are based almost exclusively on the Gilbert multiplier cell.

4-4-1 BJT Gilbert Cell

The classic active mixer (Figure 4-50), conceived by Barrie Gilbert in 1967, is the basis for most active mixers used in wireless products today. Figure 4-51 shows a Gilbert cell mixer implementation for analysis. Figure 4-52 shows its conversion gain and noise figure versus LO power. Figure 4-53 shows its IM_3 and IP_3 responses. As Gilbert wrote in a 1994 monograph [15]:

> This circuit is attractive because (a) it can be monolithically integrated with other signal-processing circuitry; (b) it can provide conversion gain, whereas a diode-ring mixer always has conversion loss; (c) it requires very low power to drive the LO port; (d) it provides excellent isolation between the signal ports; and (e) it is far less fussy about load-matching.

The two major advantages afforded by the Gilbert cell are port isolation (it should really be called a multiplicative mixer) and its significant reduction of even-order frequency combinations. As with the single-transistor mixer, the Gilbert cell does not particularly shine with high intercept point and low current consumption. Another major drawback is that it requires a high-impedance output, which gives the user headaches because most of the better filters are in 50-Ω technology, and only very simple monolithic filters offer input and output impedances around 1 kΩ. Impedances on this order can result in significant crosstalk in high-density circuit boards.

The reason the Gilbert cell can be so linear for small input signal levels is twofold: It is a differential amplifier and it makes use of the tanh function, which is linear over a wide range around the zero-crossing point [16]. Attempts have been made to increase the intercept point. The Plessey SL6440 IC was a special version of the Gilbert cell that achieved high $IP_{3,in}$ (+30 dBm) at the expense of dc current.

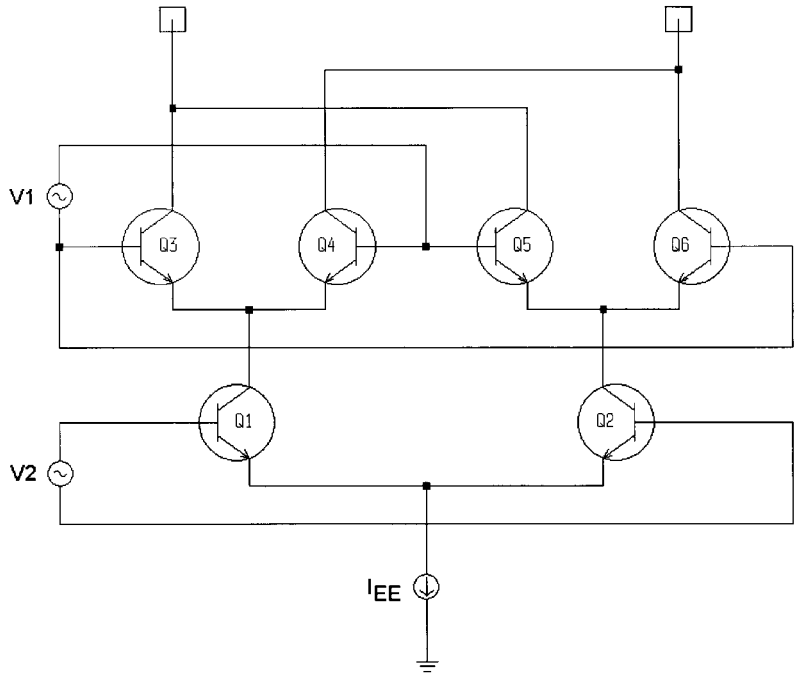


Figure 4-50 The basic Gilbert multiplier cell.

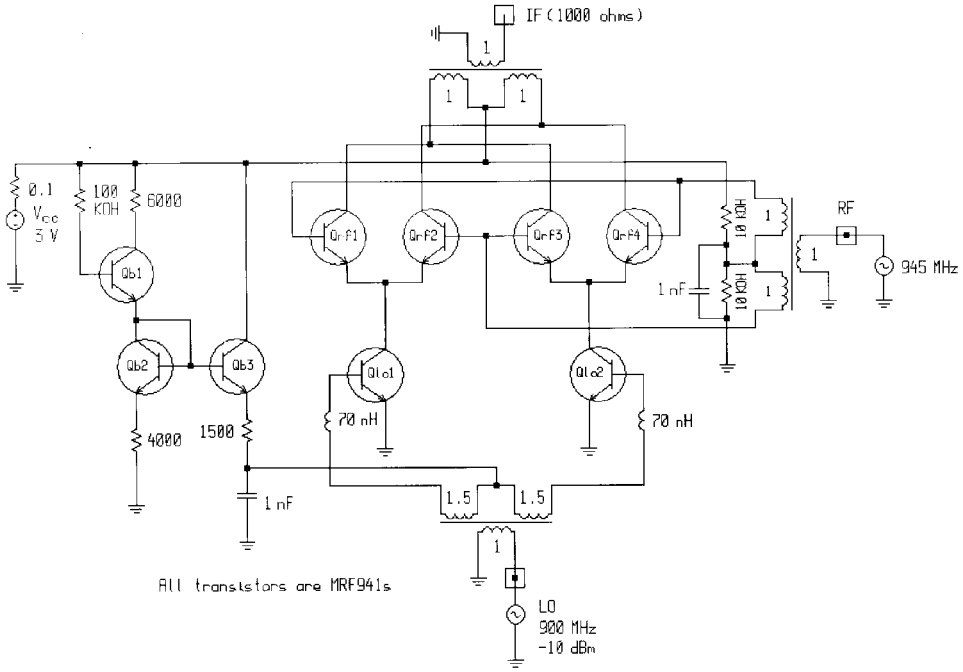


Figure 4-51 Gilbert cell mixer validation circuit.

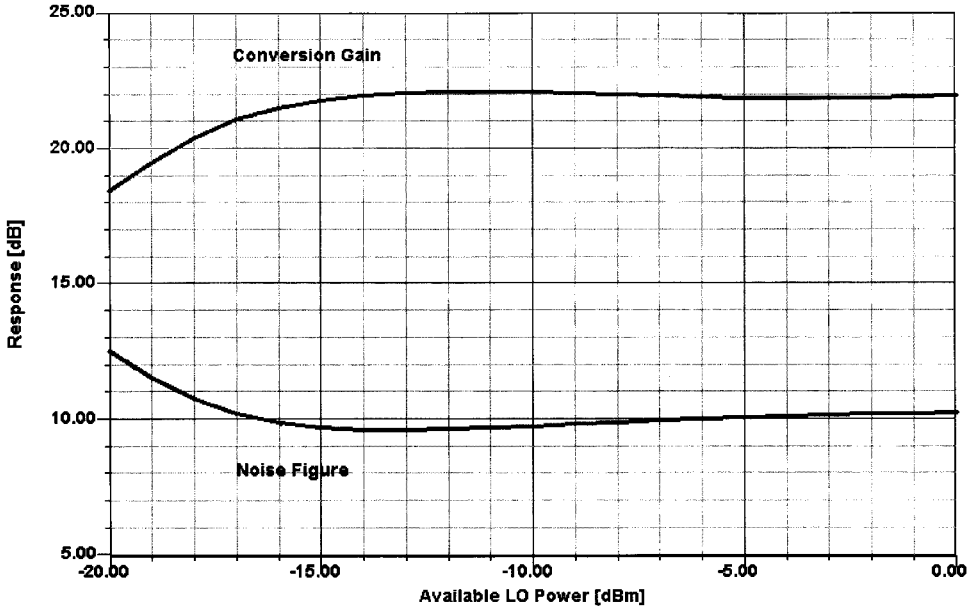


Figure 4-52 Conversion gain and noise figure versus LO power for the Gilbert cell mixer. In this analysis, LO = 900 MHz (–20 to 0 dBm) and RF = 945 MHz (–50 dBm) for an IF of 45 MHz.

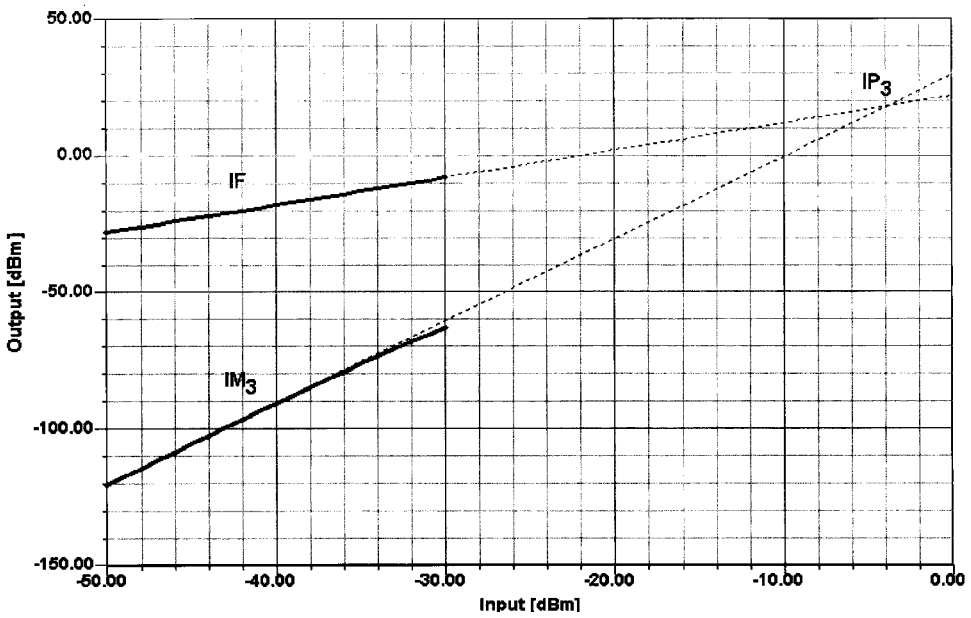


Figure 4-53 IM_3 and IP_3 responses for the BJT Gilbert cell mixer with RF signals at 945 and 946 MHz. In this analysis, LO = 900 MHz (–10 dBm), RF1 = 945 MHz (–50 to –30 dBm), and RF2 = 946 MHz (–50 to –30 dBm). Four LO harmonics and three LO sidebands were used.

Our suspicion, however, is that a combination of a modern double-balanced passive mixer using medium- to high-level diodes followed by a feedback FET amplifier, or even one of the late CATV transistors with f_T values of more than 25 GHz, would allow the achievement of a significantly higher intercept point. As pointed out in Chapter 3, CATV transistors are now available that combine a noise figure of less than 1 dB with a $IP_{3,in}$ of more than 30 dBm. Built-in push–pull structures using such transistors can achieve an $IP_{2,in}$ of 70 dBm, matching the numbers for the diode DBM. In Chapter 3 we learned that feedback always improves the dynamic range, mostly at the expense of gain, since the gain–bandwidth product gets reduced. Motorola has recently come out with a clever extension of its high-level preamplifier and modified it to become a mixer.

4-4-2 BJT Gilbert Cell with Feedback

Motorola's MC13143 low-power 2.4-GHz mixer IC uses a patented topology consisting of a Class-AB-biased Gilbert cell augmented by feedback. Its linearity can be programmed via

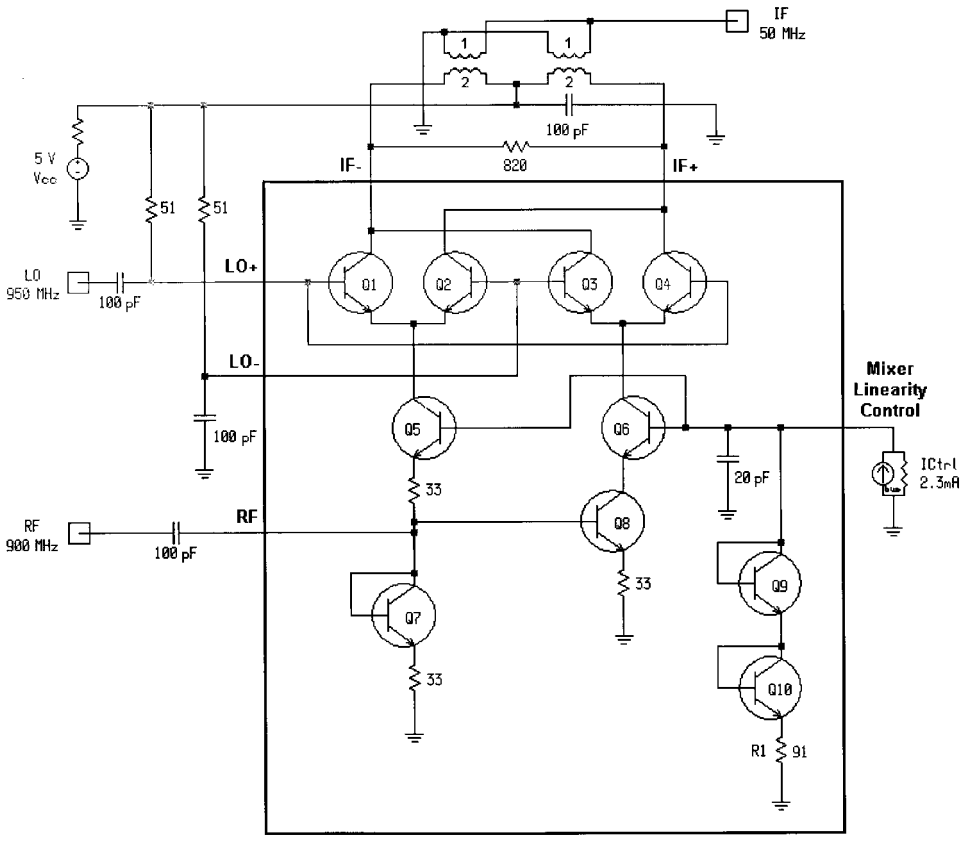


Figure 4-54 Motorola's MC13143 mixer IC uses dc feedback for improved linearity that can be programmed by applying a control current (0–2.3 mA) to its Mixer Linearity Control pin. Per Motorola, an $IP_{3,in}$ of 20 dBm may be achieved with a control current of 2.3 mA, at the expense of approximately 7 mA of additional supply current. In this validation circuit, R_1 (the resistor in Q10's emitter) has been adjusted for a total current drain (supply + linearity control) of 8 mA.

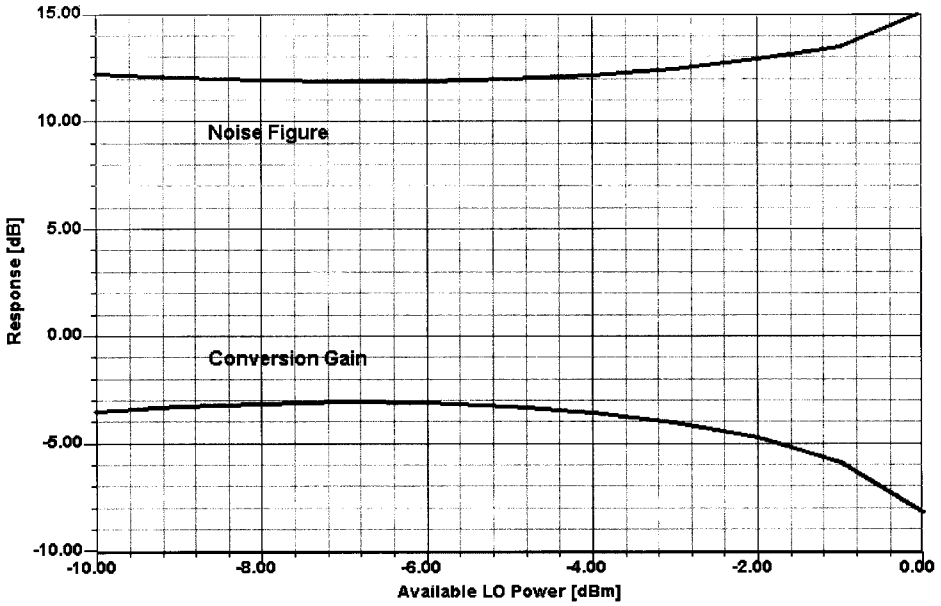


Figure 4-55 Calculated conversion gain and noise figure versus LO drive for the Motorola MC13143 IC operating at $V_{CC} = 2$. Motorola's specifications for the typical values of these characteristics are –5.0 dB and 12 dB, respectively, at an LO drive level of 0 dBm. The RF signal is at 900 MHz and the LO is at 950 MHz, for an IF of 50 MHz. In this analysis, LO = 950 MHz (–10 to 5 dBm) and RF = 901 MHz (–50 dBm) for an IF of 49 MHz.

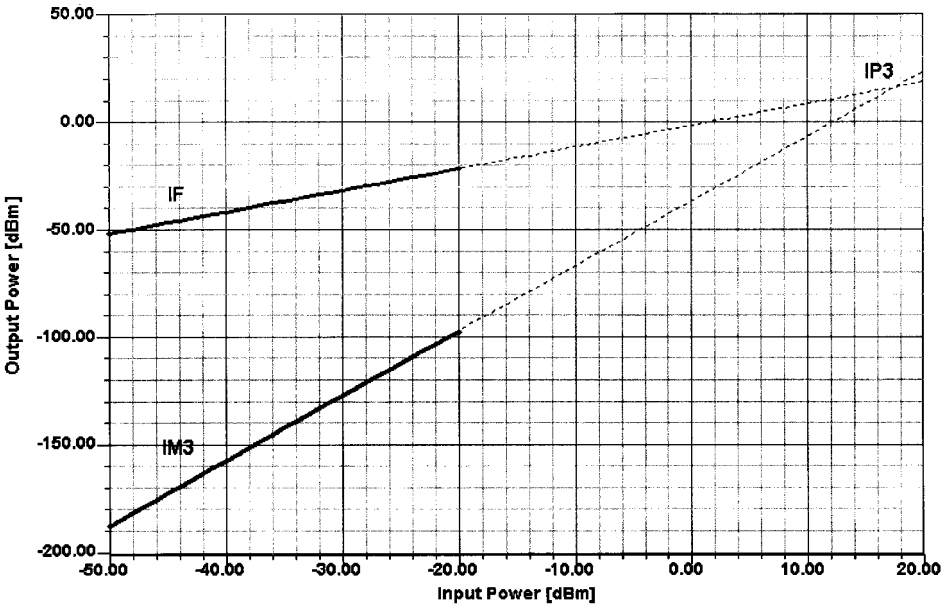


Figure 4-56 Calculated typical IF, IM_3, and IP_3 responses for the Motorola MC13143 mixer. In this analysis, LO = 950 MHz (–5 dBm), RF1 = 900 MHz (–50 to –20 dBm), and RF2 = 901 MHz (–50 to –20 dBm), and $V_{CC} = 5$. Four LO harmonics and three LO sidebands were used.

an external current source to achieve an $IP_{3,in}$ of 20 dBm at the expense of additional supply current.

The MC13143 contains 29 active transistors. Figure 4-54 shows the core of the MC13143 circuit reduced to its essentials and configured for computer analysis per Motorola's MC13143 test circuit. Figure 4-55 shows the circuit's conversion gain and noise figure versus LO drive, and Figure 4-56 shows the circuit's typical IF, IM_3, and IP_3 responses, both as calculated by Ansoft Corporation's Serenade 8.0 circuit simulator. A reproduction of the MC13143 datasheet, used by permission, follows.

In considering the approach used in the MC13143, one has to remember that feedback around many stages gives an IMD ripple, which means that the higher-order IPs are no longer a straight-line calculation. Above a certain level, the 3-dB/dB law fails, and as a result of this, the IMD characteristic can be as bad as 10 or 20 dB/dB—almost like a breakdown point.

High-frequency capability is an issue with silicon BJTs; the upper frequency limit of the technique is in the range of 3 GHz. Bipolar technology will be extended to 10 GHz by SiGe, but when these parts will be commercially available at competitive prices is currently unclear. There has not been enough published on the actual performance of Gilbert-cell-type mixers using SiGe transistors, but we assume that because of coupling on the silicon wafer, things are no longer as well behaved as we are used to at much lower frequencies.

4-4-3 FET Mixers

FETs, used in active and passive circuits, are still a popular solution for low-power integrated mixers up to 100 MHz (passive quad) and 1 GHz (dual-gate MOSFET). Cost-sensitive applications may require the use of a single-gate FET. These types of single-gate microwave mixers are frequently used to validate the quality of software but the user keeps forgetting this is a question of the quality of the model and not of the simulator. A good example of bad modeling is Figure 7.88 in Vendelin et al. [17]. Unfortunately, we have not seen the same circuit analyzed with modern tools, nor do the authors of the paper cited [18] provide sufficient details of the circuit and the transistor for us to do this ourselves.

Figure 4-57 shows a basic, active FET mixer. Like the single-diode and single-BJT mixers, it is additive, although it's common to lessen RF–LO interaction by injecting one signal at the gate and the other at the source as shown. Single-gate GaAs and LDMOS FETs are also used in this arrangement with gate biasing as appropriate to the device type. Figure 4-58 shows a dual-gate FET in an additive mixing configuration.

The linearity of FETs is based on the fact that a FET follows a square law and therefore the first derivative, its transconductance, is supposed to be constant. This is valid within a wide amplitude range. FET mixers have a noise figure similar to those found in bipolar mixers: typically around 6–9 dB, depending on the configuration. A FET mixer's intercept point is subject to load-impedance variations, with a purely resistive termination providing the best case. Terminating the mixer with a filter is problematic in that a filter looks purely resistive only within its 3-dB passband; in the transition band and beyond, the filter impedance rises rapidly and the mixer intercept point goes down. The best means devised so far to minimize this effect is to configure the FET output as a high-pass filter, using capacitive coupling from the output tuned circuit to a 50-(bandpass filter, as shown in Figure 4-59. Outside its passband, this impedance inverter acts more like a short circuit and maintains IMD products at a reasonable level. The alternative to this is the popular diplexer, which requires more components and has more insertion loss. The high-pass configuration has barely been mentioned in the literature.

Order this document by MC13143/D

MOTOROLA

Ultra Low Power DC - 2.4 GHz Linear Mixer

The MC13143 is a high compression linear mixer with single–ended RF input, differential IF output and differential LO inputs which consumes as little as 1.8 mW. A new circuit topology is used to achieve a high third order intermodulation intercept point, high linearity and high 1.0 dB output compression point while maintaining a linear 50 Ω input impedance. It is designed for Up or Down conversion anywhere from dc to 2.4 GHz.

Ultra Low Power: 1.0 mA @ V_{CC} = 1.8 to 6.5 V

- Wide Input Bandwidth: DC–2.4 GHz
- Wide Output Bandwidth: DC–2.4 GHz
- Wide LO Bandwidth: DC–2.4 GHz
- High Mixer Linearity: $P_{i1.0\ dB}$ = 3.0 dBm

Linearity Adjustment of up to IP_{3in} = 20 dBm

- 50 Ω Mixer Input
- Single–Ended Mixer Input
- Double Balanced Mixer Operation
- Differential Open Collector Mixer Output

ORDERING INFORMATION

Device	Operating Temperature Range	Package
MC13143D	T_A = –40 to 85°C	SO–8

MAXIMUM RATINGS (T_A = 25°C, unless otherwise noted.)

Rating	Symbol	Value	Unit
Power Supply Voltage	V_{CC}(max)	7.0	Vdc
Junction Temperature	T_{Jmax}	150	°C
Storage Temperature Range	T_{stg}	–65 to 150	°C

NOTE: ESD data available upon request.

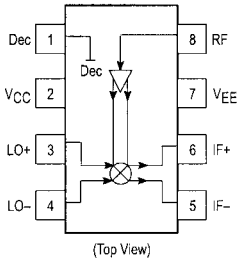
MC13143

ULTRA LOW POWER DC – 2.4 GHz LINEAR MIXER

SEMICONDUCTOR TECHNICAL DATA

D SUFFIX
PLASTIC PACKAGE
CASE 751
(SO–8)

PIN CONNECTIONS



(Top View)

This device contains 29 active transistors.

MC13143

RECOMMENDED OPERATING CONDITIONS

Rating	Symbol	Min	Typ	Max	Unit
Power Supply Voltage	V_{CC}	1.8	–	6.0	Vdc

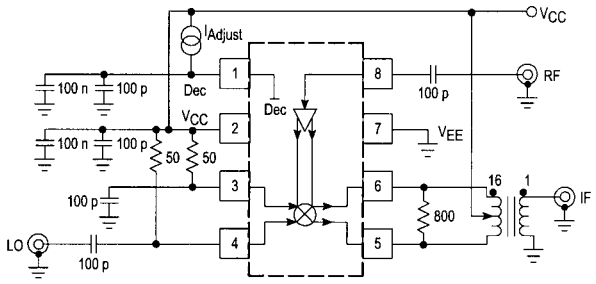
DC ELECTRICAL CHARACTERISTICS (T_A = 25°C, V_{CC} = 3.0 V, f_{RF} = 1.0 GHz, Pin = –25 dBm.)

Characteristic	Symbol	Min	Typ	Max	Unit
Supply Current (Lin Control Current = 0)	I_{CC1}	–	1.0	–	mA
Supply Current (Lin Control Current = 1.6 mA)	I_{CC2}	–	4.1	–	mA

AC ELECTRICAL CHARACTERISTICS (T_A = 25°C, V_{CC} = 3.0 V, f_{RF} = 1.0 GHz, Pin = –25 dBm.)

Characteristic	Symbol	Min	Typ	Max	Unit
Mixer Voltage Conversion Gain (R_P = R_L = 800 Ω)	VG_C	–	9.0	–	dB
Mixer Power Conversion Gain (R_P = R_L = 800 Ω)	PG_C	–3.5	–2.6	–1.5	dB
Mixer Input Return Loss	Γin_{mx}	–	–20	–	dB
Mixer SSB Noise Figure	NF_{SSB}	–	14	15	dB
Mixer 1.0 dB Compression Point (Mx Lin Control Current = 1.6 mA)	$Pin_{-1.0\ dB}$	–1	0	–	dBm
Mixer Input Third Order Intercept Point (d_f = 1.0 MHz, $I_{control}$ = 1.6 mA)	$IP3_{in}$	–	16	–	dBm
LO Drive Level	LO_{in}	–	–5.0	–	dBm
LO Leakage to Mixer IF Outputs	P_{LO-IF}	–	–33	–25	dB
Mixer Input Feedthrough Output	P_{RFm-IF}	–	–25	–	dB
LO Leakage to Mixer Input	P_{LO-RFm}	–	–40	–25	dB
Mixer Input Leakage to LO	P_{RFm-LO}	–	–35	–	dB

Figure 1. Test Circuit



MC13143

TYPICAL PERFORMANCE CURVES

Figure 2. Power Conversion Gain and Supply Current versus Supply Voltage

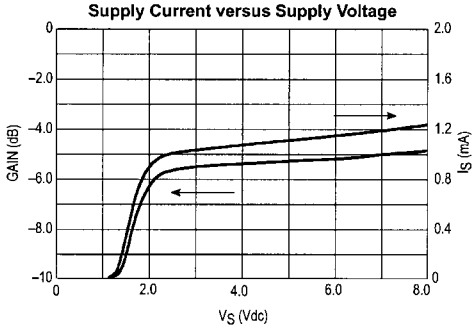


Figure 3. Noise Figure and Gain versus LO Power

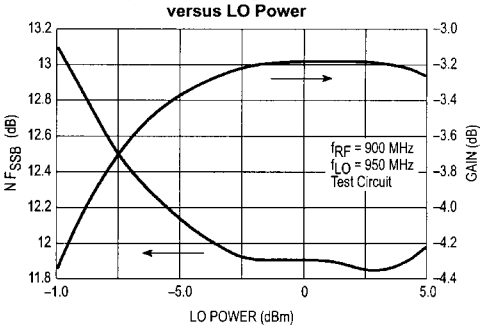


Figure 4. Mixer Input Return Loss versus RF Input Frequency

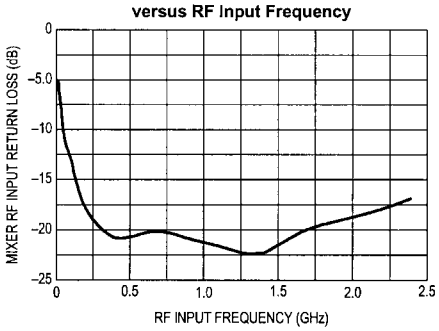


Figure 5. Power Conversion Gain and Supply Current versus RF Input Power

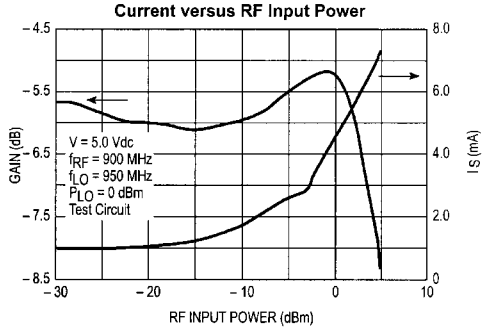


Figure 6. Noise Figure and Gain versus RF Frequency

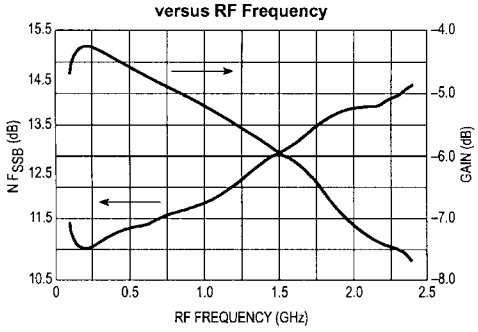
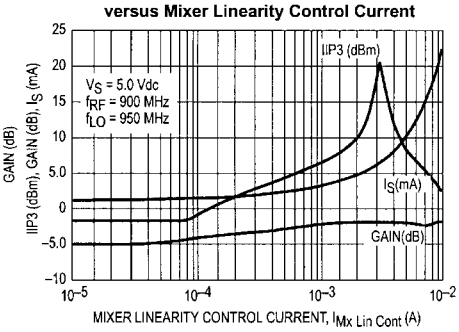


Figure 7. IIP3, Gain, Supply Current versus Mixer Linearity Control Current



MC13143

CIRCUIT DESCRIPTION

General

The MC13143 is a double–balanced Mixer. This device is designated for use as the frontend section in analog and digital FM systems such as Wireless Local Area Network (LAN), Digital European Cordless Telephone (DECT), PHS, PCS, GPS, Cellular, UHF and 800 MHz Special Mobile Radio (SMR), UHF Family Radio Services and 902 to 928 MHz cordless telephones. It features a mixer linearity control to preset or auto program the mixer dynamic range, an enable function and a wideband IF so the IC may be used either as a down converter or an up converter.

Current Regulation

Temperature compensating voltage independent current regulators provide typical supply current at 1.0 mA with no mixer linearity control current.

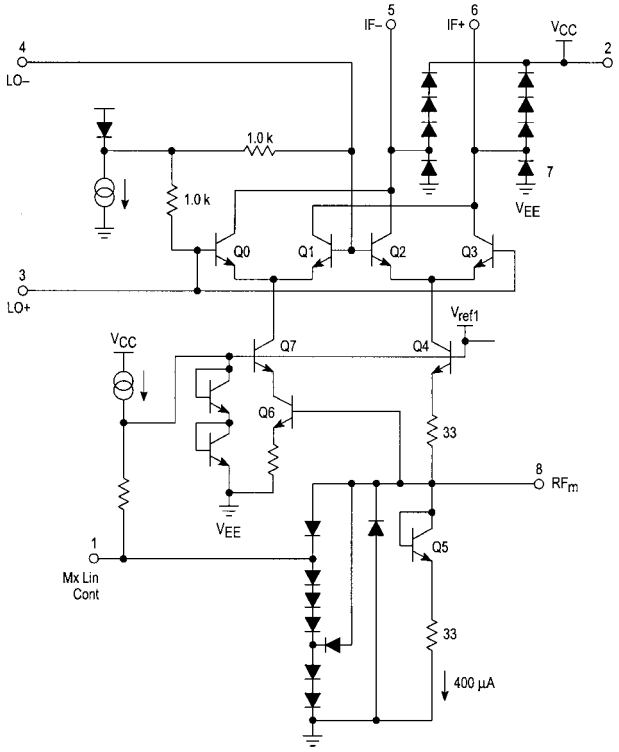
Mixer

The mixer is a unique and patented double–balanced four quadrant multiplier biased class AB allowing for programmable linearity control via an external current source. An input third order intercept point of 20 dBm may be achieved. All 3 ports of the mixer are designed to work up to 2.4 GHz. The mixer has a 50 Ω single–ended RF input and open collector differential IF outputs (see Internal Circuit Schematic for details). The linear gain of the mixer is approximately –5.0 dB with a SSB noise figure of 12 dB.

Local Oscillator

The local oscillator has differential input configuration that requires typically –10 dBm input from an external source to achieve the optimal mixer gain.

Figure 8. MC13143 Internal Circuit*



NOTE: * The MC13143 uses a unique and patented circuit topology.

MC13143

APPLICATIONS INFORMATION

Evaluation PC Board

The evaluation PCB is very versatile and is intended to be used across the entire useful frequency range of this device. The PC board is laid out to accommodate all SMT components on the circuit side (see Circuit Side Component Placement View).

Component Selection

The evaluation PC board is designed to accommodate specific components, while also being versatile enough to use components from various manufacturers. The circuit side placement view is illustrated for the components specified in the application circuit. The Component Placement View specifies particular components that were used to achieve the results shown in the typical curves and tables.

Mixer Input

The mixer input impedance is broadband 50 Ω for applications up to 2.4 GHz. It easily interfaces with a RF ceramic filter as shown in the application schematic.

Mixer Linearity Control

The mixer linearity control circuit accepts approximately 0 to 2.3 mA control current. An Input Third Order Intercept Point, IIP3 of 20 dBm may be achieved at 2.3 mA of control current (approximately 7.0 mA of additional supply current).

Local Oscillator Inputs

The differential LO inputs are internally biased at $V_{CC} - 1.0\ V_{BE}$; this is suitable for high voltage and high gain operation.

For low voltage operation, the inputs are taken to V_{CC} through 51 Ω.

IF Output

The IF is a differential open collector configuration which is designed to use over a wide frequency range for up conversion as well as down conversion.

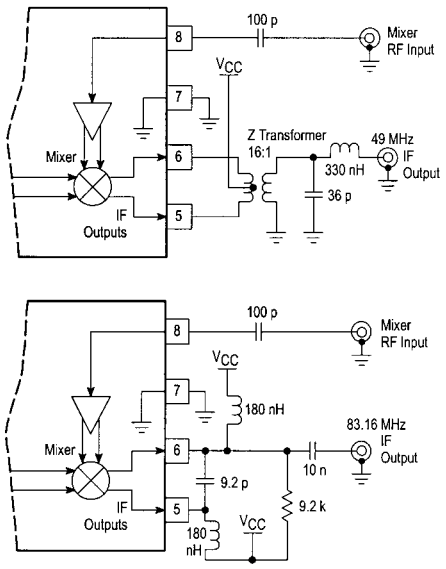
Input/Output Matching

It is desirable to use a RF ceramic or SAW filter before the mixer to provide image frequency rejection. The filter is selected based on cost, size and performance tradeoffs. Typical RF filters have 3.0 to 5.0 dB insertion loss. The PC board layout accommodates both ceramic and SAW RF filters which are offered by various suppliers such as Siemens, Toko and Murata.

Interface matching between the RF input, RF filter and the mixer will be required. The interface matching networks shown in the application circuit are designed for 50 Ω interfaces.

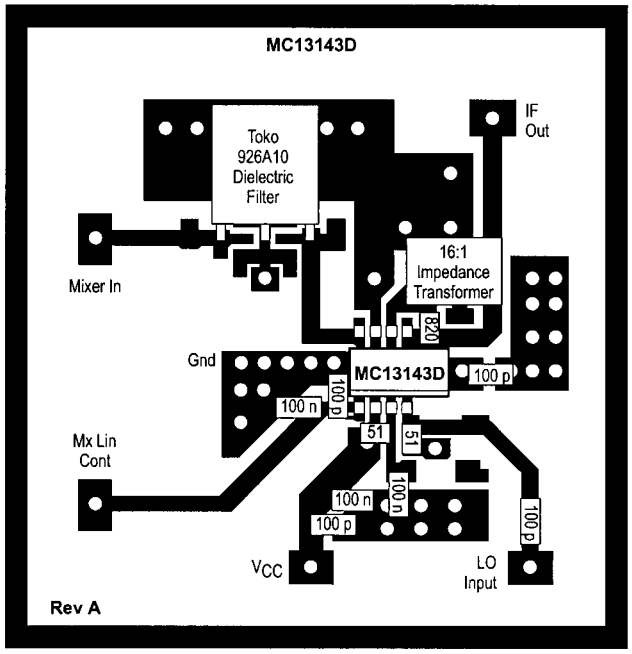
Differential to single–ended circuit configuration is shown in the test circuit. 6.0 dB of additional mixer gain can be achieved by conjugately matching the output of the MiniCircuits transformer to 50 Ω at the desired IF frequency. With narrowband IF output matching the mixer performance is 3.0 dB gain and 12 dB noise figure (see Narrowband 49 and 83 MHz IF Output Matching Options). Typical insertion loss of the Toko ceramic filter is 3.0 dB. Thus, the overall gain of the circuit is 0 dB with a 15 dB noise figure.

Figure 9. Narrowband IF Output Matching with 16:1 Z Transformer and LC Network



MC13143

Figure 10. Circuit Side Component Placement View



NOTES: 926.5 MHz preselect dielectric filter is Toko part # 4DFA–926A10; the 4DFA (2 and 3 pole SMD type) filters are available for applications in cellular and GSM, GPS, DECT, PHS, PCS and ISM bands at 902–928 MHz, 1.8–1.9 GHz at 2.4–2.5 GHz.

The PCB also accommodates a surface mount RF SAW filter in an eight or six pin ceramic package for the cellular base and handset frequencies. Recommended manufacturers are Siemens and Murata.

The PCB may also be used without a preselector filter; AC coupled to the mixer as shown in the test circuit schematic. All other external circuit components shown in the PCB layout above are the same as used in the test circuit schematic.

16:1 broadband impedance transformer is mini circuits part #TX16–R3T; it is in the leadless surface mount "TX" package. For a more selective narrowband match, a lowpass filter may be used after the transformer. The PCB is designed to accommodate lump inductors and capacitors in more selective narrowband matching of the mixer differential outputs to a single–ended output at a given IF frequency.

The local oscillator may also be driven in a differential configuration using a coaxial transformer. Recommended sources are the Toko Balun transformers type B4F, B5FL and B5F (SMD component).

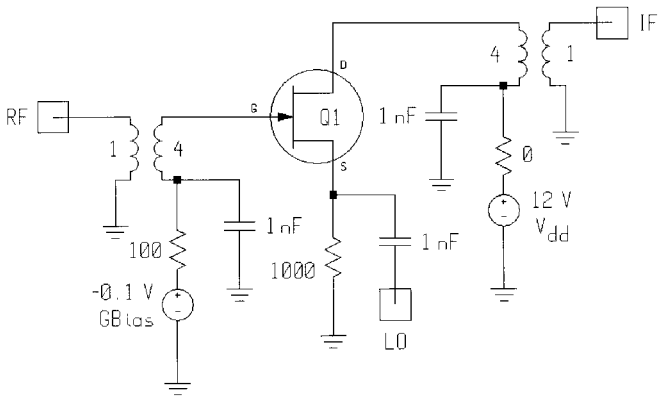


Figure 4-57 An additive JFET mixer cell using gate RF injection and source LO injection.

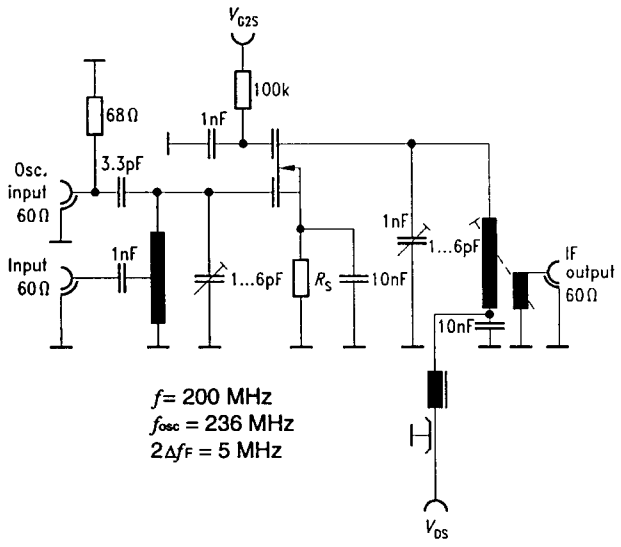


Figure 4-58 Test circuit for additive mixing using a dual-gate MOSFET.

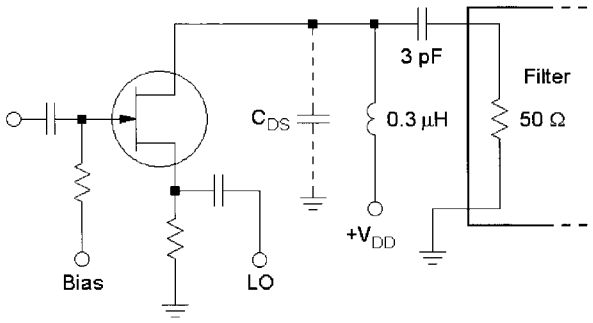


Figure 4-59 FET mixer output coupling that minimizes the effects of filter reactance outside the filter passband.

The active FET mixer achieves gain at the expense of intercept point; the difference can be as much as 20 dB. As an example, the passive MOSFET quad mixer, which is to be referenced, achieves an intercept point of greater than 40 dBm while similar active silicon FETs barely make 30 dBm. This 30-dBm performance is possible only at fairly low frequencies and requires careful selection and adjustment of the dc bias point of all transistors of the circuit. This performance also requires that the FETs be operated grounded-gate, which limits their gain. The other drawback of these active arrangements is that higher impedances at the output reduce the third-order intercept point overproportionally; as an example, a factor of two in impedance change can cost as much as 6 dB in the intercept point. In these discussions we assume that the IMD products are generated solely by the active device, and not by any passive devices, such as transformers. Actually, there are also some capacitors whose linearity varies with RF current—another unpleasant effect that is frequently overlooked.

On the other hand, one can use any FET as a passive device similar to a diode mixer, in which the source–drain channel gets switched on and off. This impedance modulation is somewhat similar to a diode mixer, but the gate electrode is isolated from both source and drain. It nonetheless falls into the category of additive mixers because there is sufficient interaction between gate and source, although the impedance at the gate changes significantly less than in an additive diode mixer. Implementation is a challenge in that building a high-performance passive FET mixer requires a pair or a quad of mixer cells that are sufficiently matched to suppress even-order IMD products. Intercept points, depending on the LO drive, vary between +20 and +45 dBm. The lower number is more applicable for microwave and RF frequencies, while the +45-dBm level is easily obtainable between and 5 and 30 MHz. It is necessary to remember that the FET is a voltage-driven switch, and LO matching becomes an issue: How do you generate 30 V peak-to-peak across a few picofarads over a wide frequency range? Physical layout can also be critical: Symmetry is not only important for the active devices, but also for the input, output, and LO circuit. We have seen cases where fractions of an inch in different lengths leading to the gate electrodes have cost up to 10 dB of IP_3—just because of the resulting lack of symmetry.

Figure 4-60 shows the circuit of Figure 4-57 modified to work with an LDMOS FET, the harmonic-balance version of the SPICE level 3 model. Figure 4-61 shows the LDMOS circuit's conversion gain and noise figure versus frequency. The "noisy" traces reflect the model's numerical problems. This model, which works up to about 0.8 μm and has been validated against the Motorola data and SPICE results, has already been made mathematically continuous but still shows the inability to properly model the real RF transistor. This is why more research needs to be done on RF and microwave applications for MOS in the modeling area. As can be seen from the references, several attempts have been made to improve the model, but the major headache remains with parameter extraction, and as the models get more complex, the modeling quality doesn't seem to improve as a linear function. The reader should be reminded of the load-pull measurement as outlined earlier.

Dual-Gate MOS/GaAs Mixers. Significantly better LO–RF isolation can be obtained by applying the RF to LO signals at separate gates of a dual-gate FET (Figure 4-62). The result is multiplicative mixing with fewer constraints on LO and RF filtering.

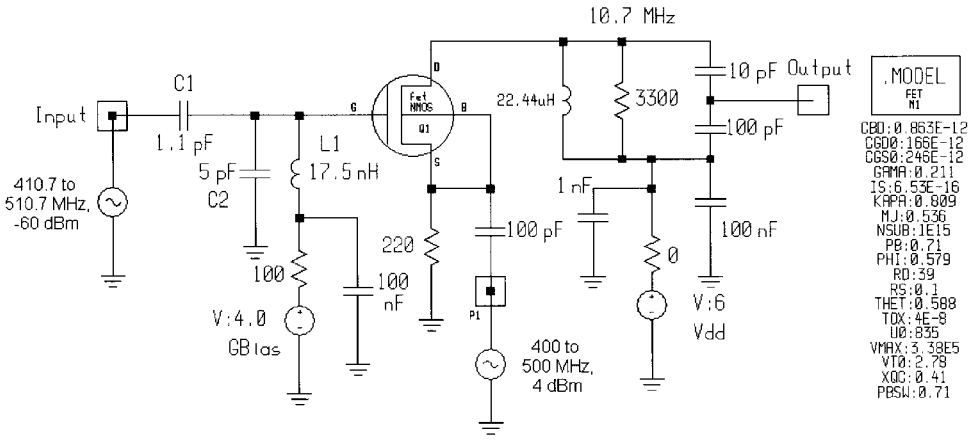


Figure 4-60 The N-JFET circuit of Figure 4-57 modified to work with an LDMOS FET.

4-4-4 MOSFET Gilbert Cell

The trend to go to smaller voltages has created some CMOS implementations of the Gilbert cell. While this may allow us to integrate reasonable mixers on the same chip, nobody should expect any miracles from these mixers. They are frequently starved in operating voltage and current and rarely fare better than a single diode mixer with proper drive level. On the other hand, the symmetry reduces some of the unwanted spurious frequencies and because of the high impedance, the actual RF power level is much less than the diode would need. There are so

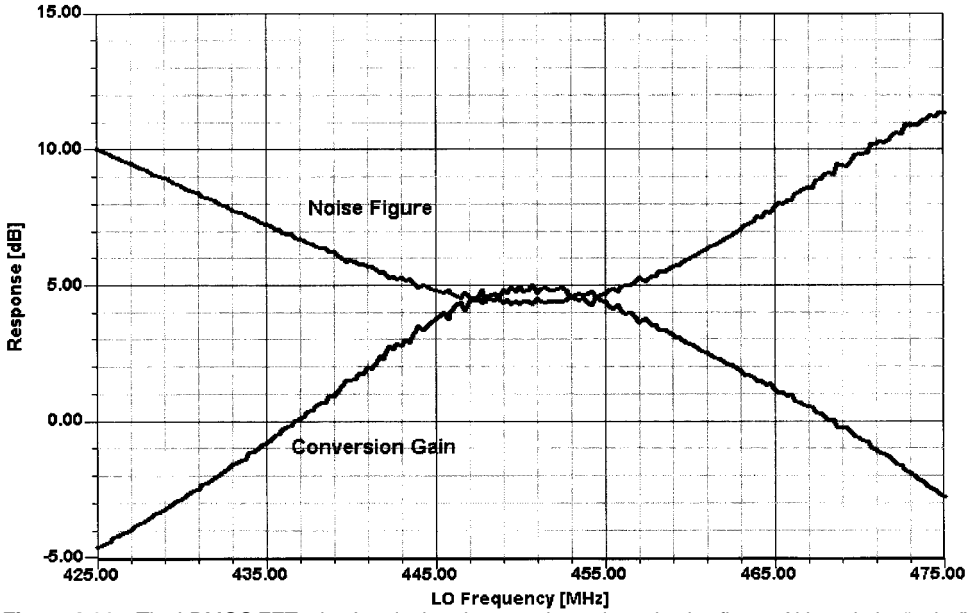


Figure 4-61 The LDMOS FET mixer's calculated conversion gain and noise figure. Although the "noisy" traces reflect numerical problems in the device model, the results can be put to good use in circuit design.

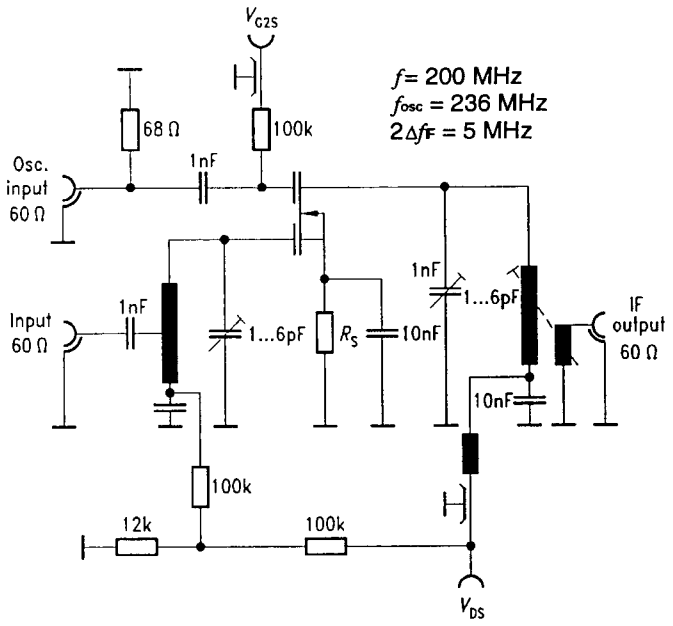


Figure 4-62 Test circuit for multiplicative mixing using a dual-gate MOSFET.

many CMOS transistors available in different processes that we cannot provide information about the general behavior, but this "low-cost" solution, as pointed out, is really only attractive because one can stay in the same technology while building the amplifier and other stages. Because of the flicker corner frequency of MOS, the noise figure of the mixer will be high compared to other devices, but not as bad as its GaAsFET brothers. In any case, to combine performance and simplicity, one needs to have a preamplifier that reduces the noise figure as well as the third-order intercept point of the mixer by the amount of preamplification.

4-4-5 GaAsFET Single-Gate Switch

In general, GaAsFET technology is the most expensive one. While GaAsFET competes with SiGe technology in frequency range at the time of this writing, it is a mature technology and well understood. The gallium arsenide wafers, however, are still unable to compete on a price basis with other technologies. There is a hybrid technology, silicon on sapphire, that also tries to go after the GaAsFET domain. The cutoff frequency of GaAsFET transistors is significantly higher than any other FETs, but the corner frequency f_c, also frequently referred to as flicker frequency, is somewhere between 10 and 100 MHz and therefore results in poor mixer performance from a noise figure point of view. The same applies, by the way, to oscillators. Another headache with GaAs is the low Q of the material, but this low Q gives transformers on GaAs a wider bandwidth. One of the pioneers of ICs for the wireless market is TriQuint, and we have been working with Wes Hayward of TriQuint in circuit analysis and validation. In doing so, we used some commercial and experimental circuits. In Chapter 3 we have already pointed out one amplifier that shows good agreement between measured and predicted data. By the time one has realized the entire circuit, including its associated biasing problems, even these circuits tend to grow. Figure 4-63 shows the circuit, which consists of an oscillator/amplifier, switching mixer, and a differential IF postamplifier.

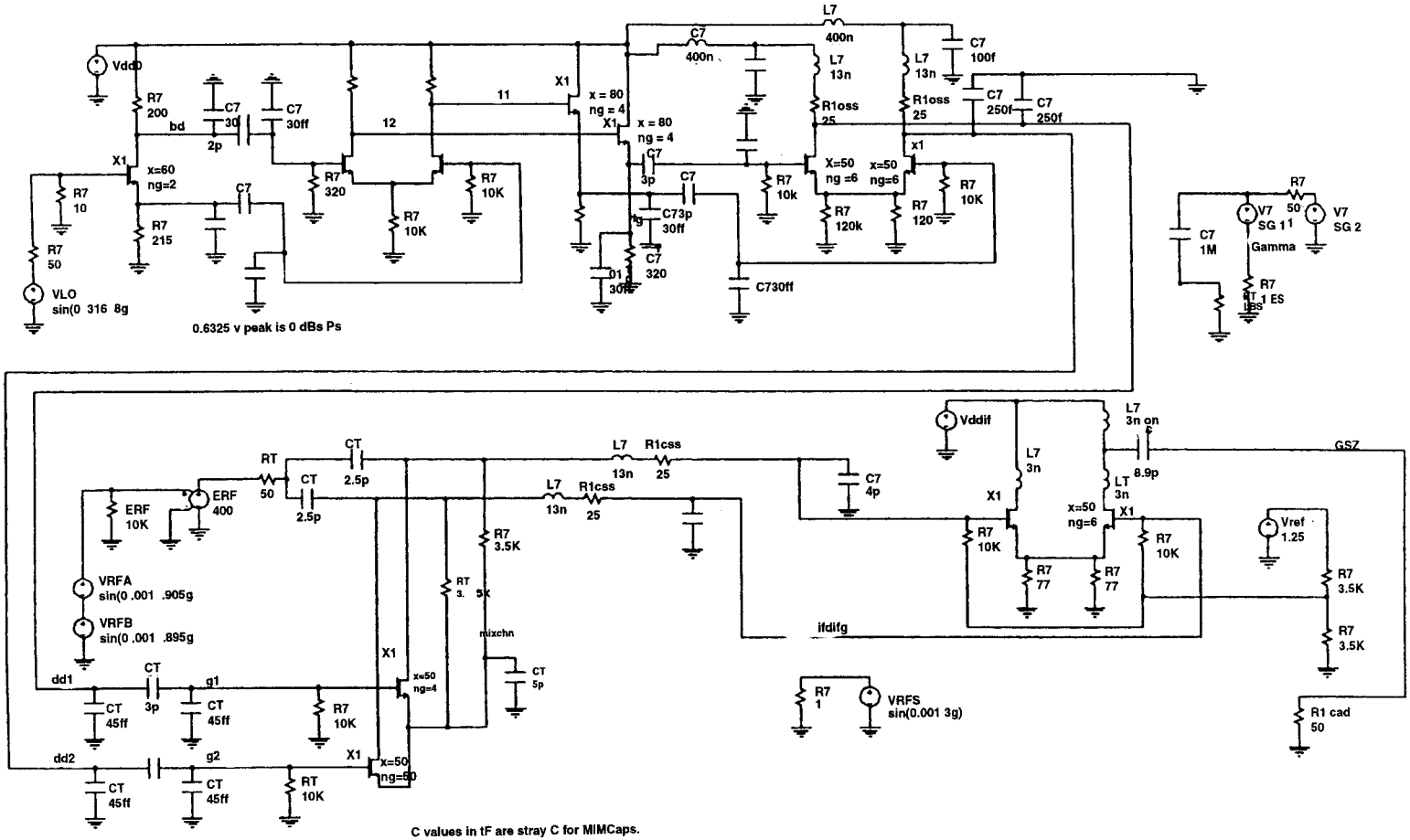
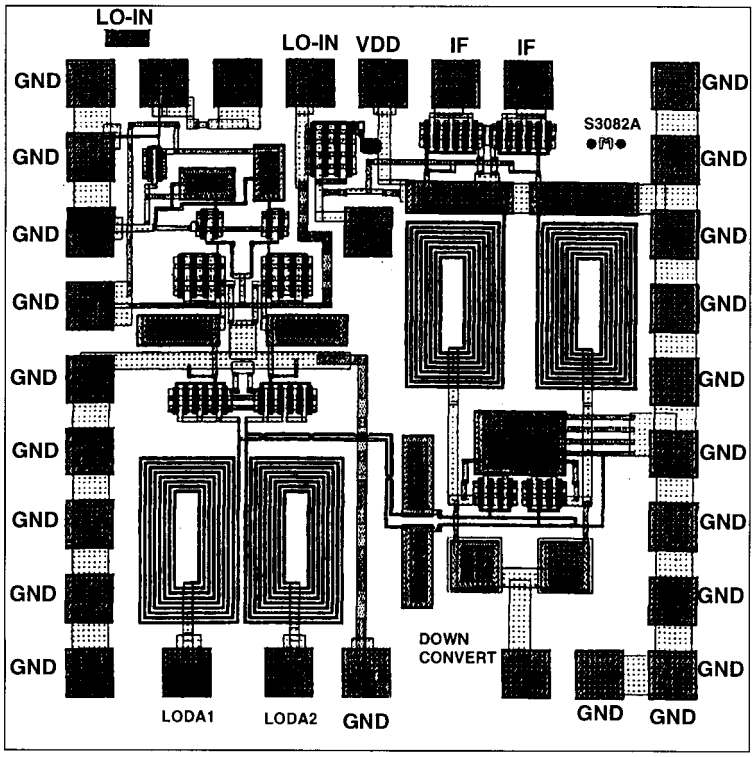


Figure 4-63 The complete switching mixer circuit.

Since the highest accuracy is needed to model these types of circuit, we found that only very recent simulators, such as Serenade 8 from Ansoft, can handle such a large number of FETs without barking at the user. Initially, we had all kinds of difficulties with these circuits because of the TriQuint TOM model, which had been developed for SPICE, and its IF/ELSE statements that made it mathematically discontinuous. The type of GaAsFETs used by TriQuint were both regular FETs and enhancement FETs, which require a positive voltage at the gate. Further details of this should be looked up in the foundry manual by TriQuint. In Chapter 2 we have shown some comparisons between different models and their prediction accuracy. Such converters need to resort to inductors in GaAs technology, and compared to other parts they become quite big (Figure 4-64).

In order to reduce some of the complexity in modeling and yet deal with the problems of mixers, a simpler mixer is needed that is a combination of a set of switching transistors and a postamplifier for the IF (Figure 4-65). The reader may notice that the small inductances of 15 nH have an extremely high loss modeled by a series resistor of 18 Ω. The first to recommend this type of switching mixers was probably Stephen Maas. The circuit diagram itself (Figure 4-65) is self-explanatory, but again we need to point out the resulting complexity.

The GaAsFET mixer giving the best performance on the market today is the Siemens model CMY210 (Figure 4-66). A reproduction of its datasheet, used by permission, follows the schematic.


LO-IN
LO-IN
VDD
IF
IF
S3082A
GND
GND
GND

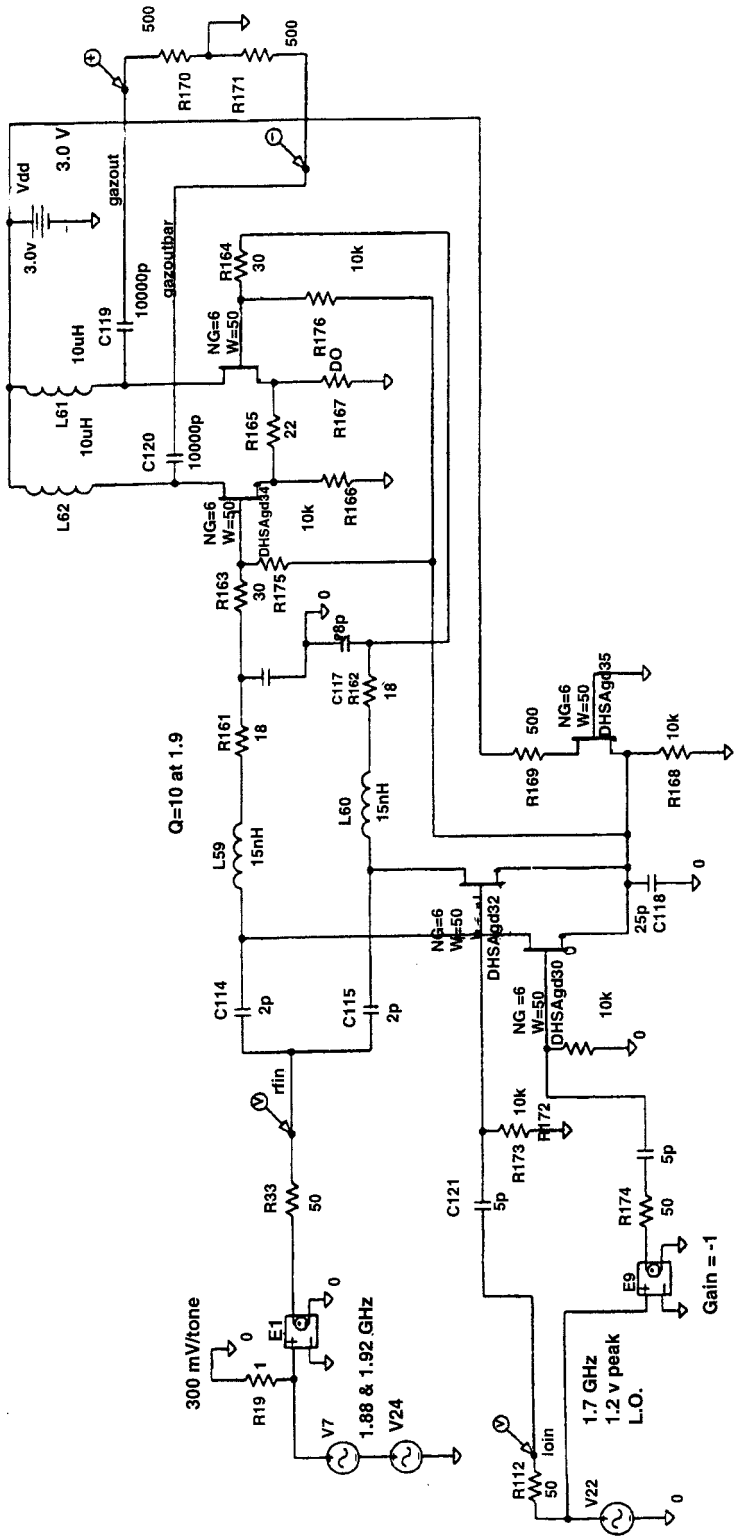


Figure 4-65 Simplified switching FET mixer.

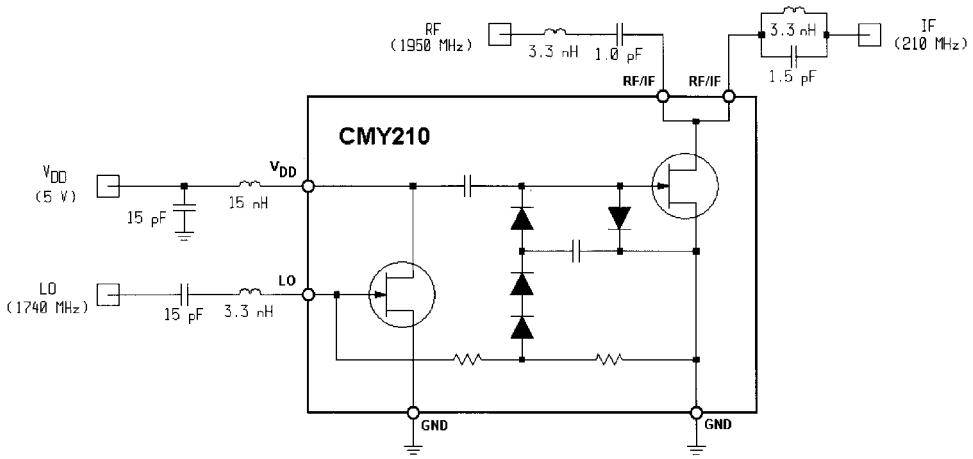


Figure 4-66 The Siemens CMY210 IC includes a shunt switch mixer and AGC-equipped LO amplifier for more constant LO drive. Despite its simplicity, this mixer achieves a typical $IP_{3,in}$ of 25 dBm.

The CMY210 circuit is a clever combination of a switching type of mixer with an automatic level control arrangement that provides the mixer with a reasonably constant drive level following the LO amplifier. Similar to cases we have shown, there is a clear relationship between the local oscillator power (invested dc power), and large-signal performance. This mixer externally has a series-tuned circuit at the input, which allows the UHF input signal

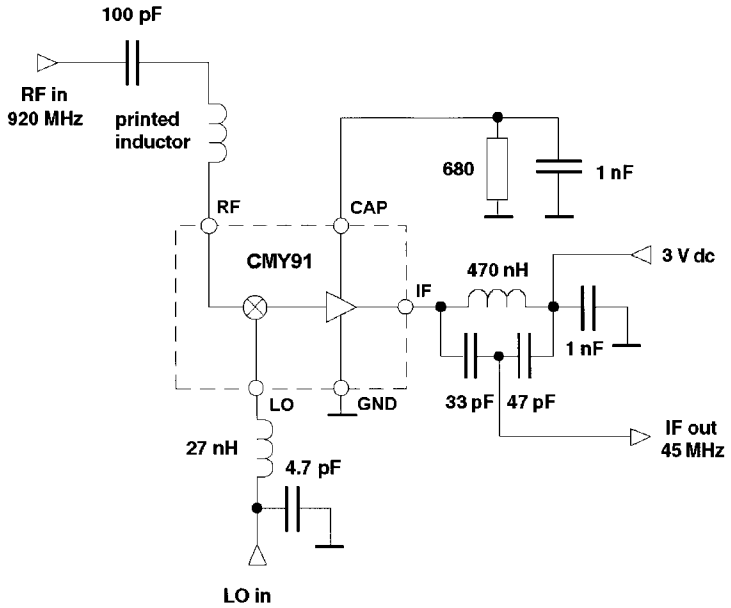


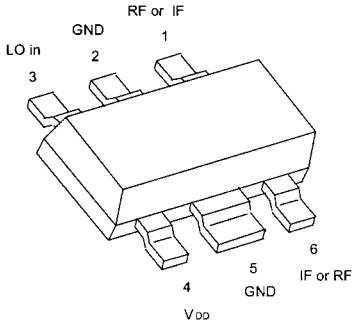
Figure 4-67 The CMY91 IC includes a postmixer amplifier instead of an AGC LO amplifier. Its $IP_{3,in}$ is typically –2 dBm (drain supply, 3 V at 1 mA) with no connection made to its CAP pin. In the configuration shown, its typical $IP_{3,in}$ is 0 dBm (drain supply, 3 V at 2.5 mA).

SIEMENS GaAs MMIC CMY 210

Preliminary Data

* **Ultralinear Mixer** with integrated LO-Buffer
* Very high **Input-IP3 of typical 25dBm**
* Very low **LO-Power demand of typ. 0dBm**
* Suited for Up- and Down-Conversion
* Wide LO-Frequency Range
  <500MHz to >2,5GHz
* Wide LO-Level Range
* Single ended Ports
* RF- and IF-Port Impedance 50 Ohm
* Operating Voltage Range: < 3 to 6V
* Very low Current Consumption of typical 7mA
* All Gold Metallisation

ESD: **E**lectro**s**tatic **d**ischarge sensitive device
Observe handling Precautions!



Type	Marking	Ordering code (tape and reel)	Package [1]
CMY210	**M3**	**Q62702 M 0016**	**MW-6**

Maximum Ratings	Port	Symbol	Value		Unit
			min	max	
Supply Voltage	4	V_{DD}	0	6	V
DC-Voltage at LO Input	3	V_3	-3	0,5	V
DC-Voltage at.RF-IF Ports [2]	1, 6	$V_{1,6}$	- 0,5	+ 0,5	V
Power into RF-IF Ports	1, 6	$P_{in,RF}$		10	dBm
Power into LO Input	3	$P_{in,LO}$		10	dBm
Channel Temperature		T_{Ch}		150	°C
Storage Temperature		T_{stg}	-55	150	°C
Thermal Resistance					
Channel to Soldering Point (GND)		R_{thChS}	≤100		K/W

1) For detailed dimensions see chapter Package Outlines
2) For DC test purposes only, no DC voltages at pins 1, 6 in application

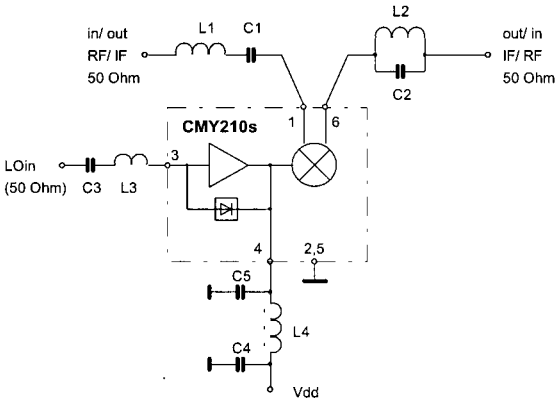
SIEMENS GaAs MMIC **CMY 210**

Electrical Characteristics

Test conditions: T_a = 25°C; V_{DD}= 3V, see test circuit; f_{RF} = 1620MHz; f_{LO} = 1500MHz; P_{LO} = 0dBm; f_{IF}= 120MHz, unless otherwise specified:

Parameter, Test Conditions	Symbol	min	typ	max	Unit
Operating Current	I_{op}	-	7	-	mA
Conversion Loss	L_c	-	5,5	-	dB
SSB Noise Figure	F_{ssb}	-	5,5	-	dB
2 Tone 3rd Order IMD $P_{RF1} = P_{RF2}$ = -3dBm f_{RF1} = 1615MHz; f_{RF2} = 1620MHz; f_{LO}=1500MHz	d_{IM3}	-	56	-	dBc
3rd Order Input Intercept Point	$IP3_{in}$	-	25	-	dBm
P_{-1dB} Input Power	P_{-1dB}	-	20	-	dBm
LO Leakage at RF/IF-Port (1,6)	$P_{LO\ 1,6}$	-	-6	-	dBm

Test circuit / application example



Notes for external elements:

L1, C1: Filter for upper frequency;
C2, L2: Filter for lower frequency;
each filter is a throughpath for the desired frequency (RF or IF) and isolates the other frequency (IF or RF) and its harmonics.
These two filters must be connected to pin 1 and pin 6 directly.
Parasitic capacitances at the ports 1 and 6 must be as small as possible.
L4 and C5 are optimized by indicating lowest I_{op} at used LO-frequency; same procedure for L3.
The ports 1, 3 and 6 must be DC open.

f *LO*	L1	C1	L2	C2	L3	C3	L4	C4	C5
MHz	nH	pF	nH	pF	nH	pF	nH	pF	pF
500	11	8.2	19	4.7	6	47	19	47	3.3
1000	7	4.7	9	3.3	6	33	16	33	0
1500	4	2.2	7,5	2.2	6	18	8.5	18	0
2000	*)	2.2	*)	1.5	4	15	*)	15	0
2500	*)	2.2	*)	1.5	3	15	*)	15	0
notes next pg	1)	1)	1)	1)	3)	3)	2)	2)	2)

SIEMENS GaAs MMIC **CMY 210**

Approximate values of used elements for down- (or up-) conversion with a lower frequency of 10MHz to 120MHz (IF or RF; in or out); parasitics will cause deviations; therefore exact values will be defined by application, especially for *).

General description and notes:

The CMY 210 is an all port single ended general purpose Up- and Down-Converter.
It combines small conversion losses and excellent intermodulation characteristics with a low demand of LO- and DC-power.
The internal level controlled LO-Buffer enables a good performance over a wide LO level range.
The internal mixers principle with one port RF and IF requires a frequency separation at pin 1 and 6 respectively.

Note 1:
Best performance with lowest conversion loss is achieved when each circuit or device for the frequency separation meets the following requirements:

Input Filter: Throughpass for the signal to be mixed; reflection of the mixed signal and the harmonics of both.

Output Filter: Throughpass for the mixed signal and reflection of the signal to be mixed and the harmonics of both.

The impedance for the reflecting frequency range of each filter toward the ports 1 and 6 should be as high as possible.
In the simplest case a series- and a parallel- resonator circuit will meet these requirements but also others as appropriate drop in filters or micro stripline elements can be used.
The two branches with filters should meet immediately at the package leads of the port 1 and 6.
Parasitic capacitances at these ports must be kept as small as possible.
The mixer also can be driven with a source- and a load impedance different to 50Ω, but performance will degrade at larger deviations.

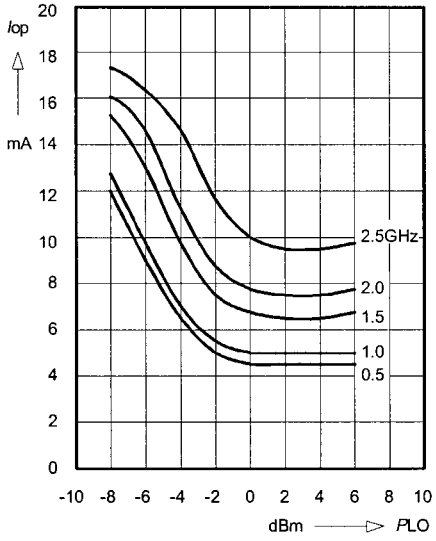
Note 2:
The LO-Buffer needs an external inductor L4 at port 4; the value of inductance depends on the LO frequency. It is tuned for minimum I_{op} consumption into port 4.
At lower LO frequencies it can be reduced by an additonal capacitor C5.

Note 3:
The LO Input impedance at Port 3 can be matched with a series inductor. It also can be tuned for a minimum current I_{op} into port 4. C3 is a DC blocking capacitor.
Since the input impedance of port 3 can be slightly negative at lower frequencies, the source reflection coefficient should be kept below 0.8 ($Z_0 = 50\ \Omega$) within this frequency range.

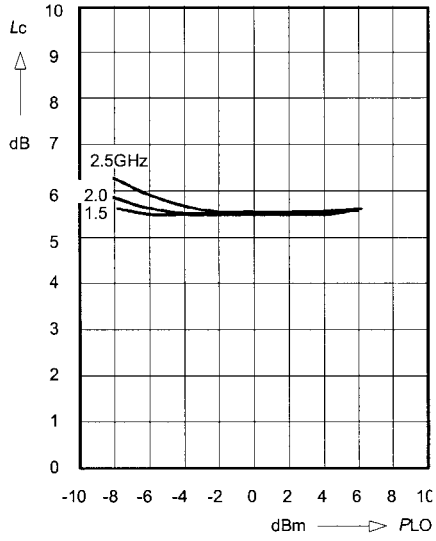
The Conversion Noise Figure Fssb is corresponding with the value of Conversion Loss L_c. The LO signal must be clean of noise and spurious at the frequencies $f_{LO} \pm f_{IF}$.

SIEMENS GaAs MMIC **CMY 210**

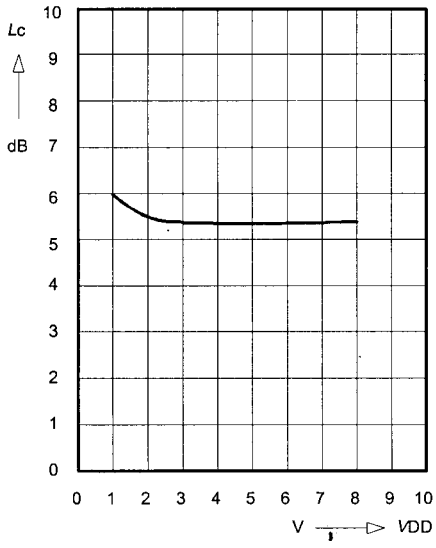
Operating Current $I_{op} = f(P_{LO})$
$V_{DD} = 3V$
f_{LO} = Parameter



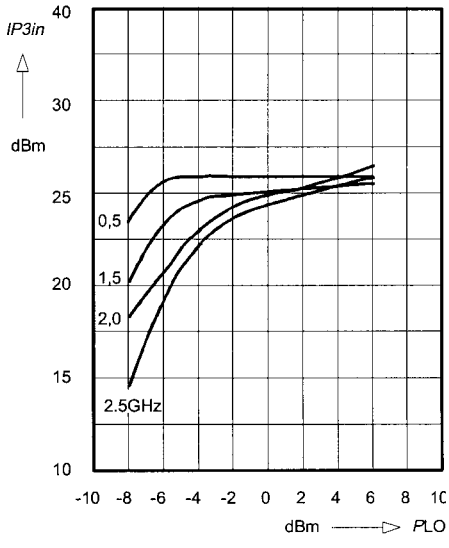
Conversion Loss $L_C = f(P_{LO})$
$V_{DD} = 3V$; $f_{IF} = 120MHz$
f_{LO} = Parameter



Conversion Loss $L_C = f(V_{DD})$
$P_{LO} = 0dBm$
$f_{LO} = 1500MHz$; $f_{IF} = 120MHz$

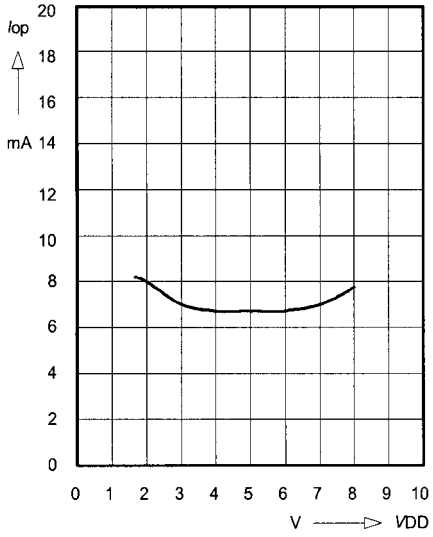


Third Order IP3 $IP3in = f(P_{LO})$
$P_{in} = 2$ x -3dBm; $f_{IF} = 40/45MHz$
$V_{DD} = 3V$; f_{LO} = Parameter;

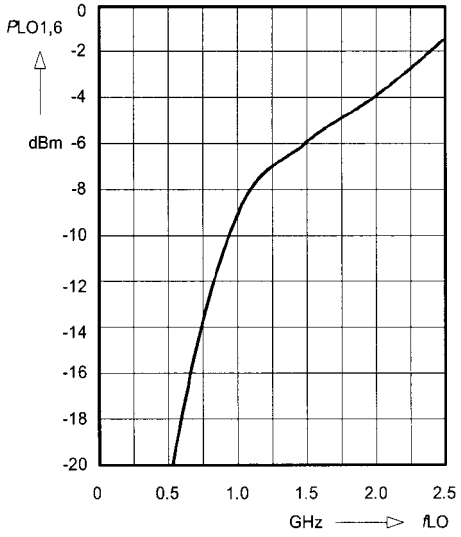


SIEMENS GaAs MMIC **CMY 210**

Operating Current $I_{op} = f(V_{DD})$
P_{LO} = 0dBm
f_{LO} = 1500MHz



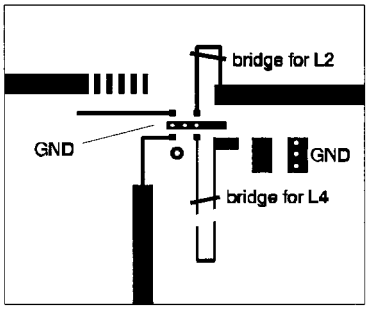
LO-Leakage at Port 1, 6 $P_{LO1,6} = f(f_{LO})$
P_{LO} = 0dBm
V_{DD} = 3V



SIEMENS GaAs MMIC **CMY 210**

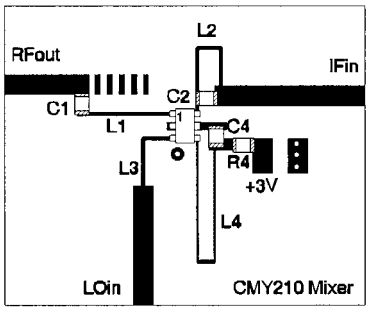
Additional informations and an example of a general purpose mixer pcb:

This general purpuse mixer demonstration board is used to show the performance of the CMY 210. The hints below will be helpful to achieve good intermodulation behaviour.



Material: Epoxy; size 24mm x 30mm; thickness 1mm

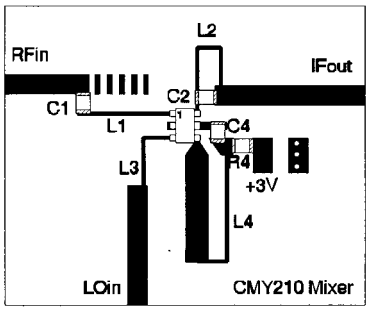
A good ground connection for CMY210 is necessary; here it is realized by 3 vias **under** the CMY 210 device. Best IM3 performance is obtained, when the capacitor C4 is grounded at the upper CMY 210 ground line without any additional vias (so the RF-Signal is encoupled from the LO-buffer best possible). Frequency tuning is done by selecting suitable capacitors C1 and C2, positioning C1 along the L1-line and by reducing the inductance of L2 and L4 by a bridge; unnecessary lines should be disconnected.



This example showes an up-converter.

The IM3 performance of upconverters mostly can be improved by tuning L4 slightly smaller then required for a minimum current consumption.

In other words, the resonant frequency of the buffer circuit at PIN4 (internal capacitor and L4) is tuned by L4 to a frequency slightly above the L.O.-frequency.



This example showes a down-converter.

Here an improved IM3-performance mostly can be obtained by a more capacitive load at CMY 210 port pin 4.
Here it is realized by a broader part of the L4 inductor line toward to port 4 and tuning L4 to a value sligthly smaller then for a minimum current consumption into port 4.

In other words, the resonant frequency of the buffer circuit at PIN4 (internal capacitor and L4) is tuned by L4 to a frequency slightly above the L.O.-frequency.

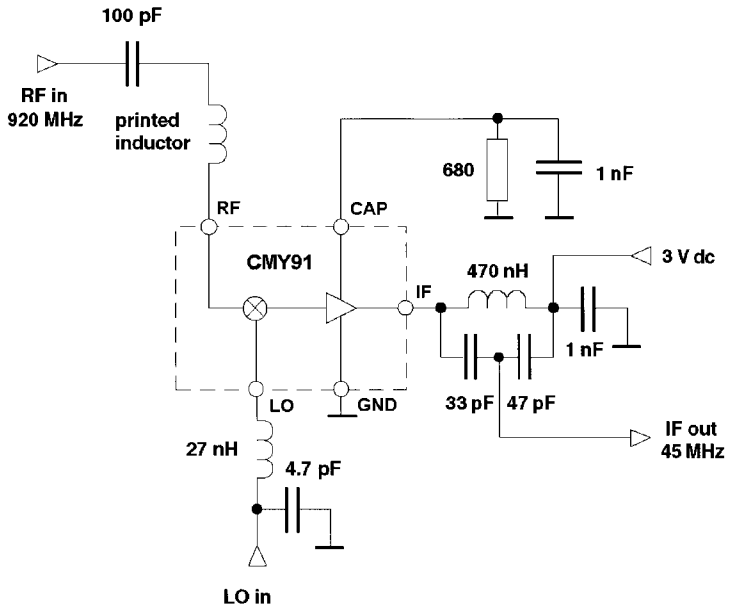


Figure 4-67 The CMY91 IC includes a postmixer amplifier instead of an AGC LO amplifier. Its $IP_{3,in}$ is typically –2 dBm (drain supply, 3 V at 1 mA) with no connection made to its CAP pin. In the configuration shown, its typical $IP_{3,in}$ is 0 dBm (drain supply, 3 V at 2.5 mA).

Table 4-6 Partial CMY91 specifications

Parameters		Symbol	Minimum	Typical	Maximum	Unit
Drain–source breakdown voltage						
I_{IF} = 500 µA	$V_{LO\text{-}GND}$ = 0 V					
$V_{RF\text{-}GND}$ = 4 V	CAP pin not connected	$V_{(BR)IF\text{-}GND}$	8	—	—	V
Drain current						
$V_{RF\text{-}GND}$ = 0 V	$V_{LO\text{-}GND}$ = 0 V					
$V_{IF\text{-}GND}$ = 3 V	CAP pin not connected	I_D	0.8	1	1.4	mA
Conversion gain						
f_{RF} = 920 MHz	f_{LO} = 965 MHz					
f_{IF} = 45 MHz	P_{LO} = –3 dBm	G_c	—	5.5	—	dB
Single-sideband noise figure						
f_{RF} = 920 MHz	f_{LO} = 965 MHz					
f_{IF} = 45 MHz	P_{LO} = –3 dBm	F_{SSB}	—	9	—	dB
Third-order intermodulation						
f_{RF} = 920 MHz	f_{LO} = 965 MHz					
f_{IF} = 45 MHz	P_{LO} = –3 dBm	IP_3	—	–2	—	dBm
LO/RF isolation						
f = 965 MHz		$Iso_{LO/RF}$	—	11	—	dB

T_A = 25 °C/V_D = 3 V; CAP pin connected to ground by 680-Ω resistor.

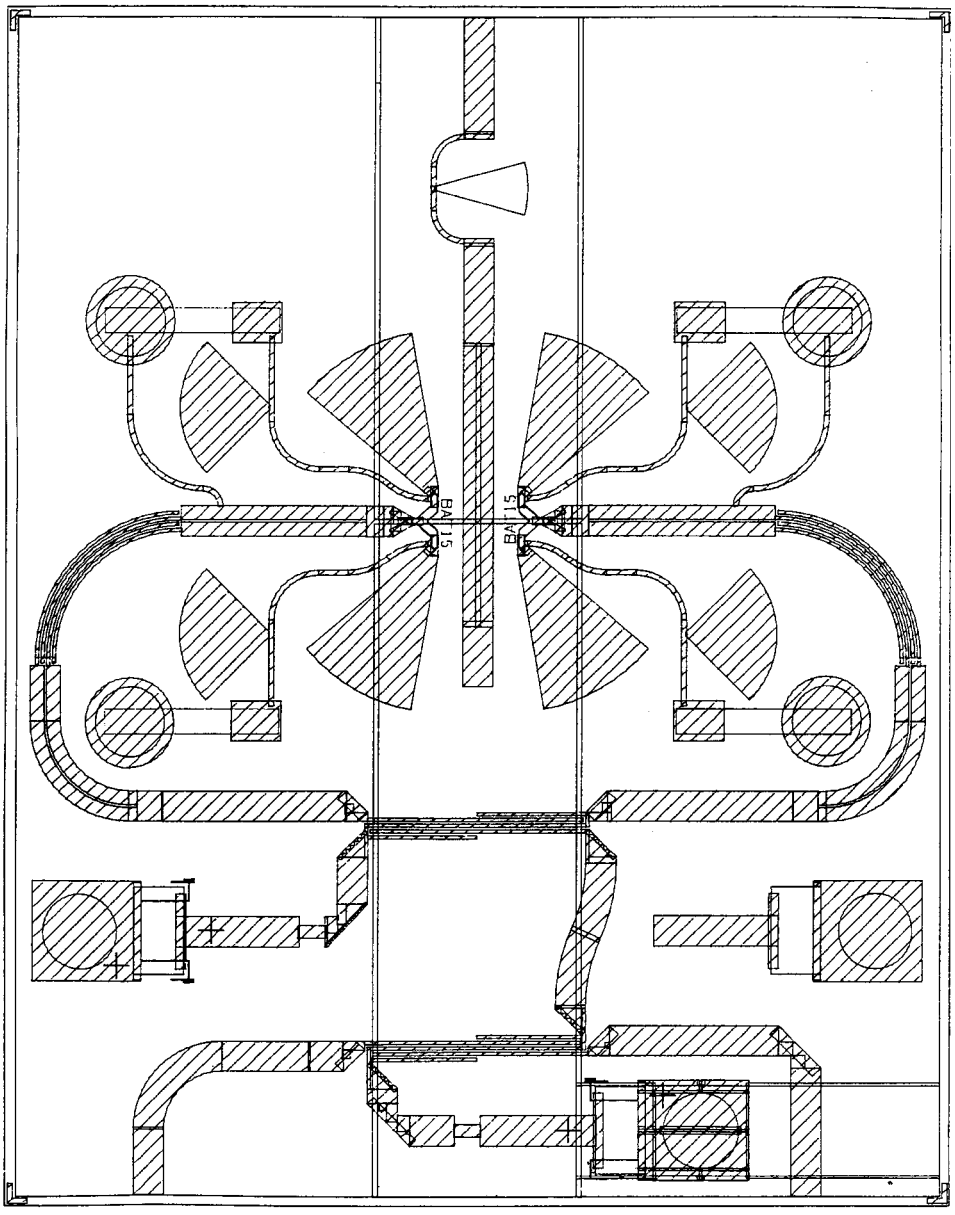


Figure 4-68 Layout of the 20-GHz diode mixer.

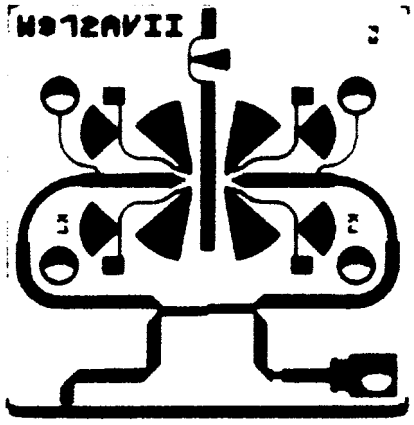


Figure 4-69 The 20-GHz diode mixer circuit board.

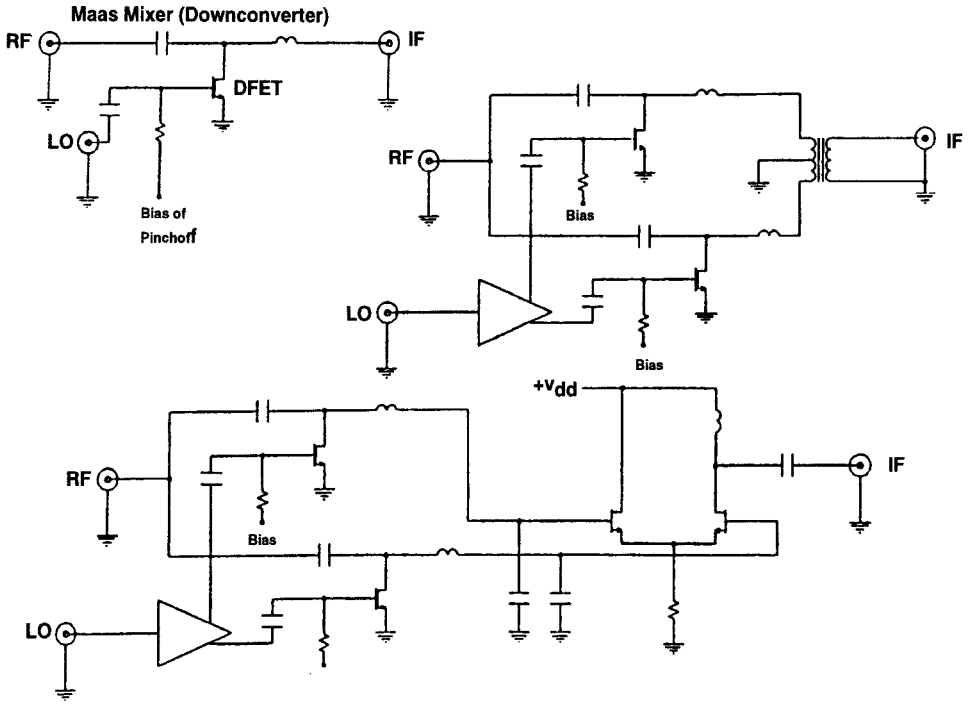


Figure 4-70a

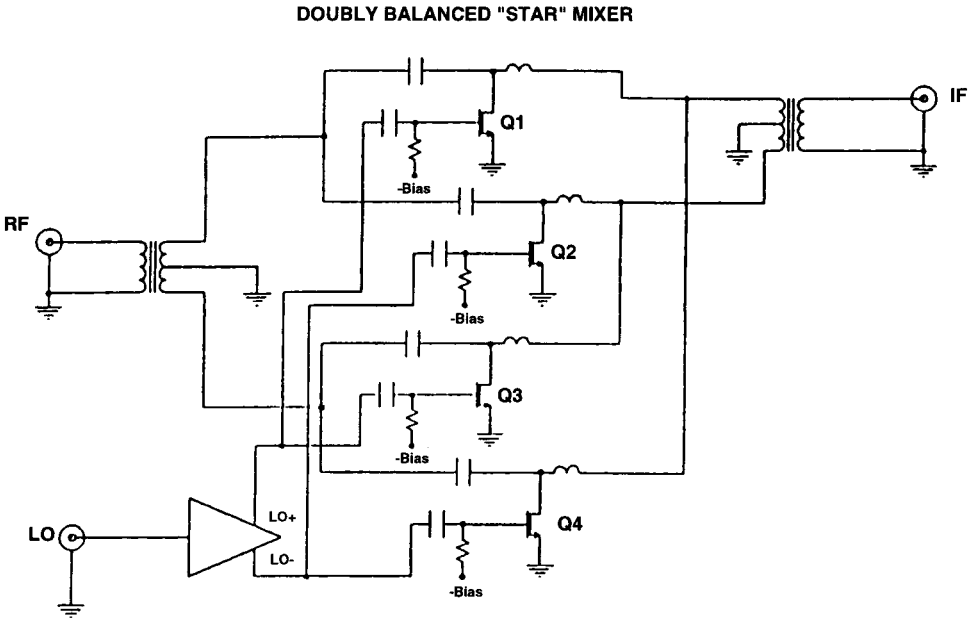


Figure 4-70b

to pass and has a parallel tuned circuit that keeps the input signal from getting into the IF stage. The resulting conversion loss of 5.5 dB is consistent with the example we presented before of 6 dB, and while the noise figure also is stated as 5.5 dB, that value is deceiving because it was determined using a 120-MHz IF. If an IF of 10 MHz or less was used, the noise figure would be tremendously worse. This is due to the frequently mentioned flicker-noise contribution. As a comparison, Figure 4-67 shows the predecessor to the CMY210, the CMY91, which includes a built-in IF amplifier. Table 4-6 shows its partial specifications. Since it has 5.5-dB gain, one can assume that the IF amplifier has something like 10 dB of gain, and the mixer portion, 6 dB losses. The resulting noise figure is the sum of the mixer noise figure and the IF amplifier noise figure. The IF transistor is probably not "noise matched" with the mixer, but the resulting 9-dB noise figure is not uncommon in this arrangement. Judging by previous examples, we would say that the noise figure of the mixer is somewhere around 6 dB, and therefore the mismatched (open at the input) IF amplifier has a noise figure of 3 dB, resulting in this total noise figure of 9 dB as stated. The third-order IP of –2 dBm is no match to the CMY210's $IP_{3,in}$ performance, which is typically 25 dBm.

For readers who are interested in what mixers look like when pushed well into the microwave region, we present the layout for a 20-GHz diode mixer. Figure 4-68 presents its layout, which includes radial stubs, circular stubs, and a Lange coupler. Figure 4-69 shows the layout implemented on a ceramic substrate.

As further food for thought and experimentation, Figure 4-70 presents mixer and related circuit notes as presented by Wes Hayward in TriQuint Semiconductor's GaAs Design Class.

HA Mixer Family Measurements

The mixers are all identical except for FET width. Mixers used the HA process and are half micron gate length. Bias was adjusted to best gain at +6 LO. The LO baluns were half wave lines of semi-rigid coax. F–LO = 1.65, RF = 1.9, and IF = 0.25 GHz. Input Power is –13 or –10 dBm RF. All measurements show downconversion results.

FET Width	P-av RF	P-LO dBm	P-out at IF	IMDR dB	Conv Gain dB	IP3-out dBm	IP3in dBm
50	–13	6	–31	44.7	–11.7	–2.35	9.35
50	–13	10	–30.5	53.5	–11.2	2.55	13.75
50	–10	10	–27.8	49	–11.5	3	14.5
100	–10	6	–25.7	41.5	–9.4	1.35	10.75
100	–10	10	–24.8	48	–8.5	5.5	14.0
100	–10	14	–24.5	68.8	–8.2	16.2	24.4
200	–10	6	–24.3	40.5	–8	2.25	10.25
200	–10	10	–23.7	45	–7.4	5.1	12.5
200	–10	14	–23.5	60.8	–7.2	13.2	20.4
200	–10	17	–24	60.7	–7.7	12.65	20.3
400	–10	6	–24	42.8	–7.7	3.7	11.4
400	–10	10	–23.5	52	–7.2	8.8	16.0
400	–10	14	–23.3	67.7	–7	16.85	23.85
400	–10	17	–23.2	57.5	–6.9	11.85	18.75
50	–10	6	–30.8	56	–14.5	3.5	18.0
50	–10	6	–28.6	42.5	–12.3	–1.05	11.25
50	–10	10	–28	49.2	–11.7	2.9	14.6
50	–10	14	–28.2	60.8	–11.9	8.5	20.4
50	–10	17	–28.3	63.5	–12	9.75	21.75
50	–10	17	–28	62.8	–11.7	9.7	21.4
100	–10	6	–29.5	42.5	–13.2	–1.95	11.25
100	–10	10	–25	47.2	–8.7	4.9	13.6
100	–10	14	–24.7	63.7	–8.4	13.45	21.85
100	–10	16	–25.5	62	–9.2	11.8	21.0

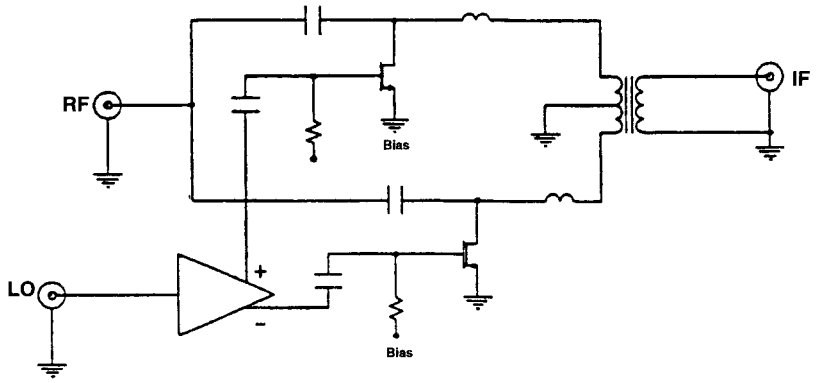
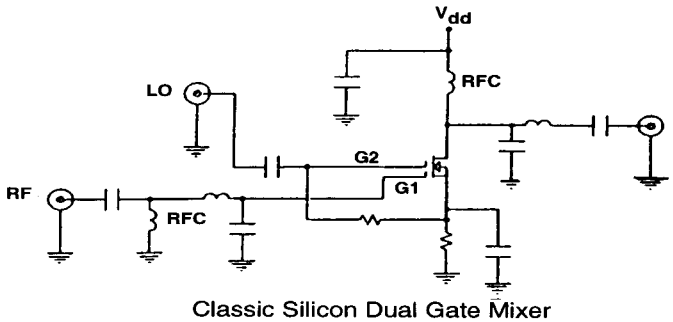


Figure 4-70c



Classic Silicon Dual Gate Mixer

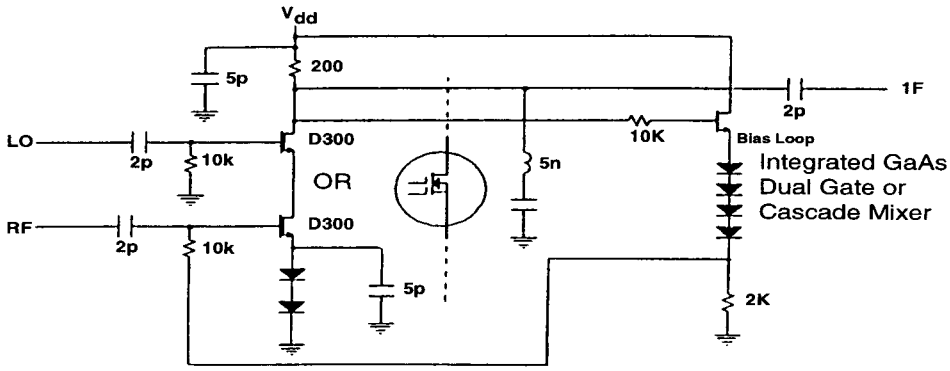


Figure 4-70d

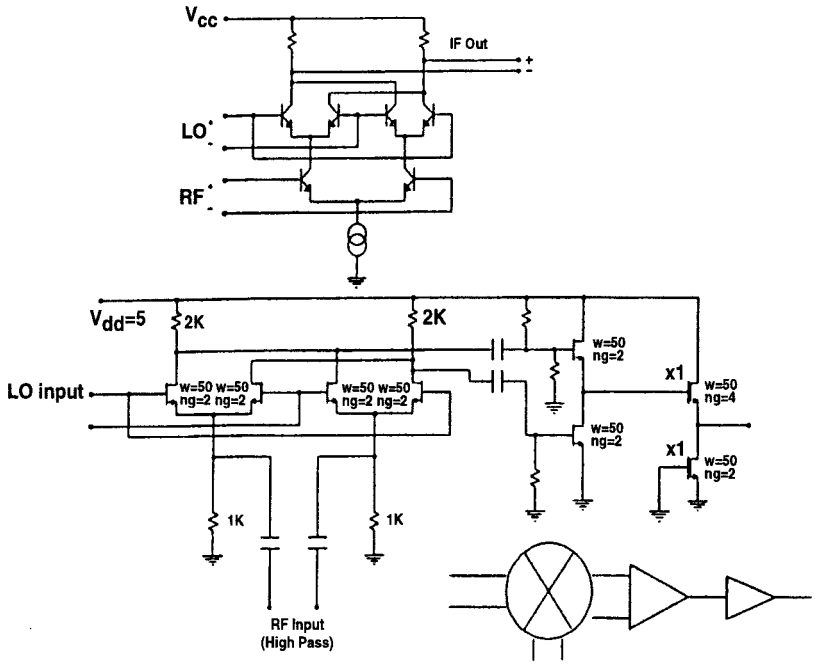


Figure 4-70e

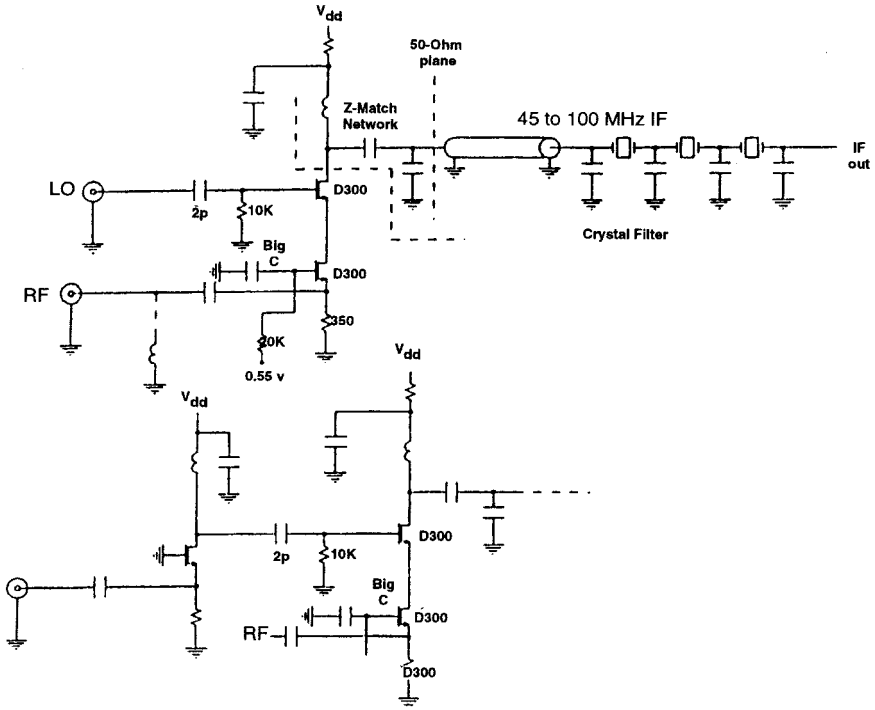


Figure 4-70f

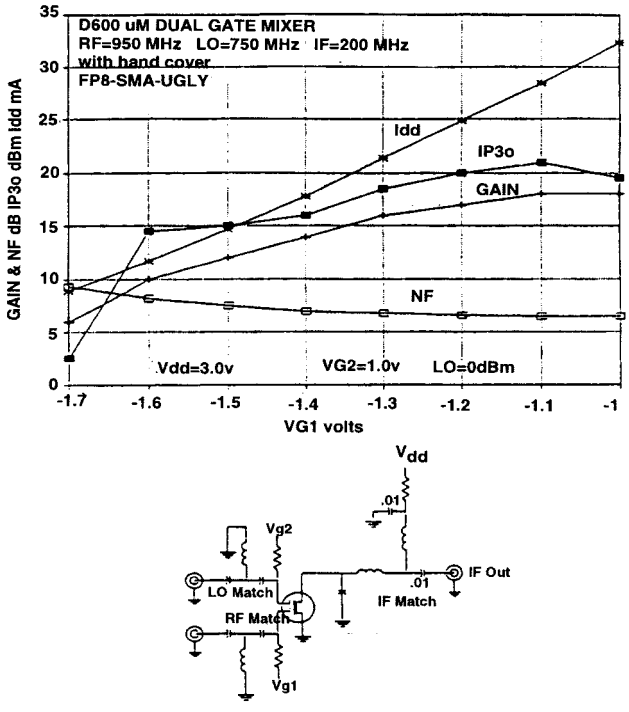


Figure 4-70g

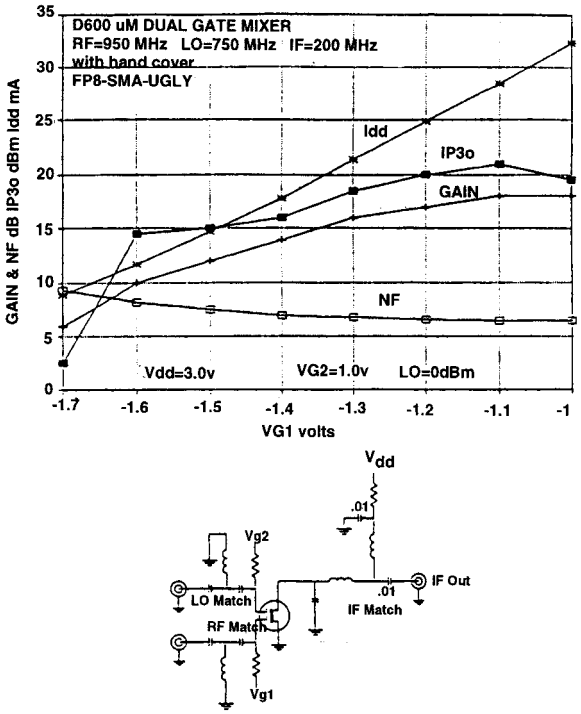


Figure 4-70h

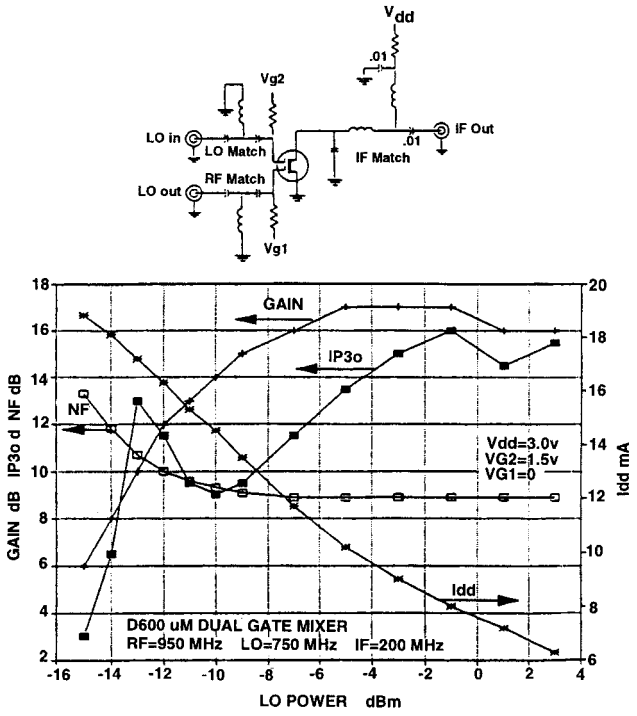


Figure 4-70i

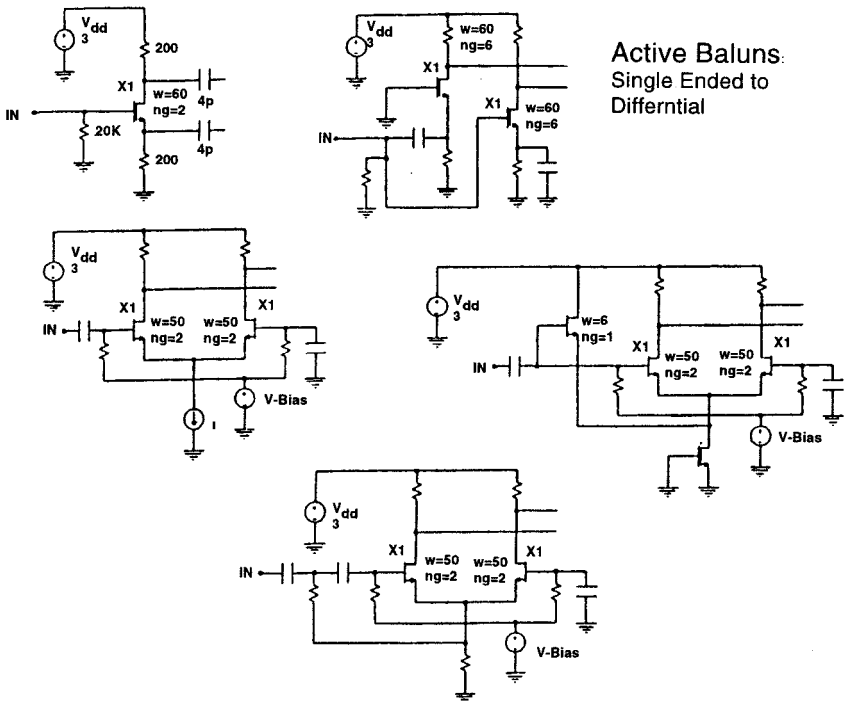


Figure 4-70j

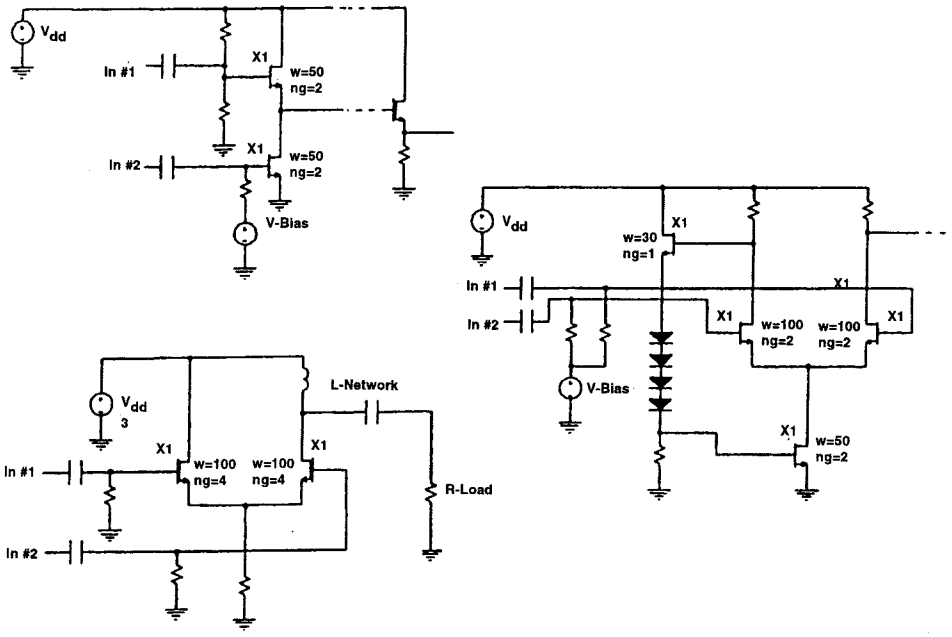


Figure 4-70k

REFERENCES

1. Otto Zinke and Hans Ludwig Hartnagel, *Lehrbuch der Hochfrequenztechnik*, *Band 2: Elektronik und Signalverarbeitung.* New York: Springer Verlag, 1987, pp. 429–430.
2. IRE Subcommittee 7.9 on Noise, "Description of Noise Performance in Receiving Systems," *Proc. IRE*, **51**, 436, 1963.
3. Stephen A. Maas, *Microwave Mixers*, 2nd ed. Boston: Artech House, 1993, pp. 153–171.
4. O. Zinke and H. Brunswig, *Lehrbuch der Hochfrequenztechnik*, Vol. 2, 3rd ed. New York: Springer Verlag, 1987, p. 435.
5. Cheng Qian, "Subharmonic Mixers Simplify Digital-Modulation Systems," *Microwaves & RF*, **37**(9), 62–70, September 1998.
6. Kenji Itoh et al., "Even-Harmonic Quadrature Modulator with Low Vector Modulation Error and Low Distortion for Microwave Digital Radio," *Microwave Theory and Techniques Society Symposium Digest*, 1996, pp. 967–970.
7. Shankar R. Joshi, "Novel I/Q Modulators Mix Cellular Signals," *Microwaves & RF*, **33**(11), 119–123, November 1994. Also see Shankar R. Joshi, "Modulator with Harmonic Mixers," Synergy Microwave Corporation, Paterson, NJ, U.S. Patent No. 5,416,449, May 16, 1995.
8. M. V. Schneider, R. Trambarulo, A. Gnauck, and M. J. Gans, "High-Speed Millimeter-Wave Modulator/Demodulator," *Microwave Theory and Techniques Society Symposium Digest*, 1990, pp. 611–614.
9. Jerzy Charamiec, "Subharmonically Pumped Schottky Diode Single-Sideband Modulator," *IEEE Trans. Microwave Theory Tech.*, **26**(9), 635–638, September 1978.
10. Ian Doyle, "A Simplified Subharmonic I/Q Modulator," *Appl. Microwave & Wireless*, **10**(8), 34–40, October 1998.
11. Peter Will, "Broadband Doubly Balanced Mixer Having Improved Termination Insensitivity Characteristics," Adams-Russell Co., Inc., Waltham, MA, U.S. Patent No. 4,224,572, September 23, 1980.
12. Joseph T. Lee, "Balanced Subharmonic Mixers," *Microwave J.*, **26**(8), August 1983.
13. Don Neuf, "Fundamental Versus Harmonic Mixing," *Microwave J.*, 1984.
14. Bert Henderson, "Full-Range Orthogonal Circuit Mixers Reach 2 to 26 GHz," *Microwave Syst. News*, January 1982.
15. Barry Gilbert, *Demystifying the Mixer* (self-published monograph), 1994, p. 27.
16. Kenneth K. Clarke and Donald T. Hess, *Communication Circuits: Analysis and Design*. Reading, MA: Addison-Wesley, 1971, p. 115.
17. George D. Vendelin, Anthony M. Pavio, and Ulrich L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990.
18. G. K. Tie and C. S. Aitchison, "Noise Figure and Associated Conversion Gain of a Microwave MESFET Gate Mixer," *Conference Proceedings of the 13th European Microwave Conference*, 1983.

FURTHER READING

Darrell L. Ash, "High-Performance TV Receiver," Texas Instruments Inc.

Sharief Babiker, Asen Asenov, Nigel Camerson, Stevfen P. Beaumont, and John R. Barker, "Complete Monte Carlo RF Analysis of 'Real' Short-Channel Compound FET's," *IEEE Trans. Electron Devices*, **Ed-45**(8), 1644–1652, August 1998.

Marc Borremans, Michel Steyaert, and Takashi Yoshitomi, "A 1.5 V, Wide Band 3 GHz, CMOS Quadrature Direct Up-Converter for Multi-Mode Wireless Communications," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference* , pp. 79–82.

A. Brunel, E. Douglas, K. Gallagher, J. Gillis, M. Goff, S. LeSage, E. Lewis, R. Pengelly, C. Ruhe, and J. Sano, "A Downconverter for Use in a Dual-Mode AMPS/CDMA Chip Set," *Microwave J.*, 20–42, February 1996.

P. Y. Chan, A. Rofougaran, K. A. Ahmed, and A. A. Abidi, "A Highly Linear 1-GHz CMOS Downconversion Mixer," *Proc. ESSCIRC*, Sevilla, September 1993, pp. 210–213.

Jan Crols and Michel S. J. Stayaert, "A 1.5-GHz Highly Linear CMOS Downconversion Mixer," *IEEE J. Solid-State Circuits*, **SC-30**(7), 736–742, July 1995.

Suman Datta, Shen Shi, Kenneth P. Roenker, Marc M. Chay, and William E. Stanchina, "Simulation and Design of InAlAs/InGaAs *pnp* Heterojunction Bipolar Transistors," *IEEE Trans. Electron Devices*, **Ed-46**(8), 1634–1641, August 1998.

Maria Luisa de la Fuente, Juan Pablo Pascual, and Eduardo Artal, "Analyze and Design a Cascode MESFET Mixer," *Microwaves & RF*, **37**(5) 129–138, May 1998.

Ravendar Goyal, *High-Frequency Analog Integrated Circuit Design* . New York: Wiley, 1995.

W. Hayward and S. Taylor, "Compensation Method and Apparatus for Enhancing Single Ended to Differential Conversion," TriQuint Semiconductor, Inc., Beaverton, OR, U.S. Patent No. 5,068,621, November 26, 1991.

Hewlett-Packard, Solutions from HP EEsof, *Low-Power Mixing Design Example Using HP Advanced Design System*, 1995.

K. Kanazawa et al., "A GaAs Double-Balanced Dual-Gate FET Mixer IC for UHF Receiver Front-End Applications," *IEEE MTT International Microwave Symposium Digest* , 1985, pp. 60–62.

K. Kimura, "Bipolar Four-Quadrant Analog Quarter-Square Multiplier," *IEEE J. Solid-State Circuits*, **29**(1), 46–55, January 1994.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

R. G. Meyer, "Intermodulation in High-Frequency Bipolar Transistor Integrated-Circuit Mixers," *IEEE J. Solid-State Circuits*, **SC-21**(4), 534–537, August 1986.

Ram Mohan Narayanan, "Multioctave, Double-Balanced Mixers Designed Using MIC Technology," *MSN & CT,* 108–115, January 1987.

D. Rabaey and J. Sevenhans, "The Challenges for Analog Circuit Design in Mobile Radio VLSI Chips," *Proceedings of the AACD Workshop* , Leuven, March 1993, Vol. 2, pp. 225–236.

J. Sevenhans, A. Vanwelsenaers, J. Wenin, and J. Baro, "An Integrated Si Bipolar Transceiver for a Zero IF 900 MHz GSM Digital Mobile Radio Front-End of a Hand Portable Phone," *Proc. CICC*, May 1991, pp. 7.7.1–7.7.4.

J. Sevenhans et al., "An Analog Radio Front-End Chip Set for a 1.9 GHz Mobile Radio Telephone Application," *Proc. ISSCC*, San Francisco, February 1994, pp. 44–45.

W. Titus, L. Holway, R. A. Pucel, P. Bernkopf, and A. Palevsky, "Broadband Control Components—FET Mixers," Final Technical Report, October 1988, Contract No. F33615-85-C-1725, PAY/RD/S-4270, Raytheon Company Research Division, Lexington, MA.

TQ9203 Mixer Noise Figure Measurement Note, TriQuint Semiconductor, Inc.

TriQuint Semicondutor GaAs Design Class Notes, Section 5 (Wes Hayward, "RFIC Lore and Philosophy" and "Non-Linear Designs") and Section 7 (Bob Hickey, "Design Example: Gilbert Cell Mixer with Active Baluns").

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

5

RF/WIRELESS OSCILLATORS

5-1 INTRODUCTION TO FREQUENCY CONTROL

Practically all modern telecommunication and test equipment uses frequency-control techniques based on *frequency synthesis*, the production of the many frequencies involved in a radiocommunication system's modulation, transmission, reception, and demodulation functions through the combination and mathematical manipulation of a very few input frequencies. Although the frequency synthesis techniques now ascendant in wireless systems—phase-locked loops and, less commonly, direct digital synthesizers—are fundamentally different, all are ultimately based on RF oscillators. This chapter covers oscillator theory, evaluation, and design.

Two types of oscillators are needed in a phase-locked-loop system (Figure 5-1). One, typically a *crystal oscillator*, generates the synthesizer's reference signal. As Figure 5-1 reflects, the reference oscillator may be built into the synthesizer IC in highly integrated systems. The other oscillator, a *voltage-controlled oscillator* (VCO), is varied in frequency by the system to produce the synthesizer's output signal. Although designing good oscillators remains somewhat like black magic or a special art, we will show that mathematics and CAD tools, applied in conjunction with practical experience, can keep the oscillator design process well under control.

5-2 BACKGROUND

The oscillator was probably first discovered by people who wanted to build an amplifier back in the vacuum-tube days. Its modern equivalent is shown in Figure 5-2. Here we see an active device, a bipolar transistor or field-effect transistor, with a tuned input and output circuit. Under ideal circumstances, the voltage at the input is 180° out of phase and the amplifier is "stable." Because of the unavoidable feedback capacitance, a certain amount of energy is transferred from the output to the input at about 90° out of phase. If one of the tuned circuits is detuned to the point where the phase shift is –180°, oscillation will occur. This means that the small energy from the output will be amplified, brought back to the input in phase and

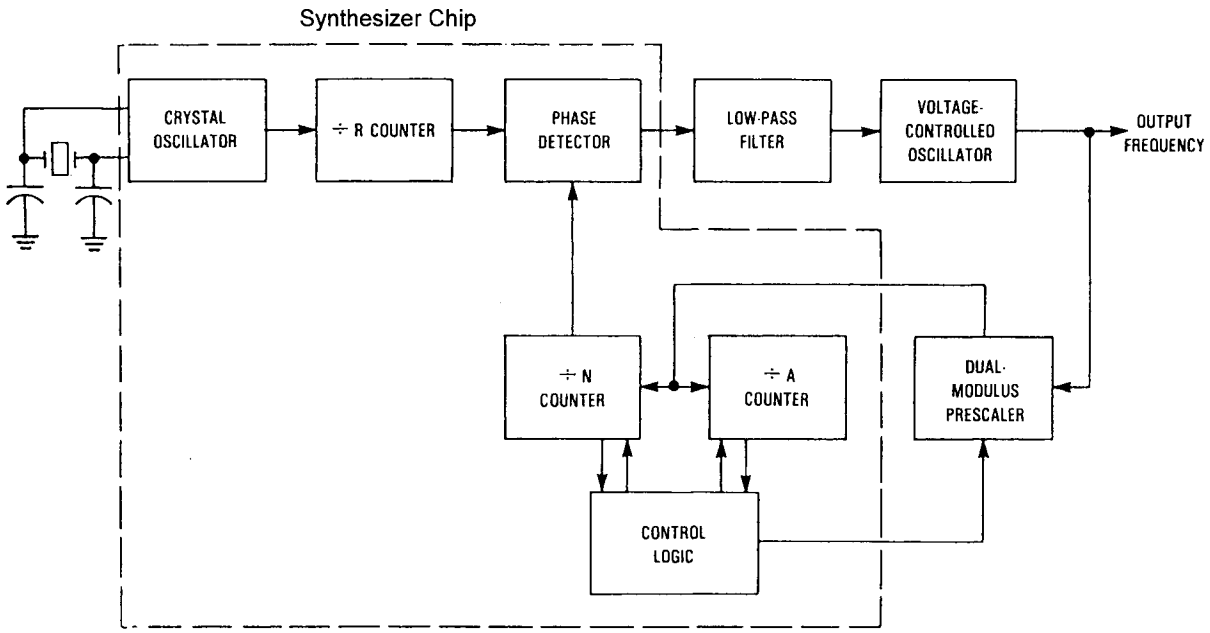


Figure 5-1 Block diagram of a modern, integrated frequency synthesizer. In this case, the designer has control over the VCO and the loop filter; the reference oscillator is part of the chip. In most cases (up to 2.5 GHz), the dual-modulus prescaler is inside the chip.

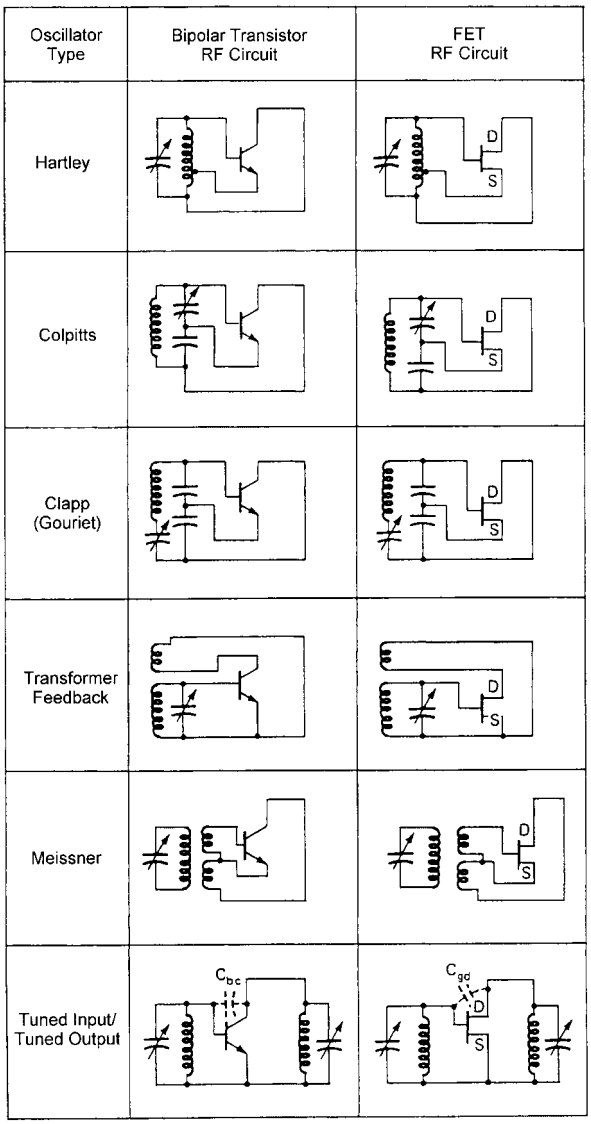


Figure 5-2 Schematic diagrams of RF connections for common oscillator circuits (dc and biasing circuits not shown) [1].

further amplified, causing the stage to "take off" and oscillate. Unless there is some mechanism to limit the amplitude of the oscillation at the input or output, its amplitude can theoretically increase to a level sufficient to destroy the device through breakdown effects or thermal runaway.

The criterion for oscillation was first described by Barkhausen [see Eq. (5-6) in the next section]. The conditions for oscillation exist when the feedback gain is high enough to cancel all losses and the difference between forward gain and reverse gain is less than zero. This implies also that oscillator circuits (Figure 5-2) provide a negative resistance, which is responsible for oscillation. The following mathematical treatment explains this.

5-3 OSCILLATOR DESIGN

A phase-locked loop generally has two oscillators: the oscillator at the output frequency and the reference oscillator. The reference oscillator at times can be another loop that is being mixed in, and the voltage-controlled oscillator (VCO) is controlled by either the reference or the oscillator loop. The VCO is one of the most important parts in a phase-locked-loop system, the performance of which is determined only by the loop filter at offsets inside the loop bandwidth, and only by the quality of the VCO design at offsets outside the loop bandwidth.

To some designers, VCO design appears to be magic. Shortly, we will go through the mathematics of the oscillator and some of its design criteria, but the results have only limited meaning. This is due to component tolerances, stray effects, and, most of all, nonlinear performance of the oscillator device, which can be modeled with only a certain degree of accuracy. However, after building oscillators for awhile, a certain feeling will be acquired for how to do this, and certain performance behavior will be predicted on a rule-of-thumb basis rather than on precise mathematical effort. For reasons of understanding, we will deal with the necessary mathematical equations, but we consider it essential to explain that these are only approximations.

5-3-1 Basics of Oscillators

An electronic oscillator is a device that converts dc power to a periodic output signal (ac power). If the output waveform is approximately sinusoidal, the oscillator is referred to as *sinusoidal.* There are many other oscillator types normally referred to as *relaxation* oscillators. For application in frequency synthesizers, we will explore only sinusoidal oscillators for reasons of purity and noise-sideband performance.

All oscillators are inherently nonlinear. Although the nonlinearity results in some distortion of the signal, linear analysis techniques can normally be used for the analysis and design of oscillators. Figure 5-3 shows, in block diagram form, the necessary components of an oscillator. It contains an amplifier with frequency-dependent forward loop gain $G(j\omega)$ and a frequency-dependent feedback network $H(j\omega)$.

The output voltage is given by

$$V_o = \frac{V_{\text{in}} G(j\omega)}{1 + G(j\omega)H(j\omega)} \tag{5-1}$$

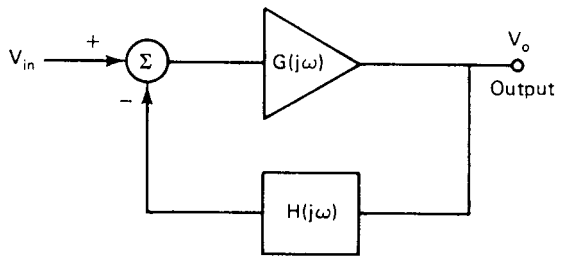


Figure 5-3 Block diagram of an oscillator showing forward and feedback loop components.

For an oscillator, the output V_o is nonzero even if the input signal $V_i = 0$. This can only be possible if the forward loop gain is infinite (which is not practical), or if the denominator

$$1 + G(j\omega)H(j\omega) = 0 \tag{5-2}$$

at some frequency ω_o. This leads to the well-known condition for oscillation (the *Nyquist criterion*), where at some frequency ω_o

$$G(j\omega_o)H(j\omega_o) = -1 \tag{5-3}$$

That is, the magnitude of the open-loop transfer function is equal to 1:

$$|G(j\omega_o)H(j\omega_o)| = 1 \tag{5-4}$$

and the phase shift is 180°:

$$\arg[G(j\omega_o)H(j\omega_o)] = 180° \tag{5-5}$$

This can be more simply expressed as follows: If in a negative-feedback system, the open-loop gain has a total phase shift of 180° at some frequency ω_o, the system will oscillate at that frequency provided that the open-loop gain is unity. If the gain is less than unity at the frequency where the phase shift is 180°, the system will be stable, whereas if the gain is greater than unity, the system will be unstable.

This statement is not correct for some complicated systems, but it is correct for those transfer functions normally encountered in oscillator design. The conditions for stability are also known as the *Barkhausen criteria*, which state that if the closed-loop transfer function is

$$\frac{V_o}{V_i} = \frac{\mu}{1 - \mu\beta} \tag{5-6}$$

the system will oscillate provided that $\mu\beta = 1$. This is equivalent to the Nyquist criterion, the difference being that the transfer function is written for a loop with positive feedback. Both versions state that the total phase shift around the loop must be 360° at the frequency of oscillation and the magnitude of the open-loop gain must be unity at that frequency.

The following analysis of the relatively simple oscillator shown in Figure 5-4 illustrates the design method. The linearized (and simplified) equivalent circuit of Figure 5-4 is given in Figure 5-5. h_{rb} has been neglected, and $1/h_{ob}$ has been assumed to be much greater than the load resistance R_L and is also ignored.

Note that the transistor is connected in the common-base configuration, which has no voltage phase inversion (the feedback is positive), so the conditions for oscillation are

$$|G(j\omega_o)H(j\omega_o) = 1| \tag{5-7}$$

and

$$\arg[G(j\omega_o)H(j\omega_o)] = 0° \tag{5-8}$$

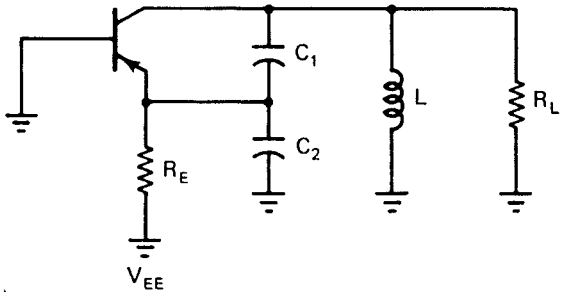


Figure 5-4 Oscillator with capacitive voltage divider.

The circuit analysis can greatly be simplified by assuming that

$$\frac{1}{\omega(C_2 + C_1)} << \frac{h_{ib}R_E}{h_{ib} + R_E} \tag{5-9}$$

and also that the Q of the load impedance is high. In this case the circuit reduces to that of Figure 5-6, where

$$V = \frac{V_o C_1}{C_1 + C_2} \tag{5-10}$$

and

$$R_{\text{eq}} = \frac{h_{ib}R_E}{h_{ib} + R_E}\left(\frac{C_1 + C_2}{C_1}\right)^2 \tag{5-11}$$

Figure 5-7 shows how the input impedance of an oscillator circuit—in this case, a BJT Colpitts oscillator minus its resonator—satisfies this condition for oscillation.

Then the forward gain

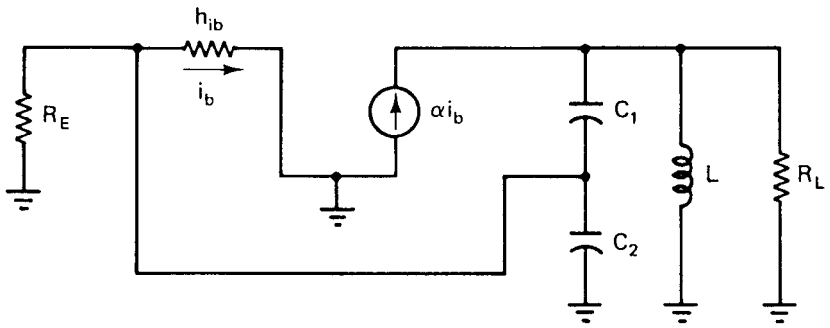


Figure 5-5 Linearized and simplified equivalent circuit of Figure 5-4.

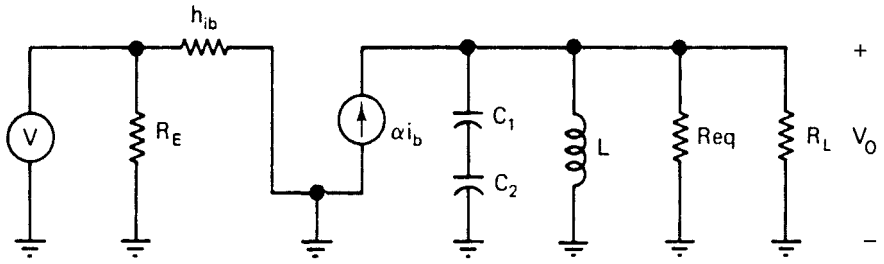


Figure 5-6 Further simplification of Figure 5-4, assuming high-impedance loads.

$$G(j\omega) = \frac{h_{fb}}{h_{ib}} Z_L = \frac{\alpha}{h_{ib}} Z_L \tag{5-12}$$

and

$$H(j\omega) = \frac{C_1}{C_1 + C_2} \tag{5-13}$$

where

$$Y_L = \frac{1}{Z_L} = \frac{1}{j\omega L} + \frac{1}{R_{\text{eq}}} + \frac{1}{R_L} + \frac{1}{j\omega C} \tag{5-14}$$

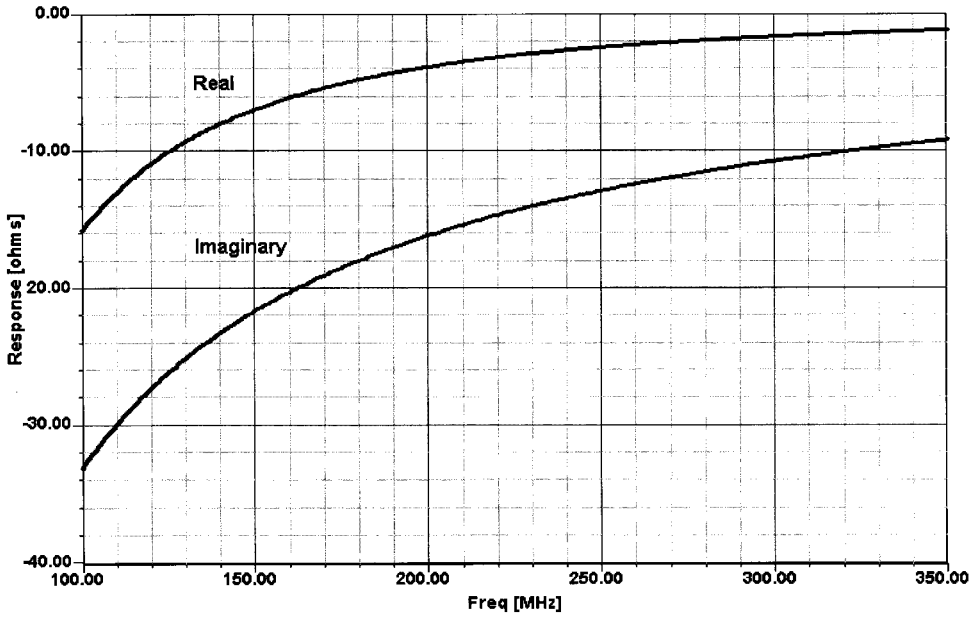


Figure 5-7 Calculated input impedance of a 200-MHz BJT Colpitts oscillator with its resonator removed. (Figure 5-46 shows an equivalent graph with the resonator present.)

A necessary condition for oscillation is that

$$\arg[G(j\omega)H(j\omega)] = 0^\circ \tag{5-15}$$

Since H does not depend on frequency in this example, if arg(GH) is zero, the phase shift of the load impedance Z_L must also be zero. This occurs only at the resonant frequency of the circuit,

$$\omega_o = \frac{1}{\sqrt{L[C_1C_2/(C_1 + C_2)]}} \tag{5-16}$$

At this frequency

$$Z_L = \frac{R_{\text{eq}}R_L}{R_{\text{eq}} + R_L} \tag{5-17}$$

and

$$GH = \frac{h_{fb}}{h_{ib}}\left(\frac{R_{\text{eq}}R_L}{R_{\text{eq}} + R_L}\right)\frac{C_1}{C_1 + C_2} \tag{5-18}$$

The other condition for oscillation is the magnitude constraint that

$$G(j\omega)H(j\omega) = \frac{\alpha}{h_{ib}}\left(\frac{R_{\text{eq}}R_L}{R_{\text{eq}} + R_L}\right)\frac{C_1}{C_1 + C_2} = 1 \tag{5-19}$$

Three-Reactance Oscillators. Although the block-diagram formulation of the stability criteria is an easy way to express stability mathematically, it is frequently not the easiest to apply since it is often difficult to identify the forward loop gain $G(j\omega)$ and the feedback ratio $H(j\omega)$ in electronic systems. A direct analysis of the circuit equations is frequently simpler than the block diagram interpretation (particularly for single-stage amplifiers). Figure 5-8 shows a generalized circuit for an electronic amplifier.

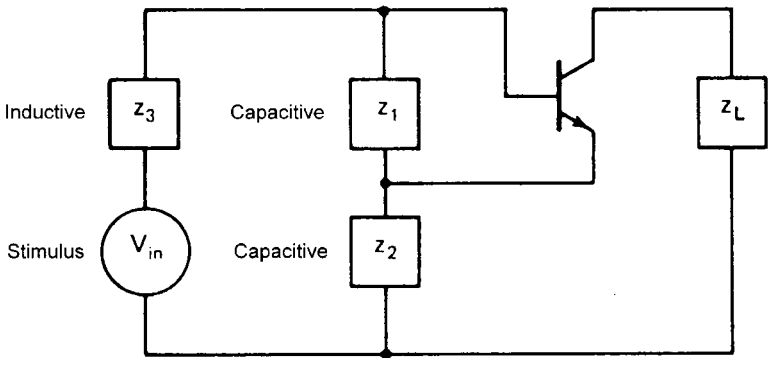


Figure 5-8 Generalized circuit for an oscillator using an amplifier model. Z_3 is inductive even with a capacitor in series with it.

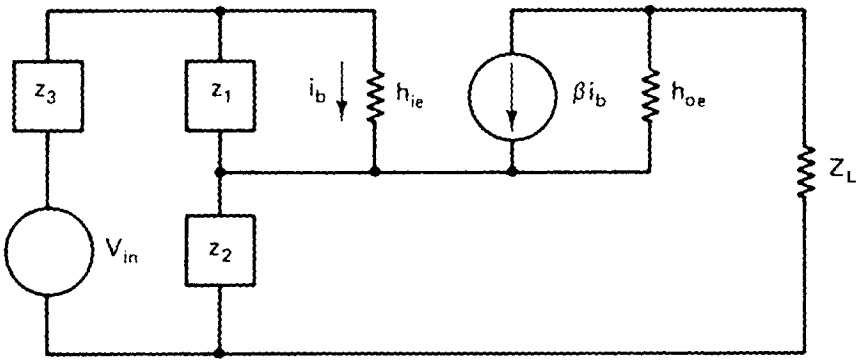


Figure 5-9 Small-signal equivalent circuit of Figure 5-8.

By inspecting Figure 5-8, one can see the three necessary reactances and the load. All other known circuits are derivatives of this by rotation. The small-signal equivalent circuit is given in Figure 5-9 (where h_{re} has been neglected).

Normally, h_{oe} can also be assumed sufficiently small and can be neglected. The loop equations are then

$$V_{\text{in}} = I_1(Z_3 + Z_1 + Z_2) - I_b Z_1 + \beta I_b Z_2 \tag{5-20}$$

$$0 = -I_1 Z_1 + I_b(h_{ie} + Z_1) \tag{5-21}$$

For the amplifier to oscillate, the currents I_b and I_1 must be nonzero even when $V_1 = 0$. This is only possible if the system determinant

$$\Delta = \begin{vmatrix} Z_3 + Z_1 + Z_2 & \beta Z_2 - Z_1 \\ -Z & h_{ie} + Z_1 \end{vmatrix} \tag{5-22}$$

is equal to 0. That is,

$$(Z_3 + Z_1 + Z_2)(h_{ie} + Z_1) - Z_1^2 + \beta Z_1 Z_2 = 0 \tag{5-23}$$

which reduces to

$$(Z_1 + Z_2 + Z_3)h_{fe} + Z_1 Z_2 \beta + Z_1(Z_2 + Z_3) = 0 \tag{5-24}$$

Assume for the moment that Z_1, Z_2, and Z_3 are purely reactive impedances. [It is easily seen that Eq. (5-24) does not have a solution if all three impedances are real.] Since both the real and imaginary parts must be zero, Eq. (5-24) is equivalent to the following equations if

$$h_{ie}(Z_1 + Z_2 + Z_3) = 0 \tag{5-25}$$

and

$$Z_1[(1 + \beta)Z_2 + Z_3] = 0 \tag{5-26}$$

Since β is real and positive, Z_2 and Z_3 must be of opposite sign for Eq. (5-26) to hold. That is,

$$(1 + \beta)Z_2 = -Z_3 \tag{5-27}$$

Therefore, since h_{ie} is nonzero, Eq. (5-25) reduces to

$$Z_1 + Z_2 - (1 + \beta)Z_2 = 0 \tag{5-28}$$

or

$$Z_1 = \beta Z_2 \qquad [C_1 < C_2; L_1 > L_2] \tag{5-29}$$

Thus, since β is positive, Z_1 and Z_2 will be reactances of the same kind. If Z_1 and Z_2 are capacitors, Z_3 is an inductor and the circuit is as shown in Figure 5-10. It is referred to as a *Colpitts* oscillator, named after the person who first described it.

If Z_1 and Z_2 are inductors and Z_3 is a capacitor as illustrated in Figure 5-11, the circuit is called a *Hartley* oscillator.

Example 1. Design a Colpitts circuit to oscillate at 200 MHz, using a transistor (operating at I_C = 6 mA dc) that has an input impedance

$$\text{Mag}(Z_{11}) = \frac{26 \text{ mV}}{I_C} \beta = 282\ \Omega (h_{ie} \triangleq Z_{11})$$

and

$$\beta_{RF} = 65; \quad \beta_{OSC} = 10 \quad \text{(large-signal condition)}$$

The oscillating transistor operates under large-signal conditions, generating harmonics in addition to the fundamental. Most of the transistor's current gain will come from dc biasing. β_{OSC}, its large-signal beta at the fundamental frequency of oscillation, must therefore be considerably less than β_{RF}, the small-signal, single-tone value, as the

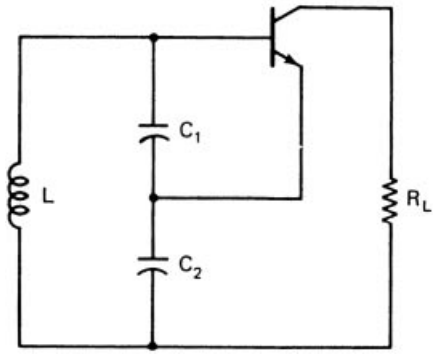


Figure 5-10 Colpitts oscillator.

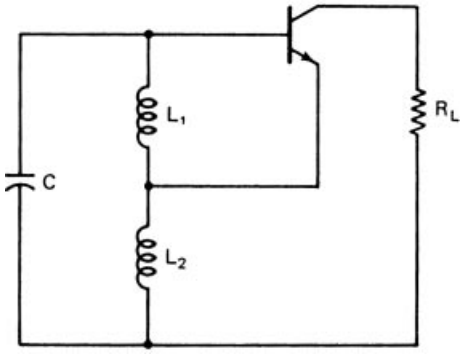


Figure 5-11 Hartley oscillator.

transistor's available current gain, β_0, is distributed among the dc, fundamental, and harmonic components present ($\beta_0 = \beta_{dc} + \beta_{OSC} + \beta_{OSC,F2} + \beta_{OSC,F3}\ldots$).

Solution: For the Colpitts circuit, Z_1 and Z_2 are capacitive reactances and Z_3 is an inductive reactance. Let $Z_2 = -j1.6\ \Omega$; then [Eq. (5-27)]

$$Z_3 = -(1 + \beta)Z_2 = -(1 + 10) \times (-1.6) \approx j17\ \Omega$$

and [Eq. (5-25)]

$$Z_1 = -(Z_2 + Z_3) \approx -j17\ \Omega$$

At 200 MHz, these impedances correspond to component values of $C_1 = 50$ pF, $C_2 = 500$ pF, and $L = 30$ nH. The oscillator will probably work with C_2 values larger than 500 pF, but at the loss of safety margin for tolerances in production, beta, temperature, and phase-noise performance. The completed circuit, except for biasing, is shown in Figure 5-12.

If Z_1, Z_2, Z_3, or h_{ie} is complex, the preceding analysis is more complicated, but the conditions for oscillation can still be obtained from Eq. (5-24). For example, if, in the Colpitts

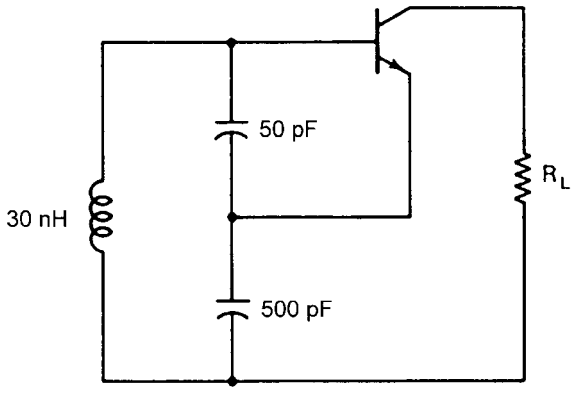


Figure 5-12 Design example of a Colpitts oscillator.

circuit, there is a resistor R in series with $L(Z_3 = R + j\omega L)$, Eq. (5-24) reduces to two equations:

$$h_{ie}\left(\omega L - \frac{1}{\omega C_1} - \frac{1}{\omega C_2}\right) - \frac{R}{\omega C_1} = 0 \tag{5-30}$$

and

$$h_{ie}R - \frac{1+\beta}{\omega^2 C_1 C_2} + \frac{1}{\omega C_1}\omega L = 0 \tag{5-31}$$

Define

$$C_1' = \frac{C_1}{1 + R/h_{ie}} \tag{5-32}$$

The resonant frequency at which oscillations will occur is found from Eq. (5-30) to be

$$\omega_o = \frac{1}{\sqrt{L[C_1' C_2/(C_1' + C_2)]}} \tag{5-33}$$

and for oscillations to occur $R_e(h_{ie}) = R_e(Z_{11})$ must be less than or equal to

$$R_e(h_{ie}) \leq \frac{1+\beta}{\omega_o^2 C_1 C_2} - \frac{L}{C_1} \tag{5-34}$$

If R becomes too large, Eq. (5-32) cannot be satisfied and oscillations will stop. In general, it is advantageous to have

$$X_{C_1} X_{C_2} = \frac{1}{\omega^2 C_1 C_2} \tag{5-35}$$

with $X \triangleq 1/\omega c$ as large as possible, since R can be large. However, if C_1 and C_2 are too small (large and X_{C_1} and X_{C_2}), the input and output capacitances of the transistor, which shunt C_1 and C_2, respectively, become important. A good, stable design will always have C_1 and C_2 much larger than the transistor capacitances they shunt.

Example 2. In Example 1, will the circuit still oscillate if the inductor now has a $Q_u = 100$? If the transistor input capacitance is 5 pF, what effect will this have on the system?

Solution: In Example 1, $X_L = 17\ \Omega$, $C_1 = 50$ pF, and $C_2 = 500$ pF. Since $C_1 = 50$ pF, adding 5 pF in parallel will change the equivalent C to 55 pF. As the inductor $Q_u = 100$, the equivalent resistance R in series with the lossless inductor is

$$R = \frac{17}{100} = 0.17\ \Omega$$

The new resonant frequency can be determined from Eqs. (5-32) and (5-33):

$$C_1' = \frac{C_1}{1 + R/h_{ie}} = \frac{55}{1 + (0.17/280)} \approx \frac{55}{1.0006} = 54.96 \text{ pF}$$

and

$$f_o = \frac{1}{2\pi L[C_1' C_2/(C_1' + C_2)]^{1/2}} \cong 191.56 \text{ MHz}$$

The effect of the finite inductor Q causes a negligible change in the oscillating frequency compared to the effect of the transistor input capacitance, which reduces the resonant frequency 4.2%. Because the large-signal beta, β_{OSC}, is 10, there is enough loop gain to maintain oscillation.

Two-Port Oscillator. Although Eqs. (5-25) and (5-26) can be used to determine the exact expressions for oscillation, they are often difficult to use and add little insight into the design process. An alternative interpretation, although not as accurate, will now be presented. It is based on the fact that an ideal tuned circuit (infinite Q), once excited, will oscillate infinitely because there is no resistance element present to dissipate the energy. In the actual case where the inductor Q is finite, the oscillations die out because energy is dissipated in the resistance. It is the function of the amplifier to maintain oscillations by supplying an amount of energy equal to that dissipated. This source of energy can be interpreted as a negative resistor in series with the tuned circuit. If the total resistance is positive, the oscillations will die out, while the oscillation amplitude will increase if the total resistance is negative. To maintain oscillations, the two resistors must be of equal magnitude. To see how a negative resistance is realized, the input impedance of the circuit in Figure 5-13 will be derived.

If h_{oe} is sufficiently small ($h_{oe} << 1/R_L$), the equivalent circuit is as shown in Figure 5-13. The steady-state loop equations are

$$V_{\text{in}} = I_{\text{in}}(X_{C_1} + X_{C_2}) - I_b(X_{C_1} - \beta X_{C_2}) \tag{5-36}$$

$$0 = -I_{\text{in}}(X_{C_1}) + I_b(X_{C_1} + h_{ie}) \tag{5-37}$$

After I_b is eliminated from these two equations, Z_{in} is obtained as

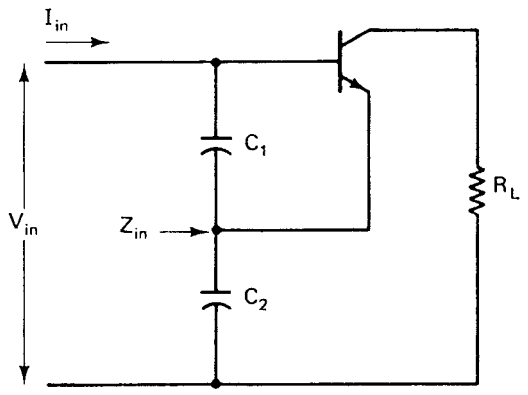


Figure 5-13 Calculation of input impedance of the negative-resistance oscillator.

$$Z_{\text{in}} = \frac{V_{\text{in}}}{I_{\text{in}}} = \frac{(1+\beta)X_{C_1}X_{C_2} + h_{ie}(X_{C_1} + X_{C_2})}{X_{C_1} + h_{ie}} \tag{5-38}$$

If $X_{C_1} << h_{ie}$, the input impedance is approximately equal to

$$Z_{\text{in}} \approx \frac{1+\beta}{h_{ie}} X_{C_1}X_{C_2} + (X_{C_1} + X_{C_2}) \tag{5-39}$$

$$Z_{\text{in}} \approx \frac{-g_m}{\omega^2 C_1 C_2} + \frac{1}{j\omega[C_1C_2/(C_1 + C_2)]} \tag{5-40}$$

That is, the input impedance of the circuit shown in Figure 5-14 is a negative resistor,

$$R = \frac{-g_m}{\omega^2 C_1 C_2} \tag{5-41}$$

in series with a capacitor,

$$C_{\text{in}} = \frac{C_1 C_2}{C_1 + C_2} \tag{5-42}$$

which is the series combination of the two capacitors.

With an inductor L (with the series resistance R_S) connected across the input, it is clear that the condition for sustained oscillation is

$$R_S = \frac{g_m}{\omega^2 C_1 C_2} \tag{5-43}$$

and the frequency of oscillation is

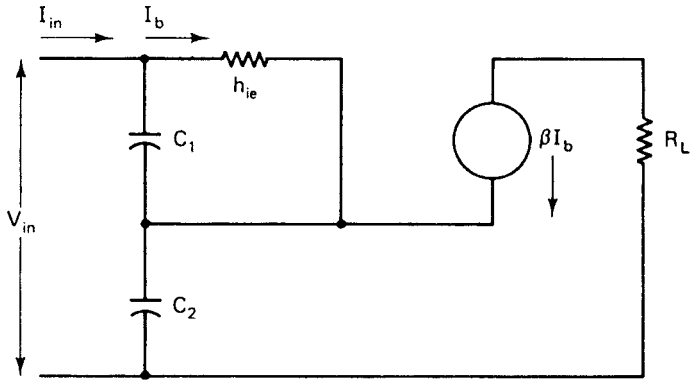


Figure 5-14 Equivalent small-signal circuit of Figure 5-13.

$$f_o = \frac{1}{2\pi\sqrt{L[C_1C_2/(C_1 + C_2)]}} \tag{5-44}$$

This interpretation of the oscillator readily provides several guidelines that can be used in the design. First, C_1 should be as large as possible so that

$$X_{C_1} << h_{ie} \tag{5-45}$$

and C_2 is to be large so that

$$X_{C_2} << \frac{1}{h_{oe}} \tag{5-46}$$

When these two capacitors are large, the transistor base-to-emitter and collector-to-emitter capacitances will have a negligible effect on the circuit's performance. However, Eq. (5-43) limits the maximum value of the capacitances since

$$r \leq \frac{g_m}{\omega^2 C_1 C_2} \leq \frac{G}{\omega^2 C_1 C_2} \tag{5-47}$$

where G is the maximum value of g_m. For a given product of C_1 and C_2, the series capacitance is at maximum when $C_1 = C_2 = C_m$. Thus Eq. (5-47) can be written

$$\frac{1}{\omega C_m} > \sqrt{\frac{r}{G}} \tag{5-48}$$

This equation is important in that it shows that for oscillations to be maintained, the minimum permissible reactance $1/\omega C_m$ is a function of the resistance of the inductor and the transistor's mutual conductance, g_m.

An oscillator circuit known as the *Clapp circuit* or *Clapp–Gouriet circuit* is shown in Figure 5-15. This oscillator is equivalent to the one just discussed, but it has the practical advantage of being able to provide another degree of design freedom by making C_o much smaller than C_1 and C_2.

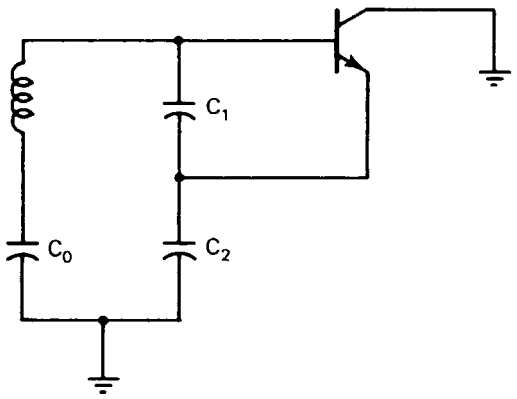


Figure 5-15 Circuit of a Clapp oscillator.

It is possible to use C_1 and C_2 to satisfy the condition of Eq. (5-47) and then adjust C_o for the desired frequency of oscillation ω_o, which is determined from

$$\omega_o L - \frac{1}{\omega_o C_o} - \frac{1}{\omega_o C_1} - \frac{1}{\omega_o C_2} = 0 \tag{5-49}$$

Amplitude Stability. Linearized analysis of the oscillator is convenient for determining the frequency but not the amplitude of the oscillation. The Nyquist stability criterion defines the frequency of oscillation as the frequency at which the loop phase shift is 360°, but it says nothing about the oscillation amplitude. If no provisions are taken to control the amplitude, it is susceptible to appreciable drift. Two frequently used methods for controlling the amplitude are operating the transistor in the nonlinear region or using a second stage for amplitude limiting. For the single-stage oscillator, amplitude limiting is accomplished by designing an unstable oscillator; that is, the loop gain is made greater than 1 at the frequency at which the phase shift is 180°. As the amplitude increases, the β of the transistor decreases, causing the loop gain to decrease until the amplitude stabilizes. This is a self-limiting oscillator. There are nonlinear analysis techniques predicting the amplitude of oscillation, but their results are approximate except in idealized cases, forcing the designer to resort to an empirical approach.

An example of a two-stage emitter-coupled oscillator is shown in Figure 5-16. In this circuit, amplitude stabilization occurs as a result of current limiting in the second stage. This circuit has the additional advantage that it has output terminals that are isolated from the feedback path. The emitter signal of Q_2, having a rich harmonic content, is normally used as output. Harmonics of the fundamental frequency can be extracted at the emitter of Q_2 by using an appropriately tuned circuit. Note that the collector of Q_2 is isolated from the feedback path.

Phase Stability. An oscillator has a frequency or phase stability that can be considered in two separate parts. First, there is *long-term* stability, in which the frequency changes over a period of minutes, hours, days, weeks, or even years. This frequency stability is normally limited by the circuit components' temperature coefficients and aging rates. The other part,

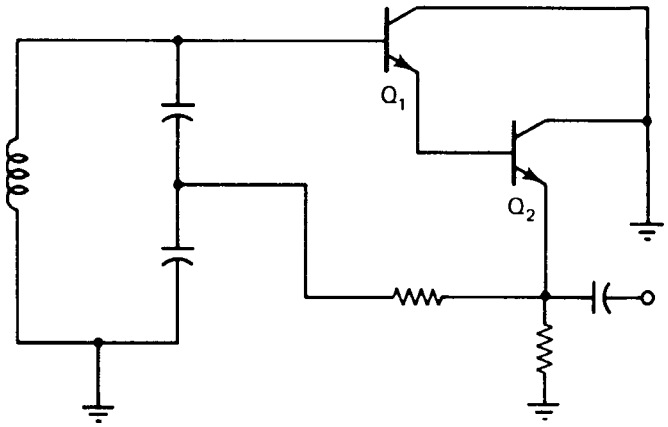


Figure 5-16 Two-stage, emitter-coupled oscillator.

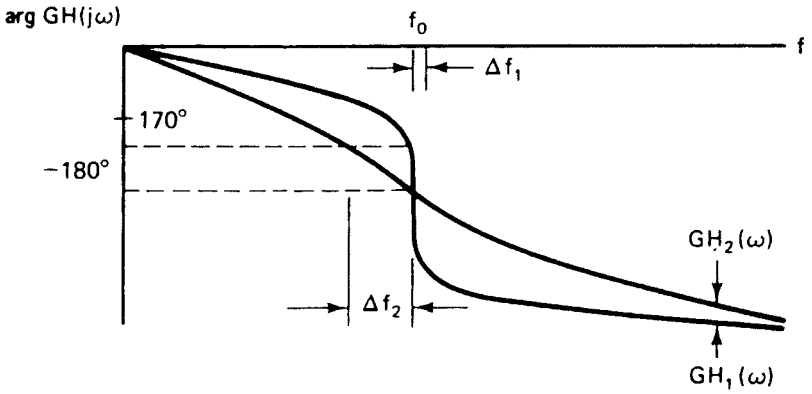


Figure 5-17 Phase plot of two open-loop systems with different resonator *Q* values.

short-term frequency stability, is measured in terms of seconds. One form of short-term instability is due to changes in phase of the system; here the term *phase stability* is used synonymously with frequency stability. It refers to how the frequency of oscillation reacts to small changes in phase shift of the open-loop system. It can be assumed that the system with the largest rate of change of phase versus frequency $d\phi/df$ will be the most stable in terms of frequency stability. Figure 5-17 shows the phase plots of two open-loop systems used in oscillators. At the system crossover frequency, the phase shift is −180°. If some external influence causes a shift in phase—say, it adds 10° of phase lag—the frequency will change until the total phase shift is again 0°. In this case, the frequency will decrease to the point where the open-loop phase shift is 170°. Figure 5-17 shows that Δf_2, the change in frequency associated with the 10° change in phase of GH_2, is greater than the change in frequency Δf_1, associated with the open-loop system GH_1, whose phase is changing more rapidly near the open-loop crossover frequency.

This qualitative discussion illustrates that $d\phi/df$ at $f = f_o$ is a measure of an oscillator's phase stability. It provides a good means of quantitatively comparing the phase stability of two oscillators. Consider the simple parallel tuned circuit shown in Figure 5-18.

For this circuit, the two-port is

$$\frac{V_o(j\omega)}{I(j\omega)} = \frac{R}{1 + jQ[(\omega/\omega_o) - (\omega_o/\omega)]} \tag{5-50}$$

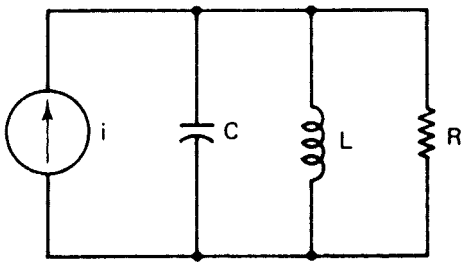


Figure 5-18 Parallel tuned circuit for phase-shift analysis.

where

$$\omega_o = \frac{1}{\sqrt{LC}} \quad \text{and} \quad Q = \frac{R}{\omega_o L} \tag{5-51}$$

The circuit phase shift is

$$\arg \frac{V_o}{I} = \theta = \tan^{-1} Q\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right) \tag{5-52}$$

and

$$\frac{d\theta}{d\omega} = \frac{1/Q}{1/Q^2 + [(\omega^2 - \omega_o^2)/\omega_o\omega]^2} \frac{\omega^2 + \omega_o^2}{(\omega_o\omega)^2} \tag{5-53}$$

At the resonant frequency ω_o,

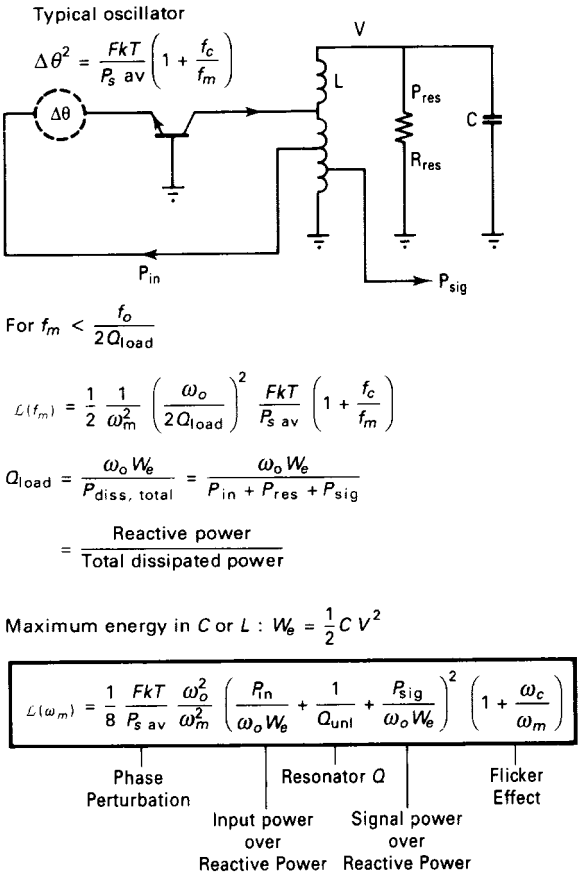


Figure 5-19 Diagram for a feedback oscillator illustrating the principles involved, and showing the key components considered in the phase noise calculation and its contribution.

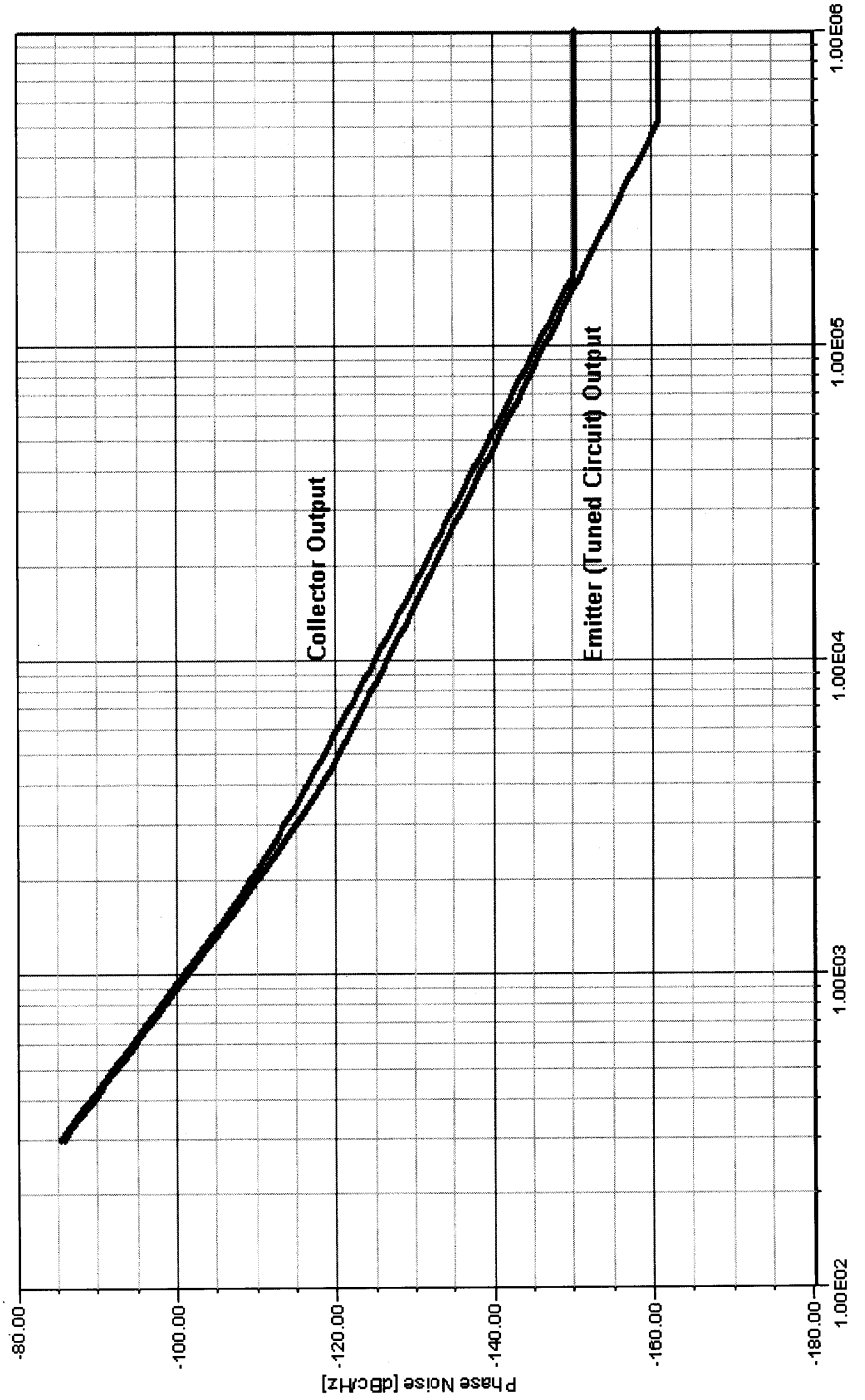


Figure 5-20 Phase noise with oscillator output taken from collector (upper trace) versus emitter (lower trace) of a BJT oscillator.

$$\left.\frac{dQ}{d\omega}\right|_{\omega=\omega_o} = \frac{2Q}{\omega_o} \tag{5-54}$$

The frequency stability factor S_F is defined as the change in phase $d\phi/d\omega$ divided by the normalized change in frequency $\Delta\omega/\omega_o$. That is,

$$S_F = 2Q \tag{5-55}$$

S_F is a measure of the short-term stability of an oscillator. Equation (5-54) indicates that the higher the circuit Q, the higher the stability factor. This is one reason for using high-Q circuits in oscillator circuits. Another reason is the ability of the tuned circuit to filter out undesired harmonics and noise.

The two oscillator types shown so far can be considered two-port oscillators because one port is responsible for oscillation and the other port supplies the output power. This type of output supposedly has better isolation (although this is not true at higher frequencies) but has a higher noise floor. A better way of extracting energy is shown in Figure 5-19.

These types of oscillators are essentially what is referred to as *two-port* oscillators. The name comes from the fact that the oscillating device has two ports, with the output energy taken from the output port, typically the collector or drain, while the tuned circuit is at the input port. For years, it was widely held that this type of design provided better isolation between the tuned circuit and termination; unfortunately, because of the Miller effect, this is not true. At the higher frequencies, changes in output-port loading affect the input port, resulting in *pulling*, a detuning effect that increases with frequency. In addition, taking the output from the collector or the drain results in a poorer signal-to-noise ratio than that achievable by extracting energy from the tuned circuit itself. The only real advantages the collector output provides are more power and higher efficiency, but definitely at the expense of phase noise, which is often referred to as *SSB phase noise* because it is expressed in terms of the decibel ratio of the noise power in a single (the upper or lower) noise sideband, in a 1-Hz bandwidth centered at a specified frequency offset from the oscillator carrier, to the carrier power. Figure 5-20 compares the phase noise of an oscillator with the energy taken off the collector and the tuned circuit; the difference is highly visible.

5-4 OSCILLATOR CIRCUITS

5-4-1 Hartley

Figure 5-21 shows the Hartley oscillator configuration, in which feedback is obtained by tapping the resonator. In practice, the resonator can consist of a single tapped inductor or two separate, magnetically uncoupled inductors, with feedback obtained at their junction.

5-4-2 Colpitts

Figure 5-22 shows the Colpitts oscillator configuration, in which feedback is obtained via a capacitive voltage divider.

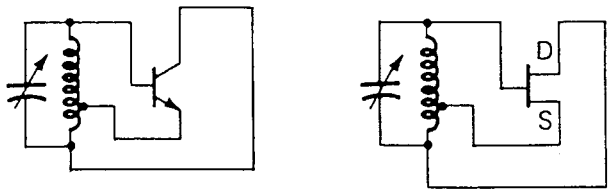


Figure 5-21 Hartley oscillator.

5-4-3 Clapp–Gouriet

Figure 5-23 shows the Clapp–Gouriet oscillator. Like the Colpitts, the Clapp–Gouriet obtains its feedback via a capacitive voltage divider; unlike the Colpitts, an additional capacitor series-tunes the resonator. The Pierce oscillator, a configuration used only with crystals, is a rotation of the Clapp–Gouriet oscillator in which the emitter is at RF ground.

5-5 DESIGN OF RF OSCILLATORS

5-5-1 General Thoughts on Transistor Oscillators

An estimate of the noise performance of an oscillator is

$$\mathcal{L}(\omega_m) = \frac{1}{8}\frac{FkT}{P_{s,\text{av}}}\frac{\omega_0^2}{\omega_m^2}\left(\frac{P_{\text{in}}}{\omega_0 W_e} + \frac{1}{Q_{\text{unl}}} + \frac{P_{\text{sig}}}{\omega_0 W_e}\right)^2\left(1 + \frac{\omega_c}{\omega_m}\right) \tag{5-56}$$

Phase perturbation

Resonator Q

Flicker effect

Input power/Reactive power ratio

Signal power/Reactive power ratio

We just state this here without presenting oscillator noise theory itself. (Also see Figure 5-19.) This equation is based on work done by Dieter Scherer of Hewlett-Packard about 1978. He was the first to introduce the flicker effect to the Leeson equation by adding the AM-to-PM

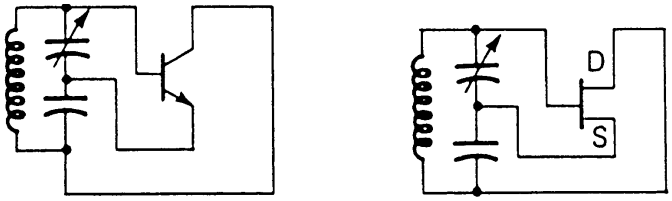


Figure 5-22 Colpitts oscillator.

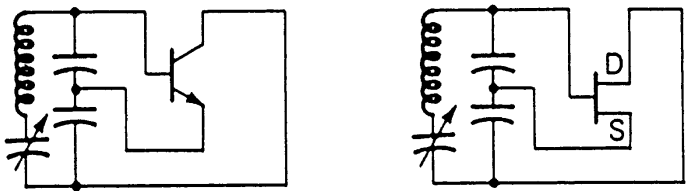


Figure 5-23 Clapp–Gouriet oscillator.

conversion effect, which is caused by the nonlinear capacitance of the active devices. This equation must be further expanded:

$$\pounds(f_m) = 10 \log\left\{\left[1 + \frac{f_0^2}{(2f_m Q_{\text{load}})^2}\right]\left(1 + \frac{f_c}{f_m}\right)\frac{FkT}{2P_{\text{sav}}} + \frac{2kTRK_0^2}{f_m^2}\right\} \tag{5-57}$$

where $\pounds(f_m)$ = ratio of sideband power in 1-Hz bandwidth at f_m to total power in dB
f_m = frequency offset
f_0 = center frequency
f_c = flicker frequency
Q_{load} = loaded Q of the tuned circuit
F = noise factor
$kT = 4.1 \times 10^{-21}$ at 300 K (room temperature)
P_{sav} = average power at oscillator output
R = equivalent noise resistance of tuning diode (typically 200 Ω to 10 kΩ)
K = oscillator voltage gain

The following table (from Motorola) shows the flicker corner frequency f_C as a function if I_C for a typical small-signal microwave BJT. $I_{C(\text{max})}$ of this transistor is about 10 mA.

I_C (mA)	f_C (kHz)
0.25	1
0.5	2.74
1	4.3
2	6.27
5	9.3

Note that f_C, which is defined by AF and KF in the SPICE model, increases with I_C. This gives us a clue about how f_C changes when a transistor oscillates. As a result of the bias-point shift that occurs during oscillation (see Example 1 of Section 5-3-1), an oscillating BJT's average I_C is higher than its small-signal I_C. KF is therefore higher for a given BJT operating as an oscillator than for the same transistor operating as a small-signal amplifier. This must be kept in mind when considering published f_C data, which are usually determined under small-signal conditions without being qualified as such. Determining a transistor's oscillating f_C is best done through measurement: Operate the device as a high-Q UHF oscillator (we

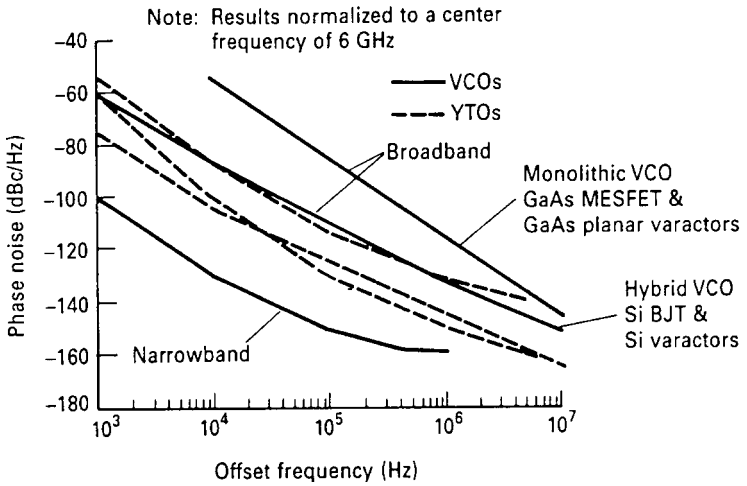


Figure 5-24 Phase noise of oscillators using different semiconductors and resonators.

suggest using a ceramic-resonator-based tank in the vicinity of 1 GHz), and measure its close-in (10 Hz to 10 kHz) phase noise versus offset from the carrier. f_C will correspond to a slight decrease in the slope of the phase noise versus offset curve. Generally, f_C varies with device type as follows: silicon JFETs, 50 Hz and higher; microwave RF BJTs, 1–10 kHz (as above); MOSFETs, 10–100 kHz; and GaAsFETs, 10–100 MHz. Figure 5-24 shows the phase noise of oscillators using different semiconductors and resonators.

Equation (5-57) is based on Rohde et al. [2]. The additional term introduces a distinction between a conventional oscillator and a VCO. Whether the voltage- or current-dependent capacitance is internal or external makes no difference; it simply affects the frequency.

For a more complete expression for a resonator oscillator's phase noise spectrum we can write

$$s_\phi(f_m) = \frac{\alpha_R F_0^4 + \alpha_E (F_0/2Q_L)^2}{f_m^3} + \frac{(2GFkT/P_0)\,(F_0/2Q_L)^2}{f_m^2} + \frac{2\alpha_R Q_L F_0^3}{f_m^2} + \frac{\alpha_E}{f_m} + \frac{2GFkT}{P_0} \tag{5-58}$$

where G = compressed power gain of the loop amplifier
F = noise factor of the loop amplifier
k = Boltzmann's constant
T = temperature in kelvins

P_0 = carrier power level (in watts) at the output of the loop amplifier
F_0 = carrier frequency in hertz
f_m = carrier offset frequency in hertz
$Q_L (= \pi F_0 \tau_g)$ = loaded Q of the resonator in the feedback loop
α_R, α_E = flicker-noise constants for the resonator and loop amplifier, respectively.

In frequency synthesizers, we have no use for *LC* oscillators that do not include a tuning diode or diodes, but it may still be of interest to analyze the low-noise, fixed-tuned *LC* oscillator first and later make both elements, inductor and capacitor, variable.

Later, we will show the performance changes if we utilize the two possible ways of getting coarse and fine tuning in oscillators:

1. Use of tuning diodes.
2. Use of switching diodes.

We will spend some time looking at the effects that switching and tuning diodes have in a circuit because they will ultimately influence the noise performance more strongly than the transistor itself.

The reason is that the noise generated in tuning diodes will be superimposed on the noise generated in the circuit, while switching diodes have losses that cause a reduction of circuit *Q*. The selection of the proper tuning and switching diodes is important, as is the proper way of connecting them. As both types are modifications of the basic *LC* oscillator, we start with the *LC* oscillator itself.

Early signal generators as were offered by several companies (namely, Rohde & Schwarz, Hewlett-Packard, Boonton Electronics, or Marconi), if they are not synthesized, use an air-dielectric variable capacitor or, as in the case of one particular Hewlett-Packard generator, the Model HP8640, a tuned cavity.

Tuning here is accomplished by changing the value of an air-dielectric variable capacitor or changing the mechanical length of a quarter-wave resonator.

Using the equations shown previously, it is fairly easy to calculate the performance of oscillators and understand how they work, but this does not necessarily optimize their design. For *LC* oscillator applications in which noise performance is crucial, the oscillator shown in Figure 5-25, used in the famous Rohde & Schwarz SMDU, is the state of the art. Its noise performance is equivalent to the noise found in the cavity-tuned oscillator used in Hewlett-Packard's popular HP8640 signal generator, and because of the unique way a tuning diode is coupled to the circuit, its modulation capabilities are substantially superior to any of the signal generators currently offered. To develop such a circuit from design equations is not possible. This circuit is the result of many years of experience and research and looks fairly simple. The grounded-gate FET circuit provides the best performance because it fulfills the important requirements of Eq. (5-56).

The tuned circuit is not connected directly to the drain, but the drain is put on a tap of the oscillator section. Therefore, the actual voltage across the tuning capacitor is higher than the supply voltage, and thus the energy stored in the capacitor is much higher than in a circuit connected between the gate electrode and ground, as it is in a normal Colpitts oscillator. In addition, the high output impedance of the FET does not load the circuit, which also provides a reduced noise contribution. Since this oscillator is optimized for best frequency-modulation

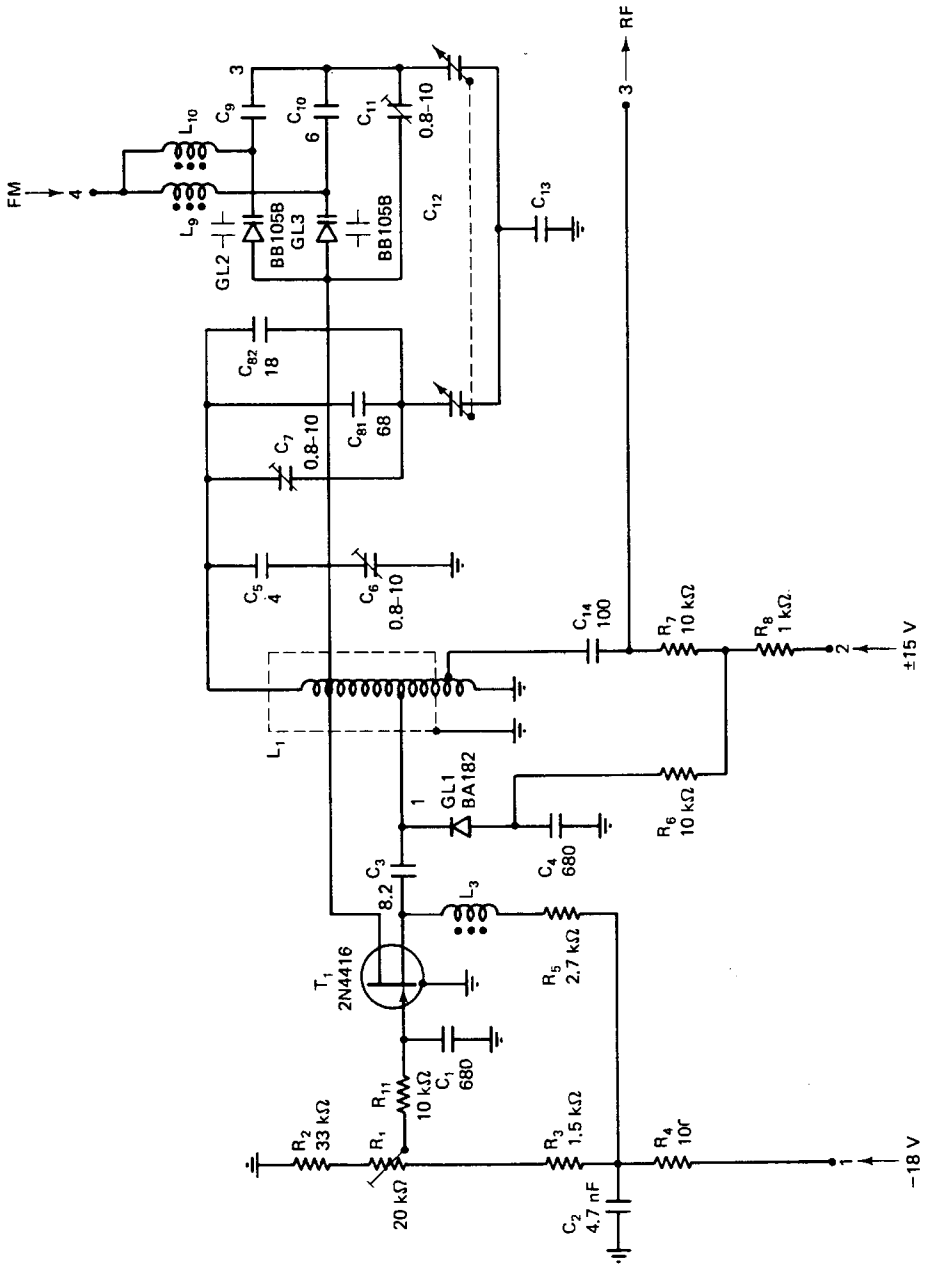


Figure 5-25 A 118–198-MHz oscillator from the Rohde & Schwarz SMDU signal generator.

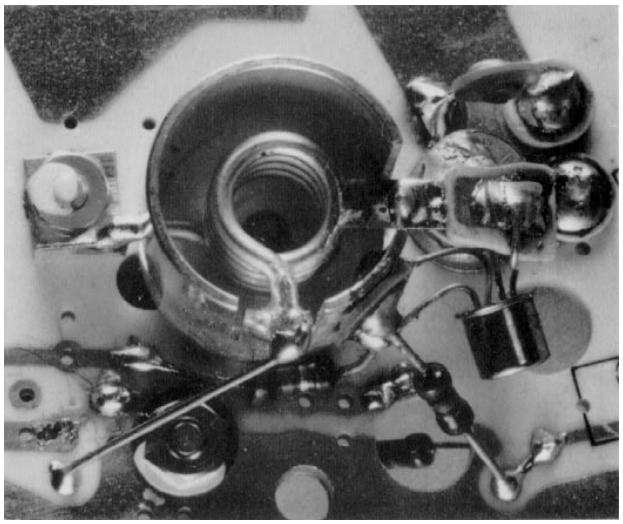


Figure 5-26 Photograph of the helical-resonator system from the Rohde & Schwarz SMDU signal generator.

performance in the FM frequency range, it becomes apparent that it fits the requirements of low-distortion stereo modulation.

For extremely critical blocking measurements in the 2-m region (140–160 MHz), the noise specification at 20 kHz off the carrier is of high importance, while the peak modulation typically does not exceed 5 kHz. Figure 5-25 shows the schematic of the oscillator section optimized for this frequency range.

Because of f_T limitations, it has been found experimentally that these *LC* oscillators, if based on silicon JFETs, should not be used above about 500 MHz; rather, a doubler stage should be employed. Analyzing the signal generators currently on the market, we find that their highest base band typically ranges from 200 to 500 MHz, with frequency doubling used to 1000 MHz. As can be seen in Figure 5-26, the mechanical layout of such an oscillator is extremely compact.

5-5-2 Two-Port Microwave/RF Oscillator Design

A common method for designing oscillators is to resonate the input port with a passive high-*Q* circuit at the desired frequency of resonance. It will be shown that if this is achieved with a load connected on the output port, the transistor is oscillating at both ports and is thus delivering power to the load port. The oscillator may be considered a two-port structure, with M_3 being the lossless resonating port and M_4 providing lossless matching such that all the external RF power is delivered to the load. See Figure 5-27. The resonating network has been described. Nominally, only parasitic resistance is present at the resonating port, since a high-*Q* resonance is desirable for minimizing oscillator noise. It is possible to have loads at both the input and output ports if such an application occurs, since the oscillator is oscillating at both ports simultaneously.

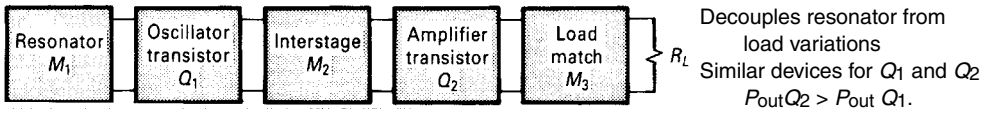


Figure 5-27 Buffered oscillator design.

The simultaneous oscillation condition is proved as follows. Assume that the oscillation condition is satisfied at port 1:

$$1/S'_{11} = \Gamma_G \tag{5-59}$$

Thus,

$$S'_{11} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L} = \frac{S_{11} - D\Gamma_L}{1 - S_{22}\Gamma_L} \tag{5-60}$$

$$\frac{1}{S'_{11}} = \frac{1 - S_{22}\Gamma_L}{S_{11} - D\Gamma_L} = \Gamma_G \tag{5-61}$$

By expanding Eq. (5-61), we find

$$\begin{gathered} \Gamma_G S_{11} - D\Gamma_L\Gamma_G = 1 - S_{22}\Gamma_L \\ \Gamma_L(S_{22} - D\Gamma_G) = 1 - S_{11}\Gamma_G \\ \Gamma_L = \frac{1 - S_{11}\Gamma_G}{S_{22} - D\Gamma_G} \end{gathered} \tag{5-62}$$

Thus,

$$S'_{22} = S_{22} + \frac{S_{12}S_{21}\Gamma_G}{1 - S_{11}\Gamma_G} \tag{5-63}$$

$$\frac{1}{S'_{22}} = \frac{1 - S_{11}\Gamma_G}{S_{22} - D\Gamma_G} \tag{5-64}$$

Comparing Eqs. (5-62) and (5-64), we find

$$1/S'_{22} = \Gamma_L \tag{5-65}$$

which means that the oscillation condition is also satisfied at port 2; this completes the proof. Thus, if either port is oscillating, the other port must be oscillating as well. A load may appear at either or both ports, but normally the load is in Γ_L, the output termination. This result can be generalized to an n-port oscillator by showing that the oscillator is simultaneously oscillating at each port:

$$\Gamma_1 S'_{11} = \Gamma_2 S'_{22} = \Gamma_3 S'_{33} = \cdots = \Gamma_n S'_{nn} \tag{5-66}$$

Before concluding this section on two-port oscillator design, the buffered oscillator shown in Figure 5-27 must be considered. This design approach is used to provide the following:

1. A reduction in load-pulling, which is the change in oscillator frequency when the load reflection coefficient changes.
2. A load impedance that is more suitable to wideband applications ($k < 1$).
3. A higher output power from a working design, although the higher output power can also be achieved by using a larger oscillator transistor.
4. This type of oscillator is not optimized for low phase noise. However, buffered oscillator designs are quite common in wideband YIG applications, where changes in the load impedance must not change the generator frequency.

Two-port oscillator design may be summarized as follows:

1. Select a transistor with sufficient gain and output-power capability for the frequency of operation. This may be based on oscillator datasheets, amplifier performance, or *S*-parameter calculation.
2. Select a topology that gives $k < 1$ at the operating frequency. Add feedback if $k < 1$ has not been achieved.
3. Select an output load-matching circuit that gives $|S'_{11}| > 1$ over the desired frequency range. In the simplest case, this could be a 50-Ω load.
4. Resonate the input port with a lossless termination so that $\Gamma_G S'_{11} = 1$. The value of S'_{22} will be greater than unity with the input properly terminated.

In all cases, the transistor delivers power to a load and the input of the transistor. Practical considerations of realizability and dc biasing will determine the best design.

For both bipolar and FET oscillators, a common topology is the common-base or common-gate, since a common-lead inductance can be used to raise S_{22} to a large value, usually greater than unity even with a 50-Ω generator resistor. However, it is not necessary for the transistor S_{22} to be greater than unity, since the 50-Ω generator is not present in the oscillator design. The requirement for oscillation is $k < 1$; then resonating the input with a lossless termination will ensure that $|S'_{11}| > 1$.

A simple example will clarify the design procedure. A common-base bipolar transistor (HP2001) was selected to design a fixed-tuned oscillator at 2 GHz. The common-base *S*-parameters and stability factors are given in Table 5-1.

Using the load circuit in Figure 5-28, we see that the reflection coefficients are

Table 5-1 *S* parameters and stability factors (V_{CE} = 15 V, I_C = 25 mA)

$L_B = 0$	$L_B = 0.5$ nH
$S_{11} = 0.94 \angle 174°$	$1.04 \angle 173°$
$S_{21} = 1.90 \angle -28°$	$2.00 \angle -30°$
$S_{12} = 0.013 \angle 98°$	$0.043 \angle 153°$
$S_{22} = 1.01 \angle -17°$	$1.05 \angle -18°$
$k = -0.09$	-0.83

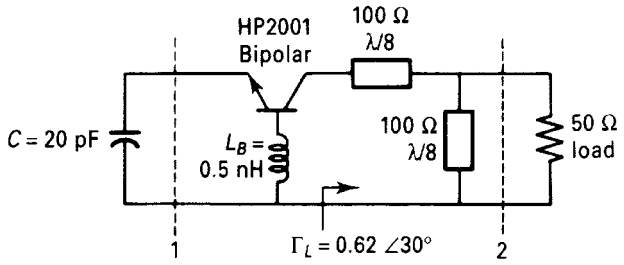


Figure 5-28 Oscillator example at 2 GHz.

$$\Gamma_L = 0.62\ \angle 30°$$

$$S'_{11} = 1.18\ \angle 173°$$

Thus, a resonating capacitance of G = 20 pF resonates the input port. In a YIG-tuned oscillator, this reactive element could be provided by the high-Q YIG element. For a dielectric-resonator oscillator (DRO), the puck would be placed to give $\Gamma_G \approx 1.0/\angle -173°$.

Another two-port design procedure is to resonate the Γ_G port and calculate S'_{22} until $|S'_{22}| > 1$, then design the load port to satisfy. This design procedure is summarized in Figure 5-29.

An example using this procedure at 4 GHz is given in Figure 5-30 using an AT41400 silicon bipolar chip in the common-base configuration with a convenient value of base and emitter

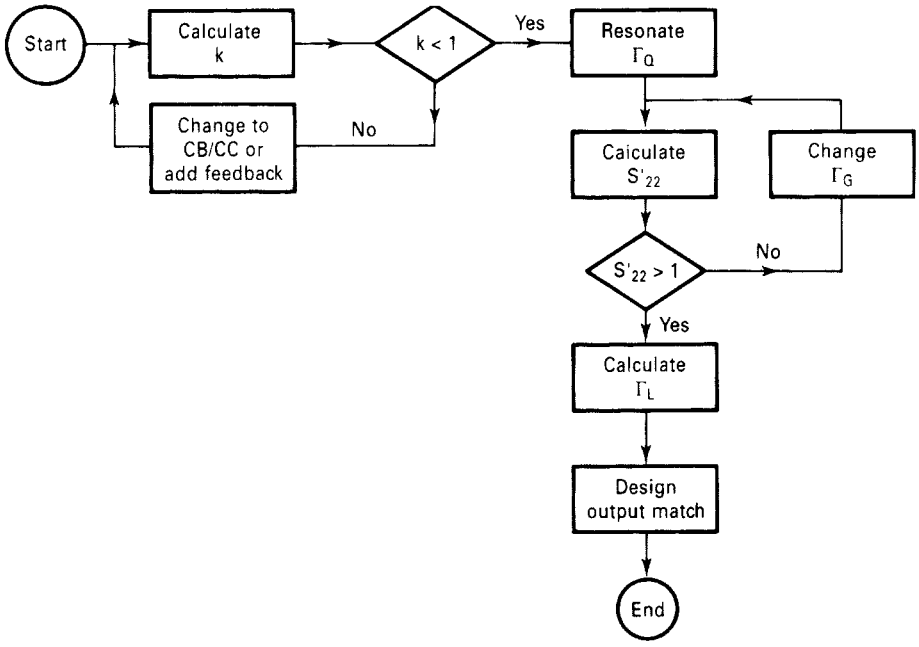


Figure 5-29 Oscillator design flowchart.

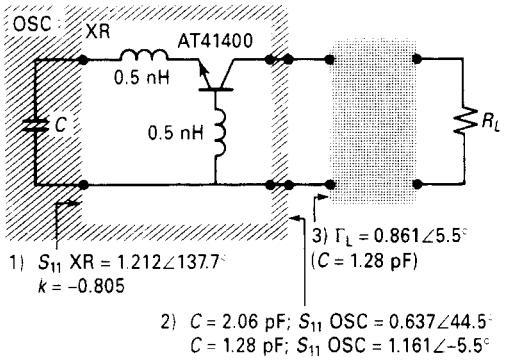


Figure 5-30 A 4-GHz lumped-resonator oscillator using the AT41400.

inductance of 0.5 nH. The feedback parameter is the base inductance, which can be varied if needed.

Since a lossless capacitor at 4 GHz of 2.06 pF gives $\Gamma_G = 1 - 0 \angle -137.7°$, this input termination is used to calculate S'_{22}, giving $S'_{22} = 0.637 \angle 44.5°$. This circuit will oscillate into any passive load. Varying the emitter capacitor about 20° on the Smith Chart to 1.28 pF gives $S'_{22} = 1.16 \angle -5.5°$, which will oscillate into a load of $\Gamma_L = 0.861 \angle 5.5°$. The completed lumped-element design is given in Figure 5-31.

5-5-3 Ceramic-Resonator Oscillators

An important application for a new class of resonators called ceramic resonators (CRs) has emerged for wireless applications. The CRs are similar to rigid coaxial cable, where the center conductor is connected at the end to the outside of the cable. These resonators are generally operating in quarter-wavelength mode and their characteristic impedance is approximately 10 Ω. Because their coaxial assemblies are made from a high-ε low-loss material with good silver plating throughout, the electromagnetic field is internally contained and therefore results in very little radiation. These resonators are therefore ideally suited for high-Q, high-density oscillators. The typical application for this resonator is VCOs ranging from not much more than 200 MHz up to about 3 or 4 GHz. At these high frequencies, the mechanical dimensions of the resonator become too tiny to offer any advantage. One of the

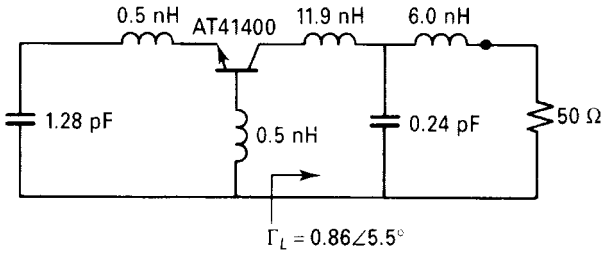


Figure 5-31 Completed lumped-resonator oscillator (LRO).

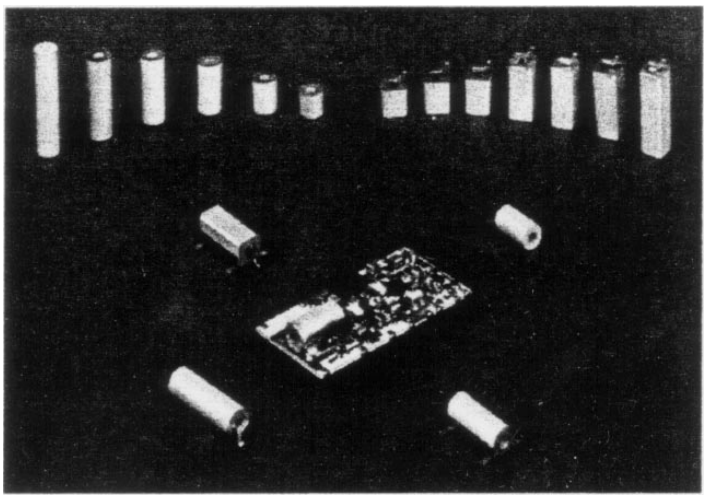


Figure 5-32 Ceramic resonators.

principal requirements is that the physical length is considerably larger than the diameter. If the frequency increases, this requirement can no longer be met.

Manufacturers supply ceramic resonators on a prefabricated basis. Figures 5-32 and 5-33 show the standard round/square packaging available and the typical dimensions for a ceramic resonator. Figure 5-34 shows a ceramic-resonator oscillator built on a ceramic substrate. The available material has a dielectric constant of 88 and is recommended for use in the 400–1500-MHz range. The next higher frequency range (800 MHz to 2.5 GHz) uses an ε of 38, while the top range (1–4.5 GHz) uses an ε of 21. Given the fact that ceramic resonators are prefabricated and have standard outside dimensions, the following quick calculation applies:

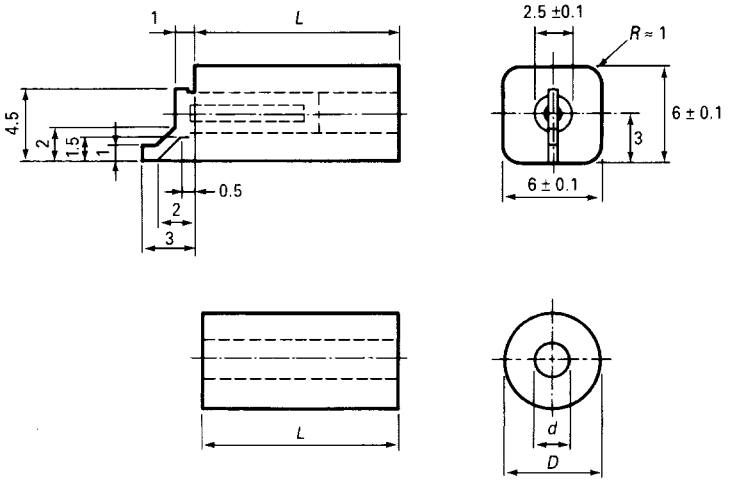


Figure 5-33 Standard round/square ceramic-resonator packaging and dimensions.

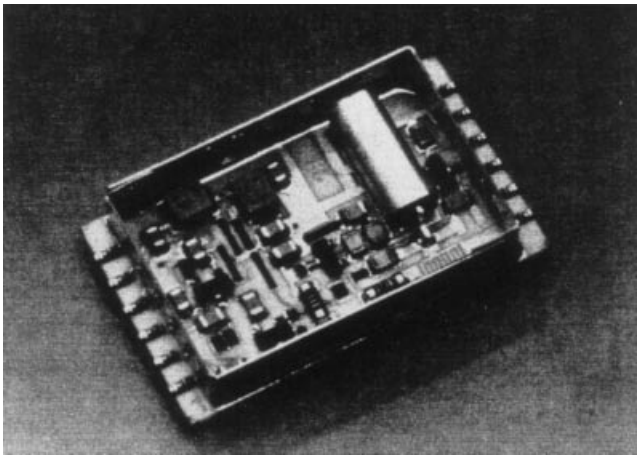


Figure 5-34 Modern ceramic-resonator oscillator for handheld cellular phones.

Relative dielectric constant of resonator material	$\varepsilon_r = 21$	$\varepsilon_r = 38$	$\varepsilon_r = 88$
Resonator length in millimeters	$l = \frac{16.6}{f}$	$l = \frac{12.6}{f}$	$l = \frac{8.2}{f}$
Temperature coefficient (ppm/°C)	10	6.5	8.5
Available temperature coefficients	–3 to +12	–3 to +12	–3 to +12
Typical resonator Q	800	500	400
Frequency range	1–4.5 GHz	800–2500 MHz	400–1500 MHz

Calculation of Equivalent Circuit. The equivalent parallel-resonant circuit afforded by a ceramic resonator has a resistance at resonance of

$$R_p = \frac{2(Z_0)^2}{R^* l} \tag{5-67}$$

where Z_0 = characteristic impedance of the resonator
l = mechanical length of the resonator
R^* = equivalent resistor due to metallization and other losses

As an example, one can calculate

$$C^* = \frac{2\pi\varepsilon_0\varepsilon_r}{\log_e(D/d)} = 55.61 \times 10^{-12}\,\frac{\varepsilon_r}{\log_e(D/d)} \tag{5-68}$$

and

$$L^* = \frac{\mu_r \mu_0}{2\pi} = \log_e\left(\frac{D}{d}\right) = 2 \times 10^{-7} \log_e\left(\frac{D}{d}\right) \tag{5-69}$$

$$Z_0 = 60\ \Omega \frac{1}{\sqrt{\varepsilon_r}} \log_e\left(\frac{D}{d}\right)$$

$$= 60 \frac{1}{\sqrt{88}} \log_e\left(\frac{6}{2.5}\right)$$

$$= 5.6\ \Omega \tag{5-70}$$

A practical example for $\varepsilon_r = 88$ and 450 MHz is

$$C_p = \frac{C^* l}{2} = 49.7\ \text{pF} \tag{5-71}$$

$$L_p = 8L^* l = 2.52\ \text{nH} \tag{5-72}$$

$$R_p = 2.5\ \text{k}\Omega \tag{5-73}$$

Figure 5-35 shows the schematic of a ceramic-resonator-based oscillator. Figures 5-36 and 5-37 show the simulated and measured phase noise of the oscillator.

By using ceramic-resonator-based oscillators in conjunction with miniature synthesizer chips, it is possible to build extremely small phase-locked-loop (PLL) systems for cellular telephone operation. Figure 5-38 shows one of the smallest currently available PLL-based synthesizers manufactured by Synergy Microwave Corporation. Because of the high-Q resonator, these types of oscillators exhibit extremely low phase noise. Values of better than –150 dBc/Hz at 1 MHz off the carrier are achievable. The ceramic resonator reduces the oscillator's sensitivity to microphonic effects and proximity effects caused by other components.

5-5-4 Using a Microstrip Inductor as the Oscillator Resonator*

A high-Q microstrip resonator [3–5] can be used to improve an oscillator's phase-noise performance. The following section, based on information in the Philips Semiconductors Application Note AN1777, describes the application of such a resonator in a BJT differential oscillator.

It can be seen from Eq. (5-56) that the phase noise is proportional to the term

$$\frac{1}{4Q_L^2} \tag{5-74}$$

This means that for any 10% increase in Q, we get something like 20% improvement in phase noise, which makes it most desirable to increase the Q. One way to reduce the loading of the

*Based on portions of the Philips Semiconductors/Signetics RF Communications Products Application Note AN1777, "Low-Voltage Front End Circuits: SA601, SA620," August 20, 1997. Used with permission.

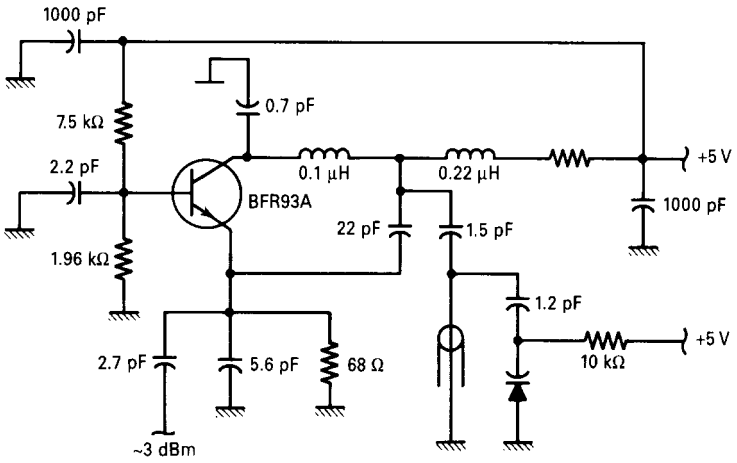


Figure 5-35 Schematic of a high-performance ceramic-resonator-based oscillator that can be used from 500 MHz to 2 GHz.

resonator is to tap the device down on the resonator. The underlying oscillator circuit is based on the differential type of oscillator found in ICs.

Increasing Loaded Q. Figure 5-39 shows the oscillator section of the Philips SA620 low-voltage front-end IC with a conventional second-order parallel-tuned tank circuit configured as the external resonator.

From the basic equations for parallel resonance, the resonator's loaded Q is given by

$$Q_L = R_P \left(\frac{C_T}{L_T} \right)^{1/2} \tag{5-75}$$

where

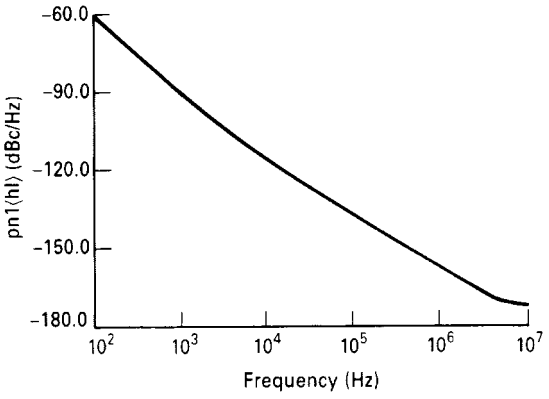


Figure 5-36 Simulated phase noise of an *npn* bipolar 1-GHz ceramic-resonator-based oscillator.

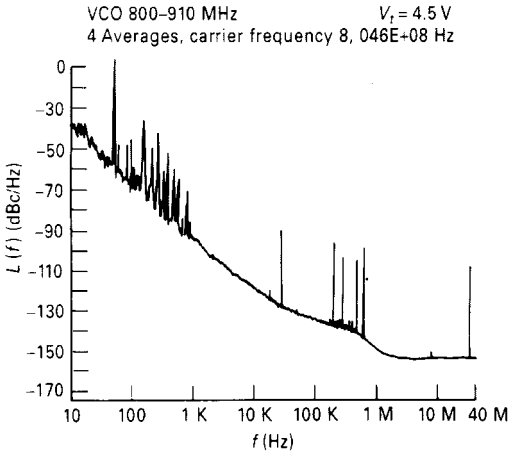


Figure 5-37 Measured phase noise of a ceramic-resonator-based oscillator.

$$R_P = R_T \| R_C = \left[\left(\frac{1}{R_T}\right) + \left(\frac{1}{R_C}\right)\right]^{-1} \tag{5-76}$$

represents the net shunt tank resistance appearing across the network at resonance. R_T represents losses in the inductance L_T and capacitance C_T (almost always dominated by inductor losses), and R_C is the active circuit's load impedance at resonance. Improving the quality of the tank components L_T and C_T (i.e., improving their Q) will increase R_T, but since R_C is low in the case of the SA620, the resulting increase in R_P is relatively small. We can

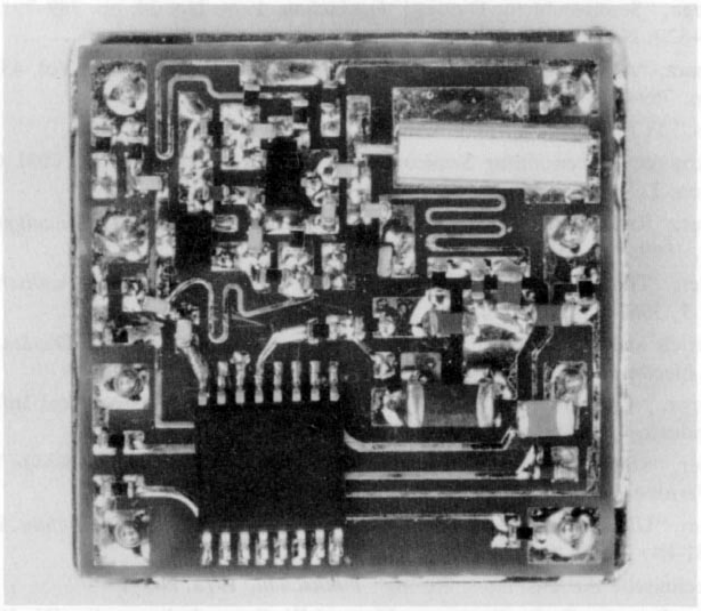


Figure 5-38 Miniature PLL-based synthesizer manufactured by Synergy Microwave Corporation.

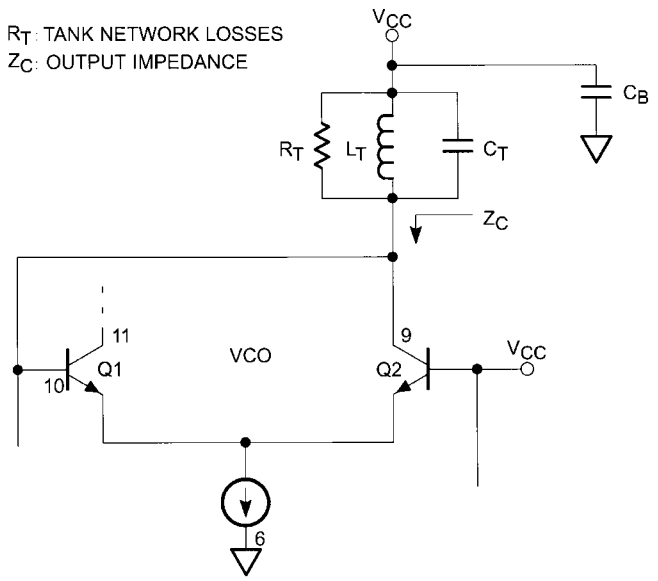


Figure 5-39 Oscillator portion of Philips SA620 front-end IC with external resonator.

increase R_P by decoupling Z_C from the tank circuit (note that R_C is the real part of Z_C at resonance). Decoupling this impedance by using either tapped-L or tapped-C tank configurations is possible. Inspection of the circuit shows that dc biasing is necessary for pins 9 and 10 (osc1 and osc2, respectively), so a tapped-C approach would require shunt-feeding these pins to V_{CC}. Thus, the most practical way is to employ a tapped-L network.

Figure 5-40 shows the basic tapped-L circuit. The effective multiplied shunt impedance at resonance across C_T is given by

$$\hat{R}_P = R_P\left[\left(\frac{L_{T2}}{L_{T1}}\right)+1\right]^{1/2} \tag{5-77}$$

with a corresponding increase in loaded Q:

$$Q_L = Q_L\left[\left(\frac{L_{T2}}{L_{T1}}\right)+1\right]^{1/2} \tag{5-78}$$

Feedback is caused by an RF voltage appearing across L_T, coupled through bypass capacitor C_B and its ground return to the cold end of the current source at pin 6. The significance of this path increases with frequency. At 900 MHz, it becomes very critical in obtaining stable oscillations and must be kept as short as possible. Attention to layout detail is pivotal here!

High-Q Microstrip Inductor. Realizing a stable tapped-L network using lumped surface-mount inductors is impractical due to their low unloaded Q and finite physical size. Another approach uses a high-Q short microstrip inductor. The conventional approach to microstrip resonators treats them as a length of transmission line terminated with a short that

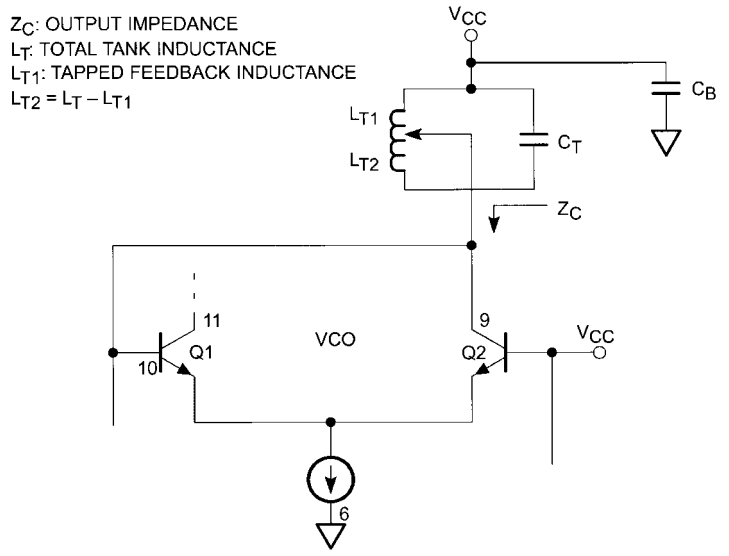


Figure 5-40 Oscillator portion of SA620 IC with tapped-microstrip resonator.

reflects back an inductance impedance that simulates a lumped inductor. The approach presented here deviates from this technique by specifically exploiting the very high unloaded Q attainable with short microstrip inductors, where we design for a large C/L ratio by making the microstrip a specific, but short, length. Resonance is achieved by replacing the short with whatever capacitance is needed for proper resonance. One equation for the inductance of a strip of metal having a length ℓ (mils) and width w (mils) is

$$L = [5.08\text{E} - 3]\,\ell\left[\ln\left(\frac{\ell}{w}\right) + 1.193\left(\frac{w}{\ell}\right)\right] \qquad \text{[3] (5-79)}$$

where L is in nH/mil. This technique results in a much shorter strip of metal for a given inductance. One intriguing property of microstrip inductors is that they are capable of very high unloaded Q values. An equation describing the quality factor for a "wide" microstrip is

$$Q_C = 0.63h[\sigma f_{\text{GHz}}]^{1/2} \qquad \text{[4] (5-80)}$$

where h is the dielectric thickness in centimeters, σ is the conductivity in S/m, and f is frequency in GHz. This equation predicts an unloaded Q exceeding 700 when silver-coated copper is used. Also, wide microstrip lines are defined as those whose strip width w to height h ratio is approximately greater than 1; that is, $w/h > 1$. The parallel capacitor formed by the metal strip over the ground plane should also be included and may be calculated by the classic formula and included in the total needed to resonate with the inductance result found from Eq. (5-79). This "strip" capacitance is given by

$$C_S = K\varepsilon_0\varepsilon_r\left(\frac{w\ell}{h}\right) \qquad \text{(5-81)}$$

where C_S is in farads, $\varepsilon_0 = 8.86\text{E} -12$ (F/cm) is the permittivity of free space, and ε_r is the relative dielectric constant of the substrate. A "fringe factor," K, is included to account for necessary fringing (estimated to be 2–15%). These equations are meant to provide insight into circuit behavior and should, therefore, be applied cautiously to specific applications.

UHF VCO Using the Tapped-Inductor Differential Oscillator at 900 MHz. The basic electrical circuit of this configuration is shown in Figure 5-41. Feedback occurs when sufficient RF voltage develops across the tap inductances L_{T1} and L_{T2} (due to tap 1 and tap 2, respectively); note that inductance is reckoned from the cold end of the tank at node A. Taps 1 and 2 should be as close as possible to pins 9 and 10, respectively. Also node A (cold end of tank) and node B (pin 6) must be as physically close together and exhibit as low an impedance as possible to discourage parasitic oscillations. This "inner loop," composed of L_{T2}, the inductance of connections between taps 1 and 2 and pins 9 and 10, respectively, and "stray" low-Q inductance between nodes A and B creates a parasitic loop that will support oscillation that no longer depends on the full tapped-L tank circuit. Oscillation based on this low-Q loop typically occurs well above 1.2 GHz and has been observed as high as 1.6 GHz, and exhibits very poor phase noise performance.

Another factor also affects this unwanted parasitic oscillation. As the full tank circuit's unloaded Q decreases, the circuit becomes conditionally stable and will eventually favor the parasitic loop exclusively. This occurs because the loaded Q of the tapped microstrip affects the magnitude of the RF voltage appearing across the entire tank. Thus, feedback at taps 1 and 2 is reduced as the Q decreases. This increases the likelihood of the parasitic inner loop controlling the oscillator, since its feedback voltage is largely independent of the microstrip tank Q.

Equation (5-80) shows that microstrip inductors are capable of very high unloaded Q values. This depends on a number of physical factors: frequency, dielectric thickness and

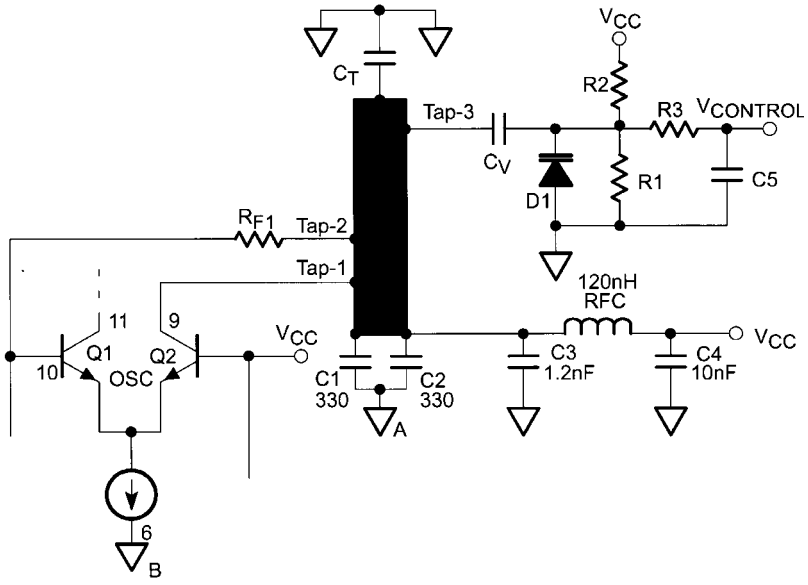


Figure 5-41 Differential oscillator with tapped-microstrip resonator.

quality, and the skin-depth conductance of the metal strip itself. For example, at 900 MHz, the unloaded Q of a silver-coated microstrip line is in the vicinity of 750. Coating the same line with lead (sloppy soldering can do it!) reduces the Q to less than 230. If the unloaded Q is too low, the loaded Q also drops. The circuit becomes conditionally stable and now may oscillate either at the higher parasitic-loop frequency or the desired lower frequency at which the entire microstrip tank resonates. Components should be connected by narrow traces closely connected to component leads to avoid soldering on the microstrip proper.

R_{F1} introduces necessary losses to control the inner-loop parasitic oscillation. Its size should be made as small as possible consistent with stability, startup, and good phase-noise performance. Since R_{F1} also decreases feedback, the oscillator's output falls as R_{F1} increases. Typically, experimental values from 4 to 30 Ω have proved sufficient.

Stability as a function of V_{CC} is a good way to assess conditional stability. Provided the microstrip tank has a high loaded Q, the oscillator's stability should be independent of V_{CC} down to less than 2.5 V or so. When loaded Q decreases sufficiently to favor the inner-loop parasitic, conditional stability will occur. This can readily be observed by increasing V_{CC} from about 1.0 V and noting whether the VCO frequency jumps back and forth unpredictably. Mixer and LO-port loading can also affect this condition and should be considered.

Dimensions for the microstrip resonator are based on several criteria. The strip must be "wide": Its width-to-height ratio must be on the order of 1 or greater. Minimum strip length seems to be about 200 mils for 62-mil-thick board. L/C ratios somewhere about 500 have yielded quite good results. Note that we usually specify the "L/C ratio" because it is always greater than 1, even though in a parallel-resonant circuit it appears in the Q equation as C/L. Thus, decreasing the "L/C ratio" by making the microstrip shorter helps the inductor's loaded Q, and the circuit's loaded Q_L, increase [see Eq. (5-75)]. However, the effect of the lower Q of C_T, even when using a high-Q surface-mount type, decreased the expected larger increase in the circuit's net loaded Q. Thus, the expected increase in Q_L may not be fully realizable. Assuming a 62-mil-thick board with FR5 dielectric, a strip 50 mils by 300 mils gave very good experimental results, where C_T was about 4.3 pF, with oscillation occurring from about 950 to 1000 MHz with various test boards. A shorter inductor designed to increase the C/L ratio requiring 7.5 pF experimentally resulted in only about 1–2-dB phase-noise improvement at 950 MHz. The reason for this is that as the microstrip inductor degenerates to a plain inductor, all of its advantages are lost.

Tap 1 may be anywhere between the cold end of the tank (where C_1 and C_2 are located) and tap 2. Tap 2 yields good results at about one-third the strip length. Making it too close to the cold end in an effort to increase Q_L will result in loss of control over the inner-loop parasitic. To keep Q_L as high as possible, C_T should be a high-Q surface-mount type. Differences greater than 5 dB in SSB phase noise have been observed between a generic NP0 surface-mount capacitor and a high-Q NP0 surface-mount capacitor.

Tap 3 can be at the end of the microstrip when C_T is connected. However, the tuning diode inevitably will cause a decrease in overall loaded Q, typically resulting in a phase-noise increase of as much as 5 dB. Moving it some distance down from the high end of the tank will decrease this effect and yield better results. A good starting point is about one-quarter down, or three-quarters of the length from the cold end of the tank. C_1 and C_2 are paralleled lower-value capacitors to yield a better low-impedance ground return to pin 6 (node B) since relatively large RF currents flow through them. Note that as Q_L increases, the peak circulating RF current will also increase, as will the RF voltage at the hot end of the microstrip. At 900 MHz, good results have been obtained when both are about 300 pF. The RF choke was chosen to be approximately series-resonant around 900 MHz, and constitutes the series feed path

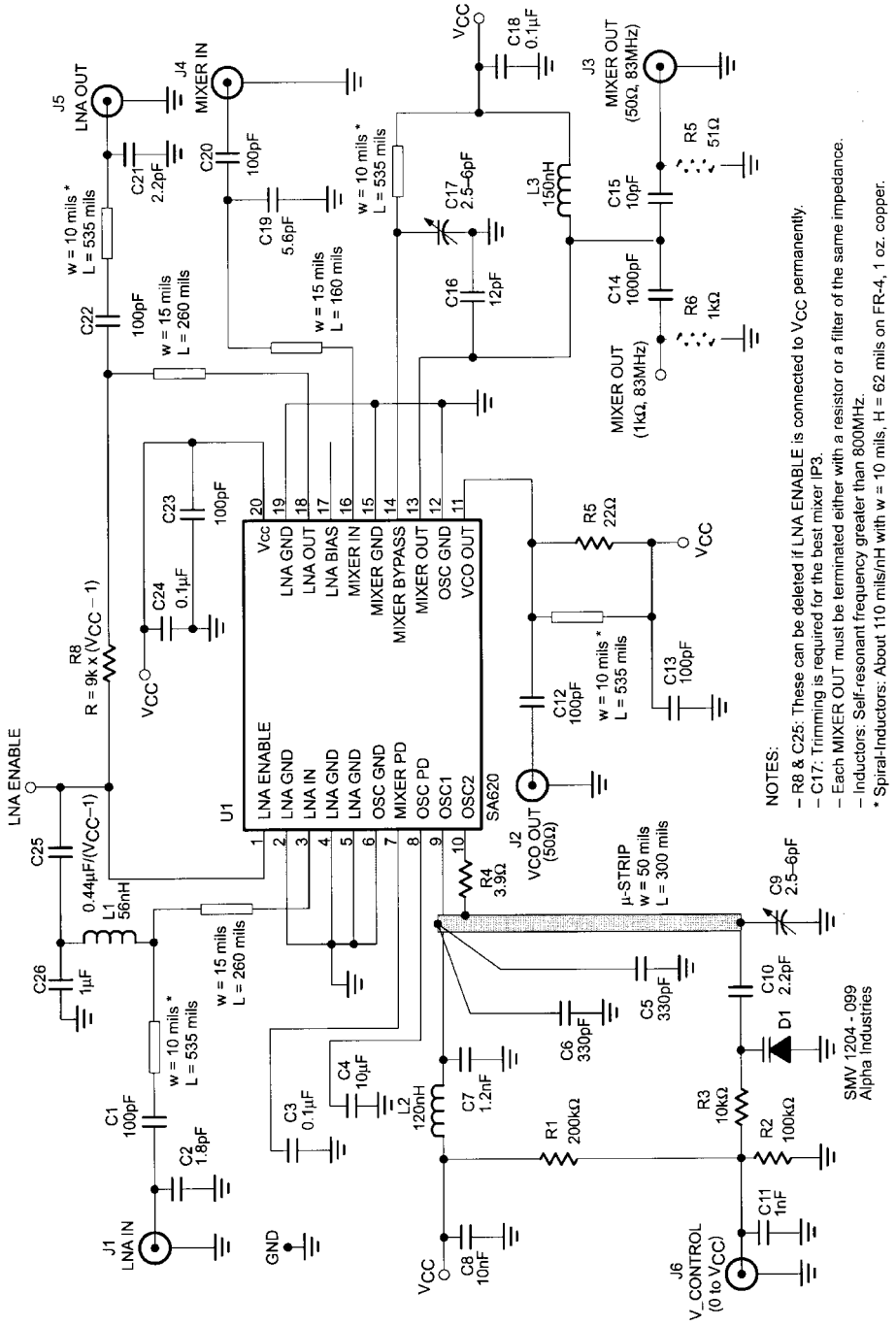


Figure 5-42 Philips SA620 application board schematic showing the tapped-microstrip resonator.

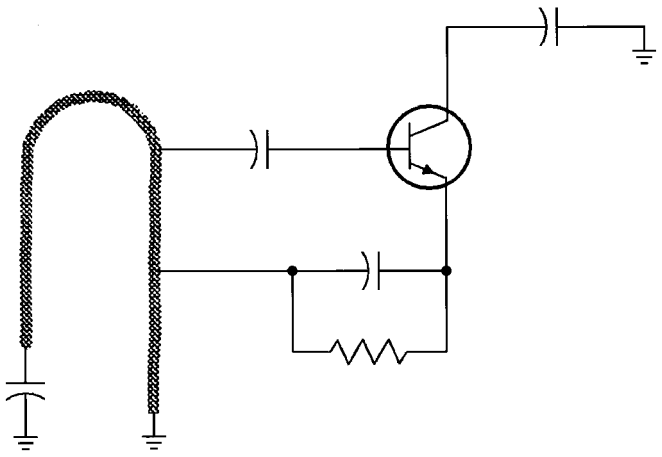


Figure 5-43 "Walking stick" microstrip Hartley resonator as used by Plextek in the 800–1500-MHz range. Below 800 MHz, the resonator's physical size becomes impractical; above 1500 MHz, the transistor's parasitic reactances make the Hartley configuration problematic.

for dc biasing. It may be possible to neglect this component entirely, provided that the PC-board connection to the cold end of the tank is very close to RF ground. Finally, note that shielding of the entire microstrip may be necessary to meet FCC Part 15 emission limitations (where applicable). Figure 5-42 shows one application of the tapped-microstrip oscillator with the Philips SA620 front-end IC.

5-5-5 Hartley Microstrip Resonator Oscillator

G. Smithson and S. Fitz of Plextek Communications Technology Consultants have reported on the construction of Hartley microstrip-resonator oscillators using thick-film techniques on ceramic (alumina) substrates for high-volume wireless products. The resonator in one such oscillator (Figure 5-43) exhibited an unloaded Q of approximately 80. Table 5-2 summarizes its phase noise versus tuning range with different tuning-diode configurations.

5-5-6 Crystal Oscillators

Most wireless applications do not require ultra-high-frequency stability that mandates the use of a primary standard, such as one based on cesium, or a secondary standard, such as one based on rubidium. For very stable oscillators at fixed or only slightly variable frequencies up to a few hundred megahertz, oscillators with piezoelectric resonators are generally used. A piezoelectric material transduces mechanical stress (deformation) to

Table 5-2 Achieved oscillator phase noise versus tuning range

Tuning Range, 0.5–3 V	Tuning Range, 1–3 V	Phase Noise at 20-kHz Offset
23%	16%	–100 dBc/Hz
16%	11%	–103 dBc/Hz
10%	8%	–106 dBc/Hz

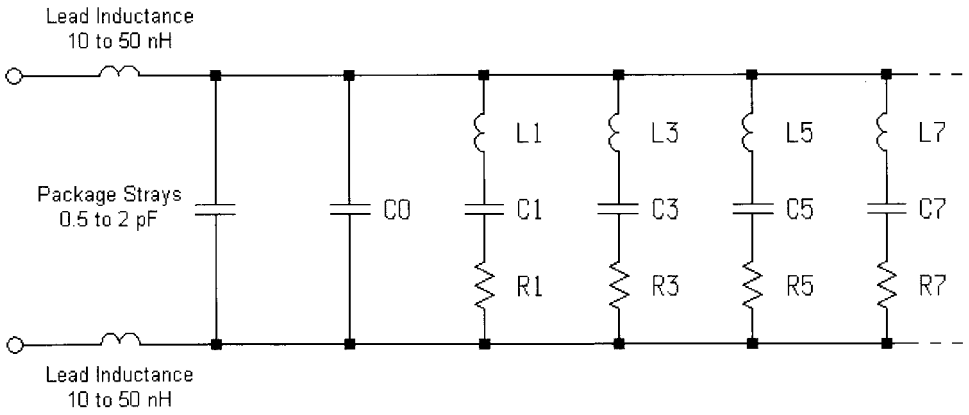


Figure 5-44 Electrical equivalent of a piezoelectric crystal resonator, showing fundamental (1) and overtone (3, 5, 7, . . .) resonances. (Courtesy of Roger Clark, Vectron International.)

Table 5-3 Approximate crystal parameters

Crystal Q Approximations ($F_{\text{resonance}}$ in MHz)

Nonprecision Fundamental Crystals:

Worst: $Q = \dfrac{2 \times 10^6}{F}$ Better: $Q = \dfrac{5 \times 10^6}{F}$

Precision Crystals Processed for High Q:

SC: $Q = \dfrac{12 \times 10^6}{F}$ AT Cut: $Q = \dfrac{10 \times 10^6}{F}$

Crystal Resistance (r) for 10-MHz, Third-Overtone Example Using Precision Crystal Approximation (F in MHz and C_1 in pF)

$$r = \frac{2\pi F C_1}{Q}$$

$$r = 6.964$$

Typical Crystal Parameters

10-MHz Fundamental: $C_1 = 0.02$ pF, $C_0 = 3.6$ pF, $Q = 350{,}000$, $r = 2.5\ \Omega$, $L = 0.012665$ H
50-MHz Third-Overtone: $C_1 = 0.0026$ pF, $C_0 = 4.212$ pF, $Q = 100{,}000$, $r = 13\ \Omega$, $L = 0.003897$ H
100-MHz Fifth-Overtone: $C_1 = 0.0005$ pF, $C_0 = 2.25$ pF, $Q = 80{,}000$, $r = 40\ \Omega$, $L = 0.005066$ H
155-MHz High-Frequency Fundamental: $C_1 = 0.005$ pF, $C_0 = 1.5$ pF, $Q = 20{,}000$, $r = 12\ \Omega$, $L = 0.002108$ H

Source: Roger Clark, Vectron International.

electrical stress (voltage), and vice versa. In noncritical applications up to 10 MHz, a piezoelectric ceramic may be used. The Q and temperature stability of such materials are relatively poor, however, so quartz is the resonator material of choice for temperature-stable, low-noise piezoelectric oscillators.

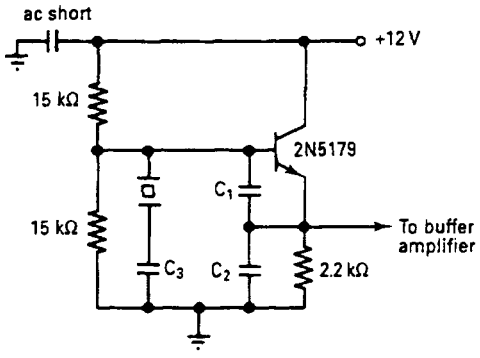
Just as an electrical resonator consisting of distributed L and C exhibits multiple electrical resonances, a physical piece of piezoelectric material, depending on its shape, exhibits multiple mechanical resonances. *Fundamental* resonance occurs at the frequency at which the crystal is one-half wavelength long; *overtone* resonances occur near odd multiples (3, 5, 7, etc.) of the fundamental. Figure 5-44 shows the electrical equivalent of a crystal. Table 5-3 presents examples of approximate parameters for several crystal types. Appropriate shaping and placement of the crystal's connecting electrodes, combined with suitable oscillator circuit design and adjustment, allow the selection or ehancement of a particular resonance for excitation.

Figure 5-45 shows a typical crystal oscillator circuit configuration, and the associated table supplies recommended values for the elements. This is in line with the chip manufacturer's recommendations for built-in oscillator circuits. Figure 5-46 shows how this oscillator's input impedance varies around resonance; Figure 5-47 shows its simulated phase noise. Placing a small resistance in series with the crystal (Figure 5-48) allows us to use the crystal as a high-Q filter, substantially improving the oscillator's spectral purity relative to that obtainable at the transistor emitter (Figure 5-49).

Some applications demand ultra-low-phase-noise crystal oscillators. Figures 5-50 and 5-51 show an appropriate circuit and its phase noise with and without the current-supplying hot carrier diode (HCD) used for noise reduction.

5-5-7 Voltage-Controlled Oscillators

A schematic of a ceramic-resonator-based VCO (Figure 5-35) has already been shown in the section on ceramic-resonator oscillators. To tune the oscillator within the required range,



	3 MHz	6 MHz	10 MHz	20 MHz	30 MHz
C_1 (pF)	330	270	180	82	43
C_2 (pF)	430	360	220	120	68
C_3 (pF)	39	43	43	36	32
Crystal C_L (pF)	32	32	30	20	15

C_3 is the approximate value needed to set oscillator on frequency using indicated crystal C_L.

Figure 5-45 A 3–30-MHz Colpitts crystal oscillator.

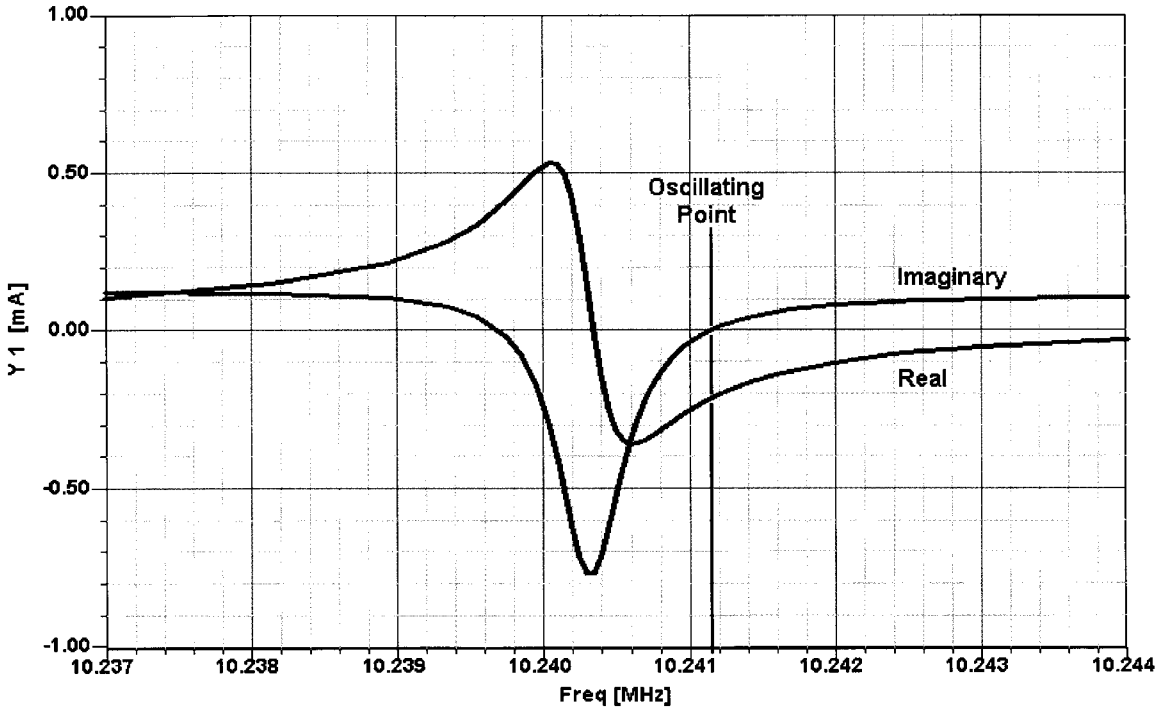


Figure 5-46 As with every oscillator, the crystal oscillator's input impedance is negative and real at resonance. This graph plots the real and imaginary components of the test voltage injected by the Oscillator Design Aid tool in Ansoft Serenade 8.0. Compare this graph with Figure 5-7, which plots an oscillator's input impedance with the resonator absent.

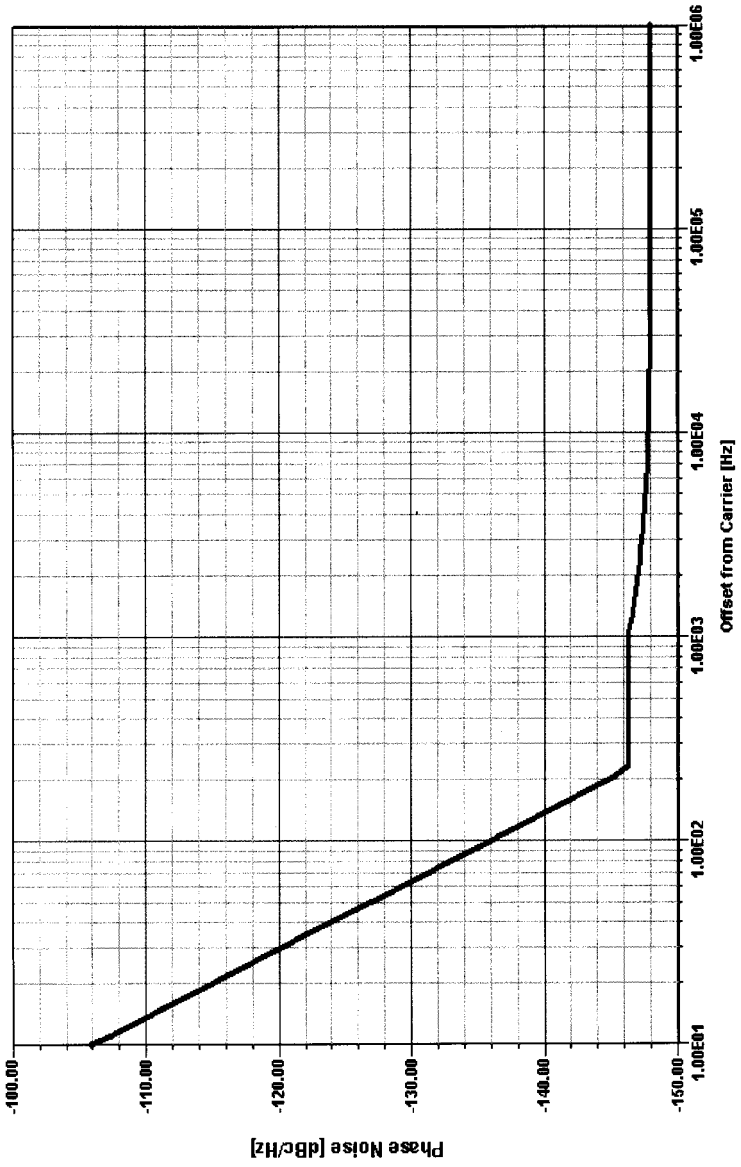


Figure 5-47 Simulated phase noise of the crystal oscillator.

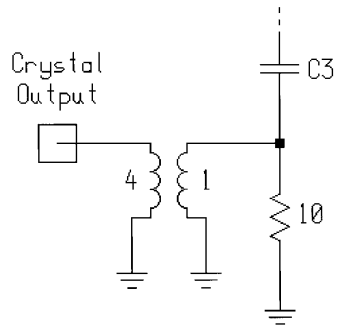


Figure 5-48 This alternative output tap uses the crystal's high selectivity for improved harmonic reduction.

tuning diodes are used. These diodes are often called *varactors* or *voltage-sensitive diodes.* By way of approximation, we can use the equation

$$C = \frac{K}{(V_R + V_D)^n} \tag{5-82}$$

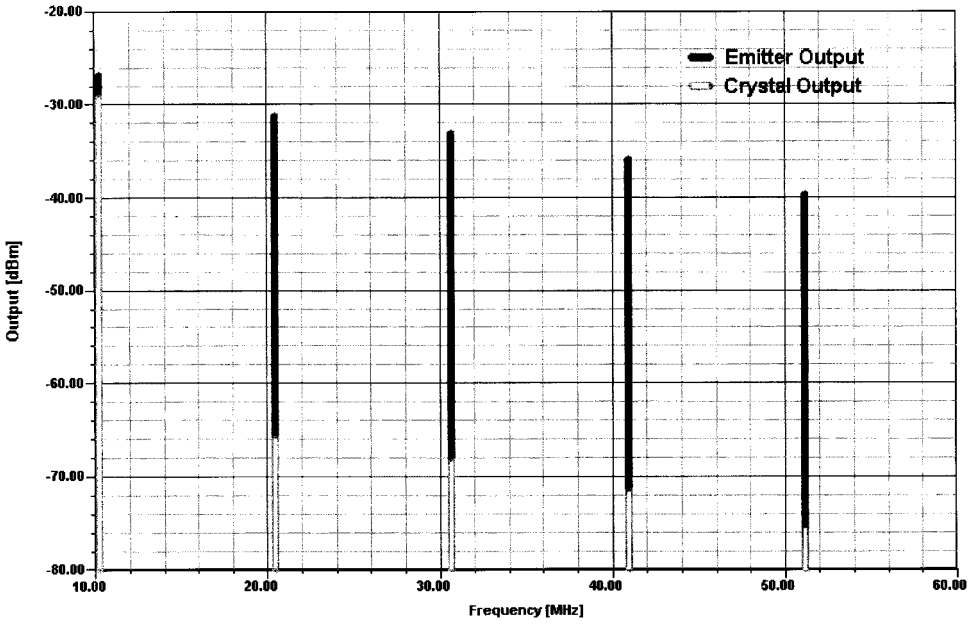


Figure 5-49 Comparison of the crystal oscillator's simulated output spectrum as obtained at the transistor emitter (dark gray, fundamental –26.7 dBm) and crystal (light gray, fundamental –29.1 dBm). At the emitter, the second harmonic is down only 4.3 dB relative to the carrier; at the crystal, the second harmonic is down 36 dB relative to the carrier.

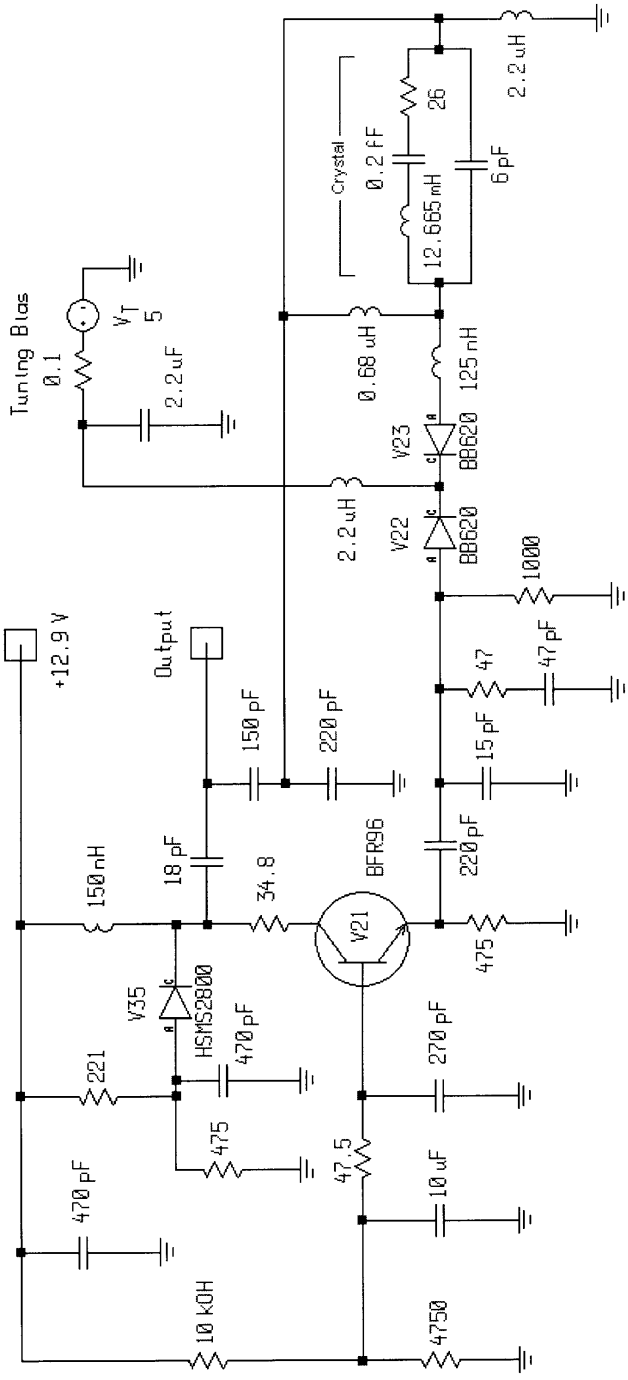


Figure 5-50 A 100-MHz VCXO suitable for ultra-low-phase-noise applications. The HSMS2800 hot-carrier diode (HCD) provides noise reduction.

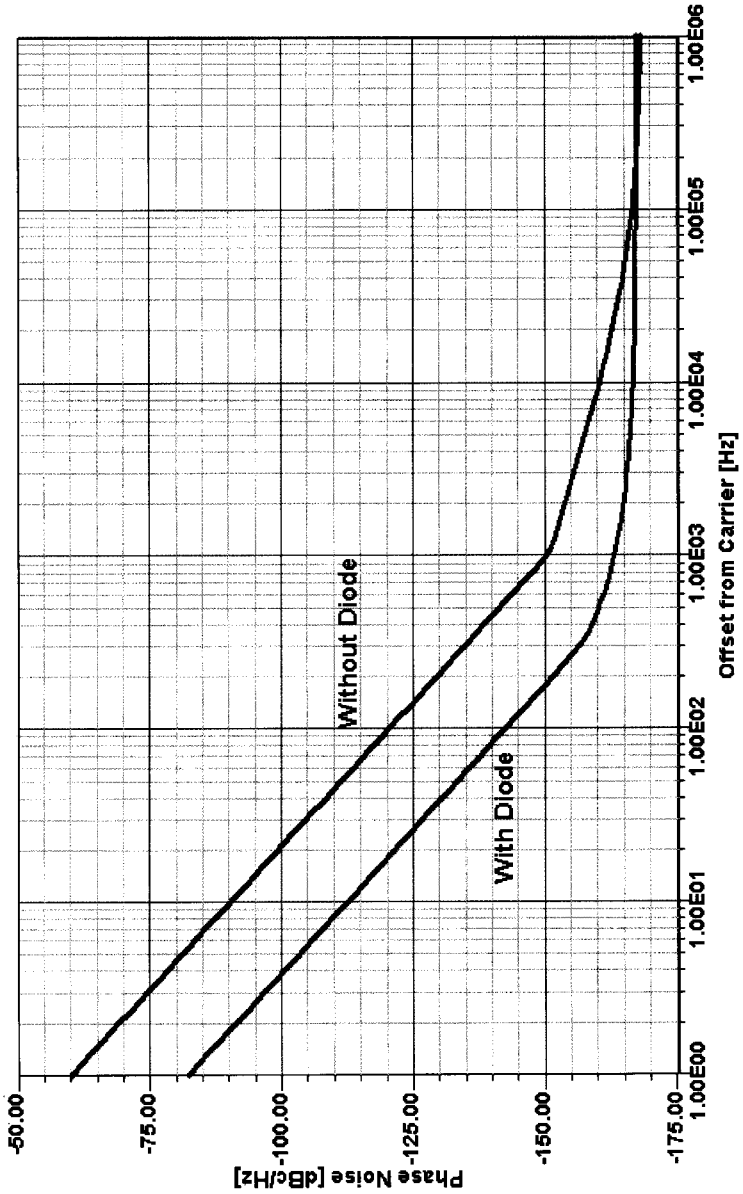


Figure 5-51 Phase noise of the VCXO with and without the noise-reduction diode.

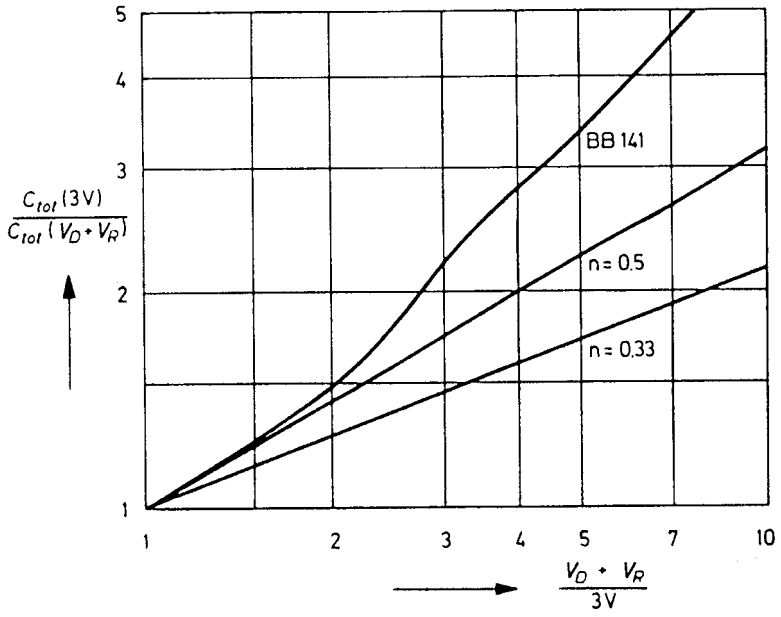


Figure 5-52 Capacitance–voltage characteristic for an alloyed capacitance diode ($n = 0.33$), a diffused capacitance diode ($n = 0.5$), and a wide-range tuning diode (BB141).

wherein all constants and all parameters determined by the manufacturing process are contained in K. The exponent is a measure of the slope of the capacitance/voltage characteristic and is 0.5 for alloyed diodes, 0.33 for single-diffused diodes, and (on average) 0.75 for tuner diodes with a hyperabrupt *pn* junction [6, 7]. Figure 5-52 shows the capacitance–voltage characteristics of an alloyed, a diffused, and a tuner diode.

Recently, an equation was developed that, although purely formal, describes the practical characteristic better than Eq. (5-82):

$$C = C_0 \left(\frac{A}{A + V_R} \right)^m \tag{5-83}$$

where C_0 is the capacitance at $V_R = 0$, and A is a constant whose dimension is the volt. The exponent m is much less dependent on voltage than the exponent n in Eq. (5-81).

The operating range of a capacitance diode or its useful capacitance ratio,

$$\frac{C_{\max}}{C_{\min}} = \frac{C_{\text{tot}}(V_{R,\min})}{C_{\text{tot}}(V_{R,\max})} \tag{5-84}$$

is limited by the fact that the diode must not be driven by the alternating voltage superimposed on the tuning voltage, either into the forward mode or the breakdown mode. Otherwise, rectification would take place, shifting the bias of the diode and considerably affecting its figure of merit.

There are several manufacturers of tuning diodes. Motorola is a typical supplier in this country; Siemens of Germany and Philips also provide good diodes. Table 5-4 and Figure 5-53 contain information for tuning diodes useful for our applications.

5-5-8 Diode-Tuned Resonant Circuits

Tuner Diode in Parallel-Resonant Circuit. Figures 5-54 and 5-55 illustrate two basic circuits for the tuning of parallel resonant circuits by means of capacitance diodes. In the circuit diagram of Figure 5-54, the tuning voltage is applied to the tuning diode via the bias resistor, R_B. Series-connected from the tuning diode to the top of the tank circuit is capacitor C_S, which completes the circuit for alternating current but isolates the cathode of the tuner diode from the coil and thus from the negative terminal of the tuning voltage. Moreover, a fixed parallel capacitance C_P is provided. The decoupling resistor preceding the bias resistor is large enough to be disregarded in the following discussion. Since for high-frequency purposes the biasing resistor is connected in parallel with the series capacitor, it is transformed into the circuit as an additional equivalent shunt resistance R_c. For the parallel loss resistance transformed into the circuit, we have the expression

$$R_c = R_B \left(1 + \frac{C_{\text{tot}}}{C_S}\right)^2 \tag{5-85}$$

Table 5-4 Example tuning diodes (Siemens)

	Maximum Ratings		Characteristics (T_A = 25 °C)								
Type	V_R (V)	I_F (mA)	C_T at (pF)	V_R (V)	C_T (pF)	V_R (V)	C_{ratio}	I_R (nA)	V_R (V)	Package	Chip Code
BB439	30	20	29	3	5.0	25	5.8	≤20	28	SOD-323	Q
BB535	30	20	18.7	1	2.1	28	8.9	≤10	30	SOD-323	A
BB545	30	20	20	1	2.0	28	10.0	≤10	30	SOD-323	C
BB639	30	20	38.3	1	2.65	28	14.5	≤10	30	SOD-323	K
BB639C	30	20	39	1	2.55	28	15.3	≤10	30	SOD-323	B
BB640	30	20	69	1	3.05	28	22.6	≤10	30	SOD-323	N
BB804 (dual)	20	50	44.75	2	26.15	8	1.71	≤20	16	SOT-23	E
BB814 (dual)	20	50	44.75	2	20.8	8	2.15	≤20	16	SOT-23	O
BB833	30	20	9.3	1	0.75	28	12.4	≤20	30	SOD-323	F
BB835	30	20	9.1	1	0.62	28	14.7	≤10	30	SOD-323	G
BB914 (dual)	20	50	43.75	2	18.7	8	2.34	≤20	16	SOT-23	P
BBY51 (dual)	7	20	5.30	1	3.10	4	1.75	≤10	6	SOT-23	H
BBY51-03W	7	20	5.30	1	3.10	4	1.75	≤10	6	SOD-323	H
BBY51-07 (dual)	7	20	5.30	1	3.10	4	1.75	≤10	6	SOT-143	H
BBY52 (dual)	7	20	1.75	1	1.25	4	1.40	≤10	6	SOT-23	I
BBY52-03W	7	20	1.75	1	1.25	4	1.40	≤10	6	SOD-323	I
BBY53 (dual)	6	20	5.30	1	2.40	3	2.20	≤10	4	SOT-23	L
BBY53-03W	6	20	5.30	1	2.40	3	2.20	≤10	4	SOD-323	L

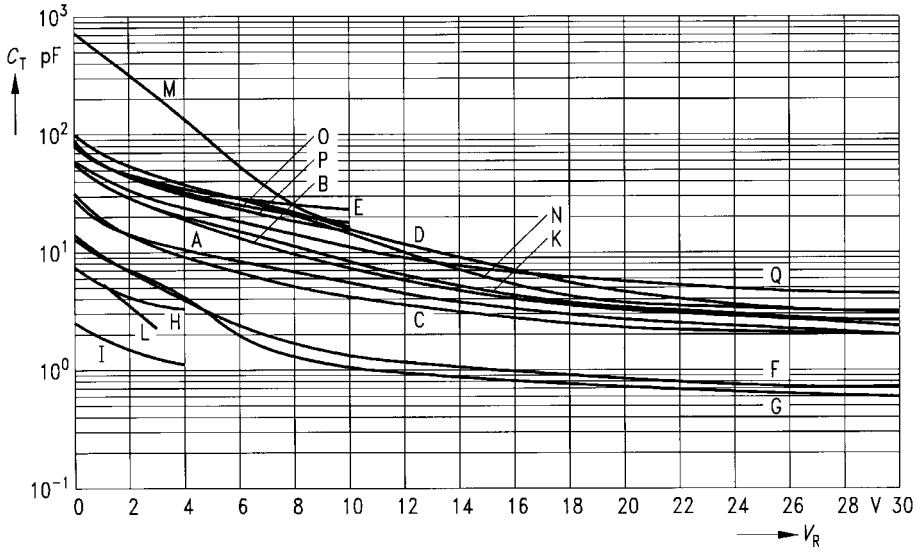


Figure 5-53 Diode capacitance C_T versus reverse voltage V_R for the diodes shown in Table 5-4. The curve labels correspond to the chip codes called out in the table.

If in this equation the diode capacitance is replaced by the resonator circuit frequency ω, we obtain

$$R_c = R_B \left(\frac{\omega^2 L C_S}{\omega^2 L (C_S + C_P) - 1} \right)^2 \tag{5-86}$$

In the circuit of Figure 5-55, the resonant circuit is tuned by two tuning diodes, which are connected in parallel via the coil for tuning purposes, but series-connected in opposition for high-frequency signals. This arrangement has the advantage that the capacitance shift caused by the ac modulation acts in opposite directions in these diodes and, therefore, cancels itself. The bias resistor R_B, which applies the tuning voltage to the tuning diodes, is transformed into the circuit at a constant ratio across the entire tuning range. Given two identical, loss-free tuning diodes, we obtain the expression

$$R_c = 4R_B \tag{5-87}$$

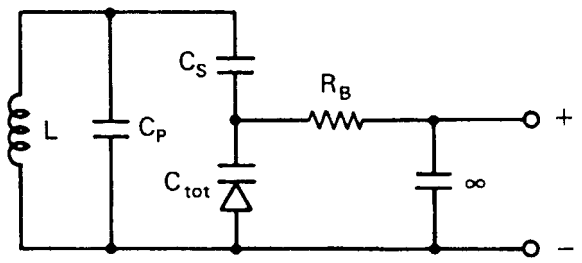


Figure 5-54 Parallel-resonant circuit with tuner diode, and bias resistor in parallel to the diode.

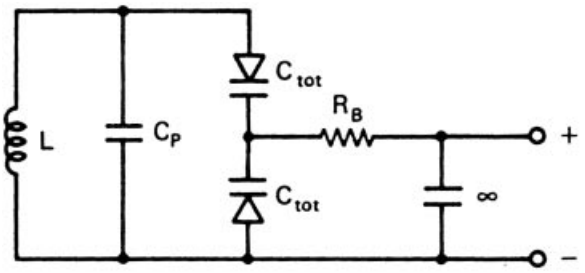


Figure 5-55 Parallel-resonant circuit with two tuning diodes.

Capacitances Connected in Parallel or in Series with the Tuning Diode. Figure 5-54 shows that a capacitor is usually in series with the tuner diode, in order to close the circuit for alternating current and, at the same time, to isolate one terminal of the tuning diode from the rest of the circuit with respect to direct current, so as to enable the tuning voltage to be applied to the diode. If possible, the value of the series resistor C_S will be chosen such that the effective capacitance variation is not restricted. However, in some cases—for example, in the oscillator circuit of receivers whose intermediate frequency is of the order of magnitude of the reception frequency—this is not possible, and the influence of the series capacitance will then have to be taken into account. By connecting the capacitor C_S, assumed to be lossless, in series with the diode capacitance C_{tot}, the tuning capacitance is reduced to the value

$$C^* = C_{tot} \frac{1}{1 + C_{tot}/C_S} \tag{5-88}$$

The Q of the effective tuning capacitance, taking into account the Q of the tuning diode, increases to

$$Q^* = Q\left(1 + \frac{C_{tot}}{C_S}\right) \tag{5-89}$$

The useful capacitance ratio is reduced to the value

$$\frac{C^*_{max}}{C^*_{min}} = \frac{C_{max}}{C_{min}} \frac{1 + C_{min}/C_S}{1 + C_{max}/C_S} \tag{5-90}$$

where C_{max} and C_{min} are the maximum and minimum capacitances, respectively, of the tuner diode.

On the other hand, the advantage is gained that, due to the capacitive potential division, the amplitude of the alternating voltage applied to the tuner diode is reduced to

$$\hat{v}^* = \hat{v} \frac{1}{1 + C_{tot}/C} \tag{5-91}$$

so that the lower value of the tuning voltage can be smaller, and this results in a higher maximum capacitance C_{max} of the tuning diode and a higher useful capacitance ratio. The

influence exerted by the series capacitor, then, can actually be kept lower than Eq. (5-89) would suggest.

The parallel capacitance C_P that appears in Figures 5-54 and 5-55 is always present, since wiring capacitances are inevitable and every coil has its self-capacitance. By treating the capacitance C_P, assumed to be lossless, as a shunt capacitance, the total tuning capacitance rises in value and, if C_S is assumed to be large enough to be disregarded, we obtain

$$C^* = C_{\text{tot}}\left(1 + \frac{C_P}{C_{\text{tot}}}\right) \tag{5-92}$$

The Q of the effective tuning capacitance, derived from the Q of the tuning diode, is

$$Q^* = Q\left(1 + \frac{C_P}{C_{\text{tot}}}\right) \tag{5-93}$$

or, in other words, it rises with the magnitude of the parallel capacitance. The useful capacitance is reduced:

$$\frac{C^*_{\text{max}}}{C^*_{\text{min}}} = \frac{C_{\text{max}}}{C_{\text{min}}}\,\frac{1 + C_P/C_{\text{max}}}{1 + C_P/C_{\text{min}}} \tag{5-94}$$

In view of the fact that even a comparatively small shunt capacitance reduces the capacitance ratio considerably, it is necessary to ensure low wiring and coil capacitances in the layout stage of the circuit design.

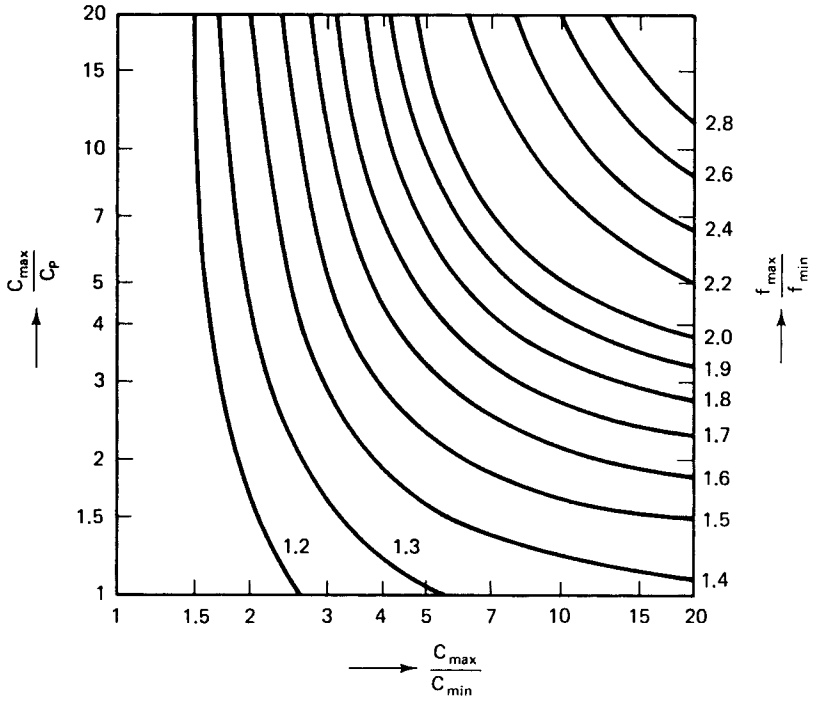


Figure 5-56 Diagram for determining capacitance ratio and maximum capacitance.

Tuning Range. The frequency range over which a parallel-resonant circuit (according to Figure 5-54) can be tuned by means of the tuning diode depends on the useful capacitance ratio of the diode and on the parallel and series capacitances present in the circuit. The ratio is

$$\frac{f_{\max}}{f_{\min}} = \left(\frac{1 + \dfrac{C_{\max}}{C_P(1 + C_{\max}/C_S)}}{1 + \dfrac{C_{\max}}{C_P(C_{\max}/C_{\min} + C_{\max}/C_S)}} \right)^{1/2} \tag{5-95}$$

In many cases, the value of the series capacitor can be chosen to be large enough for its effect to be negligible. In that case, Eq. (5-95) is simplified as follows:

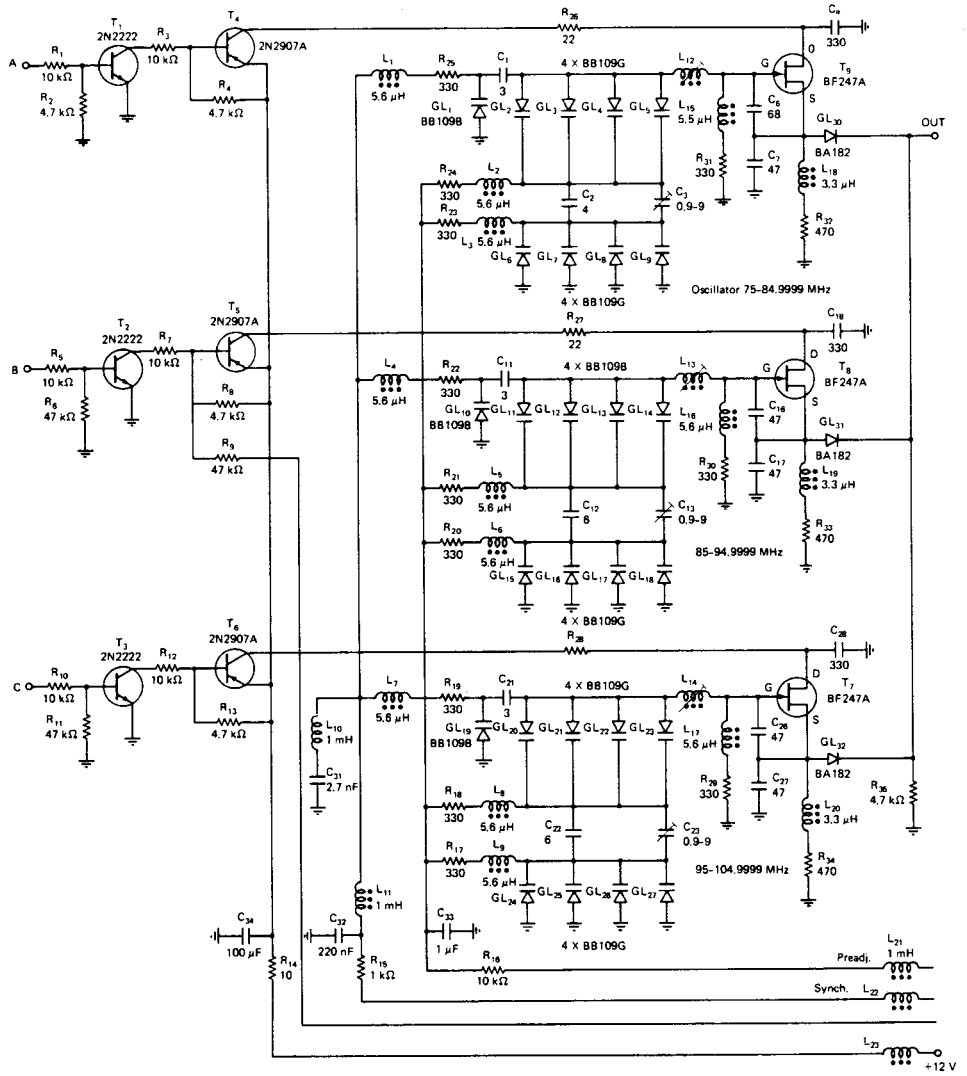


Figure 5-57 Oscillator and switching section of the Rohde & Schwarz ESH2/ESH3 test receiver.

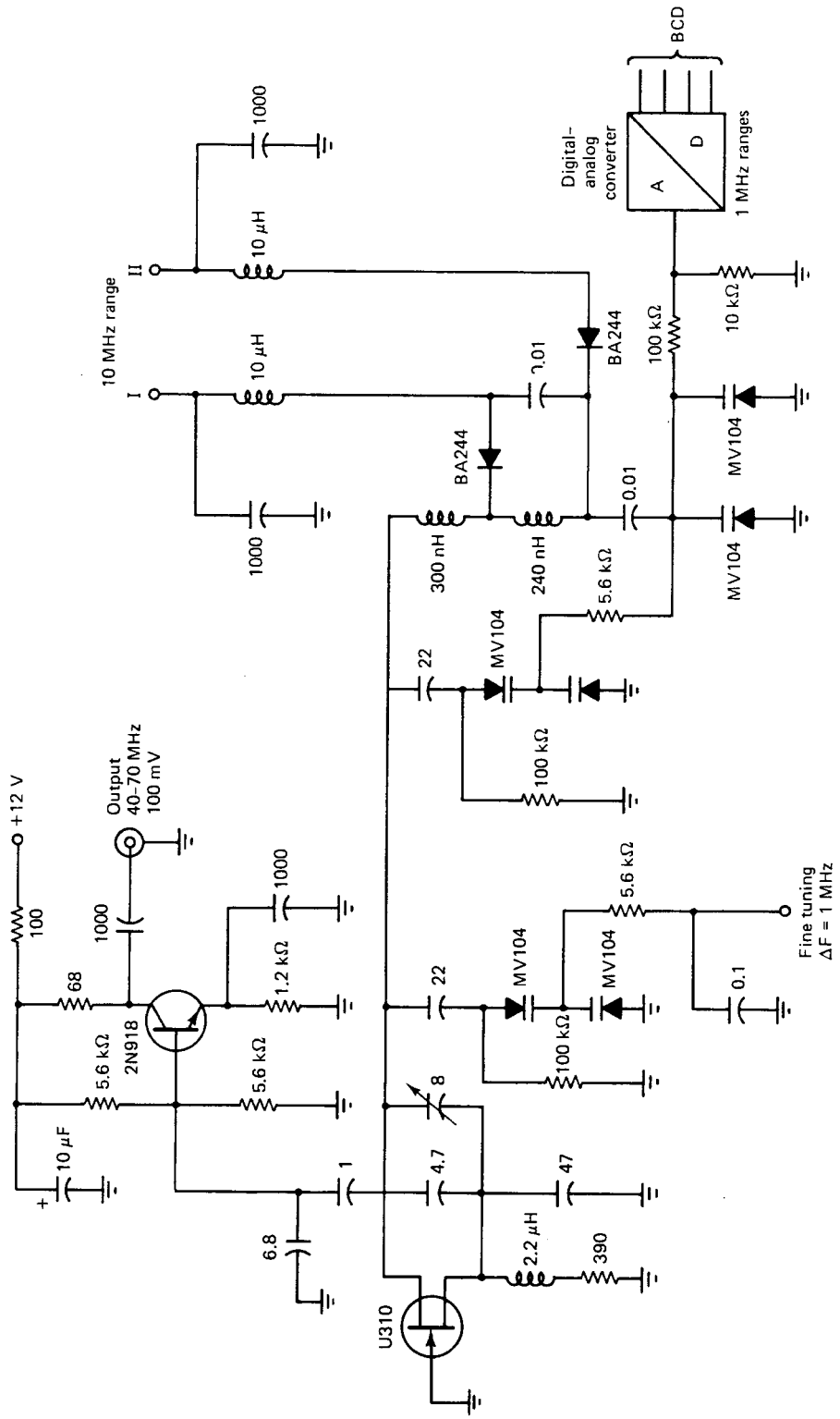


Figure 5-58 A 40–70-MHz VCO with two coarse-steering ranges and a fine-tuning range of 1 MHz.

$$\frac{f_{\max}}{f_{\min}} = \left(\frac{1 + C_{\max}/C_P}{1 + C_{\min}/C_P} \right)^{1/2} \tag{5-96}$$

From this equation, the diagram shown in Figure 5-56 is computed. With the aid of this diagram, the tuning diode parameters required for tuning a resonant circuit over a stipulated frequency range (i.e., the maximum capacitance and the capacitance ratio) can be determined. Whenever the series capacitance C_S cannot be disregarded, the effective capacitance ratio is reduced accorded to Eq. (5-90).

Tracking. When several tuned circuits are used on the same frequency, diodes must be selected for perfect tracking.

5-5-9 Practical Circuits

Oscillators with Coarse and Fine Tuning. After so much theory, it may be nice to take a look at some practical circuits, such as the one shown in Figure 5-57. This oscillator is being used in the Rohde & Schwarz ESN/ESVN40 field-strength meter, and in the HF1030 receiver produced by Cubic Communications, San Diego. This circuit combines all the various techniques shown previously. A single diode is being used for fine-tuning a narrow range of less than 1 MHz; coarse tuning is achieved with the antiseries diodes.

Several unusual properties of this circuit are apparent:

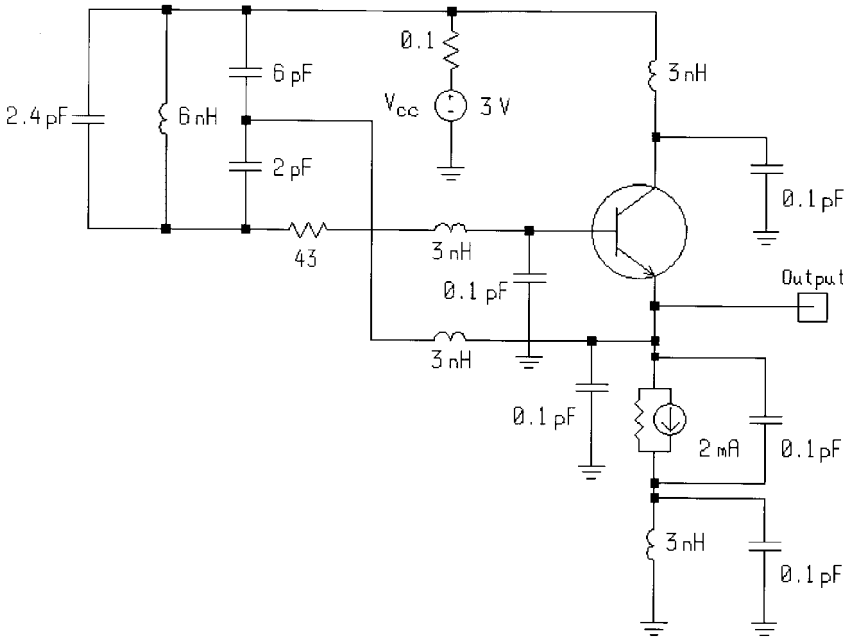


Figure 5-59 Schematic of the dc-coupled oscillator.

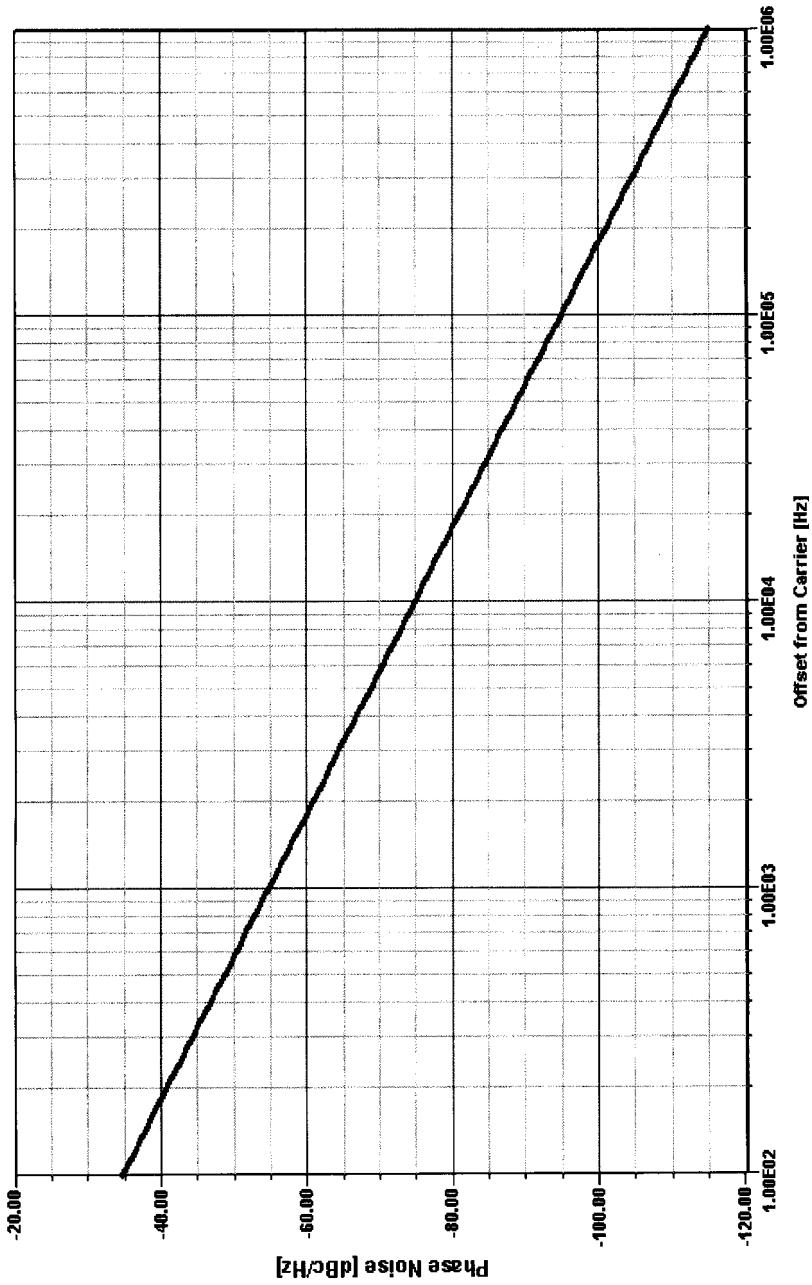


Figure 5-60 Phase noise of the dc-coupled oscillator operating at 897 MHz.

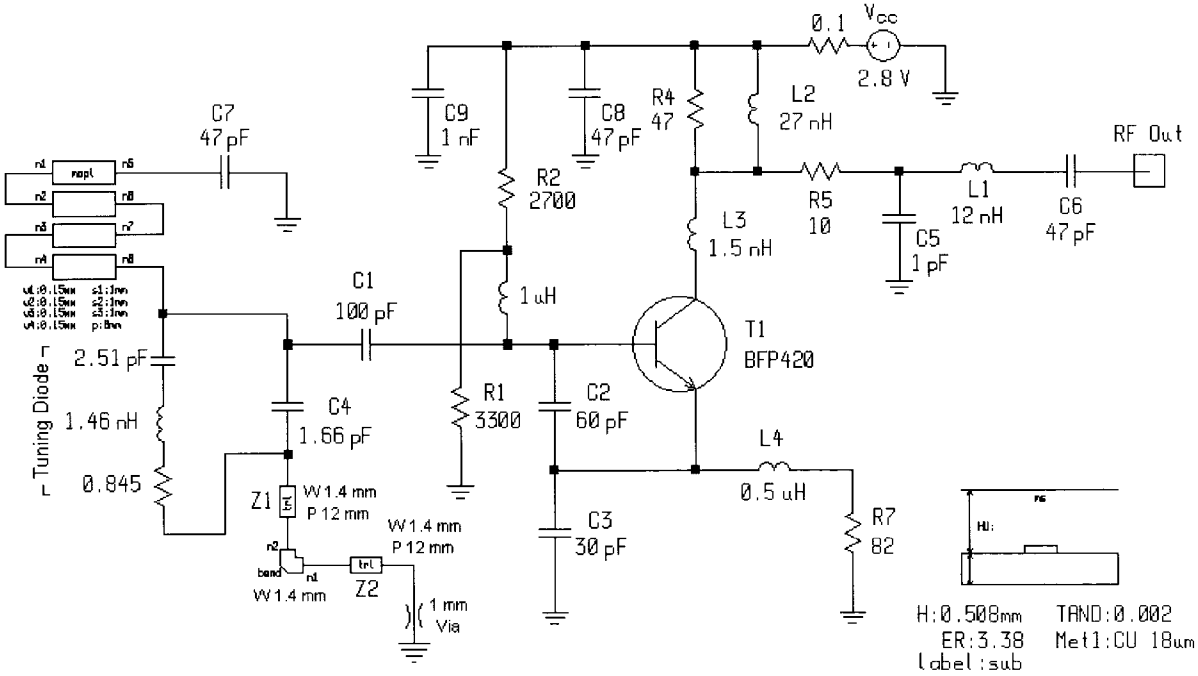


Figure 5-61 Schematic of the Siemens VCO. This oscillator (simulation) works with such high voltage-divider capacitances (C_2 and C_3) only because the circuit's large-signal transistor model, which was obtained under contract, exhibits unrealistically high gain at the frequency of oscillation. Actual capacitors with values on the order of those shown for C_2 and C_3 typically exhibit parasitic resistance and inductance of 0.3 Ω and 0.2 nH, respectively; the resulting series resonance would make any actual such capacitor act as an inductor at UHF. Obtaining practical "UHF useful" capacitances with values on the order of those shown for C_2 and C_3 would therefore necessitate the use of multiple, smaller-value capacitors in parallel—4 × 15 pF at C_2, for example.

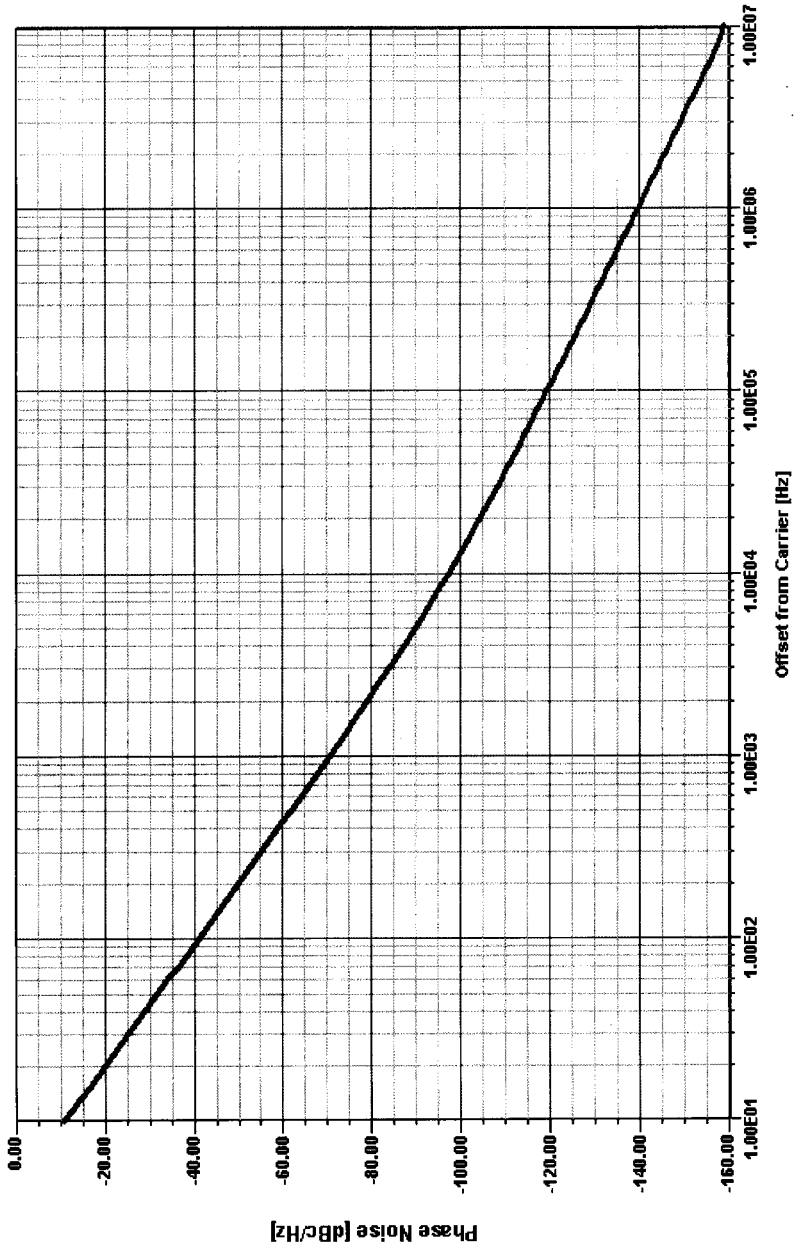


Figure 5-62 Simulated phase noise of the VCO. The frequency of oscillation is 1.052 GHz.

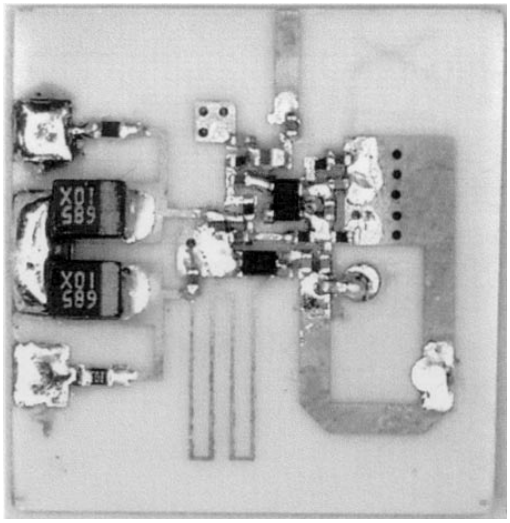


Figure 5-63 Implementation of the Siemens VCO.

1. The fine tuning is achieved with a tuning diode that has a much larger capacitance than that of the coupling capacitor in the circuit. The advantage of this technique is that the fixed capacitor and the tuning diode form a voltage divider, whereby the voltage across the tuning diode decreases as the capacitance increases. For larger values of the capacitance of the tuning diode, the Q changes and the gain K_0 increases. Because of the voltage division, the noise contribution and loading effect of the diode are reduced.
2. In the coarse-tuning circuit, several tuning diodes are used in parallel. The advantage of this circuit is a change in L/C ratio by using a higher C and storing more energy in the tuned circuit. There are no high-Q diodes available with such large capacitance values, and therefore preference is given to using several diodes in parallel rather than one tuning diode with a large capacitance, normally used only for AM (medium-frequency broadcast) tuner circuits.

We mentioned previously that, despite this, the coarse-tuning circuit will introduce noise outside the loop bandwidth, where it cannot be corrected. It is therefore preferable to incorporate switching diodes for segmenting ranges at the expense of switching current drain.

Figure 5-58 shows a circuit using a combined technique of tuning diodes for fine- and medium-resolution tuning, and coarse tuning with switching diodes.

dc-Coupled Oscillator. An interesting circuit that was cut out of an IC was supplied to us by David Lovelace of Motorola. Figure 5-59 shows the circuit and Figure 5-60 shows its phase noise. As a result of the constraints of integration, this BJT oscillator has no base bias circuitry in the conventional form. A fairly noisy performer, it ultimately was not used in production.

Siemens Colpitts Oscillator. A more elaborate approach that is essentially a standard Colpitts oscillator with a lot of detailed modeling is the oscillator of Figure 5-61, supplied

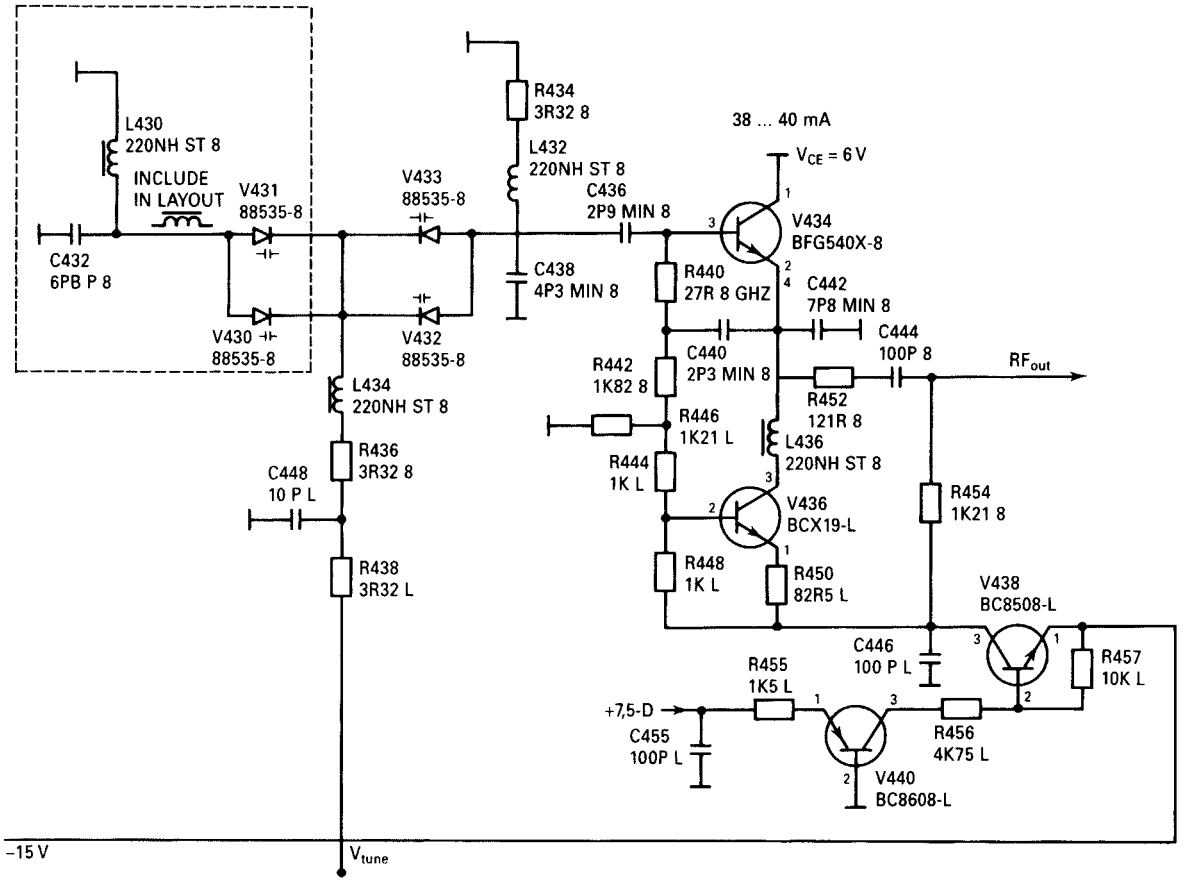


Figure 5-64 A recommended very-low-phase-noise VCO that operates in the 1-GHz region. The V436 stage is a constant-current generator responsible for improved phase-noise performance.

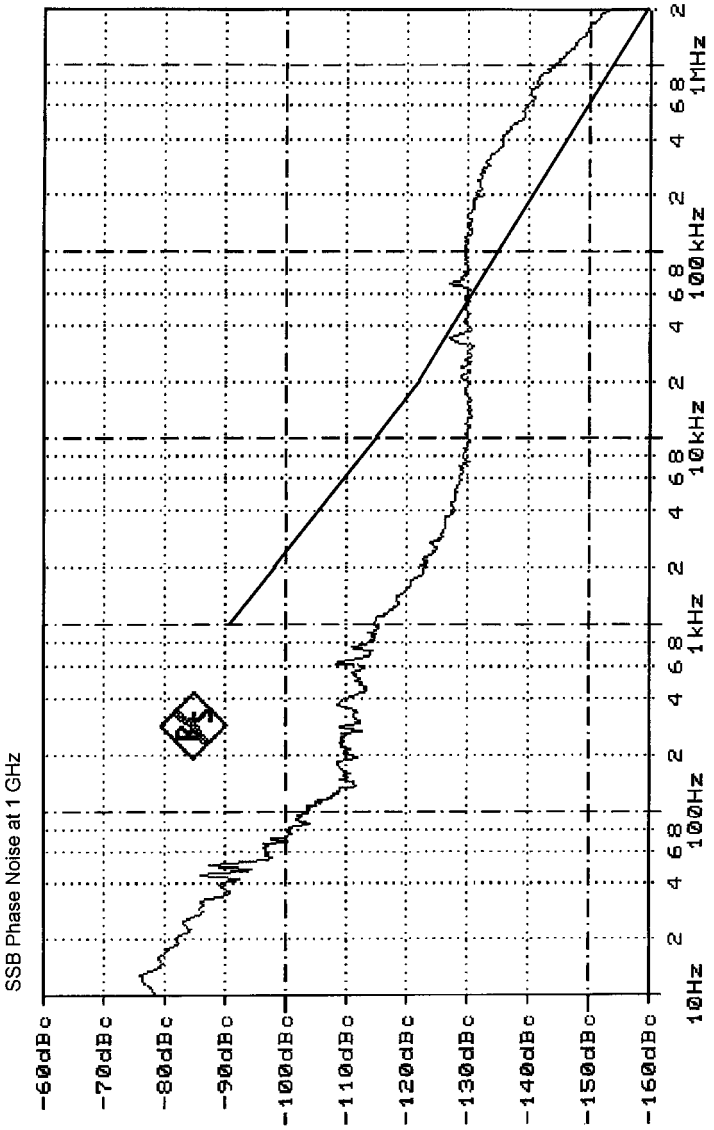


Figure 5-65 Phase noise of the low-noise VCO. The "ledges" from 100 Hz to 1 kHz and from 10 kHz to 100 kHz reflect the noise-cleanup effects of a dual-loop synthesizer; the superimposed line from 1 kHz to 2 MHz shows the oscillator's noise characteristic with both loops unlocked. The closed-loop response is noisier than the open-loop response at offsets above 60 kHz because a wide loop bandwidth is used to achieve a faster switching speed.

to us by Siemens. Considering its simplicity, its noise performance is good, as shown in Figure 5-62. Modeling this circuit involved the consideration of via holes and coupled lines. Figure 5-63 shows an actual implementation of this oscillator. The value of external emitter resistance (R7 in Figure 5-61) plays a critical role in determining a common-emitter Colpitts oscillator's upper frequency limit as described in Yeom [8].

dc-Stabilized Oscillator. For high-performance synthesizers, such as the Rohde & Schwarz series SMG/SMH, a dc-stabilized oscillator with fairly wide tuning range may be used. Figure 5-64 shows such an oscillator, the measured phase noise of which is shown in Figure 5-65. It should be noted that the output loop is magnetically coupled and therefore takes advantage of the inherent selectivity of the tuned circuit.

5-6 NOISE IN OSCILLATORS

In transmitters, oscillator noise can result in adjacent-channel interference and modulation errors; in receivers, oscillator noise can result in demodulation errors and degraded sensitivity and dynamic range. The specification, calculation, and reduction of oscillator noise are therefore of great importance in wireless system design.

5-6-1 Linear Approach to the Calculation of Oscillator Phase Noise

Since an oscillator can be viewed as an amplifier with feedback (Figure 5-66), it is helpful to examine the phase noise added to an amplifier that has a noise figure F. F is defined by [9]

$$F = \frac{(S/N)_{\text{in}}}{(S/N)_{\text{out}}} = \frac{N_{\text{out}}}{N_{\text{in}}G} = \frac{N_{\text{out}}}{GkTB} \tag{5-97}$$

$$N_{\text{out}} = FGkTB \tag{5-98}$$

$$N_{\text{in}} = kTB \tag{5-99}$$

where N_{in} is the total input noise power to a noise-free amplifier. The input phase noise in a 1-Hz bandwidth at any frequency $f_0 + f_m$ from the carrier produces a phase deviation given by (Figure 5-67)

$$\Delta\theta_{\text{peak}} = \frac{V_{n\,\text{RMS1}}}{V_{s,\text{av RMS}}} = \sqrt{\frac{fKT}{P_{s,\text{av}}}} \tag{5-100}$$

$$\Delta\theta_{\text{1RMS}} = \frac{1}{\sqrt{2}}\sqrt{\frac{FkT}{P_{s,\text{av}}}} \tag{5-101}$$

Since a correlated random phase-noise relation exists at $f_0 - f_m$, the total phase deviation becomes

$$\Delta\theta_{\text{RMS total}} = \sqrt{FkT/P_{s,\text{av}}} \tag{5-102}$$

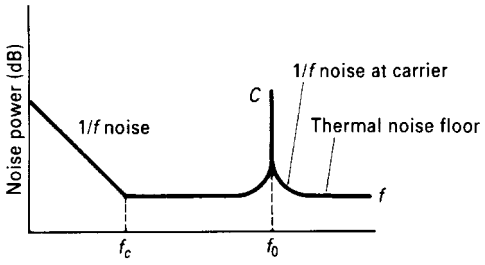


Figure 5-66 Noise power versus frequency of a transistor amplifier with an input signal applied.

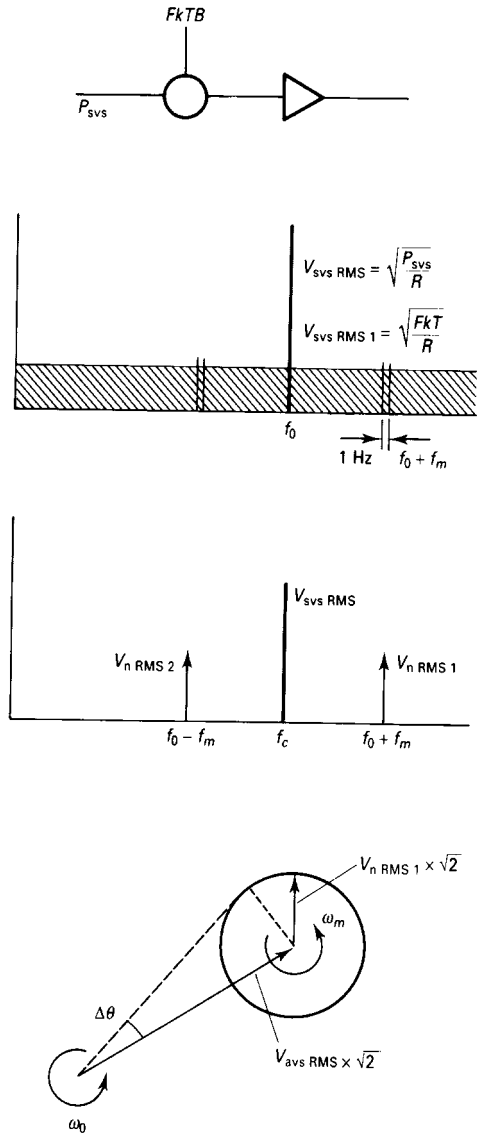


Figure 5-67 Phase noise added to carrier.

The spectral density of phase noise becomes

$$S_\theta(f_m) = \Delta\theta^2_{\text{RMS}} = FkTB/P_{s,\text{av}} \tag{5-103}$$

where $B = 1$ for a 1-Hz bandwidth. Using

$$kTB = -174 \text{ dBm/Hz} \qquad (B = 1) \tag{5-104}$$

allows a calculation of the spectral density of phase noise that is far removed from the carrier (i.e., at large values of f_m). This noise is the theoretical noise floor of the amplifier. For example, an amplifier with +10-dBm power at the input and a noise figure of 6 dB gives

$$S_\theta(f_m > f_c) = -174 \text{ dBm} + 6 \text{ dB} - 10 \text{ dBm} = -178 \text{ dBm} \tag{5-105}$$

Only if P_{OUT} is >0 dBm can we expect $\mathcal{L}$ (signal-to-noise ratio) to be greater than 174 dBc/Hz (1-Hz bandwidth). For a modulation frequency close to the carrier, $S_\theta(f_m)$ shows a flicker or $1/f$ component, which is empirically described by the corner frequency f_c. The phase noise can be modeled by a noise-free amplifier and a phase modulator at the input as shown in Figure 5-68.

The purity of the signal is degraded by the flicker noise at frequencies close to the carrier. The spectral phase noise can be described by

$$S_\theta(f_m) = \frac{FkTB}{P_{s,\text{av}}}\left(1 + \frac{f_c}{f_m}\right) \quad (B = 1) \tag{5-106}$$

No AM-to-PM conversion is considered in this equation. The oscillator may be modeled as an amplifier with feedback as shown in Figure 5-69. The phase noise at the input of the amplifier is affected by the bandwidth of the resonator in the oscillator circuit in the following way. The tank circuit or bandpass resonator has a low-pass transfer function

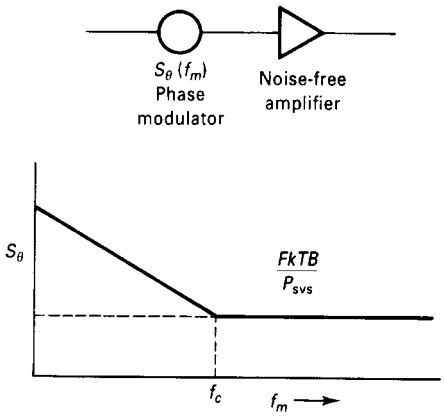


Figure 5-68 Phase noise modeled by a noise-free amplifier and phase modulator.

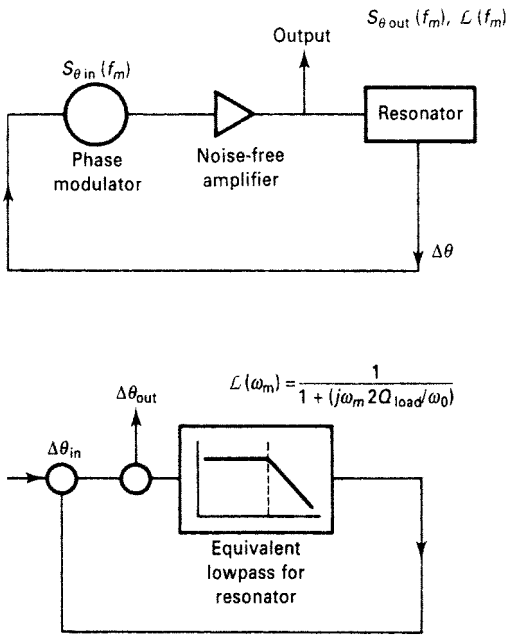


Figure 5-69 Equivalent feedback models of oscillator phase noise.

$$L(\omega_m) = \frac{1}{1 + j(2Q_L\omega_m/\omega_0)} \tag{5-107}$$

where

$$\omega_0/2Q_L = B/2 \tag{5-108}$$

is the half-bandwidth of the resonator. These equations describe the amplitude response of the bandpass resonator; the phase noise is transferred unattenuated through the resonator up to the half-bandwidth.

The closed-loop response of the phase feedback loop is given by

$$\Delta\theta_{\text{out}}(f_m) = \left(1 + \frac{\omega_0}{j2Q_L\omega_m}\right)\Delta\theta_{\text{in}}(f_m) \tag{5-109}$$

The power transfer becomes the phase spectral density

$$S_{\theta,\text{out}}(f_m) = \left[1 + \frac{1}{f_m^2}\left(\frac{f_0}{2Q_L}\right)^2\right] S_{\theta,\text{in}}(f_m) \tag{5-110}$$

where $S_{\theta,\text{in}}$ was given by Eq. (5-106). Finally, $\mathcal{L}(f_m)$ is

$$\mathcal{L}(f_m) = \frac{1}{2}\left[1 + \frac{1}{f_m^2}\left(\frac{f_0}{2Q_L}\right)^2\right] S_{\theta,\text{in}}(f_m) \tag{5-111}$$

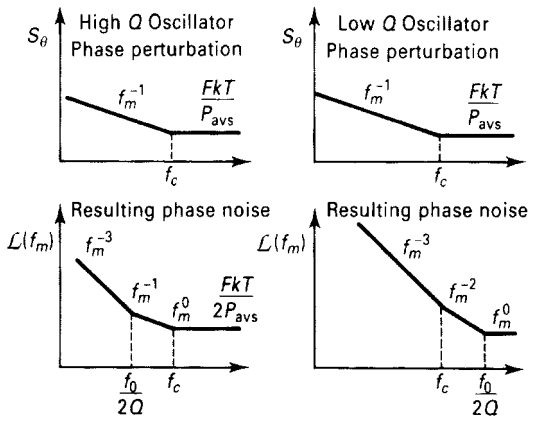


Figure 5-70 Equivalent feedback models of oscillator phase noise.

This equation describes the phase noise at the output of the amplifier (flicker corner frequency and AM-to-PM conversion are not considered). The phase perturbation $S_{\theta,\text{in}}$ at the input of the amplifier is enhanced by the positive phase feedback within the half-bandwidth of the resonator, $f_0/2Q_L$.

Depending on the relation between f_c and $f_0/2Q_L$, there are two cases of interest, as shown in Figure 5-70. For the low-Q case, the spectral phase noise is unaffected by the Q of the resonator, but the $\mathcal{L}(f_m)$ spectral density will show a $1/f^3$ and $1/f^2$ dependence close to the carrier. For the high-Q case, a region of $1/f^3$ and $1/f$ should be observed near the carrier. Substituting Eq. (5-106) into Eq. (5-111) gives an overall noise of

$$\mathcal{L}(f_m) = \frac{1}{2}\left[1 + \frac{1}{f_m^2}\left(\frac{f}{2Q_L}\right)^2\right]\frac{FkT}{P_{s,\text{av}}}\left(1 + \frac{f_c}{f_m}\right)$$

$$= \frac{FkTB}{2P_{s,\text{av}}}\left[\frac{1}{f_m^3}\frac{f^2 f_c}{4Q_L^2} + \frac{1}{f_m^2}\left(\frac{f}{2Q_L}\right)^2 + \left(1 + \frac{f_c}{f_m}\right)\right] \text{dBc/Hz} \tag{5-112}$$

Examining Eq. (5-112) gives the four major causes of oscillator noise: the upconverted $1/f$ noise or flicker FM noise, the thermal FM noise, the flicker phase noise, and the thermal noise floor, respectively.

Q_L (loaded Q) can be expressed as

$$Q_L = \frac{\omega_o W_e}{P_{\text{diss,total}}} = \frac{\omega_o W_e}{P_{\text{in}} + P_{\text{res}} + P_{\text{sig}}} = \frac{\text{Reactive power}}{\text{Total dissipated power}} \tag{5-113}$$

where W_e is the reactive energy stored in L and C,

$$W_e = \frac{1}{2} CV^2 \tag{5-114}$$

$$P_{\text{res}} = \frac{\omega_o W_e}{Q_{\text{unl}}} \tag{5-115}$$

$$\mathcal{L}(f_m) = \frac{1}{2}\left[1 + \frac{\omega_o^2}{4\omega_m^2}\left(\frac{P_{\text{in}}}{\omega_o W_e} + \frac{1}{Q_{\text{unl}}} + \frac{P_{\text{sig}}}{\omega_o W_e}\right)^2\right]\left(1 + \frac{\omega_c}{\omega_m}\right)\frac{FkT_o}{P_{\text{sav}}} \tag{5-116}$$

Input power/ Reactive power ratio

Resonator Q

Signal power/ Reactive power ratio

Flicker effect

Phase perturbation

This equation is extremely significant because it covers most of the causes of phase noise in oscillators. [AM-to-PM conversion must be added; see Eq. (5-56).] To minimize the phase noise, the following design rules apply:

1. Maximize the unloaded Q.
2. Maximize the reactive energy by means of a high RF voltage across the resonator and obtain a low L/C ratio. The limits are set by breakdown voltages of the active devices and the tuning diodes, and the forward-bias condition of the tuning diodes. If L becomes too high, the circuit degenerates into a squegging oscillator. Using lower L allows larger capacitance, which allows us to swamp the oscillating device's internal, nonlinear capacitance with external, linear capacitance, reducing the phase-noise degradation caused by the nonlinear capacitance. The ceramic-resonator oscillator best illustrates this approach, which may be hard to accomplish in discrete circuits because high-Q inductors with values above 1 nH cannot be built unless high-Q transmission lines or resonators are used.
3. Avoid saturation at all costs, and try to have either limiting or automatic gain control (AGC) without degradation of Q. Isolate the tuned circuit from the limiter or AGC circuit. Use tuning diodes in antiseries configurations to avoid forward bias.
4. Choose an active device with the lowest possible noise figure and flicker corner frequency. The noise figure of interest is the noise figure obtained with the actual impedance at which the device is operated. Using FETs rather than BJTs, it is preferable to deal with the equivalent noise voltage and noise currents rather than with the noise figure, since they are independent of source impedance. The noise figure improves as the ratio between source impedance and equivalent noise resistance increases. In addition, in a tuned circuit, the source impedance changes drastically as a function of the offset frequency, and this effect has to be considered. For low phase-noise operation, use a medium-power transistor. If you need your output power to be achieved at 6–9 mA, select a transistor with an $I_{C,\max}$ of 60–90 mA. Also avoid an f_T greater than 3× to 5× the operating frequency. To make transistors that are stable across their full frequency ranges, manufacturers add circuitry that makes the flicker corner frequency higher as f_T increases.

   The following transistors have the lowest noise figure:

   - BFG65/67 by Philips
   - BFR106/92; BFP405/420 by Siemens
   - 2SC3358/3356 by NEC
   - HXTR4105 by HP

- AT2100 (chip) by Avantek

Among the lowest-noise JFETs are the U310, 2N4416/17, and 2N5397.

In designing MMICs, one must resort to MESFETs, which are also referred to as GaAsFETs. They have a fairly high flicker frequency, which is typically around 10 MHz but can go up to as high as 100 MHz.

Table 5-5 shows equivalent circuits and measurement data for the Avantek AT2100 20-GHz *npn* silicon bipolar oscillator transistor. (Siemens of Germany is about to come out with 45-GHz f_T transistors, which will compete with SiGe transistors in the future.)

5. Phase perturbation can be minimized by using high-impedance devices such as FETs, where the signal-to-noise ratio of the signal voltage relative to the equivalent noise voltage can be made very high. This also indicates that in the case of a limiter, the limited voltage should be as high as possible.
6. Choose an active device with low flicker noise. The effect of flicker noise can be reduced through RF feedback. An unbypassed emitter resistor of 10–30 Ω in a bipolar circuit can improve the flicker noise by as much as 40 dB.

   The proper bias point of the active device is important, and precautions should be taken to prevent modulation of the input and output dynamic capacitances of the active device, which will cause amplitude-to-phase conversion and therefore introduce noise.
7. The energy should be coupled from the resonator rather than from another portion of the active device so that the resonator limits the bandwidth.
8. For microwave applications, the lowest-noise GaAsFETs are pseudomorphic HJ FETs. A reproduction of the datasheet of the NEC NE42484A, which is generally the state of the art today, follows; it is used with permission.

Equation (5-116) assumes that the phase perturbation and the flicker effect are the limiting factors, as practical use of such oscillators requires that an isolation amplifier must be used.

In the event that energy is taken directly from the resonator and the oscillator power can be increased, the signal-to-noise ratio can be increased above the theoretical limit of 174 dBc/Hz, due to the low-pass-filter effect of the tuned resonator. However, since this is mainly a theoretical assumption and does not represent a real-world system, this noise performance cannot be obtained. In an oscillator stage, even a total noise floor of –170 dBc/Hz is rarely achieved.

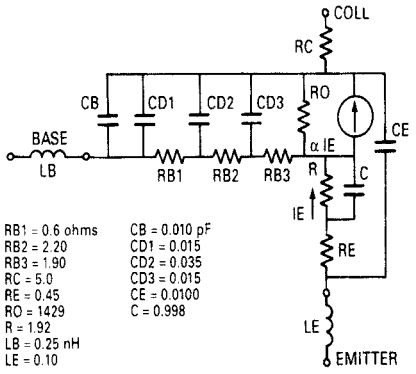
What other influences do we have that cause an oscillator's noise performance to degrade? So far, we have assumed that the Q of the tuned circuit is really determined only by the LC network and the loading effect of the transistor. In synthesizer applications, however, we find it necessary to add a tuning diode. The tuning diode has a substantially lower Q than that of a mica capacitor or even a ceramic capacitor. As a result of this, the noise sidebands change as a function of the additional loss. This is best expressed by adjusting the value for the loaded Q in Eq. (5-113).

There seems to be no precise mathematical way of predetermining the noise influence of a tuning diode, but the following approximation seems to give proper results:

$$\frac{1}{Q_{T,\text{load}}} = \frac{1}{Q_{\text{load}}} + \frac{1}{Q_{\text{diode}}} \tag{5-117}$$

Table 5-5 Linear and SPICE equivalent circuits for the Avantek AT21400 chip

LINEAR EQUIVALENT CIRCUIT[a]
V_{CE} = 8 V, I_C = 14 mA

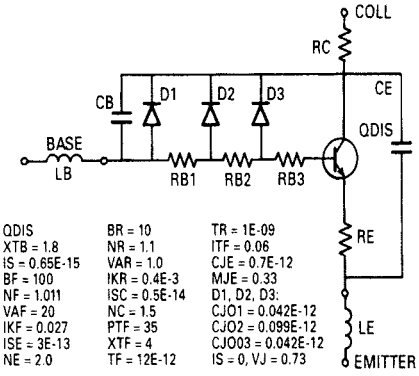


RB1 = 0.6 ohms
RB2 = 2.20
RB3 = 1.90
RC = 5.0
RE = 0.45
RO = 1429
R = 1.92
LB = 0.25 nH
LE = 0.10

CB = 0.010 pF
CD1 = 0.015
CD2 = 0.035
CD3 = 0.015
CE = 0.0100
C = 0.998

CCCS: M = .99, A = 0
R1 = R, R2 = 1E8, F = 22.7, T = 7.4

SPICE EQUIVALENT CIRCUIT[a]



QDIS
XTB = 1.8
IS = 0.65E-15
BF = 100
NF = 1.011
VAF = 20
IKF = 0.027
ISE = 3E-13
NE = 2.0

BR = 10
NR = 1.1
VAR = 1.0
IKR = 0.4E-3
ISC = 0.5E-14
NC = 1.5
PTF = 35
XTF = 4
TF = 12E-12

TR = 1E-09
ITF = 0.06
CJE = 0.7E-12
MJE = 0.33
D1, D2, D3:
CJO1 = 0.042E-12
CJO2 = 0.099E-12
CJO03 = 0.042E-12
IS = 0, VJ = 0.73

Modeled Scattering Parameters, Common Emitter[b]

T_A = 25°C, V_{CE} = 8 V, I_C = 14 mA

Freq.	S_{11}		S_{21}			S_{12}			S_{22}	
GHz	Mag	Ang	dB	Mag	Ang	dB	Mag	Ang	Mag	Ang
2	.67	−168	15.8	6.19	90	−34.0	.02	52	.65	−9
3	.67	−180	12.4	4.17	81	−32.0	.025	59	.65	−10
4	.67	172	10.0	3.15	74	−30.5	.03	64	.65	−11
5	.68	165	8.1	2.53	68	−28.0	.04	68	.65	−12
6	.68	160	6.5	2.12	62	−28.0	.04	70	.65	−14
7	.69	154	5.2	1.82	56	−26.0	.05	72	.65	−16
8	.70	149	4.1	1.60	51	−24.4	.06	73	.65	−18
9	.70	144	3.1	1.43	46	−24.4	.06	74	.65	−20
10	.71	140	2.2	1.29	41	−23.1	.07	74	.65	−22
11	.72	135	1.4	1.18	36	−21.9	.08	74	.65	−24
12	.73	131	0.7	1.08	31	−21.9	.08	75	.66	−26
13	.74	127	0	1.00	27	−20.9	.09	75	.66	−28
14	.75	123	−0.6	0.93	23	−20.0	.10	75	.66	−30
15	.76	120	−1.3	0.86	18	−19.2	.11	74	.67	−32
16	.77	116	−1.9	0.80	14	−19.2	.11	74	.67	−34
18	.79	109	−3.1	0.70	7	−17.7	.13	74	.68	−39
20	.80	103	−4.2	0.62	−1	−17.1	.14	73	.69	−44
22	.82	97	−5.4	0.54	−7	−15.9	.16	72	.70	−48
24	.84	92	−6.6	0.47	−14	−14.9	.18	71	.71	−53
26	.85	87	−8.0	0.40	−20	−14.0	.20	70	.72	−58

[a]These equivalent circuits are provided only as first-order design aids. Their accuracy for critical designs at very high frequencies has not been validated.

[b]*S* parameters are from the linear equivalent circuit. Below 10 GHz, they have been fit to measurements of die on a standard carrier with one bond wire to the base and four wires to the emitter.

NEC

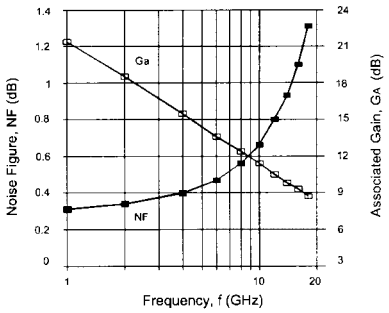
SUPER LOW NOISE PSEUDOMORPHIC HJ FET

NE42484A

FEATURES

- **VERY LOW NOISE FIGURE:** 0.8 dB TYP at 12 GHz
- **HIGH ASSOCIATED GAIN:** 10.5 dB TYP at 12 GHz
- **L_G = 0.35 µm, W_G = 200 µm**
- **LOW COST METAL CERAMIC PACKAGE**
- **TAPE & REEL PACKAGING OPTION AVAILABLE**

NOISE FIGURE & ASSOCIATED GAIN vs. FREQUENCY
V_{DS} = 2 V, I_{DS} = 10 mA



DESCRIPTION

The NE42484A is a pseudomorphic Hetero-Junction FET that uses the junction between Si-doped AlGaAs and undoped InGaAs to create very high mobility electrons. The device features mushroom shaped TiAl gates for decreased gate resistance and improved power handling capabilities. The mushroom gate also results in lower noise figure and high associated gain. This device is housed in an epoxy-sealed, metal/ceramic package and is intended for high volume consumer and industrial applications.

NEC's stringent quality assurance and test procedures ensure the highest reliability and performance.

ELECTRICAL CHARACTERISTICS (T_A = 25°C)

PART NUMBER PACKAGE OUTLINE			NE42484A 84AS		
SYMBOLS	**PARAMETERS AND CONDITIONS**	**UNITS**	**MIN**	**TYP**	**MAX**
NF_{OPT}[1]	Optimum Noise Figure, V_{DS} = 2.0 V, I_{DS} = 10 mA, f = 12 GHz	dB		0.8	1.2
G_A[1]	Associated Gain, V_{DS} = 2.0 V, I_{DS} = 10 mA, f = 12 GHz	dB	9.0	10.5	
P_{1dB}	Output Power at 1 dB Gain Compression Point, f = 12 GHz V_{DS} = 2.0 V, I_{DS} = 10 mA V_{DS} = 2.0 V, I_{DS} = 20 mA	 dBm dBm		 9.7 10.2	
G_{1dB}	Gain at P_{1dB}, f = 12 GHz V_{DS} = 2.0 V, I_{DS} = 10 mA V_{DS} = 2.0 V, I_{DS} = 20 mA	 dB dB		 10.3 10.5	
I_{DSS}	Saturated Drain Current, V_{DS} = 2.0 V, V_{GS} = 0 V	mA	15	40	70
V_P	Pinch-off Voltage, V_{DS} = 2.0 V, I_{DS} = 0.1 mA	V	-2.0	-0.8	-0.2
g_m	Transconductance, V_{DS} = 2.0 V, I_D = 10 mA	mS	45	60	
I_{GSO}	Gate to Source Leakage Current, V_{GS} = -3.0 V	µA		0.5	10.0
$R_{TH\ (CH-A)}$	Thermal Resistance (Channel to Ambient)	°C/W		750	
$R_{TH(CH-C)}$[2]	Thermal Resistance (Channel to Case)	°C/W			350

Notes:
1. Typical values of noise figures and associated gain are those obtained when 50% of the devices from a large number of lots were individually measured in a circuit with the input individually tuned to obtain the minimum value. Maximum values are criteria established on the production line as a "go-no-go" screening tuned for the "generic" type but not for each specimen.
2. R_{TH} (channel to case) for package mounted on an infinite heat sink.

NE42484A

ABSOLUTE MAXIMUM RATINGS[1] (T_A = 25°C)

SYMBOLS	PARAMETERS	UNITS	RATINGS
V_{DS}	Drain to Source Voltage	V	4.0
V_{GS}	Gate to Source Voltage	V	-3.0
I_{DS}	Drain Current	mA	I_{DSS}
I_{GRF}	Gate Current (RF Drive)	µA	200
P_{IN}	RF Input (CW)	dBm	15
T_{CH}	Channel Temperature	°C	150
T_{STG}	Storage Temperature	°C	-65 to +150
P_T	Total Power Dissipation	mW	165

Note:
1. Operation in excess of any one of these parameters may result in permanent damage.

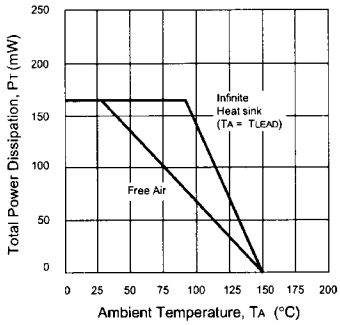
TYPICAL NOISE PARAMETERS (T_A = 25°C)

V_{DS} = 2 V, I_{DS} = 10 mA

FREQ. (GHz)	NF_{OPT} (dB)	G_A (dB)	Γ_{OPT} MAG	Γ_{OPT} ANG	Rn/50
1.0	0.31	21.4	0.78	10	0.43
2.0	0.34	18.5	0.76	28	0.38
4.0	0.40	15.5	0.72	58	0.28
6.0	0.47	13.6	0.65	84	0.21
8.0	0.56	12.4	0.57	113	0.15
10.0	0.66	11.4	0.50	141	0.10
12.0	0.80	10.5	0.44	173	0.09
14.0	0.93	9.8	0.39	-157	0.08
16.0	1.10	9.3	0.36	-125	0.08
18.0	1.31	8.7	0.35	-90	0.08

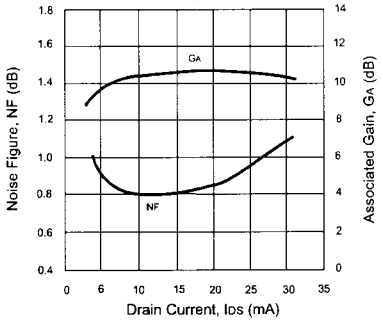
TYPICAL PERFORMANCE CURVES (T_A = 25°C)

TOTAL POWER DISSIPATION vs. AMBIENT TEMPERATURE[1]

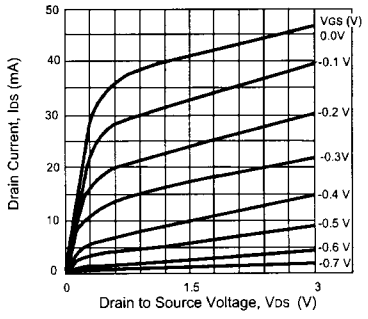


Note
1. If P_T exceeds the Free Air Value, reliable operation can be assured by measuring the worst-case temperature, $T_{(LEAD)}$, at the lead where heat flow is maximum (usually the source lead) and limiting T_A, P_T or $R_{TH\ (CKT)}$

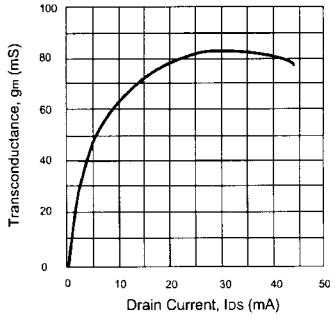
NOISE FIGURE AND ASSOCIATED GAIN vs. DRAIN CURRENT V_{DS} = 2 V, f = 12 GHz



DRAIN CURRENT vs. DRAIN TO SOURCE VOLTAGE



TRANSCONDUCTANCE vs. DRAIN CURRENT V_{DS} = 2.0 V



The tuning diode is specified to have a cutoff frequency $f_{\max}$, which is determined from the loss resistor R_S and the value of the junction capacitance as a function of voltage (e.g., measured at 3 V). This means that the voltage determines the Q and, consequently, the noise bandwidth.

5-6-2 AM-to-PM Conversion

We went into more detail in dealing with the mechanism and influence of tuning diodes in the voltage-controlled oscillators section, where we evaluated the various methods of building voltage-tunable oscillators using tuning diodes. In this section, we limit ourselves to practical results.

The loading effect of the tuning diode is due to losses, and these losses can be described by a resistor in parallel with the tuned circuit.

It is possible to define an equivalent noise $R_{a,\text{eq}}$ that, inserted in Nyquist's equation,

$$V_n = \sqrt{4kT_0R\,\Delta f} \tag{5-118}$$

where $kT_0 = 4.2 \times 10^{-21}$ at about 300 K, R is the equivalent noise resistor, and Δf is the bandwidth, determines an open-circuit noise voltage across the tuning diode. Practical values of $R_{a,\text{eq}}$ for carefully selected tuning diodes are in the vicinity of 200 Ω to 50 kΩ. If we now determine the noise voltage $V_n = \sqrt{4 \times 4.2 \times 10^{-21} \times 10{,}000}$, the resulting voltage value is 1.296×10^{-8} V $\sqrt{\text{Hz}}$.

This noise voltage generated from the tuning diode is now multiplied with the VCO gain K_0, resulting in the rms frequency deviation

$$(\Delta f_{\text{rms}}) = K_0 \times (1.296 \times 10^{-8}\ \text{V}) \text{ in 1-Hz bandwidth} \tag{5-119}$$

To translate this into an equivalent peak phase deviation,

$$\theta_d = \frac{K_0\sqrt{2}}{f_m}(1.296 \times 10^{-8}) \text{ rad in 1-Hz bandwidth} \tag{5-120}$$

or for a typical oscillator gain of 100 kHz/V,

$$\theta_d = \frac{0.00183}{f_m} \text{ rad in 1-Hz bandwidth} \tag{5-121}$$

For f_m = 25 kHz (typical spacing for adjacent-channel measurements for FM mobile radios), $\theta_c = 7.32 \times 10^{-8}$. This can be converted now into the SSB signal-to-noise ratio:

$$\mathcal{L}(f_m) = 20 \log_{10}\left(\frac{\theta_c}{2}\right) = -149 \text{ dBc/Hz} \tag{5-122}$$

For the typical oscillator gain of 10 MHz/V found in wireless applications, the resulting phase noise will be 20 dB worse [10 log (10 MHz ÷ 100 kHz)]. However, the best tuning diodes, like the BB104, have an R_n of 200 Ω instead of 10 kΩ, which again changes the picture. According to Eq. (5-118), with $kT_0 = 4.2 \times 10^{-21}$ the resulting noise voltage will be

$$V_n = \sqrt{4 \times 4.2 \times 10^{-21} \times 200}$$

$$= 1.833 \times 10^{-9} \text{ V}\sqrt{\text{Hz}} \tag{5-123}$$

From Eq. (5-114), the equivalent peak phase deviation for a gain of 10 MHz/V in a 1-Hz bandwidth is then

$$\theta_d = \frac{1 \times 10^7 \sqrt{2}}{f_m} (1.833 \times 10^{-9}) \text{ rad} \tag{5-124}$$

or

$$\theta_d = \frac{0.026}{f_m} \text{ rad in 1-Hz bandwidth} \tag{5-125}$$

With $f_m = 25$ kHz, $\theta_c = 1.04 \times 10^{-6}$. Expressing this as phase noise:

$$\mathcal{L}(f_m) = 20 \log_{10}\left(\frac{\theta_c}{2}\right) = -126 \text{ dBc/Hz} \tag{5-126}$$

The phase-noise value in Eq. (5-122) is that typically achieved in the Rohde & Schwarz SMDU or with the Hewlett-Packard 8640 signal generator and considered state-of-the-art for a free-running oscillator. It should be noted that both signal generators use a slightly different tuned circuit; the Rohde & Schwarz generator uses a helical resonator, whereas the Hewlett-Packard generator uses an electrically shortened quarter-wavelength cavity. Both generators are mechanically pretuned, and a tuning diode with a gain of about 100 kHz/V is used for AFC and frequency-modulation purposes. It is apparent that, because of the nonlinearity of the tuning diode, the gain is different for low dc voltages than for high dc voltages. The impact of this is that the noise varies with the tuning range.

If this oscillator had to be used for a synthesizer, its 1-MHz tuning range would be insufficient; therefore, a way had to be found to segment the band into the necessary ranges. In VCOs, this is typically done with switching diodes that allow the proper frequency bands to be selected. These switching diodes insert in parallel or series, depending on the circuit, with additional inductors or capacitors, depending on the design.

In low-energy-consuming circuits, the VCO frequently is divided into a coarse-tuning section using tuning diodes and a fine-tuning section with a tuning diode. In the coarse-tuning range, this results in very high gains, such as 1–10 MHz/V, for the diodes. The noise contribution of those diodes is therefore very high and can hardly be compensated by the loop. For low-noise applications, which automatically mean higher power consumption, it is unavoidable to use switching diodes.

Let us now examine some test results. Referring back to Eq. (5-106), Figure 5-71 shows the noise sideband performance as a function of Q, whereby the top curve with $Q_L = 100$ represents a somewhat poor oscillator and the lowest curve with $Q_L = 100{,}000$ probably represents a crystal oscillator where the unloaded Q of the crystal was in the vicinity of 3×10^6. Figure 5-72 shows the influence of flicker noise.

Corner frequencies of 10 Hz to 10 kHz have been selected, and it becomes apparent that around 1 kHz the influence is fairly dramatic, whereas the influence at 20 kHz off the carrier

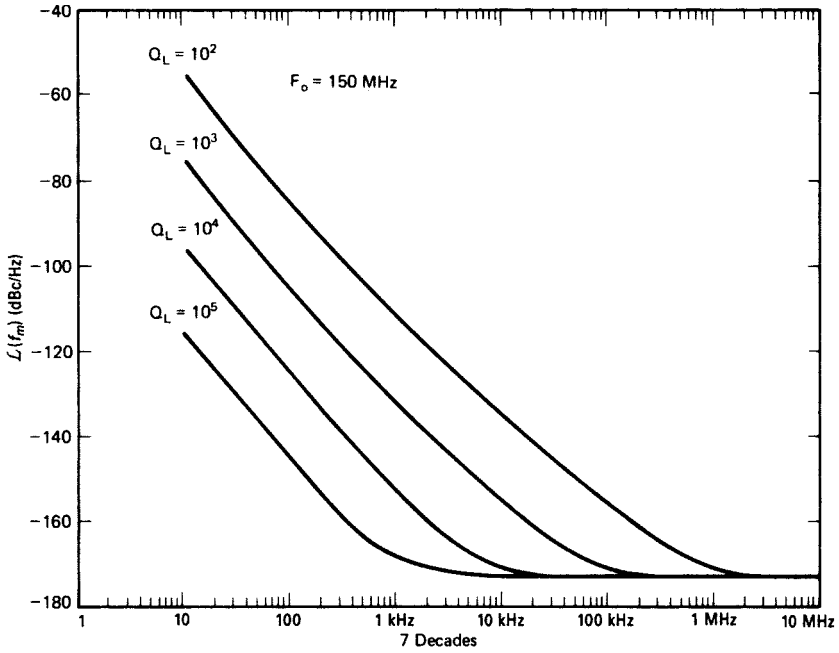


Figure 5-71 Noise sideband of an oscillator at 150 MHz as a function of the loaded Q of the resonator.

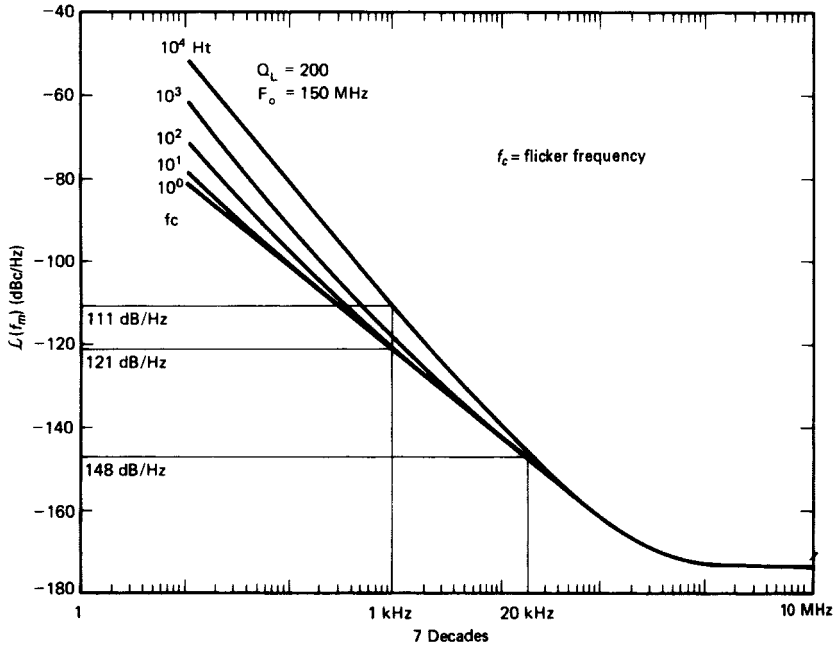


Figure 5-72 Noise-sideband performance as a function of the flicker frequency f_c varying from 10 Hz to 10 kHz.

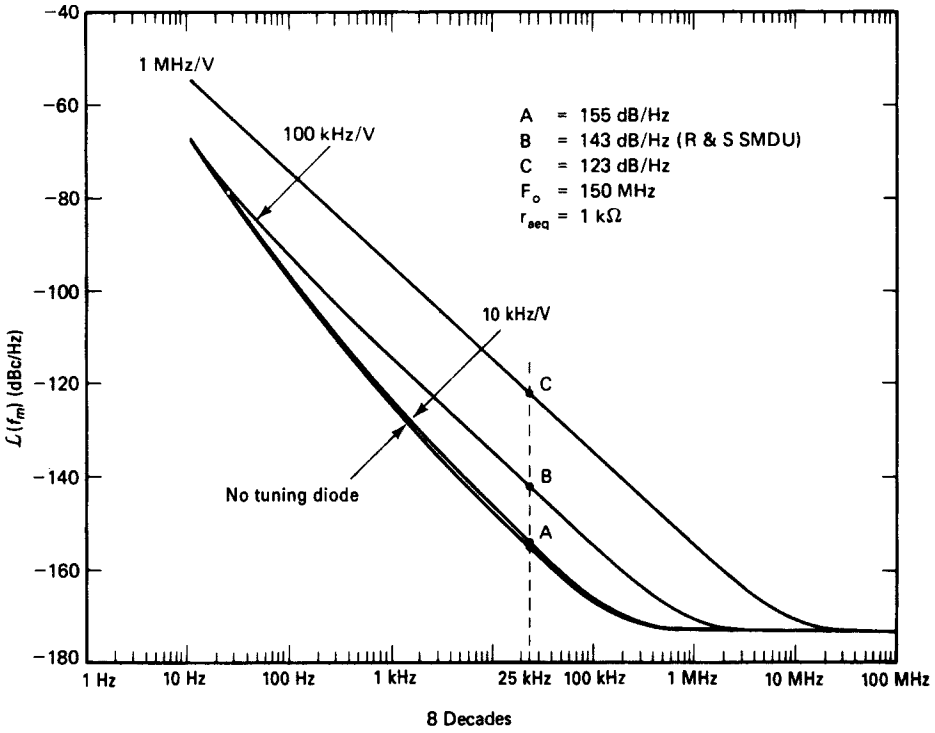


Figure 5-73 Noise-sideband performance of an oscillator at 150 MHz, showing the influence of various tuning diodes.

is not significant. Figures 5-73 and 5-74 show the influence of the tuning diodes on a high-Q oscillator.

Curve A in Figure 5-73 uses a lightly coupled tuning diode with a K_0 of 10 kHz/V; the lower curve is the noise performance without any diode. As a result, the two curves are almost identical, which can be seen from the somewhat smeared form of the graph. Curve B shows the influence of a tuning diode at 100 kHz/V and represents a value of −143 dBc/Hz from −155 dBc/Hz, already some deterioration. Curve C shows the noise if the tuning diode results in a 1-MHz/V VCO gain, and the noise sideband at 25 kHz has now deteriorated to −123 dBc/Hz. These curves speak for themselves.

Figure 5-74 shows how the phase noise of a 147-MHz VCO varies with the bias applied to its tuning diode. Because a tuning diode's Q increases with the reverse bias applied to it, we expect a VCO's phase-noise performance to improve, even if only slightly, as its tuning voltage is increased. Yet Figure 5-74 illustrates the reverse of this relationship. Inspecting this VCO's circuit (Figure 5-75), we see that the diode is in series with the circuit's 220-nH resonator. *All* of the RF current flowing through the resonator must flow through the relatively low-Q diode, and even small decreases in the diode's capacitance will significantly modify the tank's L/C ratio.

Graphing the diode's ac load line (Figure 5-76), we see that as the tuning voltage increases (i.e., as the tuning diode capacitance decreases), the RF voltage across the diode becomes disproportionately high. We note also that at every value of tuning voltage, the RF voltage across the diode goes positive (relative to the anode) during part of each cycle.

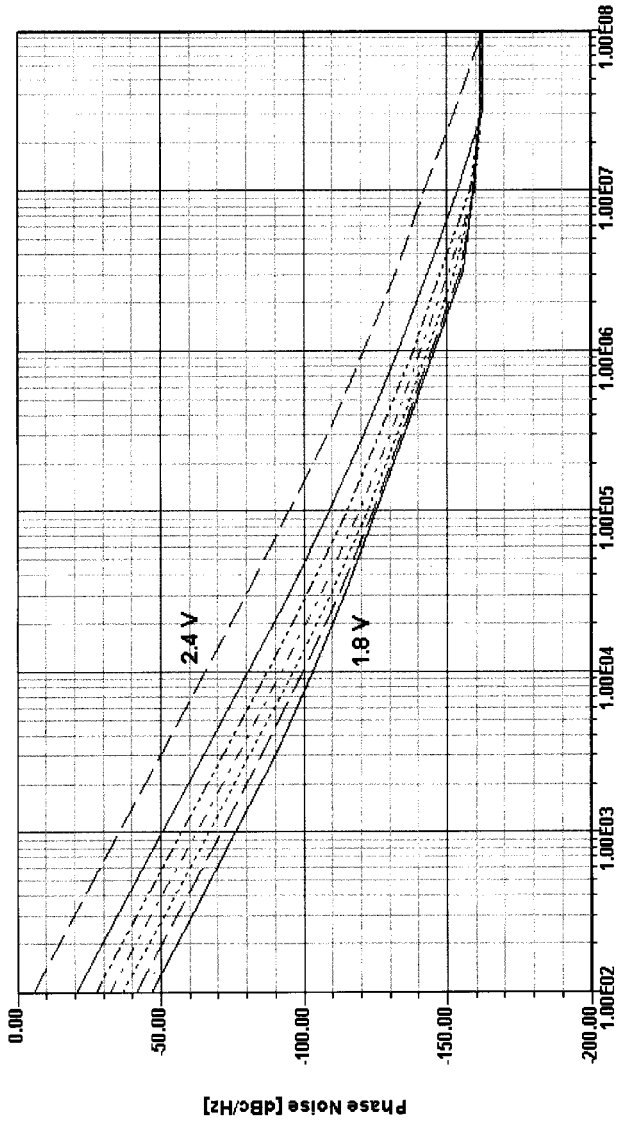


Figure 5-74 Noise-sideband performance of an oscillator at 147 MHz, showing the influence of a tuning diode operated at bias voltages from 1.8 to 2.4 V in 0.1-V increments.

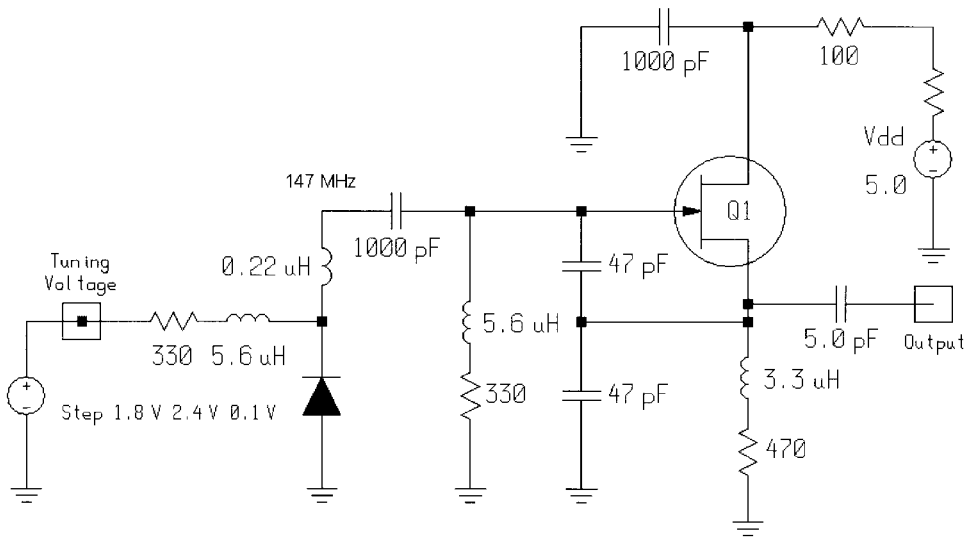


Figure 5-75 The VCO schematic. Because the tuning diode is in series with the resonator, it passes all of the tank's RF current and plays a disproportionately large role in determining the circuit's phase-noise performance.

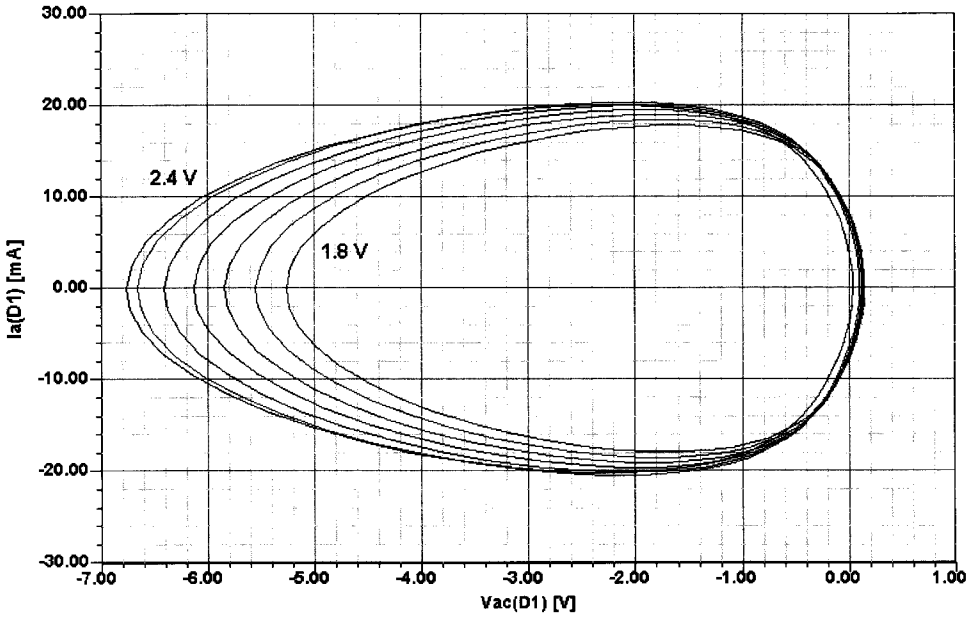


Figure 5-76 Graphing the tuning diode's ac load line reveals that increasing the diode's tuning bias (decreasing its capacitance) results in a disproportionally large increase in the RF voltage across the diode. At all tuning voltages, the RF voltage across the diode is slightly positive (relative to the anode) during part of each cycle.

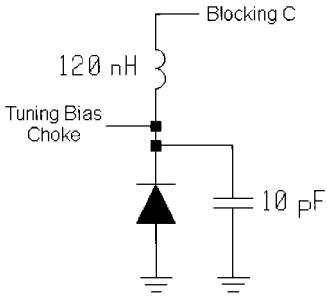


Figure 5-77 Adding capacitance across the tuning diode reduces the effect of its nonlinearities on phase-noise performance. We have also revalued the resonator to 120 nH to maintain oscillation near 147 MHz.

Modifying the oscillator tank to add capacitance in parallel with the tuning diode (Figure 5-77) reduces the tuning diode's effect on tank Q and L/C ratio, and therefore its contribution to phase noise. With this arrangement, we now see the expected relationship between tuning bias and phase noise: The higher the tuning bias, the lower the phase noise (Figure 5-78). Figure 5-79 shows that the diode's ac load line is now virtually unchanged across the tuning-voltage range. The RF voltage across the diode still goes positive (relative to the anode) during part of each cycle.

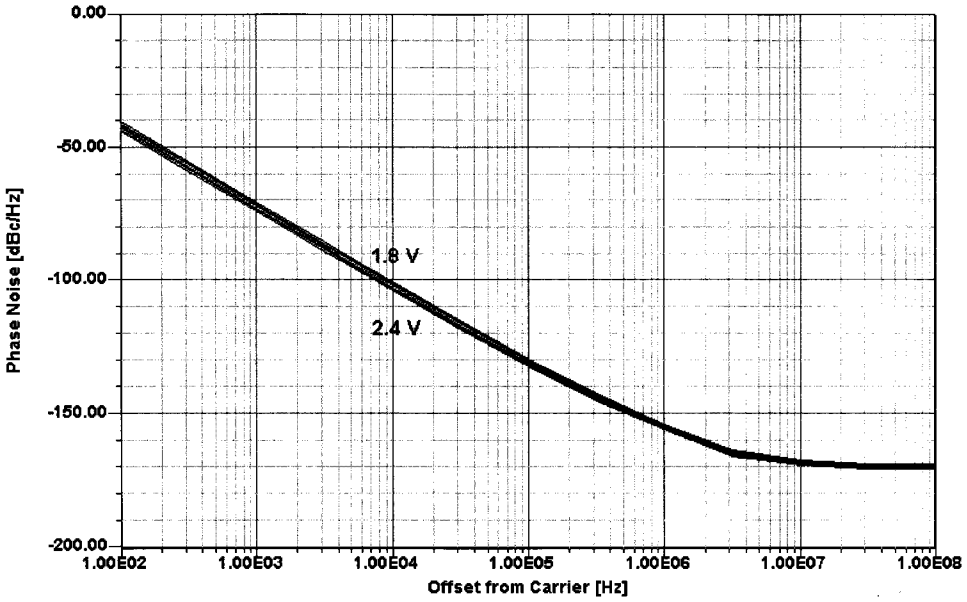


Figure 5-78 Now that the tuning diode passes only part of the tank's RF current, varying its capacitance does not result in the large change in phase-noise performance shown in Figure 5-74. Note also that the phase-noise-versus-tuning-bias relationship has reversed relative to that shown in Figure 5-74: The highest tuning voltage now corresponds to the lowest phase noise.

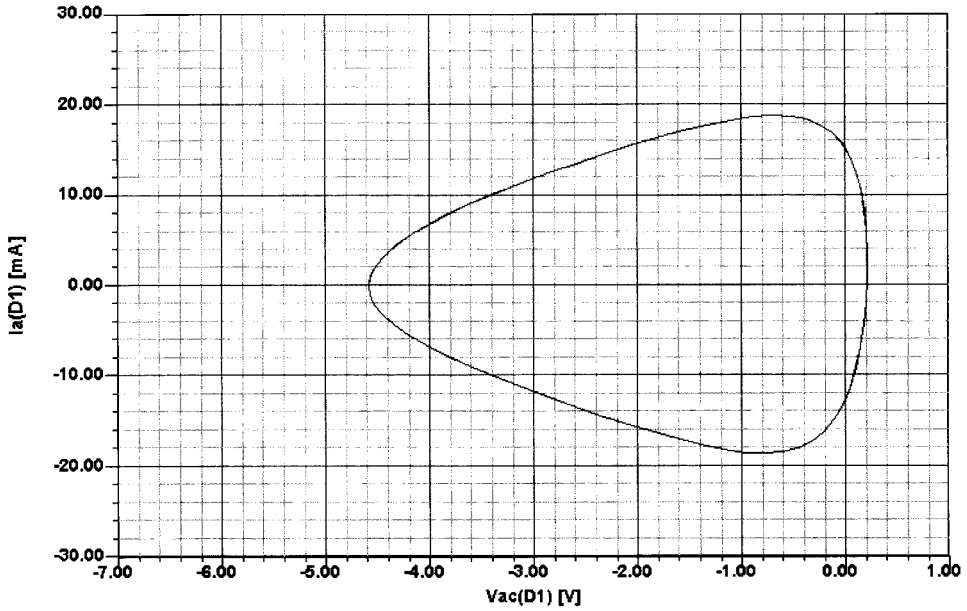


Figure 5-79 The ac load line for the tuning diode in Figure 5-77. Now the voltage swing across the diode is virtually the same at all tuning voltages. The RF voltage across the diode still goes positive (relative to the anode) during part of the cycle.

Figure 5-80 shows one more modification to the VCO tank. Now, the tuning diode is lightly coupled to the hot end of the tank, in parallel with the resonator. The result is phase-noise performance that is essentially unchanged across the 1.8–2.4-V tuning-voltage range. Figure 5-81 compares this circuit's phase-noise performance with the worst-case performance of the tank configurations in Figures 5-75 and 5-77. Figure 5-82 compares the diode ac load lines that correspond to the phase-noise curves in Figure 5-81. The best phase-noise performance corresponds to the diode ac load line that exhibits the narrowest voltage swing and never goes positive relative to the anode.

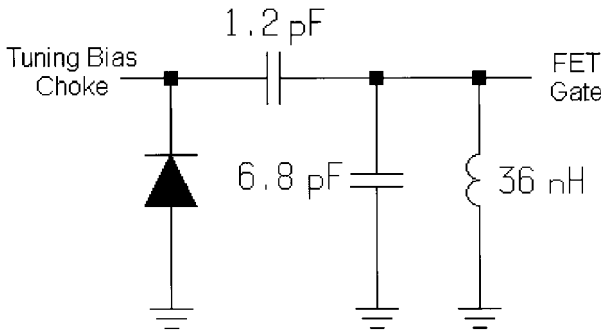


Figure 5-80 Tank circuit reconfigured as per Figure 5-54 to lightly couple the tuning diode in parallel with the resonator, with the resonator inductance again decreased to maintain oscillation near 147 MHz. The 1000-pF capacitor and gate choke of Figure 5-75 (5.6 μH in series with 330 Ω) are no longer needed.

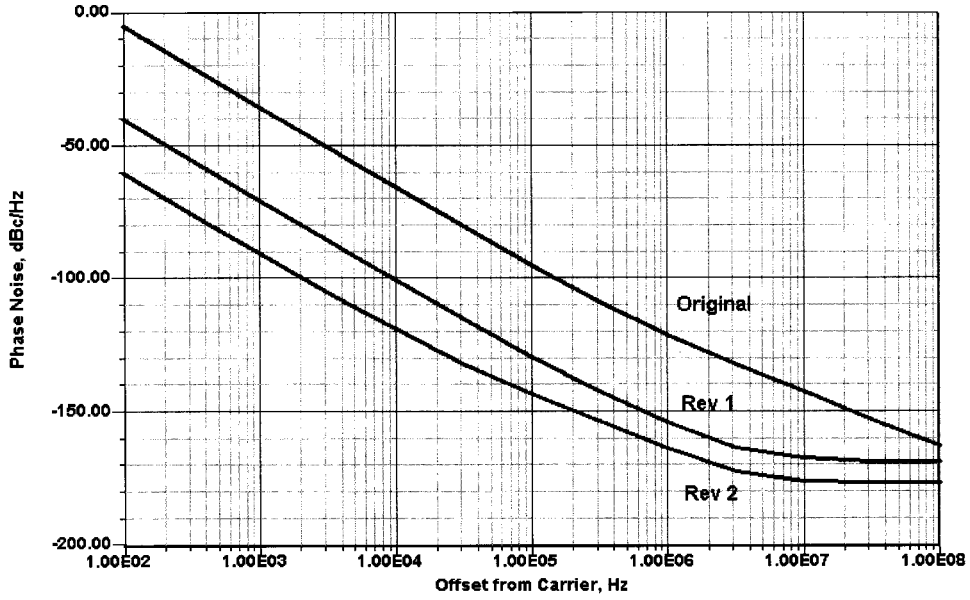


Figure 5-81 Comparison of worst-case phase-noise performance obtained with the tank configurations of Figures 5-75 (Original), 5-77 (Rev 1), and 5-80 (Rev 2).

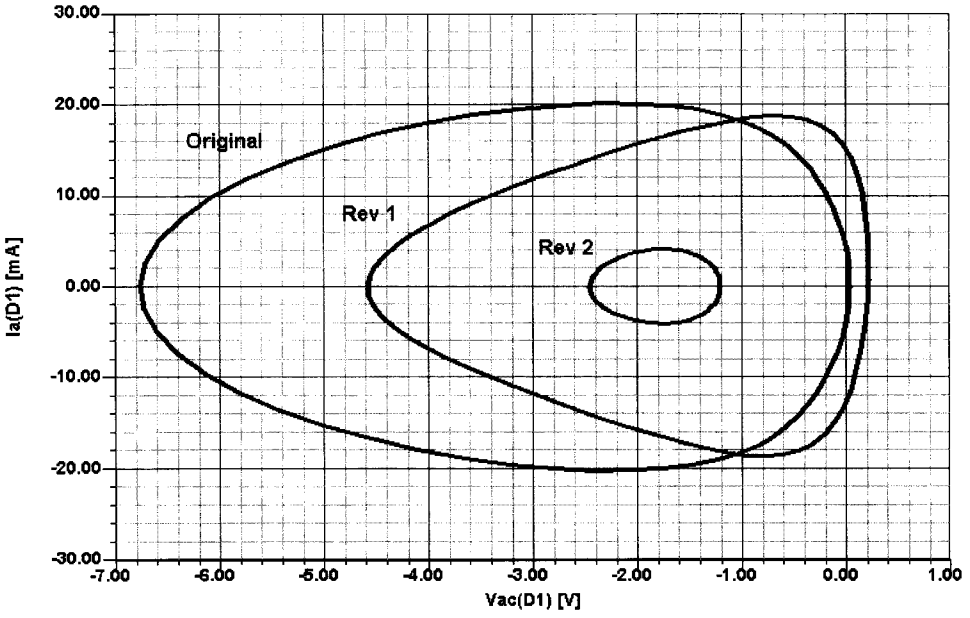


Figure 5-82 Comparison of the tuning diode's ac load line for the worst-case phase-noise curves of Figure 5-79.

It's important to keep in mind that we have not gained something for nothing in achieving the best phase-noise response in Figure 5-81. As we reduced the diode's phase-noise contribution, we also reduced the oscillator's tuning range—from 7.1% (Figure 5-75 circuit) to 1.8% (Figure 5-77 circuit) to 0.046% (Figure 5-80). Simultaneously optimizing phase-noise performance and tuning range requires a combination of techniques. First, we would maximize diode Q by using as high a tuning voltage as feasible. Second, we would use antiseries-connected diodes to minimize the RF voltage swing across them (Figure 5-55). Both of these techniques work against the achievement of wide tuning ranges by minimizing the capacitance swing that the diodes can contribute. As a result of this, and because hyperabrupt diodes are not available for the frequencies at which wireless VCOs operate, we would therefore use multiple antiseries diode pairs in parallel.

It is of interest to compare various oscillators. Figure 5-83 shows the performance of a 10-MHz crystal oscillator, a 40-MHz *LC* oscillator, the HP8640 cavity-tuned oscillator at 500 MHz, the 310–640-MHz switched-reactance oscillator of the HP8662 oscillator, and a 2–6-GHz YIG oscillator at 6 GHz.

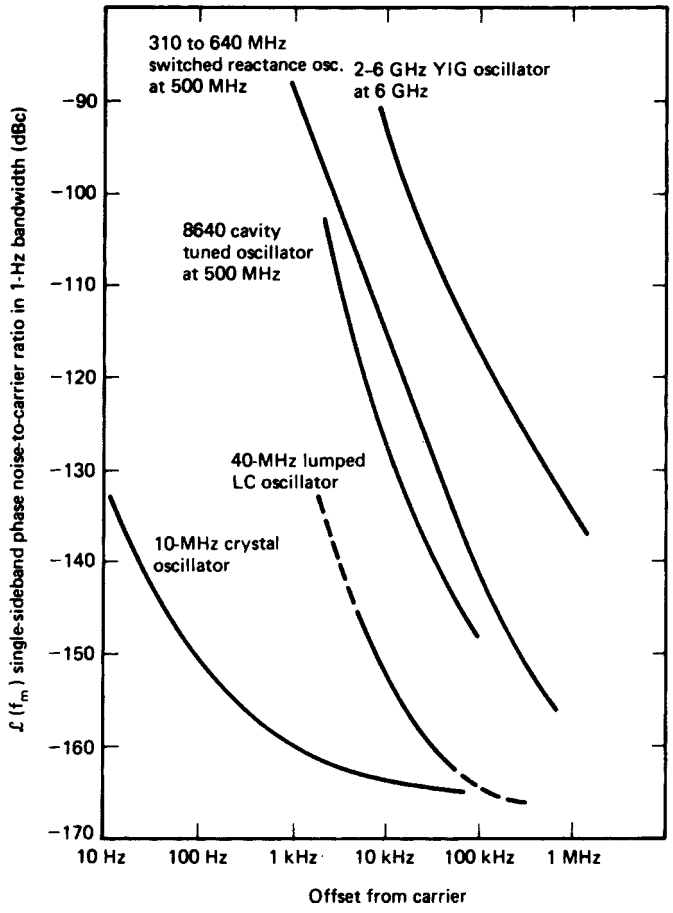


Figure 5-83 Comparison of noise-sideband performances of a crystal oscillator, *LC* oscillator, cavity-tuned oscillator, switched-reactance oscillator, and YIG oscillator.

5-6-3 Nonlinear Approach to the Calculation of Oscillator Phase Noise

The mechanism of noise generation in an oscillator combines the equivalent of frequency conversion (mixing) with the effect of AM-to-PM conversion. Therefore, to calculate oscillator phase noise, we must first be able to calculate the noise figure of a mixer. The following nonlinear approach is based on Ulrich L. Rohde, Chao-Ren Chang, and Jason Gerber, "Design and Optimization of Low-Noise Oscillators Using Nonlinear CAD Tools," *IEEE Frequency Control Symposium Proceedings*, 1994, pp. 548–554. This section presents the use of a novel algorithm for the computation of SSB carrier noise in free-running oscillators using the harmonic-balance (HB) nonlinear technique.

Traditional approaches relying on frequency conversion analysis are not sufficient to describe the complex physical behavior of a noisy oscillator. The accuracy of this nonlinear approach is based on the dynamic range of the harmonic-balance simulator and the quality of the parameter extraction for the active device. The algorithm described has also been verified with several examples up to millimeter wavelengths. This is the only algorithm that provides a complete and rigorous treatment of noise analysis for autonomous circuits.

Noise Generation in Oscillators. As shown earlier, the qualitative linearized picture of noise generation in oscillators is very well known [9]. The physical effects of random fluctuations taking place in the circuit are different depending on their spectral allocation with respect to the carrier:

- Noise components at low-frequency deviations result in frequency modulation of the carrier through mean square frequency fluctuation proportional to the available noise power.
- Noise components at high-frequency deviations result in phase modulation of the carrier through mean square phase fluctuation proportional to the available noise power.

We will demonstrate that the same conclusions can be derived quantitatively from the HB equations for an autonomous circuit.

Equivalent Representation of a Noisy Nonlinear Circuit. A general noisy nonlinear network can be described by the equivalent circuit shown in Figure 5-84. The circuit is divided into linear and nonlinear subnetworks as noise-free multiports. Noise generation is accounted for by connecting a set of noise voltage and noise current sources at the ports of the linear subnetwork [9–18].

Frequency Conversion Approach. The circuit supports a large-signal time-periodic steady state of fundamental angular frequency ω_0(carrier). Noise signals are small perturbations superimposed on the steady state, represented by families of pseudosinusoids located at the sidebands of the carrier harmonics. Therefore, the noise performance of the circuit is determined by the exchange of power among the sidebands of the unperturbed steady state through frequency conversion in the nonlinear subnetwork. Due to the perturbative assumption, the nonlinear subnetwork can be replaced with a multifrequency linear multiport described by a conversion matrix. The flow of noise signals can be computed by means of conventional linear circuit techniques.

The frequency conversion approach frequently used has limitations. The frequency-conversion approach is not sufficient to predict the noise performance of an autonomous

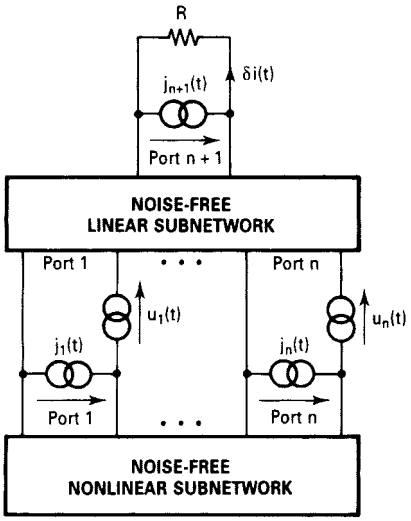


Figure 5-84 Equivalent circuit of a general noisy nonlinear network.

circuit. The spectral density of the output noise power and, consequently, the PM noise computed by the conversion analysis are proportional to the available power of the noise sources.

- In the presence of both thermal and flicker noise sources, PM noise increases as ω^{-1} for $\omega \to 0$ and tends to a finite limit for $\omega \to \infty$.
- Frequency conversion analysis correctly predicts the far-carrier noise behavior of an oscillator and, in particular, the oscillator noise floor; it does not provide results consistent with the physical observations at low deviations from the carrier.

This inconsistency can be removed by adding the modulation noise analysis. In order to determine the far-away noise using the autonomous circuit perturbation analysis, the following applies.

The circuit supports a large-signal time-periodic autonomous regime. The circuit is perturbed by a set of small sources located at the carrier harmonics and at the sidebands at a deviation ω from carrier harmonics. The perturbation of the circuit state (δX_B, δX_H) is given by the uncoupled sets of equations:

$$\frac{\partial E_B}{\partial E_B}\,\delta X_B = J_B(\omega) \tag{5-127}$$

$$\frac{\partial E_H}{\partial E_H}\,\delta X_H = J_H(\omega) \tag{5-128}$$

where E_B, E_H = vectors of HB errors
X_B, X_H = vectors of state-variable (SV) harmonics (since the circuit is autonomous, one of the entries X is replaced by the fundamental frequency ω_0)
J_B, J_H = vectors of forcing terms

The subscripts *B* and *H* denote sidebands and carrier harmonics, respectively.

For a spot noise analysis at a frequency ω, the noise sources can be interpreted in either of two ways:

- Pseudosinusoids with random amplitude and phase located at the sidebands. This noise generation mechanism is described by Eq. (5-127), which is essentially a frequency conversion equation relating the sideband harmonics of the state variables and of the noise sources. This description is exactly equivalent to the one provided by the frequency conversion approach. This mechanism is referred to as *conversion noise* [19–21].
- Sinusoids located at the carrier harmonics and randomly phase and amplitude modulated by pseudosinusoidal noise at frequency ω. This noise generation mechanism is described by Eq. (5-128), which describes noise-induced jitter of the circuit-state, represented by the vector δX_H. The modulated perturbing signals are represented by replacing the entries of J_H with the complex modulation laws. This mechanism is referred to as *modulation noise.* One of the entries of δX_H is $\delta\omega_0$, where $\delta\omega_0(\omega)$ is the phasor of the pseudosinusoidal components of the fundamental frequency fluctuations in a 1-Hz band at frequency ω. Equation (5-128) provides a frequency jitter with a mean square value proportional to the available noise power. In the presence of both thermal and flicker noise, PM noise increases as ω^{-3} for $\omega \to 0$ and tends to 0 for $\omega \to \infty$. Modulation noise analysis correctly describes the noise behavior of an oscillator at low deviations from the carrier and does not provide results consistent with physical observations at high deviations from the carrier.

The combination of both phenomena explains the noise in the oscillator shown in Figure 5-85, where the near-carrier noise dominates below ω_X and far-carrier noise dominates above ω_X. Figure 5-86 (itemized form) shows the noise sources as they are applied at the IF. We have arbitrarily defined the low oscillator output as IF. This applies to the conversion matrix calculation.

Figure 5-87 shows the contributions that have to be taken into consideration for calculation of the noise at the output. The accuracy of the calculation of the phase noise depends highly on the quality of the parameter extraction for the nonlinear device; in particular, high-frequency phenomena must be modeled properly. In addition, the flicker noise contribution is essential. This is also valid for mixer noise analysis.

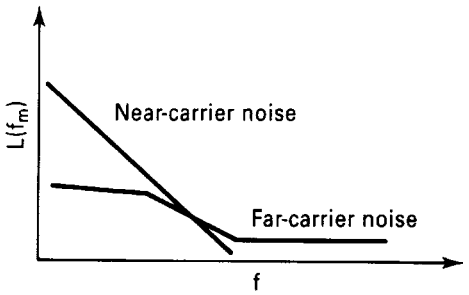


Figure 5-85 Oscillator noise components.

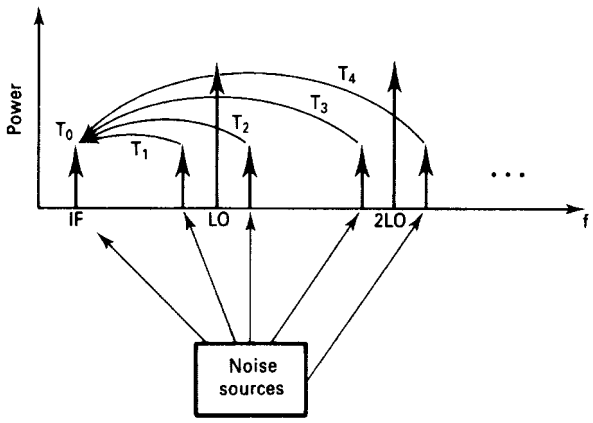


Figure 5-86 Noise sources where the noise at each sideband contributes to the output noise at the IF through frequency conversion.

Figure 5-88 shows the measured phase noise at 10 kHz of the carrier of an 800-MHz oscillator and the simulated result. The good agreement between measured and predicted data is worth noticing. At currents above 5 mA, the simulation becomes inaccurate as breakdown effects occur that were not part of the simulation.

Conversion Noise Analysis. The actual mathematics used to calculate the noise result (Ansoft Serenade 8.*x*) are as follows.

kth Harmonic PM Noise

$$\langle |\delta\Phi_k(\omega)|^2 \rangle = \frac{N_k(\omega) - N_{-k}(\omega) - 2\,\mathrm{Re}[C_k(\omega)]}{R|I_k^{SS}|^2} \tag{5-129}$$

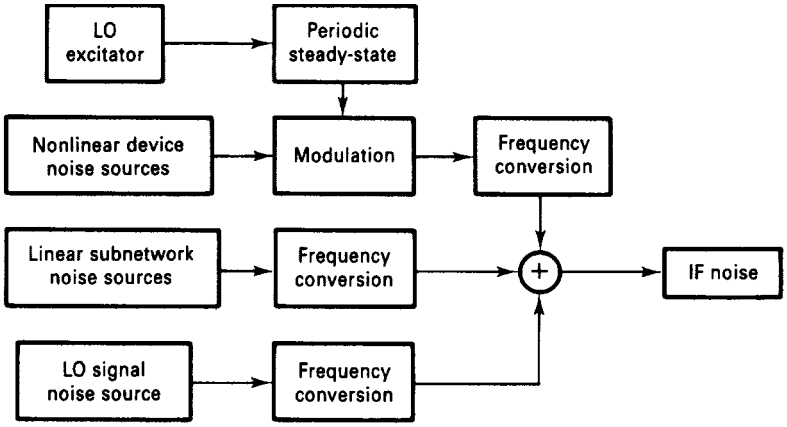


Figure 5-87 Noise mechanisms.

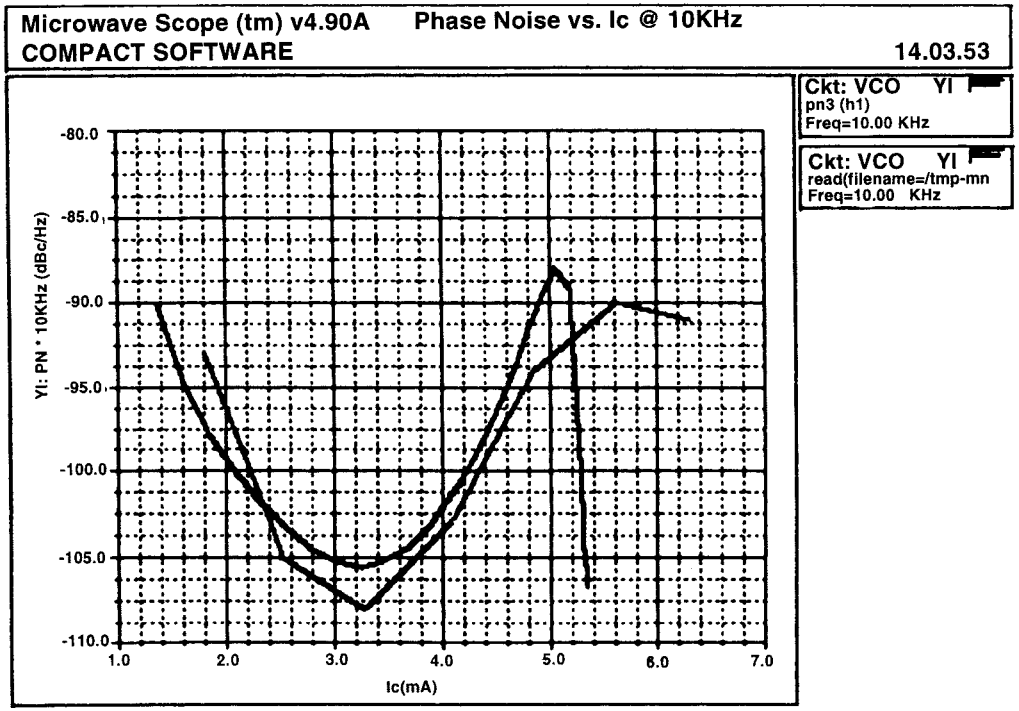


Figure 5-88 Bias-dependent measured and predicted phase noise at 10 kHz off the carrier of an 800-MHz oscillator.

kth Harmonic AM Noise

$$\langle |\delta A_k(\omega)|^2 \rangle = 2\,\frac{N_k(\omega) - N_{-k}(\omega) + 2\,\mathrm{Re}[C_k(\omega)]}{R|I_k^{SS}|^2} \tag{5-130}$$

kth Harmonic PM–AM Correlation Coefficient

$$C_k^{\mathrm{PMAM}}(\omega) = \langle \delta\Phi_k(\omega)\delta A_k(\omega)^* \rangle = -\sqrt{2}\,\frac{2\,\mathrm{Im}[C_k(\omega)] + j[N_k(\omega) - N_{-k}(\omega)]}{R|I_k^{SS|^2}} \tag{5-131}$$

where

$N_k(\omega)$, $N_{-k}(\omega)$ = noise power spectral densities at the upper and lower sidebands of the kth harmonic
$C_k(\omega)$ = normalized correlation coefficient of the upper and lower sidebands of the kth carrier harmonic
R = load resistance
I_k^{SS} = kth harmonic of the steady-state current through the load

Modulation Noise Analysis

kth Harmonic PM Noise

$$\langle |\delta\Phi_k(\omega)|^2 \rangle = \frac{k^2}{\omega^2}\,\mathrm{T}_F\,\langle \mathrm{J}_H(\omega)\mathrm{J}_H^t(\omega)\rangle \mathrm{T}_F^t \tag{5-132}$$

kth Harmonic AM Noise

$$\langle |\delta A_k(\omega)|^2 \rangle = \frac{2}{|I_k^{SS}|^2}\,\mathrm{T}_{Ak}\,\langle \mathrm{J}_H(\omega)\mathrm{J}_H^t(\omega)\rangle \mathrm{T}_{Ak}^t \tag{5-133}$$

kth Harmonic PM–AM Correlation Coefficient

$$C_k^{\mathrm{PMAM}}(\omega) = \langle \delta\Phi_k(\omega)\delta A_k(\omega)^* \rangle = \frac{k\sqrt{2}}{j\omega|I_K^{SS}|^2}\,\mathrm{T}_F\langle \mathrm{J}_H(\omega)\mathrm{J}_H^t(\omega)\rangle \mathrm{T}_{Ak}^t \tag{5-134}$$

where $\mathrm{J}_H(\omega)$ = vector of Norton equivalent of the noise sources
T_F = frequency transfer matrix
I_k^{SS} = kth harmonic of the steady-state current through the load

Experimental Variations. We will look at two examples where we compare predicted and measured data. Figure 5-89 shows the abbreviated circuit of a 10-MHz crystal oscillator. It uses a high-precision, high-Q crystal made by companies such as Bliley. Oscillators like this are intended for use as frequency and low-phase-noise standards. In this case, the circuit under consideration is part of the HP3048 phase-noise measurement system.

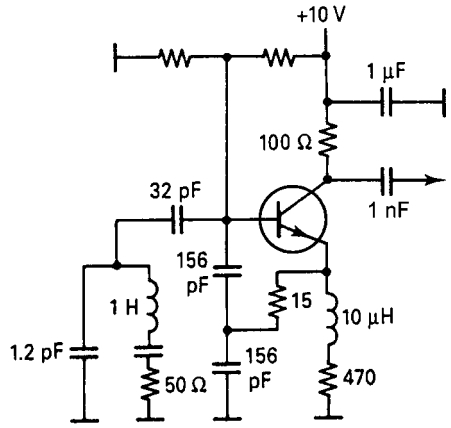


Figure 5-89 Abbreviated circuit of a 10-MHz crystal oscillator.

Figure 5-90 shows the measured phase noise of this HP frequency standard, and Figure 5-91 shows the phase noise predicted using the mathematical approach outlined above.

In cooperation with Motorola, we also analyzed an 800-MHz VCO. In this case, we also did the parameter extraction for the Motorola transistor. Figure 5-92 shows the circuit, a Colpitts oscillator that uses RF feedback in the form of a 15-Ω resistor and a capacitive

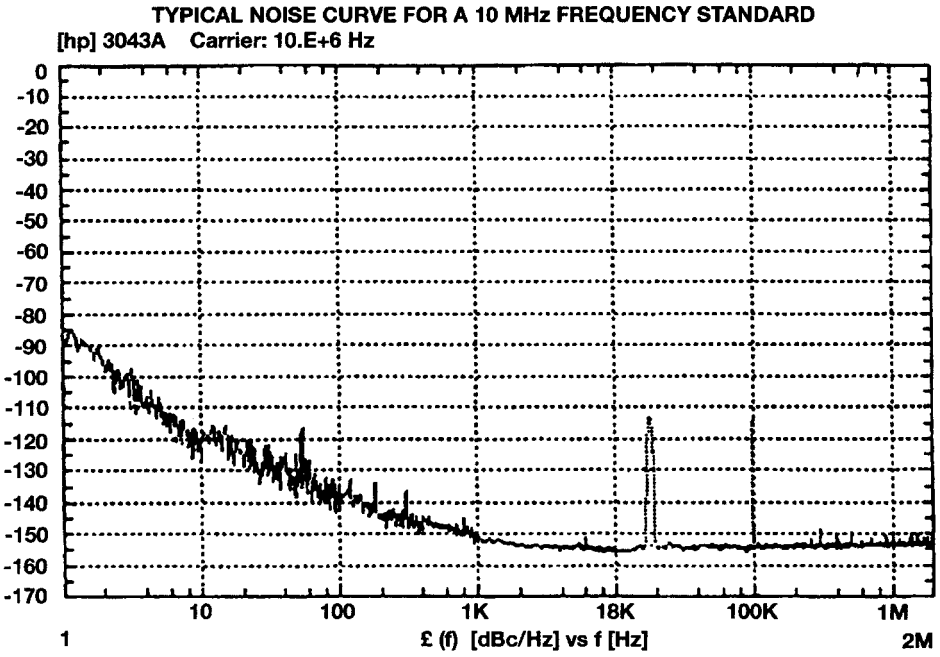


Figure 5-90 Measured phase noise for this frequency standard by HP.

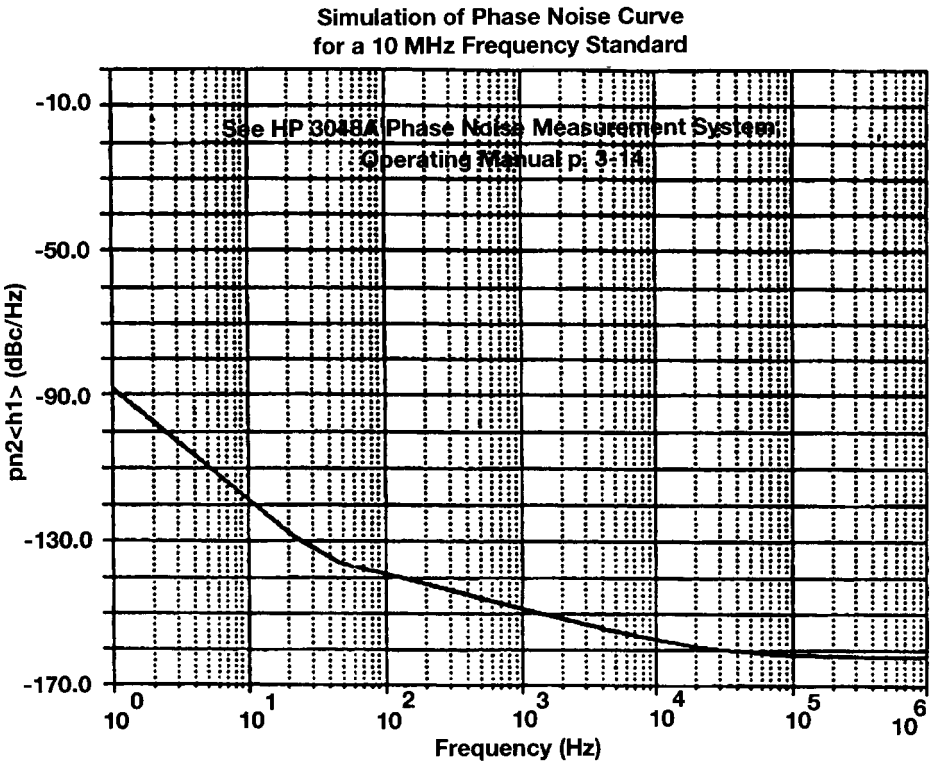


Figure 5-91 Simulated phase noise of the oscillator shown in Figure 5-89.

voltage divider consisting of 1 pF between the BJT's base and the feedback resistor, and 1 pF between the feedback resistor and common. Also, the tuned circuit is loosely coupled to this part of the transistor circuit. Figure 5-93 shows a comparison between predicted and measured phase noise for this oscillator. Figures 5-94, 5-95, and 5-96 show the oscillator's predicted output waveform, spectrum, and dc I–V and ac load line responses, respectively.

Phase Noise Optimization. By allowing the simulator's optimizer to vary the oscillator's feedback and dc operating point, phase noise can be improved. The values that were allowed to vary were the values of the capacitors in the feedback voltage divider, the value of a negative-feedback resistor (starting value, 15 Ω) that reduces AM–PM conversion, and the emitter resistor, which changes the transistor's dc bias and therefore affects its bias-dependent flicker noise. Figure 5-97 shows the improvement of the 800-MHz VCO as previously shown. While the close-in phase noise can be improved drastically by approximately 32 dB, the far-out noise at 20 MHz and beyond is deteriorated. This is due to the resistive feedback, which reduces AM-to-PM conversion [22–24].

In testing an oscillator's resistance to frequency pushing (frequency shift with supply-voltage changes), variations in phase noise can be observed. Phase noise is also bias dependent,

and this fact can be used for phase-noise optimization. Figure 5-98 shows the phase noise of an oscillator as its supply voltage is varied.

Summary. This section has demonstrated the power of harmonic-balance simulation in the design of low-phase-noise oscillators. The simulator can be used to accurately predict close-in and far-out phase noise based on the contribution of the transistor, the loaded Q of the resonator, and, in the case of a VCO, the contribution of the tuning diode.

This is the first time that a bias-dependent noise model for both FETs and bipolar transistors has been developed and used for this type of application, in which optimization is allowed to vary circuit components and the bias point for best performance.

Practical examples ranging from a 10-MHz crystal oscillator to a 1-GHz CR oscillator have been supplied. The theoretical background for the nonlinear oscillator and its validation have been presented.

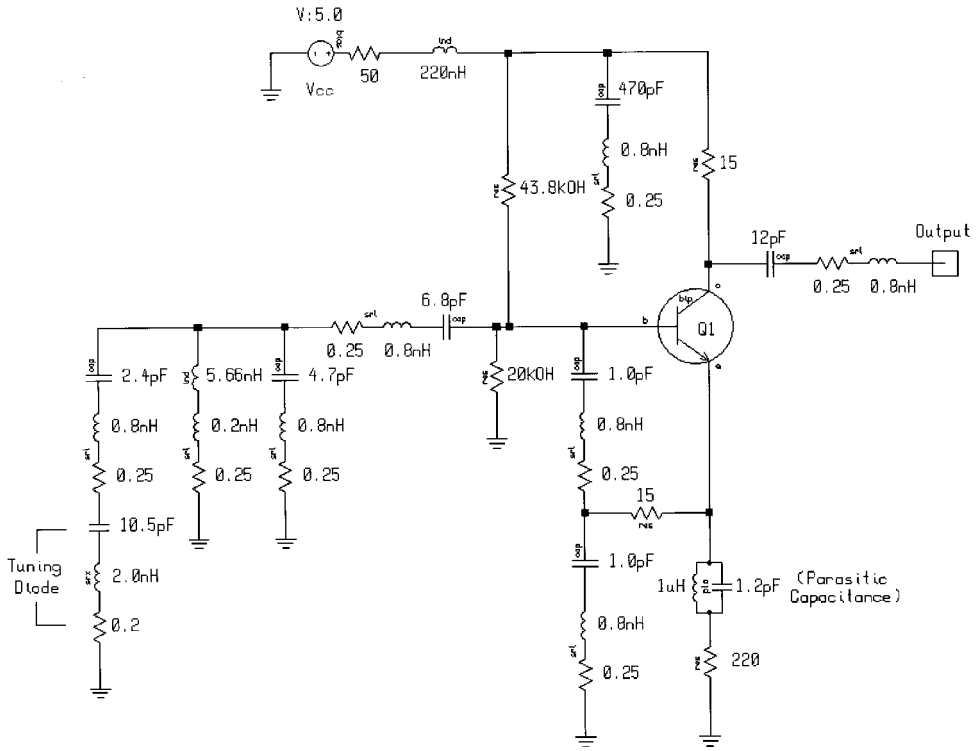


Figure 5-92 Colpitts oscillator that uses RF negative feedback between the emitter and capacitive voltage divider. To be realistic, we used actual components rather than ideal ones. The suppliers for the capacitors and inductors provide some typical values for the parasitics. The major changes are 0.8 nH and 0.25 Ω in series with the capacitors. The same thing applies for the main inductance, which has a parasitic connection inductance of 0.2 nH in series with a 0.25-Ω resistance. These types of parasitics are valid for a fairly large range of components assembled in surface-mount applications. Most engineers model the circuit only by assuming lossy devices, not adding these important parasitics. One of the side-effects we have noticed is that the output power is more realistic and, needless to say, the simulated phase noise agrees quite well with measured data. This circuit can also serve as an example for modeling amplifiers and mixers using surface-mount components.

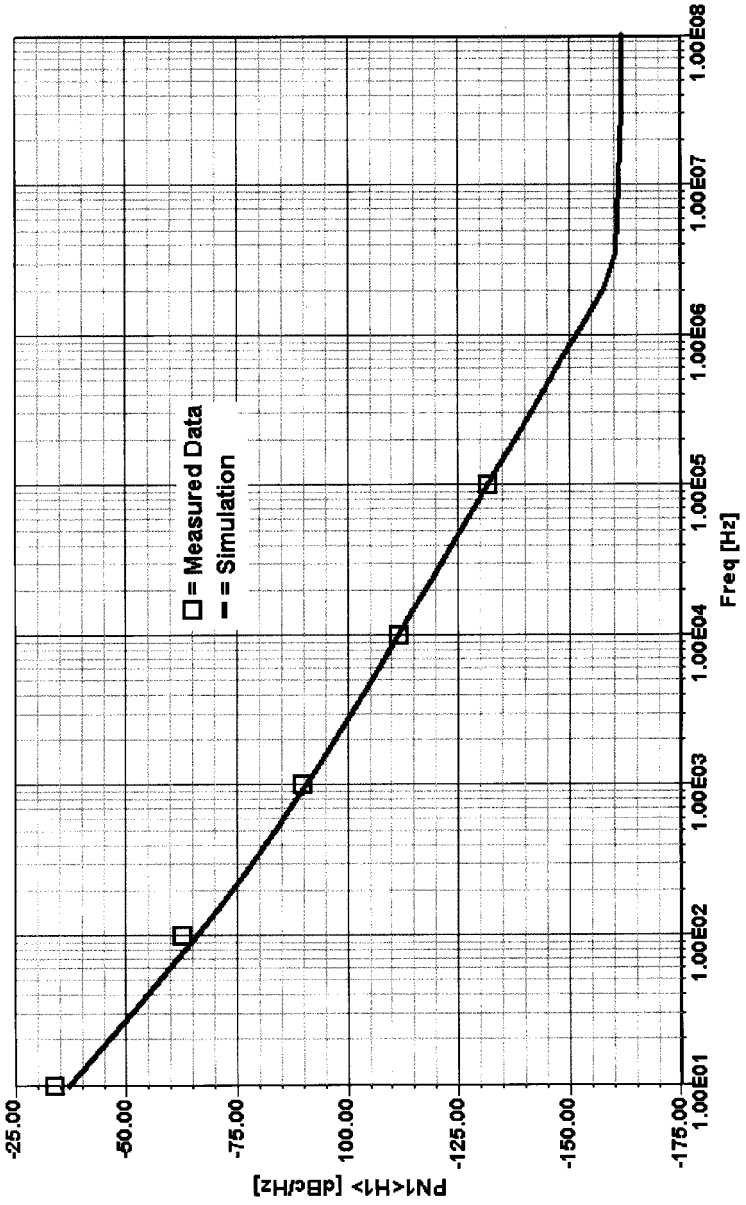


Figure 5-93 Comparison between predicted and measured phase noise for the oscillator shown in Figure 5-92.

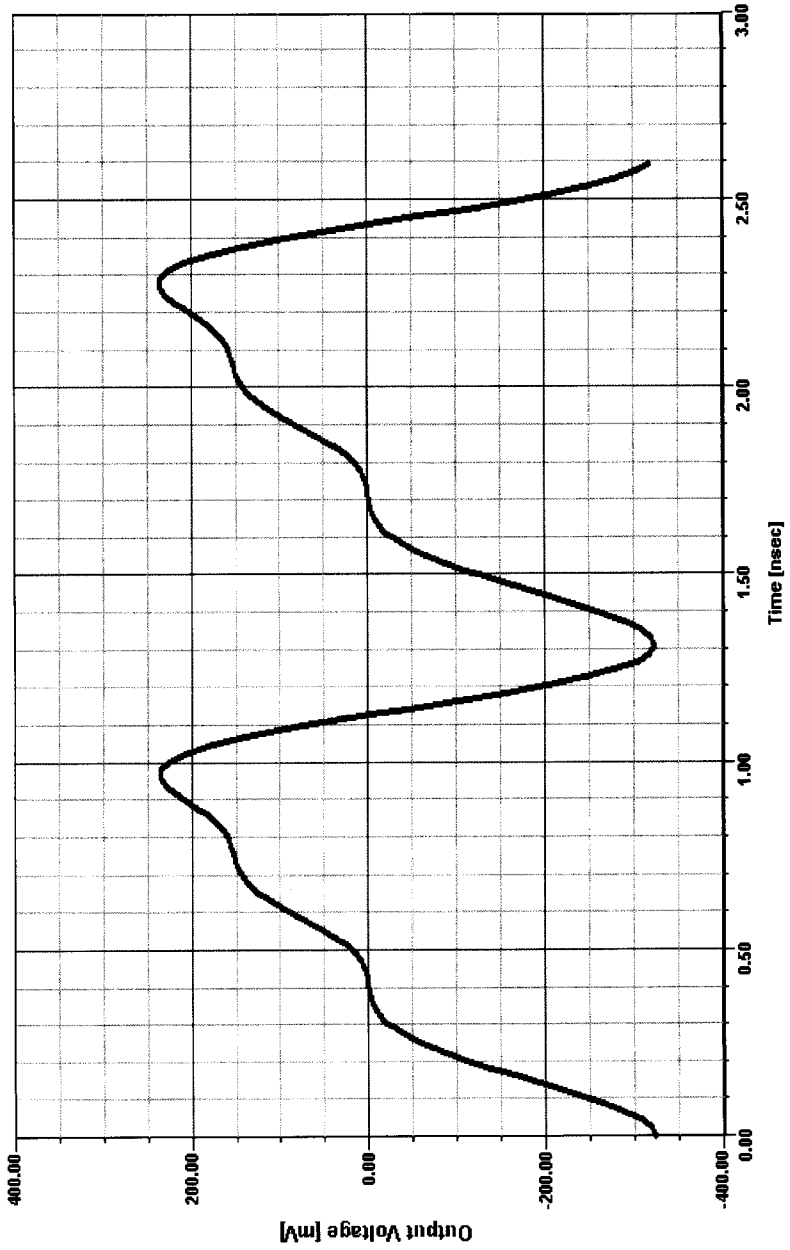


Figure 5-94 Predicted output waveform for the oscillator shown in Figure 5-92.

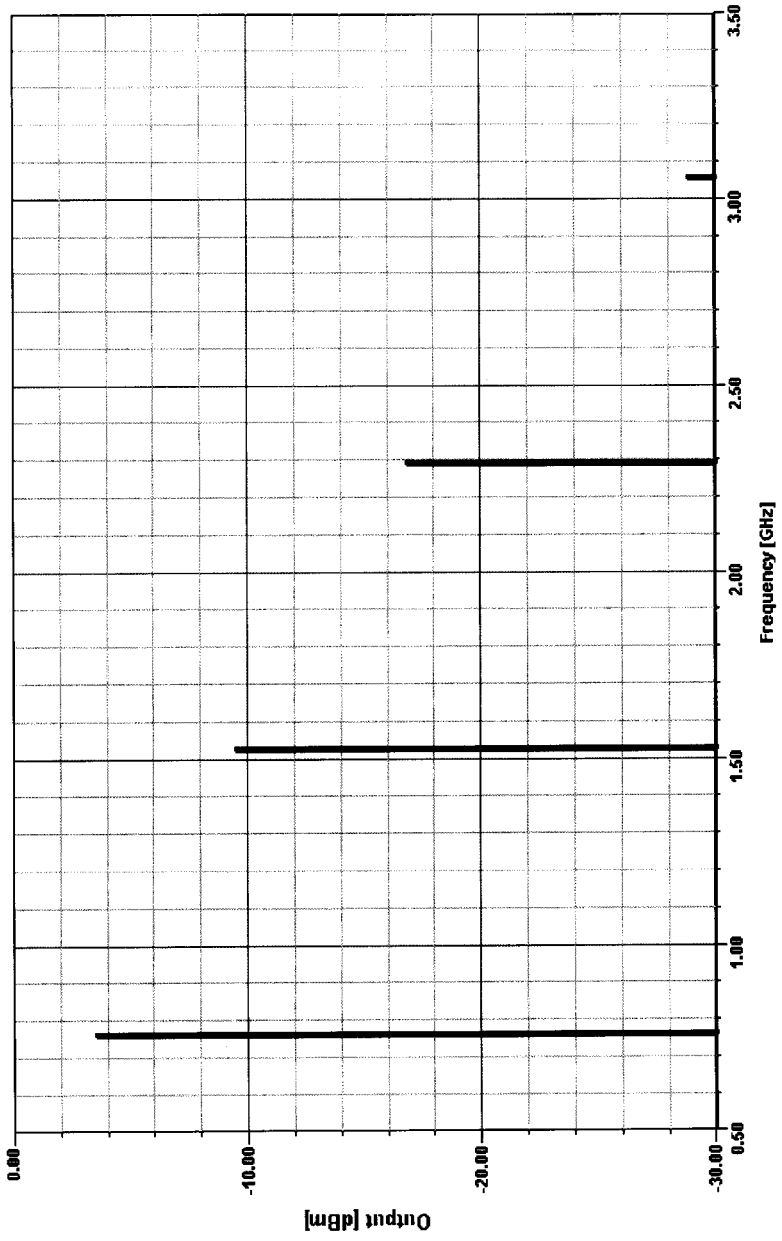


Figure 5-95 Predicted output spectrum for the oscillator shown in Figure 5-92.

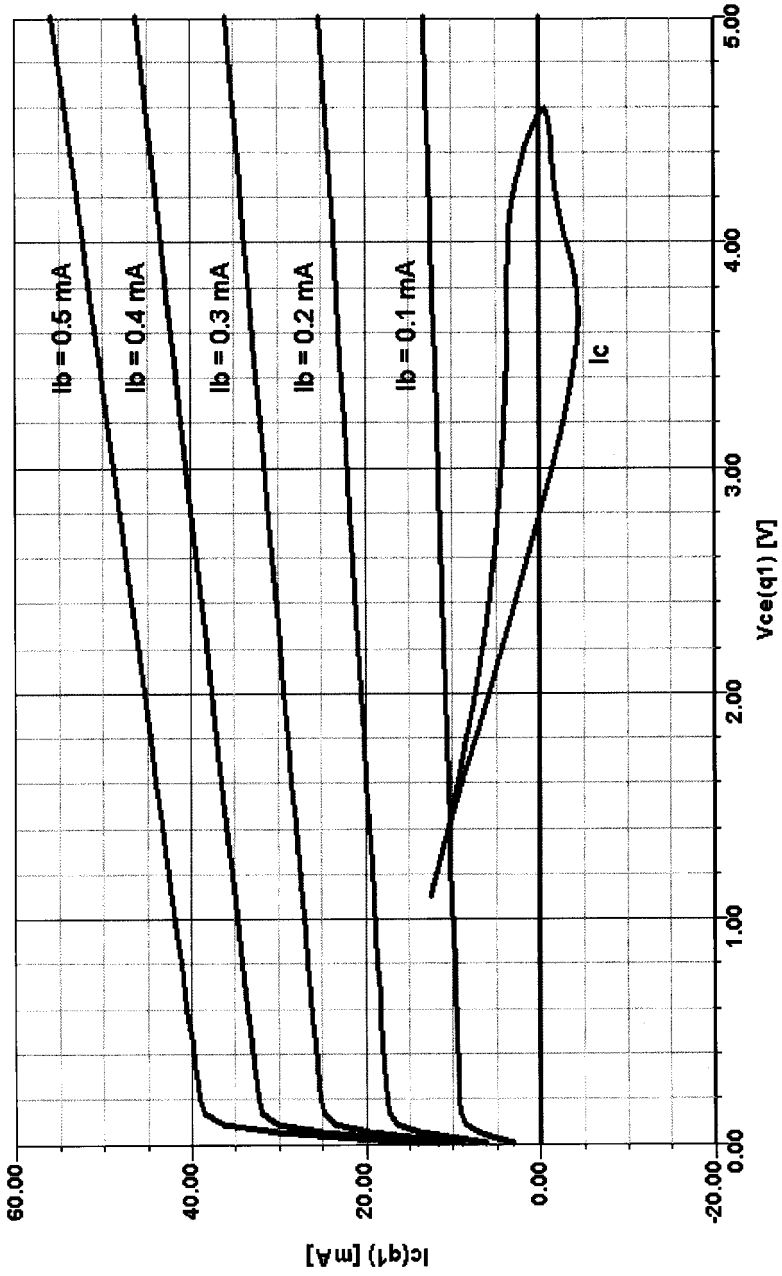


Figure 5-96 The ac load line and dc *I–V* curves for the transistor in the oscillator shown in Figure 5-92.

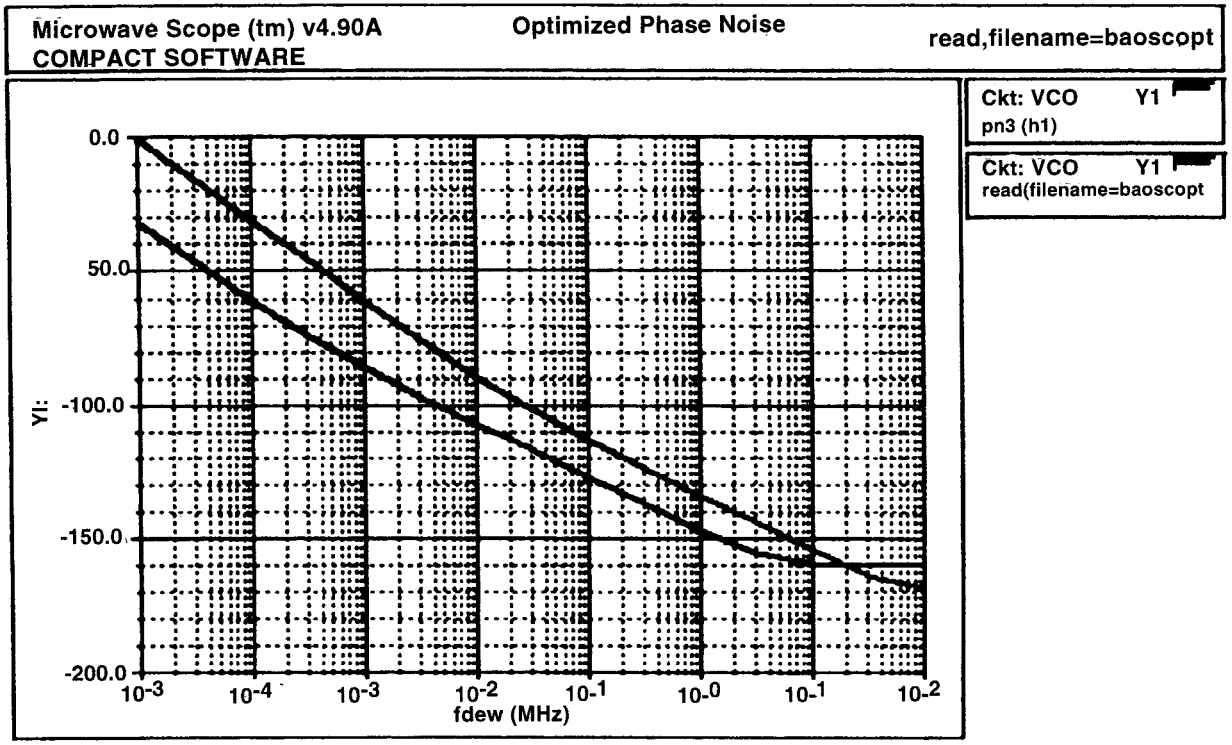


Figure 5-97 Phase noise of the 800-MHz VCO before and after optimization.

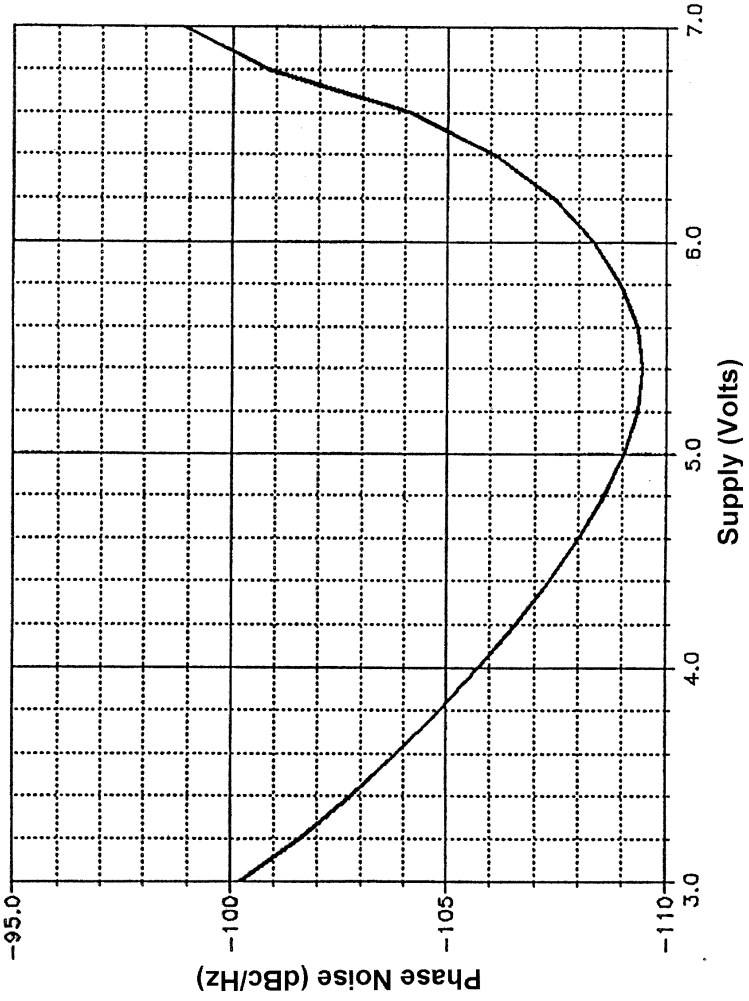


Figure 5-98 Calculated phase noise as a function of the supply voltage for a BJT oscillator. The graph shows a distinct minimum for a particular bias, which could be translated into a collector current change. Such phase-noise minimization could also be accomplished by changing just the voltage applied to the bias portion of the oscillator.

Acknowledgment. This research work for nonlinear operation in oscillators, partly sponsored by the U.S. Department of Defense MIMIC program, was performed in cooperation with Professor Vittorio Rizzoli's group of the Department of Electrical Engineering, Information and Systems Engineering at the University of Bologna, Italy, and the Motorola Portable Products Division, Florida, as well as the engineering team of Ansoft Corporation in New Jersey.

5-7 OSCILLATORS IN PRACTICE

5-7-1 Oscillator Specifications

Although we have tended to emphasize phase noise as a critical indicator of oscillator quality, phase noise is only one of a number of oscillator performance criteria that must be considered in wireless system design. In addition to phase noise, the following characteristics are used by commercial companies to describe oscillator performance:

Frequency Pushing. Frequency pushing characterizes the degree to which an oscillator's frequency is affected by its supply voltage. A sudden current surge caused by activating a transceiver's RF power amplifier may produce a spike on the VCO's dc power supply and a consequent frequency jump. Like tuning sensitivity (below), pushing is specified in frequency/voltage form and is tested by varying the VCO's dc supply voltage (typically ± 1 V) with its tuning voltage held constant.

Harmonic Output Power. The harmonic content is measured relative to the output power. Typical values are 20 dB or more suppression relative to the fundamental. This suppression can be improved by additional filtering.

Output Power. The output power of the oscillator, typically expressed in dBm, is measured into a 50-Ω load. The output power is always combined with a specification for flatness or variation. A typical specification would be 0 dBm ± 1 dB.

Output Power as a Function of Temperature. All active circuits vary in performance as a function of temperature. An oscillator's output power over a temperature range should vary less than a specified value, such as 1 dB.

Post-tuning Drift. After a voltage step is applied to the tuning diode input, the oscillator frequency may continue to change until it settles to a final value. This post-tuning drift is one of the parameters that limits the bandwidth of the VCO input.

Power Consumption. This characteristic conveys the dc power, usually specified in milliwatts and sometimes qualified by operating voltage, required by the oscillator.

Sensitivity to Load Changes. To keep costs down, many wireless applications use a VCO alone, without the buffering action of a high reverse-isolation amplifier stage. In such applications, *frequency pulling*, the change of frequency resulting from partially reactive loads, is an important oscillator characteristic. Pulling is commonly specified in terms of the frequency shift that occurs when the oscillator is connected to a load that exhibits a non-unity VSWR (such as 1.75, usually referenced to 50 Ω), compared to the frequency that results with a unity VSWR load (usually 50 Ω). Frequency pulling must be minimized, especially in cases where power stages are close to the VCO unit

and short pulses may affect the output frequency. Such feedback may make phase locking impossible.

Spurious Outputs. A VCO's spurious output specification, expressed in decibels, enumerates the strength of unwanted and nonharmonically related components relative to the oscillator fundamental. Because a stable, properly designed oscillator is inherently clean, such *spurs* are typically introduced only by external sources in the form of radiated or conducted interference. See *harmonic output power.*

Temperature Drift. Although the synthesizer is responsible for locking and maintaining the oscillator's frequency, the VCO's frequency change as a function of temperature is a critical parameter and must be specified. Its value varies between 10 kHz/°C to several hundred kHz/°C depending on the center frequency and tuning range.

Tuning Characteristic. This specification shows the relationship, depicted as a graph, between a VCO's operating frequency and the tuning voltage applied. Ideally, the correspondence between operating frequency and tuning voltage is linear. See *tuning linearity.*

Tuning Linearity. For stable synthesizers, a constant deviation of frequency versus tuning voltage is desirable. It is also important to make sure that there are no breaks in the tuning range—for example, that the oscillator does not stop operating with a tuning voltage of 0 V. See *tuning characteristic.*

Tuning Sensitivity, Tuning Performance. This characteristic, typically expressed in megahertz per volt (MHz/V), enumerates how much a VCO's frequency changes per unit of tuning-voltage change.

Tuning Speed. This characteristic is defined as the time necessary for the VCO to reach 90% of its final frequency on the application of a tuning-voltage step. Tuning speed depends on the internal components between the input pin and tuning diode—including, among other things, the capacitance present at the input port. The input port's parasitic elements determine the VCO's maximum possible modulation bandwidth.

5-7-2 More Practical Circuits

Because of the need for integration, we will present three IC oscillators of increasing complexity. The circuit in Figure 5-99 is found in early Siemens ICs. Figure 5-100 shows its phase-noise performance.

Figure 5-101 shows the basic push–pull oscillator topology underlying the circuit in Figure 5-99. The omission of the current source and active bias components improves the simpler circuit's phase-noise performance over that of the circuit shown in Figure 5-99. Figure 5-102 compares the phase noise of this basic circuit to the integrated version in Figure 5-99.

Figure 5-103 is for those who speculate on the possibility of replacing the BJTs in Figure 5-101 with LDMOS FETs or JFETs. Figure 5-104 compares its phase noise to that of the BJT case. The rule of thumb is that FETs do not have enough gain for high-Q oscillation in oscillators above 400 MHz, and bipolar transistors are a better choice. Because of their higher flicker noise contribution, GaAsFETs should be considered only at frequencies above 4 or 5 GHz.

Figure 5-105 shows a circuit based on two differential amplifiers and frequently found in Gilbert cell IC applications. Figure 5-106 shows its phase-noise performance.

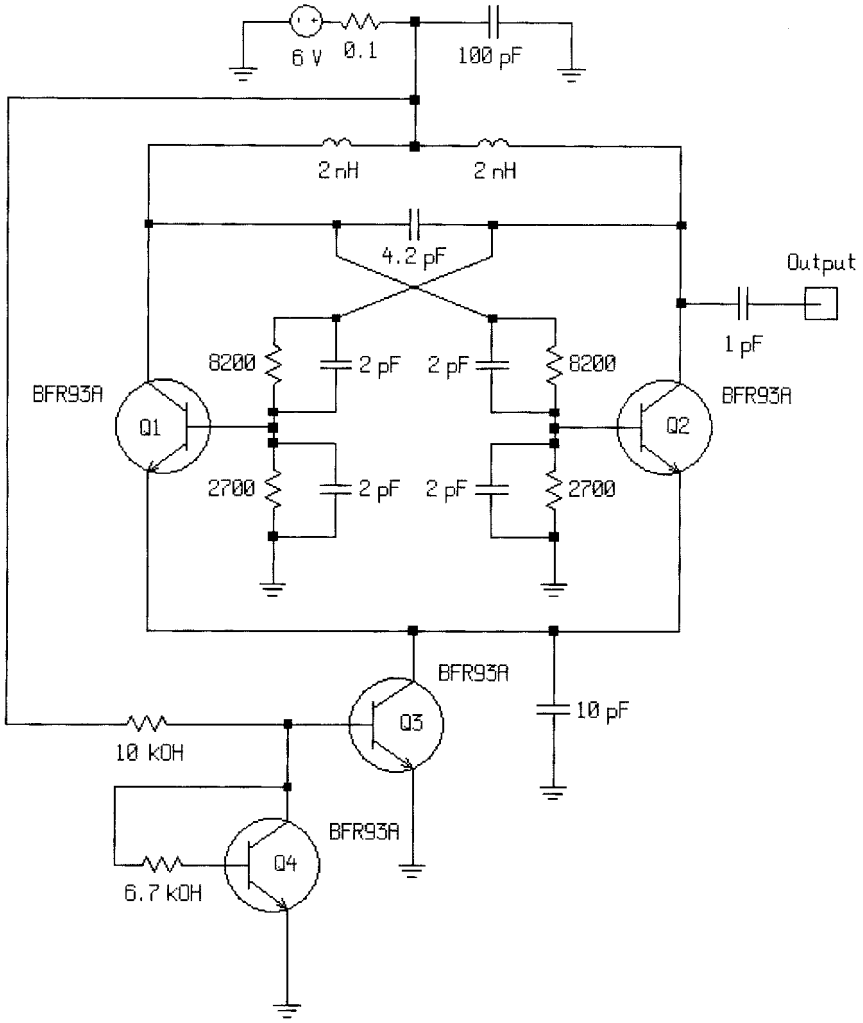


Figure 5-99 Schematic of the Siemens IC oscillator.

Figure 5-107 is a validation circuit that demonstrates a comparison of measured phase-noise performance versus its prediction by circuit simulation software. The circuit is based on the core of Motorola's MC12148 ECL oscillator IC, the topology of which is derived from the superseded MC1648. Motorola specifies the MC12148's typical phase-noise performance as –90 dBc/Hz at an offset of 25 kHz. The close agreement between the measured and simulated results highlights the value of state-of-the-art CAD tools in wireless oscillator design (Figure 5-108).

Figure 5-108 presents the results of *two* simulations: one with ideal L and C components, and another with realistic values specified for these reactances. In using a circuit

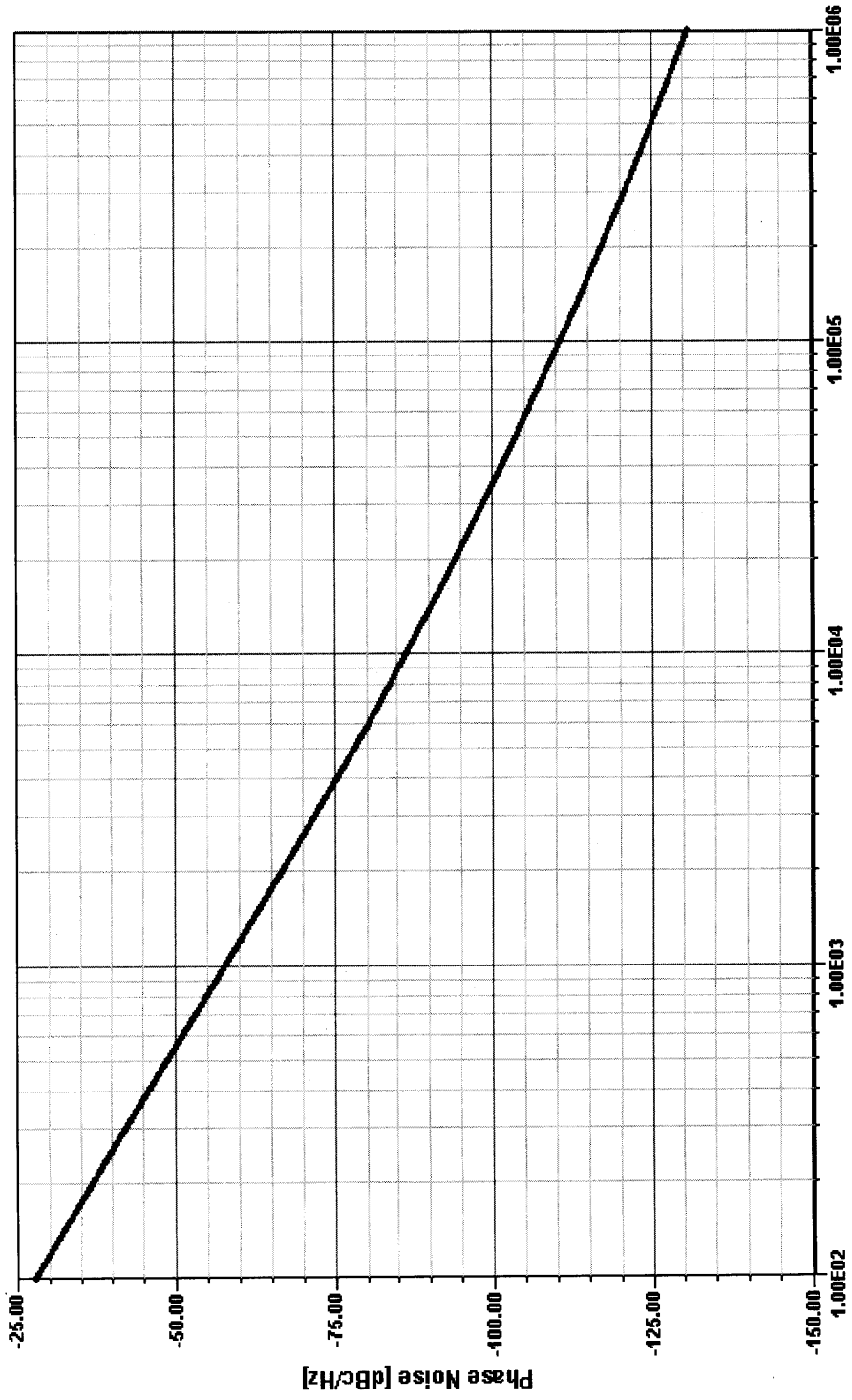


Figure 5-100 Phase noise of the Siemens IC oscillator. The frequency of oscillation is 1.051 GHz.

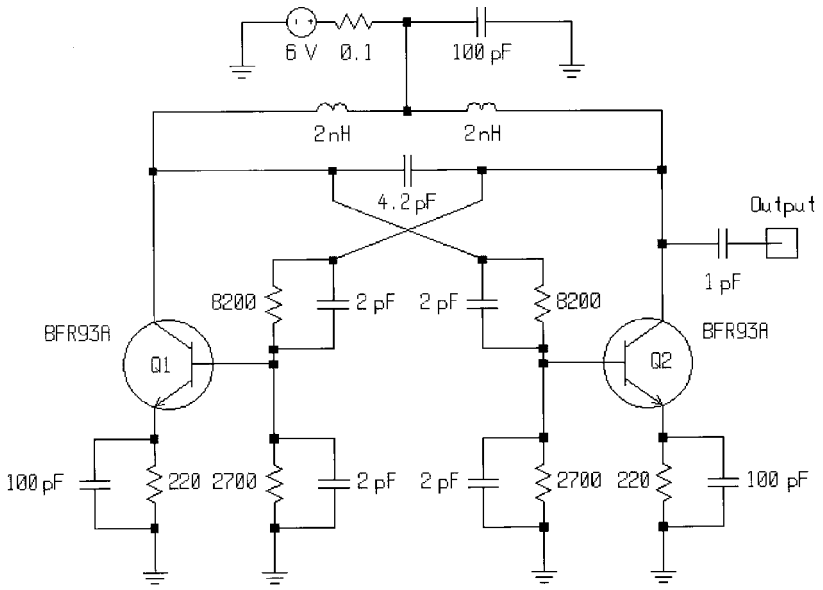


Figure 5-101 Basic push–pull oscillator underlying the circuit shown in Figure 5-99.

simulator to validate and compare oscillator topologies, it's important to start with ideal components because we are interested in comparing the inherent phase-noise performance of various circuits—particularly the role played by the nonlinear reactances in active devices. Because the Q values of practical components are so low in the microwave range (very generally, 200–300 for capacitors and 60–70 for inductors), the inherent merits of one topology, or one device, over another would be masked by loading effects if we conducted our investigations using nonideal reactances. Once initial evaluation is completed, realistic Q values should be specified for increased accuracy in phase-noise, output-power, and output-spectrum simulations.

Finally, as evidence of how moving from lumped to distributed techniques can improve oscillator performance at frequencies where *LC* tanks become problematic, Figure 5-109 compares, for a simulated BJT Colpitts oscillator operating at 2.3 GHz, the difference in phase-noise performance obtainable with a resonator consisting of an ideal 2-nH inductor and a $\frac{1}{4}\lambda$ transmission line (11 Ω, 90° long at 2.6 GHz, attenuation 0.1 dB/m) with the transistor biased by constant-current and constant-voltage sources.

Silicon/GaAs-Based Integrated VCOs and Possible Difficulties. Due to process variations, the center frequency of the VCO cannot be repeated without trimming. Trimming can be avoided by making the VCO gain large enough to compensate for tank circuit variations, but this degrades phase-noise performance and increases the VCO susceptibility to unwanted spurs. For example, a VCO with a gain of 10 MHz/V has a phase noise of –80 dBc/Hz at 10 kHz, while a VCO with a gain of 75 MHz/V has a phase noise of only –60 dBc/Hz at 10 kHz. All but the tank inductors are integrated.

Tank inductors are generally not integrated for reasons of size and Q. The area of a 1-nH inductor is approximately 42,000 mm^2, which is ~3.5% of a typical IC; the area of a 10-nH

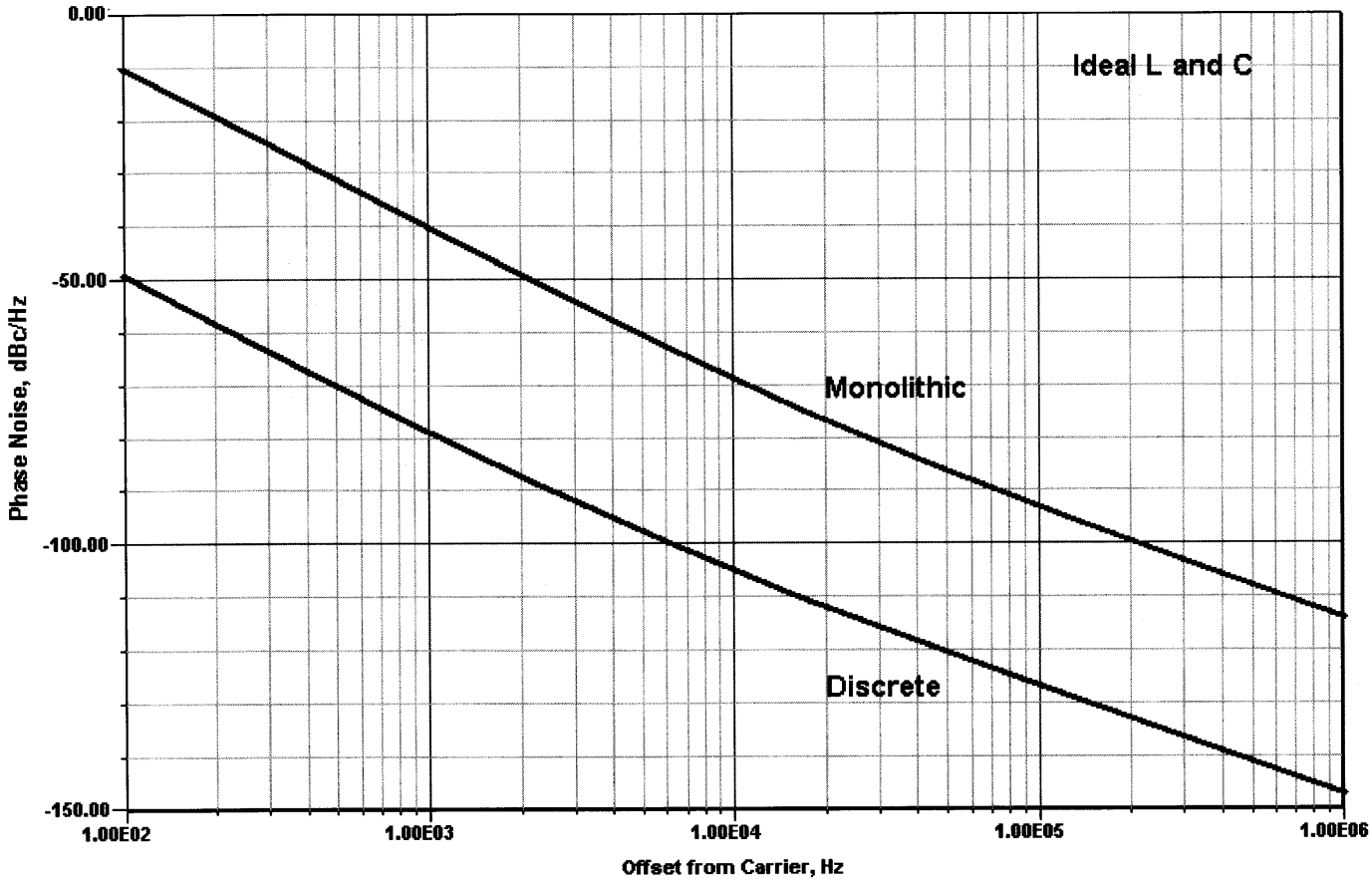


Figure 5-102 The basic version of the oscillator (Discrete curve) is a better phase-noise performer than the integrated version (Monolithic curve). The frequency of oscillation is 1.039 GHz.

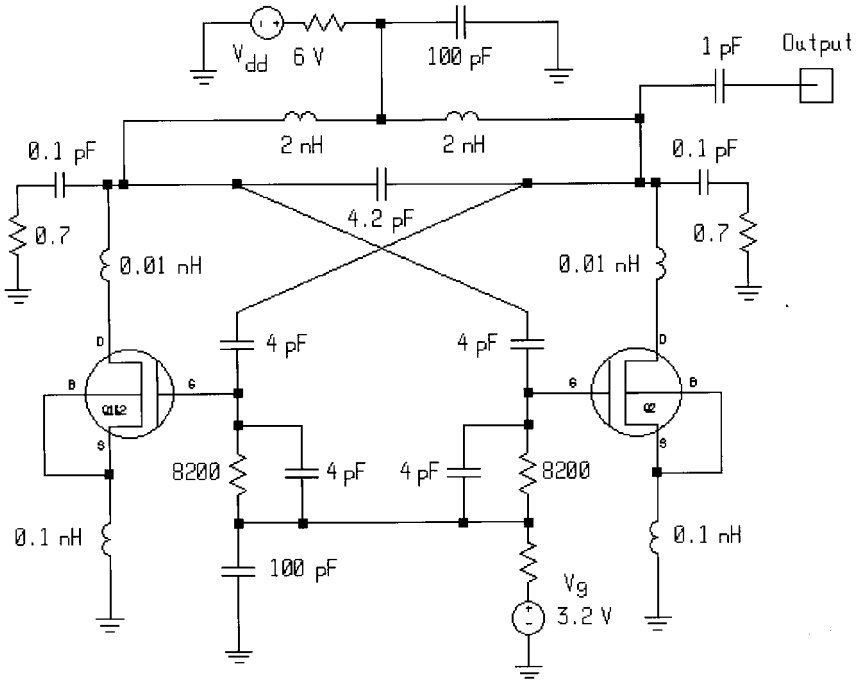


Figure 5-103 A 1.04-GHz push–pull oscillator using LDMOS FETs instead of BJTs. (LDMOS FET parameters courtesy of David K. Lovelace, Motorola, Inc.)

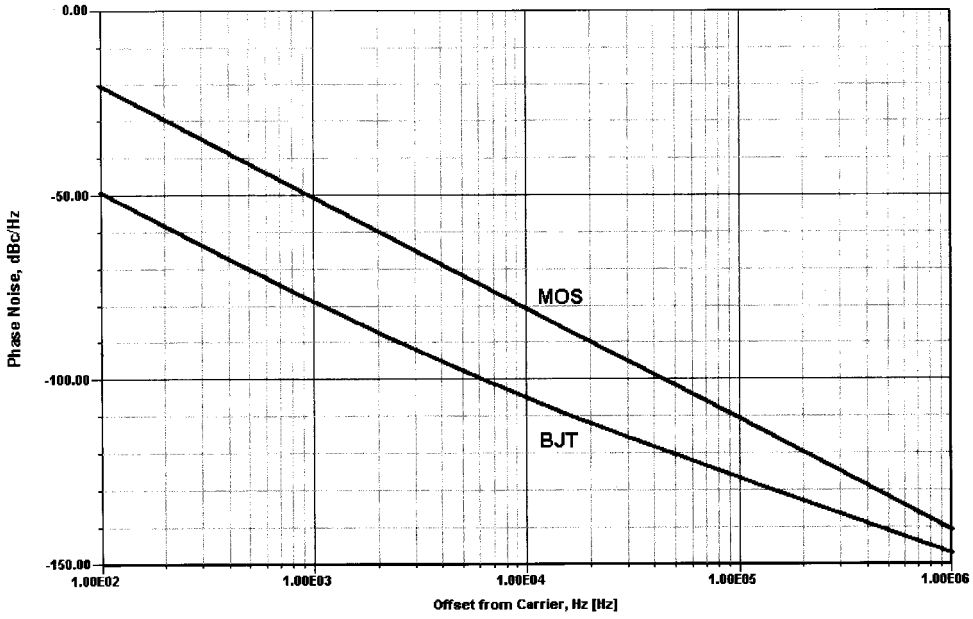


Figure 5-104 Phase-noise comparison of the BJT and MOSFET oscillators.

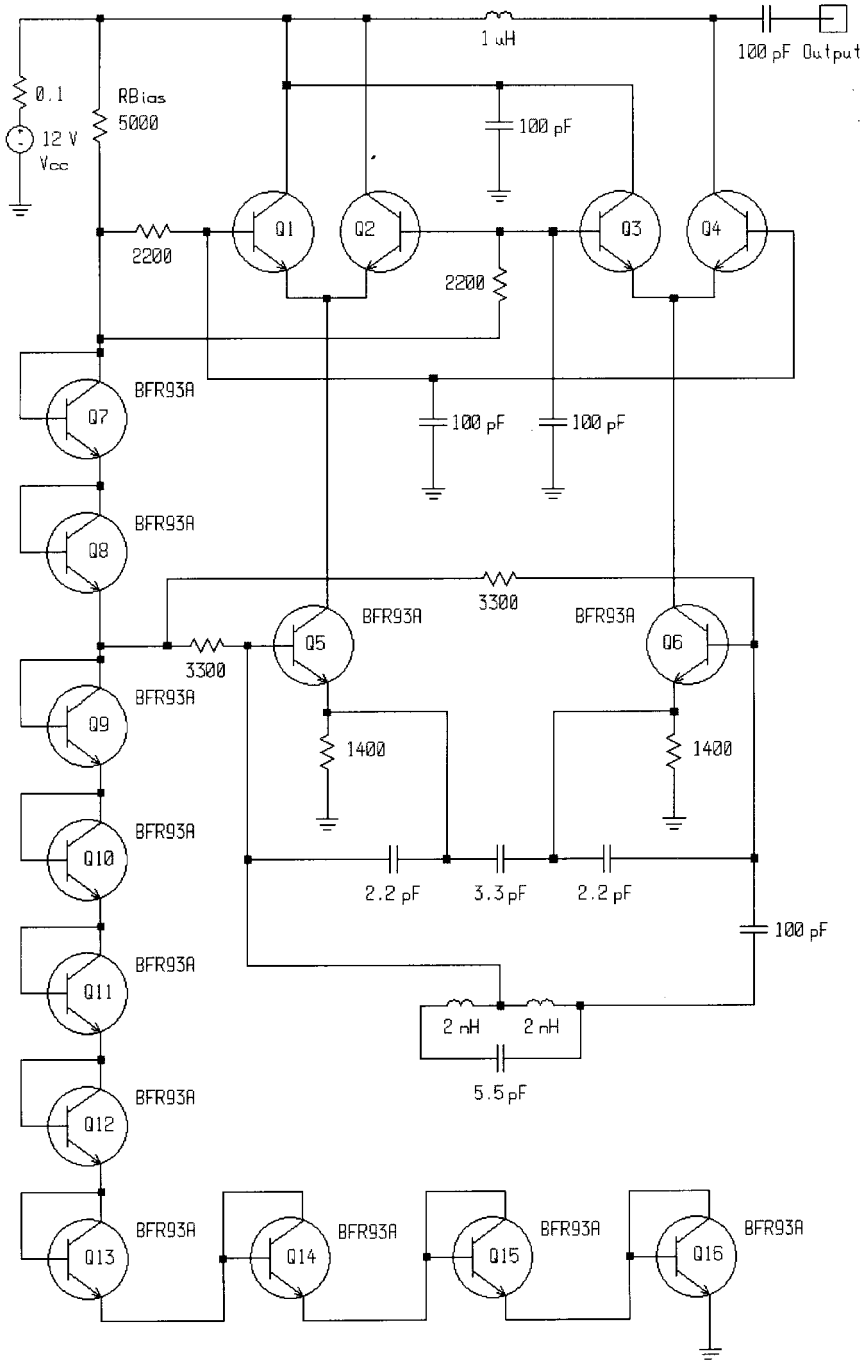


Figure 5-105 A 1.022-GHz IC oscillator based on two differential amplifiers. With an RF signal fed to the bases of Q_1–Q_4 and Q_2–Q_3 in push–pull, and IF output extracted from the collectors of Q_1–Q_3 and Q_2–Q_4 in push–pull, the circuit would act as a Gilbert cell converter. In the circuit shown, we have modified the differential amplifiers' collector circuitry to give single-ended output and avoid cancellation of the oscillator signal.

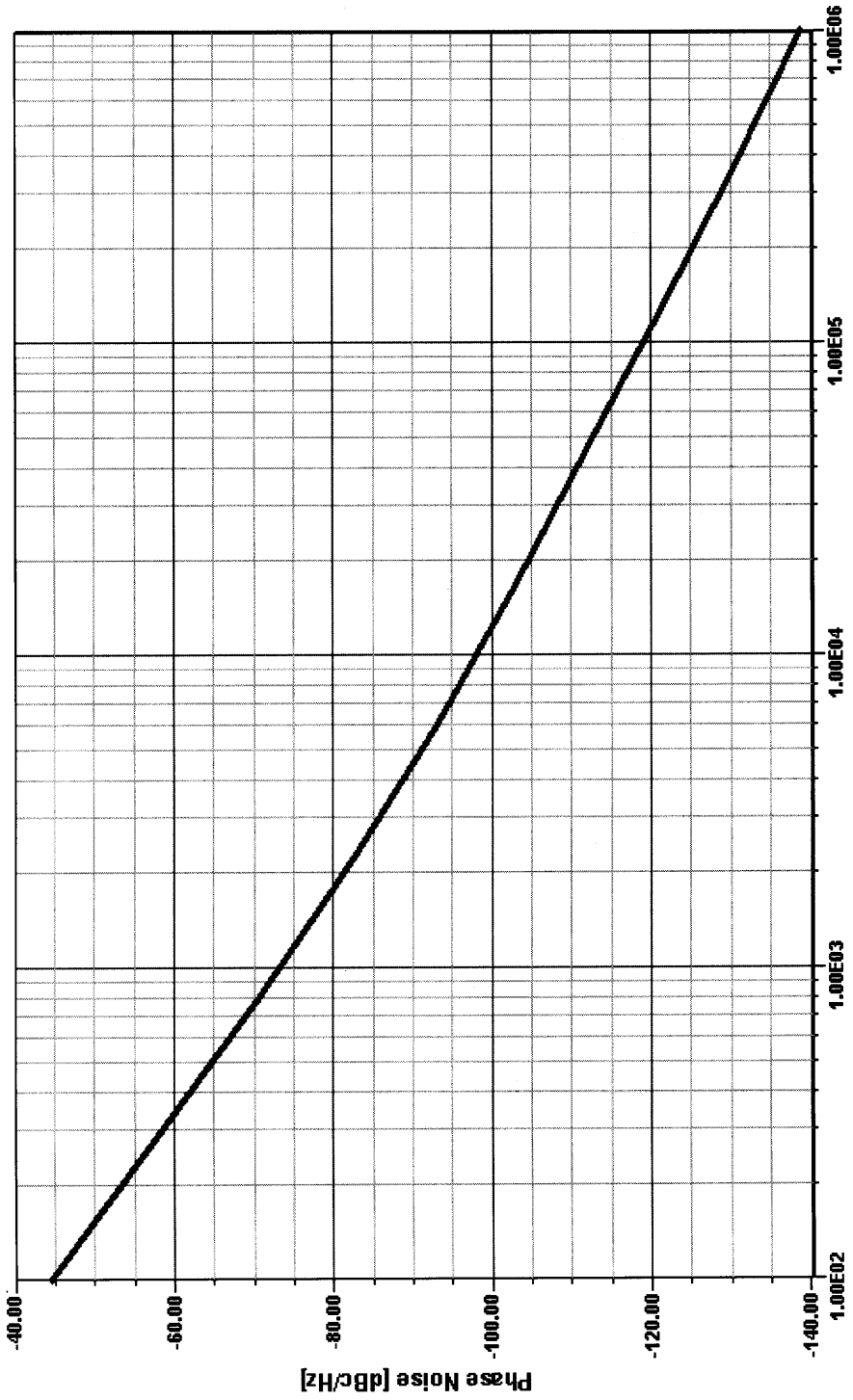


Figure 5-106 Phase noise of the two-differential-amplifier oscillator circuit.

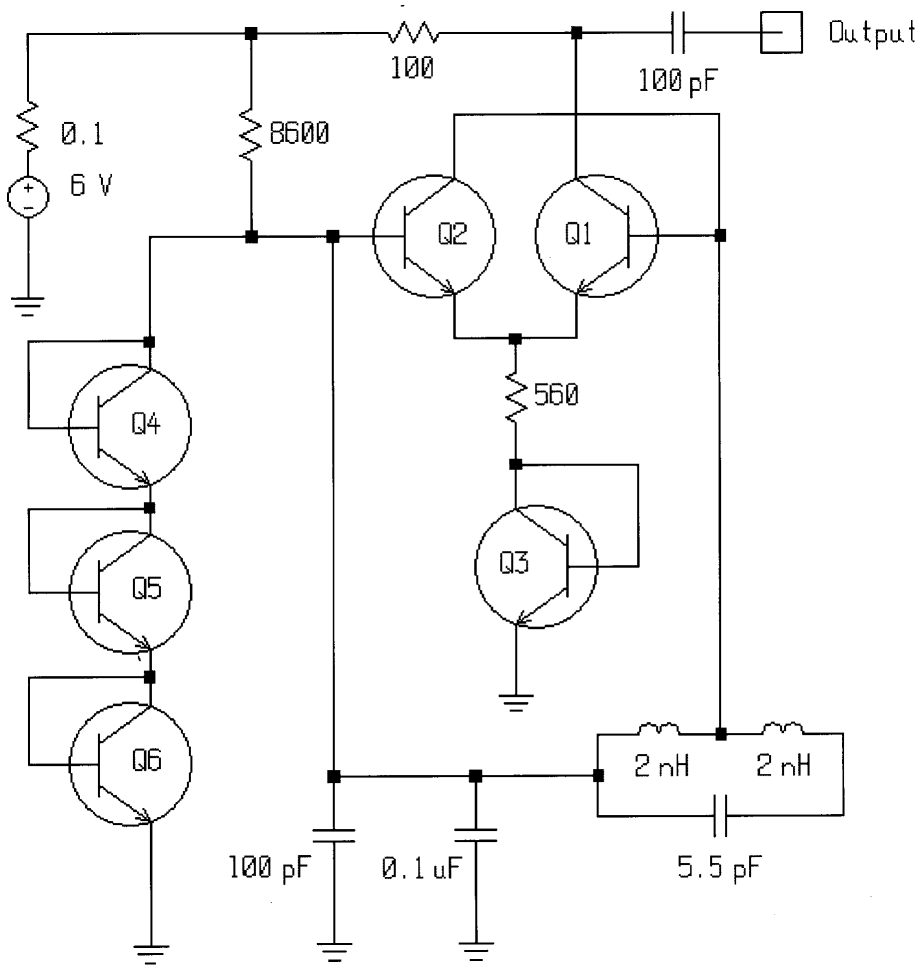


Figure 5-107 This validation circuit models the oscillator topology at the core of Motorola's MC12148 ECL oscillator IC.

inductor is approximately 167,000 mm^2, which is ~14% of a typical IC. For GaAs substrate, the resulting Q would be about 5; for silicon, as we already saw, 20 is already a very high value.

The customer can use PCB traces for the required inductance instead of external surface-mount inductors (cost: $0.03 each, most expensive passive component!). It must be pointed out that even for a nonintegrated approach, the higher VCO gain is noisier by definition. This was explained in Section 5-6-2. These problems are not unique to silicon; they apply to GaAs also. Figure 5-110 shows the size relationship of on-chip capacitors and inductors compared to the actual transistor; therefore, on-chip inductors should really be avoided for reasons of space and Q. Just imagine how the picture would change if the 1.1-nH inductor had to be 11 nH instead! It would dominate the IC area, and therefore the IC cost as well.

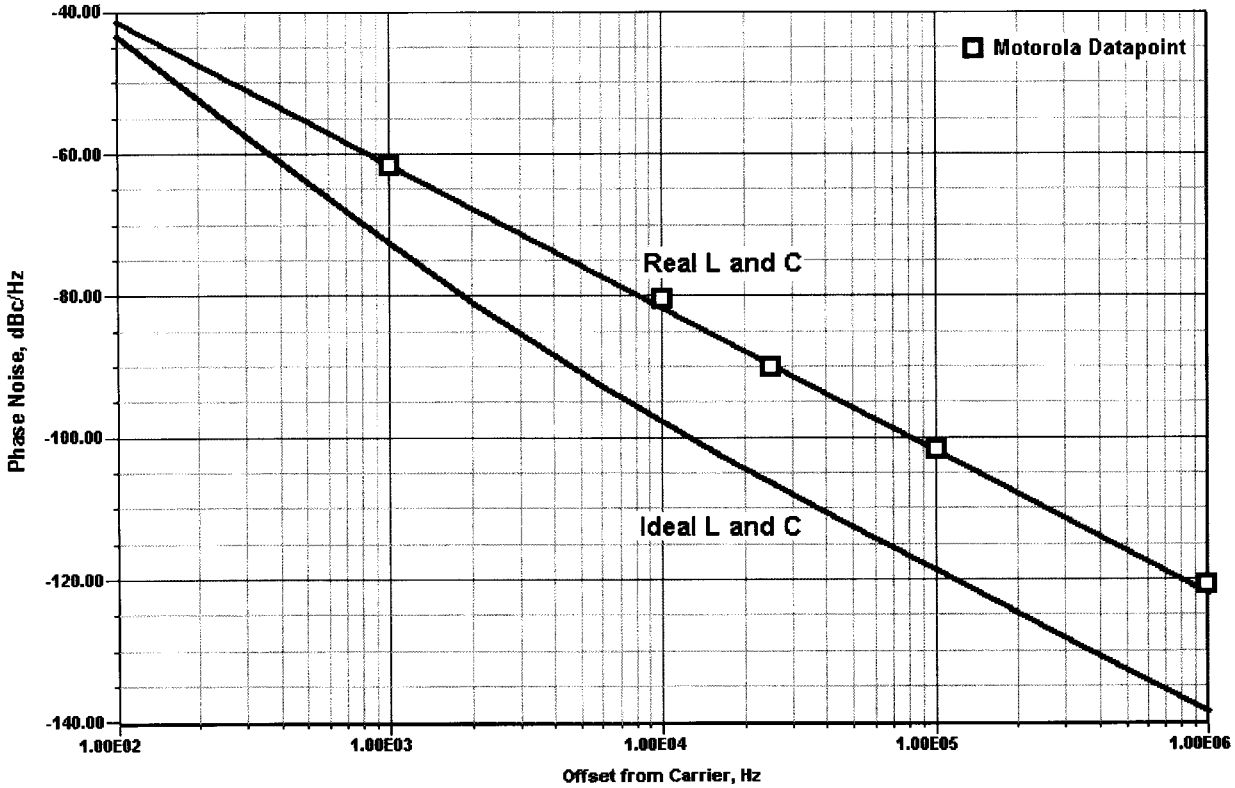


Figure 5-108 Comparison of measured versus simulated phase noise of the MC12148 differential oscillator. The simulation (1.016 GHz) predicts a phase noise of –89.54 dBc/Hz at an offset of 10 kHz (Real L and C curve), which agrees well with Motorola's specification of –89.54 dBc/Hz (typical at 930 MHz) at this offset. The Ideal L and C curve plots the simulation's phase-noise performance with ideal reactances, as opposed to the Real L and C curve, which graphs the results with realistic Q values specified for its reactive components ($Q = 250$ for its C values and $Q = 65$ for its L, all at 1 GHz).

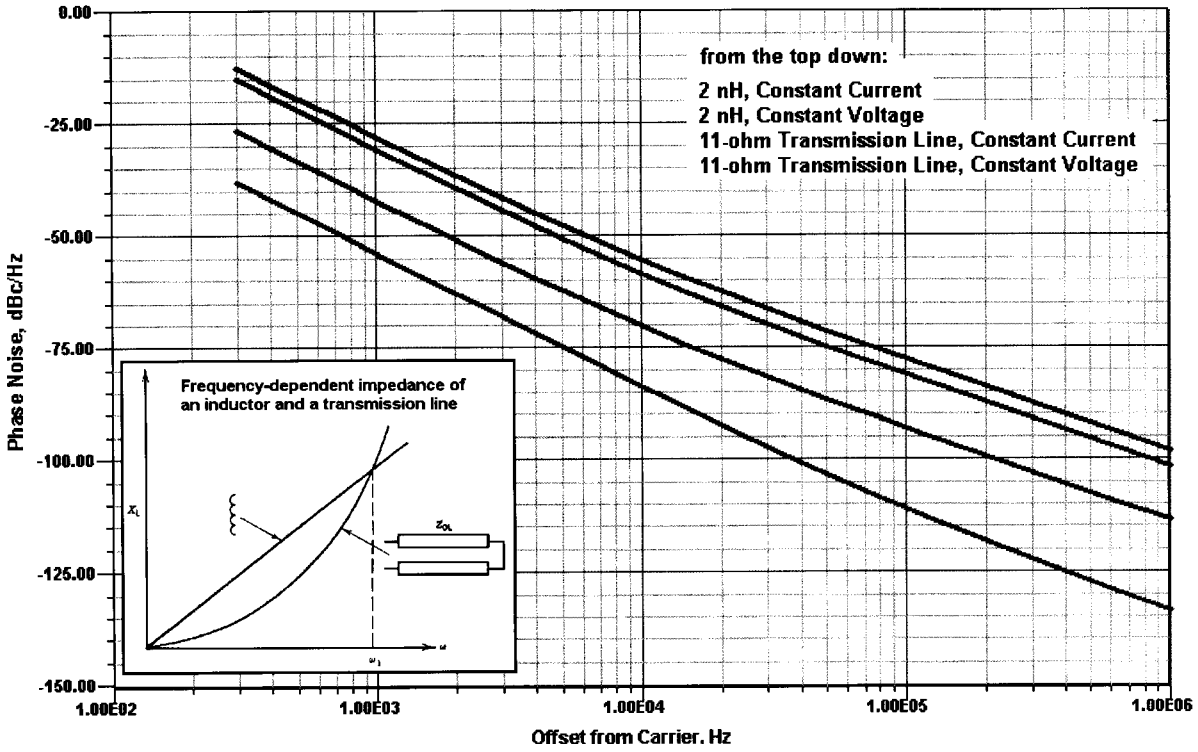


Figure 5-109 Phase-noise performance of a 2.3-GHz BJT oscillator with a resonator consisting of an inductor (2 nH) and a 1/4 λ transmission line (11 Ω, approximating the behavior of a dielectric resonator) with bias from a constant-current source and a low-impedance, resistive constant-voltage source.

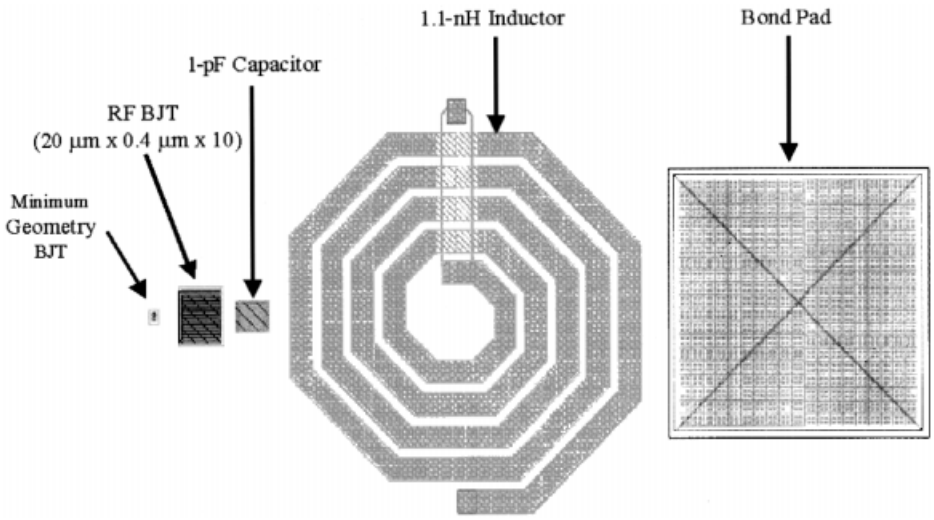


Figure 5-110 Same-scale comparison of space requirements for transistors, capacitors, inductors, and bond pads. For reference purposes, the coil diameter is 0.2 mm.

5-8 DESIGN OF RF OSCILLATORS USING CAD

5-8-1 Harmonic-Balance Simulation

HB analysis performs steady-state analysis of periodically excited circuits. The circuit to be analyzed is split into linear and nonlinear subcircuits. The linear subcircuit is calculated in the frequency domain. Features of this aspect of the HB process include:

- Use of distributed models in the spectral domain
- Matrix formulation that can enable reduction of internal nodes
- Major speed advantage
- Straightforward intermodulation and mixer analysis

The nonlinear subcircuit is calculated in the time domain. Features of this aspect of the HB process include:

- Nonlinear models derived directly from device physics
- Intuitive, easy, and logical circuit representation

Figure 5-111 diagrams this approach for a MESFET amplifier. Figure 5-112 charts a general-purpose nonlinear design algorithm that includes optimization. Modern analysis tools that must provide accurate phase-noise calculation should be based on the principle of harmonic balance. Figure 5-113 shows a BJT microwave oscillator entered into the schematic-capture module of a commercially available HB simulator (Ansoft Serenade 8.0); Figure 5-114 shows this oscillator's simulated phase noise. By the way, HB analysis can also handle mixers.

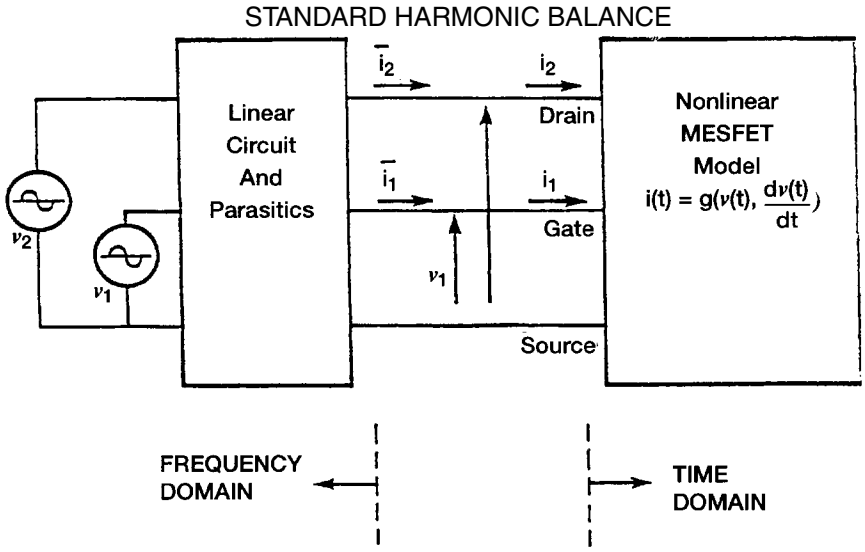


Figure 5-111 MESFET circuit partitioned into linear and nonlinear subcircuits for harmonic-balance analysis. Applied gate and drain voltages, and relevant terminal voltages and currents, are indicated.

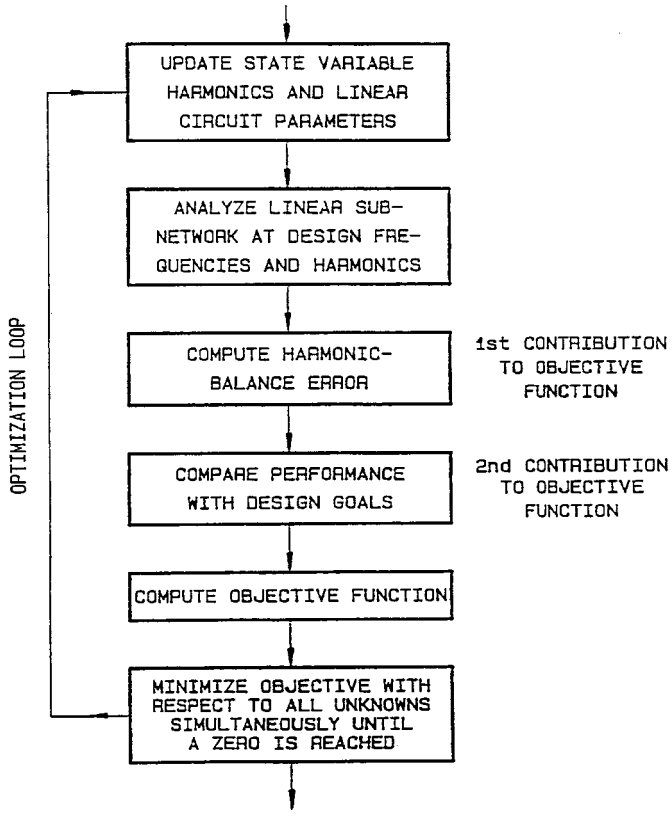


Figure 5-112 Flowchart of a general-purpose harmonic-balance design algorithm that includes optimization.

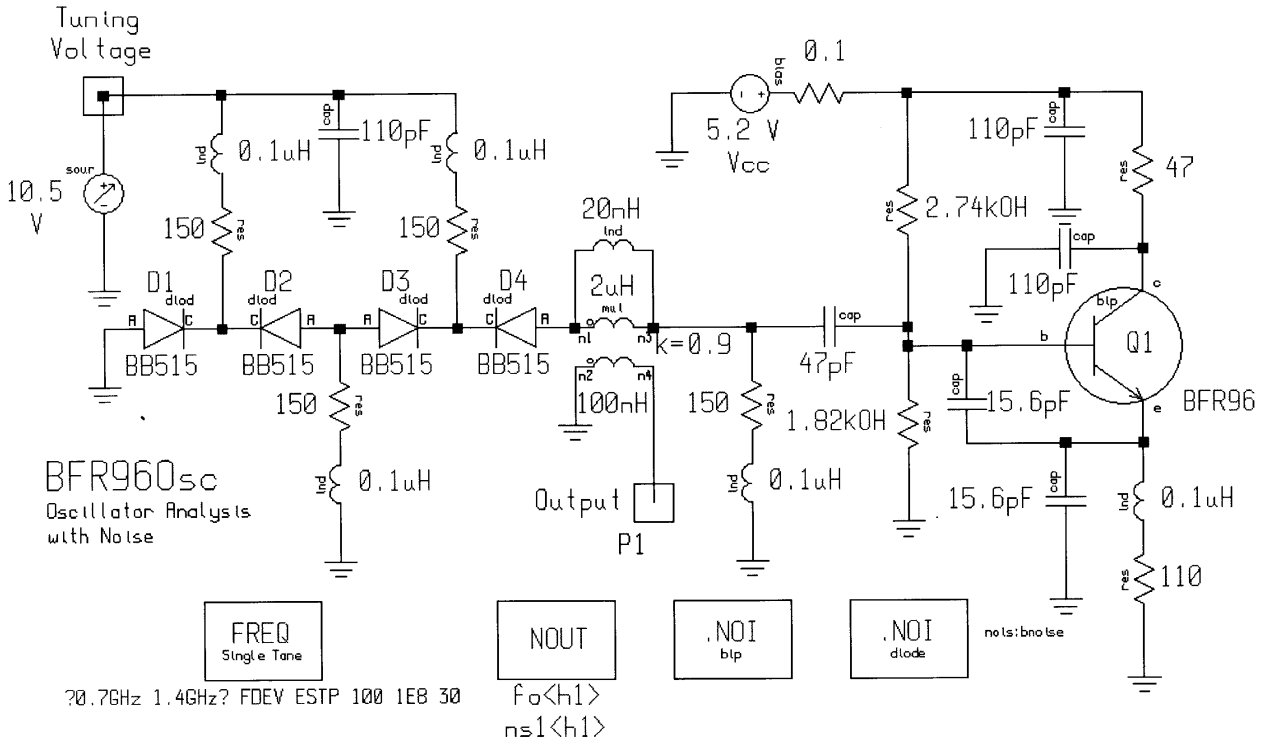


Figure 5-113 BJT microwave oscillator entered into the schematic-capture module of a commercial HB analysis program.

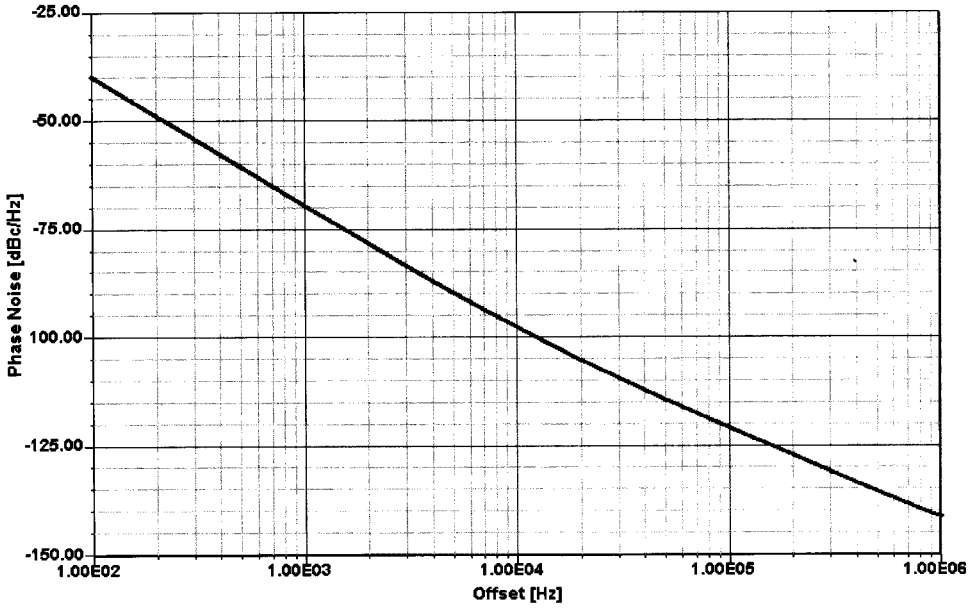


Figure 5-114 Calculated phase noise of the microwave BJT oscillator shown in Figure 5-113.

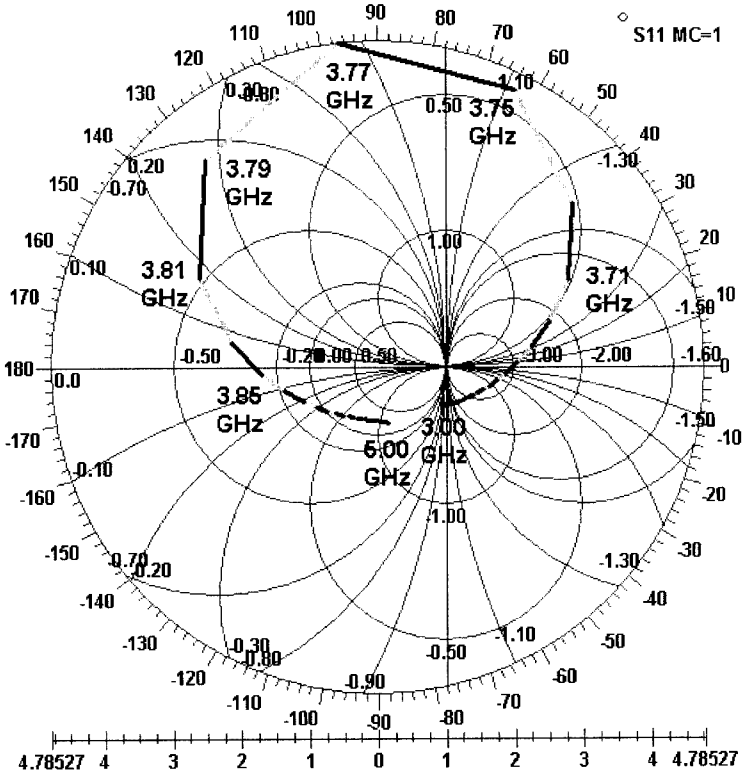


Figure 5-115 S_{11} response of a 3.8-GHz BJT DRO as modeled by Ansoft's Super-Spice time-domain simulator.

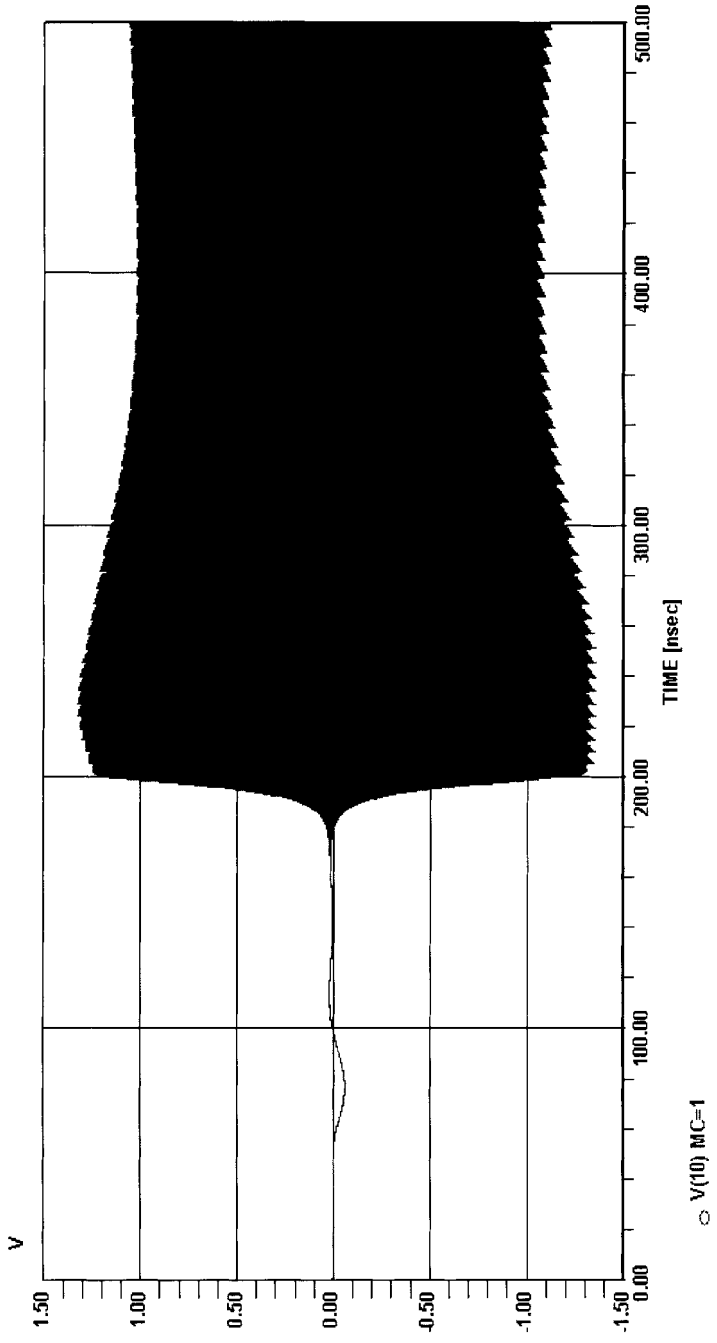


Figure 5-116 Output voltage of the BJT DRO as simulated by Super-Spice.

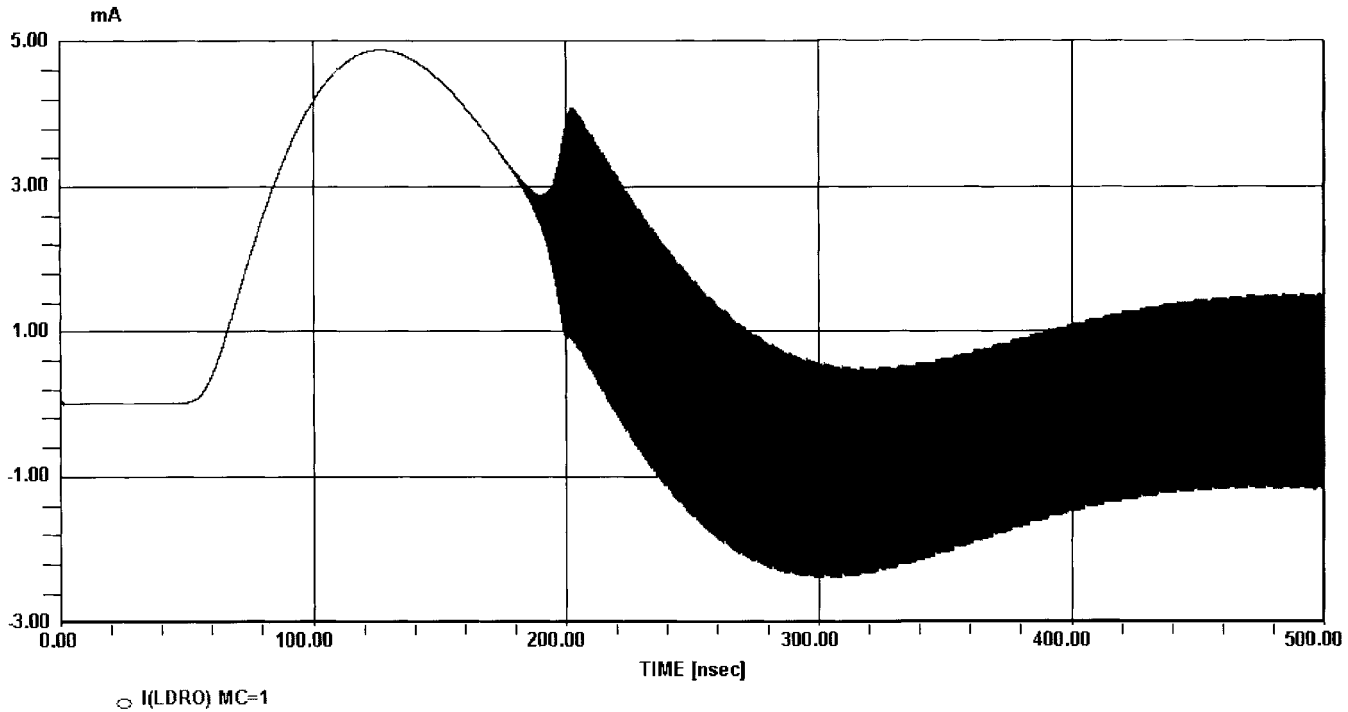


Figure 5-117 Current through the BJT DRO's equivalent resonator inductance from startup to amplitude stabilization.

5-8-2 Time-Domain Simulation

SPICE, a general-purpose simulation program capable of nonlinear dc, linear ac, and nonlinear transient analysis, is probably the most common of the time-domain methods used for nonlinear circuit analysis. Its ability to represent arbitrary time-domain waveforms and model time delays, level shifts, and the sinking and sourcing characteristics of such circuits have made it the simulator of choice for digital designers concerned with analog issues. SPICE's application in RF, microwave, and wireless circuit design is limited by its difficulties with simulating multitone excitations, and coupling and transmission-line effects—including losses, parasitics, and dispersion for distributed elements—and its inefficiency in simulating steady-state conditions.

One area in which SPICE's transient analysis capabilities can be particularly useful to RF and wireless designers is the simulation of oscillator startup. To illustrate this, we present a 3.8-GHz dielectric resonator oscillator modeled with Ansoft Corporation's Super-Spice simulator. Figure 5-115 shows how the oscillator's S_{11} response, positive throughout the simulation's frequency span of 3–5 GHz, peaks at 3.77 GHz. Figure 5-116 shows the oscillator's output voltage as the circuit starts oscillating and amplitude-stabilizes; Figure 5-117 shows the current through the equivalent inductance of the circuit's dielectric resonator across the same time span.

5-9 PHASE-NOISE IMPROVEMENTS OF INTEGRATED RF AND MILLIMETER-WAVE OSCILLATORS *

5-9-1 Introduction

The generation of microwave and millimeter-wave frequencies can be done by lower-frequency VCOs multiplied up in frequency, such as that provided by comb generators; by using hybrid GaAsFET-based oscillators with external resonators; or—for the best phase noise obtainable—YIG oscillators whose physical size and cost do not always provide a practical solution. A significant improvement in oscillator noise performance can be obtained with a novel feedback circuit that uses internally generated noise and cancels the close-in noise in a bandwidth of up to 1 MHz. This results in a phase noise improvement of more than 15 dB at microwave frequencies for the same topology, and the feedback system can be made part of the biasing circuit. This feedback circuit can be used over a wide frequency range and works well at VHF/UHF frequencies and in the millimeter-wave area.

5-9-2 Review of Noise Analysis

Section 5-6 of this chapter covered noise in oscillators in detail, so we will merely review oscillator noise issues here. The first linear model for a basic oscillator, without considering semiconductor noise, was developed by Leeson in 1966 [9] and has been quoted in numerous applications.

The phase noise of a VCO is completely determined by

*This section is based on a paper given at the Fifth International Workshop on Integrated Nonlinear Microwave and Millimeterwave Circuits, sponsored by IEEE at Gerhard-Mercator University, Duisburg, Germany, October 1–2, 1998.

$$£(f_m) = 10 \log\left\{\left[1 + \frac{f_0^2}{(2f_m Q_{\text{load}})^2}\right]\left(1 + \frac{f_c}{f_m}\right)\frac{FkT}{2P_{s,\text{av}}} + \frac{2kTRK_0^2}{f_m^2}\right\} \tag{5-135}$$

where $£(f_m)$ = ratio of sideband power in a 1-Hz bandwidth at f_m to total power in dB
f_m = frequency offset
f_0 = center frequency
f_c = flicker frequency
Q_{load} = loaded Q of the tuned circuit
F = noise factor
$kT = 4.1 \times 10^{-21}$ at 300 K (room temperature)
$P_{s,\text{av}}$ = average power at oscillator output
R = equivalent noise resistance of tuning diode (typically 200 Ω to 10 kΩ)
K_0 = oscillator voltage gain

When adding an isolating amplifier the noise of an *LC* oscillator is determined by

$$S_\phi(f_m) = [a_R F_0^4 + a_E(F_0/2Q_L)^2]/f_m^3$$
$$+ [(2GFkT/P_0)(F_0/2Q_L)^2]/f_m^2$$
$$+ 2a_R Q_L F_0^3/f_m^2$$
$$+ a_E/f_m + 2GFkT/P_0$$

where G = compressed power gain of the loop amplifier
F = noise factor of the loop amplifier
k = Boltzmann's constant
T = temperature in kelvins
P_0 = carrier power level (in watts) at the output of the loop amplifier
F_0 = carrier frequency in hertz
f_m = carrier offset frequency in hertz
$Q_L (= \pi F_0 \tau_g)$ = loaded Q of the resonator in the feedback loop
a_R, a_E = flicker noise constants for the resonator and loop amplifier, respectively

In 1978, Dieter Scherer from Hewlett-Packard added the $(1 + f_c/f_m)$ term, which addresses the 1/*f* noise or flicker corner frequency [25]. The 1/*f* frequency phenomenon is based on the surface effects inside the semiconductor and varies from 50 Hz for silicon FETs, to 5 kHz and more for silicon microwave bipolar transistors, to 1 MHz and higher for GaAsFETs. As a rule of thumb, one can state that the higher the frequency of operation, the higher the f_T for the active device, the higher the flicker corner frequency will be. Figure 5-118 shows the spectral distribution as a function of offset frequency. This linearization was first discovered by individuals from NIST, formerly the National Bureau of Standards, in Boulder, Colorado. The left corner is dominated by the flicker corner frequency and the oscillator itself shows at the output a combination of modulation and conversion

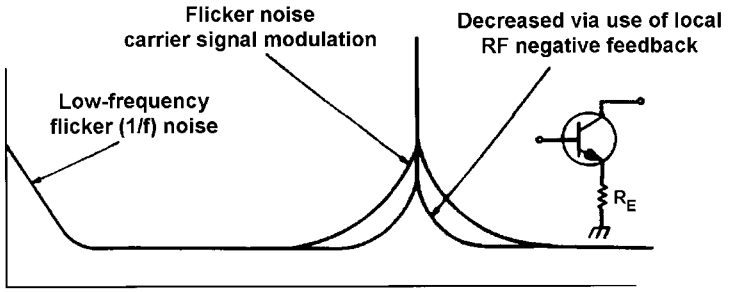


Figure 5-118 Reduction of transistor flicker-of-phase noise via use of local negative feedback (emitter resistance), showing, for a BJT, the flicker noise or corner frequency, as well as the phase noise that results from the flicker noise. The negative feedback from the unbypassed emitter resistor reduces the flicker noise by up to 40 dB [26, 27].

noise. In addition to the $1/f$ noise, the real-life oscillator exhibits a wideband noise, which is due to nonlinearities—specifically, the AM-to-PM conversion resulting from nonlinear capacitance and transconductance of the transistor. This complete linear "noise equation" was first published by Rohde et al. [2]. Therefore, the noise model for the transistor had to be enhanced by describing this mechanism. Essentially, it is equal to a tuning diode coupled to the resonator, which is modulated by thermal noise, and therefore results in a higher phase-noise contribution than the oscillator itself. This noise equation also has Q playing a major role. The phase noise, therefore, ignoring other frequencies, is proportional to $1/Q^2$. In monolithic circuits, the loss of transmission lines, which determines the Q, is related to the substrate material. One has little choice of material, particularly in microwave and millimeter-wave oscillators. Printed resonators sadly provide low Q at these frequencies. A similar problem related to the material is the choices of tuning diodes. In many, if not all, cases, one uses a GaAsFET with the gate connected to the drain or to the source as a tuning diode, and the resulting Q is also disappointing.

The example in Figure 5-118 takes advantage of a microwave bipolar transistor and also is applicable for heterojunction bipolar transistors (HBTs). The transconductance of the bipolar transistor, which is approximately 39 mA/V, increases to values up to 20 times larger than exhibited by GaAsFETs. This method of using negative feedback reduces the gain of the GaAsFET too much and is, therefore, not applicable.

5-9-3 Workarounds

When looking at the modes of operation of test sets for phase-noise measurement, the one method that requires only one signal generator is most desirable. Based on the frequency discriminator method, it uses a cable of appropriate length that acts as a delay line. The oscillator under test drives a double-balanced mixer [one port directly (LO port)] and drives the RF port "delayed" via the delay line cable, and the mixer is therefore driven by the same frequency with a different phase. This setup has been used for phase-noise measurements very successfully. One can easily connect a low-noise FFT-based audio spectrum analyzer to the mixer output and measure the signal-to-noise ratio as a function of offset. With modern FFT analyzers available up to 10 MHz and higher, a fast and reliable measurement can be done. The length of the cable determines the frequency range over which an accurate

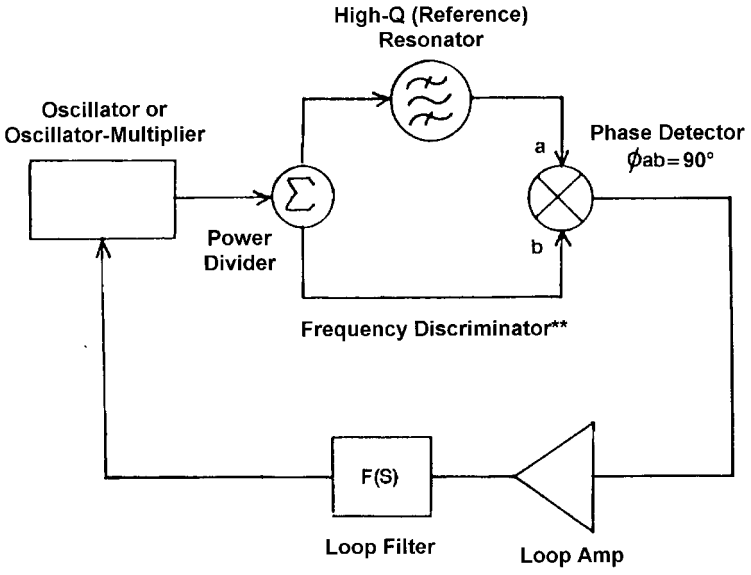
measurement is possible. This phase-noise information (dc and ac voltage) at the output can be used to modulate the VCO and improve the phase noise in the frequency range over which this method is applicable. Fluke and other companies have built signal generators based on this principle to improve the phase noise. The drawback of this method is that the delay line acts like a band-limiting filter, and, therefore, one cannot compensate over a large range of frequencies. The delay line in question, responsible for the phase shift, can also be substituted by a high-Q resonator. By closing the loop, one can "clean up" the phase noise of the device under test (DUT). The difficulty with this is that it requires an external high-Q resonator. One might as well use the same resonator for the oscillator. Therefore, this method is not suitable for MMICs. See Figure 5-119 for more details.

As far as the actual measurement is concerned, one has to observe certain limitations of the test setup as explained in Figure 5-120. When using the delay line instead of a high-Q resonator (low loss is assumed for coaxial cable), one needs to determine the physical length of the cable. The physical length of the cable determines the delay and, therefore, the discriminator noise floor as a function of the offset frequency. Typically, 8 feet of RG-8/U cable results in a delay of 12 ns. As can be seen from Figure 5-121, it is advisable to have several lengths of cable depending on the range of frequency for which one wants to do the measurements under different lengths of cable [28].

Since we are primarily concerned about the flicker noise, here is another, similar, approach to "cleaning up" the flicker noise, as shown in Figure 5-122. The bandwidth over which this technique works depends on the bandwidth of the phase detector system, which can be made fairly wide. A derivative of this will provide a useful solution.

5-9-4 Reduction of Flicker Noise

From inspecting a device datasheet, one is always pleasantly surprised about very low spot noise figures, while the 1/f corner frequency is rarely specified at all. The mechanism that



**Resonator may be operated in reflection mode as well as transmission mode.

Figure 5-119 AFC stabilization of an external oscillator.

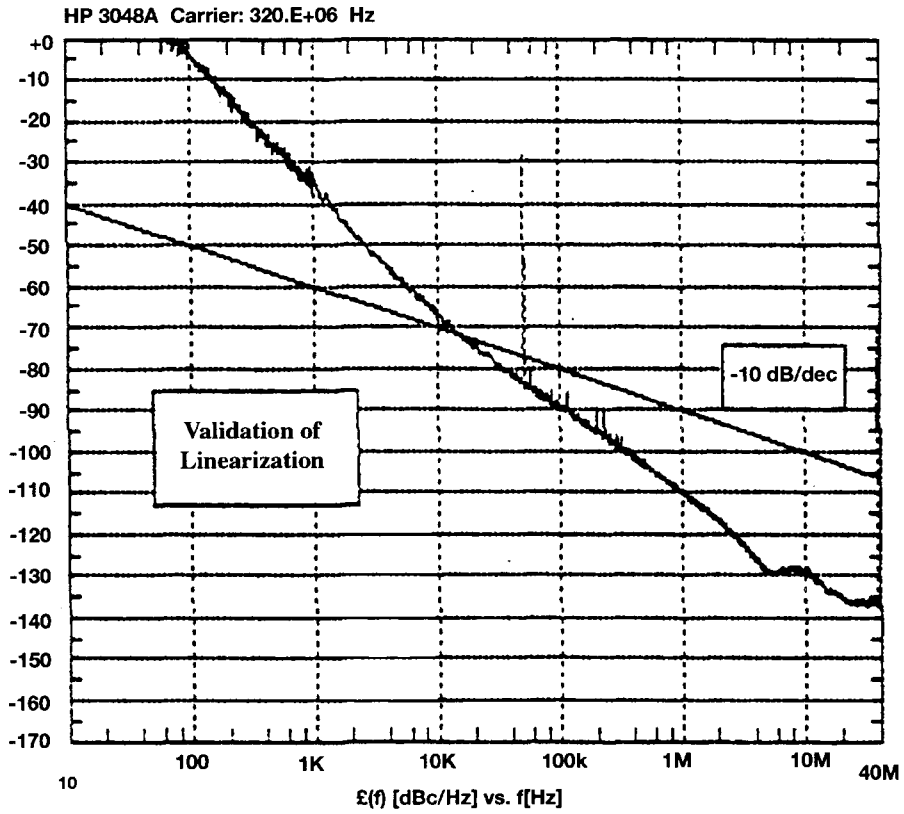


Figure 5-120 Display of a typical phase-noise measurement using the delay line principle. This method is only applicable where $x = \sin(x)$. The measurements above the solid line violate this relationship and therefore are not valid.

transforms the noise, which is generated by the $1/f$ contribution into what we referred to as modulation noise, is the AM-to-PM conversion. If one could "linearize" the transistor, the flicker noise effect would drastically be reduced.

Figure 5-123 shows two variations on an approach that is consistent with the previous method. In Figure 5-123a, the noise across the 56-Ω emitter resistor of oscillator transistor Q1 is sampled, amplified and inverted in Q2, and used to modulate Q1's base voltage. A low-frequency, general-purpose *npn* transistor (BC848C) is used at Q2. A high-frequency, higher-power transistor (BFR93A) is used at Q1; an HBT could also be used. Figure 5-123b shows the same approach using a variation of the *pnp* active biasing circuit shown in Figure 3-192. Figure 5-124 shows the phase-noise performance for the circuit of Figure 5-123a.

5-9-5 Applications to Integrated Oscillators

Figure 5-125 shows a silicon-based BiCMOS technology wireless oscillator. The three big inductors can be seen easily. This design by Motorola is a good candidate for using this innovative technology. Applying this feedback technology, we measure the phase noise shown in Figure 5-126.

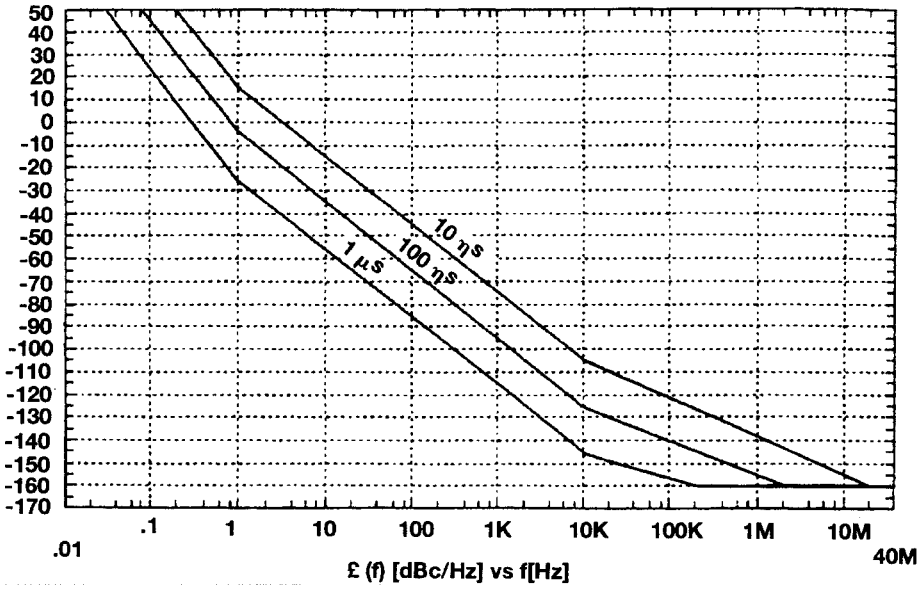


Figure 5-121 Dynamic range as a function of cable delay. A delay of 1 μs is ideal for microwave frequencies.

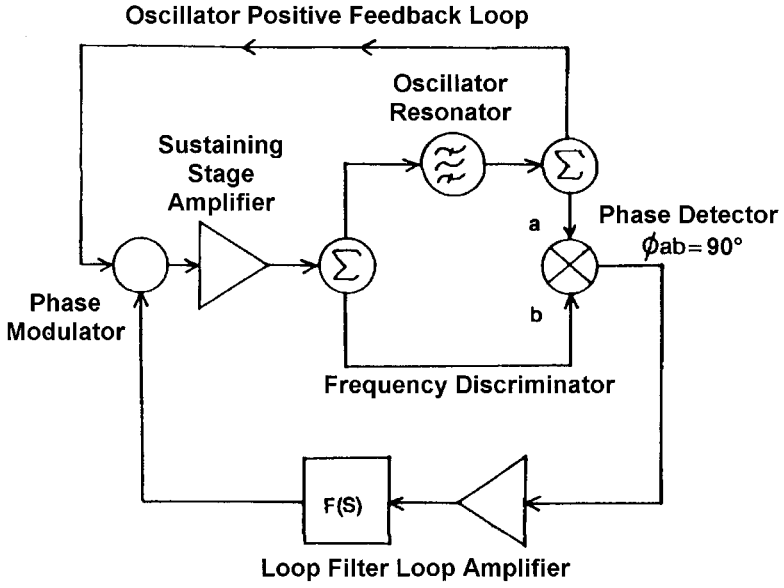
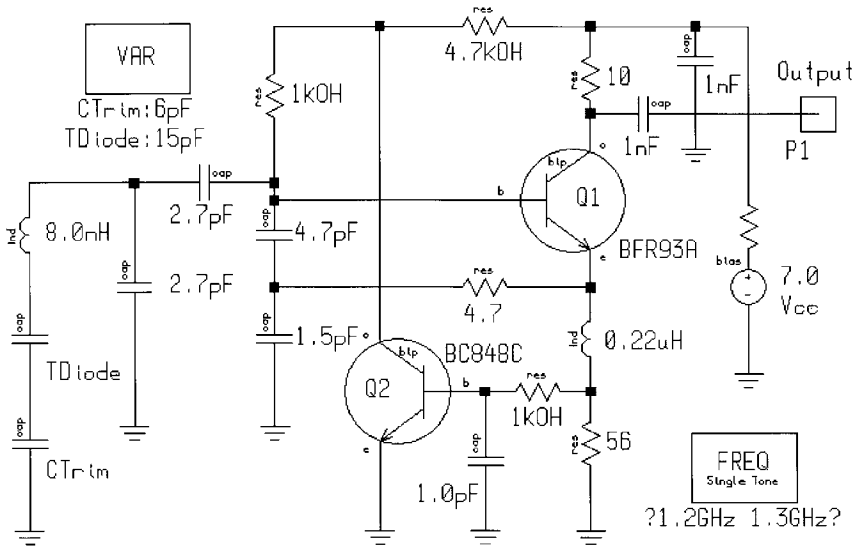
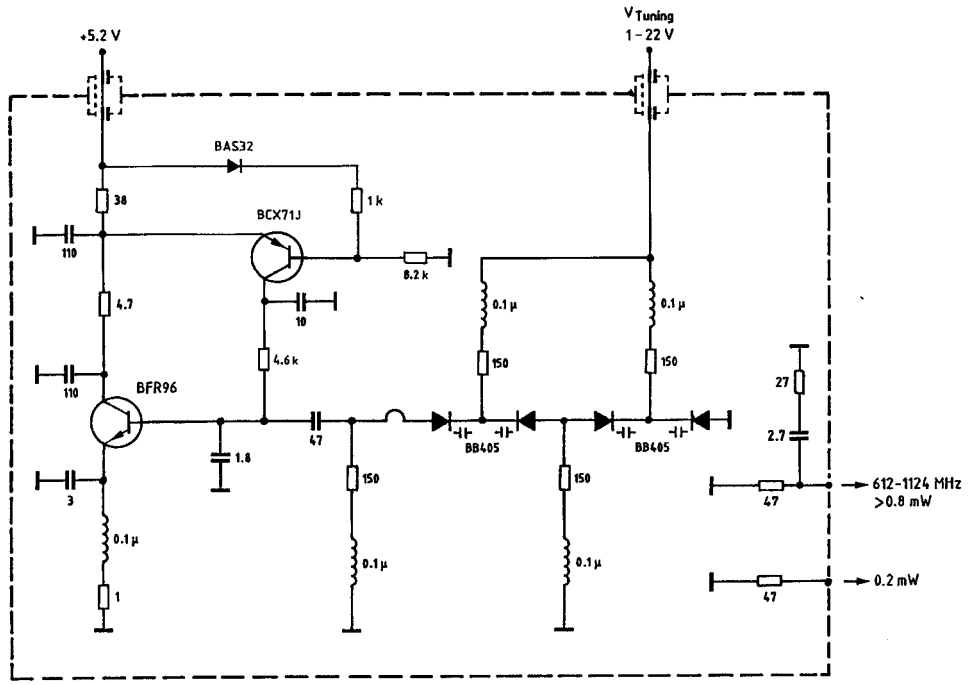


Figure 5-122 Noise reduction in an oscillator whose signal flicker-of-frequency noise is primarily due to sustaining stage flicker-of-phase noise. A phase perturbation in the sustaining stage produces a frequency change in the oscillator, which produces a change in the signal phase shift through the resonator detected as the phase detector. The resonator may be operated simultaneously in transmission line and reflection modes in the oscillator in the oscillator and discriminator portions of the circuit, respectively.



(a)



(b)

Figure 5-123 BJT-based oscillators with noise feedback. (a) The noise sampling is done in the transistor emitter; (b) the noise sampling is done in the collector using a variation of the active biasing circuit shown in Figure 5-124.

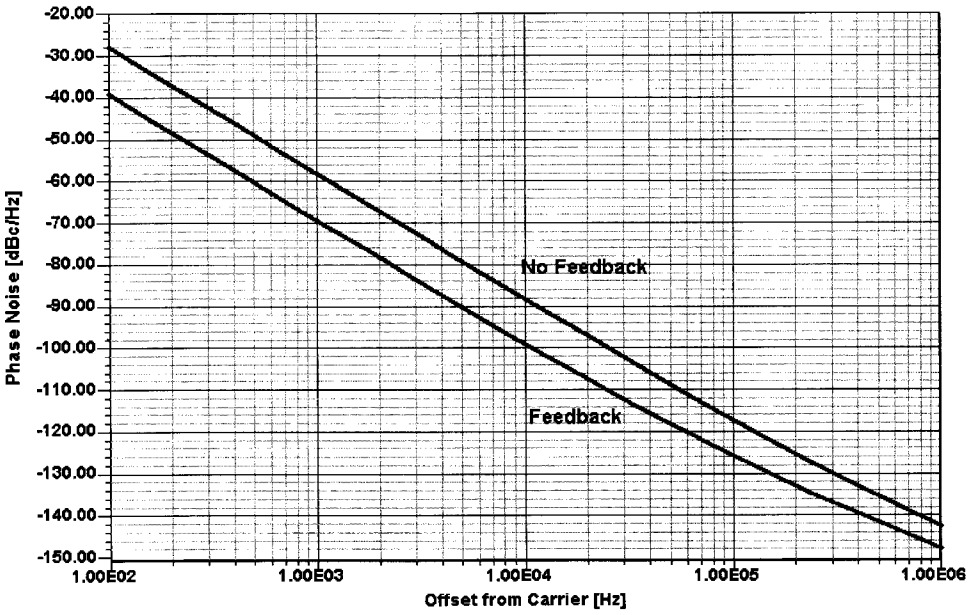


Figure 5-124 Phase noise of the oscillator in Figure 5-123a with and without noise feedback.

Moving up into the millimeter-wave area, one of the target oscillators is the Ka-band MMIC voltage-controlled oscillator designed for Ka-band smart munitions applications (Figure 5-127). This voltage-controlled oscillator MMIC employs 0.25-μm gate length, double-heterojunction, pseudomorphic, high-electron-mobility transistor (PHEMT) technology. This is a custom chip developed by Martin Marietta under the U.S. Government MMIC program. Ansoft Serenade tools were used in this design.

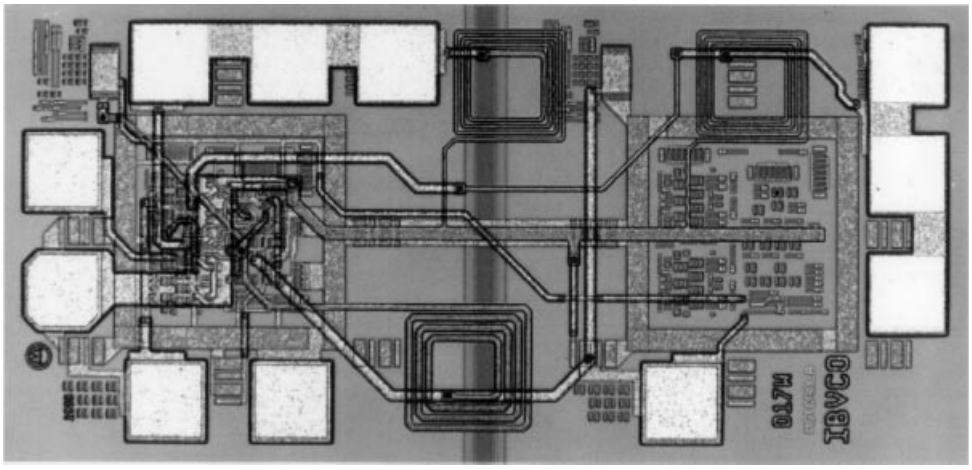


Figure 5-125 Layout of the Motorola 800-MHz monolithic differential oscillator IC in Si technology.

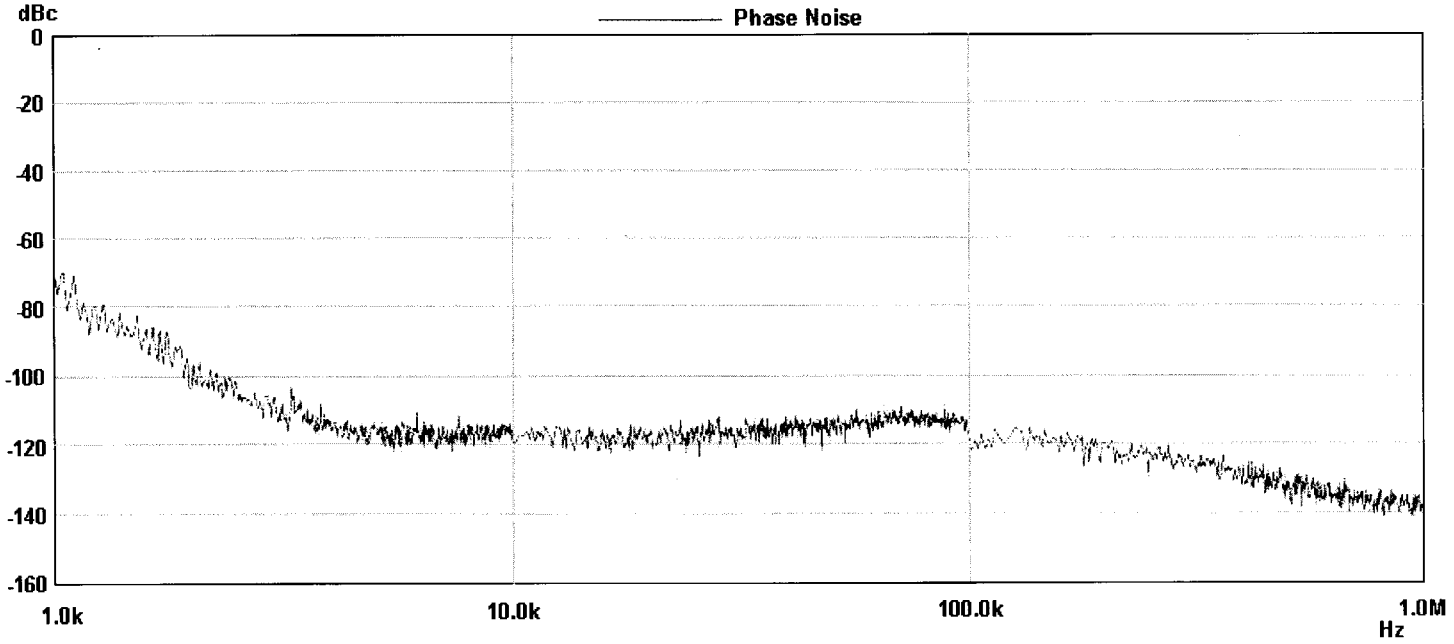


Figure 5-126 Phase noise of the oscillator of Figure 5-125. The noise floor is limited by the on-board isolation amplifier.

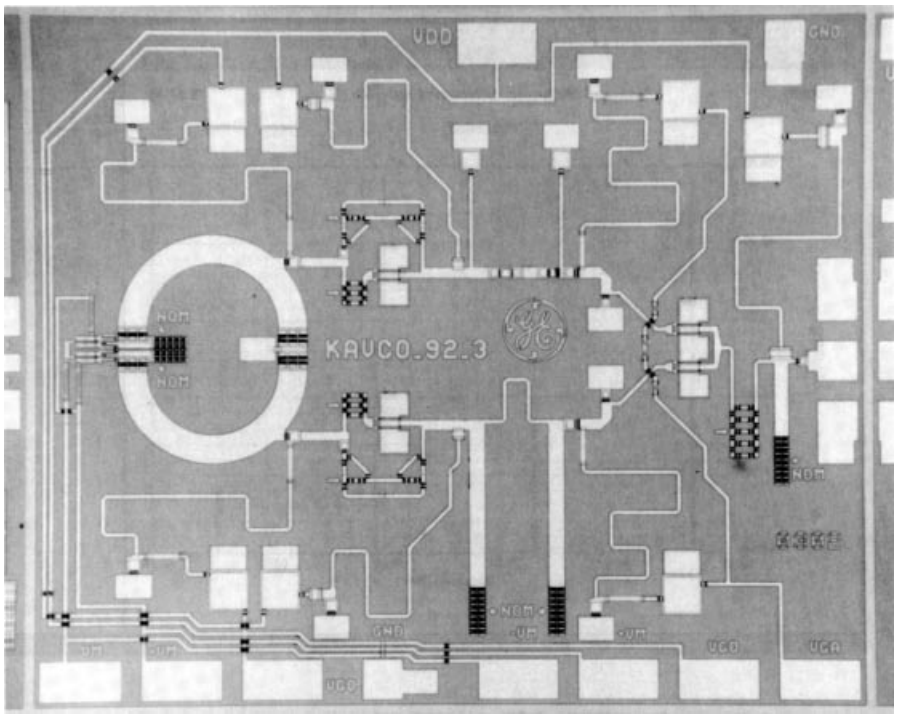


Figure 5-127 Layout of the Ka-band voltage-controlled oscillator.

Features

- Fundamental-mode, differential-ring VCO
- Electronically tunable
- Output power greater than 16 dBm
- Power-added efficiency (PAE) of 13%
- Compact size for easy integration with power amplifier
- 0.25-μm pseudomorphic HEMTs

Specifications

- Frequency range: Ka band
- Output power: 16 dBm minimum

When applying the same technology as shown in the previous bipolar example, the phase noise can be reduced significantly. Since field-effect transistors require a negative gate voltage, it is possible to sample the noise in the drain current using a resistor, amplify the noise with an appropriate loop amplifier, and modulate the gate voltage within a 1-MHz bandwidth.

A spectrum analyzer measurement of this oscillator type, as shown in Figure 5-128, indicates a phase noise of −78.6 dBc/Hz at an offset frequency of 100 kHz at a center frequency of 38 GHz, while the closed-loop phase noise of Figure 5-129 using the compen-

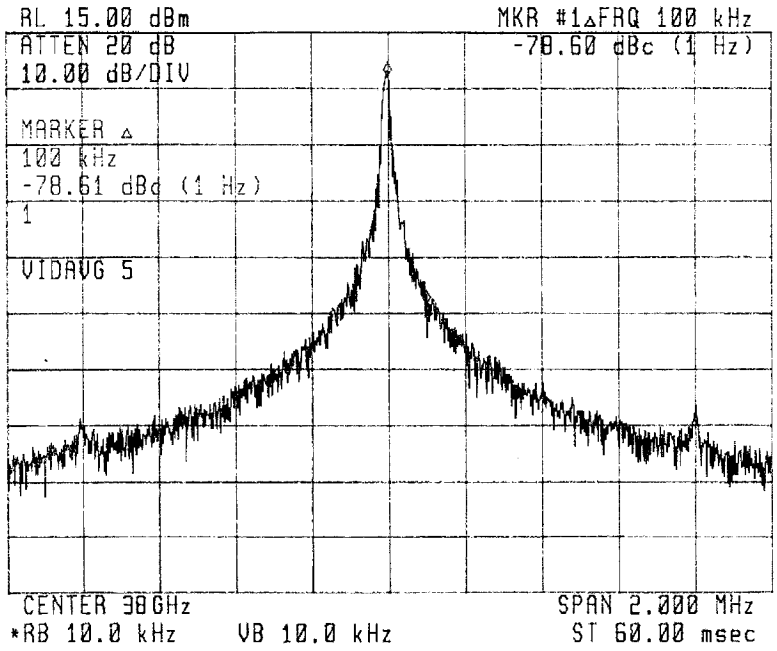


Figure 5-128 Spectrogram of the 38-GHz free-running push–pull VCO shown in Figure 5-127.

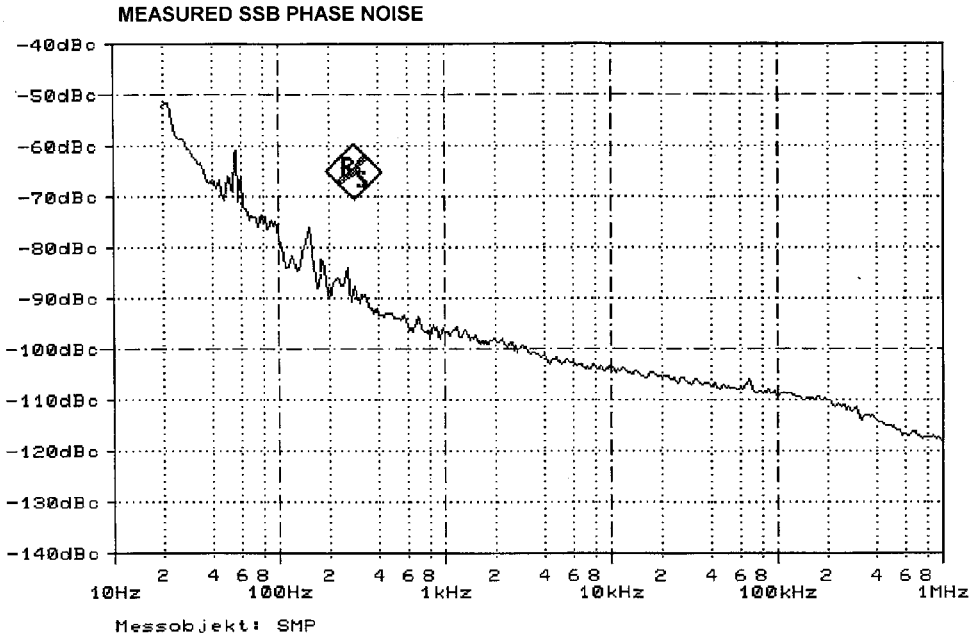


Figure 5-129 Phase-noise performance of a 38-GHz oscillator with phase-noise “clean-up” system.

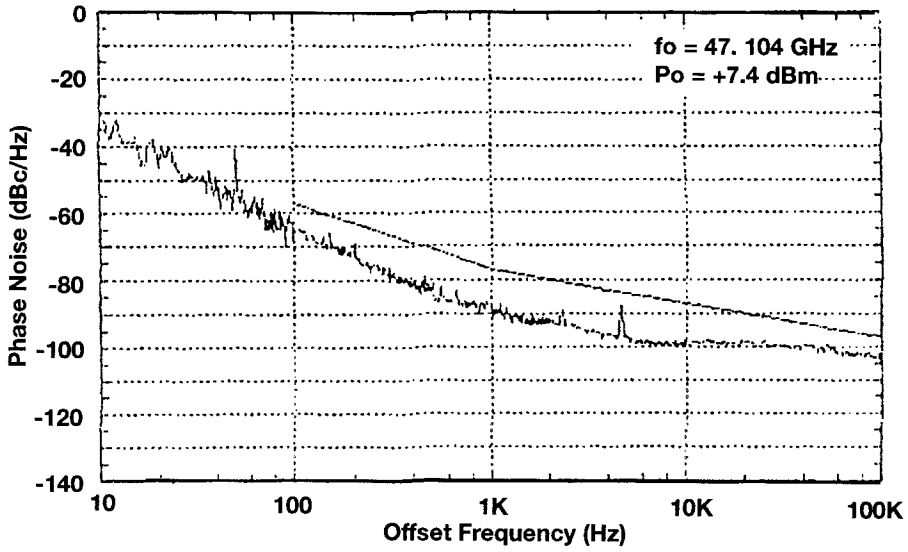


Figure 5-130 A 47-GHz loop-stabilized VCO with phase-noise performance similar to that of Figure 5-129.

sation scheme indicates a phase noise of about −108 dBc/Hz at 100 kHz. This is consistent with reported results from other researchers who have used complicated phase-locked loops to achieve similar performance as seen in Figure 5-130.

5-9-6 Summary

A novel method of improving the phase noise of oscillators has been demonstrated. This approach works well for low-frequency oscillators such as 1000 MHz up to the millimeter-wave range. It only requires an additional wideband dc biasing circuit, which applies noise feedback. The correctness of the theory has been demonstrated by appropriate validation.

REFERENCES

1. Paul Danzer, ed., *The ARRL Handbook for Radio Amateurs,* 75th ed. Newington, CT: American Radio Relay League, 1997, Chap. 14, AC/RF Sources (Oscillators and Synthesizers).
2. Ulrich L. Rohde, Jerry C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*, 2nd ed. New York: McGraw-Hill, 1997, p. 380.
3. G. Vendelin, A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design: Using Linear and Nonlinear Techniques.* New York: Wiley, New York, 1990.
4. Samuel Y. Liao, *Microwave Devices and Circuits,* 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1985.
5. S. C. Peterson, "900-MHz μ-Strip," Philips Semiconductors application document.
6. H. Keller, "Electronic UHF Tuning in TV Receivers," *Radio-Fernsch-Phono-Praxis,* No. 3, 1967.
7. L. Micic, "The Tuner Diode," *Radio Mentor Electron.,* **32**(5), 404–405, 1966.

8. Kyung-Whan Yeom, "An Investigation of the High-Frequency Limit of a Miniaturized Commercial Voltage-Controlled Oscillator Used in 900-MHz Band Mobile-Communication Handset," *IEEE Trans. Microwave Theory Tech.,* **MTT-46**, 1165–1168, August 1998.
9. D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," *Proc. IEEE,* **54**, 329–330, 1966.
10. U. L. Rohde, C. R. Chang, and J. Gerber, "Parameter Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits Like Mixers and Oscillators," *Proceedings of the 23rd European Microwave Conference,* Madrid, Spain, September 6–9, 1993.
11. C. R. Chang, "Mixer Noise Analysis Using the Enhanced Microwave Harmonica," Compact Software *Transmission Line News*, **6**(2), 4–9, June 1992.
12. T. Antognetti and G. Massobrio, *Semiconductor Device Modeling with SPICE.* New York: McGraw-Hill, 1988, p. 91.
13. R. J. Hawkins, "Limitation of Nielsen's and Related Noise Equations Applied to Microwave Bipolar Transistors, and a New Expression for the Frequency and Current Dependent Noise Figure," *Solid State Electron.*, **20**, 191–196, 1977.
14. Tzu-Hwa Hus and Craig P. Snapp, "Low Noise Microwave Bipolar Transistor Sub-Hall-Micrometer Emitter Width," *IEEE Trans. Electron Devices*, **ED-25**, 723–730, June 1978.
15. R. A. Pucel and U. L. Rohde, "An Accurate Expression for the Noise Resistance R_n of a Bipolar Transistor for Use with the Hawkins Noise Model," *IEEE Microwave Guided Wave Lett.*, **3**(2), 35–37, February 1993.
16. See reference 3.
17. R. A. Pucel, W. Struble, Robert Hallgren, and U. L. Rohde, "A General Noise De-embedding Procedure for Packaged Two-Port Linear Active Devices," *IEEE Trans. Microwave Theory Tech.,* **MTT-40**(11), 2013–2024, November 1993.
18. U. L. Rohde, "Improved Noise Modeling of GaAs FETS: Using an Enhanced Equivalent Circuit Technique," *Microwave J.,* Part I: 87–101, November 1991; Part II: 87–95, December 1991.
19. V. Rizzoli, F. Mastri, and C. Cecchefti, "Computer-Aided Noise Analysis of MESFET and HEMT Mixers," *IEEE Trans. Microwave Theory Tech.,* **MTT-37**, 1401–1410, September 1989.
20. V. Rizzoli and A. Lippadni, "Computer-Aided Noise Analysis of Linear Multiport Networks of Arbitrary Topology," *IEEE Trans. Microwave Theory Tech.,* **MTT-33**, 1507–1512, December 1985.
21. V. Rizzoli, F. Mastri, and D. Masotti, "General-Purpose Noise Analysis of Forced Nonlinear Microwave Circuits," *Military Microwave,* 1992.
22. U. L. Rohde, "Key Components of Modern Receiver Design," *VHF Conference*, Weinheim, Germany, September 17–18, 1993.
23. C. R. Chang and U. L. Rohde, "The Accurate Simulation of Oscillator and PLL Phase Noise in RF Sources," *Wireless '94 Symposium,* Santa Clara, California, February 15–18, 1994.
24. U. L. Rohde, "Oscillator Design for Lowest Phase Noise," *Microwave Eng. Eur.*, 31–40, May 1994.
25. Dieter Scherer, "Design Principles and Test Methods for Low Phase Noise RF and Microwave Sources," *RF & Microwave Measurement Symposium and Exhibition*, Hewlett-Packard.
26. D. J. Healy III, "Flicker of Frequency and Phase and White Frequency and Phase Fluctuations in Frequency Sources," *Proceedings of the 26th Annual Symposium on Frequency Control*, Fort Monmouth, NJ, June 1972, pp. 43–49.
27. Michael M. Driscoll, "Low Noise Oscillator Design Using Acoustic and Other High *Q* Resonators," *44th Annual Symposium on Frequency Control*, Baltimore, MD, May 1990.
28. Ulrich L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design*. New York: Wiley, 1997.

INTERESTING PATENTS

Michel Fache and Claude Juliot, "Tunable UHF Oscillator with Harmonic Limitation," Lignes Telegraphoques et Telephoniques (Societe), Paris Cedex, France, U.S. Patent No. 4,146,850, March 27, 1979.

Thomas M. Higgins, Jr., "Negative Resistance Oscillator with Electronically Tunable Base Inductance," Hewlett-Packard Company, Palo Alto, California, U.S. Patent No. 5,373,264, December 13, 1994.

Osamu Ichiyoshi, "Voltage-Control Oscillator Which Suppresses Phase Noise Caused by Internal Noise of the Oscillator," NEC Corporation, Japan, U.S. Patent No. 5,351,014, September 27, 1994.

Shankar R. Joshi, Ulrich L. Rohde, and Klaus Eichel, "Phase-Locked Loop Circuits and Voltage Controlled Oscillator Circuits," Synergy Microwave Corp., Paterson, New Jersey, U.S. Patent No. 5,650,754, July 22, 1997.

Joseph H. Kiser, "Wide-Range Electronic Oscillator," Vari-L Co., Denver, Colorado, U.S. Patent No. 4,621,241, November 4, 1986.

Joseph H. Kiser, "Oscillator Voltage Regulator," Vari-L Co., Denver, Colorado, U.S. Patent No. 5,675,478, October 7, 1997.

Charles Lewis, "Low Phase Noise UHF and Microwave Oscillator," Z-Communications, Inc., San Diego, California, U.S. Patent No. 5,748,051, May 5, 1998.

Larry R. Lockwood, "Low-Noise Oscillator," Tektronix, Inc., Beaverton, Oregon, U.S. Patent No. 4,580,109, April 1, 1986.

Benedict J. Nardi, "Low Phase Noise Reference Oscillator," Watkins-Johnson Company, Palo Alto, California, U.S. Patent No. 5,341,110, August 23, 1994.

Leon Riebman, "Low-Noise Oscillator," AEL Defense Corp., Lansdale, Pennsylvania, U.S. Patent No. 5,166,647, November 24, 1992.

Takeshi Saitoh and Hiroshi Hatashita, "TV Tuner Oscillator with Feedback for More Low-Frequency Power," Hitachi Ltd., Tokyo, Japan, U.S. Patent No. 4,564,822, January 14, 1986.

FURTHER READING

Stanislaw Alechno, "Analysis Method Characterizes Microwave Oscillators," Part 1, *Microwaves & RF*, **36**(11), 82–86, November 1997.

P. Antognetti and G. Massobrio, *Semiconductor Device Modeling with SPICE.* New York: McGraw-Hill, 1993.

W. Anzill, F. X. Kärtner, and P. Russer, "Simulation of the Single-Sideband Phase Noise of Oscillators," *Second International Workshop of Integrated Nonlinear Microwave and Millimeterwave Circuits*, 1992.

Behagi, "Piece-wise Linear Modeling of Solid-State Varactor-Tuned Microwave Oscillators," *Proceedings of the IEEE Frequency Control Symposium*, 1992, pp. 415–519.

Noël Boutin, "RF Oscillator Analysis and Design by the Loop Gain Method," *Appl. Microwave & Wireless*, **11**(8), 32–48, August 1999.

P. Braun, B. Roth, and A. Beyer, "A Measurement Setup for the Analysis of Integrated Microwave Oscillators," *Second International Workshop of Integrated Nonlinear Microwave and Millimeterwave Circuits,* 1992.

C. R. Chang, M. B. Steer, S. Martin, and E. Reese, "Computer-Aided Analysis of Free-Running Microwave Oscillators," *IEEE Trans. Microwave Theory Tech.,* **MTT-39**, 1735–1745, October 1991.

Y. Cheng, K. Chen, K. Imai, and C. Hu, "A Unified MOSFET Channel Charge Model for Device Modeling in Circuit Simulation," *IEEE Trans. Computer-Aided Des.*, **17**, 641–644, August 1998.

J. Craninckx and M. Steyaert, "A 1.8-GHz Low-Phase-Noise CMOS VCO Using Optimized Hollow Spiral Inductors," *IEEE J. Solid-State Circuits*, **SC-32**(5), 736–744, May 1997.

P. Davis, P. Smith, E. Campbell, J. Lin, K. Gross, G. Bath, Y. Low, M. Lau, Y. Degani, J. Gregus, R. Frye, and K. Tai, "Silicon-on-Silicon Integration of a GSM Transceiver with VCO Resonator," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers,* pp. 248–249.

Alper Demir, Amit Mehrotra, and Jaijeet Roychowdhury, "Phase Noise and Timing Jitter in Oscillators," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 45–48.

P. J. Garner, M. H. Howes, and C. M. Snowden, "*Ka*-Band and MMIC pHEMT-Based VCOs with Low Phase-Noise Properties," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1531–1536, October 1998.

G. Gonzalez, *Microwave Transistor Amplifiers—Analysis and Design,* 2nd ed. New York: Prentice-Hall, 1997.

A. V. Grebennikov and V. V. Nikiforov, "An Analytic Method of Microwave Transistor Oscillator Design," *Int. J. Electron.,* **83**, 849–858, December 1997.

Andrey V. Grebennikov, "Microwave Transistor Oscillators: An Analytic Approach to Simplify Computer-Aided Design," *Microwave J.*, **42**(5), 292–300, May 1999.

Ali Hajimiri and Thomas H. Lee, "A General Theory of Phase Noise in Electrical Oscillators," *IEEE J. Solid-State Circuits*, **SC-33**, 179–194, February 1998.

Ali Hajimiri, Sotiriosw Limotyrakis, and Thomas H. Lee, "Phase Noise in Multi-gigahertz CMOS Ring Oscillators," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 49–52.

Frank Herzel and Behzad Razavi, "Oscillator Jitter Due to Supply and Substrate Noise," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 489–492.

Qiuting Huang, "On the Exact Design of RF Oscillators," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference*, pp. 41–44.

Kamran Iravani and Gary Miller, "VCOs with Very Low Sensitivity to Noise on the Power Supply," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 515–518.

K. M. Johnson, "Microwave Varactor-Tuned Transistor Oscillator Design," *IEEE Trans. Microwave Theory Tech.*, **MTT-14**(11), 564–572, September 1966.

K. M. Johnson, "Large-Signal GaAs MESFET Oscillator Design," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**(3), 217–226, March 1979.

W. E. Kamali, J. P. Grimm, R. Meierer, and C. Trirons, "New Design Approach for Wideband FET Voltage-Controlled Oscillators," *IEEE Trans. Microwave Theory Tech.,* **MTT-34**(10), 1059–1063, October 1986.

F. X. Kärtner, "Noise in Oscillating Systems," *Second International Workshop of Integrated Nonlinear Microwave and Millimeterwave Circuits*, 1992.

T. Kashiwa, T. Ishida, T. Katho, H. Kurusu, H. Hoshi, and Y. Mitsui, "V-Band High-Power Low Phase-Noise Monolithic Oscillators and Investigation of Low Phase-Noise Performance at High Drain Bias," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1559–1565, October 1998.

Peter Kinget, "A Fully Integrated 2.7V 0.35 μm CMOS VCO for 5GHz Wireless Applications," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 226–228.

J. Kitchen, "Octave Bandwidth Varactor-Tuned Oscillators," *Microwave J.*, **30**(5), 347–353, May 1987.

A. Kral, F. Behbahani, and A. A. Abidi, "RF-CMOS Oscillators with Switched Tuning," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference*, pp. 555–558.

Vênceslav F. Kroupa, ed., *Direct Digital Frequency Synthesizers.* New York: IEEE Press, 1999.

K. Kurokawa, "Some Basic Characteristics of Broadband Negative Resistance Oscillator Circuits," *Bell Syst. Tech. J.*, 1937–1955, July 1969.

Fred Labaar, "New Discriminator Boosts Phase-Noise Testing," *Microwaves*, 65–69, March 1982.

A. L. Lacaita and C. Samori, "Phase-Noise Performance of Crystal-Like *LC* Tanks," *IEEE Trans. Circuits Syst. Analog Digital Signal Processing*, **45**, 898–900, 1998.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

M. Madihian and T. Noguchi, "An Analytical Approach to Microwave GaAsFET Oscillator Design," *NEC Res. Dev.*, No. 82, 49–56, July 1986.

J. O. McSpadden, L. Fan, and K. Chang, "High-Efficiency *Ku*-Band Oscillators," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1566–1571, October 1998.

Mini-Circuits Engineering, "Designers' Considerations for VCOs Used in Phase-Locked Loops," *Microwave Prod. Digest*, 24, 34, 36, 37, 40, February 1999.

T. Nakagawa and T. Ohira, "A Novel Phase Noise Reduction Technique for MMIC Frequency Synthesizers that Uses Newly Developed Pulse Generator LSI," NTT Radio Communication Systems Laboratories, 1-2356 Take, Yokosuka-shi, Kanagawa 238-03, Japan.

Mohamed K. Nezami, "Evaluate the Impact of Phase Noise on Receiver Performance," *Microwaves & RF*, **41**(5), 165–172, May 1998.

E. C. Niehenke and R. D. Hess, "A Microstrip Low Noise X-Band Voltage-Controlled Oscillator," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**(12), 1075–1079, October 1986.

Chris O'Connor, "Develop Trimless Voltage-Controlled Oscillators," *Microwaves & RF*, **38**(7), 68–78, July 1999.

T. Ohira et al., "A Compact Full MMIC Module for Ku-Band Phase-Locked Oscillators," *IEEE Trans. Microwave Theory Tech.*, **MTT-37**(4), 723–728, April 1989.

D. P. Owen, "Measurement of Phase Noise in Signal Generators," *Marconi Instrum.,* **15**(6), 117–122, 19xx.

D. F. Page, "A Design Basis for Junction Transistor Oscillator Circuits," *Proc. IRE,* **46**, 1271–1280, November 1958.

Byeong-Ha Park, "1 GHz, Low-Power, Low-Phase-Noise, CMOS Voltage-Controlled Oscillator with an Integrated LC Resonator," private communication with the author.

D. F. Paterson, "Varactor Properties for Wideband Linear-Tuning Microwave VCOs," *IEEE Trans. Microwave Theory Tech.*, **MTT-28**(2), 110–119, February 1980.

Matthew Pope, Russell Crouch, and Jeff Gudewicz, "Dual Band Voltage Controlled Oscillators for Personal Communication," *Microwave Prod. Digest,* 52, 84, 85, May 1999.

Nenad Popovic, "Review of Some Types of Varactor Tuned DROs," *Appl. Microwave & Wireless,* **11**(8), 62–70, August 1999.

R. A. Pucel and J. Curtis, "Near-Carrier Noise in FET Oscillators," *IEEE MTT-S International Microwave Symposium Digest*, 1983, pp. 282–284.

S. Qi, K. Wu, and Z. Ou, "Hybrid Integrated HEMT Oscillator with a Multiple-Ring Nonradiative Dielectric (NRD) Resonator Feedback Circuit," *IEEE Trans. Microwave Theory Tech.,* **MTT-46**, 1552–1558, October 1998.

M. Regis, O. Llopis, and J. Graffeuil, "Nonlinear Modeling and Design of Bipolar Transistors Ultra-Low Phase-Noise Dielectric-Resonator Oscillators," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1589–1593, October 1998.

V. Rizzoli, F. Mastri, and D. Masotti, "A General Purpose Harmonic Balance Approach to the Computation of Near-Carrier Noise in Free-Running Microwave Oscillators," *MTT-S,* **x**, 309–312, 1993.

M. Runkel, "Analyse und Vergleich Spannungsgesteuerter Oszillatoren für Mobilfunkanwendungen," Master Thesis, Gerhard-Mercator-Universistät Gesamthochschule Duisburg, Chaired by Prof. Dr. Ingo Wolff, August 1994.

Carlo Samori, Andrea L. Lacaita, Francesco Vill, and Franco Zappa, "Spectrum Folding and Phase Noise in *LC* Tuned Oscillators," *IEEE Trans. Circuits Syst. Analog Digital Signal Processing,* **45**, 781–790, 1998.

Sarafian and B. Z. Kaplan, "A New Approach to the Modeling of the Dynamics of RF VCOs and Some of Its Practical Implications," *IEEE Trans. Circuits Syst. Fundam. Theory Applic.,* **40**(12), 895–901, December 1993.

G. Sauvage, "Phase Noise in Oscillators: A Mathematical Analysis of Leeson's Model," *IEEE Trans. Instrum. Meas.*, **IM-26**(4), x–x, December 1977.

Jwo-Shiun Sun, "Design and Analysis of Microwave Varactor-Tuned Oscillators," *Microwave J.,* 302–310, May 1999.

R. J. Trew, "Octave-Band GaAs FET YIG-Tuned Oscillators," *Electron. Lett.,* **13**(21), October 1977.

R. J. Trew, "MESFET Models for Microwave Computer-Aided Design," *Microwave J.,* **33**(5) 115–128, May 1990.

T. Yamasaki and E. Nagata, *Microwave Synthesizer for Terrestrial and Satellite Communications,* NEC Corporation, June 1989.

N. L. Wang and W. J. Ho, "X-Band HBT VCO with High-Efficiency CB Buffer Amplifier," *IEEE J. Solid-State Circuits,* **SC-27**(10), 1439–1443, October 1992.

Shing-Tak Yan and Howard C. Luong, "A 3-V 1.3-to-1.8-GHz CMOS Voltage-Controlled Oscillator with 0.3-ps Jitter," *IEEE Trans. Circuits Syst. Analog Digital Signal Processing,* **45**, 876–880, 1998.

D. A Warren, J. M. Golio, and W. L. Seely, "Large and Small Signal Oscillator Analysis," *Microwave J.*, **32**(5), 229–246, May 1989.

R. Winch, "Wideband Varactor-Tuned Oscillators," *IEEE J. Solid-State Circuits,* **SC-17**, 1214–1219, December 19, 1982.

Markus Zannoth, Bernd Kolb, Josef Fenk, and Robert Weigel, "A Fully Integrated VCO at 2 GHz," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers,* pp. 224–225.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

5

RF/WIRELESS OSCILLATORS

5-1 INTRODUCTION TO FREQUENCY CONTROL

Practically all modern telecommunication and test equipment uses frequency-control techniques based on *frequency synthesis*, the production of the many frequencies involved in a radiocommunication system's modulation, transmission, reception, and demodulation functions through the combination and mathematical manipulation of a very few input frequencies. Although the frequency synthesis techniques now ascendant in wireless systems—phase-locked loops and, less commonly, direct digital synthesizers—are fundamentally different, all are ultimately based on RF oscillators. This chapter covers oscillator theory, evaluation, and design.

Two types of oscillators are needed in a phase-locked-loop system (Figure 5-1). One, typically a *crystal oscillator*, generates the synthesizer's reference signal. As Figure 5-1 reflects, the reference oscillator may be built into the synthesizer IC in highly integrated systems. The other oscillator, a *voltage-controlled oscillator* (VCO), is varied in frequency by the system to produce the synthesizer's output signal. Although designing good oscillators remains somewhat like black magic or a special art, we will show that mathematics and CAD tools, applied in conjunction with practical experience, can keep the oscillator design process well under control.

5-2 BACKGROUND

The oscillator was probably first discovered by people who wanted to build an amplifier back in the vacuum-tube days. Its modern equivalent is shown in Figure 5-2. Here we see an active device, a bipolar transistor or field-effect transistor, with a tuned input and output circuit. Under ideal circumstances, the voltage at the input is 180° out of phase and the amplifier is "stable." Because of the unavoidable feedback capacitance, a certain amount of energy is transferred from the output to the input at about 90° out of phase. If one of the tuned circuits is detuned to the point where the phase shift is –180°, oscillation will occur. This means that the small energy from the output will be amplified, brought back to the input in phase and

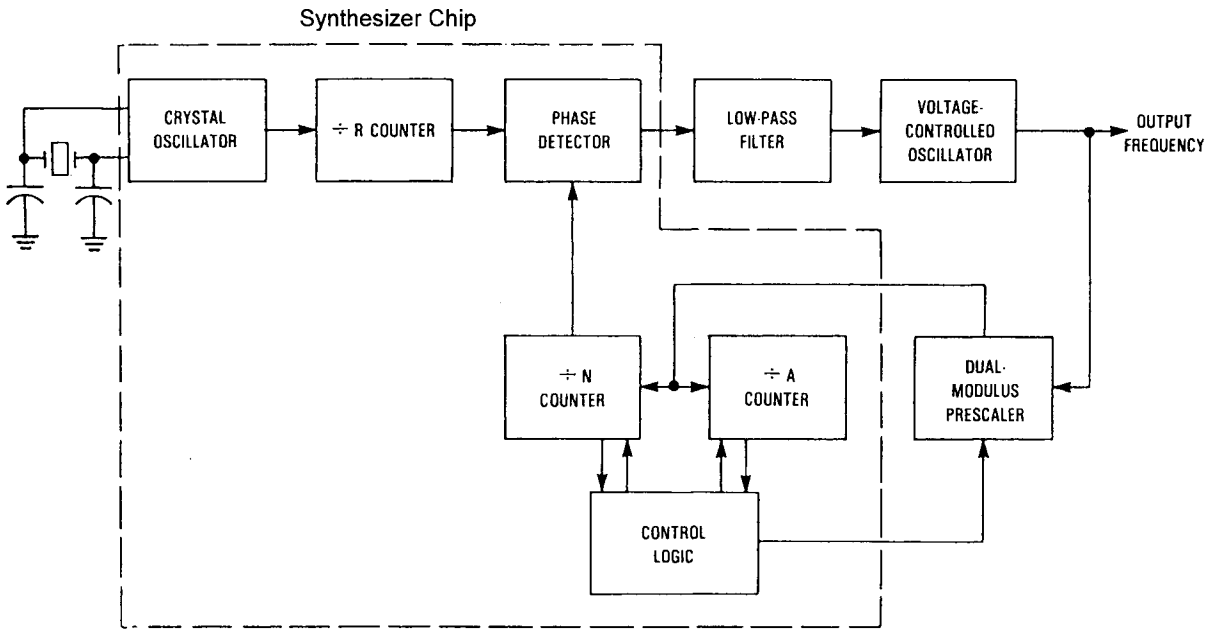


Figure 5-1 Block diagram of a modern, integrated frequency synthesizer. In this case, the designer has control over the VCO and the loop filter; the reference oscillator is part of the chip. In most cases (up to 2.5 GHz), the dual-modulus prescaler is inside the chip.

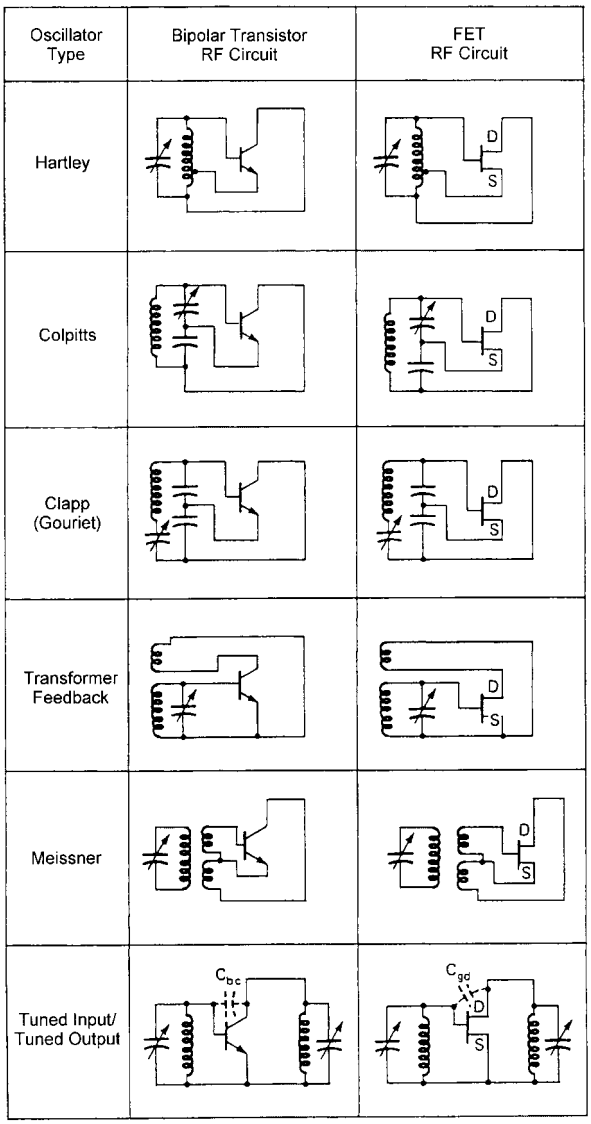


Figure 5-2 Schematic diagrams of RF connections for common oscillator circuits (dc and biasing circuits not shown) [1].

further amplified, causing the stage to "take off" and oscillate. Unless there is some mechanism to limit the amplitude of the oscillation at the input or output, its amplitude can theoretically increase to a level sufficient to destroy the device through breakdown effects or thermal runaway.

The criterion for oscillation was first described by Barkhausen [see Eq. (5-6) in the next section]. The conditions for oscillation exist when the feedback gain is high enough to cancel all losses and the difference between forward gain and reverse gain is less than zero. This implies also that oscillator circuits (Figure 5-2) provide a negative resistance, which is responsible for oscillation. The following mathematical treatment explains this.

5-3 OSCILLATOR DESIGN

A phase-locked loop generally has two oscillators: the oscillator at the output frequency and the reference oscillator. The reference oscillator at times can be another loop that is being mixed in, and the voltage-controlled oscillator (VCO) is controlled by either the reference or the oscillator loop. The VCO is one of the most important parts in a phase-locked-loop system, the performance of which is determined only by the loop filter at offsets inside the loop bandwidth, and only by the quality of the VCO design at offsets outside the loop bandwidth.

To some designers, VCO design appears to be magic. Shortly, we will go through the mathematics of the oscillator and some of its design criteria, but the results have only limited meaning. This is due to component tolerances, stray effects, and, most of all, nonlinear performance of the oscillator device, which can be modeled with only a certain degree of accuracy. However, after building oscillators for awhile, a certain feeling will be acquired for how to do this, and certain performance behavior will be predicted on a rule-of-thumb basis rather than on precise mathematical effort. For reasons of understanding, we will deal with the necessary mathematical equations, but we consider it essential to explain that these are only approximations.

5-3-1 Basics of Oscillators

An electronic oscillator is a device that converts dc power to a periodic output signal (ac power). If the output waveform is approximately sinusoidal, the oscillator is referred to as *sinusoidal.* There are many other oscillator types normally referred to as *relaxation* oscillators. For application in frequency synthesizers, we will explore only sinusoidal oscillators for reasons of purity and noise-sideband performance.

All oscillators are inherently nonlinear. Although the nonlinearity results in some distortion of the signal, linear analysis techniques can normally be used for the analysis and design of oscillators. Figure 5-3 shows, in block diagram form, the necessary components of an oscillator. It contains an amplifier with frequency-dependent forward loop gain $G(j\omega)$ and a frequency-dependent feedback network $H(j\omega)$.

The output voltage is given by

$$V_o = \frac{V_{\text{in}} G(j\omega)}{1 + G(j\omega)H(j\omega)} \tag{5-1}$$

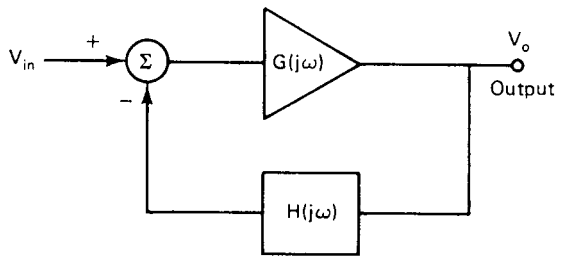


Figure 5-3 Block diagram of an oscillator showing forward and feedback loop components.

For an oscillator, the output V_o is nonzero even if the input signal $V_i = 0$. This can only be possible if the forward loop gain is infinite (which is not practical), or if the denominator

$$1 + G(j\omega)H(j\omega) = 0 \tag{5-2}$$

at some frequency ω_o. This leads to the well-known condition for oscillation (the *Nyquist criterion*), where at some frequency ω_o

$$G(j\omega_o)H(j\omega_o) = -1 \tag{5-3}$$

That is, the magnitude of the open-loop transfer function is equal to 1:

$$|G(j\omega_o)H(j\omega_o)| = 1 \tag{5-4}$$

and the phase shift is 180°:

$$\arg[G(j\omega_o)H(j\omega_o)] = 180° \tag{5-5}$$

This can be more simply expressed as follows: If in a negative-feedback system, the open-loop gain has a total phase shift of 180° at some frequency ω_o, the system will oscillate at that frequency provided that the open-loop gain is unity. If the gain is less than unity at the frequency where the phase shift is 180°, the system will be stable, whereas if the gain is greater than unity, the system will be unstable.

This statement is not correct for some complicated systems, but it is correct for those transfer functions normally encountered in oscillator design. The conditions for stability are also known as the *Barkhausen criteria*, which state that if the closed-loop transfer function is

$$\frac{V_o}{V_i} = \frac{\mu}{1 - \mu\beta} \tag{5-6}$$

the system will oscillate provided that $\mu\beta = 1$. This is equivalent to the Nyquist criterion, the difference being that the transfer function is written for a loop with positive feedback. Both versions state that the total phase shift around the loop must be 360° at the frequency of oscillation and the magnitude of the open-loop gain must be unity at that frequency.

The following analysis of the relatively simple oscillator shown in Figure 5-4 illustrates the design method. The linearized (and simplified) equivalent circuit of Figure 5-4 is given in Figure 5-5. h_{rb} has been neglected, and $1/h_{ob}$ has been assumed to be much greater than the load resistance R_L and is also ignored.

Note that the transistor is connected in the common-base configuration, which has no voltage phase inversion (the feedback is positive), so the conditions for oscillation are

$$|G(j\omega_o)H(j\omega_o) = 1| \tag{5-7}$$

and

$$\arg[G(j\omega_o)H(j\omega_o)] = 0° \tag{5-8}$$

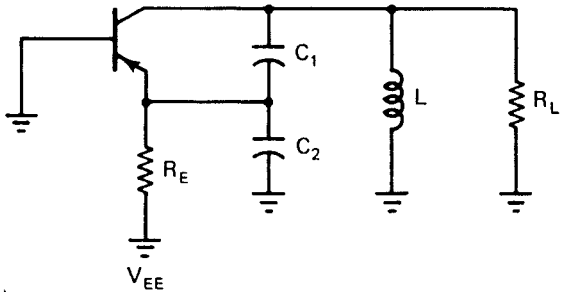


Figure 5-4 Oscillator with capacitive voltage divider.

The circuit analysis can greatly be simplified by assuming that

$$\frac{1}{\omega(C_2 + C_1)} << \frac{h_{ib}R_E}{h_{ib} + R_E} \tag{5-9}$$

and also that the Q of the load impedance is high. In this case the circuit reduces to that of Figure 5-6, where

$$V = \frac{V_o C_1}{C_1 + C_2} \tag{5-10}$$

and

$$R_{\text{eq}} = \frac{h_{ib}R_E}{h_{ib} + R_E}\left(\frac{C_1 + C_2}{C_1}\right)^2 \tag{5-11}$$

Figure 5-7 shows how the input impedance of an oscillator circuit—in this case, a BJT Colpitts oscillator minus its resonator—satisfies this condition for oscillation.

Then the forward gain

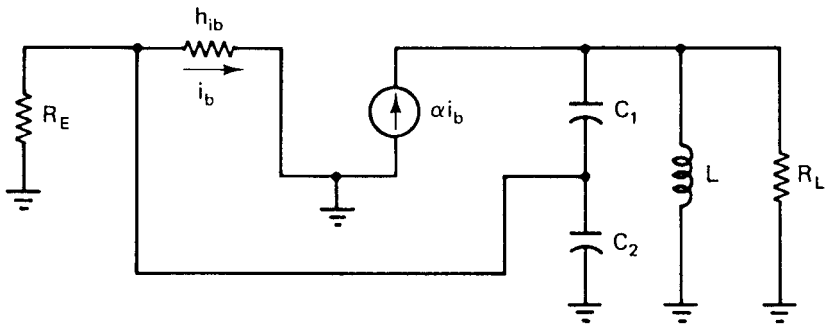


Figure 5-5 Linearized and simplified equivalent circuit of Figure 5-4.

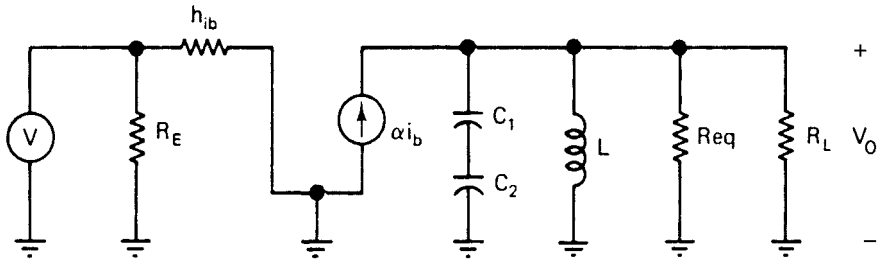


Figure 5-6 Further simplification of Figure 5-4, assuming high-impedance loads.

$$G(j\omega) = \frac{h_{fb}}{h_{ib}} Z_L = \frac{\alpha}{h_{ib}} Z_L \tag{5-12}$$

and

$$H(j\omega) = \frac{C_1}{C_1 + C_2} \tag{5-13}$$

where

$$Y_L = \frac{1}{Z_L} = \frac{1}{j\omega L} + \frac{1}{R_{\text{eq}}} + \frac{1}{R_L} + \frac{1}{j\omega C} \tag{5-14}$$

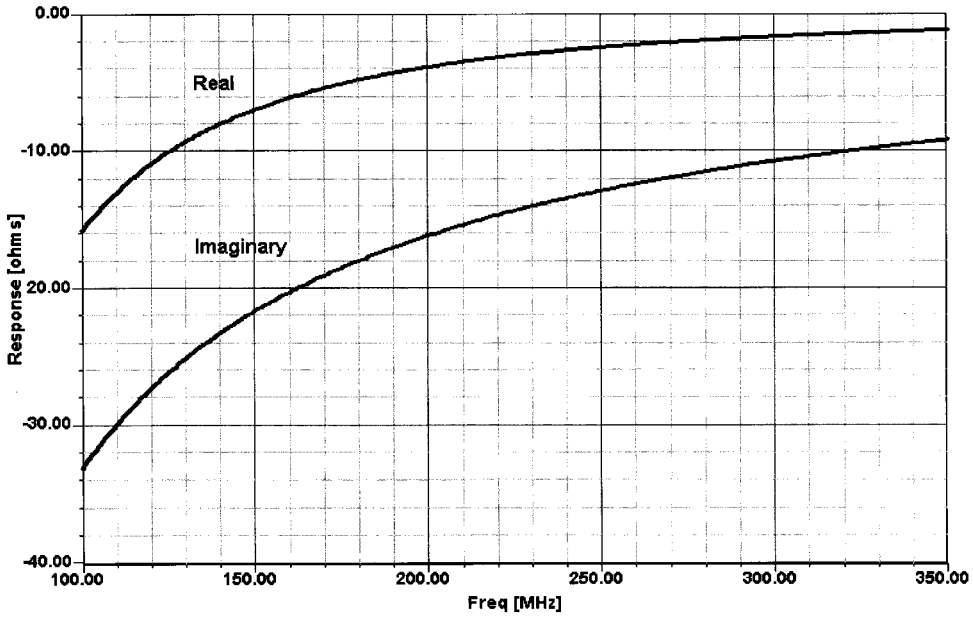


Figure 5-7 Calculated input impedance of a 200-MHz BJT Colpitts oscillator with its resonator removed. (Figure 5-46 shows an equivalent graph with the resonator present.)

A necessary condition for oscillation is that

$$\arg[G(j\omega)H(j\omega)] = 0^\circ \tag{5-15}$$

Since H does not depend on frequency in this example, if arg(GH) is zero, the phase shift of the load impedance Z_L must also be zero. This occurs only at the resonant frequency of the circuit,

$$\omega_o = \frac{1}{\sqrt{L[C_1C_2/(C_1 + C_2)]}} \tag{5-16}$$

At this frequency

$$Z_L = \frac{R_{\text{eq}}R_L}{R_{\text{eq}} + R_L} \tag{5-17}$$

and

$$GH = \frac{h_{fb}}{h_{ib}}\left(\frac{R_{\text{eq}}R_L}{R_{\text{eq}} + R_L}\right)\frac{C_1}{C_1 + C_2} \tag{5-18}$$

The other condition for oscillation is the magnitude constraint that

$$G(j\omega)H(j\omega) = \frac{\alpha}{h_{ib}}\left(\frac{R_{\text{eq}}R_L}{R_{\text{eq}} + R_L}\right)\frac{C_1}{C_1 + C_2} = 1 \tag{5-19}$$

Three-Reactance Oscillators. Although the block-diagram formulation of the stability criteria is an easy way to express stability mathematically, it is frequently not the easiest to apply since it is often difficult to identify the forward loop gain $G(j\omega)$ and the feedback ratio $H(j\omega)$ in electronic systems. A direct analysis of the circuit equations is frequently simpler than the block diagram interpretation (particularly for single-stage amplifiers). Figure 5-8 shows a generalized circuit for an electronic amplifier.

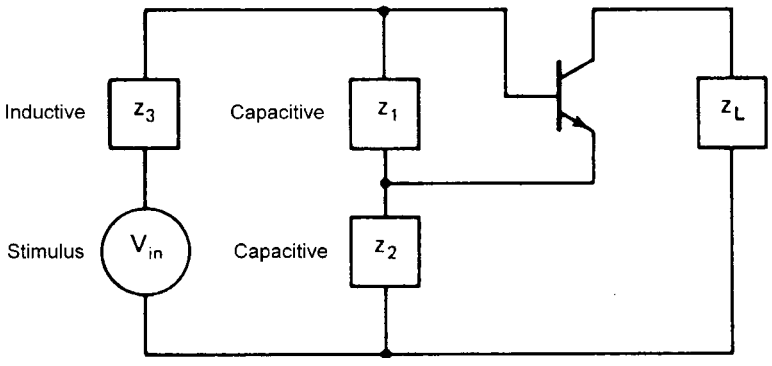


Figure 5-8 Generalized circuit for an oscillator using an amplifier model. Z_3 is inductive even with a capacitor in series with it.

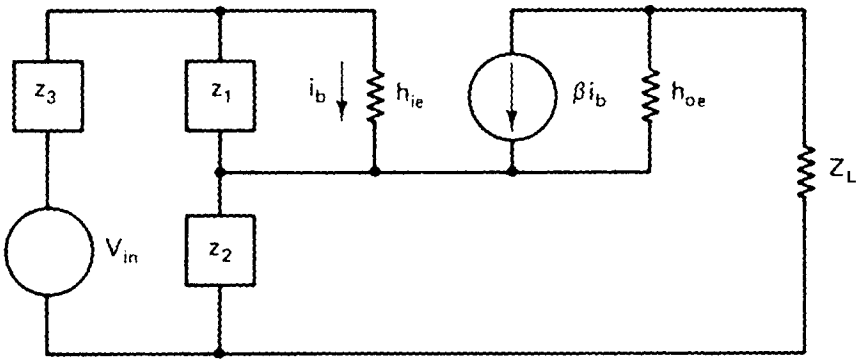


Figure 5-9 Small-signal equivalent circuit of Figure 5-8.

By inspecting Figure 5-8, one can see the three necessary reactances and the load. All other known circuits are derivatives of this by rotation. The small-signal equivalent circuit is given in Figure 5-9 (where h_{re} has been neglected).

Normally, h_{oe} can also be assumed sufficiently small and can be neglected. The loop equations are then

$$V_{\text{in}} = I_1(Z_3 + Z_1 + Z_2) - I_b Z_1 + \beta I_b Z_2 \tag{5-20}$$

$$0 = -I_1 Z_1 + I_b(h_{ie} + Z_1) \tag{5-21}$$

For the amplifier to oscillate, the currents I_b and I_1 must be nonzero even when $V_1 = 0$. This is only possible if the system determinant

$$\Delta = \begin{vmatrix} Z_3 + Z_1 + Z_2 & \beta Z_2 - Z_1 \\ -Z & h_{ie} + Z_1 \end{vmatrix} \tag{5-22}$$

is equal to 0. That is,

$$(Z_3 + Z_1 + Z_2)(h_{ie} + Z_1) - Z_1^2 + \beta Z_1 Z_2 = 0 \tag{5-23}$$

which reduces to

$$(Z_1 + Z_2 + Z_3)h_{fe} + Z_1 Z_2 \beta + Z_1(Z_2 + Z_3) = 0 \tag{5-24}$$

Assume for the moment that Z_1, Z_2, and Z_3 are purely reactive impedances. [It is easily seen that Eq. (5-24) does not have a solution if all three impedances are real.] Since both the real and imaginary parts must be zero, Eq. (5-24) is equivalent to the following equations if

$$h_{ie}(Z_1 + Z_2 + Z_3) = 0 \tag{5-25}$$

and

$$Z_1[(1 + \beta)Z_2 + Z_3] = 0 \tag{5-26}$$

Since β is real and positive, Z_2 and Z_3 must be of opposite sign for Eq. (5-26) to hold. That is,

$$(1 + \beta)Z_2 = -Z_3 \tag{5-27}$$

Therefore, since h_{ie} is nonzero, Eq. (5-25) reduces to

$$Z_1 + Z_2 - (1 + \beta)Z_2 = 0 \tag{5-28}$$

or

$$Z_1 = \beta Z_2 \qquad [C_1 < C_2; L_1 > L_2] \tag{5-29}$$

Thus, since β is positive, Z_1 and Z_2 will be reactances of the same kind. If Z_1 and Z_2 are capacitors, Z_3 is an inductor and the circuit is as shown in Figure 5-10. It is referred to as a *Colpitts* oscillator, named after the person who first described it.

If Z_1 and Z_2 are inductors and Z_3 is a capacitor as illustrated in Figure 5-11, the circuit is called a *Hartley* oscillator.

Example 1. Design a Colpitts circuit to oscillate at 200 MHz, using a transistor (operating at I_C = 6 mA dc) that has an input impedance

$$\text{Mag}(Z_{11}) = \frac{26 \text{ mV}}{I_C} \beta = 282\ \Omega (h_{ie} \triangleq Z_{11})$$

and

$$\beta_{RF} = 65; \quad \beta_{OSC} = 10 \quad \text{(large-signal condition)}$$

The oscillating transistor operates under large-signal conditions, generating harmonics in addition to the fundamental. Most of the transistor's current gain will come from dc biasing. β_{OSC}, its large-signal beta at the fundamental frequency of oscillation, must therefore be considerably less than β_{RF}, the small-signal, single-tone value, as the

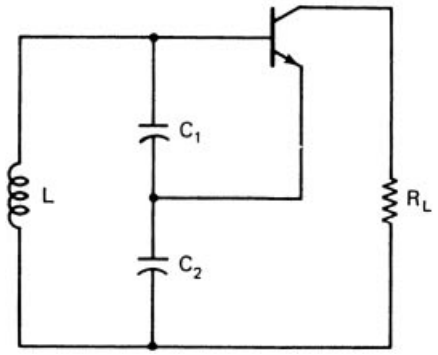


Figure 5-10 Colpitts oscillator.

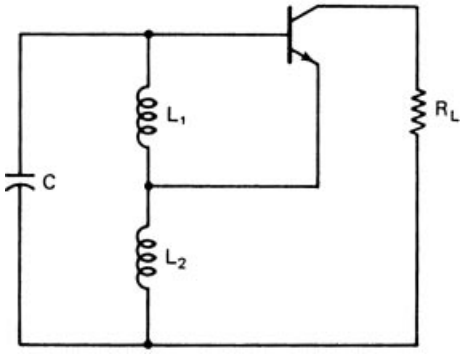


Figure 5-11 Hartley oscillator.

transistor's available current gain, β_0, is distributed among the dc, fundamental, and harmonic components present ($\beta_0 = \beta_{dc} + \beta_{OSC} + \beta_{OSC,F2} + \beta_{OSC,F3}\ldots$).

Solution: For the Colpitts circuit, Z_1 and Z_2 are capacitive reactances and Z_3 is an inductive reactance. Let $Z_2 = -j1.6\ \Omega$; then [Eq. (5-27)]

$$Z_3 = -(1 + \beta)Z_2 = -(1 + 10) \times (-1.6) \approx j17\ \Omega$$

and [Eq. (5-25)]

$$Z_1 = -(Z_2 + Z_3) \approx -j17\ \Omega$$

At 200 MHz, these impedances correspond to component values of $C_1 = 50$ pF, $C_2 = 500$ pF, and $L = 30$ nH. The oscillator will probably work with C_2 values larger than 500 pF, but at the loss of safety margin for tolerances in production, beta, temperature, and phase-noise performance. The completed circuit, except for biasing, is shown in Figure 5-12.

If Z_1, Z_2, Z_3, or h_{ie} is complex, the preceding analysis is more complicated, but the conditions for oscillation can still be obtained from Eq. (5-24). For example, if, in the Colpitts

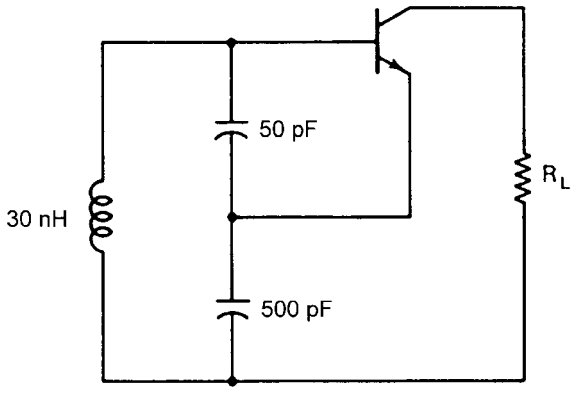


Figure 5-12 Design example of a Colpitts oscillator.

circuit, there is a resistor R in series with $L(Z_3 = R + j\omega L)$, Eq. (5-24) reduces to two equations:

$$h_{ie}\left(\omega L - \frac{1}{\omega C_1} - \frac{1}{\omega C_2}\right) - \frac{R}{\omega C_1} = 0 \tag{5-30}$$

and

$$h_{ie}R - \frac{1+\beta}{\omega^2 C_1 C_2} + \frac{1}{\omega C_1}\omega L = 0 \tag{5-31}$$

Define

$$C_1' = \frac{C_1}{1 + R/h_{ie}} \tag{5-32}$$

The resonant frequency at which oscillations will occur is found from Eq. (5-30) to be

$$\omega_o = \frac{1}{\sqrt{L[C_1' C_2/(C_1' + C_2)]}} \tag{5-33}$$

and for oscillations to occur $R_e(h_{ie}) = R_e(Z_{11})$ must be less than or equal to

$$R_e(h_{ie}) \leq \frac{1+\beta}{\omega_o^2 C_1 C_2} - \frac{L}{C_1} \tag{5-34}$$

If R becomes too large, Eq. (5-32) cannot be satisfied and oscillations will stop. In general, it is advantageous to have

$$X_{C_1} X_{C_2} = \frac{1}{\omega^2 C_1 C_2} \tag{5-35}$$

with $X \triangleq 1/\omega c$ as large as possible, since R can be large. However, if C_1 and C_2 are too small (large and X_{C_1} and X_{C_2}), the input and output capacitances of the transistor, which shunt C_1 and C_2, respectively, become important. A good, stable design will always have C_1 and C_2 much larger than the transistor capacitances they shunt.

Example 2. In Example 1, will the circuit still oscillate if the inductor now has a $Q_u = 100$? If the transistor input capacitance is 5 pF, what effect will this have on the system?

Solution: In Example 1, $X_L = 17\ \Omega$, $C_1 = 50$ pF, and $C_2 = 500$ pF. Since $C_1 = 50$ pF, adding 5 pF in parallel will change the equivalent C to 55 pF. As the inductor $Q_u = 100$, the equivalent resistance R in series with the lossless inductor is

$$R = \frac{17}{100} = 0.17\ \Omega$$

The new resonant frequency can be determined from Eqs. (5-32) and (5-33):

$$C_1' = \frac{C_1}{1 + R/h_{ie}} = \frac{55}{1 + (0.17/280)} \approx \frac{55}{1.0006} = 54.96 \text{ pF}$$

and

$$f_o = \frac{1}{2\pi L[C_1' C_2/(C_1' + C_2)]^{1/2}} \cong 191.56 \text{ MHz}$$

The effect of the finite inductor Q causes a negligible change in the oscillating frequency compared to the effect of the transistor input capacitance, which reduces the resonant frequency 4.2%. Because the large-signal beta, β_{OSC}, is 10, there is enough loop gain to maintain oscillation.

Two-Port Oscillator. Although Eqs. (5-25) and (5-26) can be used to determine the exact expressions for oscillation, they are often difficult to use and add little insight into the design process. An alternative interpretation, although not as accurate, will now be presented. It is based on the fact that an ideal tuned circuit (infinite Q), once excited, will oscillate infinitely because there is no resistance element present to dissipate the energy. In the actual case where the inductor Q is finite, the oscillations die out because energy is dissipated in the resistance. It is the function of the amplifier to maintain oscillations by supplying an amount of energy equal to that dissipated. This source of energy can be interpreted as a negative resistor in series with the tuned circuit. If the total resistance is positive, the oscillations will die out, while the oscillation amplitude will increase if the total resistance is negative. To maintain oscillations, the two resistors must be of equal magnitude. To see how a negative resistance is realized, the input impedance of the circuit in Figure 5-13 will be derived.

If h_{oe} is sufficiently small ($h_{oe} << 1/R_L$), the equivalent circuit is as shown in Figure 5-13. The steady-state loop equations are

$$V_{\text{in}} = I_{\text{in}}(X_{C_1} + X_{C_2}) - I_b(X_{C_1} - \beta X_{C_2}) \tag{5-36}$$

$$0 = -I_{\text{in}}(X_{C_1}) + I_b(X_{C_1} + h_{ie}) \tag{5-37}$$

After I_b is eliminated from these two equations, Z_{in} is obtained as

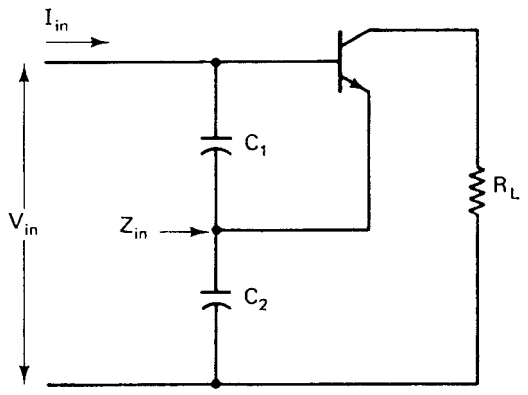


Figure 5-13 Calculation of input impedance of the negative-resistance oscillator.

$$Z_{\text{in}} = \frac{V_{\text{in}}}{I_{\text{in}}} = \frac{(1+\beta)X_{C_1}X_{C_2} + h_{ie}(X_{C_1} + X_{C_2})}{X_{C_1} + h_{ie}} \tag{5-38}$$

If $X_{C_1} << h_{ie}$, the input impedance is approximately equal to

$$Z_{\text{in}} \approx \frac{1+\beta}{h_{ie}} X_{C_1}X_{C_2} + (X_{C_1} + X_{C_2}) \tag{5-39}$$

$$Z_{\text{in}} \approx \frac{-g_m}{\omega^2 C_1 C_2} + \frac{1}{j\omega[C_1C_2/(C_1 + C_2)]} \tag{5-40}$$

That is, the input impedance of the circuit shown in Figure 5-14 is a negative resistor,

$$R = \frac{-g_m}{\omega^2 C_1 C_2} \tag{5-41}$$

in series with a capacitor,

$$C_{\text{in}} = \frac{C_1 C_2}{C_1 + C_2} \tag{5-42}$$

which is the series combination of the two capacitors.

With an inductor L (with the series resistance R_S) connected across the input, it is clear that the condition for sustained oscillation is

$$R_S = \frac{g_m}{\omega^2 C_1 C_2} \tag{5-43}$$

and the frequency of oscillation is

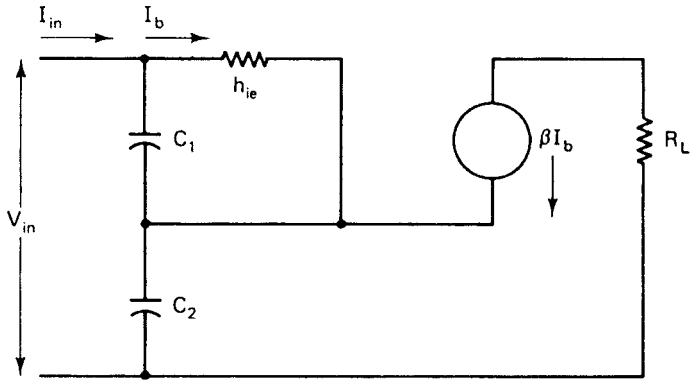


Figure 5-14 Equivalent small-signal circuit of Figure 5-13.

$$f_o = \frac{1}{2\pi\sqrt{L[C_1C_2/(C_1 + C_2)]}} \tag{5-44}$$

This interpretation of the oscillator readily provides several guidelines that can be used in the design. First, C_1 should be as large as possible so that

$$X_{C_1} << h_{ie} \tag{5-45}$$

and C_2 is to be large so that

$$X_{C_2} << \frac{1}{h_{oe}} \tag{5-46}$$

When these two capacitors are large, the transistor base-to-emitter and collector-to-emitter capacitances will have a negligible effect on the circuit's performance. However, Eq. (5-43) limits the maximum value of the capacitances since

$$r \leq \frac{g_m}{\omega^2 C_1 C_2} \leq \frac{G}{\omega^2 C_1 C_2} \tag{5-47}$$

where G is the maximum value of g_m. For a given product of C_1 and C_2, the series capacitance is at maximum when $C_1 = C_2 = C_m$. Thus Eq. (5-47) can be written

$$\frac{1}{\omega C_m} > \sqrt{\frac{r}{G}} \tag{5-48}$$

This equation is important in that it shows that for oscillations to be maintained, the minimum permissible reactance $1/\omega C_m$ is a function of the resistance of the inductor and the transistor's mutual conductance, g_m.

An oscillator circuit known as the *Clapp circuit* or *Clapp–Gouriet circuit* is shown in Figure 5-15. This oscillator is equivalent to the one just discussed, but it has the practical advantage of being able to provide another degree of design freedom by making C_o much smaller than C_1 and C_2.

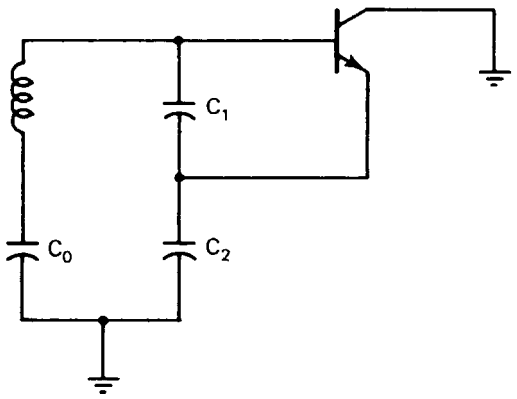


Figure 5-15 Circuit of a Clapp oscillator.

It is possible to use C_1 and C_2 to satisfy the condition of Eq. (5-47) and then adjust C_o for the desired frequency of oscillation ω_o, which is determined from

$$\omega_o L - \frac{1}{\omega_o C_o} - \frac{1}{\omega_o C_1} - \frac{1}{\omega_o C_2} = 0 \tag{5-49}$$

Amplitude Stability. Linearized analysis of the oscillator is convenient for determining the frequency but not the amplitude of the oscillation. The Nyquist stability criterion defines the frequency of oscillation as the frequency at which the loop phase shift is 360°, but it says nothing about the oscillation amplitude. If no provisions are taken to control the amplitude, it is susceptible to appreciable drift. Two frequently used methods for controlling the amplitude are operating the transistor in the nonlinear region or using a second stage for amplitude limiting. For the single-stage oscillator, amplitude limiting is accomplished by designing an unstable oscillator; that is, the loop gain is made greater than 1 at the frequency at which the phase shift is 180°. As the amplitude increases, the β of the transistor decreases, causing the loop gain to decrease until the amplitude stabilizes. This is a self-limiting oscillator. There are nonlinear analysis techniques predicting the amplitude of oscillation, but their results are approximate except in idealized cases, forcing the designer to resort to an empirical approach.

An example of a two-stage emitter-coupled oscillator is shown in Figure 5-16. In this circuit, amplitude stabilization occurs as a result of current limiting in the second stage. This circuit has the additional advantage that it has output terminals that are isolated from the feedback path. The emitter signal of Q_2, having a rich harmonic content, is normally used as output. Harmonics of the fundamental frequency can be extracted at the emitter of Q_2 by using an appropriately tuned circuit. Note that the collector of Q_2 is isolated from the feedback path.

Phase Stability. An oscillator has a frequency or phase stability that can be considered in two separate parts. First, there is *long-term* stability, in which the frequency changes over a period of minutes, hours, days, weeks, or even years. This frequency stability is normally limited by the circuit components' temperature coefficients and aging rates. The other part,

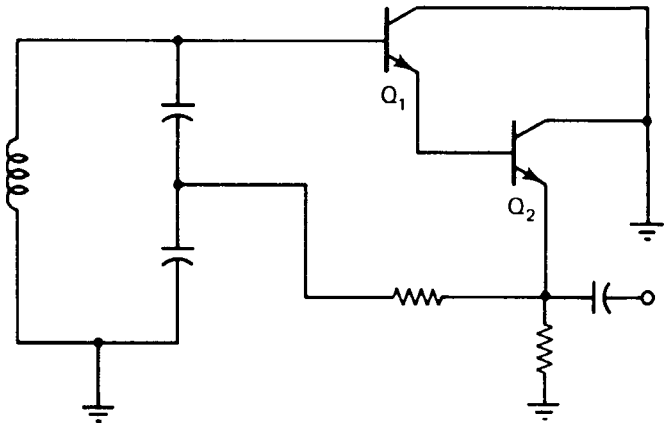


Figure 5-16 Two-stage, emitter-coupled oscillator.

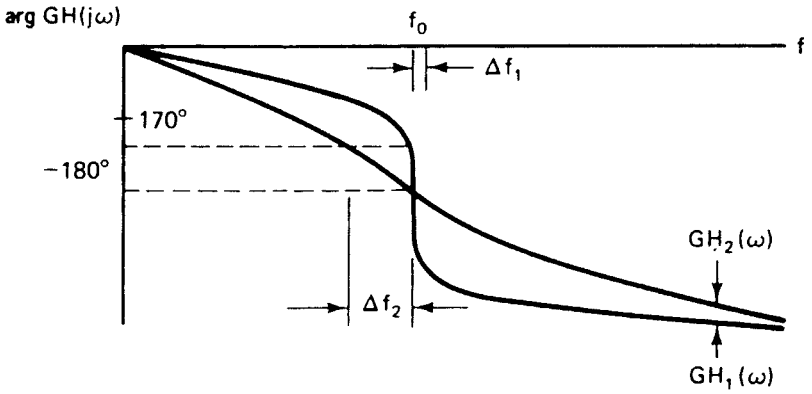


Figure 5-17 Phase plot of two open-loop systems with different resonator *Q* values.

short-term frequency stability, is measured in terms of seconds. One form of short-term instability is due to changes in phase of the system; here the term *phase stability* is used synonymously with frequency stability. It refers to how the frequency of oscillation reacts to small changes in phase shift of the open-loop system. It can be assumed that the system with the largest rate of change of phase versus frequency $d\phi/df$ will be the most stable in terms of frequency stability. Figure 5-17 shows the phase plots of two open-loop systems used in oscillators. At the system crossover frequency, the phase shift is −180°. If some external influence causes a shift in phase—say, it adds 10° of phase lag—the frequency will change until the total phase shift is again 0°. In this case, the frequency will decrease to the point where the open-loop phase shift is 170°. Figure 5-17 shows that Δf_2, the change in frequency associated with the 10° change in phase of GH_2, is greater than the change in frequency Δf_1, associated with the open-loop system GH_1, whose phase is changing more rapidly near the open-loop crossover frequency.

This qualitative discussion illustrates that $d\phi/df$ at $f = f_o$ is a measure of an oscillator's phase stability. It provides a good means of quantitatively comparing the phase stability of two oscillators. Consider the simple parallel tuned circuit shown in Figure 5-18.

For this circuit, the two-port is

$$\frac{V_o(j\omega)}{I(j\omega)} = \frac{R}{1 + jQ[(\omega/\omega_o) - (\omega_o/\omega)]} \tag{5-50}$$

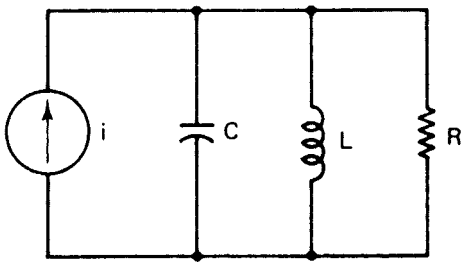


Figure 5-18 Parallel tuned circuit for phase-shift analysis.

where

$$\omega_o = \frac{1}{\sqrt{LC}} \quad \text{and} \quad Q = \frac{R}{\omega_o L} \tag{5-51}$$

The circuit phase shift is

$$\arg \frac{V_o}{I} = \theta = \tan^{-1} Q\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right) \tag{5-52}$$

and

$$\frac{d\theta}{d\omega} = \frac{1/Q}{1/Q^2 + [(\omega^2 - \omega_o^2)/\omega_o\omega]^2} \frac{\omega^2 + \omega_o^2}{(\omega_o\omega)^2} \tag{5-53}$$

At the resonant frequency ω_o,

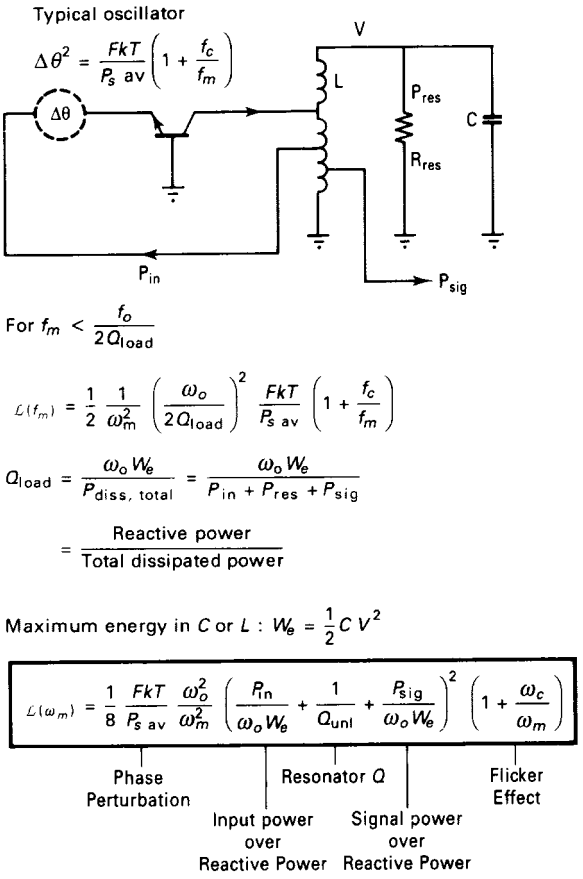


Figure 5-19 Diagram for a feedback oscillator illustrating the principles involved, and showing the key components considered in the phase noise calculation and its contribution.

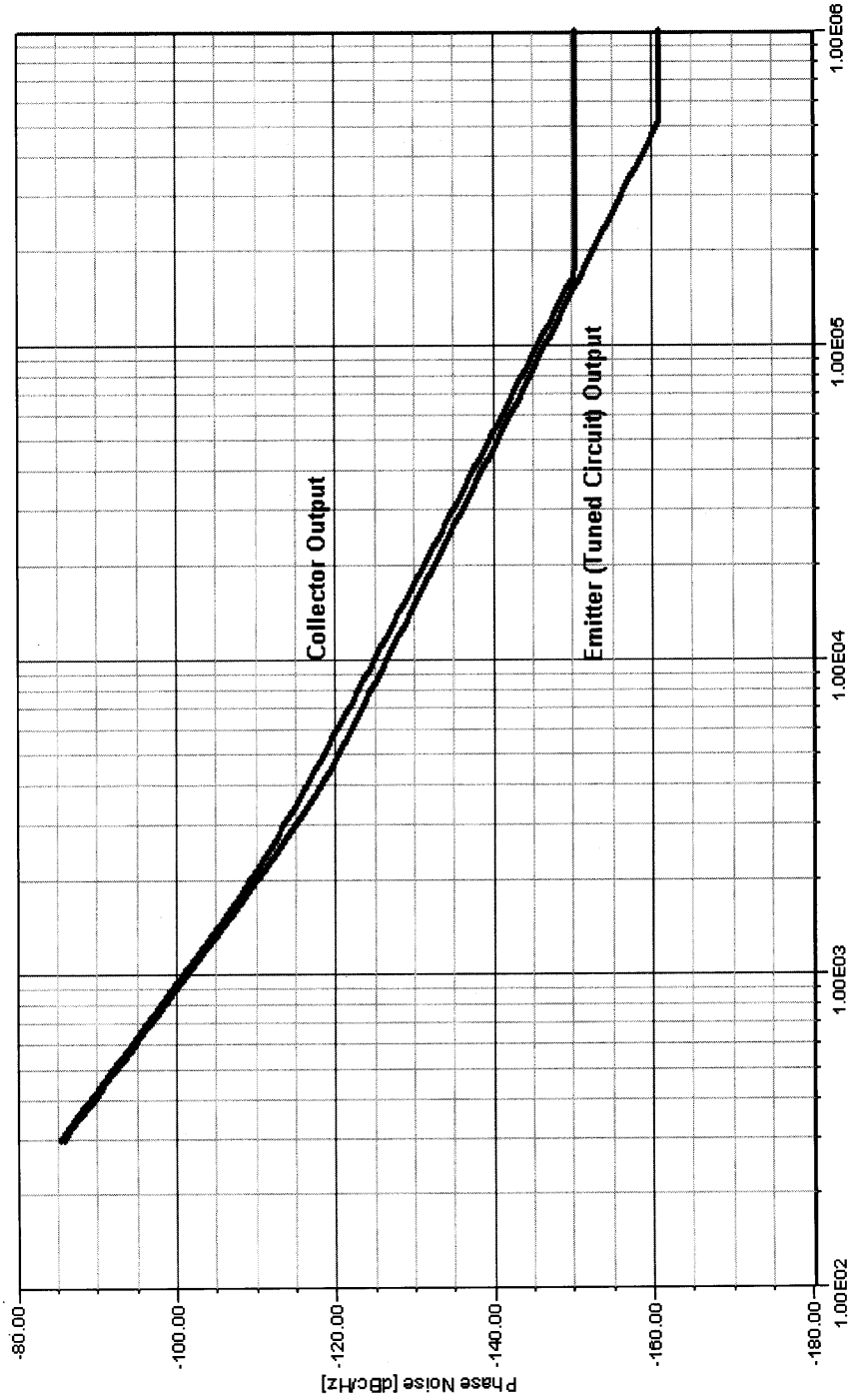


Figure 5-20 Phase noise with oscillator output taken from collector (upper trace) versus emitter (lower trace) of a BJT oscillator.

$$\left.\frac{dQ}{d\omega}\right|_{\omega=\omega_o} = \frac{2Q}{\omega_o} \tag{5-54}$$

The frequency stability factor S_F is defined as the change in phase $d\phi/d\omega$ divided by the normalized change in frequency $\Delta\omega/\omega_o$. That is,

$$S_F = 2Q \tag{5-55}$$

S_F is a measure of the short-term stability of an oscillator. Equation (5-54) indicates that the higher the circuit Q, the higher the stability factor. This is one reason for using high-Q circuits in oscillator circuits. Another reason is the ability of the tuned circuit to filter out undesired harmonics and noise.

The two oscillator types shown so far can be considered two-port oscillators because one port is responsible for oscillation and the other port supplies the output power. This type of output supposedly has better isolation (although this is not true at higher frequencies) but has a higher noise floor. A better way of extracting energy is shown in Figure 5-19.

These types of oscillators are essentially what is referred to as *two-port* oscillators. The name comes from the fact that the oscillating device has two ports, with the output energy taken from the output port, typically the collector or drain, while the tuned circuit is at the input port. For years, it was widely held that this type of design provided better isolation between the tuned circuit and termination; unfortunately, because of the Miller effect, this is not true. At the higher frequencies, changes in output-port loading affect the input port, resulting in *pulling*, a detuning effect that increases with frequency. In addition, taking the output from the collector or the drain results in a poorer signal-to-noise ratio than that achievable by extracting energy from the tuned circuit itself. The only real advantages the collector output provides are more power and higher efficiency, but definitely at the expense of phase noise, which is often referred to as *SSB phase noise* because it is expressed in terms of the decibel ratio of the noise power in a single (the upper or lower) noise sideband, in a 1-Hz bandwidth centered at a specified frequency offset from the oscillator carrier, to the carrier power. Figure 5-20 compares the phase noise of an oscillator with the energy taken off the collector and the tuned circuit; the difference is highly visible.

5-4 OSCILLATOR CIRCUITS

5-4-1 Hartley

Figure 5-21 shows the Hartley oscillator configuration, in which feedback is obtained by tapping the resonator. In practice, the resonator can consist of a single tapped inductor or two separate, magnetically uncoupled inductors, with feedback obtained at their junction.

5-4-2 Colpitts

Figure 5-22 shows the Colpitts oscillator configuration, in which feedback is obtained via a capacitive voltage divider.

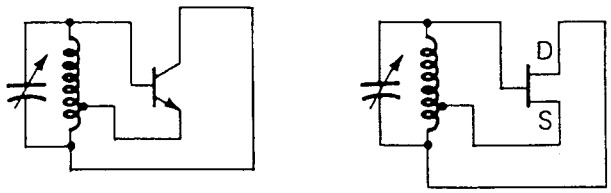


Figure 5-21 Hartley oscillator.

5-4-3 Clapp–Gouriet

Figure 5-23 shows the Clapp–Gouriet oscillator. Like the Colpitts, the Clapp–Gouriet obtains its feedback via a capacitive voltage divider; unlike the Colpitts, an additional capacitor series-tunes the resonator. The Pierce oscillator, a configuration used only with crystals, is a rotation of the Clapp–Gouriet oscillator in which the emitter is at RF ground.

5-5 DESIGN OF RF OSCILLATORS

5-5-1 General Thoughts on Transistor Oscillators

An estimate of the noise performance of an oscillator is

$$\mathcal{L}(\omega_m) = \frac{1}{8}\frac{FkT}{P_{s,\text{av}}}\frac{\omega_0^2}{\omega_m^2}\left(\frac{P_{\text{in}}}{\omega_0 W_e} + \frac{1}{Q_{\text{unl}}} + \frac{P_{\text{sig}}}{\omega_0 W_e}\right)^2\left(1 + \frac{\omega_c}{\omega_m}\right) \tag{5-56}$$

Phase perturbation

Resonator Q

Flicker effect

Input power/Reactive power ratio

Signal power/Reactive power ratio

We just state this here without presenting oscillator noise theory itself. (Also see Figure 5-19.) This equation is based on work done by Dieter Scherer of Hewlett-Packard about 1978. He was the first to introduce the flicker effect to the Leeson equation by adding the AM-to-PM

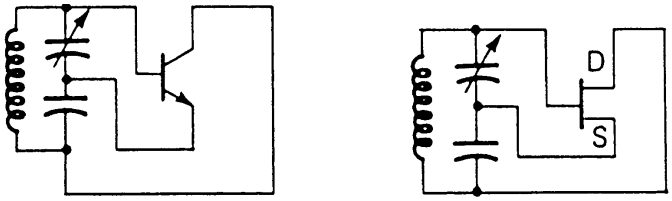


Figure 5-22 Colpitts oscillator.

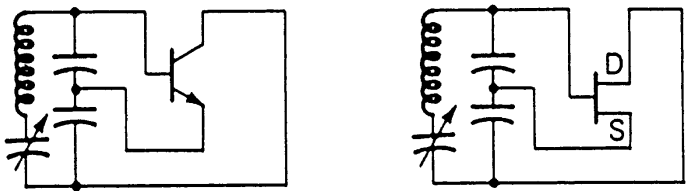


Figure 5-23 Clapp–Gouriet oscillator.

conversion effect, which is caused by the nonlinear capacitance of the active devices. This equation must be further expanded:

$$\pounds(f_m) = 10 \log\left\{\left[1 + \frac{f_0^2}{(2f_m Q_{\text{load}})^2}\right]\left(1 + \frac{f_c}{f_m}\right)\frac{FkT}{2P_{\text{sav}}} + \frac{2kTRK_0^2}{f_m^2}\right\} \tag{5-57}$$

where $\pounds(f_m)$ = ratio of sideband power in 1-Hz bandwidth at f_m to total power in dB
f_m = frequency offset
f_0 = center frequency
f_c = flicker frequency
Q_{load} = loaded Q of the tuned circuit
F = noise factor
$kT = 4.1 \times 10^{-21}$ at 300 K (room temperature)
P_{sav} = average power at oscillator output
R = equivalent noise resistance of tuning diode (typically 200 Ω to 10 kΩ)
K = oscillator voltage gain

The following table (from Motorola) shows the flicker corner frequency f_C as a function if I_C for a typical small-signal microwave BJT. $I_{C(\text{max})}$ of this transistor is about 10 mA.

I_C (mA)	f_C (kHz)
0.25	1
0.5	2.74
1	4.3
2	6.27
5	9.3

Note that f_C, which is defined by AF and KF in the SPICE model, increases with I_C. This gives us a clue about how f_C changes when a transistor oscillates. As a result of the bias-point shift that occurs during oscillation (see Example 1 of Section 5-3-1), an oscillating BJT's average I_C is higher than its small-signal I_C. KF is therefore higher for a given BJT operating as an oscillator than for the same transistor operating as a small-signal amplifier. This must be kept in mind when considering published f_C data, which are usually determined under small-signal conditions without being qualified as such. Determining a transistor's oscillating f_C is best done through measurement: Operate the device as a high-Q UHF oscillator (we

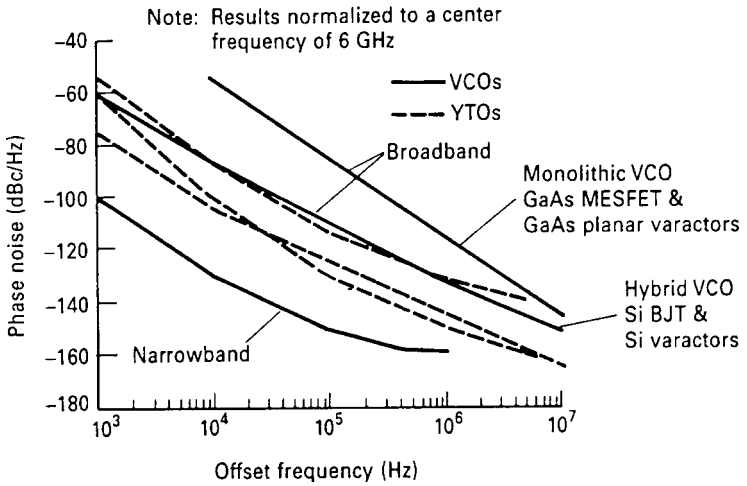


Figure 5-24 Phase noise of oscillators using different semiconductors and resonators.

suggest using a ceramic-resonator-based tank in the vicinity of 1 GHz), and measure its close-in (10 Hz to 10 kHz) phase noise versus offset from the carrier. f_C will correspond to a slight decrease in the slope of the phase noise versus offset curve. Generally, f_C varies with device type as follows: silicon JFETs, 50 Hz and higher; microwave RF BJTs, 1–10 kHz (as above); MOSFETs, 10–100 kHz; and GaAsFETs, 10–100 MHz. Figure 5-24 shows the phase noise of oscillators using different semiconductors and resonators.

Equation (5-57) is based on Rohde et al. [2]. The additional term introduces a distinction between a conventional oscillator and a VCO. Whether the voltage- or current-dependent capacitance is internal or external makes no difference; it simply affects the frequency.

For a more complete expression for a resonator oscillator's phase noise spectrum we can write

$$s_\phi(f_m) = \frac{\alpha_R F_0^4 + \alpha_E (F_0/2Q_L)^2}{f_m^3} + \frac{(2GFkT/P_0)\,(F_0/2Q_L)^2}{f_m^2} + \frac{2\alpha_R Q_L F_0^3}{f_m^2} + \frac{\alpha_E}{f_m} + \frac{2GFkT}{P_0} \tag{5-58}$$

where G = compressed power gain of the loop amplifier
F = noise factor of the loop amplifier
k = Boltzmann's constant
T = temperature in kelvins

P_0 = carrier power level (in watts) at the output of the loop amplifier
F_0 = carrier frequency in hertz
f_m = carrier offset frequency in hertz
$Q_L (= \pi F_0 \tau_g)$ = loaded Q of the resonator in the feedback loop
α_R, α_E = flicker-noise constants for the resonator and loop amplifier, respectively.

In frequency synthesizers, we have no use for *LC* oscillators that do not include a tuning diode or diodes, but it may still be of interest to analyze the low-noise, fixed-tuned *LC* oscillator first and later make both elements, inductor and capacitor, variable.

Later, we will show the performance changes if we utilize the two possible ways of getting coarse and fine tuning in oscillators:

1. Use of tuning diodes.
2. Use of switching diodes.

We will spend some time looking at the effects that switching and tuning diodes have in a circuit because they will ultimately influence the noise performance more strongly than the transistor itself.

The reason is that the noise generated in tuning diodes will be superimposed on the noise generated in the circuit, while switching diodes have losses that cause a reduction of circuit *Q*. The selection of the proper tuning and switching diodes is important, as is the proper way of connecting them. As both types are modifications of the basic *LC* oscillator, we start with the *LC* oscillator itself.

Early signal generators as were offered by several companies (namely, Rohde & Schwarz, Hewlett-Packard, Boonton Electronics, or Marconi), if they are not synthesized, use an air-dielectric variable capacitor or, as in the case of one particular Hewlett-Packard generator, the Model HP8640, a tuned cavity.

Tuning here is accomplished by changing the value of an air-dielectric variable capacitor or changing the mechanical length of a quarter-wave resonator.

Using the equations shown previously, it is fairly easy to calculate the performance of oscillators and understand how they work, but this does not necessarily optimize their design. For *LC* oscillator applications in which noise performance is crucial, the oscillator shown in Figure 5-25, used in the famous Rohde & Schwarz SMDU, is the state of the art. Its noise performance is equivalent to the noise found in the cavity-tuned oscillator used in Hewlett-Packard's popular HP8640 signal generator, and because of the unique way a tuning diode is coupled to the circuit, its modulation capabilities are substantially superior to any of the signal generators currently offered. To develop such a circuit from design equations is not possible. This circuit is the result of many years of experience and research and looks fairly simple. The grounded-gate FET circuit provides the best performance because it fulfills the important requirements of Eq. (5-56).

The tuned circuit is not connected directly to the drain, but the drain is put on a tap of the oscillator section. Therefore, the actual voltage across the tuning capacitor is higher than the supply voltage, and thus the energy stored in the capacitor is much higher than in a circuit connected between the gate electrode and ground, as it is in a normal Colpitts oscillator. In addition, the high output impedance of the FET does not load the circuit, which also provides a reduced noise contribution. Since this oscillator is optimized for best frequency-modulation

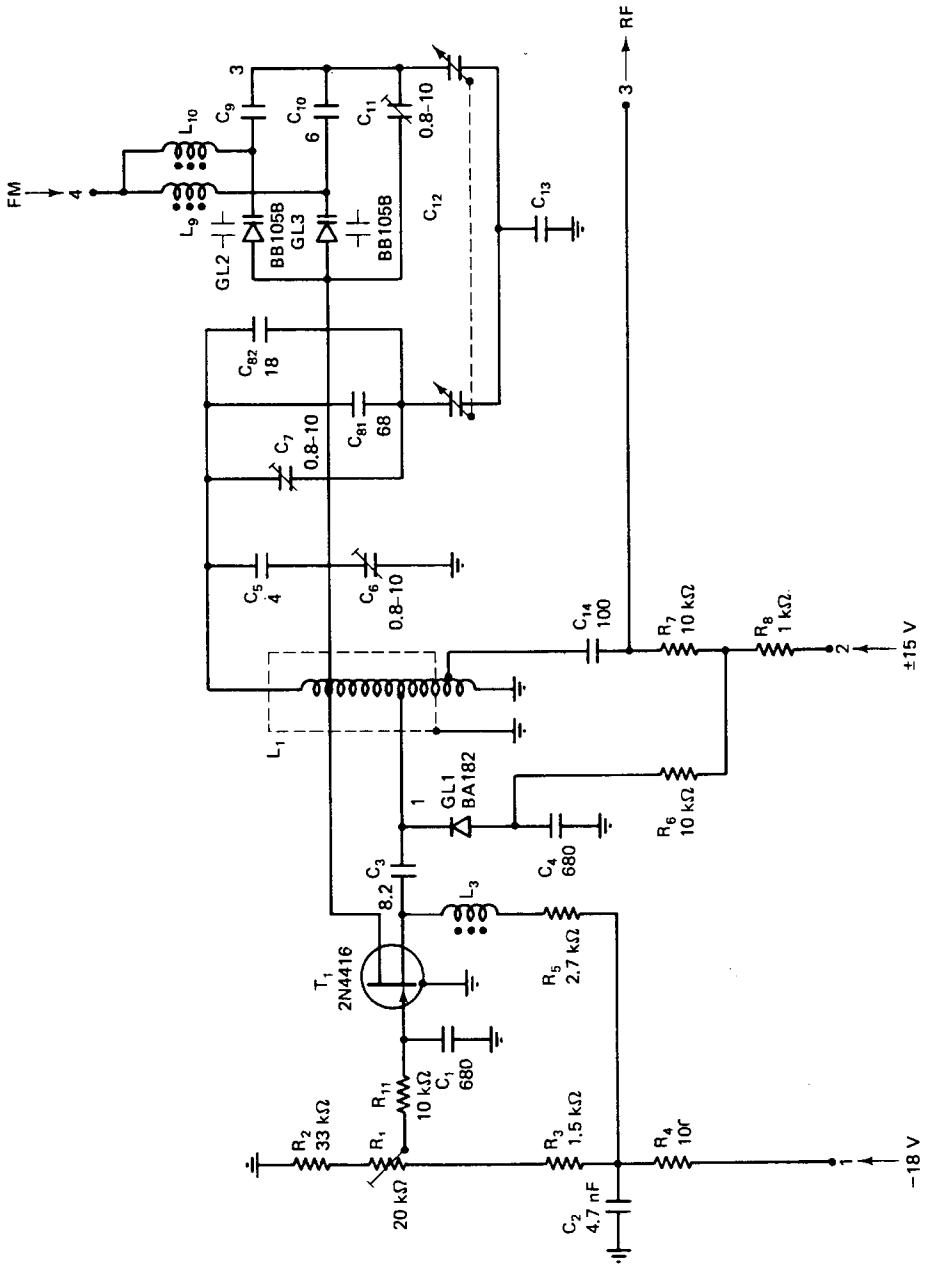


Figure 5-25 A 118–198-MHz oscillator from the Rohde & Schwarz SMDU signal generator.

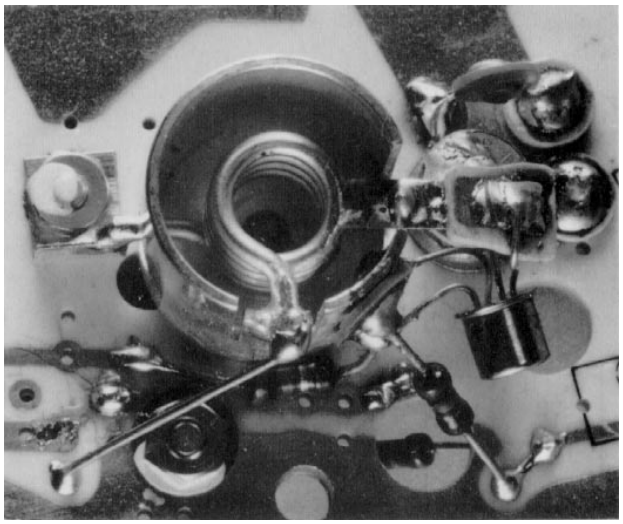


Figure 5-26 Photograph of the helical-resonator system from the Rohde & Schwarz SMDU signal generator.

performance in the FM frequency range, it becomes apparent that it fits the requirements of low-distortion stereo modulation.

For extremely critical blocking measurements in the 2-m region (140–160 MHz), the noise specification at 20 kHz off the carrier is of high importance, while the peak modulation typically does not exceed 5 kHz. Figure 5-25 shows the schematic of the oscillator section optimized for this frequency range.

Because of f_T limitations, it has been found experimentally that these *LC* oscillators, if based on silicon JFETs, should not be used above about 500 MHz; rather, a doubler stage should be employed. Analyzing the signal generators currently on the market, we find that their highest base band typically ranges from 200 to 500 MHz, with frequency doubling used to 1000 MHz. As can be seen in Figure 5-26, the mechanical layout of such an oscillator is extremely compact.

5-5-2 Two-Port Microwave/RF Oscillator Design

A common method for designing oscillators is to resonate the input port with a passive high-*Q* circuit at the desired frequency of resonance. It will be shown that if this is achieved with a load connected on the output port, the transistor is oscillating at both ports and is thus delivering power to the load port. The oscillator may be considered a two-port structure, with M_3 being the lossless resonating port and M_4 providing lossless matching such that all the external RF power is delivered to the load. See Figure 5-27. The resonating network has been described. Nominally, only parasitic resistance is present at the resonating port, since a high-*Q* resonance is desirable for minimizing oscillator noise. It is possible to have loads at both the input and output ports if such an application occurs, since the oscillator is oscillating at both ports simultaneously.

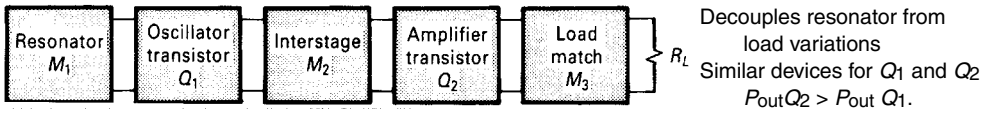


Figure 5-27 Buffered oscillator design.

The simultaneous oscillation condition is proved as follows. Assume that the oscillation condition is satisfied at port 1:

$$1/S'_{11} = \Gamma_G \tag{5-59}$$

Thus,

$$S'_{11} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L} = \frac{S_{11} - D\Gamma_L}{1 - S_{22}\Gamma_L} \tag{5-60}$$

$$\frac{1}{S'_{11}} = \frac{1 - S_{22}\Gamma_L}{S_{11} - D\Gamma_L} = \Gamma_G \tag{5-61}$$

By expanding Eq. (5-61), we find

$$\begin{gathered} \Gamma_G S_{11} - D\Gamma_L\Gamma_G = 1 - S_{22}\Gamma_L \\ \Gamma_L(S_{22} - D\Gamma_G) = 1 - S_{11}\Gamma_G \\ \Gamma_L = \frac{1 - S_{11}\Gamma_G}{S_{22} - D\Gamma_G} \end{gathered} \tag{5-62}$$

Thus,

$$S'_{22} = S_{22} + \frac{S_{12}S_{21}\Gamma_G}{1 - S_{11}\Gamma_G} \tag{5-63}$$

$$\frac{1}{S'_{22}} = \frac{1 - S_{11}\Gamma_G}{S_{22} - D\Gamma_G} \tag{5-64}$$

Comparing Eqs. (5-62) and (5-64), we find

$$1/S'_{22} = \Gamma_L \tag{5-65}$$

which means that the oscillation condition is also satisfied at port 2; this completes the proof. Thus, if either port is oscillating, the other port must be oscillating as well. A load may appear at either or both ports, but normally the load is in Γ_L, the output termination. This result can be generalized to an n-port oscillator by showing that the oscillator is simultaneously oscillating at each port:

$$\Gamma_1 S'_{11} = \Gamma_2 S'_{22} = \Gamma_3 S'_{33} = \cdots = \Gamma_n S'_{nn} \tag{5-66}$$

Before concluding this section on two-port oscillator design, the buffered oscillator shown in Figure 5-27 must be considered. This design approach is used to provide the following:

1. A reduction in load-pulling, which is the change in oscillator frequency when the load reflection coefficient changes.
2. A load impedance that is more suitable to wideband applications ($k < 1$).
3. A higher output power from a working design, although the higher output power can also be achieved by using a larger oscillator transistor.
4. This type of oscillator is not optimized for low phase noise. However, buffered oscillator designs are quite common in wideband YIG applications, where changes in the load impedance must not change the generator frequency.

Two-port oscillator design may be summarized as follows:

1. Select a transistor with sufficient gain and output-power capability for the frequency of operation. This may be based on oscillator datasheets, amplifier performance, or *S*-parameter calculation.
2. Select a topology that gives $k < 1$ at the operating frequency. Add feedback if $k < 1$ has not been achieved.
3. Select an output load-matching circuit that gives $|S'_{11}| > 1$ over the desired frequency range. In the simplest case, this could be a 50-Ω load.
4. Resonate the input port with a lossless termination so that $\Gamma_G S'_{11} = 1$. The value of S'_{22} will be greater than unity with the input properly terminated.

In all cases, the transistor delivers power to a load and the input of the transistor. Practical considerations of realizability and dc biasing will determine the best design.

For both bipolar and FET oscillators, a common topology is the common-base or common-gate, since a common-lead inductance can be used to raise S_{22} to a large value, usually greater than unity even with a 50-Ω generator resistor. However, it is not necessary for the transistor S_{22} to be greater than unity, since the 50-Ω generator is not present in the oscillator design. The requirement for oscillation is $k < 1$; then resonating the input with a lossless termination will ensure that $|S'_{11}| > 1$.

A simple example will clarify the design procedure. A common-base bipolar transistor (HP2001) was selected to design a fixed-tuned oscillator at 2 GHz. The common-base *S*-parameters and stability factors are given in Table 5-1.

Using the load circuit in Figure 5-28, we see that the reflection coefficients are

Table 5-1 *S* parameters and stability factors (V_{CE} = 15 V, I_C = 25 mA)

$L_B = 0$	$L_B = 0.5$ nH
$S_{11} = 0.94 \angle 174°$	$1.04 \angle 173°$
$S_{21} = 1.90 \angle -28°$	$2.00 \angle -30°$
$S_{12} = 0.013 \angle 98°$	$0.043 \angle 153°$
$S_{22} = 1.01 \angle -17°$	$1.05 \angle -18°$
$k = -0.09$	-0.83

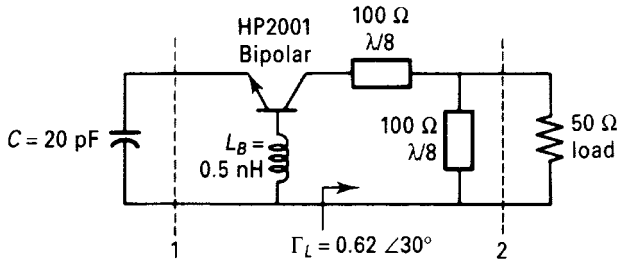


Figure 5-28 Oscillator example at 2 GHz.

$$\Gamma_L = 0.62\ \angle 30°$$

$$S'_{11} = 1.18\ \angle 173°$$

Thus, a resonating capacitance of G = 20 pF resonates the input port. In a YIG-tuned oscillator, this reactive element could be provided by the high-Q YIG element. For a dielectric-resonator oscillator (DRO), the puck would be placed to give $\Gamma_G \approx 1.0/\angle -173°$.

Another two-port design procedure is to resonate the Γ_G port and calculate S'_{22} until $|S'_{22}| > 1$, then design the load port to satisfy. This design procedure is summarized in Figure 5-29.

An example using this procedure at 4 GHz is given in Figure 5-30 using an AT41400 silicon bipolar chip in the common-base configuration with a convenient value of base and emitter

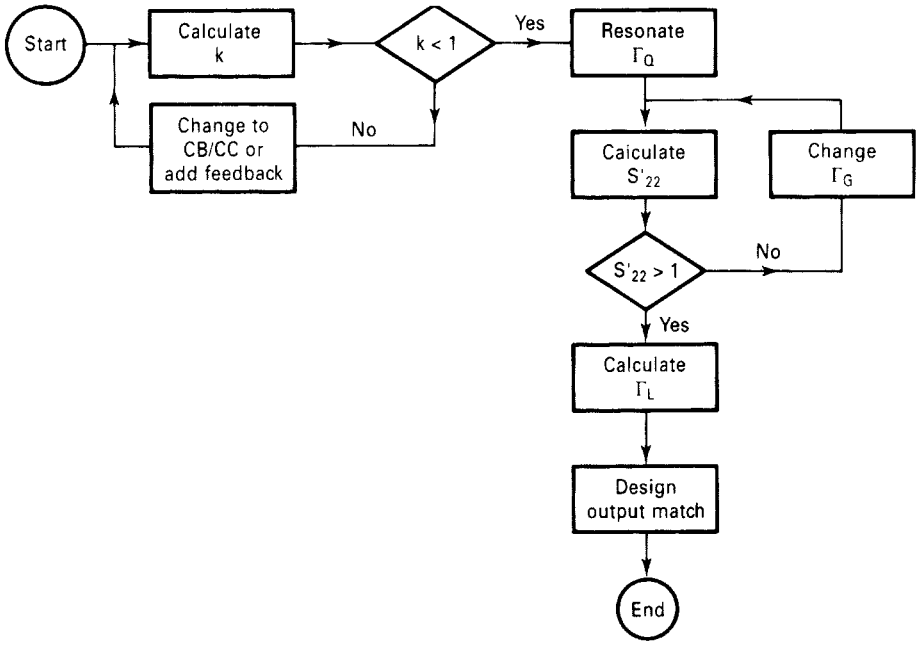


Figure 5-29 Oscillator design flowchart.

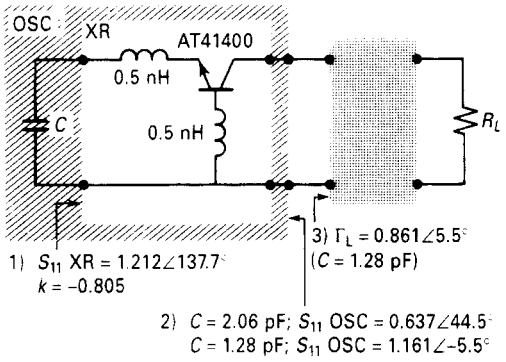


Figure 5-30 A 4-GHz lumped-resonator oscillator using the AT41400.

inductance of 0.5 nH. The feedback parameter is the base inductance, which can be varied if needed.

Since a lossless capacitor at 4 GHz of 2.06 pF gives $\Gamma_G = 1 - 0 \angle -137.7°$, this input termination is used to calculate S'_{22}, giving $S'_{22} = 0.637 \angle 44.5°$. This circuit will oscillate into any passive load. Varying the emitter capacitor about 20° on the Smith Chart to 1.28 pF gives $S'_{22} = 1.16 \angle -5.5°$, which will oscillate into a load of $\Gamma_L = 0.861 \angle 5.5°$. The completed lumped-element design is given in Figure 5-31.

5-5-3 Ceramic-Resonator Oscillators

An important application for a new class of resonators called ceramic resonators (CRs) has emerged for wireless applications. The CRs are similar to rigid coaxial cable, where the center conductor is connected at the end to the outside of the cable. These resonators are generally operating in quarter-wavelength mode and their characteristic impedance is approximately 10 Ω. Because their coaxial assemblies are made from a high-ε low-loss material with good silver plating throughout, the electromagnetic field is internally contained and therefore results in very little radiation. These resonators are therefore ideally suited for high-Q, high-density oscillators. The typical application for this resonator is VCOs ranging from not much more than 200 MHz up to about 3 or 4 GHz. At these high frequencies, the mechanical dimensions of the resonator become too tiny to offer any advantage. One of the

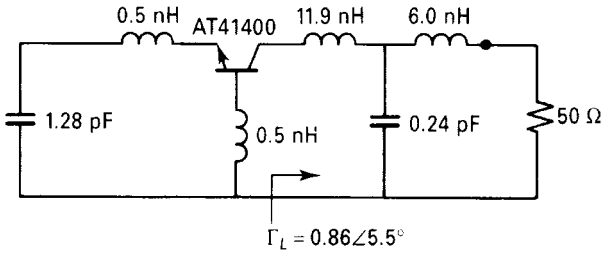


Figure 5-31 Completed lumped-resonator oscillator (LRO).

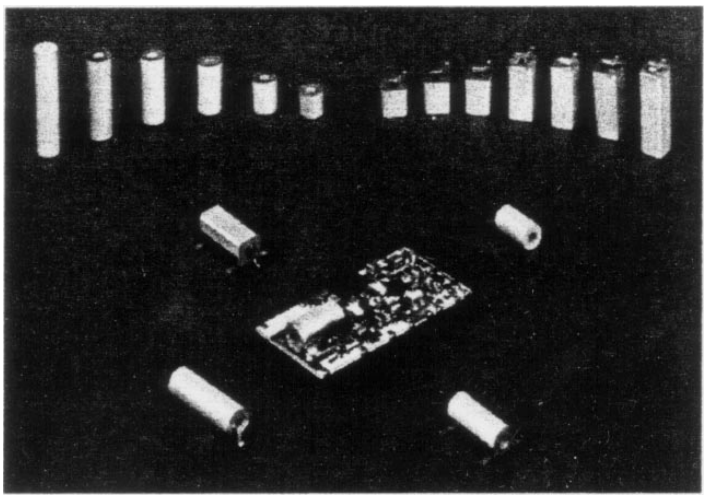


Figure 5-32 Ceramic resonators.

principal requirements is that the physical length is considerably larger than the diameter. If the frequency increases, this requirement can no longer be met.

Manufacturers supply ceramic resonators on a prefabricated basis. Figures 5-32 and 5-33 show the standard round/square packaging available and the typical dimensions for a ceramic resonator. Figure 5-34 shows a ceramic-resonator oscillator built on a ceramic substrate. The available material has a dielectric constant of 88 and is recommended for use in the 400–1500-MHz range. The next higher frequency range (800 MHz to 2.5 GHz) uses an ε of 38, while the top range (1–4.5 GHz) uses an ε of 21. Given the fact that ceramic resonators are prefabricated and have standard outside dimensions, the following quick calculation applies:

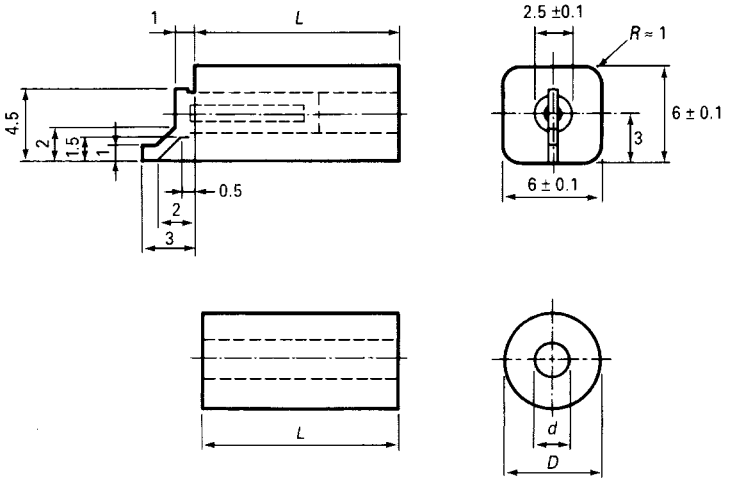


Figure 5-33 Standard round/square ceramic-resonator packaging and dimensions.

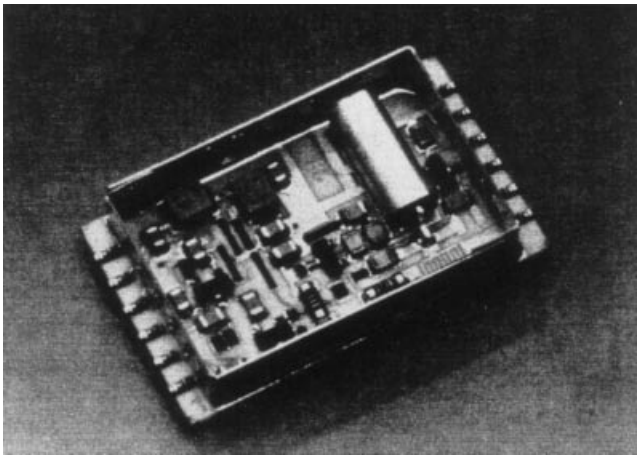


Figure 5-34 Modern ceramic-resonator oscillator for handheld cellular phones.

Relative dielectric constant of resonator material	$\varepsilon_r = 21$	$\varepsilon_r = 38$	$\varepsilon_r = 88$
Resonator length in millimeters	$l = \frac{16.6}{f}$	$l = \frac{12.6}{f}$	$l = \frac{8.2}{f}$
Temperature coefficient (ppm/°C)	10	6.5	8.5
Available temperature coefficients	–3 to +12	–3 to +12	–3 to +12
Typical resonator Q	800	500	400
Frequency range	1–4.5 GHz	800–2500 MHz	400–1500 MHz

Calculation of Equivalent Circuit. The equivalent parallel-resonant circuit afforded by a ceramic resonator has a resistance at resonance of

$$R_p = \frac{2(Z_0)^2}{R^* l} \tag{5-67}$$

where Z_0 = characteristic impedance of the resonator
l = mechanical length of the resonator
R^* = equivalent resistor due to metallization and other losses

As an example, one can calculate

$$C^* = \frac{2\pi\varepsilon_0\varepsilon_r}{\log_e(D/d)} = 55.61 \times 10^{-12}\,\frac{\varepsilon_r}{\log_e(D/d)} \tag{5-68}$$

and

$$L^* = \frac{\mu_r \mu_0}{2\pi} = \log_e\left(\frac{D}{d}\right) = 2 \times 10^{-7} \log_e\left(\frac{D}{d}\right) \tag{5-69}$$

$$Z_0 = 60\ \Omega \frac{1}{\sqrt{\varepsilon_r}} \log_e\left(\frac{D}{d}\right)$$

$$= 60 \frac{1}{\sqrt{88}} \log_e\left(\frac{6}{2.5}\right)$$

$$= 5.6\ \Omega \tag{5-70}$$

A practical example for $\varepsilon_r = 88$ and 450 MHz is

$$C_p = \frac{C^* l}{2} = 49.7\ \text{pF} \tag{5-71}$$

$$L_p = 8L^* l = 2.52\ \text{nH} \tag{5-72}$$

$$R_p = 2.5\ \text{k}\Omega \tag{5-73}$$

Figure 5-35 shows the schematic of a ceramic-resonator-based oscillator. Figures 5-36 and 5-37 show the simulated and measured phase noise of the oscillator.

By using ceramic-resonator-based oscillators in conjunction with miniature synthesizer chips, it is possible to build extremely small phase-locked-loop (PLL) systems for cellular telephone operation. Figure 5-38 shows one of the smallest currently available PLL-based synthesizers manufactured by Synergy Microwave Corporation. Because of the high-Q resonator, these types of oscillators exhibit extremely low phase noise. Values of better than –150 dBc/Hz at 1 MHz off the carrier are achievable. The ceramic resonator reduces the oscillator's sensitivity to microphonic effects and proximity effects caused by other components.

5-5-4 Using a Microstrip Inductor as the Oscillator Resonator*

A high-Q microstrip resonator [3–5] can be used to improve an oscillator's phase-noise performance. The following section, based on information in the Philips Semiconductors Application Note AN1777, describes the application of such a resonator in a BJT differential oscillator.

It can be seen from Eq. (5-56) that the phase noise is proportional to the term

$$\frac{1}{4Q_L^2} \tag{5-74}$$

This means that for any 10% increase in Q, we get something like 20% improvement in phase noise, which makes it most desirable to increase the Q. One way to reduce the loading of the

*Based on portions of the Philips Semiconductors/Signetics RF Communications Products Application Note AN1777, "Low-Voltage Front End Circuits: SA601, SA620," August 20, 1997. Used with permission.

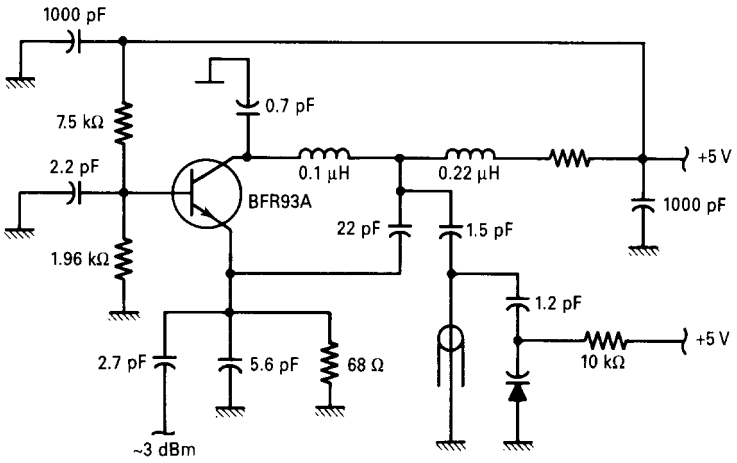


Figure 5-35 Schematic of a high-performance ceramic-resonator-based oscillator that can be used from 500 MHz to 2 GHz.

resonator is to tap the device down on the resonator. The underlying oscillator circuit is based on the differential type of oscillator found in ICs.

Increasing Loaded Q. Figure 5-39 shows the oscillator section of the Philips SA620 low-voltage front-end IC with a conventional second-order parallel-tuned tank circuit configured as the external resonator.

From the basic equations for parallel resonance, the resonator's loaded Q is given by

$$Q_L = R_P \left(\frac{C_T}{L_T} \right)^{1/2} \tag{5-75}$$

where

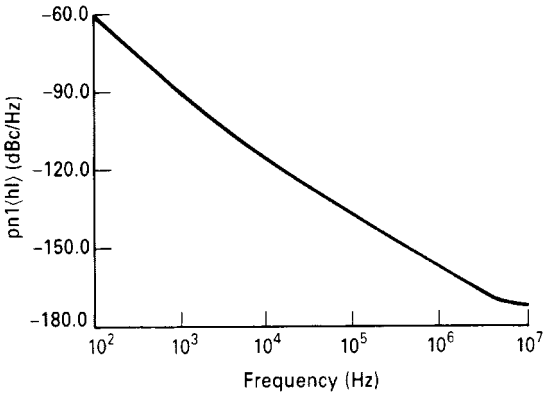


Figure 5-36 Simulated phase noise of an *npn* bipolar 1-GHz ceramic-resonator-based oscillator.

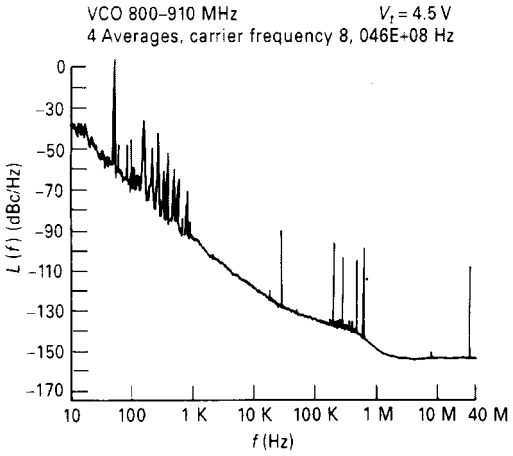


Figure 5-37 Measured phase noise of a ceramic-resonator-based oscillator.

$$R_P = R_T \| R_C = \left[\left(\frac{1}{R_T}\right) + \left(\frac{1}{R_C}\right)\right]^{-1} \tag{5-76}$$

represents the net shunt tank resistance appearing across the network at resonance. R_T represents losses in the inductance L_T and capacitance C_T (almost always dominated by inductor losses), and R_C is the active circuit's load impedance at resonance. Improving the quality of the tank components L_T and C_T (i.e., improving their Q) will increase R_T, but since R_C is low in the case of the SA620, the resulting increase in R_P is relatively small. We can

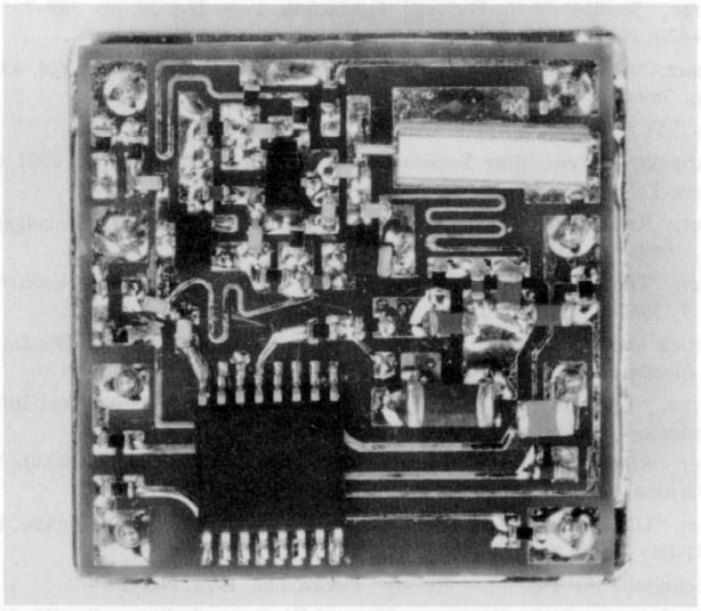


Figure 5-38 Miniature PLL-based synthesizer manufactured by Synergy Microwave Corporation.

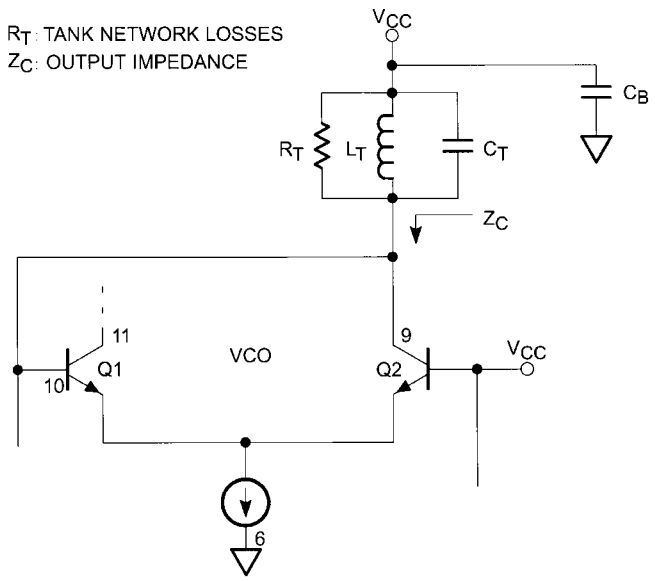


Figure 5-39 Oscillator portion of Philips SA620 front-end IC with external resonator.

increase R_P by decoupling Z_C from the tank circuit (note that R_C is the real part of Z_C at resonance). Decoupling this impedance by using either tapped-L or tapped-C tank configurations is possible. Inspection of the circuit shows that dc biasing is necessary for pins 9 and 10 (osc1 and osc2, respectively), so a tapped-C approach would require shunt-feeding these pins to V_{CC}. Thus, the most practical way is to employ a tapped-L network.

Figure 5-40 shows the basic tapped-L circuit. The effective multiplied shunt impedance at resonance across C_T is given by

$$\hat{R}_P = R_P\left[\left(\frac{L_{T2}}{L_{T1}}\right)+1\right]^{1/2} \tag{5-77}$$

with a corresponding increase in loaded Q:

$$Q_L = Q_L\left[\left(\frac{L_{T2}}{L_{T1}}\right)+1\right]^{1/2} \tag{5-78}$$

Feedback is caused by an RF voltage appearing across L_T, coupled through bypass capacitor C_B and its ground return to the cold end of the current source at pin 6. The significance of this path increases with frequency. At 900 MHz, it becomes very critical in obtaining stable oscillations and must be kept as short as possible. Attention to layout detail is pivotal here!

High-Q Microstrip Inductor. Realizing a stable tapped-L network using lumped surface-mount inductors is impractical due to their low unloaded Q and finite physical size. Another approach uses a high-Q short microstrip inductor. The conventional approach to microstrip resonators treats them as a length of transmission line terminated with a short that

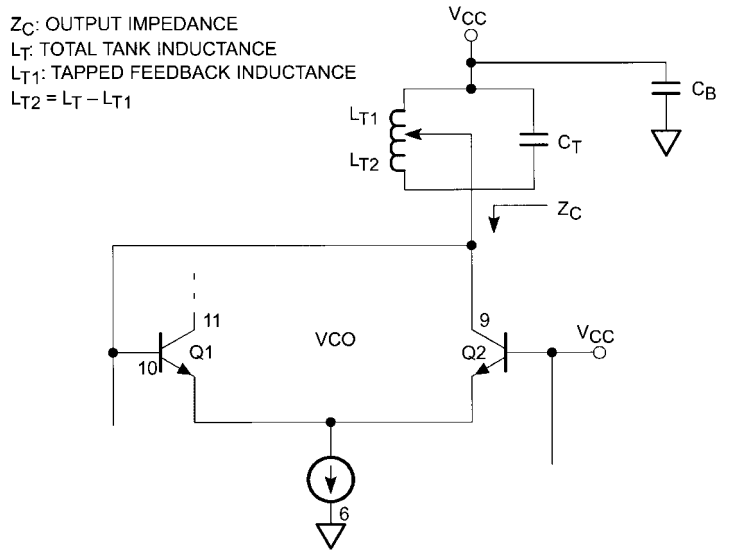


Figure 5-40 Oscillator portion of SA620 IC with tapped-microstrip resonator.

reflects back an inductance impedance that simulates a lumped inductor. The approach presented here deviates from this technique by specifically exploiting the very high unloaded Q attainable with short microstrip inductors, where we design for a large C/L ratio by making the microstrip a specific, but short, length. Resonance is achieved by replacing the short with whatever capacitance is needed for proper resonance. One equation for the inductance of a strip of metal having a length ℓ (mils) and width w (mils) is

$$L = [5.08\text{E} - 3]\,\ell\left[\ln\left(\frac{\ell}{w}\right) + 1.193\left(\frac{w}{\ell}\right)\right] \qquad \text{[3] (5-79)}$$

where L is in nH/mil. This technique results in a much shorter strip of metal for a given inductance. One intriguing property of microstrip inductors is that they are capable of very high unloaded Q values. An equation describing the quality factor for a "wide" microstrip is

$$Q_C = 0.63h[\sigma f_{\text{GHz}}]^{1/2} \qquad \text{[4] (5-80)}$$

where h is the dielectric thickness in centimeters, σ is the conductivity in S/m, and f is frequency in GHz. This equation predicts an unloaded Q exceeding 700 when silver-coated copper is used. Also, wide microstrip lines are defined as those whose strip width w to height h ratio is approximately greater than 1; that is, $w/h > 1$. The parallel capacitor formed by the metal strip over the ground plane should also be included and may be calculated by the classic formula and included in the total needed to resonate with the inductance result found from Eq. (5-79). This "strip" capacitance is given by

$$C_S = K\varepsilon_0\varepsilon_r\left(\frac{w\ell}{h}\right) \qquad \text{(5-81)}$$

where C_S is in farads, $\varepsilon_0 = 8.86\text{E} -12$ (F/cm) is the permittivity of free space, and ε_r is the relative dielectric constant of the substrate. A "fringe factor," K, is included to account for necessary fringing (estimated to be 2–15%). These equations are meant to provide insight into circuit behavior and should, therefore, be applied cautiously to specific applications.

UHF VCO Using the Tapped-Inductor Differential Oscillator at 900 MHz. The basic electrical circuit of this configuration is shown in Figure 5-41. Feedback occurs when sufficient RF voltage develops across the tap inductances L_{T1} and L_{T2} (due to tap 1 and tap 2, respectively); note that inductance is reckoned from the cold end of the tank at node A. Taps 1 and 2 should be as close as possible to pins 9 and 10, respectively. Also node A (cold end of tank) and node B (pin 6) must be as physically close together and exhibit as low an impedance as possible to discourage parasitic oscillations. This "inner loop," composed of L_{T2}, the inductance of connections between taps 1 and 2 and pins 9 and 10, respectively, and "stray" low-Q inductance between nodes A and B creates a parasitic loop that will support oscillation that no longer depends on the full tapped-L tank circuit. Oscillation based on this low-Q loop typically occurs well above 1.2 GHz and has been observed as high as 1.6 GHz, and exhibits very poor phase noise performance.

Another factor also affects this unwanted parasitic oscillation. As the full tank circuit's unloaded Q decreases, the circuit becomes conditionally stable and will eventually favor the parasitic loop exclusively. This occurs because the loaded Q of the tapped microstrip affects the magnitude of the RF voltage appearing across the entire tank. Thus, feedback at taps 1 and 2 is reduced as the Q decreases. This increases the likelihood of the parasitic inner loop controlling the oscillator, since its feedback voltage is largely independent of the microstrip tank Q.

Equation (5-80) shows that microstrip inductors are capable of very high unloaded Q values. This depends on a number of physical factors: frequency, dielectric thickness and

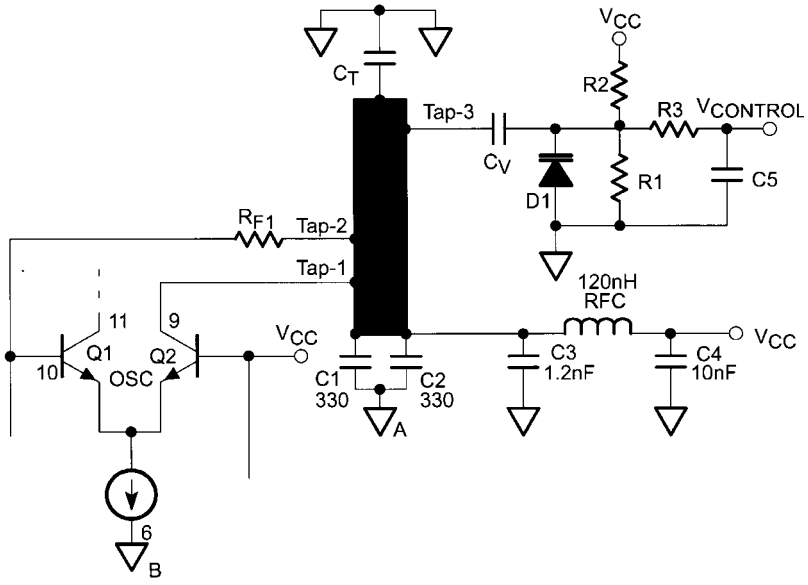


Figure 5-41 Differential oscillator with tapped-microstrip resonator.

quality, and the skin-depth conductance of the metal strip itself. For example, at 900 MHz, the unloaded Q of a silver-coated microstrip line is in the vicinity of 750. Coating the same line with lead (sloppy soldering can do it!) reduces the Q to less than 230. If the unloaded Q is too low, the loaded Q also drops. The circuit becomes conditionally stable and now may oscillate either at the higher parasitic-loop frequency or the desired lower frequency at which the entire microstrip tank resonates. Components should be connected by narrow traces closely connected to component leads to avoid soldering on the microstrip proper.

R_{F1} introduces necessary losses to control the inner-loop parasitic oscillation. Its size should be made as small as possible consistent with stability, startup, and good phase-noise performance. Since R_{F1} also decreases feedback, the oscillator's output falls as R_{F1} increases. Typically, experimental values from 4 to 30 Ω have proved sufficient.

Stability as a function of V_{CC} is a good way to assess conditional stability. Provided the microstrip tank has a high loaded Q, the oscillator's stability should be independent of V_{CC} down to less than 2.5 V or so. When loaded Q decreases sufficiently to favor the inner-loop parasitic, conditional stability will occur. This can readily be observed by increasing V_{CC} from about 1.0 V and noting whether the VCO frequency jumps back and forth unpredictably. Mixer and LO-port loading can also affect this condition and should be considered.

Dimensions for the microstrip resonator are based on several criteria. The strip must be "wide": Its width-to-height ratio must be on the order of 1 or greater. Minimum strip length seems to be about 200 mils for 62-mil-thick board. L/C ratios somewhere about 500 have yielded quite good results. Note that we usually specify the "L/C ratio" because it is always greater than 1, even though in a parallel-resonant circuit it appears in the Q equation as C/L. Thus, decreasing the "L/C ratio" by making the microstrip shorter helps the inductor's loaded Q, and the circuit's loaded Q_L, increase [see Eq. (5-75)]. However, the effect of the lower Q of C_T, even when using a high-Q surface-mount type, decreased the expected larger increase in the circuit's net loaded Q. Thus, the expected increase in Q_L may not be fully realizable. Assuming a 62-mil-thick board with FR5 dielectric, a strip 50 mils by 300 mils gave very good experimental results, where C_T was about 4.3 pF, with oscillation occurring from about 950 to 1000 MHz with various test boards. A shorter inductor designed to increase the C/L ratio requiring 7.5 pF experimentally resulted in only about 1–2-dB phase-noise improvement at 950 MHz. The reason for this is that as the microstrip inductor degenerates to a plain inductor, all of its advantages are lost.

Tap 1 may be anywhere between the cold end of the tank (where C_1 and C_2 are located) and tap 2. Tap 2 yields good results at about one-third the strip length. Making it too close to the cold end in an effort to increase Q_L will result in loss of control over the inner-loop parasitic. To keep Q_L as high as possible, C_T should be a high-Q surface-mount type. Differences greater than 5 dB in SSB phase noise have been observed between a generic NP0 surface-mount capacitor and a high-Q NP0 surface-mount capacitor.

Tap 3 can be at the end of the microstrip when C_T is connected. However, the tuning diode inevitably will cause a decrease in overall loaded Q, typically resulting in a phase-noise increase of as much as 5 dB. Moving it some distance down from the high end of the tank will decrease this effect and yield better results. A good starting point is about one-quarter down, or three-quarters of the length from the cold end of the tank. C_1 and C_2 are paralleled lower-value capacitors to yield a better low-impedance ground return to pin 6 (node B) since relatively large RF currents flow through them. Note that as Q_L increases, the peak circulating RF current will also increase, as will the RF voltage at the hot end of the microstrip. At 900 MHz, good results have been obtained when both are about 300 pF. The RF choke was chosen to be approximately series-resonant around 900 MHz, and constitutes the series feed path

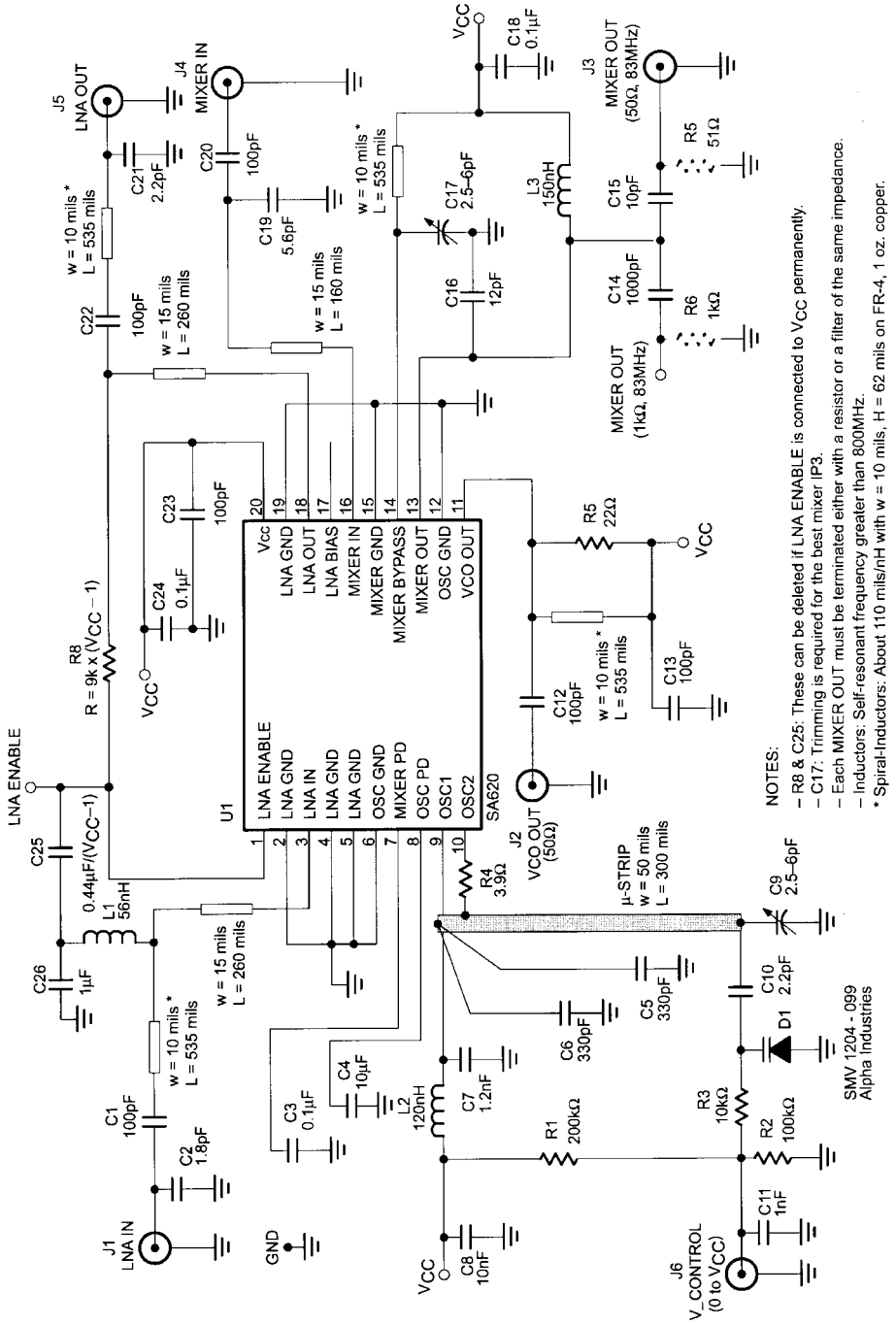


Figure 5-42 Philips SA620 application board schematic showing the tapped-microstrip resonator.

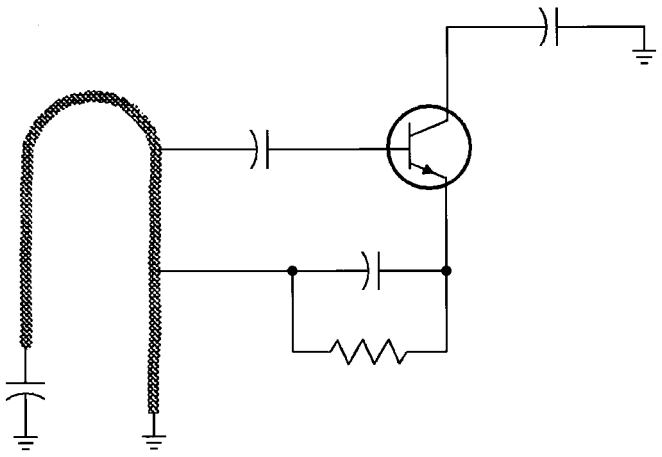


Figure 5-43 "Walking stick" microstrip Hartley resonator as used by Plextek in the 800–1500-MHz range. Below 800 MHz, the resonator's physical size becomes impractical; above 1500 MHz, the transistor's parasitic reactances make the Hartley configuration problematic.

for dc biasing. It may be possible to neglect this component entirely, provided that the PC-board connection to the cold end of the tank is very close to RF ground. Finally, note that shielding of the entire microstrip may be necessary to meet FCC Part 15 emission limitations (where applicable). Figure 5-42 shows one application of the tapped-microstrip oscillator with the Philips SA620 front-end IC.

5-5-5 Hartley Microstrip Resonator Oscillator

G. Smithson and S. Fitz of Plextek Communications Technology Consultants have reported on the construction of Hartley microstrip-resonator oscillators using thick-film techniques on ceramic (alumina) substrates for high-volume wireless products. The resonator in one such oscillator (Figure 5-43) exhibited an unloaded Q of approximately 80. Table 5-2 summarizes its phase noise versus tuning range with different tuning-diode configurations.

5-5-6 Crystal Oscillators

Most wireless applications do not require ultra-high-frequency stability that mandates the use of a primary standard, such as one based on cesium, or a secondary standard, such as one based on rubidium. For very stable oscillators at fixed or only slightly variable frequencies up to a few hundred megahertz, oscillators with piezoelectric resonators are generally used. A piezoelectric material transduces mechanical stress (deformation) to

Table 5-2 Achieved oscillator phase noise versus tuning range

Tuning Range, 0.5–3 V	Tuning Range, 1–3 V	Phase Noise at 20-kHz Offset
23%	16%	–100 dBc/Hz
16%	11%	–103 dBc/Hz
10%	8%	–106 dBc/Hz

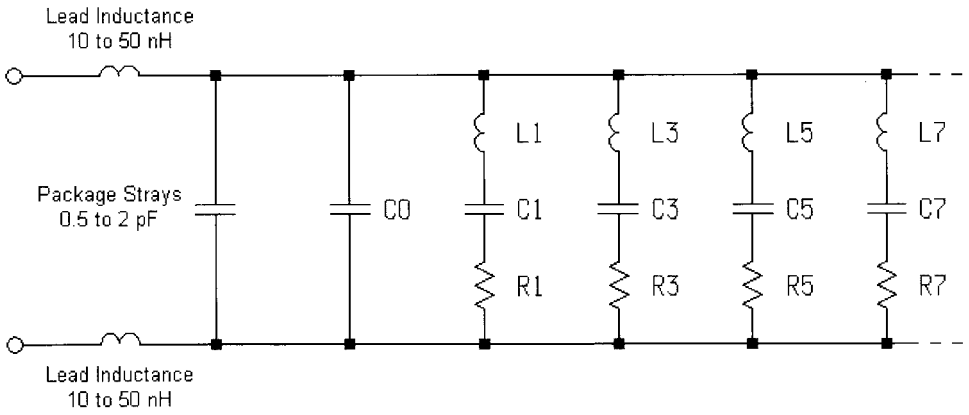


Figure 5-44 Electrical equivalent of a piezoelectric crystal resonator, showing fundamental (1) and overtone (3, 5, 7, . . .) resonances. (Courtesy of Roger Clark, Vectron International.)

Table 5-3 Approximate crystal parameters

Crystal Q Approximations ($F_{\text{resonance}}$ in MHz)

Nonprecision Fundamental Crystals:

Worst: $Q = \dfrac{2 \times 10^6}{F}$ Better: $Q = \dfrac{5 \times 10^6}{F}$

Precision Crystals Processed for High Q:

SC: $Q = \dfrac{12 \times 10^6}{F}$ AT Cut: $Q = \dfrac{10 \times 10^6}{F}$

Crystal Resistance (r) for 10-MHz, Third-Overtone Example Using Precision Crystal Approximation (F in MHz and C_1 in pF)

$$r = \frac{2\pi F C_1}{Q}$$

$$r = 6.964$$

Typical Crystal Parameters

10-MHz Fundamental: $C_1 = 0.02$ pF, $C_0 = 3.6$ pF, $Q = 350{,}000$, $r = 2.5\ \Omega$, $L = 0.012665$ H
50-MHz Third-Overtone: $C_1 = 0.0026$ pF, $C_0 = 4.212$ pF, $Q = 100{,}000$, $r = 13\ \Omega$, $L = 0.003897$ H
100-MHz Fifth-Overtone: $C_1 = 0.0005$ pF, $C_0 = 2.25$ pF, $Q = 80{,}000$, $r = 40\ \Omega$, $L = 0.005066$ H
155-MHz High-Frequency Fundamental: $C_1 = 0.005$ pF, $C_0 = 1.5$ pF, $Q = 20{,}000$, $r = 12\ \Omega$, $L = 0.002108$ H

Source: Roger Clark, Vectron International.

electrical stress (voltage), and vice versa. In noncritical applications up to 10 MHz, a piezoelectric ceramic may be used. The Q and temperature stability of such materials are relatively poor, however, so quartz is the resonator material of choice for temperature-stable, low-noise piezoelectric oscillators.

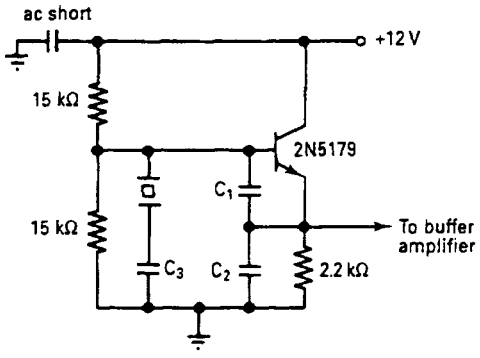
Just as an electrical resonator consisting of distributed L and C exhibits multiple electrical resonances, a physical piece of piezoelectric material, depending on its shape, exhibits multiple mechanical resonances. *Fundamental* resonance occurs at the frequency at which the crystal is one-half wavelength long; *overtone* resonances occur near odd multiples (3, 5, 7, etc.) of the fundamental. Figure 5-44 shows the electrical equivalent of a crystal. Table 5-3 presents examples of approximate parameters for several crystal types. Appropriate shaping and placement of the crystal's connecting electrodes, combined with suitable oscillator circuit design and adjustment, allow the selection or ehancement of a particular resonance for excitation.

Figure 5-45 shows a typical crystal oscillator circuit configuration, and the associated table supplies recommended values for the elements. This is in line with the chip manufacturer's recommendations for built-in oscillator circuits. Figure 5-46 shows how this oscillator's input impedance varies around resonance; Figure 5-47 shows its simulated phase noise. Placing a small resistance in series with the crystal (Figure 5-48) allows us to use the crystal as a high-Q filter, substantially improving the oscillator's spectral purity relative to that obtainable at the transistor emitter (Figure 5-49).

Some applications demand ultra-low-phase-noise crystal oscillators. Figures 5-50 and 5-51 show an appropriate circuit and its phase noise with and without the current-supplying hot carrier diode (HCD) used for noise reduction.

5-5-7 Voltage-Controlled Oscillators

A schematic of a ceramic-resonator-based VCO (Figure 5-35) has already been shown in the section on ceramic-resonator oscillators. To tune the oscillator within the required range,



	3 MHz	6 MHz	10 MHz	20 MHz	30 MHz
C_1 (pF)	330	270	180	82	43
C_2 (pF)	430	360	220	120	68
C_3 (pF)	39	43	43	36	32
Crystal C_L (pF)	32	32	30	20	15

C_3 is the approximate value needed to set oscillator on frequency using indicated crystal C_L.

Figure 5-45 A 3–30-MHz Colpitts crystal oscillator.

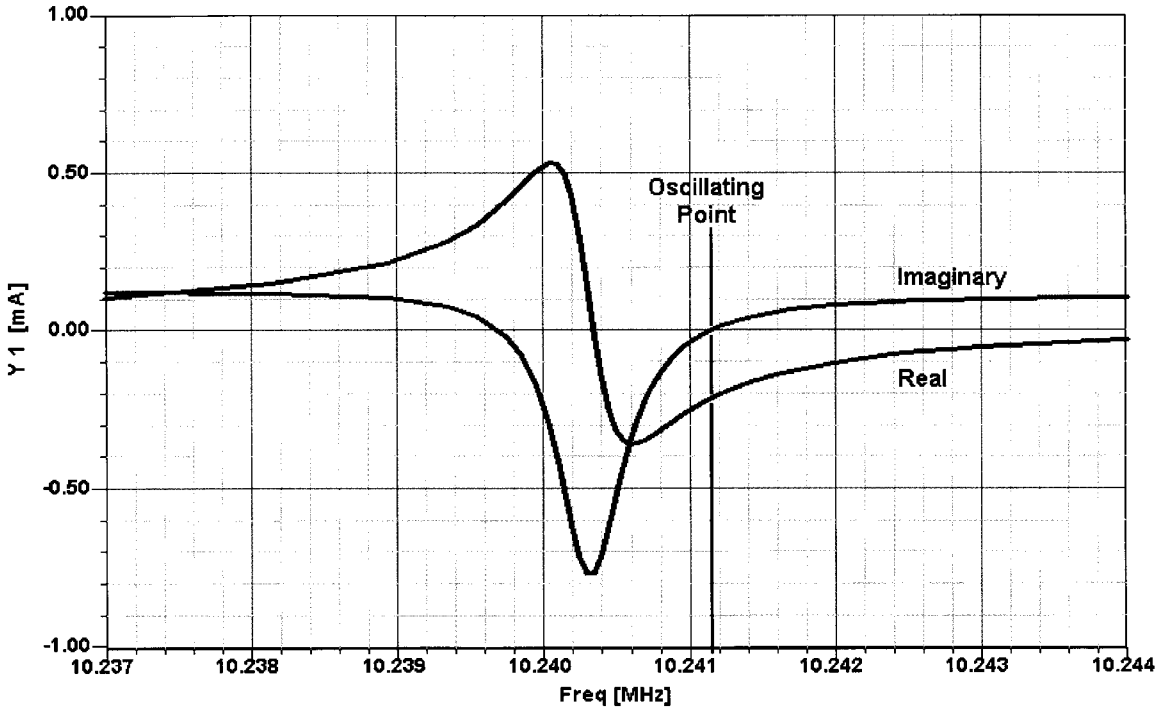


Figure 5-46 As with every oscillator, the crystal oscillator's input impedance is negative and real at resonance. This graph plots the real and imaginary components of the test voltage injected by the Oscillator Design Aid tool in Ansoft Serenade 8.0. Compare this graph with Figure 5-7, which plots an oscillator's input impedance with the resonator absent.

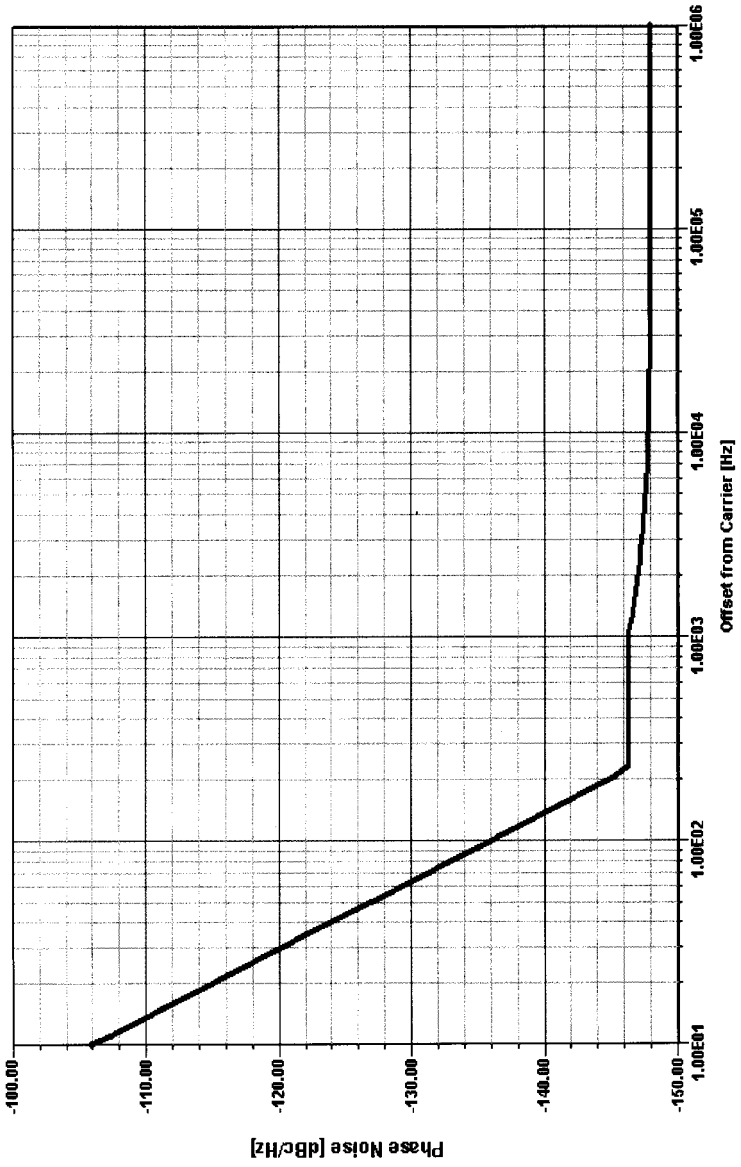


Figure 5-47 Simulated phase noise of the crystal oscillator.

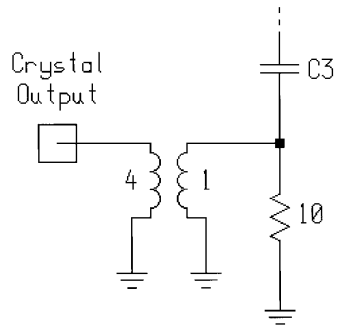


Figure 5-48 This alternative output tap uses the crystal's high selectivity for improved harmonic reduction.

tuning diodes are used. These diodes are often called *varactors* or *voltage-sensitive diodes.* By way of approximation, we can use the equation

$$C = \frac{K}{(V_R + V_D)^n} \tag{5-82}$$

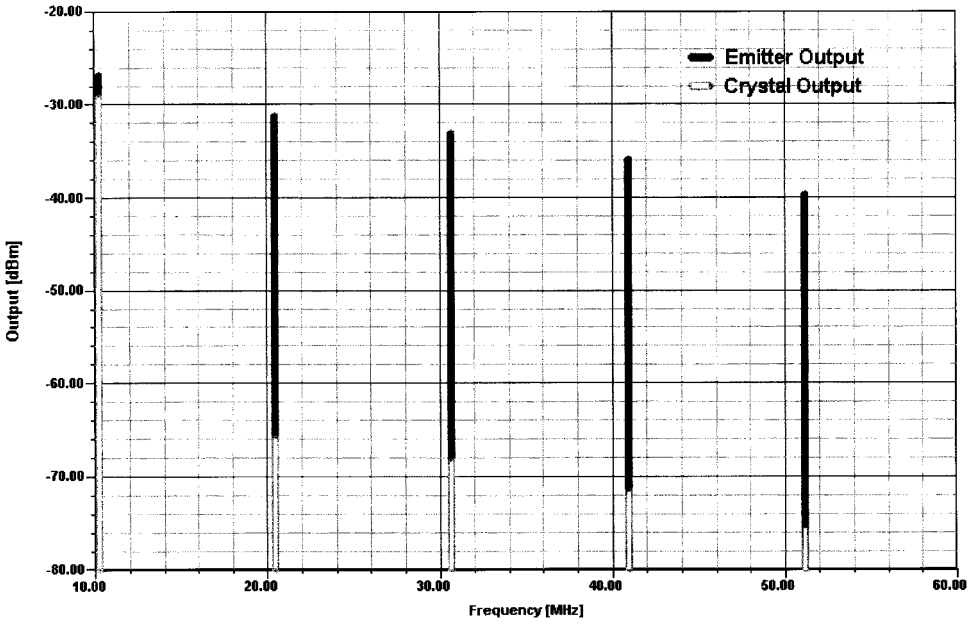


Figure 5-49 Comparison of the crystal oscillator's simulated output spectrum as obtained at the transistor emitter (dark gray, fundamental –26.7 dBm) and crystal (light gray, fundamental –29.1 dBm). At the emitter, the second harmonic is down only 4.3 dB relative to the carrier; at the crystal, the second harmonic is down 36 dB relative to the carrier.

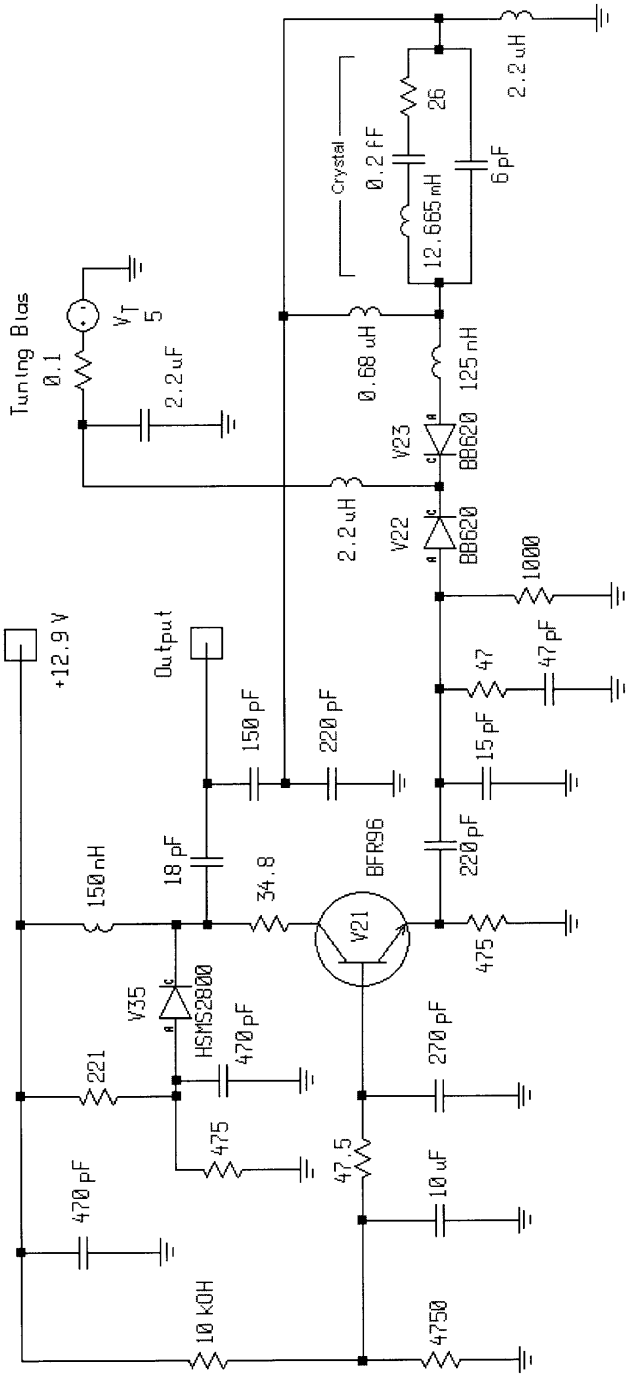


Figure 5-50 A 100-MHz VCXO suitable for ultra-low-phase-noise applications. The HSMS2800 hot-carrier diode (HCD) provides noise reduction.

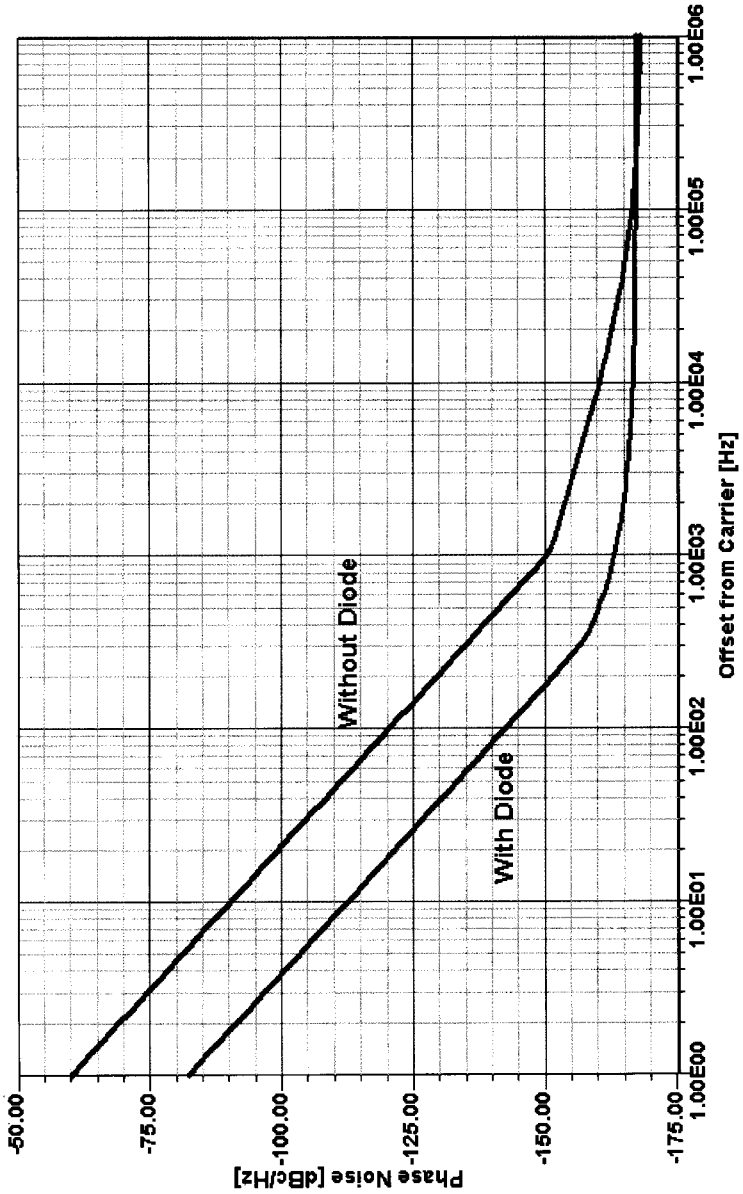


Figure 5-51 Phase noise of the VCXO with and without the noise-reduction diode.

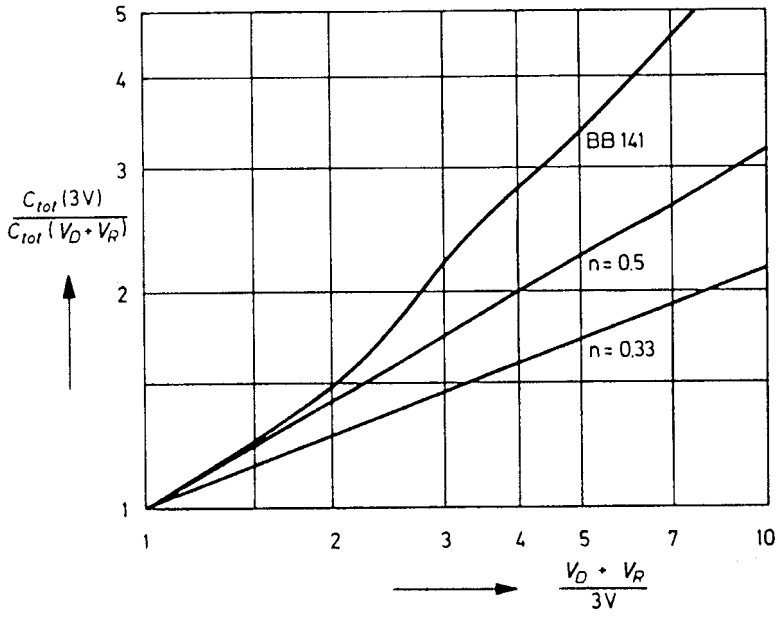


Figure 5-52 Capacitance–voltage characteristic for an alloyed capacitance diode ($n = 0.33$), a diffused capacitance diode ($n = 0.5$), and a wide-range tuning diode (BB141).

wherein all constants and all parameters determined by the manufacturing process are contained in K. The exponent is a measure of the slope of the capacitance/voltage characteristic and is 0.5 for alloyed diodes, 0.33 for single-diffused diodes, and (on average) 0.75 for tuner diodes with a hyperabrupt *pn* junction [6, 7]. Figure 5-52 shows the capacitance–voltage characteristics of an alloyed, a diffused, and a tuner diode.

Recently, an equation was developed that, although purely formal, describes the practical characteristic better than Eq. (5-82):

$$C = C_0 \left(\frac{A}{A + V_R} \right)^m \tag{5-83}$$

where C_0 is the capacitance at $V_R = 0$, and A is a constant whose dimension is the volt. The exponent m is much less dependent on voltage than the exponent n in Eq. (5-81).

The operating range of a capacitance diode or its useful capacitance ratio,

$$\frac{C_{\max}}{C_{\min}} = \frac{C_{\text{tot}}(V_{R,\min})}{C_{\text{tot}}(V_{R,\max})} \tag{5-84}$$

is limited by the fact that the diode must not be driven by the alternating voltage superimposed on the tuning voltage, either into the forward mode or the breakdown mode. Otherwise, rectification would take place, shifting the bias of the diode and considerably affecting its figure of merit.

There are several manufacturers of tuning diodes. Motorola is a typical supplier in this country; Siemens of Germany and Philips also provide good diodes. Table 5-4 and Figure 5-53 contain information for tuning diodes useful for our applications.

5-5-8 Diode-Tuned Resonant Circuits

Tuner Diode in Parallel-Resonant Circuit. Figures 5-54 and 5-55 illustrate two basic circuits for the tuning of parallel resonant circuits by means of capacitance diodes. In the circuit diagram of Figure 5-54, the tuning voltage is applied to the tuning diode via the bias resistor, R_B. Series-connected from the tuning diode to the top of the tank circuit is capacitor C_S, which completes the circuit for alternating current but isolates the cathode of the tuner diode from the coil and thus from the negative terminal of the tuning voltage. Moreover, a fixed parallel capacitance C_P is provided. The decoupling resistor preceding the bias resistor is large enough to be disregarded in the following discussion. Since for high-frequency purposes the biasing resistor is connected in parallel with the series capacitor, it is transformed into the circuit as an additional equivalent shunt resistance R_c. For the parallel loss resistance transformed into the circuit, we have the expression

$$R_c = R_B \left(1 + \frac{C_{\text{tot}}}{C_S}\right)^2 \tag{5-85}$$

Table 5-4 Example tuning diodes (Siemens)

	Maximum Ratings		Characteristics (T_A = 25 °C)								
Type	V_R (V)	I_F (mA)	C_T at (pF)	V_R (V)	C_T (pF)	V_R (V)	C_{ratio}	I_R (nA)	V_R (V)	Package	Chip Code
BB439	30	20	29	3	5.0	25	5.8	≤20	28	SOD-323	Q
BB535	30	20	18.7	1	2.1	28	8.9	≤10	30	SOD-323	A
BB545	30	20	20	1	2.0	28	10.0	≤10	30	SOD-323	C
BB639	30	20	38.3	1	2.65	28	14.5	≤10	30	SOD-323	K
BB639C	30	20	39	1	2.55	28	15.3	≤10	30	SOD-323	B
BB640	30	20	69	1	3.05	28	22.6	≤10	30	SOD-323	N
BB804 (dual)	20	50	44.75	2	26.15	8	1.71	≤20	16	SOT-23	E
BB814 (dual)	20	50	44.75	2	20.8	8	2.15	≤20	16	SOT-23	O
BB833	30	20	9.3	1	0.75	28	12.4	≤20	30	SOD-323	F
BB835	30	20	9.1	1	0.62	28	14.7	≤10	30	SOD-323	G
BB914 (dual)	20	50	43.75	2	18.7	8	2.34	≤20	16	SOT-23	P
BBY51 (dual)	7	20	5.30	1	3.10	4	1.75	≤10	6	SOT-23	H
BBY51-03W	7	20	5.30	1	3.10	4	1.75	≤10	6	SOD-323	H
BBY51-07 (dual)	7	20	5.30	1	3.10	4	1.75	≤10	6	SOT-143	H
BBY52 (dual)	7	20	1.75	1	1.25	4	1.40	≤10	6	SOT-23	I
BBY52-03W	7	20	1.75	1	1.25	4	1.40	≤10	6	SOD-323	I
BBY53 (dual)	6	20	5.30	1	2.40	3	2.20	≤10	4	SOT-23	L
BBY53-03W	6	20	5.30	1	2.40	3	2.20	≤10	4	SOD-323	L

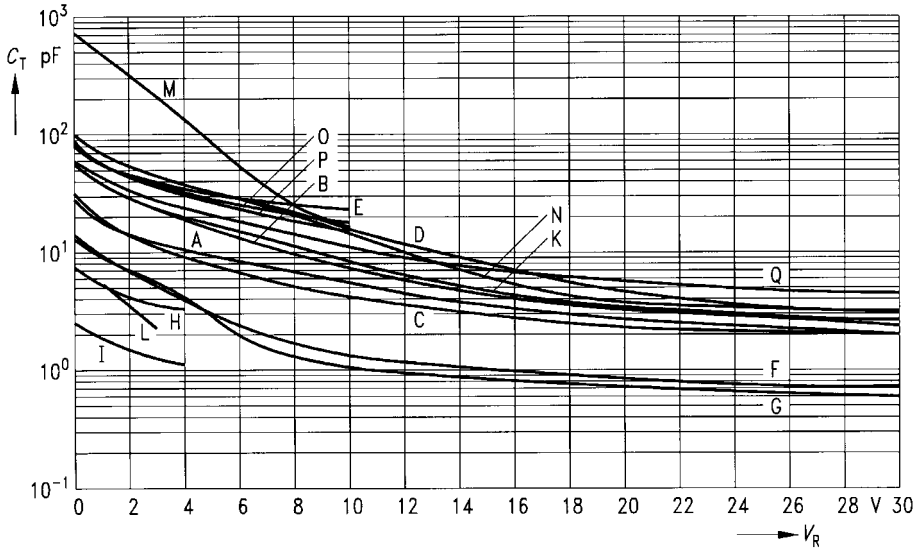


Figure 5-53 Diode capacitance C_T versus reverse voltage V_R for the diodes shown in Table 5-4. The curve labels correspond to the chip codes called out in the table.

If in this equation the diode capacitance is replaced by the resonator circuit frequency ω, we obtain

$$R_c = R_B \left(\frac{\omega^2 L C_S}{\omega^2 L (C_S + C_P) - 1} \right)^2 \tag{5-86}$$

In the circuit of Figure 5-55, the resonant circuit is tuned by two tuning diodes, which are connected in parallel via the coil for tuning purposes, but series-connected in opposition for high-frequency signals. This arrangement has the advantage that the capacitance shift caused by the ac modulation acts in opposite directions in these diodes and, therefore, cancels itself. The bias resistor R_B, which applies the tuning voltage to the tuning diodes, is transformed into the circuit at a constant ratio across the entire tuning range. Given two identical, loss-free tuning diodes, we obtain the expression

$$R_c = 4R_B \tag{5-87}$$

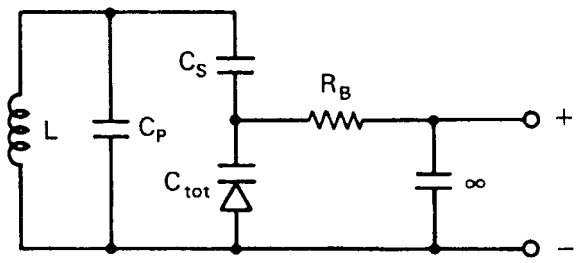


Figure 5-54 Parallel-resonant circuit with tuner diode, and bias resistor in parallel to the diode.

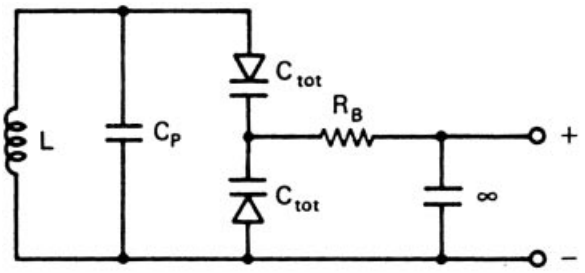


Figure 5-55 Parallel-resonant circuit with two tuning diodes.

Capacitances Connected in Parallel or in Series with the Tuning Diode. Figure 5-54 shows that a capacitor is usually in series with the tuner diode, in order to close the circuit for alternating current and, at the same time, to isolate one terminal of the tuning diode from the rest of the circuit with respect to direct current, so as to enable the tuning voltage to be applied to the diode. If possible, the value of the series resistor C_S will be chosen such that the effective capacitance variation is not restricted. However, in some cases—for example, in the oscillator circuit of receivers whose intermediate frequency is of the order of magnitude of the reception frequency—this is not possible, and the influence of the series capacitance will then have to be taken into account. By connecting the capacitor C_S, assumed to be lossless, in series with the diode capacitance C_{tot}, the tuning capacitance is reduced to the value

$$C^* = C_{tot} \frac{1}{1 + C_{tot}/C_S} \tag{5-88}$$

The Q of the effective tuning capacitance, taking into account the Q of the tuning diode, increases to

$$Q^* = Q\left(1 + \frac{C_{tot}}{C_S}\right) \tag{5-89}$$

The useful capacitance ratio is reduced to the value

$$\frac{C^*_{max}}{C^*_{min}} = \frac{C_{max}}{C_{min}} \frac{1 + C_{min}/C_S}{1 + C_{max}/C_S} \tag{5-90}$$

where C_{max} and C_{min} are the maximum and minimum capacitances, respectively, of the tuner diode.

On the other hand, the advantage is gained that, due to the capacitive potential division, the amplitude of the alternating voltage applied to the tuner diode is reduced to

$$\hat{v}^* = \hat{v} \frac{1}{1 + C_{tot}/C} \tag{5-91}$$

so that the lower value of the tuning voltage can be smaller, and this results in a higher maximum capacitance C_{max} of the tuning diode and a higher useful capacitance ratio. The

influence exerted by the series capacitor, then, can actually be kept lower than Eq. (5-89) would suggest.

The parallel capacitance C_P that appears in Figures 5-54 and 5-55 is always present, since wiring capacitances are inevitable and every coil has its self-capacitance. By treating the capacitance C_P, assumed to be lossless, as a shunt capacitance, the total tuning capacitance rises in value and, if C_S is assumed to be large enough to be disregarded, we obtain

$$C^* = C_{\text{tot}}\left(1 + \frac{C_P}{C_{\text{tot}}}\right) \tag{5-92}$$

The Q of the effective tuning capacitance, derived from the Q of the tuning diode, is

$$Q^* = Q\left(1 + \frac{C_P}{C_{\text{tot}}}\right) \tag{5-93}$$

or, in other words, it rises with the magnitude of the parallel capacitance. The useful capacitance is reduced:

$$\frac{C^*_{\text{max}}}{C^*_{\text{min}}} = \frac{C_{\text{max}}}{C_{\text{min}}}\,\frac{1 + C_P/C_{\text{max}}}{1 + C_P/C_{\text{min}}} \tag{5-94}$$

In view of the fact that even a comparatively small shunt capacitance reduces the capacitance ratio considerably, it is necessary to ensure low wiring and coil capacitances in the layout stage of the circuit design.

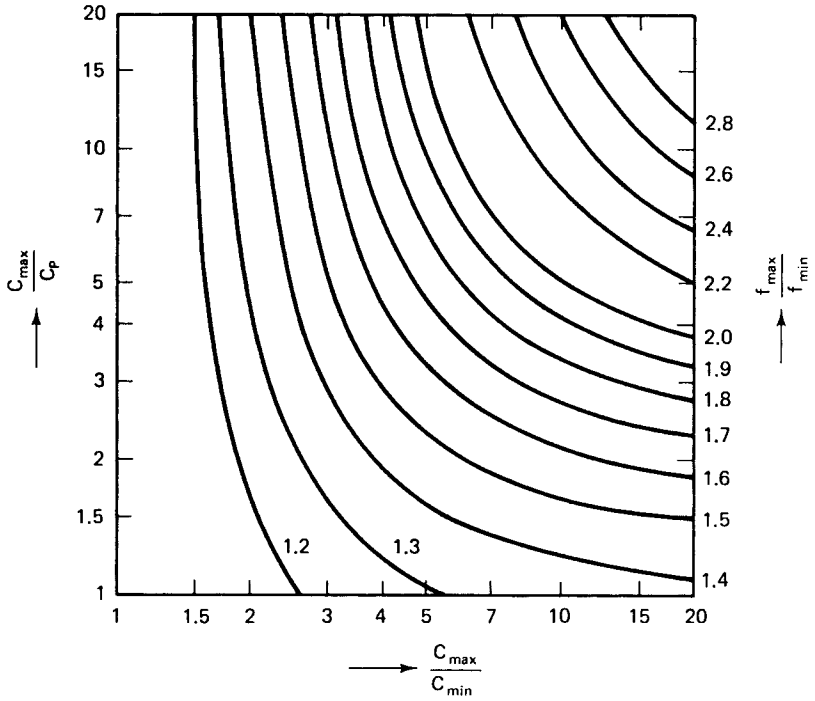


Figure 5-56 Diagram for determining capacitance ratio and maximum capacitance.

Tuning Range. The frequency range over which a parallel-resonant circuit (according to Figure 5-54) can be tuned by means of the tuning diode depends on the useful capacitance ratio of the diode and on the parallel and series capacitances present in the circuit. The ratio is

$$\frac{f_{\max}}{f_{\min}} = \left(\frac{1 + \dfrac{C_{\max}}{C_P(1 + C_{\max}/C_S)}}{1 + \dfrac{C_{\max}}{C_P(C_{\max}/C_{\min} + C_{\max}/C_S)}} \right)^{1/2} \tag{5-95}$$

In many cases, the value of the series capacitor can be chosen to be large enough for its effect to be negligible. In that case, Eq. (5-95) is simplified as follows:

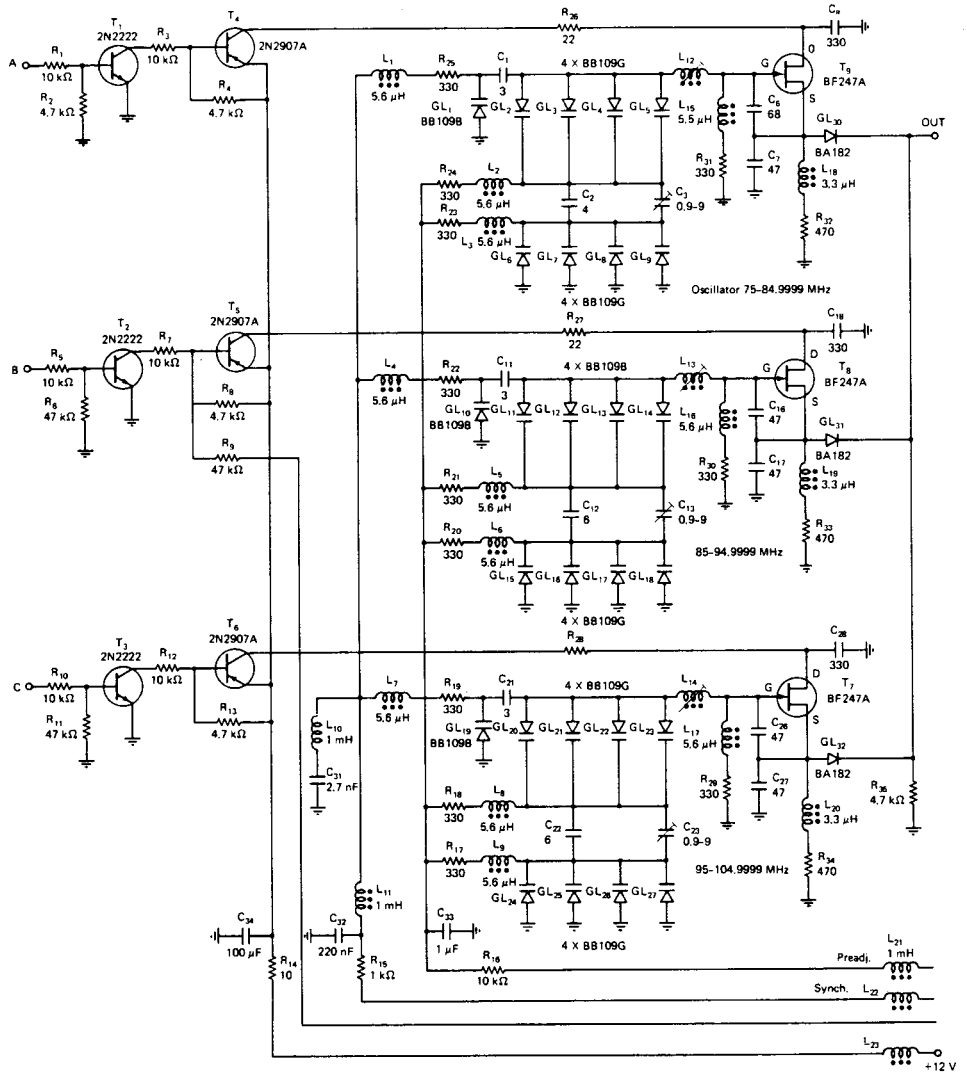


Figure 5-57 Oscillator and switching section of the Rohde & Schwarz ESH2/ESH3 test receiver.

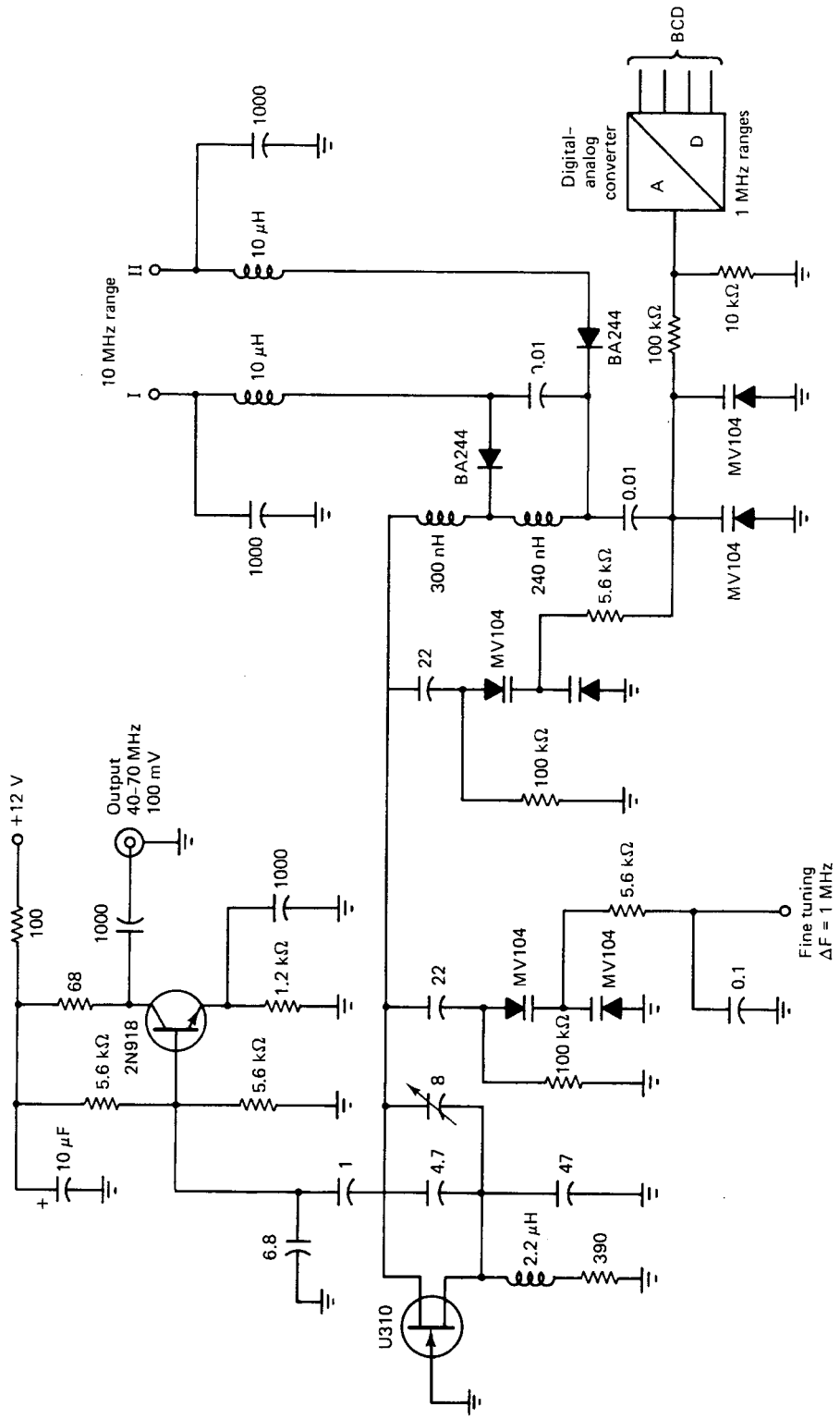


Figure 5-58 A 40–70-MHz VCO with two coarse-steering ranges and a fine-tuning range of 1 MHz.

$$\frac{f_{\max}}{f_{\min}} = \left(\frac{1 + C_{\max}/C_P}{1 + C_{\min}/C_P} \right)^{1/2} \tag{5-96}$$

From this equation, the diagram shown in Figure 5-56 is computed. With the aid of this diagram, the tuning diode parameters required for tuning a resonant circuit over a stipulated frequency range (i.e., the maximum capacitance and the capacitance ratio) can be determined. Whenever the series capacitance C_S cannot be disregarded, the effective capacitance ratio is reduced accorded to Eq. (5-90).

Tracking. When several tuned circuits are used on the same frequency, diodes must be selected for perfect tracking.

5-5-9 Practical Circuits

Oscillators with Coarse and Fine Tuning. After so much theory, it may be nice to take a look at some practical circuits, such as the one shown in Figure 5-57. This oscillator is being used in the Rohde & Schwarz ESN/ESVN40 field-strength meter, and in the HF1030 receiver produced by Cubic Communications, San Diego. This circuit combines all the various techniques shown previously. A single diode is being used for fine-tuning a narrow range of less than 1 MHz; coarse tuning is achieved with the antiseries diodes.

Several unusual properties of this circuit are apparent:

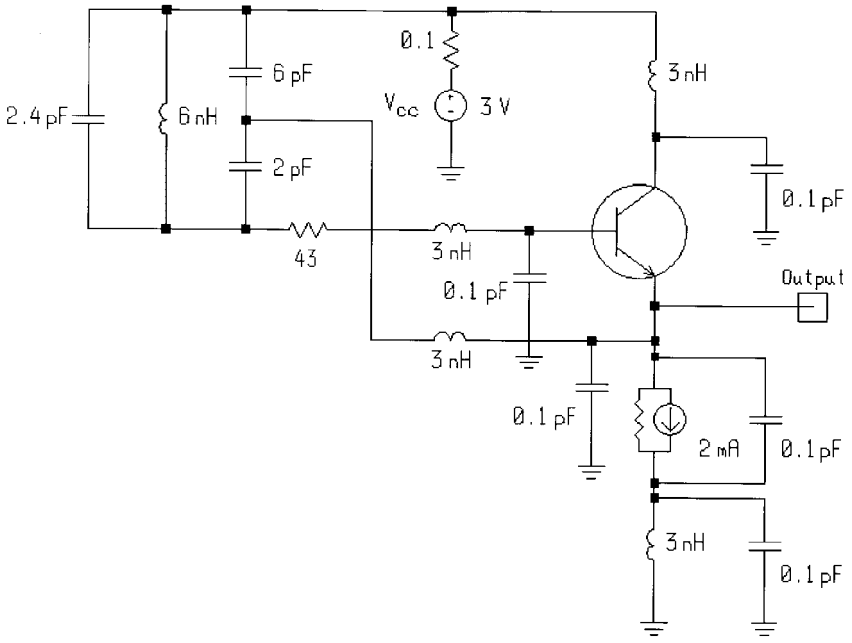


Figure 5-59 Schematic of the dc-coupled oscillator.

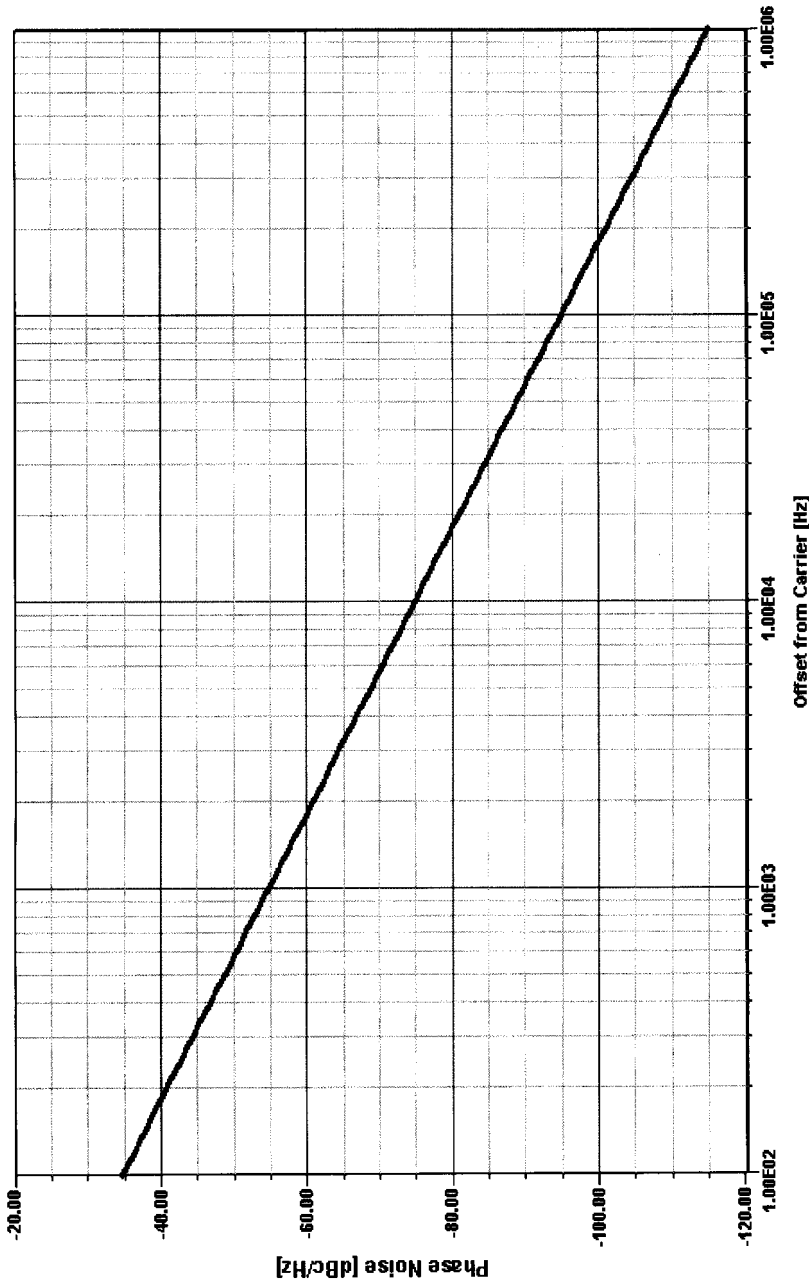


Figure 5-60 Phase noise of the dc-coupled oscillator operating at 897 MHz.

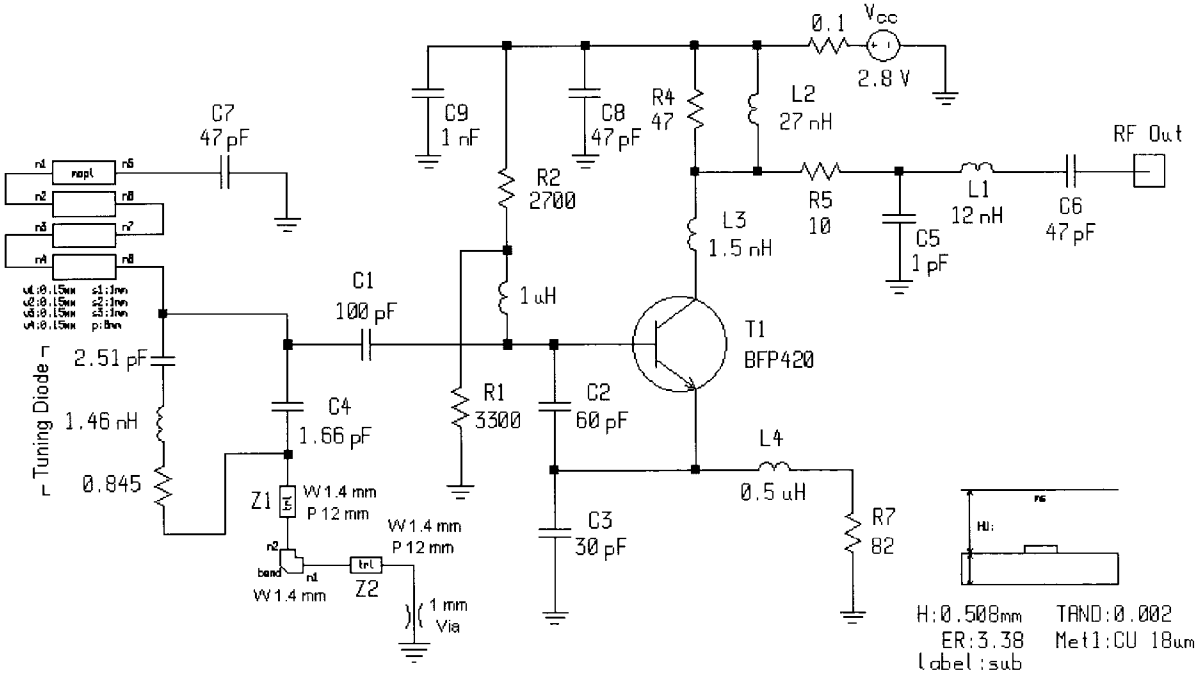


Figure 5-61 Schematic of the Siemens VCO. This oscillator (simulation) works with such high voltage-divider capacitances (C_2 and C_3) only because the circuit's large-signal transistor model, which was obtained under contract, exhibits unrealistically high gain at the frequency of oscillation. Actual capacitors with values on the order of those shown for C_2 and C_3 typically exhibit parasitic resistance and inductance of 0.3 Ω and 0.2 nH, respectively; the resulting series resonance would make any actual such capacitor act as an inductor at UHF. Obtaining practical "UHF useful" capacitances with values on the order of those shown for C_2 and C_3 would therefore necessitate the use of multiple, smaller-value capacitors in parallel—4 × 15 pF at C_2, for example.

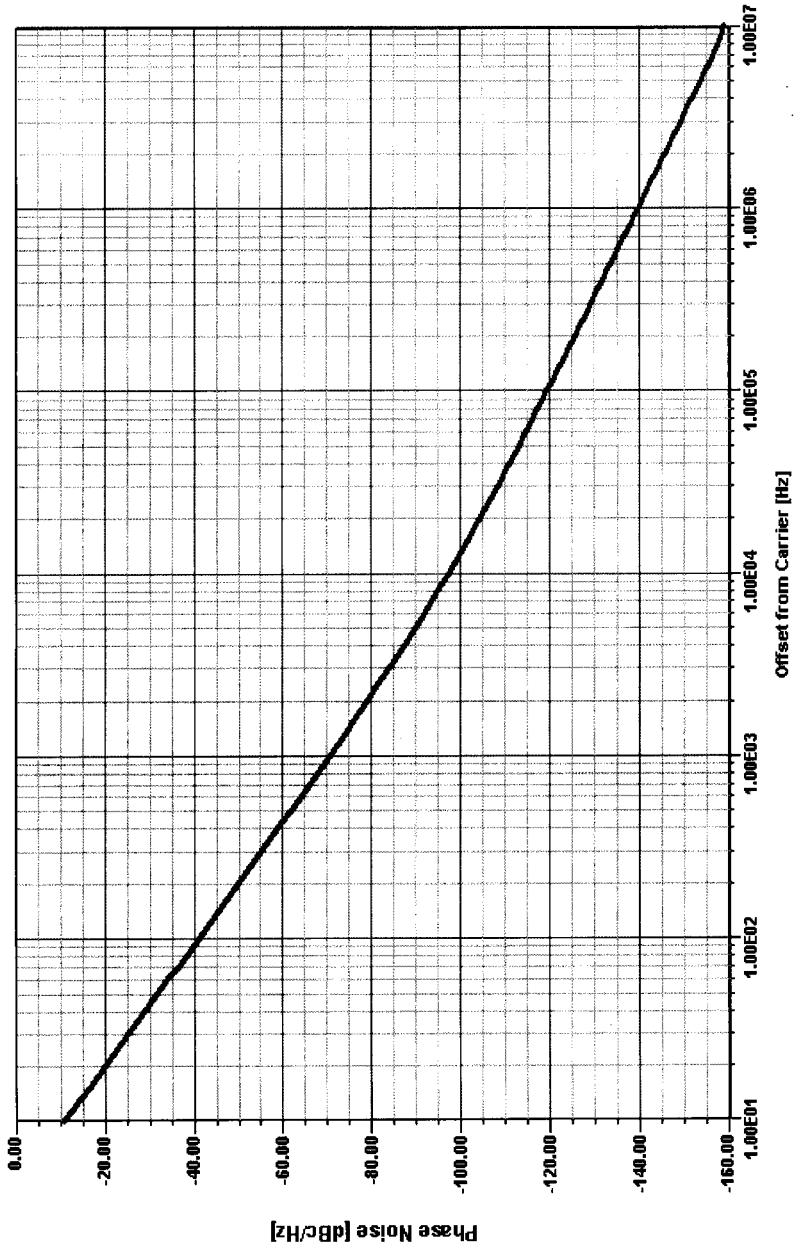


Figure 5-62 Simulated phase noise of the VCO. The frequency of oscillation is 1.052 GHz.

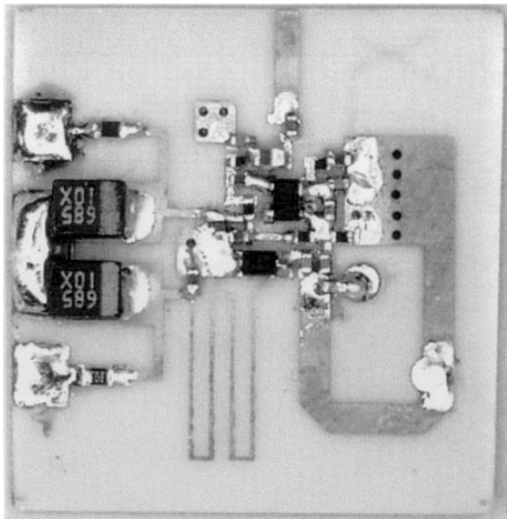


Figure 5-63 Implementation of the Siemens VCO.

1. The fine tuning is achieved with a tuning diode that has a much larger capacitance than that of the coupling capacitor in the circuit. The advantage of this technique is that the fixed capacitor and the tuning diode form a voltage divider, whereby the voltage across the tuning diode decreases as the capacitance increases. For larger values of the capacitance of the tuning diode, the Q changes and the gain K_0 increases. Because of the voltage division, the noise contribution and loading effect of the diode are reduced.
2. In the coarse-tuning circuit, several tuning diodes are used in parallel. The advantage of this circuit is a change in L/C ratio by using a higher C and storing more energy in the tuned circuit. There are no high-Q diodes available with such large capacitance values, and therefore preference is given to using several diodes in parallel rather than one tuning diode with a large capacitance, normally used only for AM (medium-frequency broadcast) tuner circuits.

We mentioned previously that, despite this, the coarse-tuning circuit will introduce noise outside the loop bandwidth, where it cannot be corrected. It is therefore preferable to incorporate switching diodes for segmenting ranges at the expense of switching current drain.

Figure 5-58 shows a circuit using a combined technique of tuning diodes for fine- and medium-resolution tuning, and coarse tuning with switching diodes.

dc-Coupled Oscillator. An interesting circuit that was cut out of an IC was supplied to us by David Lovelace of Motorola. Figure 5-59 shows the circuit and Figure 5-60 shows its phase noise. As a result of the constraints of integration, this BJT oscillator has no base bias circuitry in the conventional form. A fairly noisy performer, it ultimately was not used in production.

Siemens Colpitts Oscillator. A more elaborate approach that is essentially a standard Colpitts oscillator with a lot of detailed modeling is the oscillator of Figure 5-61, supplied

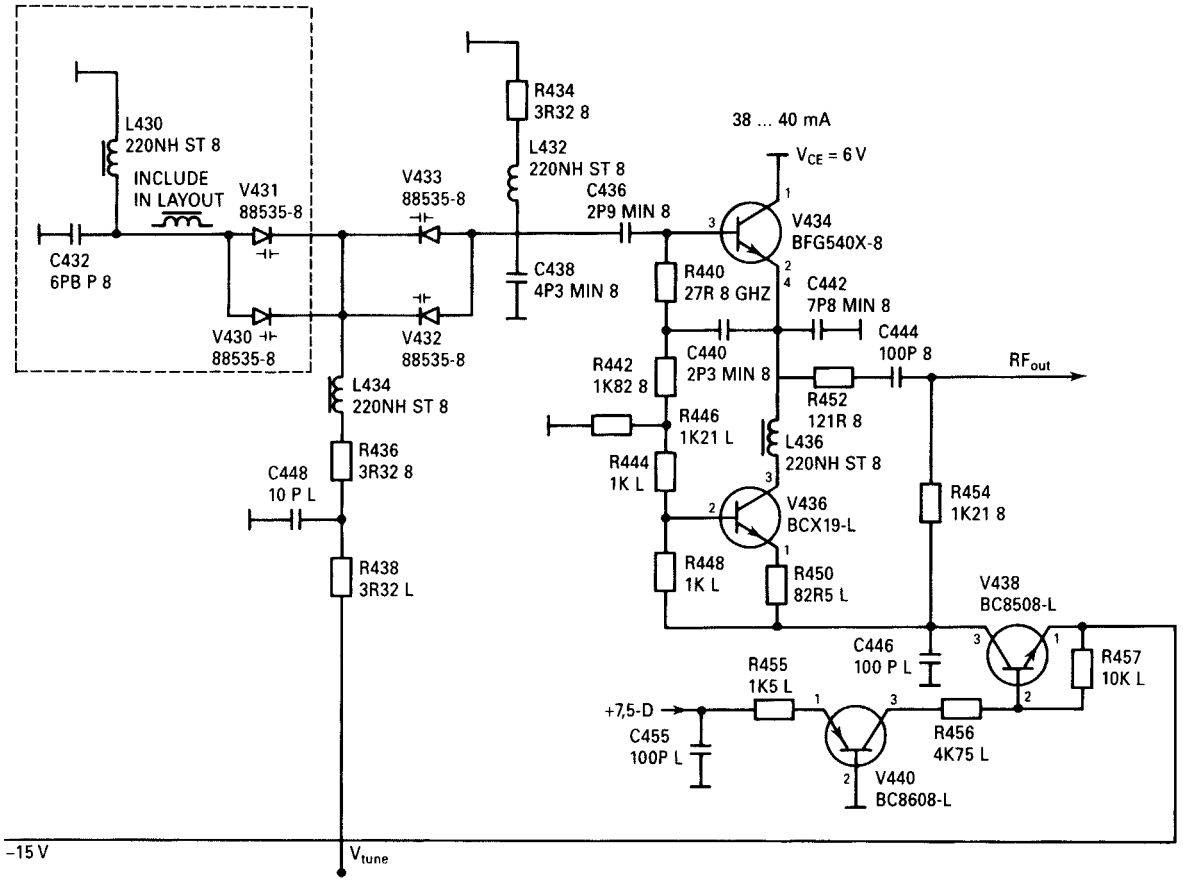


Figure 5-64 A recommended very-low-phase-noise VCO that operates in the 1-GHz region. The V436 stage is a constant-current generator responsible for improved phase-noise performance.

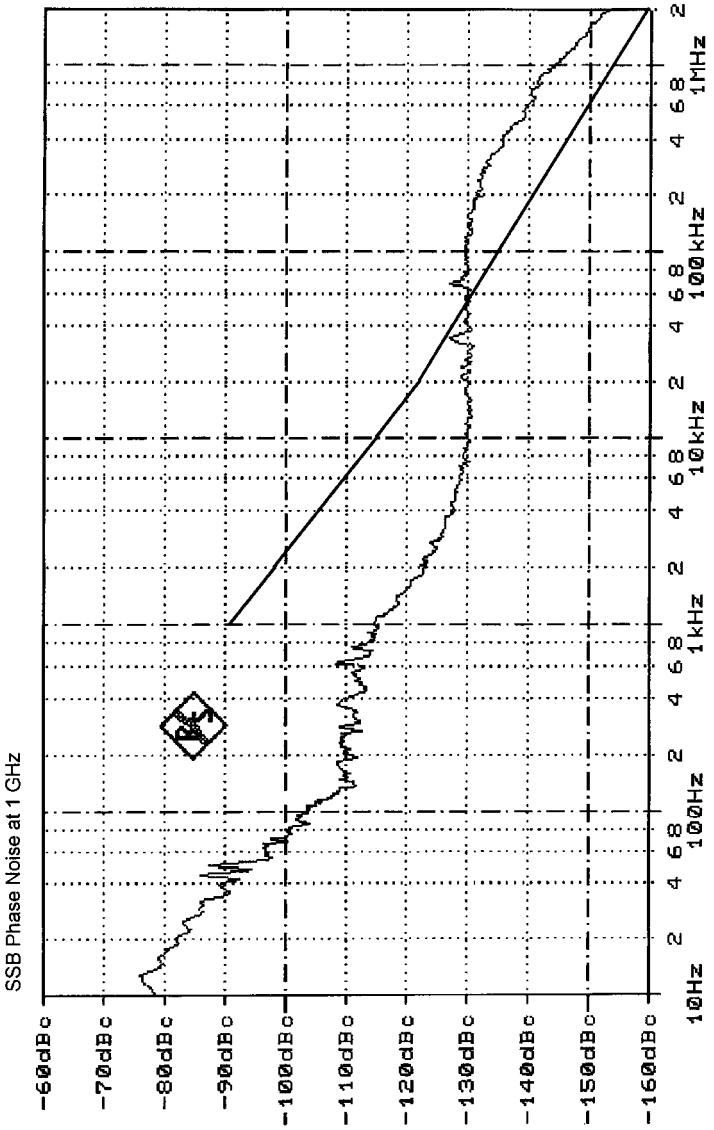


Figure 5-65 Phase noise of the low-noise VCO. The "ledges" from 100 Hz to 1 kHz and from 10 kHz to 100 kHz reflect the noise-cleanup effects of a dual-loop synthesizer; the superimposed line from 1 kHz to 2 MHz shows the oscillator's noise characteristic with both loops unlocked. The closed-loop response is noisier than the open-loop response at offsets above 60 kHz because a wide loop bandwidth is used to achieve a faster switching speed.

to us by Siemens. Considering its simplicity, its noise performance is good, as shown in Figure 5-62. Modeling this circuit involved the consideration of via holes and coupled lines. Figure 5-63 shows an actual implementation of this oscillator. The value of external emitter resistance (R7 in Figure 5-61) plays a critical role in determining a common-emitter Colpitts oscillator's upper frequency limit as described in Yeom [8].

dc-Stabilized Oscillator. For high-performance synthesizers, such as the Rohde & Schwarz series SMG/SMH, a dc-stabilized oscillator with fairly wide tuning range may be used. Figure 5-64 shows such an oscillator, the measured phase noise of which is shown in Figure 5-65. It should be noted that the output loop is magnetically coupled and therefore takes advantage of the inherent selectivity of the tuned circuit.

5-6 NOISE IN OSCILLATORS

In transmitters, oscillator noise can result in adjacent-channel interference and modulation errors; in receivers, oscillator noise can result in demodulation errors and degraded sensitivity and dynamic range. The specification, calculation, and reduction of oscillator noise are therefore of great importance in wireless system design.

5-6-1 Linear Approach to the Calculation of Oscillator Phase Noise

Since an oscillator can be viewed as an amplifier with feedback (Figure 5-66), it is helpful to examine the phase noise added to an amplifier that has a noise figure F. F is defined by [9]

$$F = \frac{(S/N)_{\text{in}}}{(S/N)_{\text{out}}} = \frac{N_{\text{out}}}{N_{\text{in}}G} = \frac{N_{\text{out}}}{GkTB} \tag{5-97}$$

$$N_{\text{out}} = FGkTB \tag{5-98}$$

$$N_{\text{in}} = kTB \tag{5-99}$$

where N_{in} is the total input noise power to a noise-free amplifier. The input phase noise in a 1-Hz bandwidth at any frequency $f_0 + f_m$ from the carrier produces a phase deviation given by (Figure 5-67)

$$\Delta\theta_{\text{peak}} = \frac{V_{n\,\text{RMS1}}}{V_{s,\text{av RMS}}} = \sqrt{\frac{fKT}{P_{s,\text{av}}}} \tag{5-100}$$

$$\Delta\theta_{\text{1RMS}} = \frac{1}{\sqrt{2}}\sqrt{\frac{FkT}{P_{s,\text{av}}}} \tag{5-101}$$

Since a correlated random phase-noise relation exists at $f_0 - f_m$, the total phase deviation becomes

$$\Delta\theta_{\text{RMS total}} = \sqrt{FkT/P_{s,\text{av}}} \tag{5-102}$$

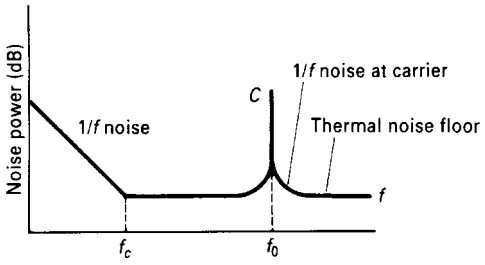


Figure 5-66 Noise power versus frequency of a transistor amplifier with an input signal applied.

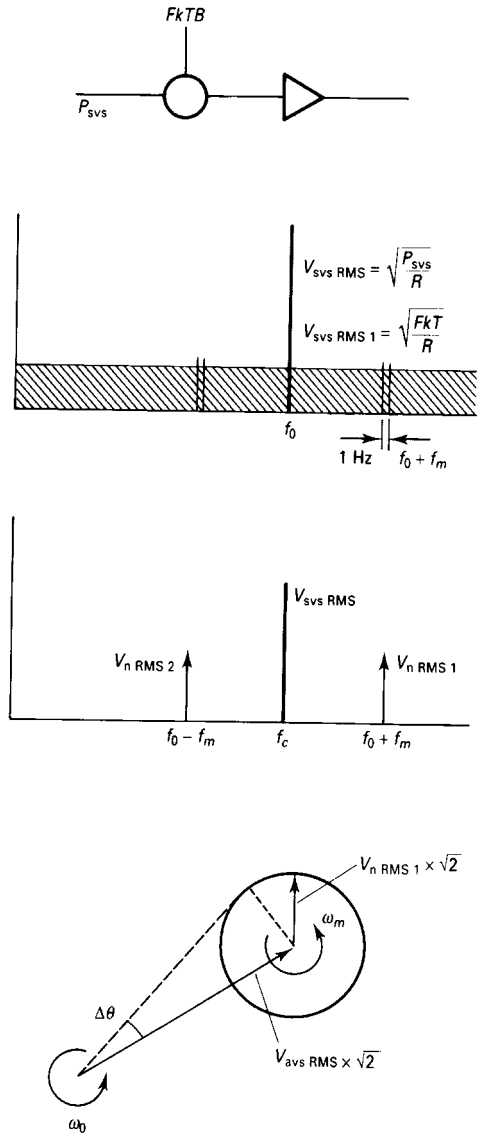


Figure 5-67 Phase noise added to carrier.

The spectral density of phase noise becomes

$$S_\theta(f_m) = \Delta\theta^2_{\text{RMS}} = FkTB/P_{s,\text{av}} \tag{5-103}$$

where $B = 1$ for a 1-Hz bandwidth. Using

$$kTB = -174 \text{ dBm/Hz} \qquad (B = 1) \tag{5-104}$$

allows a calculation of the spectral density of phase noise that is far removed from the carrier (i.e., at large values of f_m). This noise is the theoretical noise floor of the amplifier. For example, an amplifier with +10-dBm power at the input and a noise figure of 6 dB gives

$$S_\theta(f_m > f_c) = -174 \text{ dBm} + 6 \text{ dB} - 10 \text{ dBm} = -178 \text{ dBm} \tag{5-105}$$

Only if P_{OUT} is >0 dBm can we expect $\mathcal{L}$ (signal-to-noise ratio) to be greater than 174 dBc/Hz (1-Hz bandwidth). For a modulation frequency close to the carrier, $S_\theta(f_m)$ shows a flicker or $1/f$ component, which is empirically described by the corner frequency f_c. The phase noise can be modeled by a noise-free amplifier and a phase modulator at the input as shown in Figure 5-68.

The purity of the signal is degraded by the flicker noise at frequencies close to the carrier. The spectral phase noise can be described by

$$S_\theta(f_m) = \frac{FkTB}{P_{s,\text{av}}}\left(1 + \frac{f_c}{f_m}\right) \quad (B = 1) \tag{5-106}$$

No AM-to-PM conversion is considered in this equation. The oscillator may be modeled as an amplifier with feedback as shown in Figure 5-69. The phase noise at the input of the amplifier is affected by the bandwidth of the resonator in the oscillator circuit in the following way. The tank circuit or bandpass resonator has a low-pass transfer function

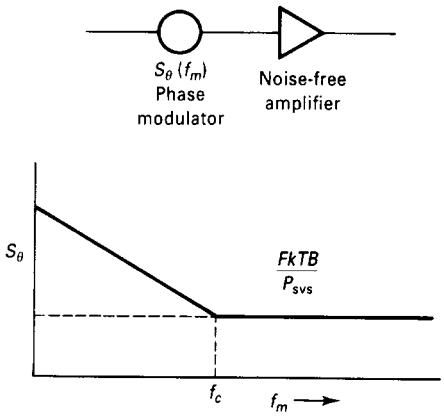


Figure 5-68 Phase noise modeled by a noise-free amplifier and phase modulator.

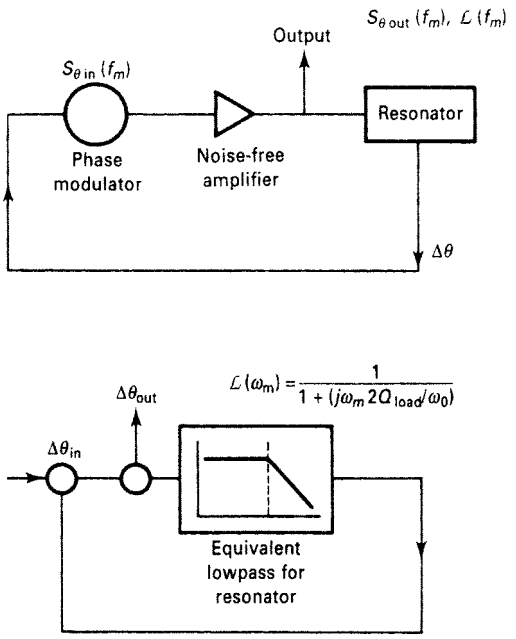


Figure 5-69 Equivalent feedback models of oscillator phase noise.

$$L(\omega_m) = \frac{1}{1 + j(2Q_L\omega_m/\omega_0)} \tag{5-107}$$

where

$$\omega_0/2Q_L = B/2 \tag{5-108}$$

is the half-bandwidth of the resonator. These equations describe the amplitude response of the bandpass resonator; the phase noise is transferred unattenuated through the resonator up to the half-bandwidth.

The closed-loop response of the phase feedback loop is given by

$$\Delta\theta_{\text{out}}(f_m) = \left(1 + \frac{\omega_0}{j2Q_L\omega_m}\right)\Delta\theta_{\text{in}}(f_m) \tag{5-109}$$

The power transfer becomes the phase spectral density

$$S_{\theta,\text{out}}(f_m) = \left[1 + \frac{1}{f_m^2}\left(\frac{f_0}{2Q_L}\right)^2\right] S_{\theta,\text{in}}(f_m) \tag{5-110}$$

where $S_{\theta,\text{in}}$ was given by Eq. (5-106). Finally, $\mathcal{L}(f_m)$ is

$$\mathcal{L}(f_m) = \frac{1}{2}\left[1 + \frac{1}{f_m^2}\left(\frac{f_0}{2Q_L}\right)^2\right] S_{\theta,\text{in}}(f_m) \tag{5-111}$$

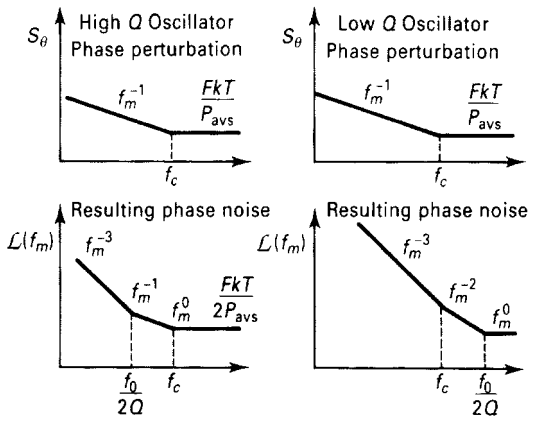


Figure 5-70 Equivalent feedback models of oscillator phase noise.

This equation describes the phase noise at the output of the amplifier (flicker corner frequency and AM-to-PM conversion are not considered). The phase perturbation $S_{\theta,\text{in}}$ at the input of the amplifier is enhanced by the positive phase feedback within the half-bandwidth of the resonator, $f_0/2Q_L$.

Depending on the relation between f_c and $f_0/2Q_L$, there are two cases of interest, as shown in Figure 5-70. For the low-Q case, the spectral phase noise is unaffected by the Q of the resonator, but the $\mathcal{L}(f_m)$ spectral density will show a $1/f^3$ and $1/f^2$ dependence close to the carrier. For the high-Q case, a region of $1/f^3$ and $1/f$ should be observed near the carrier. Substituting Eq. (5-106) into Eq. (5-111) gives an overall noise of

$$\mathcal{L}(f_m) = \frac{1}{2}\left[1 + \frac{1}{f_m^2}\left(\frac{f}{2Q_L}\right)^2\right]\frac{FkT}{P_{s,\text{av}}}\left(1 + \frac{f_c}{f_m}\right)$$

$$= \frac{FkTB}{2P_{s,\text{av}}}\left[\frac{1}{f_m^3}\frac{f^2 f_c}{4Q_L^2} + \frac{1}{f_m^2}\left(\frac{f}{2Q_L}\right)^2 + \left(1 + \frac{f_c}{f_m}\right)\right] \text{dBc/Hz} \tag{5-112}$$

Examining Eq. (5-112) gives the four major causes of oscillator noise: the upconverted $1/f$ noise or flicker FM noise, the thermal FM noise, the flicker phase noise, and the thermal noise floor, respectively.

Q_L (loaded Q) can be expressed as

$$Q_L = \frac{\omega_o W_e}{P_{\text{diss,total}}} = \frac{\omega_o W_e}{P_{\text{in}} + P_{\text{res}} + P_{\text{sig}}} = \frac{\text{Reactive power}}{\text{Total dissipated power}} \tag{5-113}$$

where W_e is the reactive energy stored in L and C,

$$W_e = \frac{1}{2} CV^2 \tag{5-114}$$

$$P_{\text{res}} = \frac{\omega_o W_e}{Q_{\text{unl}}} \tag{5-115}$$

$$\mathcal{L}(f_m) = \frac{1}{2}\left[1 + \frac{\omega_o^2}{4\omega_m^2}\left(\frac{P_{\text{in}}}{\omega_o W_e} + \frac{1}{Q_{\text{unl}}} + \frac{P_{\text{sig}}}{\omega_o W_e}\right)^2\right]\left(1 + \frac{\omega_c}{\omega_m}\right)\frac{FkT_o}{P_{\text{sav}}} \tag{5-116}$$

Input power/ Reactive power ratio

Resonator Q

Signal power/ Reactive power ratio

Flicker effect

Phase perturbation

This equation is extremely significant because it covers most of the causes of phase noise in oscillators. [AM-to-PM conversion must be added; see Eq. (5-56).] To minimize the phase noise, the following design rules apply:

1. Maximize the unloaded Q.
2. Maximize the reactive energy by means of a high RF voltage across the resonator and obtain a low L/C ratio. The limits are set by breakdown voltages of the active devices and the tuning diodes, and the forward-bias condition of the tuning diodes. If L becomes too high, the circuit degenerates into a squegging oscillator. Using lower L allows larger capacitance, which allows us to swamp the oscillating device's internal, nonlinear capacitance with external, linear capacitance, reducing the phase-noise degradation caused by the nonlinear capacitance. The ceramic-resonator oscillator best illustrates this approach, which may be hard to accomplish in discrete circuits because high-Q inductors with values above 1 nH cannot be built unless high-Q transmission lines or resonators are used.
3. Avoid saturation at all costs, and try to have either limiting or automatic gain control (AGC) without degradation of Q. Isolate the tuned circuit from the limiter or AGC circuit. Use tuning diodes in antiseries configurations to avoid forward bias.
4. Choose an active device with the lowest possible noise figure and flicker corner frequency. The noise figure of interest is the noise figure obtained with the actual impedance at which the device is operated. Using FETs rather than BJTs, it is preferable to deal with the equivalent noise voltage and noise currents rather than with the noise figure, since they are independent of source impedance. The noise figure improves as the ratio between source impedance and equivalent noise resistance increases. In addition, in a tuned circuit, the source impedance changes drastically as a function of the offset frequency, and this effect has to be considered. For low phase-noise operation, use a medium-power transistor. If you need your output power to be achieved at 6–9 mA, select a transistor with an $I_{C,\max}$ of 60–90 mA. Also avoid an f_T greater than 3× to 5× the operating frequency. To make transistors that are stable across their full frequency ranges, manufacturers add circuitry that makes the flicker corner frequency higher as f_T increases.

   The following transistors have the lowest noise figure:

   - BFG65/67 by Philips
   - BFR106/92; BFP405/420 by Siemens
   - 2SC3358/3356 by NEC
   - HXTR4105 by HP

- AT2100 (chip) by Avantek

Among the lowest-noise JFETs are the U310, 2N4416/17, and 2N5397.

In designing MMICs, one must resort to MESFETs, which are also referred to as GaAsFETs. They have a fairly high flicker frequency, which is typically around 10 MHz but can go up to as high as 100 MHz.

Table 5-5 shows equivalent circuits and measurement data for the Avantek AT2100 20-GHz *npn* silicon bipolar oscillator transistor. (Siemens of Germany is about to come out with 45-GHz f_T transistors, which will compete with SiGe transistors in the future.)

5. Phase perturbation can be minimized by using high-impedance devices such as FETs, where the signal-to-noise ratio of the signal voltage relative to the equivalent noise voltage can be made very high. This also indicates that in the case of a limiter, the limited voltage should be as high as possible.
6. Choose an active device with low flicker noise. The effect of flicker noise can be reduced through RF feedback. An unbypassed emitter resistor of 10–30 Ω in a bipolar circuit can improve the flicker noise by as much as 40 dB.

   The proper bias point of the active device is important, and precautions should be taken to prevent modulation of the input and output dynamic capacitances of the active device, which will cause amplitude-to-phase conversion and therefore introduce noise.
7. The energy should be coupled from the resonator rather than from another portion of the active device so that the resonator limits the bandwidth.
8. For microwave applications, the lowest-noise GaAsFETs are pseudomorphic HJ FETs. A reproduction of the datasheet of the NEC NE42484A, which is generally the state of the art today, follows; it is used with permission.

Equation (5-116) assumes that the phase perturbation and the flicker effect are the limiting factors, as practical use of such oscillators requires that an isolation amplifier must be used.

In the event that energy is taken directly from the resonator and the oscillator power can be increased, the signal-to-noise ratio can be increased above the theoretical limit of 174 dBc/Hz, due to the low-pass-filter effect of the tuned resonator. However, since this is mainly a theoretical assumption and does not represent a real-world system, this noise performance cannot be obtained. In an oscillator stage, even a total noise floor of –170 dBc/Hz is rarely achieved.

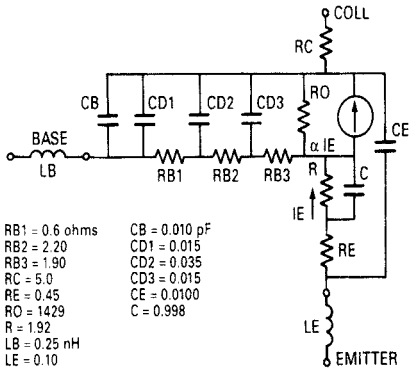
What other influences do we have that cause an oscillator's noise performance to degrade? So far, we have assumed that the Q of the tuned circuit is really determined only by the LC network and the loading effect of the transistor. In synthesizer applications, however, we find it necessary to add a tuning diode. The tuning diode has a substantially lower Q than that of a mica capacitor or even a ceramic capacitor. As a result of this, the noise sidebands change as a function of the additional loss. This is best expressed by adjusting the value for the loaded Q in Eq. (5-113).

There seems to be no precise mathematical way of predetermining the noise influence of a tuning diode, but the following approximation seems to give proper results:

$$\frac{1}{Q_{T,\text{load}}} = \frac{1}{Q_{\text{load}}} + \frac{1}{Q_{\text{diode}}} \tag{5-117}$$

Table 5-5 Linear and SPICE equivalent circuits for the Avantek AT21400 chip

LINEAR EQUIVALENT CIRCUIT[a]
V_{CE} = 8 V, I_C = 14 mA

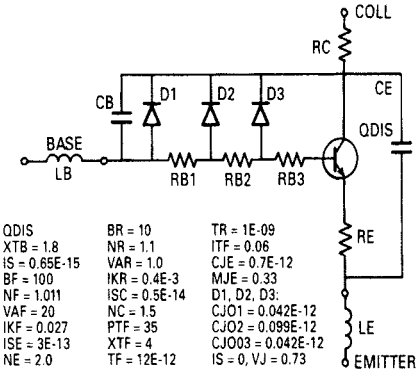


RB1 = 0.6 ohms
RB2 = 2.20
RB3 = 1.90
RC = 5.0
RE = 0.45
RO = 1429
R = 1.92
LB = 0.25 nH
LE = 0.10

CB = 0.010 pF
CD1 = 0.015
CD2 = 0.035
CD3 = 0.015
CE = 0.0100
C = 0.998

CCCS: M = .99, A = 0
R1 = R, R2 = 1E8, F = 22.7, T = 7.4

SPICE EQUIVALENT CIRCUIT[a]



QDIS
XTB = 1.8
IS = 0.65E-15
BF = 100
NF = 1.011
VAF = 20
IKF = 0.027
ISE = 3E-13
NE = 2.0

BR = 10
NR = 1.1
VAR = 1.0
IKR = 0.4E-3
ISC = 0.5E-14
NC = 1.5
PTF = 35
XTF = 4
TF = 12E-12

TR = 1E-09
ITF = 0.06
CJE = 0.7E-12
MJE = 0.33
D1, D2, D3:
CJO1 = 0.042E-12
CJO2 = 0.099E-12
CJO03 = 0.042E-12
IS = 0, VJ = 0.73

Modeled Scattering Parameters, Common Emitter[b]

T_A = 25°C, V_{CE} = 8 V, I_C = 14 mA

Freq.	S_{11}		S_{21}			S_{12}			S_{22}	
GHz	Mag	Ang	dB	Mag	Ang	dB	Mag	Ang	Mag	Ang
2	.67	−168	15.8	6.19	90	−34.0	.02	52	.65	−9
3	.67	−180	12.4	4.17	81	−32.0	.025	59	.65	−10
4	.67	172	10.0	3.15	74	−30.5	.03	64	.65	−11
5	.68	165	8.1	2.53	68	−28.0	.04	68	.65	−12
6	.68	160	6.5	2.12	62	−28.0	.04	70	.65	−14
7	.69	154	5.2	1.82	56	−26.0	.05	72	.65	−16
8	.70	149	4.1	1.60	51	−24.4	.06	73	.65	−18
9	.70	144	3.1	1.43	46	−24.4	.06	74	.65	−20
10	.71	140	2.2	1.29	41	−23.1	.07	74	.65	−22
11	.72	135	1.4	1.18	36	−21.9	.08	74	.65	−24
12	.73	131	0.7	1.08	31	−21.9	.08	75	.66	−26
13	.74	127	0	1.00	27	−20.9	.09	75	.66	−28
14	.75	123	−0.6	0.93	23	−20.0	.10	75	.66	−30
15	.76	120	−1.3	0.86	18	−19.2	.11	74	.67	−32
16	.77	116	−1.9	0.80	14	−19.2	.11	74	.67	−34
18	.79	109	−3.1	0.70	7	−17.7	.13	74	.68	−39
20	.80	103	−4.2	0.62	−1	−17.1	.14	73	.69	−44
22	.82	97	−5.4	0.54	−7	−15.9	.16	72	.70	−48
24	.84	92	−6.6	0.47	−14	−14.9	.18	71	.71	−53
26	.85	87	−8.0	0.40	−20	−14.0	.20	70	.72	−58

[a]These equivalent circuits are provided only as first-order design aids. Their accuracy for critical designs at very high frequencies has not been validated.

[b]*S* parameters are from the linear equivalent circuit. Below 10 GHz, they have been fit to measurements of die on a standard carrier with one bond wire to the base and four wires to the emitter.

NEC

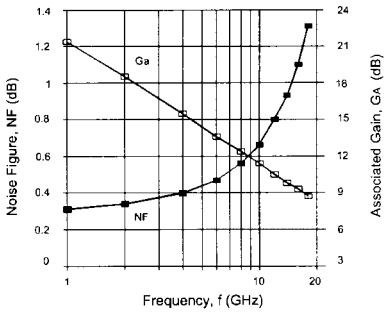
SUPER LOW NOISE PSEUDOMORPHIC HJ FET

NE42484A

FEATURES

- **VERY LOW NOISE FIGURE:** 0.8 dB TYP at 12 GHz
- **HIGH ASSOCIATED GAIN:** 10.5 dB TYP at 12 GHz
- **L_G = 0.35 µm, W_G = 200 µm**
- **LOW COST METAL CERAMIC PACKAGE**
- **TAPE & REEL PACKAGING OPTION AVAILABLE**

NOISE FIGURE & ASSOCIATED GAIN vs. FREQUENCY
V_{DS} = 2 V, I_{DS} = 10 mA



DESCRIPTION

The NE42484A is a pseudomorphic Hetero-Junction FET that uses the junction between Si-doped AlGaAs and undoped InGaAs to create very high mobility electrons. The device features mushroom shaped TiAl gates for decreased gate resistance and improved power handling capabilities. The mushroom gate also results in lower noise figure and high associated gain. This device is housed in an epoxy-sealed, metal/ceramic package and is intended for high volume consumer and industrial applications.

NEC's stringent quality assurance and test procedures ensure the highest reliability and performance.

ELECTRICAL CHARACTERISTICS (T_A = 25°C)

PART NUMBER PACKAGE OUTLINE			NE42484A 84AS		
SYMBOLS	**PARAMETERS AND CONDITIONS**	**UNITS**	**MIN**	**TYP**	**MAX**
NF_{OPT}[1]	Optimum Noise Figure, V_{DS} = 2.0 V, I_{DS} = 10 mA, f = 12 GHz	dB		0.8	1.2
G_A[1]	Associated Gain, V_{DS} = 2.0 V, I_{DS} = 10 mA, f = 12 GHz	dB	9.0	10.5	
P_{1dB}	Output Power at 1 dB Gain Compression Point, f = 12 GHz V_{DS} = 2.0 V, I_{DS} = 10 mA V_{DS} = 2.0 V, I_{DS} = 20 mA	 dBm dBm		 9.7 10.2	
G_{1dB}	Gain at P_{1dB}, f = 12 GHz V_{DS} = 2.0 V, I_{DS} = 10 mA V_{DS} = 2.0 V, I_{DS} = 20 mA	 dB dB		 10.3 10.5	
I_{DSS}	Saturated Drain Current, V_{DS} = 2.0 V, V_{GS} = 0 V	mA	15	40	70
V_P	Pinch-off Voltage, V_{DS} = 2.0 V, I_{DS} = 0.1 mA	V	-2.0	-0.8	-0.2
g_m	Transconductance, V_{DS} = 2.0 V, I_D = 10 mA	mS	45	60	
I_{GSO}	Gate to Source Leakage Current, V_{GS} = -3.0 V	µA		0.5	10.0
$R_{TH\ (CH-A)}$	Thermal Resistance (Channel to Ambient)	°C/W		750	
$R_{TH(CH-C)}$[2]	Thermal Resistance (Channel to Case)	°C/W			350

Notes:
1. Typical values of noise figures and associated gain are those obtained when 50% of the devices from a large number of lots were individually measured in a circuit with the input individually tuned to obtain the minimum value. Maximum values are criteria established on the production line as a "go-no-go" screening tuned for the "generic" type but not for each specimen.
2. R_{TH} (channel to case) for package mounted on an infinite heat sink.

NE42484A

ABSOLUTE MAXIMUM RATINGS[1] (T_A = 25°C)

SYMBOLS	PARAMETERS	UNITS	RATINGS
V_{DS}	Drain to Source Voltage	V	4.0
V_{GS}	Gate to Source Voltage	V	-3.0
I_{DS}	Drain Current	mA	I_{DSS}
I_{GRF}	Gate Current (RF Drive)	µA	200
P_{IN}	RF Input (CW)	dBm	15
T_{CH}	Channel Temperature	°C	150
T_{STG}	Storage Temperature	°C	-65 to +150
P_T	Total Power Dissipation	mW	165

Note:
1. Operation in excess of any one of these parameters may result in permanent damage.

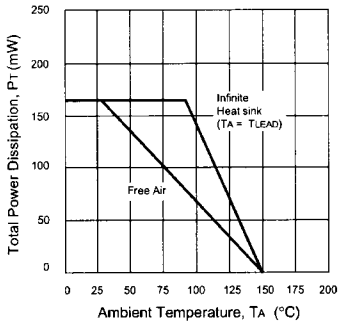
TYPICAL NOISE PARAMETERS (T_A = 25°C)

V_{DS} = 2 V, I_{DS} = 10 mA

FREQ. (GHz)	NF_{OPT} (dB)	G_A (dB)	Γ_{OPT} MAG	Γ_{OPT} ANG	Rn/50
1.0	0.31	21.4	0.78	10	0.43
2.0	0.34	18.5	0.76	28	0.38
4.0	0.40	15.5	0.72	58	0.28
6.0	0.47	13.6	0.65	84	0.21
8.0	0.56	12.4	0.57	113	0.15
10.0	0.66	11.4	0.50	141	0.10
12.0	0.80	10.5	0.44	173	0.09
14.0	0.93	9.8	0.39	-157	0.08
16.0	1.10	9.3	0.36	-125	0.08
18.0	1.31	8.7	0.35	-90	0.08

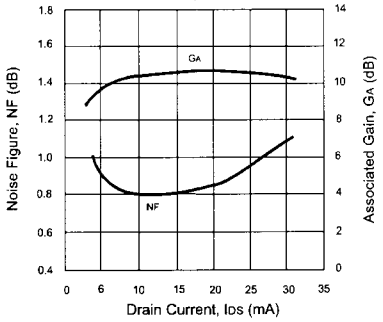
TYPICAL PERFORMANCE CURVES (T_A = 25°C)

TOTAL POWER DISSIPATION vs. AMBIENT TEMPERATURE[1]

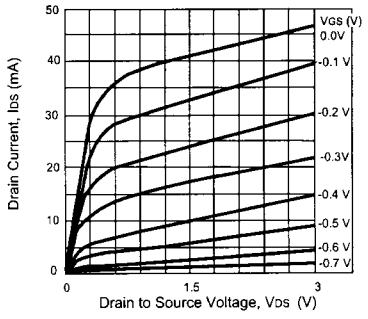


Note
1. If P_T exceeds the Free Air Value, reliable operation can be assured by measuring the worst-case temperature, $T_{(LEAD)}$, at the lead where heat flow is maximum (usually the source lead) and limiting T_A, P_T or $R_{TH\ (CKT)}$

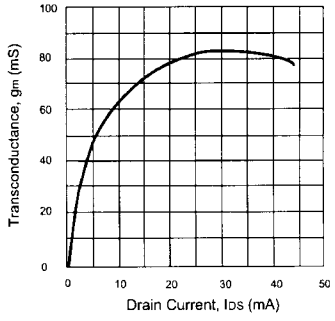
NOISE FIGURE AND ASSOCIATED GAIN vs. DRAIN CURRENT V_{DS} = 2 V, f = 12 GHz



DRAIN CURRENT vs. DRAIN TO SOURCE VOLTAGE



TRANSCONDUCTANCE vs. DRAIN CURRENT V_{DS} = 2.0 V



The tuning diode is specified to have a cutoff frequency $f_{\max}$, which is determined from the loss resistor R_S and the value of the junction capacitance as a function of voltage (e.g., measured at 3 V). This means that the voltage determines the Q and, consequently, the noise bandwidth.

5-6-2 AM-to-PM Conversion

We went into more detail in dealing with the mechanism and influence of tuning diodes in the voltage-controlled oscillators section, where we evaluated the various methods of building voltage-tunable oscillators using tuning diodes. In this section, we limit ourselves to practical results.

The loading effect of the tuning diode is due to losses, and these losses can be described by a resistor in parallel with the tuned circuit.

It is possible to define an equivalent noise $R_{a,\text{eq}}$ that, inserted in Nyquist's equation,

$$V_n = \sqrt{4kT_0R\,\Delta f} \tag{5-118}$$

where $kT_0 = 4.2 \times 10^{-21}$ at about 300 K, R is the equivalent noise resistor, and Δf is the bandwidth, determines an open-circuit noise voltage across the tuning diode. Practical values of $R_{a,\text{eq}}$ for carefully selected tuning diodes are in the vicinity of 200 Ω to 50 kΩ. If we now determine the noise voltage $V_n = \sqrt{4 \times 4.2 \times 10^{-21} \times 10{,}000}$, the resulting voltage value is 1.296×10^{-8} V $\sqrt{\text{Hz}}$.

This noise voltage generated from the tuning diode is now multiplied with the VCO gain K_0, resulting in the rms frequency deviation

$$(\Delta f_{\text{rms}}) = K_0 \times (1.296 \times 10^{-8}\ \text{V}) \text{ in 1-Hz bandwidth} \tag{5-119}$$

To translate this into an equivalent peak phase deviation,

$$\theta_d = \frac{K_0\sqrt{2}}{f_m}(1.296 \times 10^{-8}) \text{ rad in 1-Hz bandwidth} \tag{5-120}$$

or for a typical oscillator gain of 100 kHz/V,

$$\theta_d = \frac{0.00183}{f_m} \text{ rad in 1-Hz bandwidth} \tag{5-121}$$

For f_m = 25 kHz (typical spacing for adjacent-channel measurements for FM mobile radios), $\theta_c = 7.32 \times 10^{-8}$. This can be converted now into the SSB signal-to-noise ratio:

$$\mathcal{L}(f_m) = 20 \log_{10}\left(\frac{\theta_c}{2}\right) = -149 \text{ dBc/Hz} \tag{5-122}$$

For the typical oscillator gain of 10 MHz/V found in wireless applications, the resulting phase noise will be 20 dB worse [10 log (10 MHz ÷ 100 kHz)]. However, the best tuning diodes, like the BB104, have an R_n of 200 Ω instead of 10 kΩ, which again changes the picture. According to Eq. (5-118), with $kT_0 = 4.2 \times 10^{-21}$ the resulting noise voltage will be

$$V_n = \sqrt{4 \times 4.2 \times 10^{-21} \times 200}$$

$$= 1.833 \times 10^{-9} \text{ V}\sqrt{\text{Hz}} \tag{5-123}$$

From Eq. (5-114), the equivalent peak phase deviation for a gain of 10 MHz/V in a 1-Hz bandwidth is then

$$\theta_d = \frac{1 \times 10^7 \sqrt{2}}{f_m} (1.833 \times 10^{-9}) \text{ rad} \tag{5-124}$$

or

$$\theta_d = \frac{0.026}{f_m} \text{ rad in 1-Hz bandwidth} \tag{5-125}$$

With $f_m = 25$ kHz, $\theta_c = 1.04 \times 10^{-6}$. Expressing this as phase noise:

$$\mathcal{L}(f_m) = 20 \log_{10}\left(\frac{\theta_c}{2}\right) = -126 \text{ dBc/Hz} \tag{5-126}$$

The phase-noise value in Eq. (5-122) is that typically achieved in the Rohde & Schwarz SMDU or with the Hewlett-Packard 8640 signal generator and considered state-of-the-art for a free-running oscillator. It should be noted that both signal generators use a slightly different tuned circuit; the Rohde & Schwarz generator uses a helical resonator, whereas the Hewlett-Packard generator uses an electrically shortened quarter-wavelength cavity. Both generators are mechanically pretuned, and a tuning diode with a gain of about 100 kHz/V is used for AFC and frequency-modulation purposes. It is apparent that, because of the nonlinearity of the tuning diode, the gain is different for low dc voltages than for high dc voltages. The impact of this is that the noise varies with the tuning range.

If this oscillator had to be used for a synthesizer, its 1-MHz tuning range would be insufficient; therefore, a way had to be found to segment the band into the necessary ranges. In VCOs, this is typically done with switching diodes that allow the proper frequency bands to be selected. These switching diodes insert in parallel or series, depending on the circuit, with additional inductors or capacitors, depending on the design.

In low-energy-consuming circuits, the VCO frequently is divided into a coarse-tuning section using tuning diodes and a fine-tuning section with a tuning diode. In the coarse-tuning range, this results in very high gains, such as 1–10 MHz/V, for the diodes. The noise contribution of those diodes is therefore very high and can hardly be compensated by the loop. For low-noise applications, which automatically mean higher power consumption, it is unavoidable to use switching diodes.

Let us now examine some test results. Referring back to Eq. (5-106), Figure 5-71 shows the noise sideband performance as a function of Q, whereby the top curve with $Q_L = 100$ represents a somewhat poor oscillator and the lowest curve with $Q_L = 100{,}000$ probably represents a crystal oscillator where the unloaded Q of the crystal was in the vicinity of 3×10^6. Figure 5-72 shows the influence of flicker noise.

Corner frequencies of 10 Hz to 10 kHz have been selected, and it becomes apparent that around 1 kHz the influence is fairly dramatic, whereas the influence at 20 kHz off the carrier

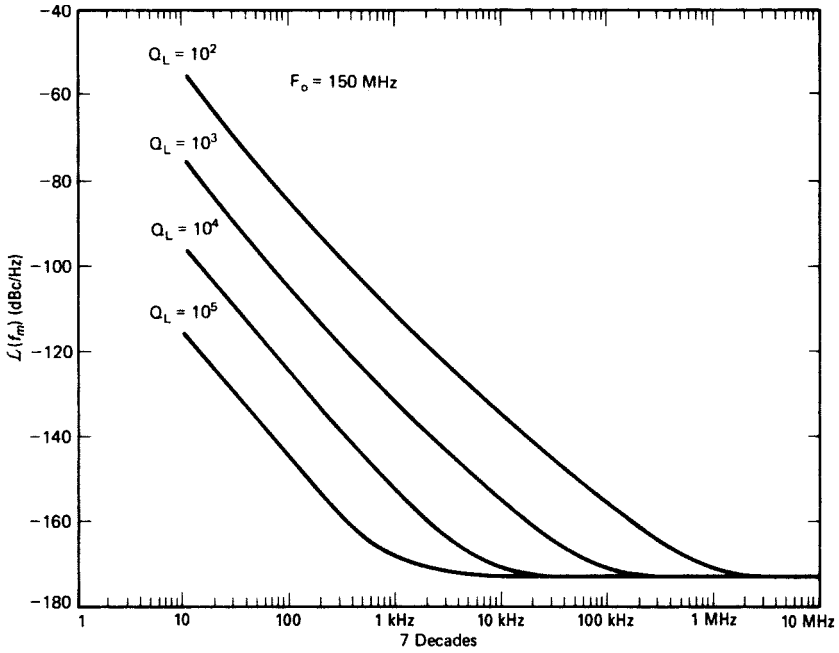


Figure 5-71 Noise sideband of an oscillator at 150 MHz as a function of the loaded Q of the resonator.

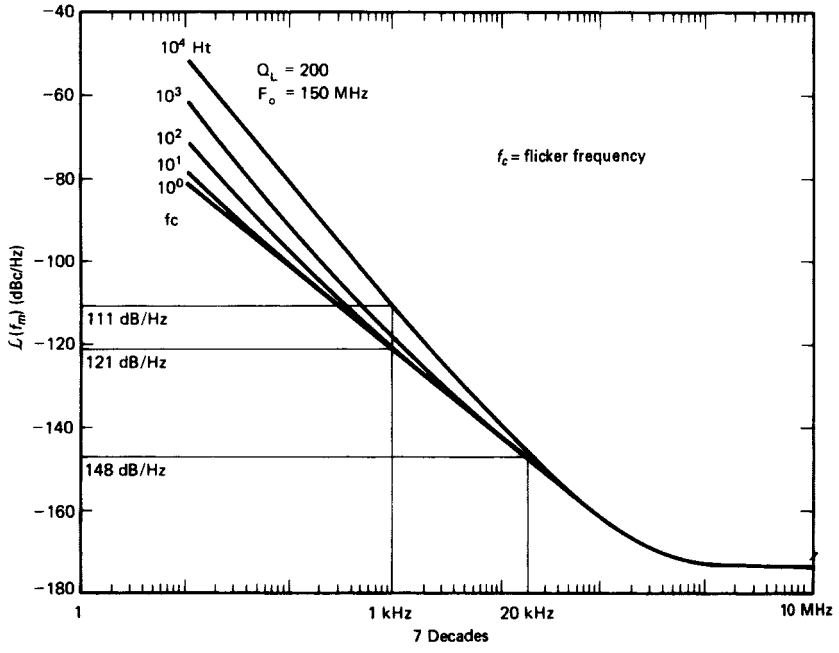


Figure 5-72 Noise-sideband performance as a function of the flicker frequency f_c varying from 10 Hz to 10 kHz.

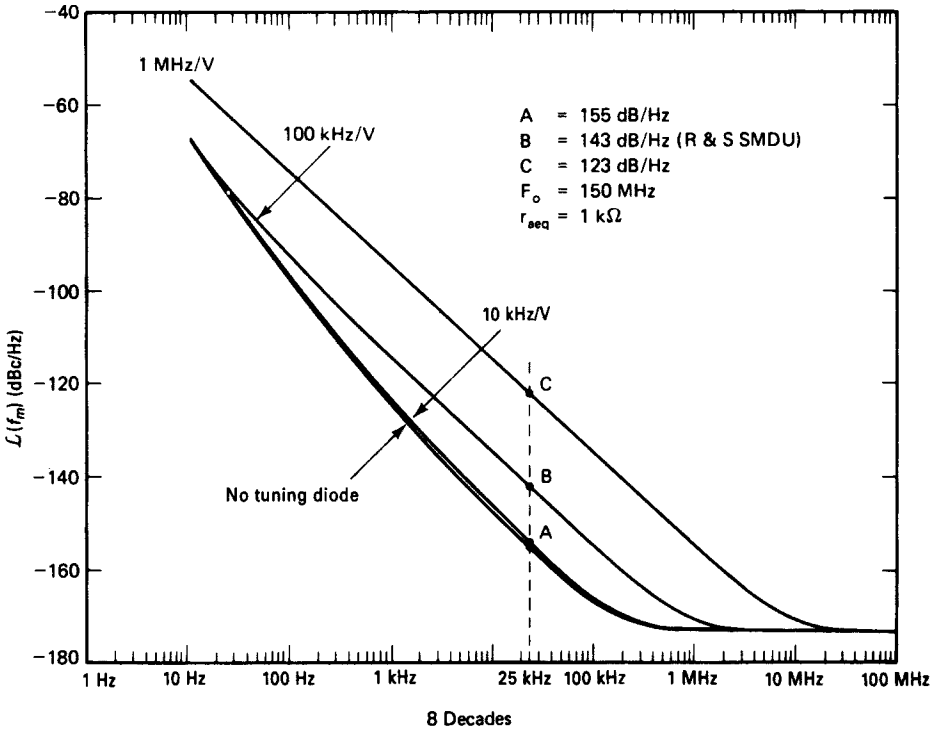


Figure 5-73 Noise-sideband performance of an oscillator at 150 MHz, showing the influence of various tuning diodes.

is not significant. Figures 5-73 and 5-74 show the influence of the tuning diodes on a high-Q oscillator.

Curve A in Figure 5-73 uses a lightly coupled tuning diode with a K_0 of 10 kHz/V; the lower curve is the noise performance without any diode. As a result, the two curves are almost identical, which can be seen from the somewhat smeared form of the graph. Curve B shows the influence of a tuning diode at 100 kHz/V and represents a value of −143 dBc/Hz from −155 dBc/Hz, already some deterioration. Curve C shows the noise if the tuning diode results in a 1-MHz/V VCO gain, and the noise sideband at 25 kHz has now deteriorated to −123 dBc/Hz. These curves speak for themselves.

Figure 5-74 shows how the phase noise of a 147-MHz VCO varies with the bias applied to its tuning diode. Because a tuning diode's Q increases with the reverse bias applied to it, we expect a VCO's phase-noise performance to improve, even if only slightly, as its tuning voltage is increased. Yet Figure 5-74 illustrates the reverse of this relationship. Inspecting this VCO's circuit (Figure 5-75), we see that the diode is in series with the circuit's 220-nH resonator. *All* of the RF current flowing through the resonator must flow through the relatively low-Q diode, and even small decreases in the diode's capacitance will significantly modify the tank's L/C ratio.

Graphing the diode's ac load line (Figure 5-76), we see that as the tuning voltage increases (i.e., as the tuning diode capacitance decreases), the RF voltage across the diode becomes disproportionately high. We note also that at every value of tuning voltage, the RF voltage across the diode goes positive (relative to the anode) during part of each cycle.

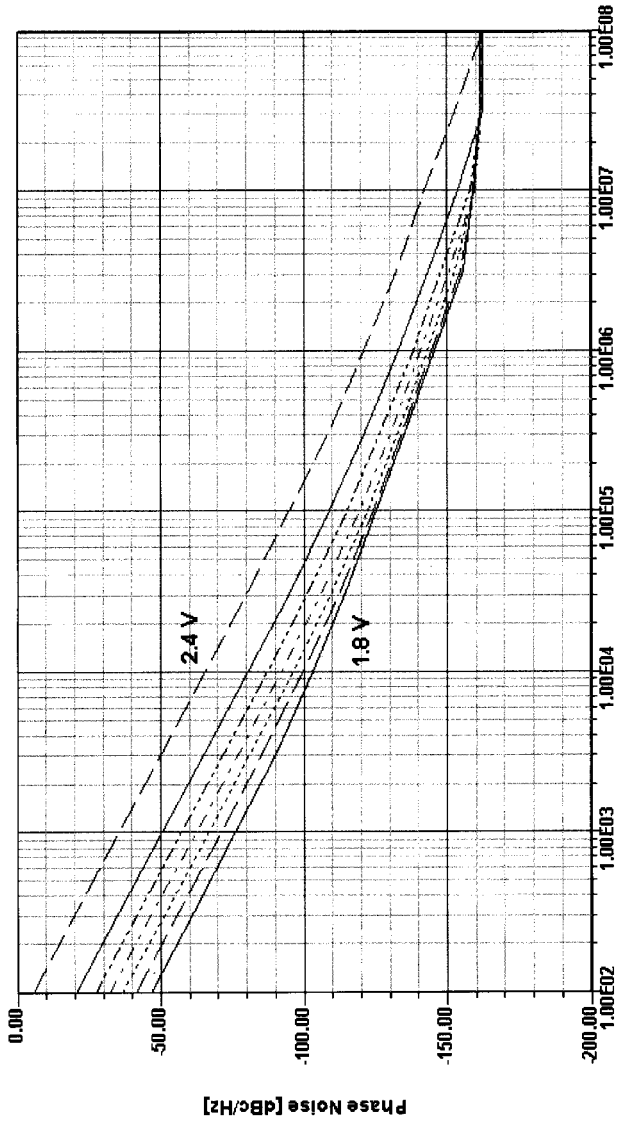


Figure 5-74 Noise-sideband performance of an oscillator at 147 MHz, showing the influence of a tuning diode operated at bias voltages from 1.8 to 2.4 V in 0.1-V increments.

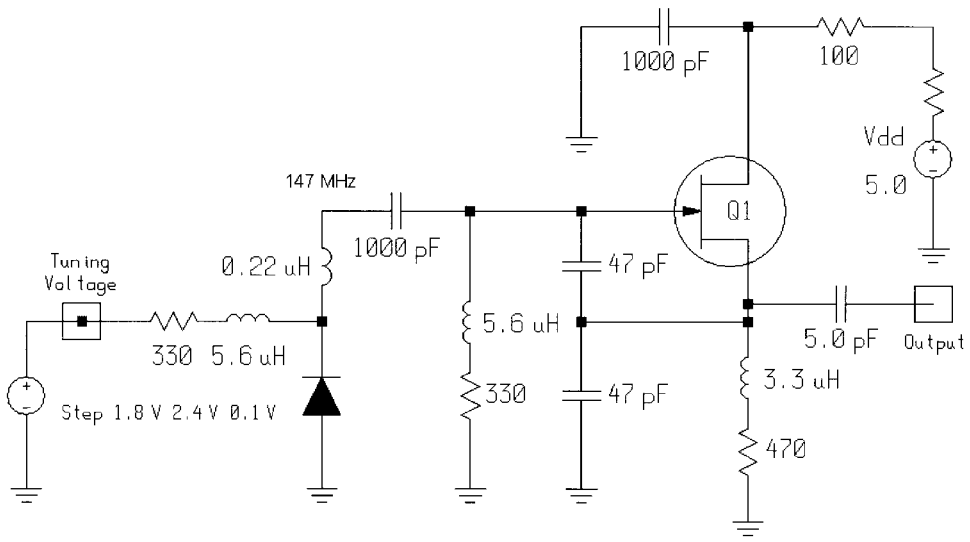


Figure 5-75 The VCO schematic. Because the tuning diode is in series with the resonator, it passes all of the tank's RF current and plays a disproportionately large role in determining the circuit's phase-noise performance.

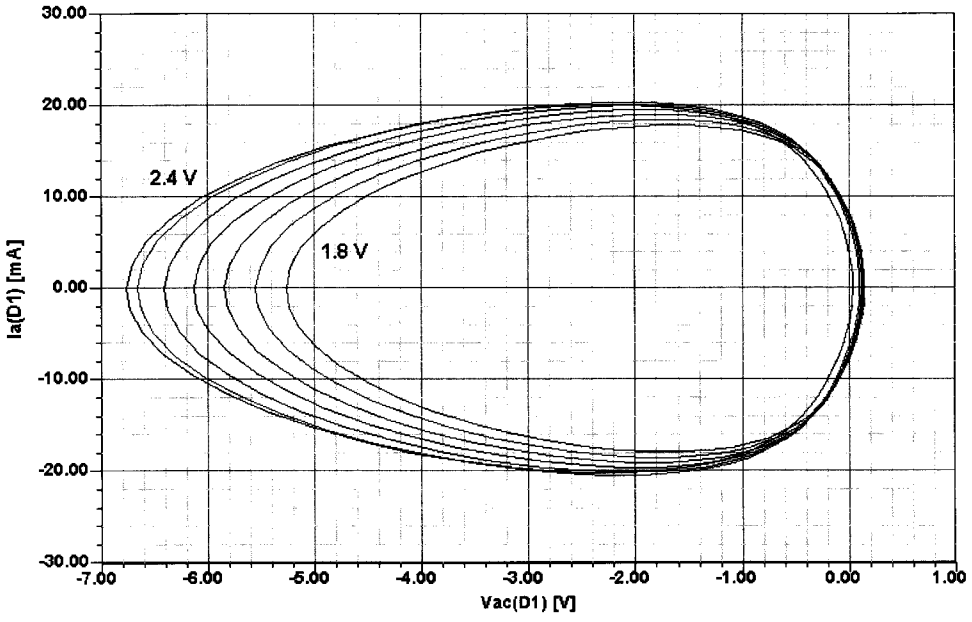


Figure 5-76 Graphing the tuning diode's ac load line reveals that increasing the diode's tuning bias (decreasing its capacitance) results in a disproportionally large increase in the RF voltage across the diode. At all tuning voltages, the RF voltage across the diode is slightly positive (relative to the anode) during part of each cycle.

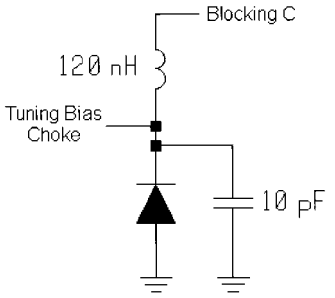


Figure 5-77 Adding capacitance across the tuning diode reduces the effect of its nonlinearities on phase-noise performance. We have also revalued the resonator to 120 nH to maintain oscillation near 147 MHz.

Modifying the oscillator tank to add capacitance in parallel with the tuning diode (Figure 5-77) reduces the tuning diode's effect on tank Q and L/C ratio, and therefore its contribution to phase noise. With this arrangement, we now see the expected relationship between tuning bias and phase noise: The higher the tuning bias, the lower the phase noise (Figure 5-78). Figure 5-79 shows that the diode's ac load line is now virtually unchanged across the tuning-voltage range. The RF voltage across the diode still goes positive (relative to the anode) during part of each cycle.

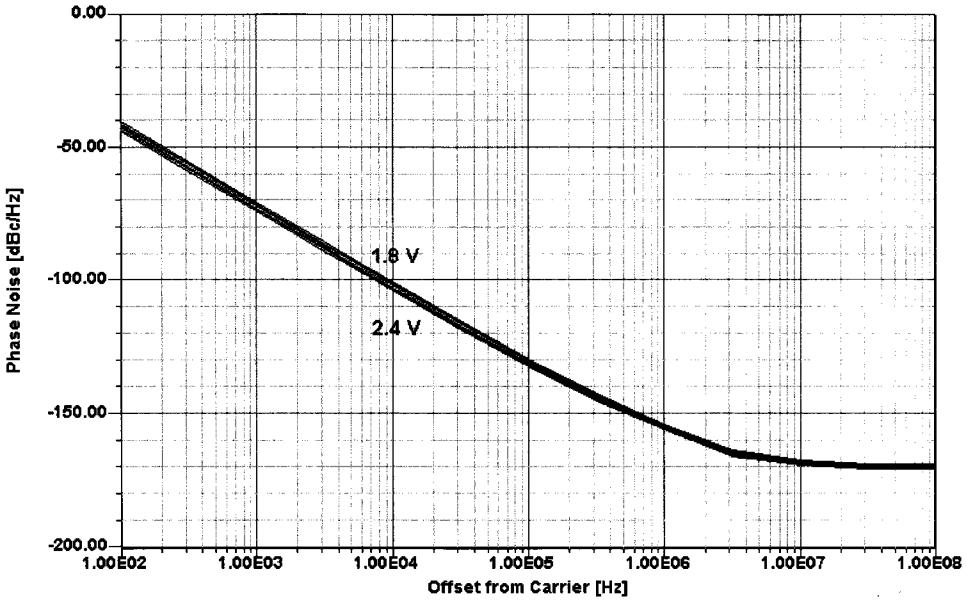


Figure 5-78 Now that the tuning diode passes only part of the tank's RF current, varying its capacitance does not result in the large change in phase-noise performance shown in Figure 5-74. Note also that the phase-noise-versus-tuning-bias relationship has reversed relative to that shown in Figure 5-74: The highest tuning voltage now corresponds to the lowest phase noise.

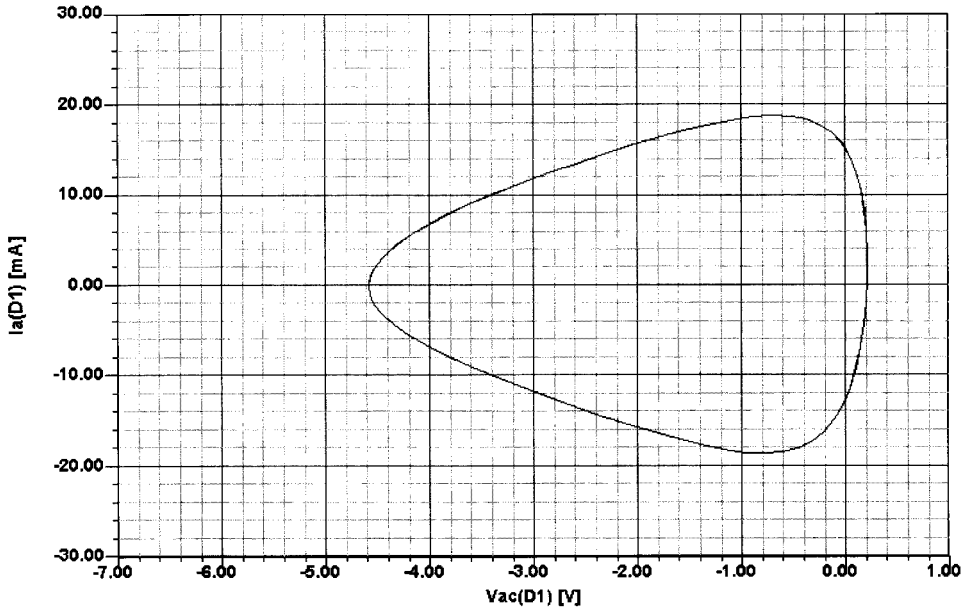


Figure 5-79 The ac load line for the tuning diode in Figure 5-77. Now the voltage swing across the diode is virtually the same at all tuning voltages. The RF voltage across the diode still goes positive (relative to the anode) during part of the cycle.

Figure 5-80 shows one more modification to the VCO tank. Now, the tuning diode is lightly coupled to the hot end of the tank, in parallel with the resonator. The result is phase-noise performance that is essentially unchanged across the 1.8–2.4-V tuning-voltage range. Figure 5-81 compares this circuit's phase-noise performance with the worst-case performance of the tank configurations in Figures 5-75 and 5-77. Figure 5-82 compares the diode ac load lines that correspond to the phase-noise curves in Figure 5-81. The best phase-noise performance corresponds to the diode ac load line that exhibits the narrowest voltage swing and never goes positive relative to the anode.

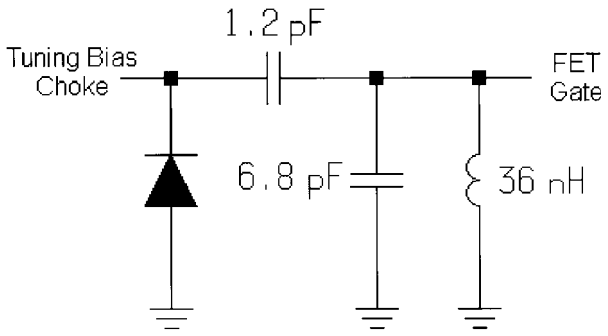


Figure 5-80 Tank circuit reconfigured as per Figure 5-54 to lightly couple the tuning diode in parallel with the resonator, with the resonator inductance again decreased to maintain oscillation near 147 MHz. The 1000-pF capacitor and gate choke of Figure 5-75 (5.6 μH in series with 330 Ω) are no longer needed.

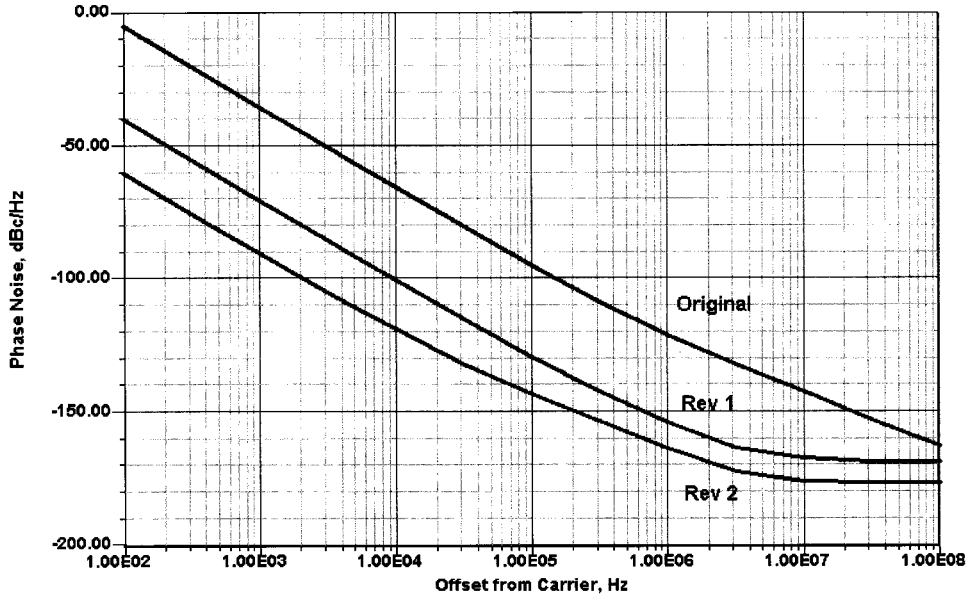


Figure 5-81 Comparison of worst-case phase-noise performance obtained with the tank configurations of Figures 5-75 (Original), 5-77 (Rev 1), and 5-80 (Rev 2).

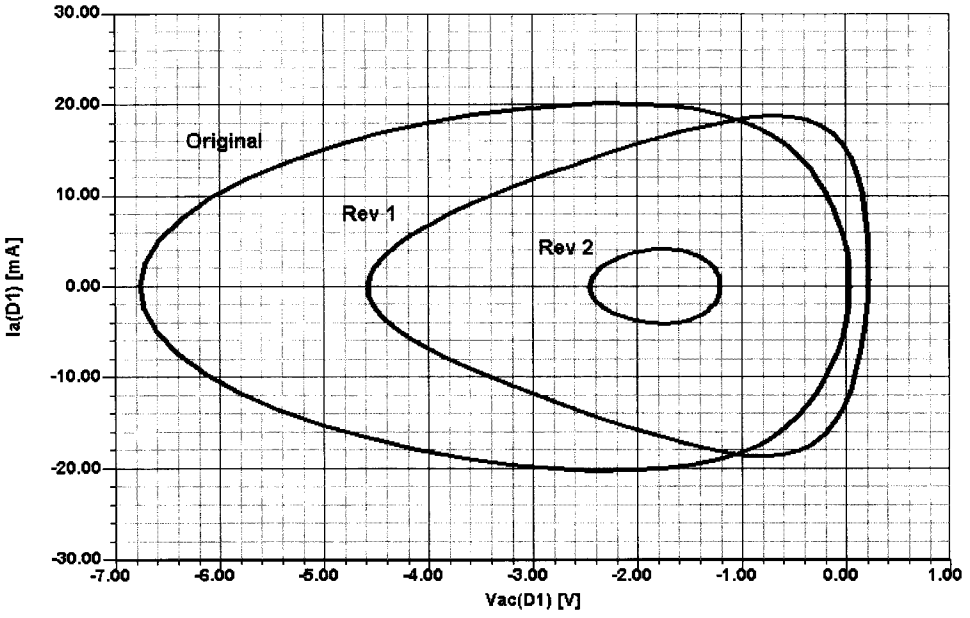


Figure 5-82 Comparison of the tuning diode's ac load line for the worst-case phase-noise curves of Figure 5-79.

It's important to keep in mind that we have not gained something for nothing in achieving the best phase-noise response in Figure 5-81. As we reduced the diode's phase-noise contribution, we also reduced the oscillator's tuning range—from 7.1% (Figure 5-75 circuit) to 1.8% (Figure 5-77 circuit) to 0.046% (Figure 5-80). Simultaneously optimizing phase-noise performance and tuning range requires a combination of techniques. First, we would maximize diode Q by using as high a tuning voltage as feasible. Second, we would use antiseries-connected diodes to minimize the RF voltage swing across them (Figure 5-55). Both of these techniques work against the achievement of wide tuning ranges by minimizing the capacitance swing that the diodes can contribute. As a result of this, and because hyperabrupt diodes are not available for the frequencies at which wireless VCOs operate, we would therefore use multiple antiseries diode pairs in parallel.

It is of interest to compare various oscillators. Figure 5-83 shows the performance of a 10-MHz crystal oscillator, a 40-MHz *LC* oscillator, the HP8640 cavity-tuned oscillator at 500 MHz, the 310–640-MHz switched-reactance oscillator of the HP8662 oscillator, and a 2–6-GHz YIG oscillator at 6 GHz.

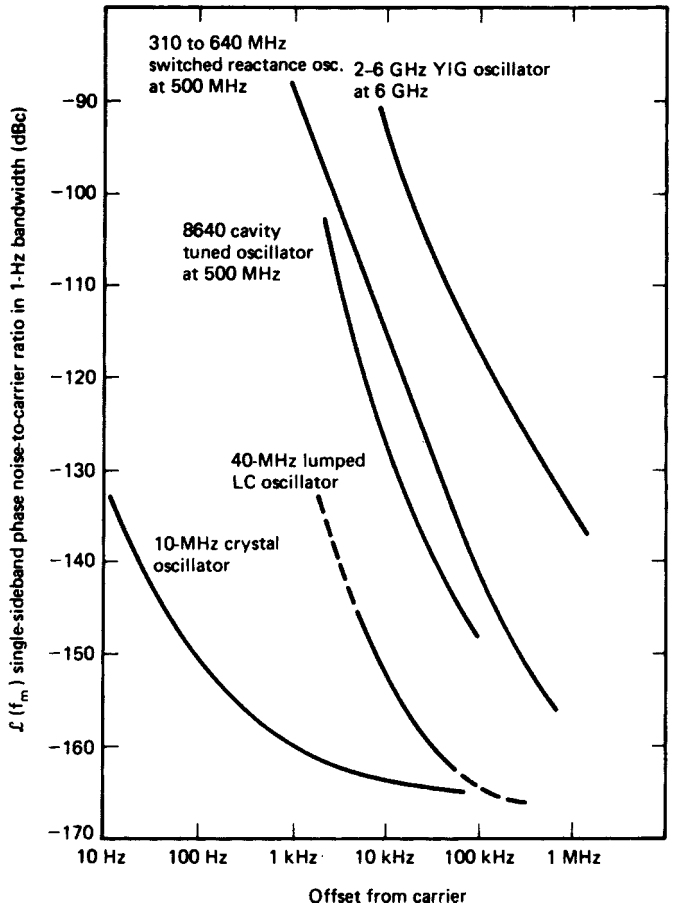


Figure 5-83 Comparison of noise-sideband performances of a crystal oscillator, *LC* oscillator, cavity-tuned oscillator, switched-reactance oscillator, and YIG oscillator.

5-6-3 Nonlinear Approach to the Calculation of Oscillator Phase Noise

The mechanism of noise generation in an oscillator combines the equivalent of frequency conversion (mixing) with the effect of AM-to-PM conversion. Therefore, to calculate oscillator phase noise, we must first be able to calculate the noise figure of a mixer. The following nonlinear approach is based on Ulrich L. Rohde, Chao-Ren Chang, and Jason Gerber, "Design and Optimization of Low-Noise Oscillators Using Nonlinear CAD Tools," *IEEE Frequency Control Symposium Proceedings*, 1994, pp. 548–554. This section presents the use of a novel algorithm for the computation of SSB carrier noise in free-running oscillators using the harmonic-balance (HB) nonlinear technique.

Traditional approaches relying on frequency conversion analysis are not sufficient to describe the complex physical behavior of a noisy oscillator. The accuracy of this nonlinear approach is based on the dynamic range of the harmonic-balance simulator and the quality of the parameter extraction for the active device. The algorithm described has also been verified with several examples up to millimeter wavelengths. This is the only algorithm that provides a complete and rigorous treatment of noise analysis for autonomous circuits.

Noise Generation in Oscillators. As shown earlier, the qualitative linearized picture of noise generation in oscillators is very well known [9]. The physical effects of random fluctuations taking place in the circuit are different depending on their spectral allocation with respect to the carrier:

- Noise components at low-frequency deviations result in frequency modulation of the carrier through mean square frequency fluctuation proportional to the available noise power.
- Noise components at high-frequency deviations result in phase modulation of the carrier through mean square phase fluctuation proportional to the available noise power.

We will demonstrate that the same conclusions can be derived quantitatively from the HB equations for an autonomous circuit.

Equivalent Representation of a Noisy Nonlinear Circuit. A general noisy nonlinear network can be described by the equivalent circuit shown in Figure 5-84. The circuit is divided into linear and nonlinear subnetworks as noise-free multiports. Noise generation is accounted for by connecting a set of noise voltage and noise current sources at the ports of the linear subnetwork [9–18].

Frequency Conversion Approach. The circuit supports a large-signal time-periodic steady state of fundamental angular frequency ω_0(carrier). Noise signals are small perturbations superimposed on the steady state, represented by families of pseudosinusoids located at the sidebands of the carrier harmonics. Therefore, the noise performance of the circuit is determined by the exchange of power among the sidebands of the unperturbed steady state through frequency conversion in the nonlinear subnetwork. Due to the perturbative assumption, the nonlinear subnetwork can be replaced with a multifrequency linear multiport described by a conversion matrix. The flow of noise signals can be computed by means of conventional linear circuit techniques.

The frequency conversion approach frequently used has limitations. The frequency-conversion approach is not sufficient to predict the noise performance of an autonomous

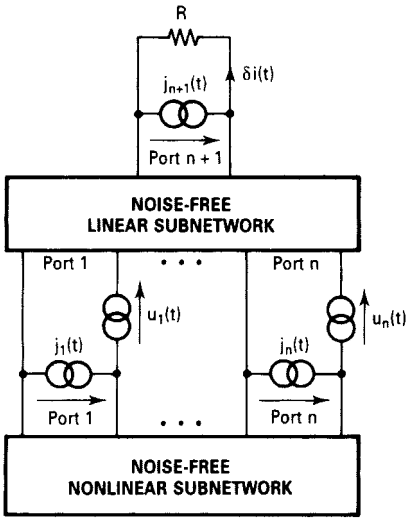


Figure 5-84 Equivalent circuit of a general noisy nonlinear network.

circuit. The spectral density of the output noise power and, consequently, the PM noise computed by the conversion analysis are proportional to the available power of the noise sources.

- In the presence of both thermal and flicker noise sources, PM noise increases as ω^{-1} for $\omega \to 0$ and tends to a finite limit for $\omega \to \infty$.
- Frequency conversion analysis correctly predicts the far-carrier noise behavior of an oscillator and, in particular, the oscillator noise floor; it does not provide results consistent with the physical observations at low deviations from the carrier.

This inconsistency can be removed by adding the modulation noise analysis. In order to determine the far-away noise using the autonomous circuit perturbation analysis, the following applies.

The circuit supports a large-signal time-periodic autonomous regime. The circuit is perturbed by a set of small sources located at the carrier harmonics and at the sidebands at a deviation ω from carrier harmonics. The perturbation of the circuit state (δX_B, δX_H) is given by the uncoupled sets of equations:

$$\frac{\partial E_B}{\partial E_B}\,\delta X_B = J_B(\omega) \tag{5-127}$$

$$\frac{\partial E_H}{\partial E_H}\,\delta X_H = J_H(\omega) \tag{5-128}$$

where E_B, E_H = vectors of HB errors
X_B, X_H = vectors of state-variable (SV) harmonics (since the circuit is autonomous, one of the entries X is replaced by the fundamental frequency ω_0)
J_B, J_H = vectors of forcing terms

The subscripts *B* and *H* denote sidebands and carrier harmonics, respectively.

For a spot noise analysis at a frequency ω, the noise sources can be interpreted in either of two ways:

- Pseudosinusoids with random amplitude and phase located at the sidebands. This noise generation mechanism is described by Eq. (5-127), which is essentially a frequency conversion equation relating the sideband harmonics of the state variables and of the noise sources. This description is exactly equivalent to the one provided by the frequency conversion approach. This mechanism is referred to as *conversion noise* [19–21].
- Sinusoids located at the carrier harmonics and randomly phase and amplitude modulated by pseudosinusoidal noise at frequency ω. This noise generation mechanism is described by Eq. (5-128), which describes noise-induced jitter of the circuit-state, represented by the vector δX_H. The modulated perturbing signals are represented by replacing the entries of J_H with the complex modulation laws. This mechanism is referred to as *modulation noise.* One of the entries of δX_H is $\delta\omega_0$, where $\delta\omega_0(\omega)$ is the phasor of the pseudosinusoidal components of the fundamental frequency fluctuations in a 1-Hz band at frequency ω. Equation (5-128) provides a frequency jitter with a mean square value proportional to the available noise power. In the presence of both thermal and flicker noise, PM noise increases as ω^{-3} for $\omega \to 0$ and tends to 0 for $\omega \to \infty$. Modulation noise analysis correctly describes the noise behavior of an oscillator at low deviations from the carrier and does not provide results consistent with physical observations at high deviations from the carrier.

The combination of both phenomena explains the noise in the oscillator shown in Figure 5-85, where the near-carrier noise dominates below ω_X and far-carrier noise dominates above ω_X. Figure 5-86 (itemized form) shows the noise sources as they are applied at the IF. We have arbitrarily defined the low oscillator output as IF. This applies to the conversion matrix calculation.

Figure 5-87 shows the contributions that have to be taken into consideration for calculation of the noise at the output. The accuracy of the calculation of the phase noise depends highly on the quality of the parameter extraction for the nonlinear device; in particular, high-frequency phenomena must be modeled properly. In addition, the flicker noise contribution is essential. This is also valid for mixer noise analysis.

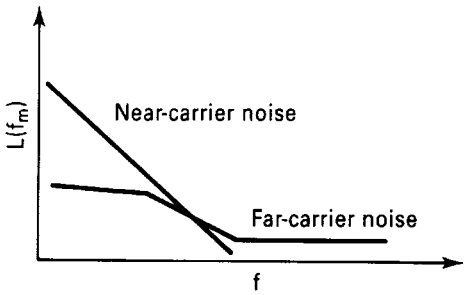


Figure 5-85 Oscillator noise components.

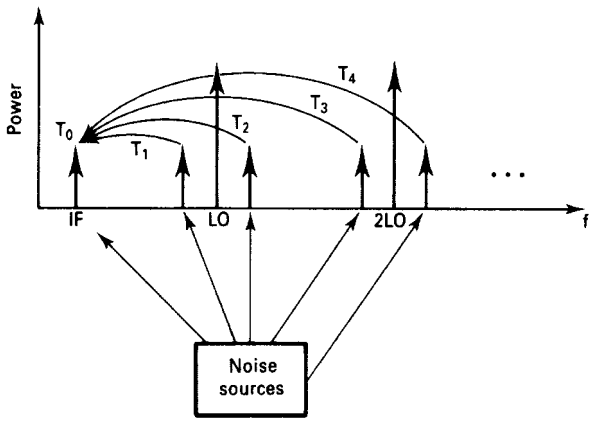


Figure 5-86 Noise sources where the noise at each sideband contributes to the output noise at the IF through frequency conversion.

Figure 5-88 shows the measured phase noise at 10 kHz of the carrier of an 800-MHz oscillator and the simulated result. The good agreement between measured and predicted data is worth noticing. At currents above 5 mA, the simulation becomes inaccurate as breakdown effects occur that were not part of the simulation.

Conversion Noise Analysis. The actual mathematics used to calculate the noise result (Ansoft Serenade 8.*x*) are as follows.

kth Harmonic PM Noise

$$\langle |\delta\Phi_k(\omega)|^2 \rangle = \frac{N_k(\omega) - N_{-k}(\omega) - 2\,\mathrm{Re}[C_k(\omega)]}{R|I_k^{SS}|^2} \tag{5-129}$$

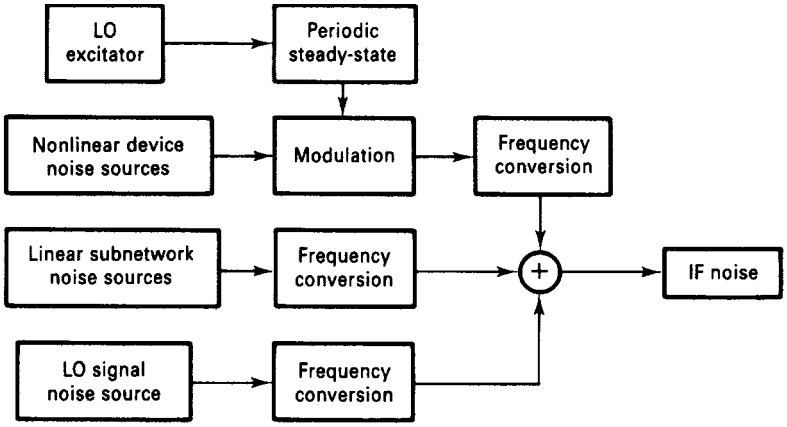


Figure 5-87 Noise mechanisms.

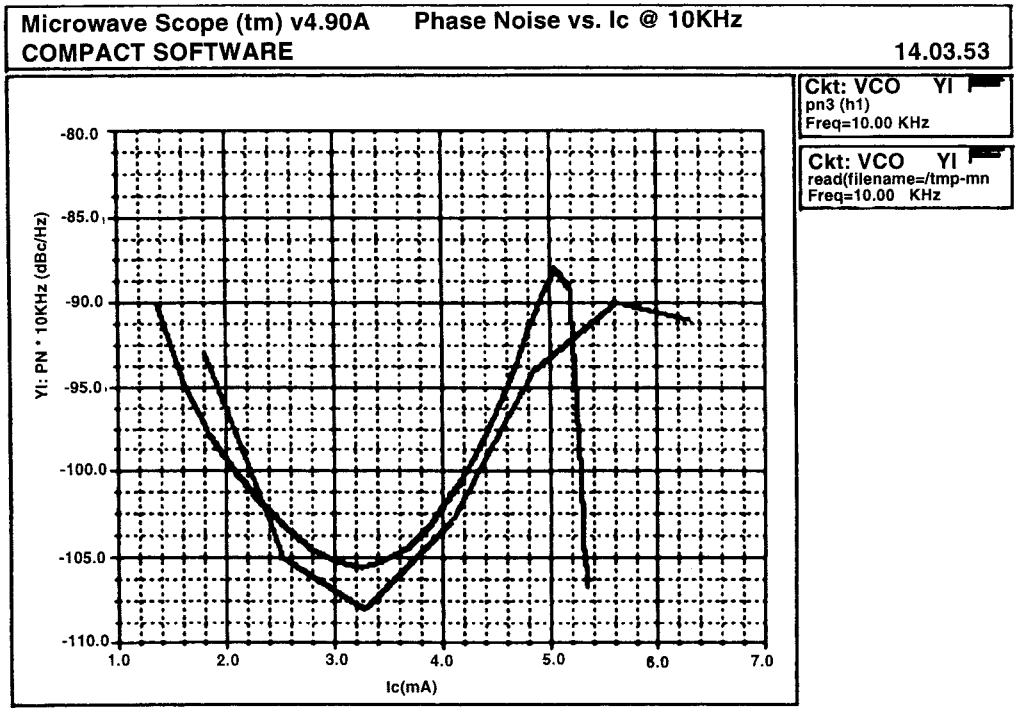


Figure 5-88 Bias-dependent measured and predicted phase noise at 10 kHz off the carrier of an 800-MHz oscillator.

kth Harmonic AM Noise

$$\langle |\delta A_k(\omega)|^2 \rangle = 2\,\frac{N_k(\omega) - N_{-k}(\omega) + 2\,\mathrm{Re}[C_k(\omega)]}{R|I_k^{SS}|^2} \tag{5-130}$$

kth Harmonic PM–AM Correlation Coefficient

$$C_k^{\mathrm{PMAM}}(\omega) = \langle \delta\Phi_k(\omega)\delta A_k(\omega)^* \rangle = -\sqrt{2}\,\frac{2\,\mathrm{Im}[C_k(\omega)] + j[N_k(\omega) - N_{-k}(\omega)]}{R|I_k^{SS|^2}} \tag{5-131}$$

where

$N_k(\omega)$, $N_{-k}(\omega)$ = noise power spectral densities at the upper and lower sidebands of the kth harmonic
$C_k(\omega)$ = normalized correlation coefficient of the upper and lower sidebands of the kth carrier harmonic
R = load resistance
I_k^{SS} = kth harmonic of the steady-state current through the load

Modulation Noise Analysis

kth Harmonic PM Noise

$$\langle |\delta\Phi_k(\omega)|^2 \rangle = \frac{k^2}{\omega^2}\,\mathrm{T}_F\,\langle \mathrm{J}_H(\omega)\mathrm{J}_H^t(\omega)\rangle \mathrm{T}_F^t \tag{5-132}$$

kth Harmonic AM Noise

$$\langle |\delta A_k(\omega)|^2 \rangle = \frac{2}{|I_k^{SS}|^2}\,\mathrm{T}_{Ak}\,\langle \mathrm{J}_H(\omega)\mathrm{J}_H^t(\omega)\rangle \mathrm{T}_{Ak}^t \tag{5-133}$$

kth Harmonic PM–AM Correlation Coefficient

$$C_k^{\mathrm{PMAM}}(\omega) = \langle \delta\Phi_k(\omega)\delta A_k(\omega)^* \rangle = \frac{k\sqrt{2}}{j\omega|I_K^{SS}|^2}\,\mathrm{T}_F\langle \mathrm{J}_H(\omega)\mathrm{J}_H^t(\omega)\rangle \mathrm{T}_{Ak}^t \tag{5-134}$$

where $\mathrm{J}_H(\omega)$ = vector of Norton equivalent of the noise sources
T_F = frequency transfer matrix
I_k^{SS} = kth harmonic of the steady-state current through the load

Experimental Variations. We will look at two examples where we compare predicted and measured data. Figure 5-89 shows the abbreviated circuit of a 10-MHz crystal oscillator. It uses a high-precision, high-Q crystal made by companies such as Bliley. Oscillators like this are intended for use as frequency and low-phase-noise standards. In this case, the circuit under consideration is part of the HP3048 phase-noise measurement system.

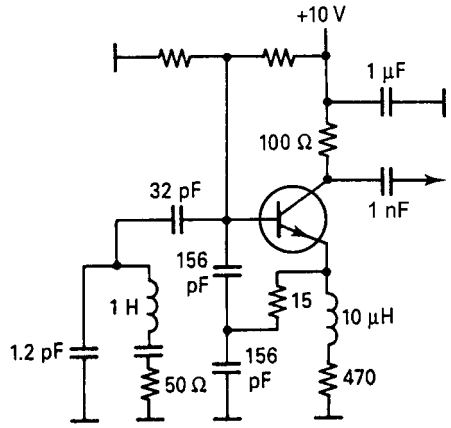


Figure 5-89 Abbreviated circuit of a 10-MHz crystal oscillator.

Figure 5-90 shows the measured phase noise of this HP frequency standard, and Figure 5-91 shows the phase noise predicted using the mathematical approach outlined above.

In cooperation with Motorola, we also analyzed an 800-MHz VCO. In this case, we also did the parameter extraction for the Motorola transistor. Figure 5-92 shows the circuit, a Colpitts oscillator that uses RF feedback in the form of a 15-Ω resistor and a capacitive

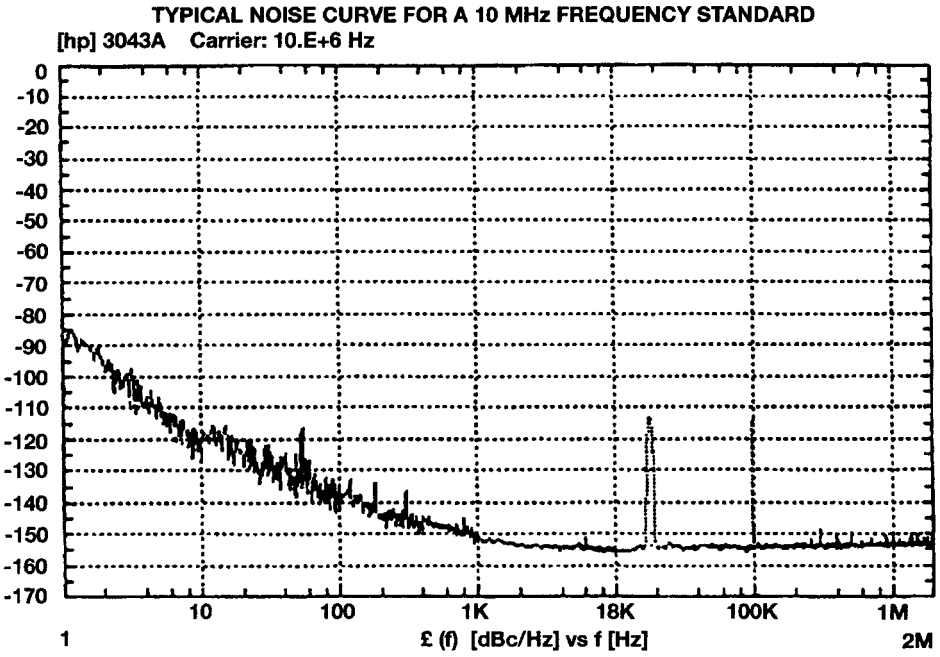


Figure 5-90 Measured phase noise for this frequency standard by HP.

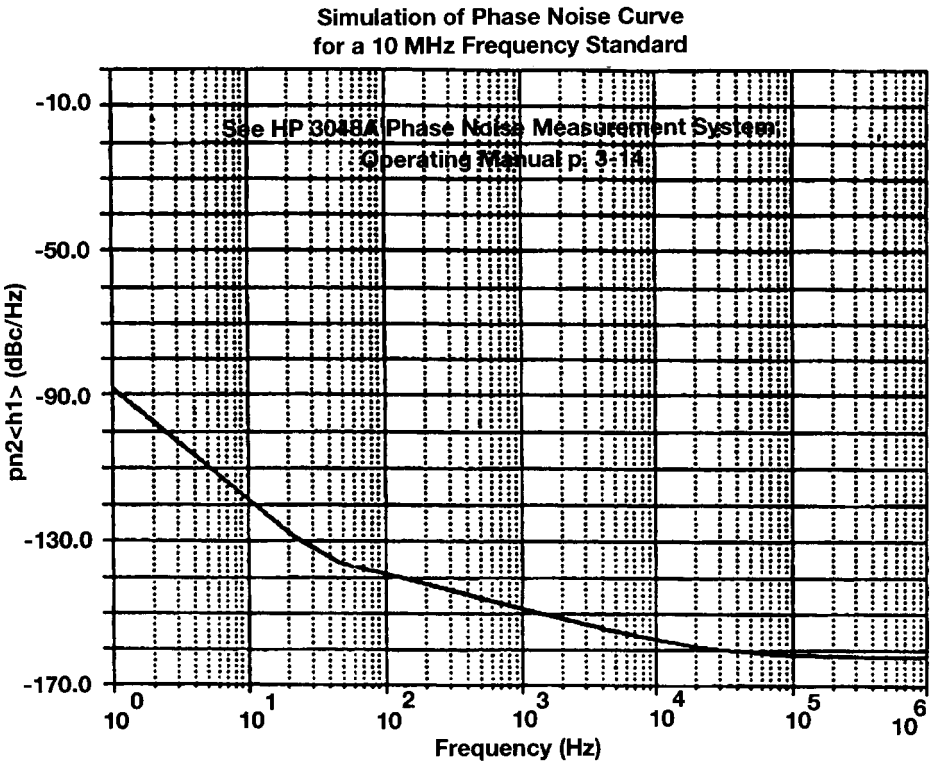


Figure 5-91 Simulated phase noise of the oscillator shown in Figure 5-89.

voltage divider consisting of 1 pF between the BJT's base and the feedback resistor, and 1 pF between the feedback resistor and common. Also, the tuned circuit is loosely coupled to this part of the transistor circuit. Figure 5-93 shows a comparison between predicted and measured phase noise for this oscillator. Figures 5-94, 5-95, and 5-96 show the oscillator's predicted output waveform, spectrum, and dc I–V and ac load line responses, respectively.

Phase Noise Optimization. By allowing the simulator's optimizer to vary the oscillator's feedback and dc operating point, phase noise can be improved. The values that were allowed to vary were the values of the capacitors in the feedback voltage divider, the value of a negative-feedback resistor (starting value, 15 Ω) that reduces AM–PM conversion, and the emitter resistor, which changes the transistor's dc bias and therefore affects its bias-dependent flicker noise. Figure 5-97 shows the improvement of the 800-MHz VCO as previously shown. While the close-in phase noise can be improved drastically by approximately 32 dB, the far-out noise at 20 MHz and beyond is deteriorated. This is due to the resistive feedback, which reduces AM-to-PM conversion [22–24].

In testing an oscillator's resistance to frequency pushing (frequency shift with supply-voltage changes), variations in phase noise can be observed. Phase noise is also bias dependent,

and this fact can be used for phase-noise optimization. Figure 5-98 shows the phase noise of an oscillator as its supply voltage is varied.

Summary. This section has demonstrated the power of harmonic-balance simulation in the design of low-phase-noise oscillators. The simulator can be used to accurately predict close-in and far-out phase noise based on the contribution of the transistor, the loaded Q of the resonator, and, in the case of a VCO, the contribution of the tuning diode.

This is the first time that a bias-dependent noise model for both FETs and bipolar transistors has been developed and used for this type of application, in which optimization is allowed to vary circuit components and the bias point for best performance.

Practical examples ranging from a 10-MHz crystal oscillator to a 1-GHz CR oscillator have been supplied. The theoretical background for the nonlinear oscillator and its validation have been presented.

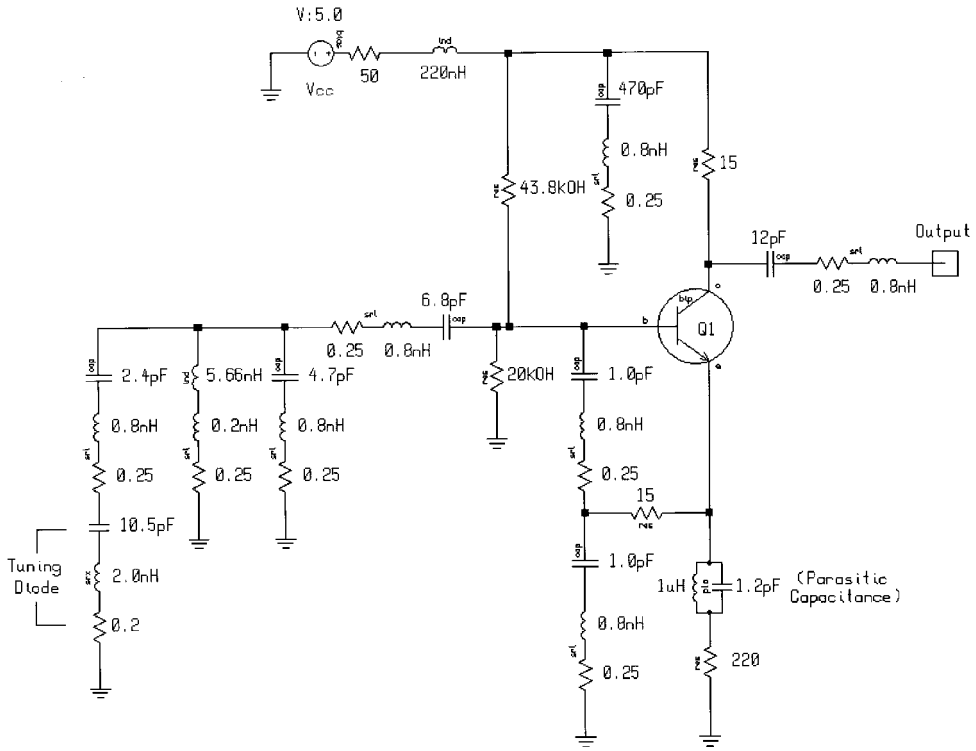


Figure 5-92 Colpitts oscillator that uses RF negative feedback between the emitter and capacitive voltage divider. To be realistic, we used actual components rather than ideal ones. The suppliers for the capacitors and inductors provide some typical values for the parasitics. The major changes are 0.8 nH and 0.25 Ω in series with the capacitors. The same thing applies for the main inductance, which has a parasitic connection inductance of 0.2 nH in series with a 0.25-Ω resistance. These types of parasitics are valid for a fairly large range of components assembled in surface-mount applications. Most engineers model the circuit only by assuming lossy devices, not adding these important parasitics. One of the side-effects we have noticed is that the output power is more realistic and, needless to say, the simulated phase noise agrees quite well with measured data. This circuit can also serve as an example for modeling amplifiers and mixers using surface-mount components.

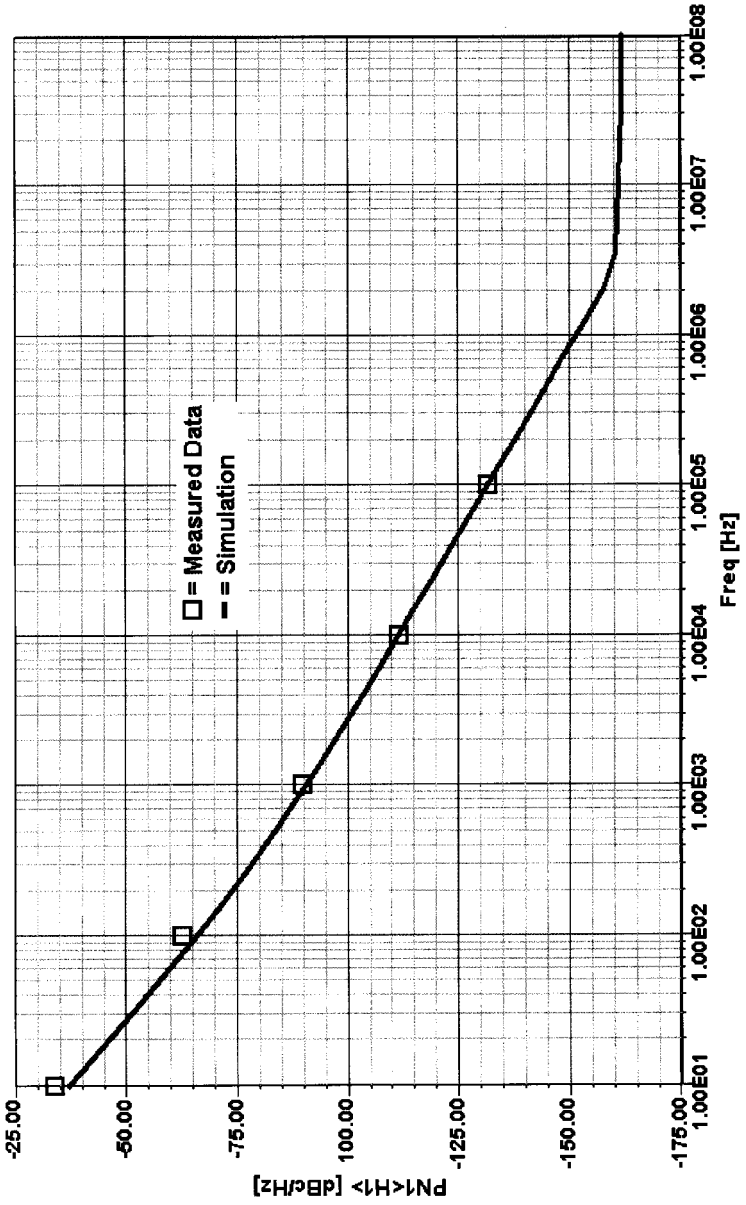


Figure 5-93 Comparison between predicted and measured phase noise for the oscillator shown in Figure 5-92.

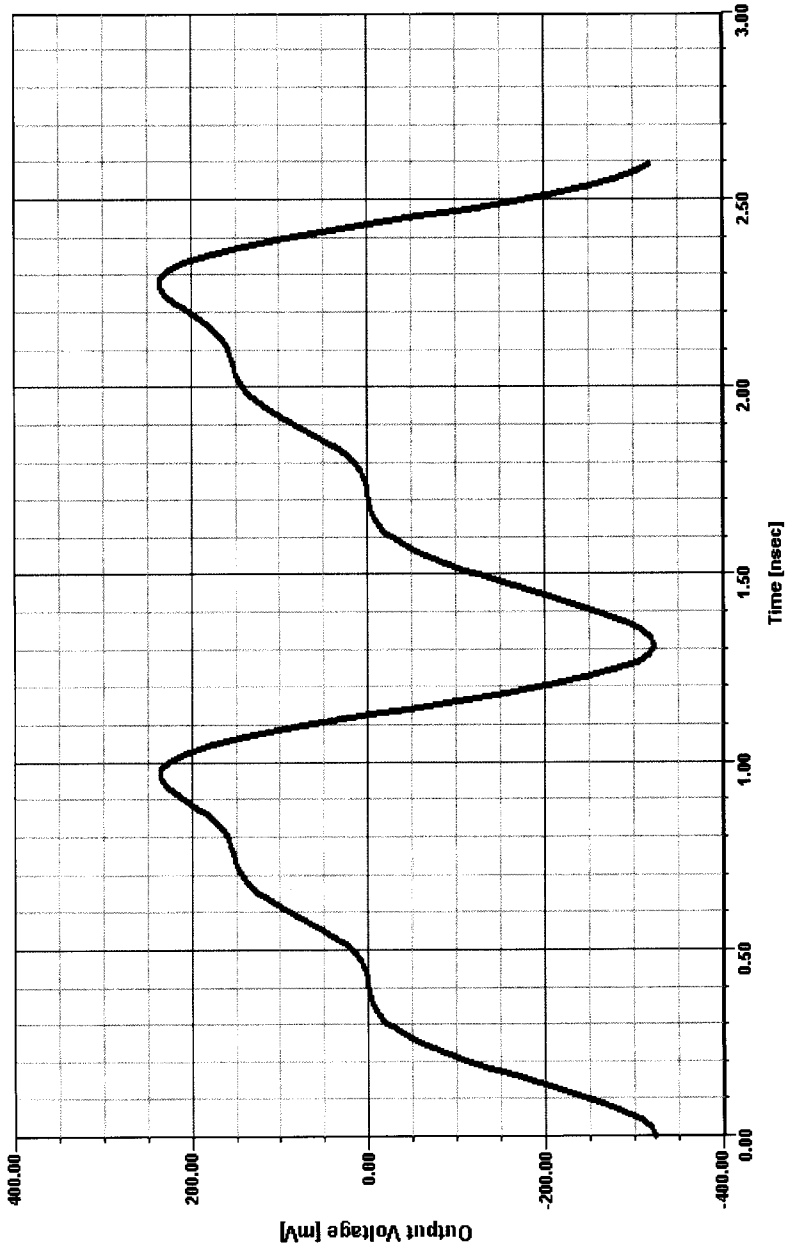


Figure 5-94 Predicted output waveform for the oscillator shown in Figure 5-92.

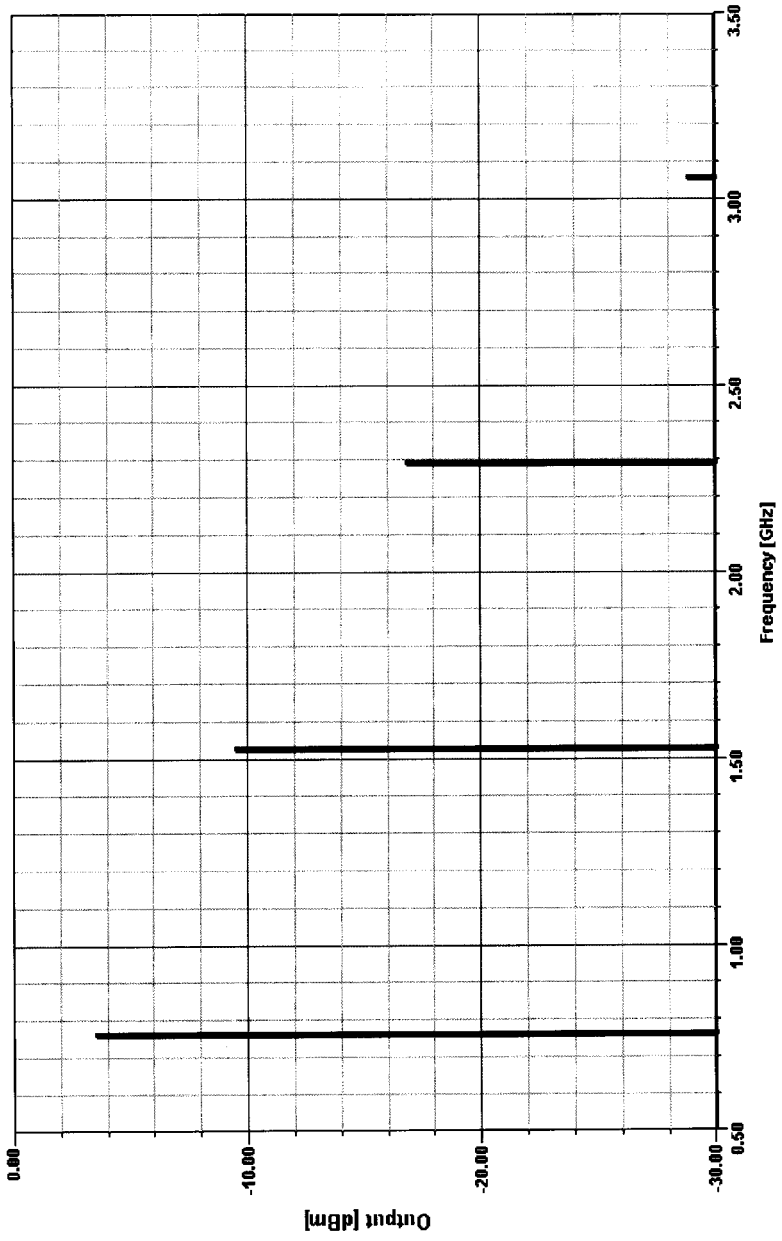


Figure 5-95 Predicted output spectrum for the oscillator shown in Figure 5-92.

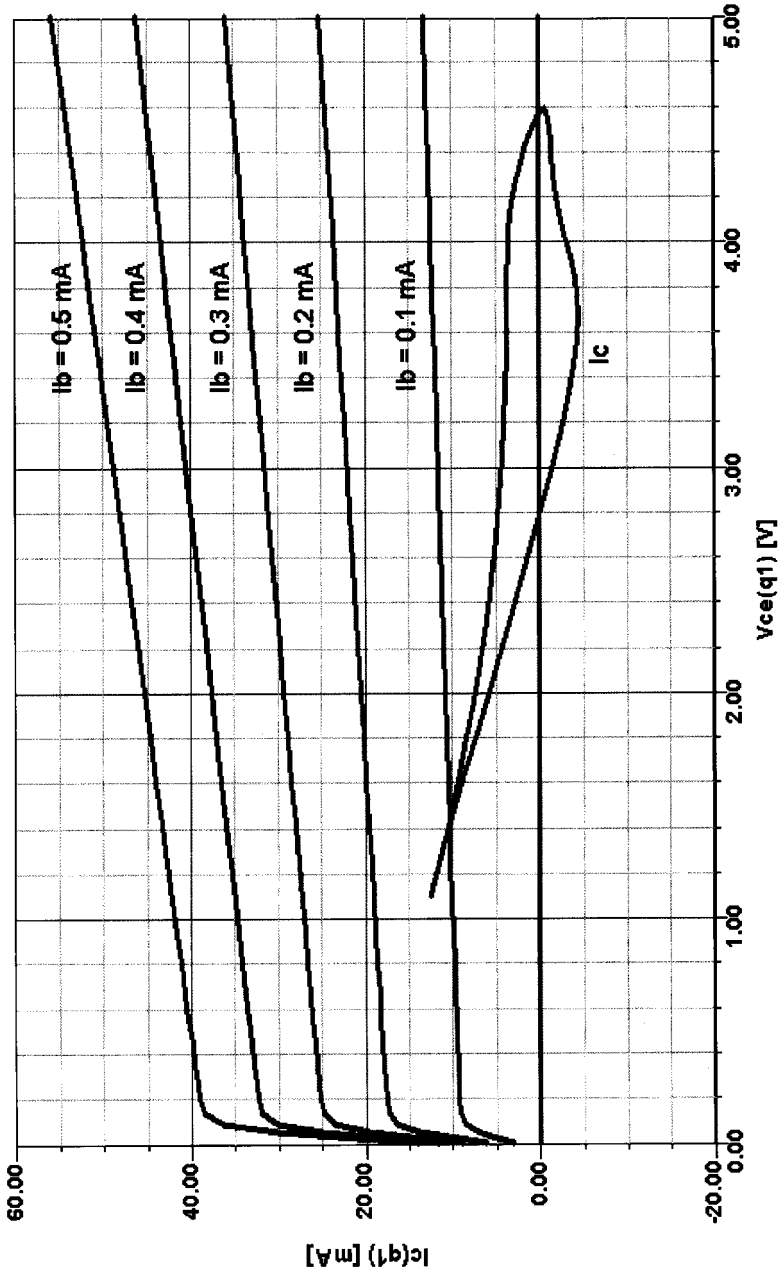


Figure 5-96 The ac load line and dc *I–V* curves for the transistor in the oscillator shown in Figure 5-92.

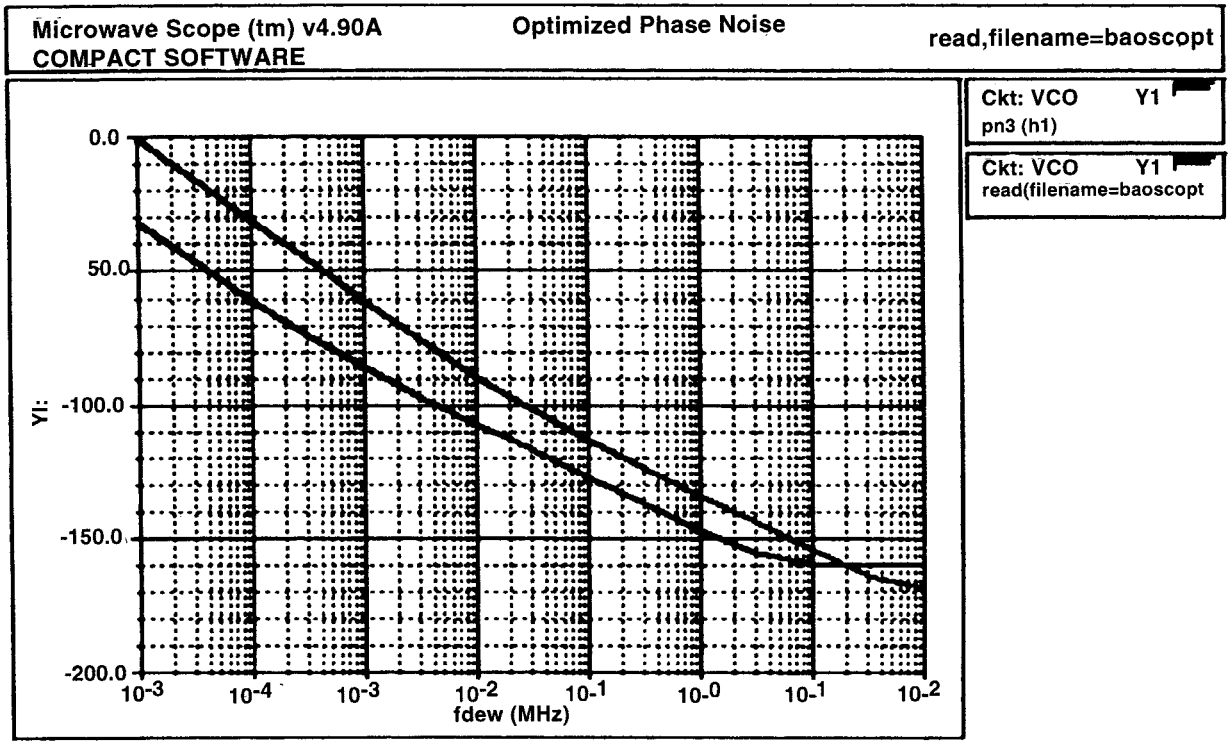


Figure 5-97 Phase noise of the 800-MHz VCO before and after optimization.

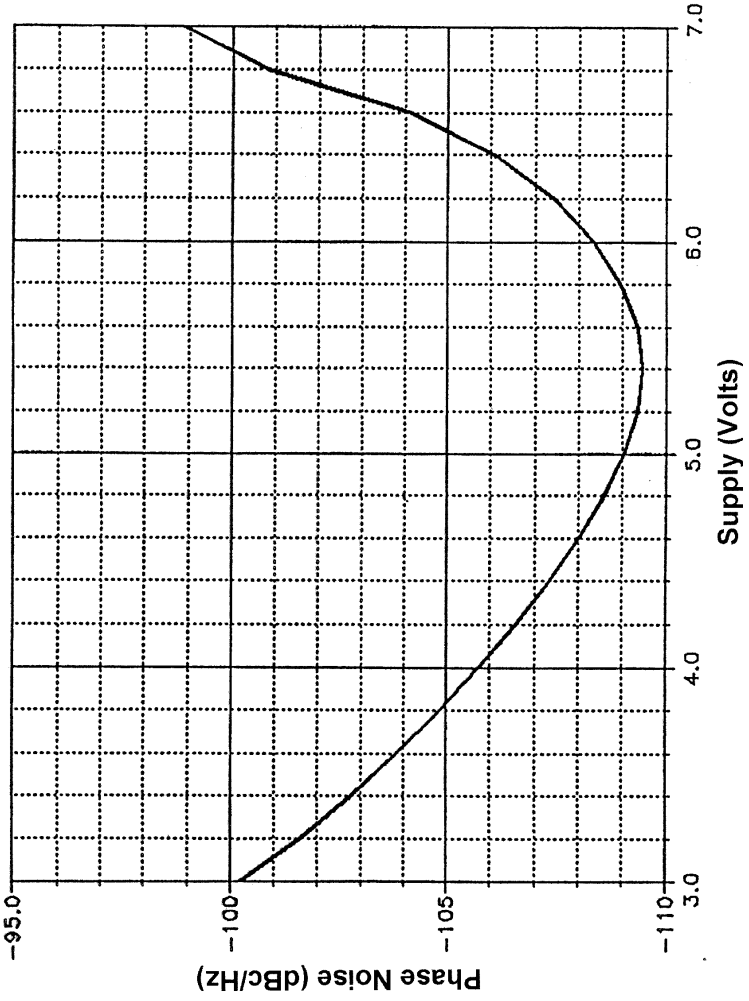


Figure 5-98 Calculated phase noise as a function of the supply voltage for a BJT oscillator. The graph shows a distinct minimum for a particular bias, which could be translated into a collector current change. Such phase-noise minimization could also be accomplished by changing just the voltage applied to the bias portion of the oscillator.

Acknowledgment. This research work for nonlinear operation in oscillators, partly sponsored by the U.S. Department of Defense MIMIC program, was performed in cooperation with Professor Vittorio Rizzoli's group of the Department of Electrical Engineering, Information and Systems Engineering at the University of Bologna, Italy, and the Motorola Portable Products Division, Florida, as well as the engineering team of Ansoft Corporation in New Jersey.

5-7 OSCILLATORS IN PRACTICE

5-7-1 Oscillator Specifications

Although we have tended to emphasize phase noise as a critical indicator of oscillator quality, phase noise is only one of a number of oscillator performance criteria that must be considered in wireless system design. In addition to phase noise, the following characteristics are used by commercial companies to describe oscillator performance:

Frequency Pushing. Frequency pushing characterizes the degree to which an oscillator's frequency is affected by its supply voltage. A sudden current surge caused by activating a transceiver's RF power amplifier may produce a spike on the VCO's dc power supply and a consequent frequency jump. Like tuning sensitivity (below), pushing is specified in frequency/voltage form and is tested by varying the VCO's dc supply voltage (typically ± 1 V) with its tuning voltage held constant.

Harmonic Output Power. The harmonic content is measured relative to the output power. Typical values are 20 dB or more suppression relative to the fundamental. This suppression can be improved by additional filtering.

Output Power. The output power of the oscillator, typically expressed in dBm, is measured into a 50-Ω load. The output power is always combined with a specification for flatness or variation. A typical specification would be 0 dBm ± 1 dB.

Output Power as a Function of Temperature. All active circuits vary in performance as a function of temperature. An oscillator's output power over a temperature range should vary less than a specified value, such as 1 dB.

Post-tuning Drift. After a voltage step is applied to the tuning diode input, the oscillator frequency may continue to change until it settles to a final value. This post-tuning drift is one of the parameters that limits the bandwidth of the VCO input.

Power Consumption. This characteristic conveys the dc power, usually specified in milliwatts and sometimes qualified by operating voltage, required by the oscillator.

Sensitivity to Load Changes. To keep costs down, many wireless applications use a VCO alone, without the buffering action of a high reverse-isolation amplifier stage. In such applications, *frequency pulling*, the change of frequency resulting from partially reactive loads, is an important oscillator characteristic. Pulling is commonly specified in terms of the frequency shift that occurs when the oscillator is connected to a load that exhibits a non-unity VSWR (such as 1.75, usually referenced to 50 Ω), compared to the frequency that results with a unity VSWR load (usually 50 Ω). Frequency pulling must be minimized, especially in cases where power stages are close to the VCO unit

and short pulses may affect the output frequency. Such feedback may make phase locking impossible.

Spurious Outputs. A VCO's spurious output specification, expressed in decibels, enumerates the strength of unwanted and nonharmonically related components relative to the oscillator fundamental. Because a stable, properly designed oscillator is inherently clean, such *spurs* are typically introduced only by external sources in the form of radiated or conducted interference. See *harmonic output power.*

Temperature Drift. Although the synthesizer is responsible for locking and maintaining the oscillator's frequency, the VCO's frequency change as a function of temperature is a critical parameter and must be specified. Its value varies between 10 kHz/°C to several hundred kHz/°C depending on the center frequency and tuning range.

Tuning Characteristic. This specification shows the relationship, depicted as a graph, between a VCO's operating frequency and the tuning voltage applied. Ideally, the correspondence between operating frequency and tuning voltage is linear. See *tuning linearity.*

Tuning Linearity. For stable synthesizers, a constant deviation of frequency versus tuning voltage is desirable. It is also important to make sure that there are no breaks in the tuning range—for example, that the oscillator does not stop operating with a tuning voltage of 0 V. See *tuning characteristic.*

Tuning Sensitivity, Tuning Performance. This characteristic, typically expressed in megahertz per volt (MHz/V), enumerates how much a VCO's frequency changes per unit of tuning-voltage change.

Tuning Speed. This characteristic is defined as the time necessary for the VCO to reach 90% of its final frequency on the application of a tuning-voltage step. Tuning speed depends on the internal components between the input pin and tuning diode—including, among other things, the capacitance present at the input port. The input port's parasitic elements determine the VCO's maximum possible modulation bandwidth.

5-7-2 More Practical Circuits

Because of the need for integration, we will present three IC oscillators of increasing complexity. The circuit in Figure 5-99 is found in early Siemens ICs. Figure 5-100 shows its phase-noise performance.

Figure 5-101 shows the basic push–pull oscillator topology underlying the circuit in Figure 5-99. The omission of the current source and active bias components improves the simpler circuit's phase-noise performance over that of the circuit shown in Figure 5-99. Figure 5-102 compares the phase noise of this basic circuit to the integrated version in Figure 5-99.

Figure 5-103 is for those who speculate on the possibility of replacing the BJTs in Figure 5-101 with LDMOS FETs or JFETs. Figure 5-104 compares its phase noise to that of the BJT case. The rule of thumb is that FETs do not have enough gain for high-Q oscillation in oscillators above 400 MHz, and bipolar transistors are a better choice. Because of their higher flicker noise contribution, GaAsFETs should be considered only at frequencies above 4 or 5 GHz.

Figure 5-105 shows a circuit based on two differential amplifiers and frequently found in Gilbert cell IC applications. Figure 5-106 shows its phase-noise performance.

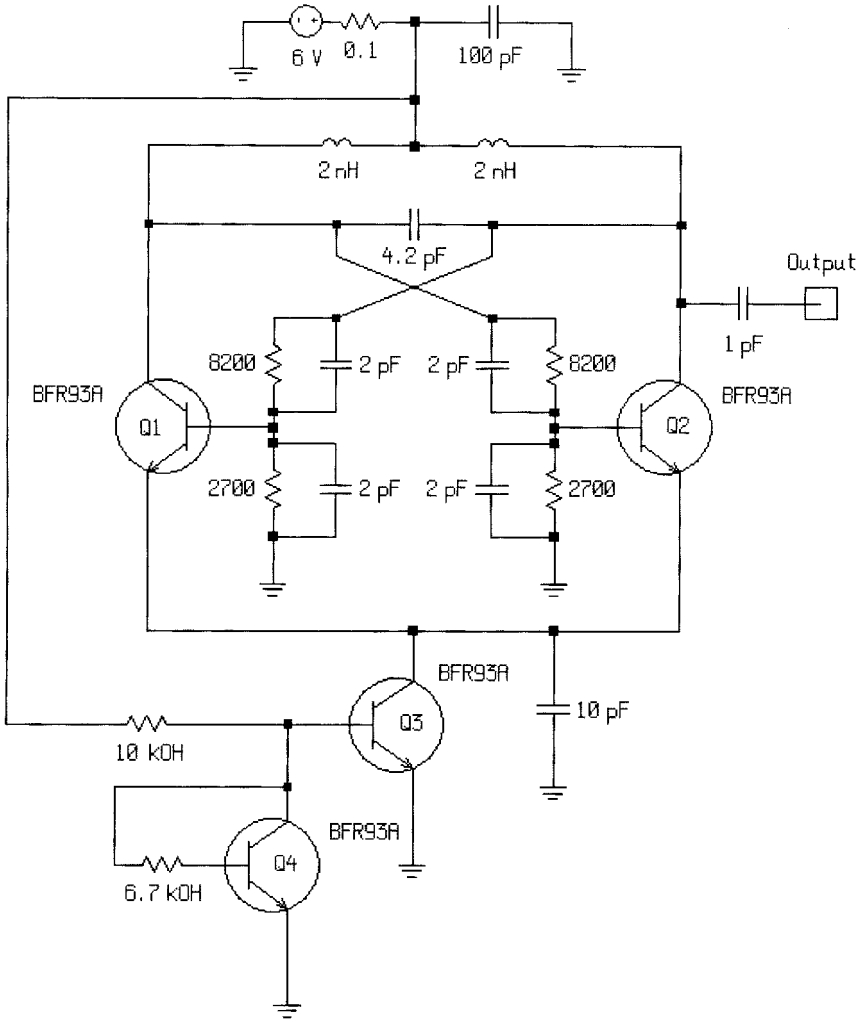


Figure 5-99 Schematic of the Siemens IC oscillator.

Figure 5-107 is a validation circuit that demonstrates a comparison of measured phase-noise performance versus its prediction by circuit simulation software. The circuit is based on the core of Motorola's MC12148 ECL oscillator IC, the topology of which is derived from the superseded MC1648. Motorola specifies the MC12148's typical phase-noise performance as –90 dBc/Hz at an offset of 25 kHz. The close agreement between the measured and simulated results highlights the value of state-of-the-art CAD tools in wireless oscillator design (Figure 5-108).

Figure 5-108 presents the results of *two* simulations: one with ideal L and C components, and another with realistic values specified for these reactances. In using a circuit

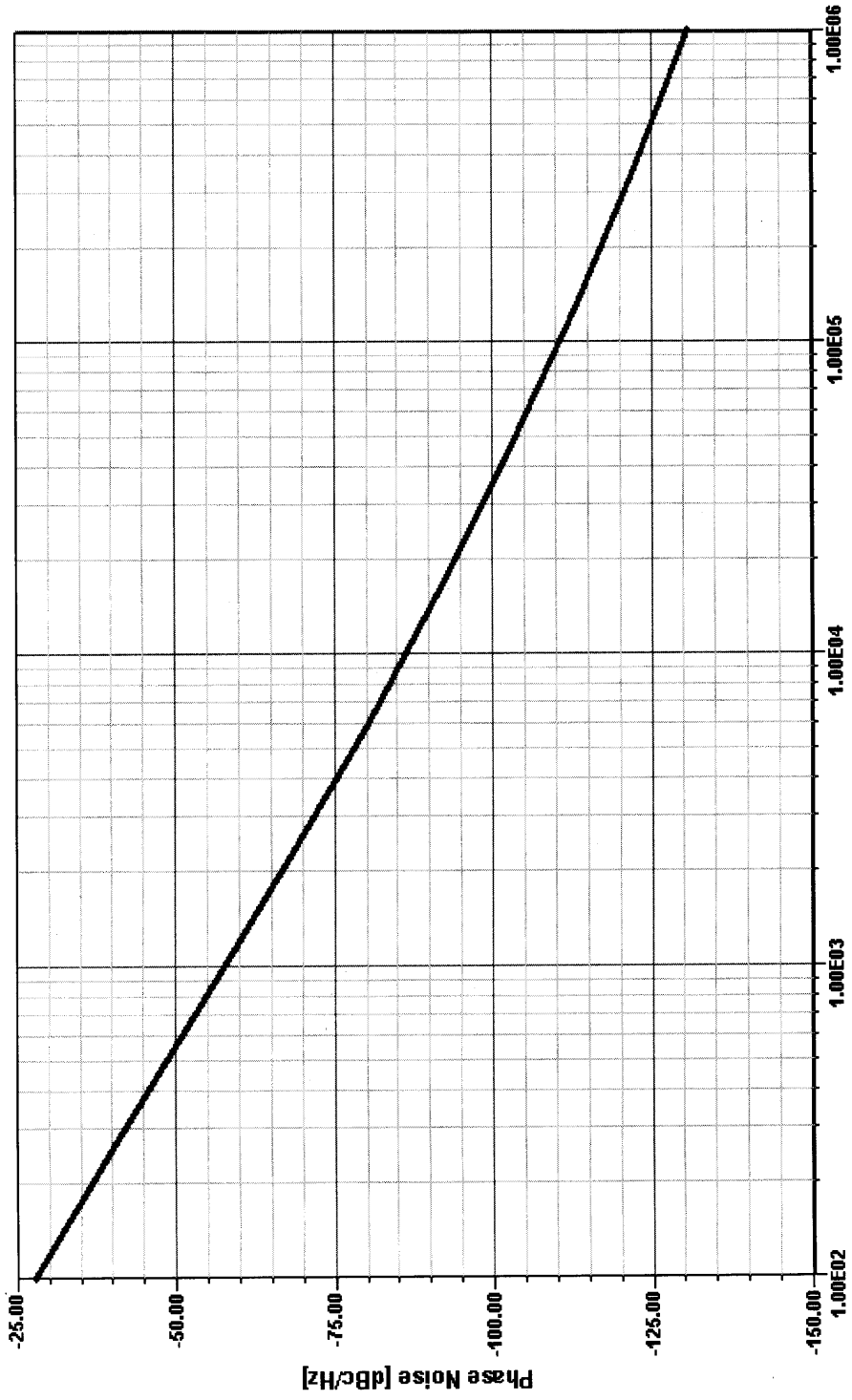


Figure 5-100 Phase noise of the Siemens IC oscillator. The frequency of oscillation is 1.051 GHz.

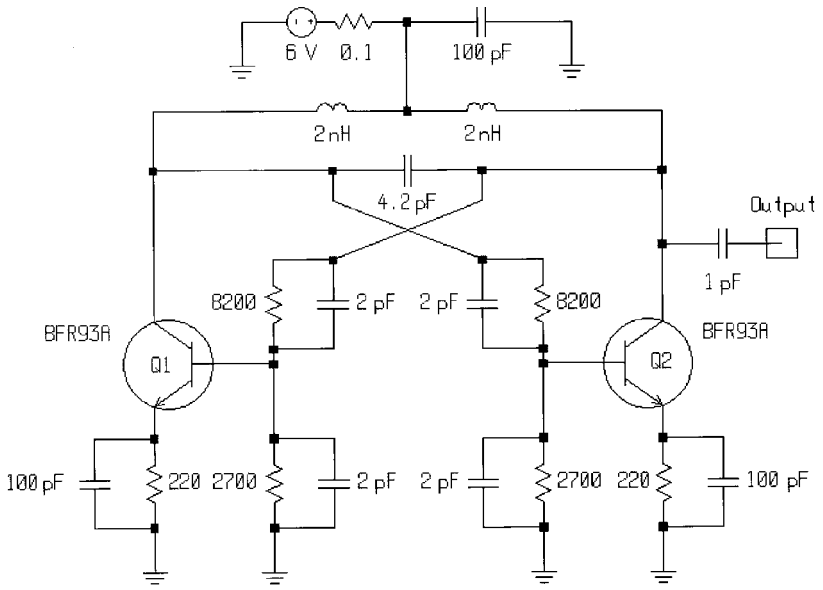


Figure 5-101 Basic push–pull oscillator underlying the circuit shown in Figure 5-99.

simulator to validate and compare oscillator topologies, it's important to start with ideal components because we are interested in comparing the inherent phase-noise performance of various circuits—particularly the role played by the nonlinear reactances in active devices. Because the Q values of practical components are so low in the microwave range (very generally, 200–300 for capacitors and 60–70 for inductors), the inherent merits of one topology, or one device, over another would be masked by loading effects if we conducted our investigations using nonideal reactances. Once initial evaluation is completed, realistic Q values should be specified for increased accuracy in phase-noise, output-power, and output-spectrum simulations.

Finally, as evidence of how moving from lumped to distributed techniques can improve oscillator performance at frequencies where *LC* tanks become problematic, Figure 5-109 compares, for a simulated BJT Colpitts oscillator operating at 2.3 GHz, the difference in phase-noise performance obtainable with a resonator consisting of an ideal 2-nH inductor and a $\frac{1}{4}\lambda$ transmission line (11 Ω, 90° long at 2.6 GHz, attenuation 0.1 dB/m) with the transistor biased by constant-current and constant-voltage sources.

Silicon/GaAs-Based Integrated VCOs and Possible Difficulties. Due to process variations, the center frequency of the VCO cannot be repeated without trimming. Trimming can be avoided by making the VCO gain large enough to compensate for tank circuit variations, but this degrades phase-noise performance and increases the VCO susceptibility to unwanted spurs. For example, a VCO with a gain of 10 MHz/V has a phase noise of –80 dBc/Hz at 10 kHz, while a VCO with a gain of 75 MHz/V has a phase noise of only –60 dBc/Hz at 10 kHz. All but the tank inductors are integrated.

Tank inductors are generally not integrated for reasons of size and Q. The area of a 1-nH inductor is approximately 42,000 mm^2, which is ~3.5% of a typical IC; the area of a 10-nH

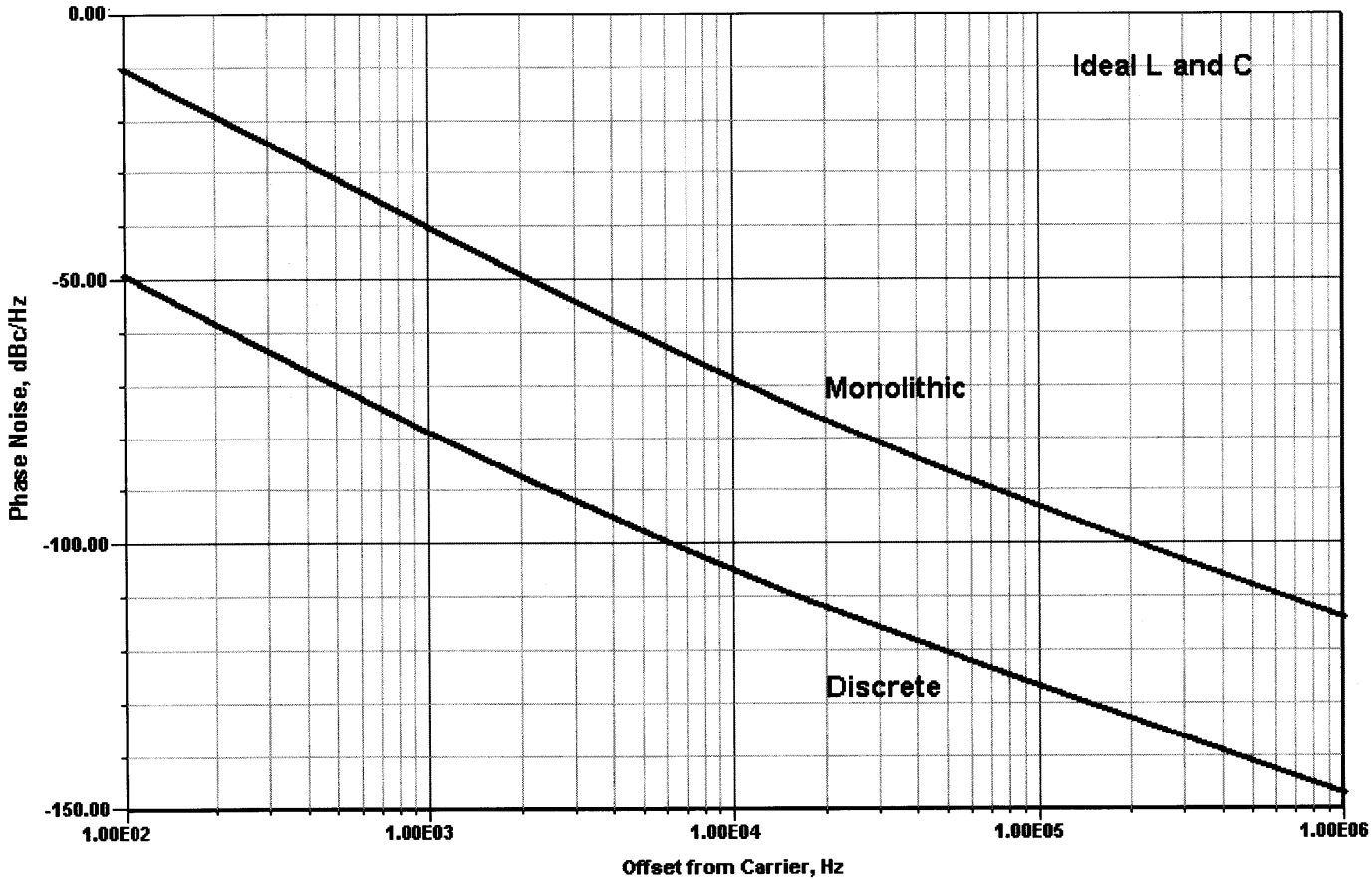


Figure 5-102 The basic version of the oscillator (Discrete curve) is a better phase-noise performer than the integrated version (Monolithic curve). The frequency of oscillation is 1.039 GHz.

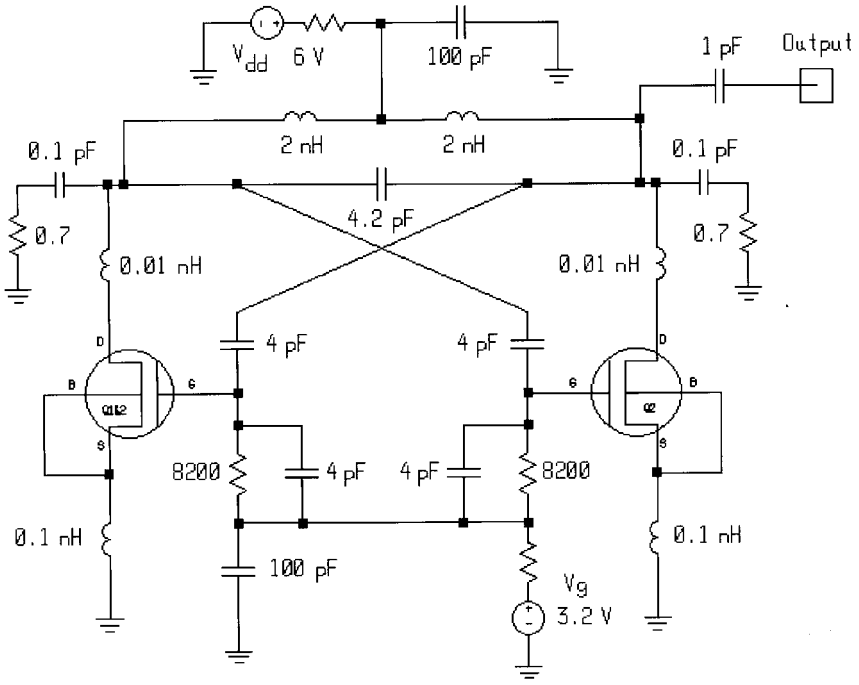


Figure 5-103 A 1.04-GHz push–pull oscillator using LDMOS FETs instead of BJTs. (LDMOS FET parameters courtesy of David K. Lovelace, Motorola, Inc.)

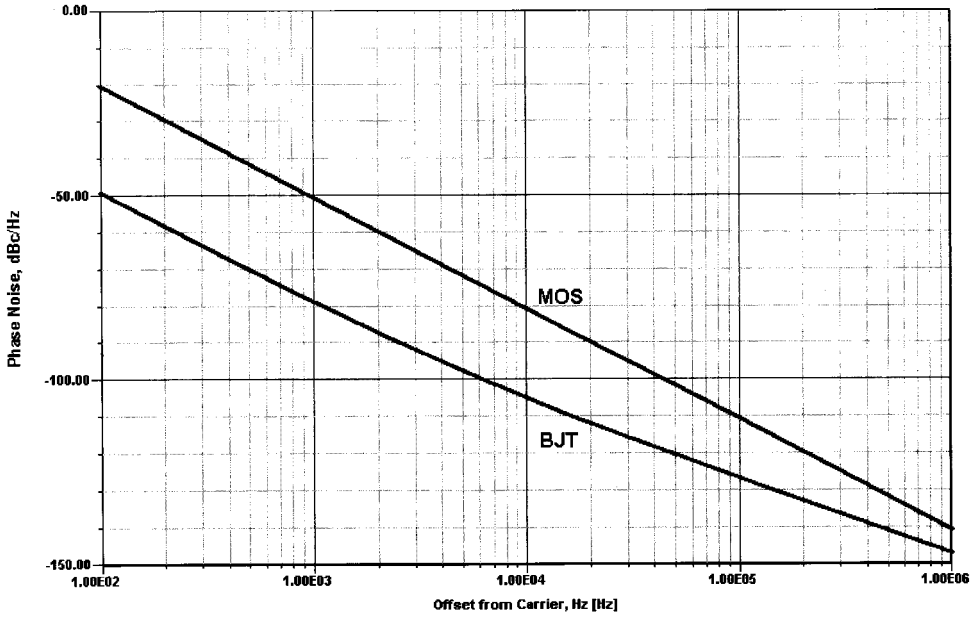


Figure 5-104 Phase-noise comparison of the BJT and MOSFET oscillators.

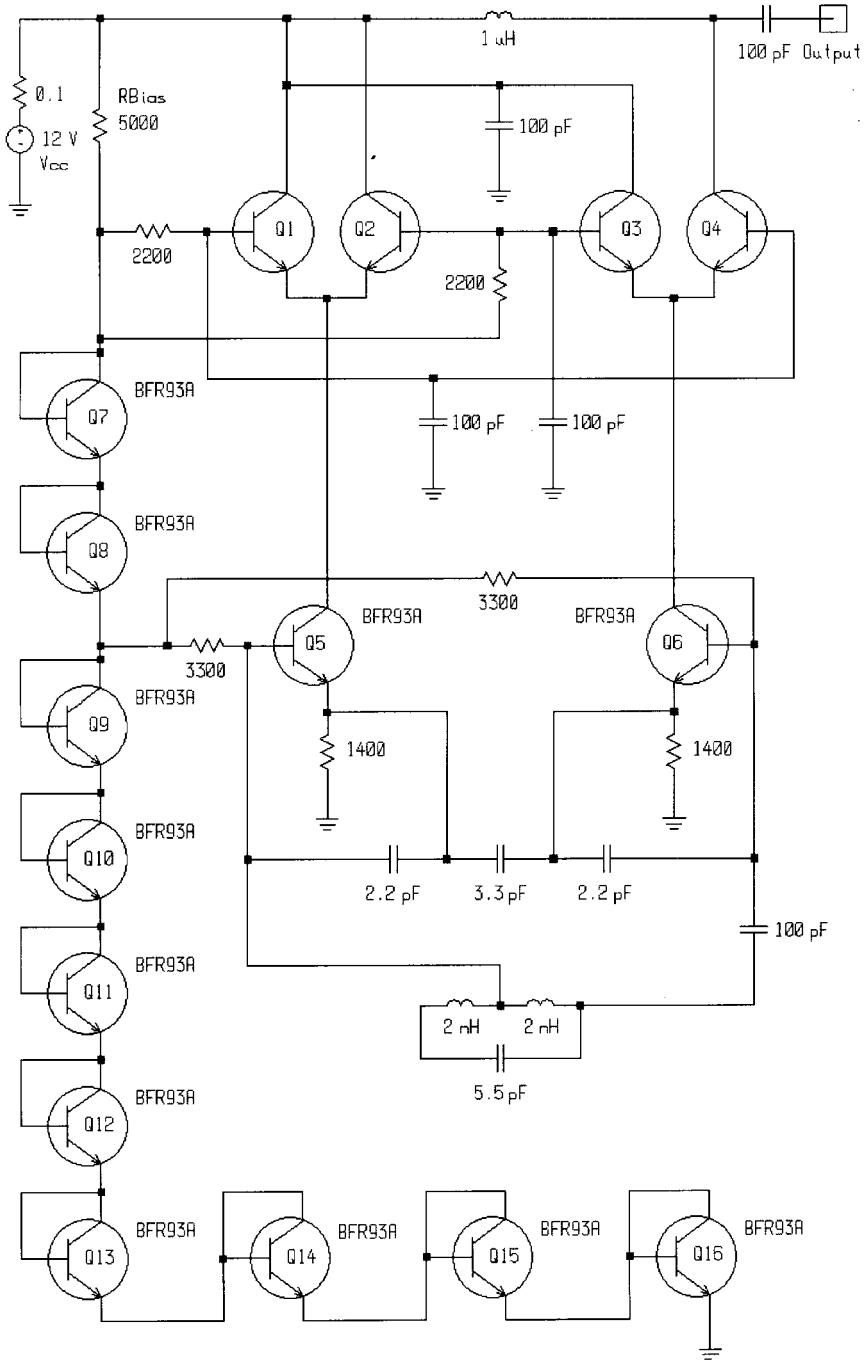


Figure 5-105 A 1.022-GHz IC oscillator based on two differential amplifiers. With an RF signal fed to the bases of Q_1–Q_4 and Q_2–Q_3 in push–pull, and IF output extracted from the collectors of Q_1–Q_3 and Q_2–Q_4 in push–pull, the circuit would act as a Gilbert cell converter. In the circuit shown, we have modified the differential amplifiers' collector circuitry to give single-ended output and avoid cancellation of the oscillator signal.

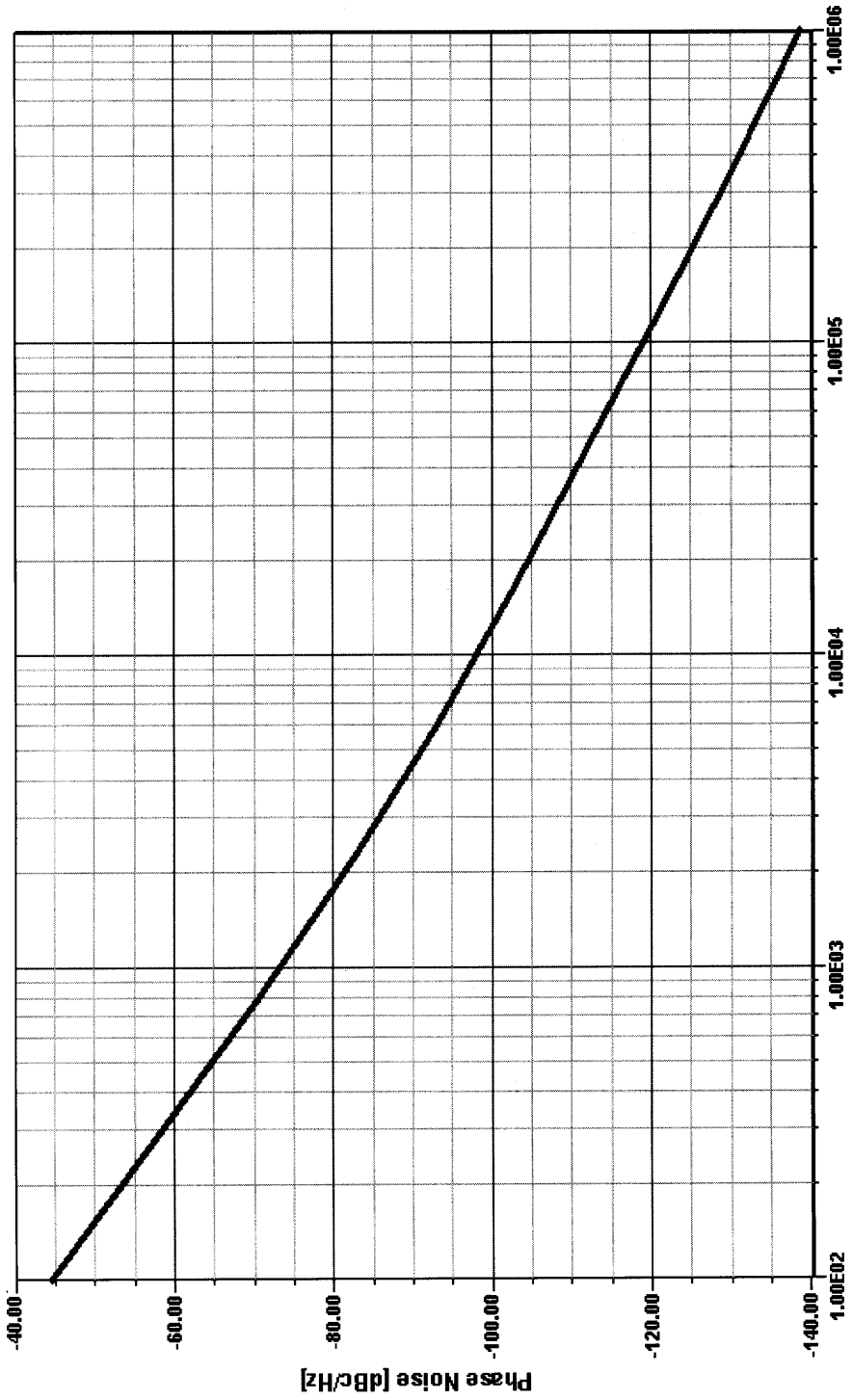


Figure 5-106 Phase noise of the two-differential-amplifier oscillator circuit.

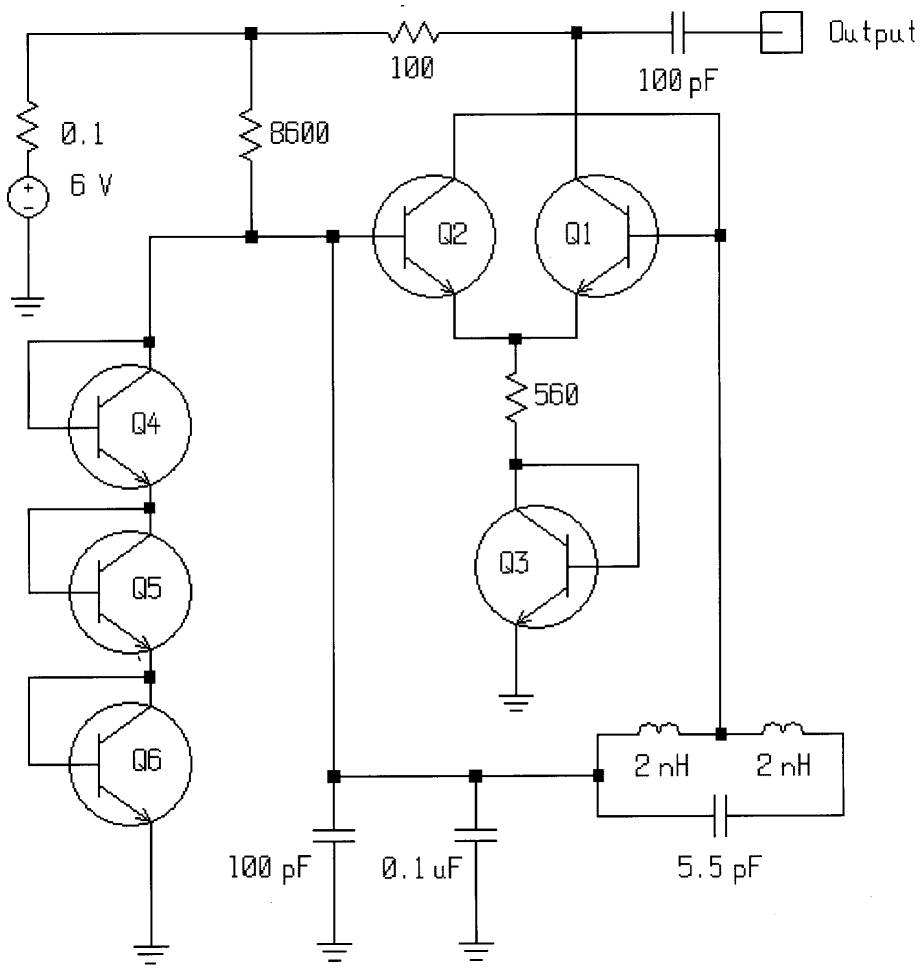


Figure 5-107 This validation circuit models the oscillator topology at the core of Motorola's MC12148 ECL oscillator IC.

inductor is approximately 167,000 mm^2, which is ~14% of a typical IC. For GaAs substrate, the resulting Q would be about 5; for silicon, as we already saw, 20 is already a very high value.

The customer can use PCB traces for the required inductance instead of external surface-mount inductors (cost: $0.03 each, most expensive passive component!). It must be pointed out that even for a nonintegrated approach, the higher VCO gain is noisier by definition. This was explained in Section 5-6-2. These problems are not unique to silicon; they apply to GaAs also. Figure 5-110 shows the size relationship of on-chip capacitors and inductors compared to the actual transistor; therefore, on-chip inductors should really be avoided for reasons of space and Q. Just imagine how the picture would change if the 1.1-nH inductor had to be 11 nH instead! It would dominate the IC area, and therefore the IC cost as well.

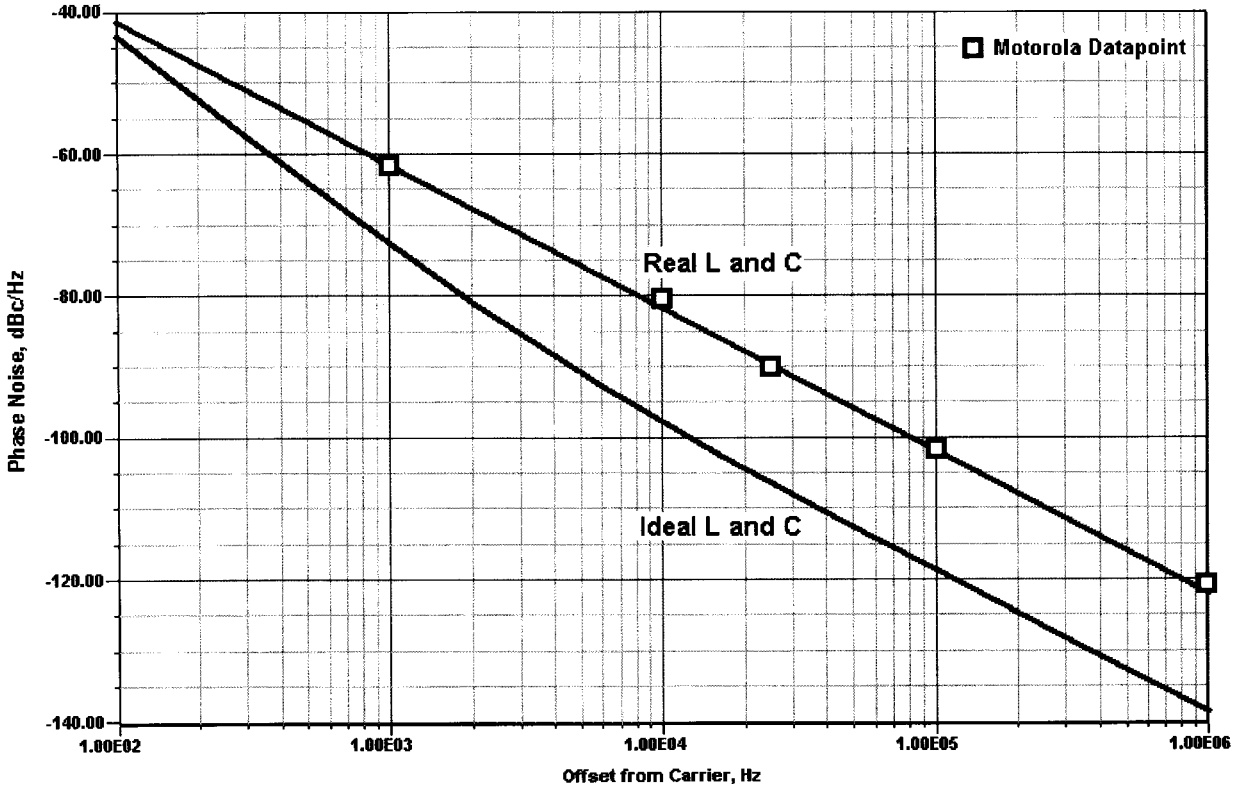


Figure 5-108 Comparison of measured versus simulated phase noise of the MC12148 differential oscillator. The simulation (1.016 GHz) predicts a phase noise of –89.54 dBc/Hz at an offset of 10 kHz (Real L and C curve), which agrees well with Motorola's specification of –89.54 dBc/Hz (typical at 930 MHz) at this offset. The Ideal L and C curve plots the simulation's phase-noise performance with ideal reactances, as opposed to the Real L and C curve, which graphs the results with realistic Q values specified for its reactive components ($Q = 250$ for its C values and $Q = 65$ for its L, all at 1 GHz).

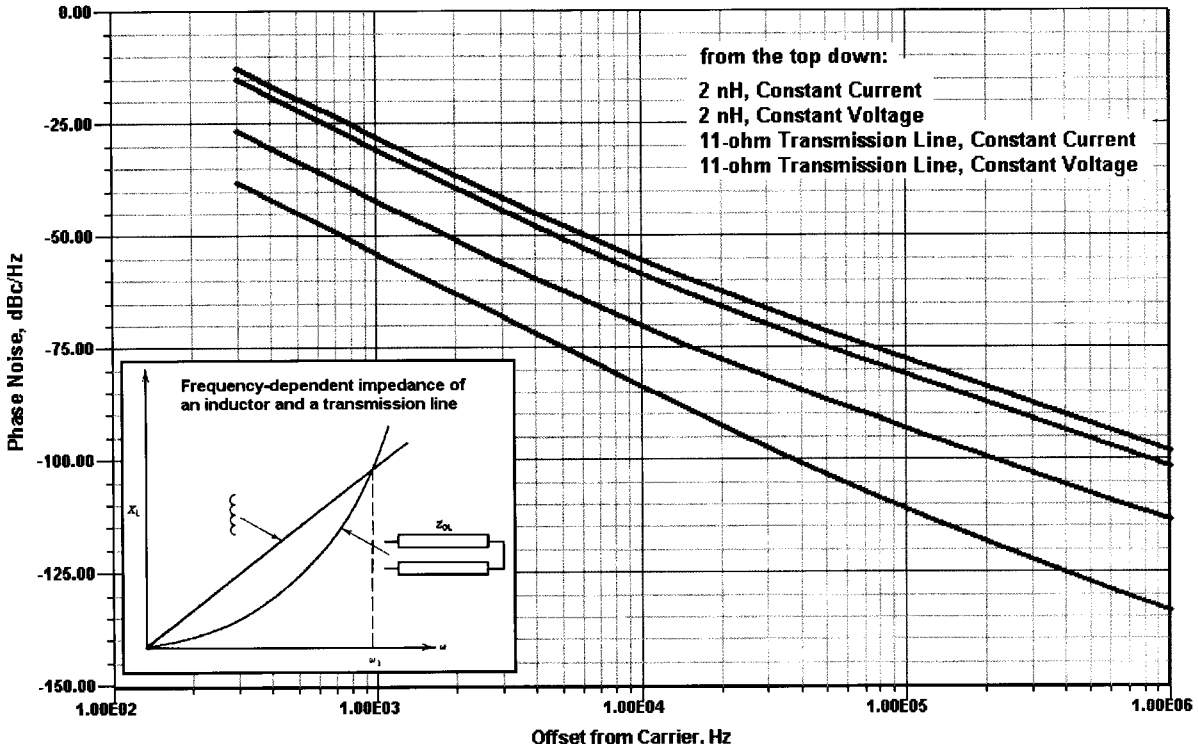


Figure 5-109 Phase-noise performance of a 2.3-GHz BJT oscillator with a resonator consisting of an inductor (2 nH) and a 1/4 λ transmission line (11 Ω, approximating the behavior of a dielectric resonator) with bias from a constant-current source and a low-impedance, resistive constant-voltage source.

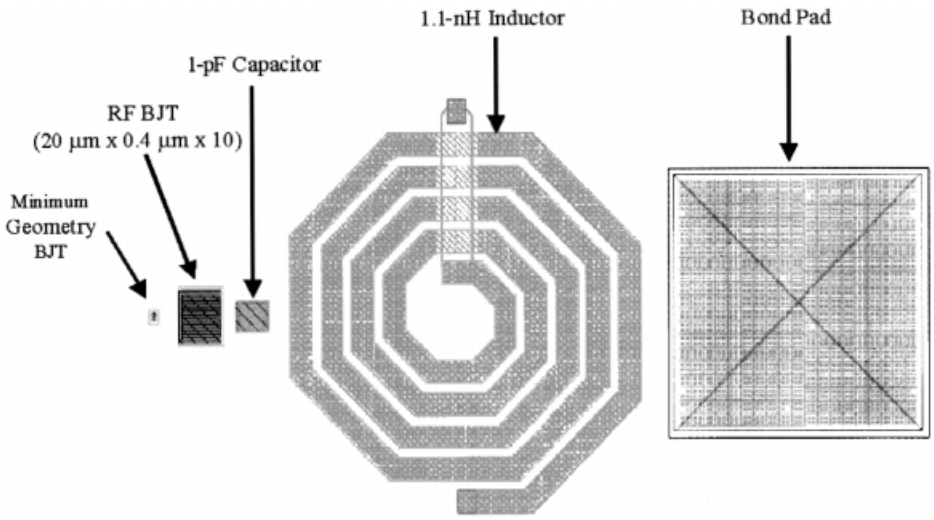


Figure 5-110 Same-scale comparison of space requirements for transistors, capacitors, inductors, and bond pads. For reference purposes, the coil diameter is 0.2 mm.

5-8 DESIGN OF RF OSCILLATORS USING CAD

5-8-1 Harmonic-Balance Simulation

HB analysis performs steady-state analysis of periodically excited circuits. The circuit to be analyzed is split into linear and nonlinear subcircuits. The linear subcircuit is calculated in the frequency domain. Features of this aspect of the HB process include:

- Use of distributed models in the spectral domain
- Matrix formulation that can enable reduction of internal nodes
- Major speed advantage
- Straightforward intermodulation and mixer analysis

The nonlinear subcircuit is calculated in the time domain. Features of this aspect of the HB process include:

- Nonlinear models derived directly from device physics
- Intuitive, easy, and logical circuit representation

Figure 5-111 diagrams this approach for a MESFET amplifier. Figure 5-112 charts a general-purpose nonlinear design algorithm that includes optimization. Modern analysis tools that must provide accurate phase-noise calculation should be based on the principle of harmonic balance. Figure 5-113 shows a BJT microwave oscillator entered into the schematic-capture module of a commercially available HB simulator (Ansoft Serenade 8.0); Figure 5-114 shows this oscillator's simulated phase noise. By the way, HB analysis can also handle mixers.

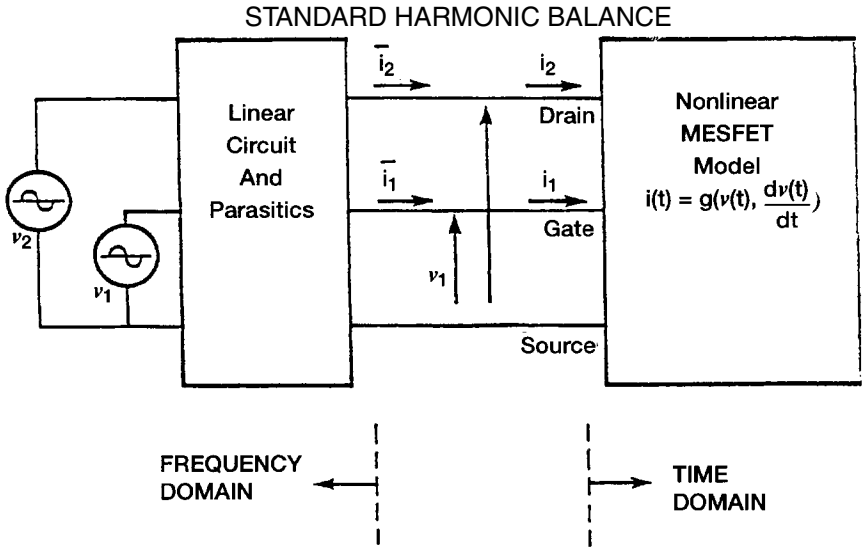


Figure 5-111 MESFET circuit partitioned into linear and nonlinear subcircuits for harmonic-balance analysis. Applied gate and drain voltages, and relevant terminal voltages and currents, are indicated.

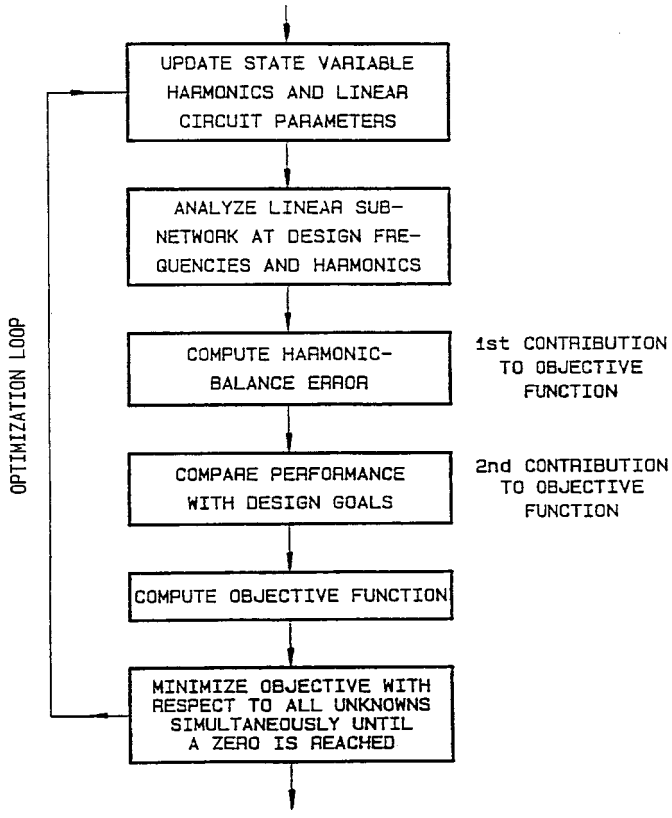


Figure 5-112 Flowchart of a general-purpose harmonic-balance design algorithm that includes optimization.

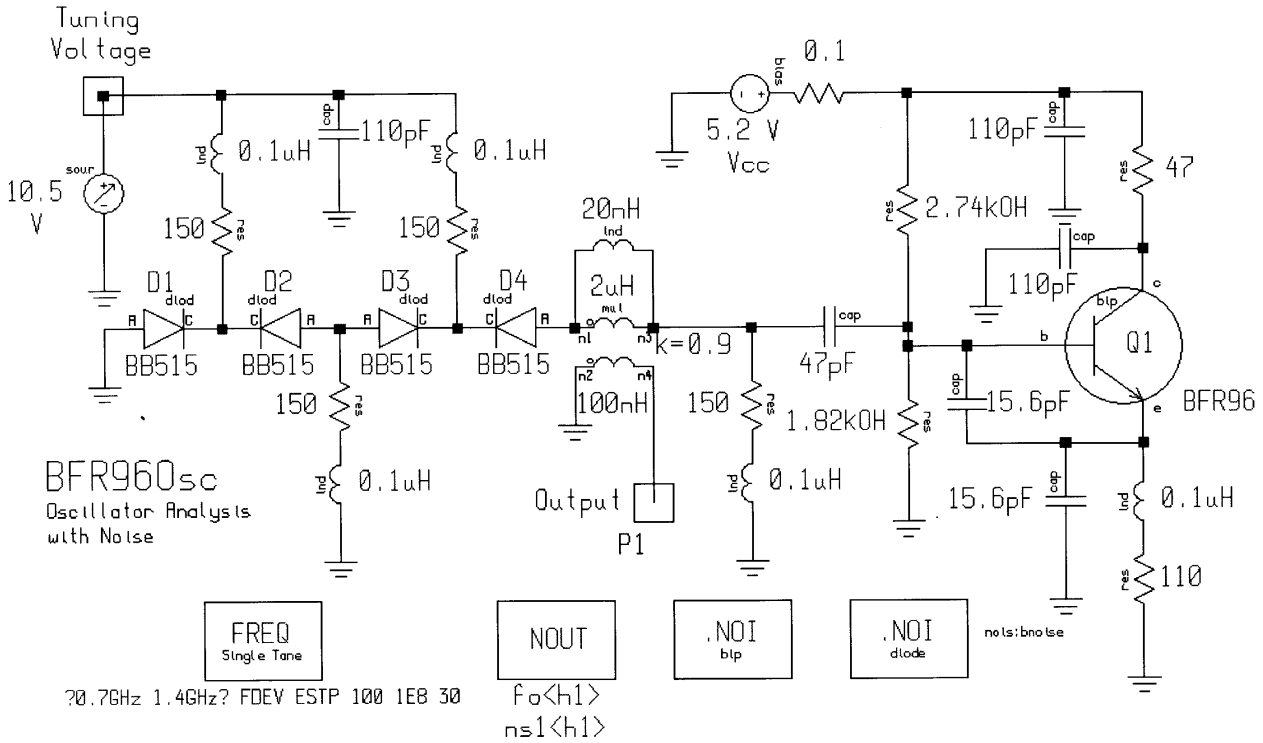


Figure 5-113 BJT microwave oscillator entered into the schematic-capture module of a commercial HB analysis program.

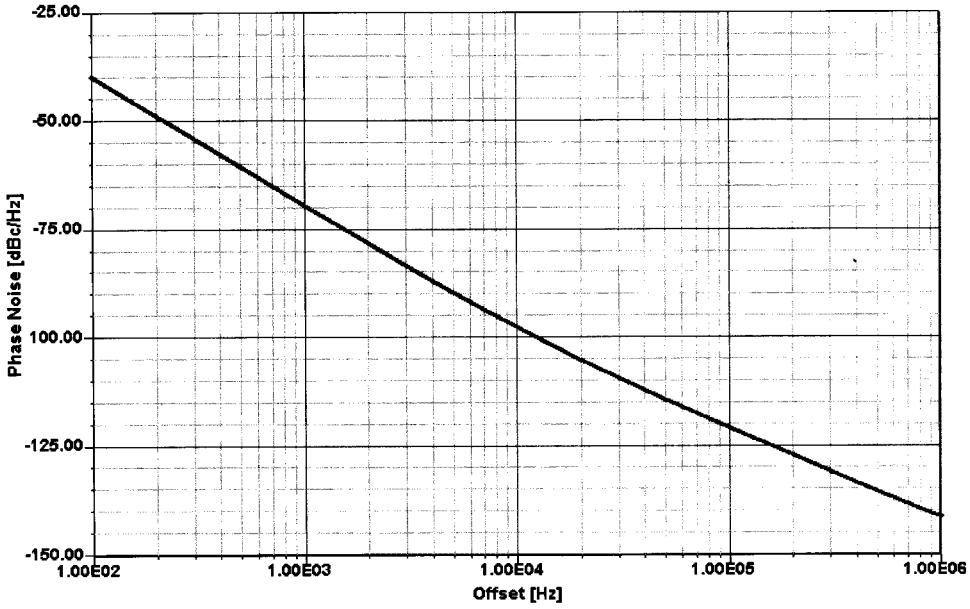


Figure 5-114 Calculated phase noise of the microwave BJT oscillator shown in Figure 5-113.

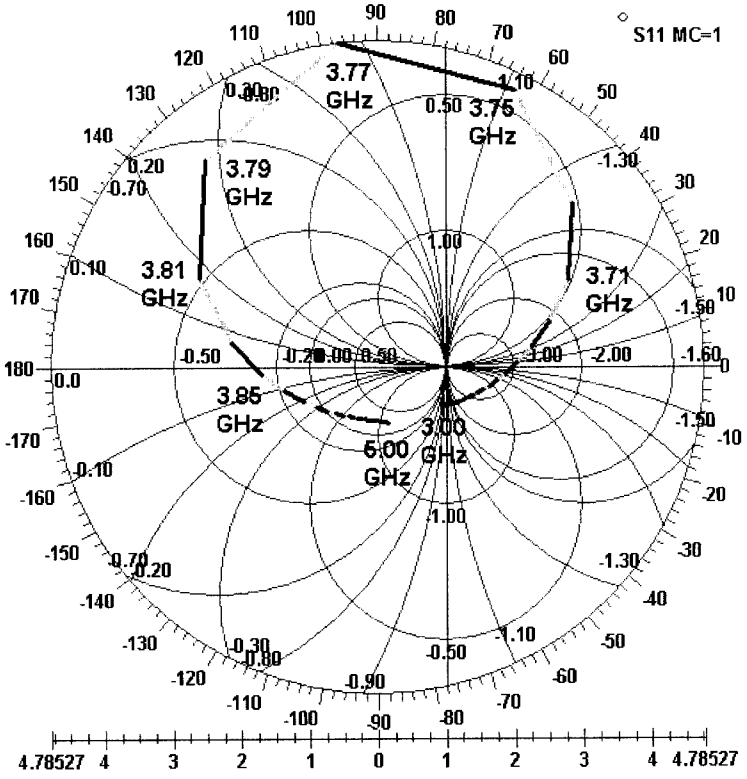


Figure 5-115 S_{11} response of a 3.8-GHz BJT DRO as modeled by Ansoft's Super-Spice time-domain simulator.

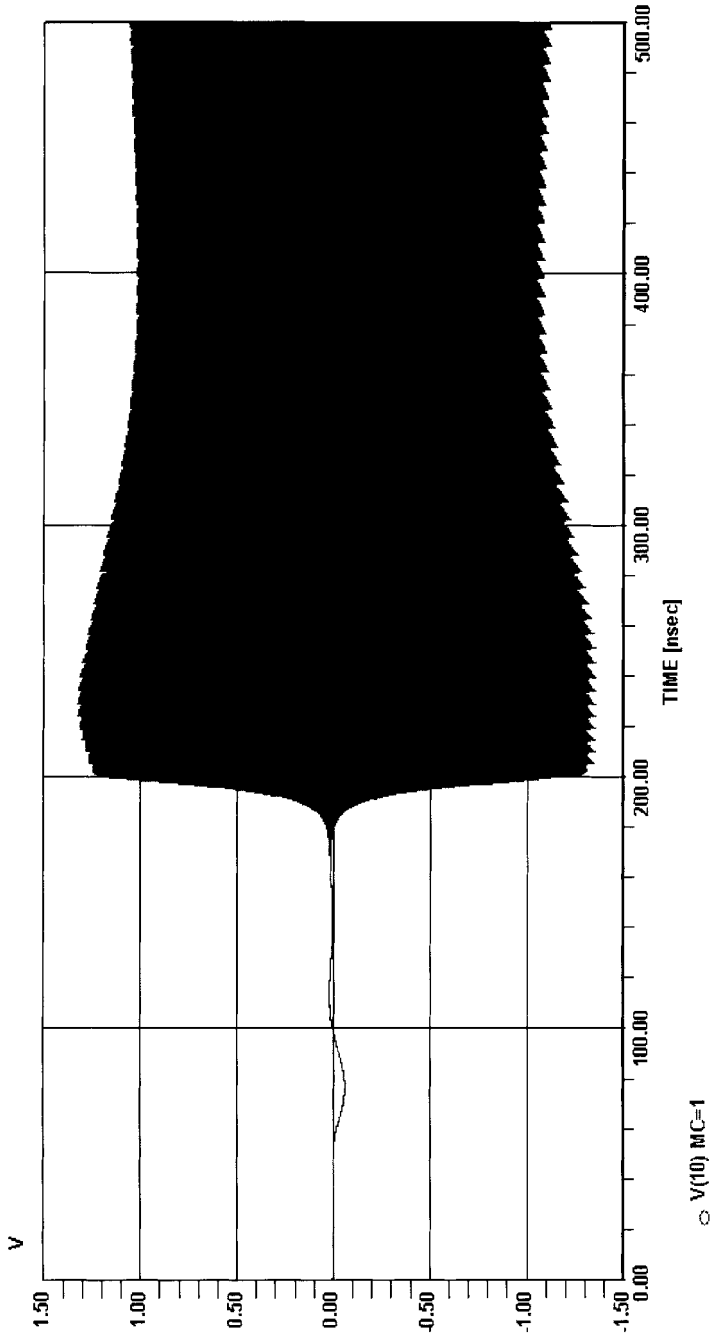


Figure 5-116 Output voltage of the BJT DRO as simulated by Super-Spice.

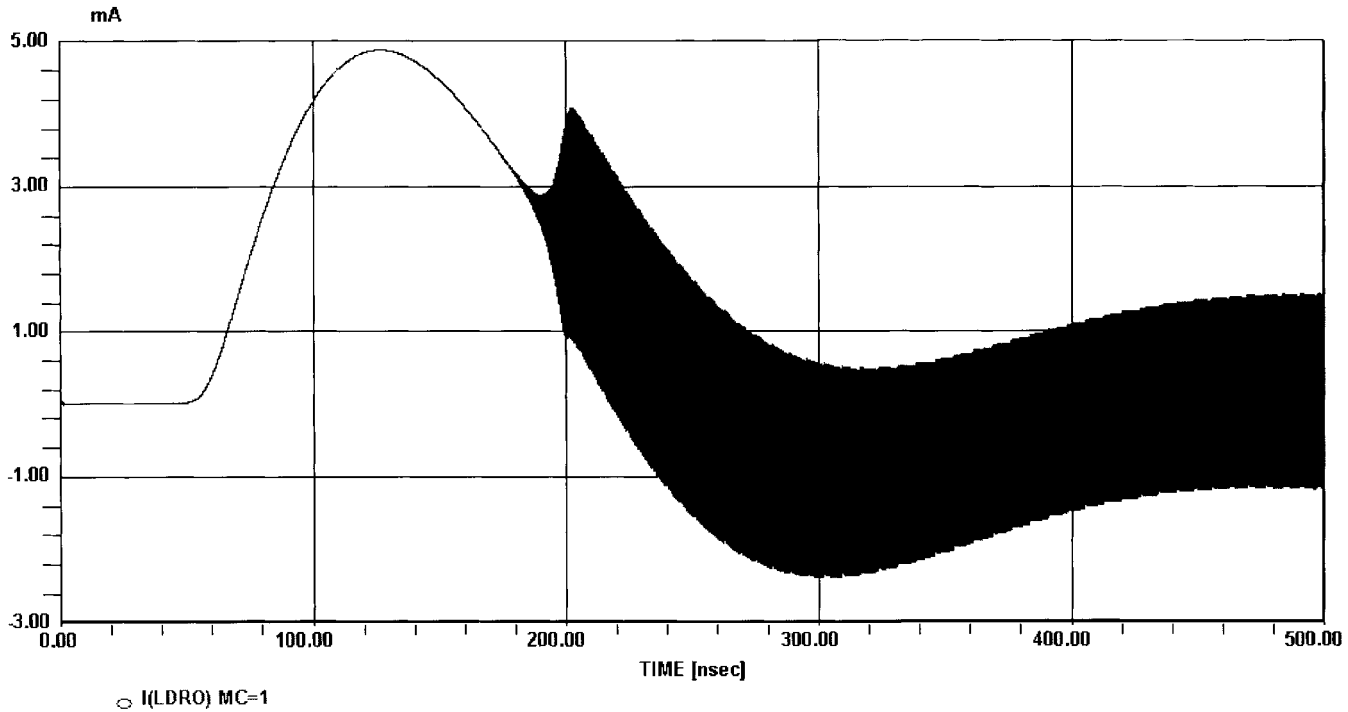


Figure 5-117 Current through the BJT DRO's equivalent resonator inductance from startup to amplitude stabilization.

5-8-2 Time-Domain Simulation

SPICE, a general-purpose simulation program capable of nonlinear dc, linear ac, and nonlinear transient analysis, is probably the most common of the time-domain methods used for nonlinear circuit analysis. Its ability to represent arbitrary time-domain waveforms and model time delays, level shifts, and the sinking and sourcing characteristics of such circuits have made it the simulator of choice for digital designers concerned with analog issues. SPICE's application in RF, microwave, and wireless circuit design is limited by its difficulties with simulating multitone excitations, and coupling and transmission-line effects—including losses, parasitics, and dispersion for distributed elements—and its inefficiency in simulating steady-state conditions.

One area in which SPICE's transient analysis capabilities can be particularly useful to RF and wireless designers is the simulation of oscillator startup. To illustrate this, we present a 3.8-GHz dielectric resonator oscillator modeled with Ansoft Corporation's Super-Spice simulator. Figure 5-115 shows how the oscillator's S_{11} response, positive throughout the simulation's frequency span of 3–5 GHz, peaks at 3.77 GHz. Figure 5-116 shows the oscillator's output voltage as the circuit starts oscillating and amplitude-stabilizes; Figure 5-117 shows the current through the equivalent inductance of the circuit's dielectric resonator across the same time span.

5-9 PHASE-NOISE IMPROVEMENTS OF INTEGRATED RF AND MILLIMETER-WAVE OSCILLATORS *

5-9-1 Introduction

The generation of microwave and millimeter-wave frequencies can be done by lower-frequency VCOs multiplied up in frequency, such as that provided by comb generators; by using hybrid GaAsFET-based oscillators with external resonators; or—for the best phase noise obtainable—YIG oscillators whose physical size and cost do not always provide a practical solution. A significant improvement in oscillator noise performance can be obtained with a novel feedback circuit that uses internally generated noise and cancels the close-in noise in a bandwidth of up to 1 MHz. This results in a phase noise improvement of more than 15 dB at microwave frequencies for the same topology, and the feedback system can be made part of the biasing circuit. This feedback circuit can be used over a wide frequency range and works well at VHF/UHF frequencies and in the millimeter-wave area.

5-9-2 Review of Noise Analysis

Section 5-6 of this chapter covered noise in oscillators in detail, so we will merely review oscillator noise issues here. The first linear model for a basic oscillator, without considering semiconductor noise, was developed by Leeson in 1966 [9] and has been quoted in numerous applications.

The phase noise of a VCO is completely determined by

*This section is based on a paper given at the Fifth International Workshop on Integrated Nonlinear Microwave and Millimeterwave Circuits, sponsored by IEEE at Gerhard-Mercator University, Duisburg, Germany, October 1–2, 1998.

$$£(f_m) = 10 \log\left\{\left[1 + \frac{f_0^2}{(2f_m Q_{\text{load}})^2}\right]\left(1 + \frac{f_c}{f_m}\right)\frac{FkT}{2P_{s,\text{av}}} + \frac{2kTRK_0^2}{f_m^2}\right\} \tag{5-135}$$

where $£(f_m)$ = ratio of sideband power in a 1-Hz bandwidth at f_m to total power in dB
f_m = frequency offset
f_0 = center frequency
f_c = flicker frequency
Q_{load} = loaded Q of the tuned circuit
F = noise factor
$kT = 4.1 \times 10^{-21}$ at 300 K (room temperature)
$P_{s,\text{av}}$ = average power at oscillator output
R = equivalent noise resistance of tuning diode (typically 200 Ω to 10 kΩ)
K_0 = oscillator voltage gain

When adding an isolating amplifier the noise of an *LC* oscillator is determined by

$$S_\phi(f_m) = [a_R F_0^4 + a_E(F_0/2Q_L)^2]/f_m^3$$
$$+ [(2GFkT/P_0)(F_0/2Q_L)^2]/f_m^2$$
$$+ 2a_R Q_L F_0^3/f_m^2$$
$$+ a_E/f_m + 2GFkT/P_0$$

where G = compressed power gain of the loop amplifier
F = noise factor of the loop amplifier
k = Boltzmann's constant
T = temperature in kelvins
P_0 = carrier power level (in watts) at the output of the loop amplifier
F_0 = carrier frequency in hertz
f_m = carrier offset frequency in hertz
$Q_L (= \pi F_0 \tau_g)$ = loaded Q of the resonator in the feedback loop
a_R, a_E = flicker noise constants for the resonator and loop amplifier, respectively

In 1978, Dieter Scherer from Hewlett-Packard added the $(1 + f_c/f_m)$ term, which addresses the 1/*f* noise or flicker corner frequency [25]. The 1/*f* frequency phenomenon is based on the surface effects inside the semiconductor and varies from 50 Hz for silicon FETs, to 5 kHz and more for silicon microwave bipolar transistors, to 1 MHz and higher for GaAsFETs. As a rule of thumb, one can state that the higher the frequency of operation, the higher the f_T for the active device, the higher the flicker corner frequency will be. Figure 5-118 shows the spectral distribution as a function of offset frequency. This linearization was first discovered by individuals from NIST, formerly the National Bureau of Standards, in Boulder, Colorado. The left corner is dominated by the flicker corner frequency and the oscillator itself shows at the output a combination of modulation and conversion

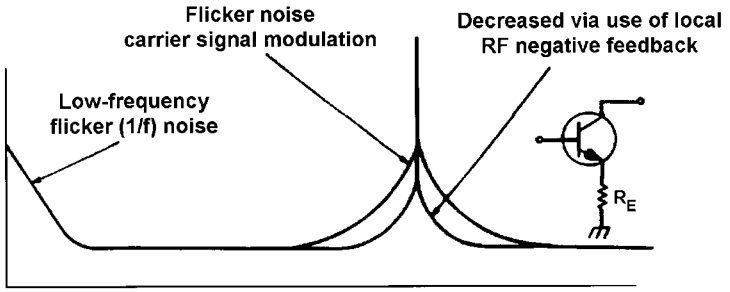


Figure 5-118 Reduction of transistor flicker-of-phase noise via use of local negative feedback (emitter resistance), showing, for a BJT, the flicker noise or corner frequency, as well as the phase noise that results from the flicker noise. The negative feedback from the unbypassed emitter resistor reduces the flicker noise by up to 40 dB [26, 27].

noise. In addition to the $1/f$ noise, the real-life oscillator exhibits a wideband noise, which is due to nonlinearities—specifically, the AM-to-PM conversion resulting from nonlinear capacitance and transconductance of the transistor. This complete linear "noise equation" was first published by Rohde et al. [2]. Therefore, the noise model for the transistor had to be enhanced by describing this mechanism. Essentially, it is equal to a tuning diode coupled to the resonator, which is modulated by thermal noise, and therefore results in a higher phase-noise contribution than the oscillator itself. This noise equation also has Q playing a major role. The phase noise, therefore, ignoring other frequencies, is proportional to $1/Q^2$. In monolithic circuits, the loss of transmission lines, which determines the Q, is related to the substrate material. One has little choice of material, particularly in microwave and millimeter-wave oscillators. Printed resonators sadly provide low Q at these frequencies. A similar problem related to the material is the choices of tuning diodes. In many, if not all, cases, one uses a GaAsFET with the gate connected to the drain or to the source as a tuning diode, and the resulting Q is also disappointing.

The example in Figure 5-118 takes advantage of a microwave bipolar transistor and also is applicable for heterojunction bipolar transistors (HBTs). The transconductance of the bipolar transistor, which is approximately 39 mA/V, increases to values up to 20 times larger than exhibited by GaAsFETs. This method of using negative feedback reduces the gain of the GaAsFET too much and is, therefore, not applicable.

5-9-3 Workarounds

When looking at the modes of operation of test sets for phase-noise measurement, the one method that requires only one signal generator is most desirable. Based on the frequency discriminator method, it uses a cable of appropriate length that acts as a delay line. The oscillator under test drives a double-balanced mixer [one port directly (LO port)] and drives the RF port "delayed" via the delay line cable, and the mixer is therefore driven by the same frequency with a different phase. This setup has been used for phase-noise measurements very successfully. One can easily connect a low-noise FFT-based audio spectrum analyzer to the mixer output and measure the signal-to-noise ratio as a function of offset. With modern FFT analyzers available up to 10 MHz and higher, a fast and reliable measurement can be done. The length of the cable determines the frequency range over which an accurate

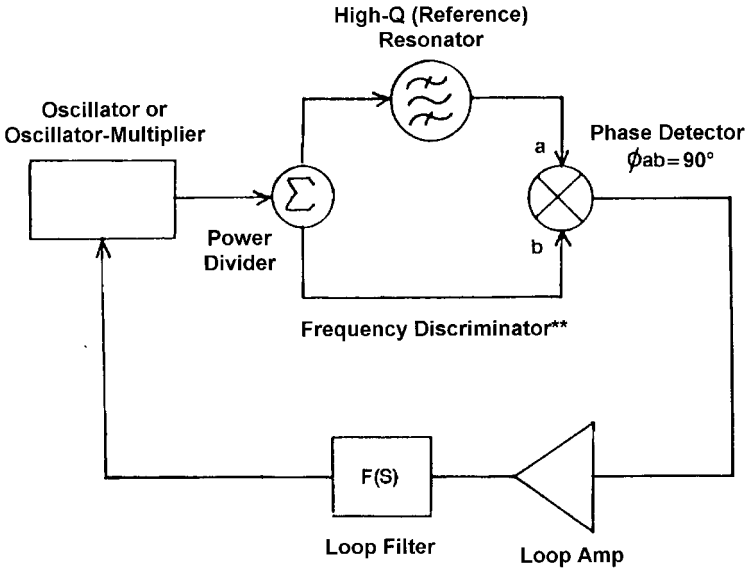
measurement is possible. This phase-noise information (dc and ac voltage) at the output can be used to modulate the VCO and improve the phase noise in the frequency range over which this method is applicable. Fluke and other companies have built signal generators based on this principle to improve the phase noise. The drawback of this method is that the delay line acts like a band-limiting filter, and, therefore, one cannot compensate over a large range of frequencies. The delay line in question, responsible for the phase shift, can also be substituted by a high-Q resonator. By closing the loop, one can "clean up" the phase noise of the device under test (DUT). The difficulty with this is that it requires an external high-Q resonator. One might as well use the same resonator for the oscillator. Therefore, this method is not suitable for MMICs. See Figure 5-119 for more details.

As far as the actual measurement is concerned, one has to observe certain limitations of the test setup as explained in Figure 5-120. When using the delay line instead of a high-Q resonator (low loss is assumed for coaxial cable), one needs to determine the physical length of the cable. The physical length of the cable determines the delay and, therefore, the discriminator noise floor as a function of the offset frequency. Typically, 8 feet of RG-8/U cable results in a delay of 12 ns. As can be seen from Figure 5-121, it is advisable to have several lengths of cable depending on the range of frequency for which one wants to do the measurements under different lengths of cable [28].

Since we are primarily concerned about the flicker noise, here is another, similar, approach to "cleaning up" the flicker noise, as shown in Figure 5-122. The bandwidth over which this technique works depends on the bandwidth of the phase detector system, which can be made fairly wide. A derivative of this will provide a useful solution.

5-9-4 Reduction of Flicker Noise

From inspecting a device datasheet, one is always pleasantly surprised about very low spot noise figures, while the 1/f corner frequency is rarely specified at all. The mechanism that



**Resonator may be operated in reflection mode as well as transmission mode.

Figure 5-119 AFC stabilization of an external oscillator.

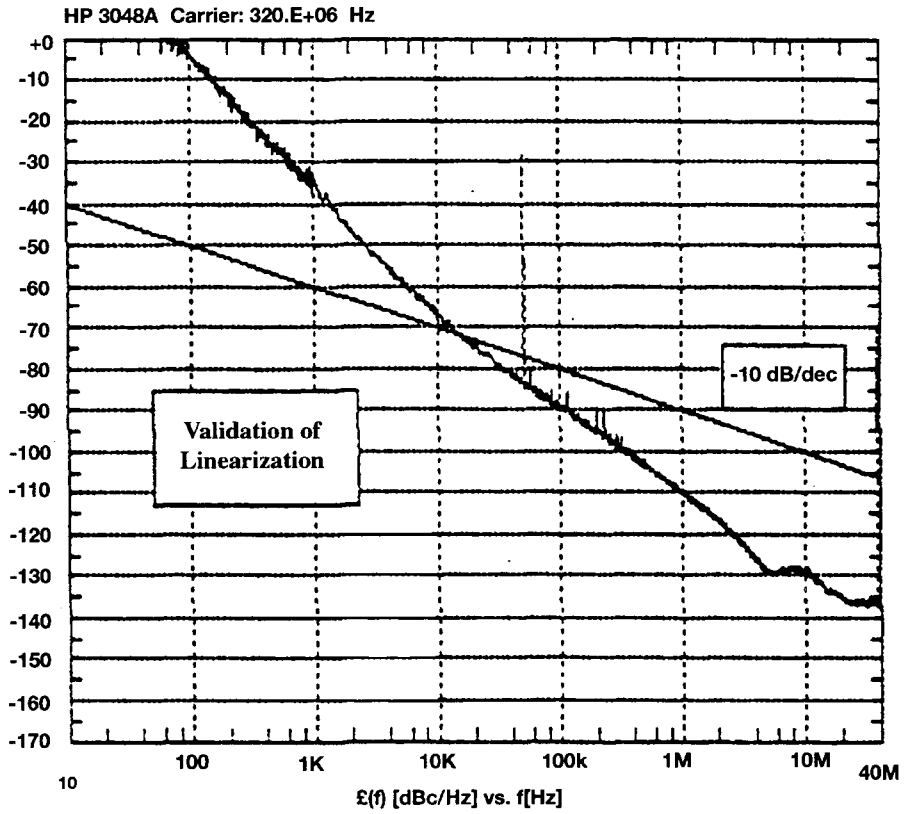


Figure 5-120 Display of a typical phase-noise measurement using the delay line principle. This method is only applicable where $x = \sin(x)$. The measurements above the solid line violate this relationship and therefore are not valid.

transforms the noise, which is generated by the $1/f$ contribution into what we referred to as modulation noise, is the AM-to-PM conversion. If one could "linearize" the transistor, the flicker noise effect would drastically be reduced.

Figure 5-123 shows two variations on an approach that is consistent with the previous method. In Figure 5-123a, the noise across the 56-Ω emitter resistor of oscillator transistor Q1 is sampled, amplified and inverted in Q2, and used to modulate Q1's base voltage. A low-frequency, general-purpose *npn* transistor (BC848C) is used at Q2. A high-frequency, higher-power transistor (BFR93A) is used at Q1; an HBT could also be used. Figure 5-123b shows the same approach using a variation of the *pnp* active biasing circuit shown in Figure 3-192. Figure 5-124 shows the phase-noise performance for the circuit of Figure 5-123a.

5-9-5 Applications to Integrated Oscillators

Figure 5-125 shows a silicon-based BiCMOS technology wireless oscillator. The three big inductors can be seen easily. This design by Motorola is a good candidate for using this innovative technology. Applying this feedback technology, we measure the phase noise shown in Figure 5-126.

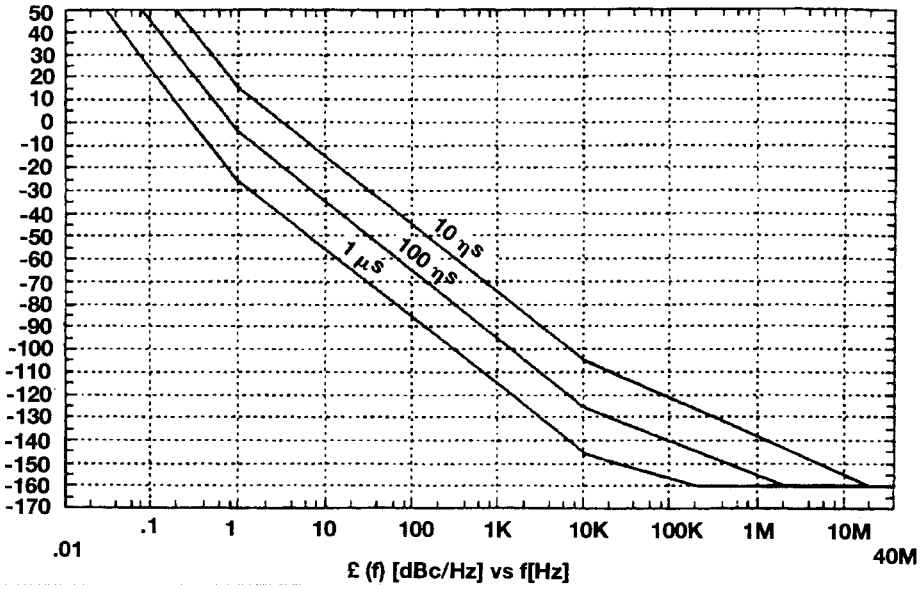


Figure 5-121 Dynamic range as a function of cable delay. A delay of 1 μs is ideal for microwave frequencies.

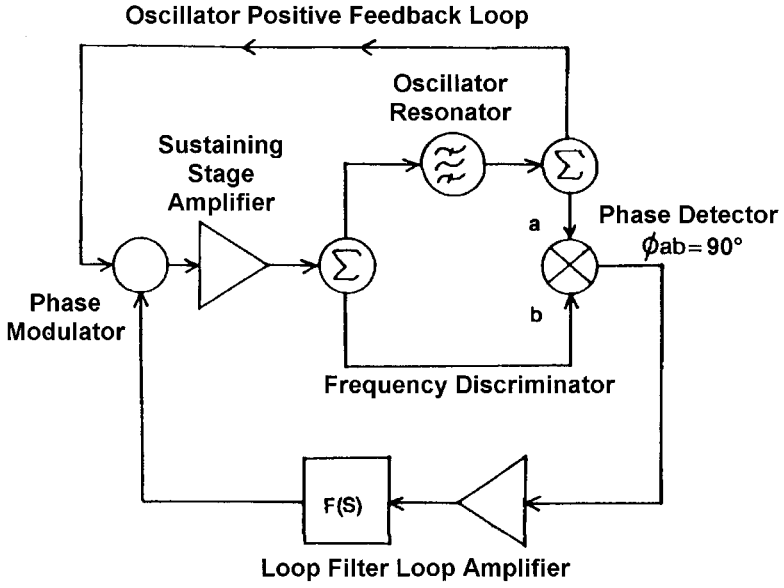
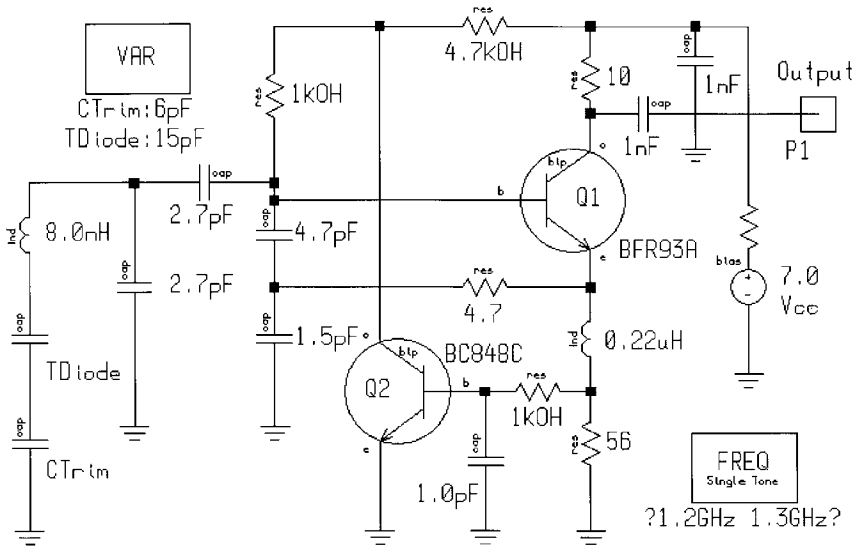
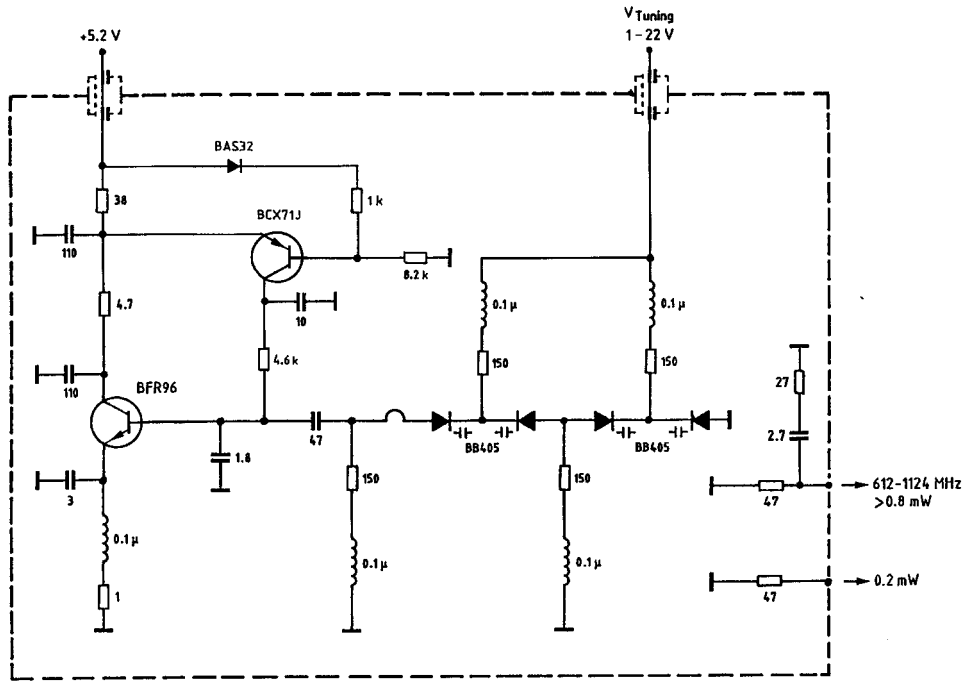


Figure 5-122 Noise reduction in an oscillator whose signal flicker-of-frequency noise is primarily due to sustaining stage flicker-of-phase noise. A phase perturbation in the sustaining stage produces a frequency change in the oscillator, which produces a change in the signal phase shift through the resonator detected as the phase detector. The resonator may be operated simultaneously in transmission line and reflection modes in the oscillator in the oscillator and discriminator portions of the circuit, respectively.



(a)



(b)

Figure 5-123 BJT-based oscillators with noise feedback. (a) The noise sampling is done in the transistor emitter; (b) the noise sampling is done in the collector using a variation of the active biasing circuit shown in Figure 5-124.

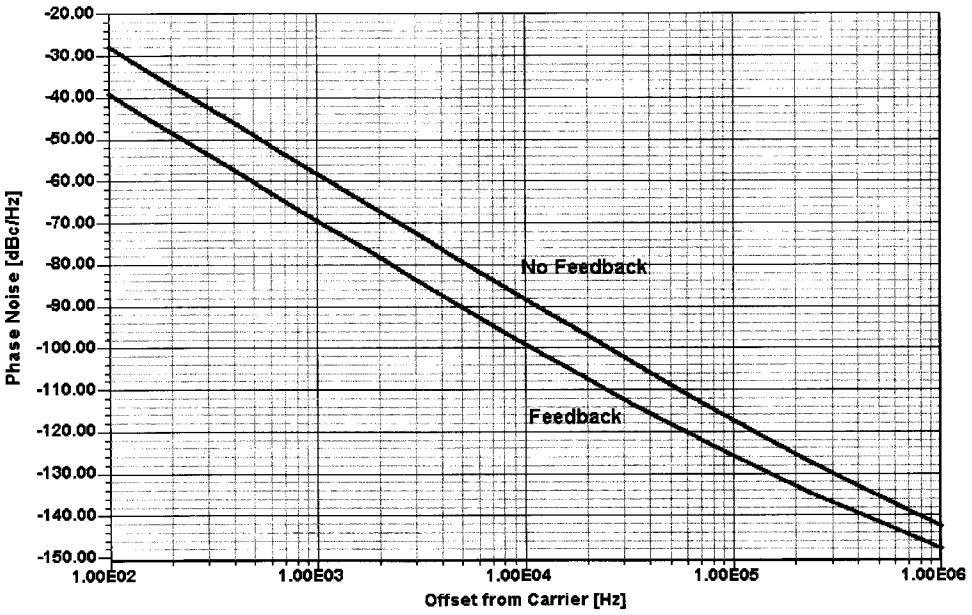


Figure 5-124 Phase noise of the oscillator in Figure 5-123a with and without noise feedback.

Moving up into the millimeter-wave area, one of the target oscillators is the Ka-band MMIC voltage-controlled oscillator designed for Ka-band smart munitions applications (Figure 5-127). This voltage-controlled oscillator MMIC employs 0.25-μm gate length, double-heterojunction, pseudomorphic, high-electron-mobility transistor (PHEMT) technology. This is a custom chip developed by Martin Marietta under the U.S. Government MMIC program. Ansoft Serenade tools were used in this design.

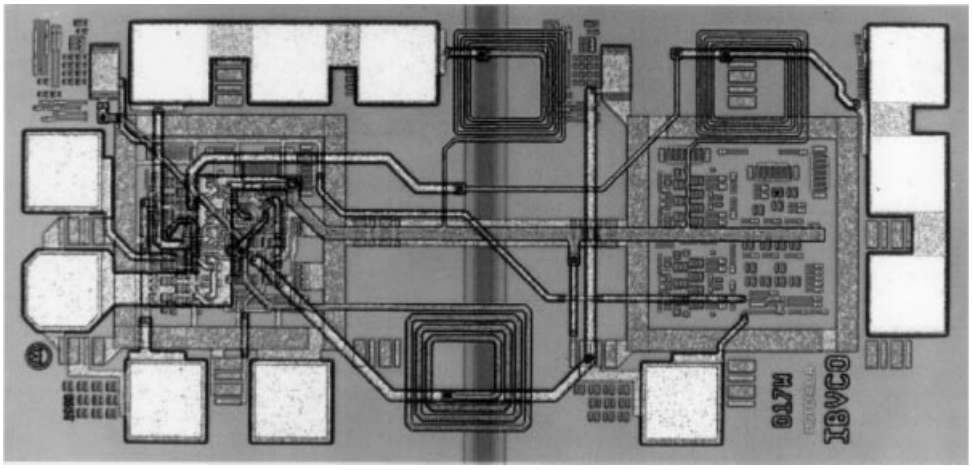


Figure 5-125 Layout of the Motorola 800-MHz monolithic differential oscillator IC in Si technology.

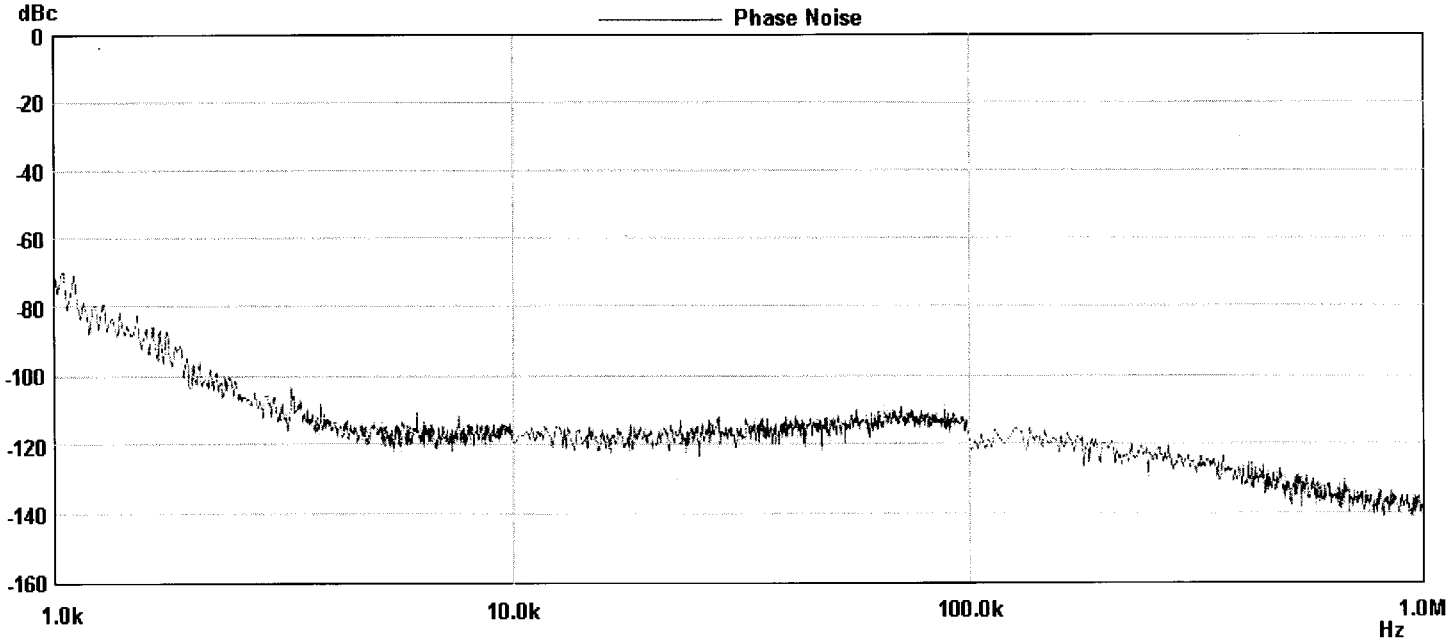


Figure 5-126 Phase noise of the oscillator of Figure 5-125. The noise floor is limited by the on-board isolation amplifier.

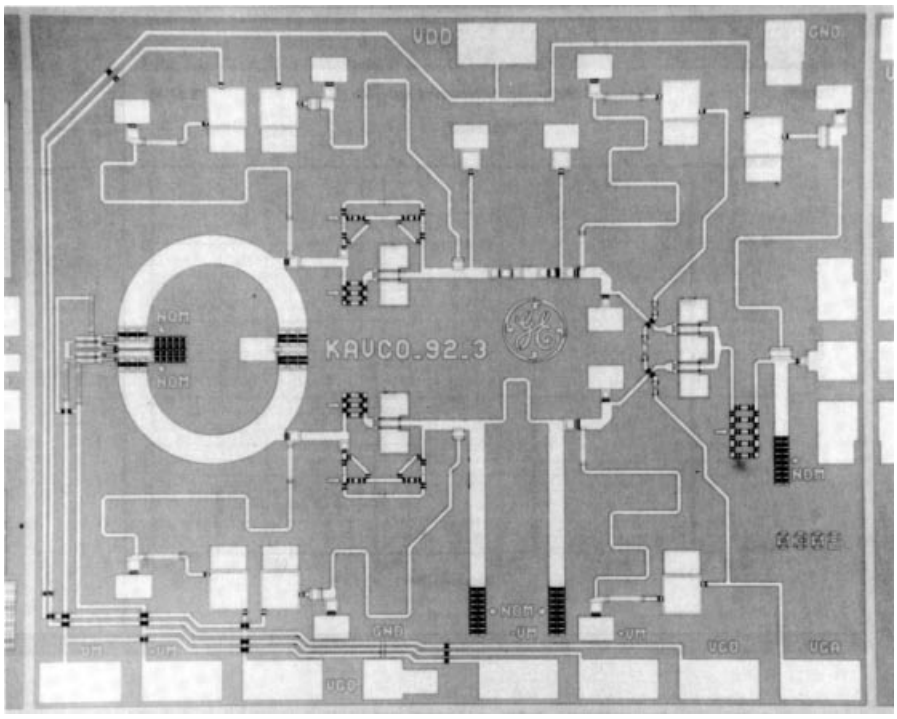


Figure 5-127 Layout of the Ka-band voltage-controlled oscillator.

Features

- Fundamental-mode, differential-ring VCO
- Electronically tunable
- Output power greater than 16 dBm
- Power-added efficiency (PAE) of 13%
- Compact size for easy integration with power amplifier
- 0.25-μm pseudomorphic HEMTs

Specifications

- Frequency range: Ka band
- Output power: 16 dBm minimum

When applying the same technology as shown in the previous bipolar example, the phase noise can be reduced significantly. Since field-effect transistors require a negative gate voltage, it is possible to sample the noise in the drain current using a resistor, amplify the noise with an appropriate loop amplifier, and modulate the gate voltage within a 1-MHz bandwidth.

A spectrum analyzer measurement of this oscillator type, as shown in Figure 5-128, indicates a phase noise of −78.6 dBc/Hz at an offset frequency of 100 kHz at a center frequency of 38 GHz, while the closed-loop phase noise of Figure 5-129 using the compen-

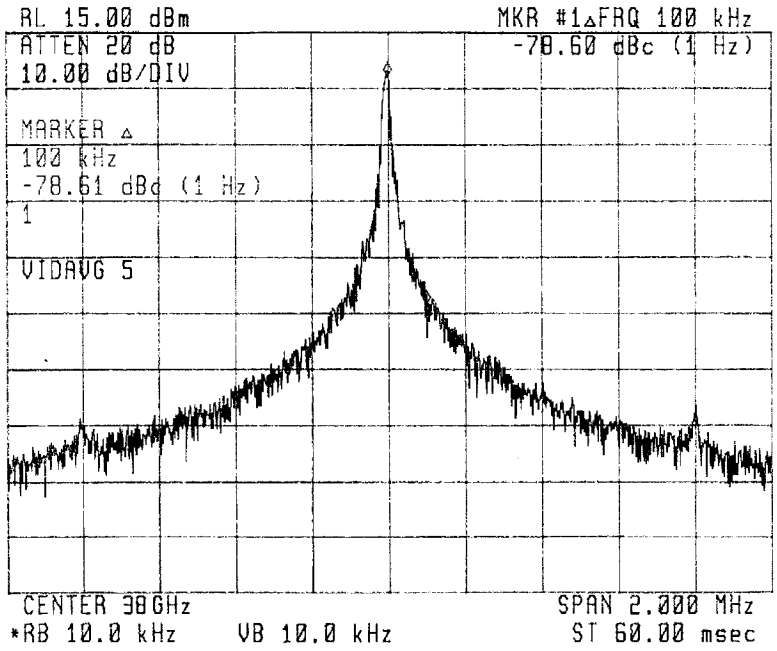


Figure 5-128 Spectrogram of the 38-GHz free-running push–pull VCO shown in Figure 5-127.

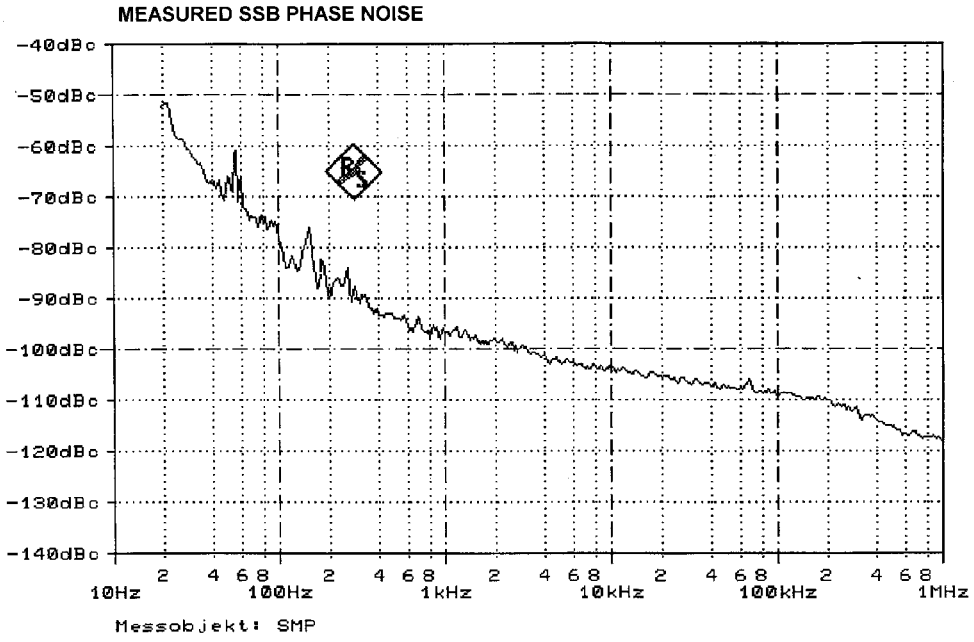


Figure 5-129 Phase-noise performance of a 38-GHz oscillator with phase-noise “clean-up” system.

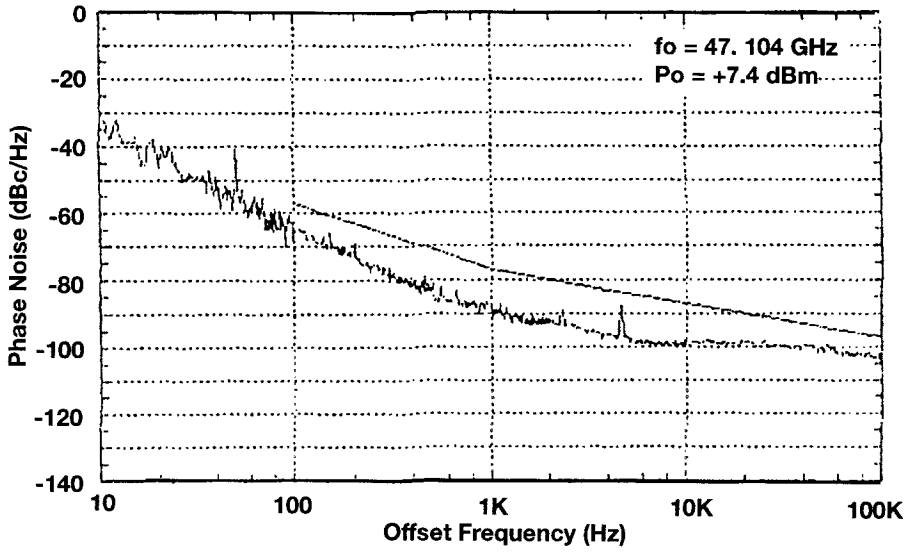


Figure 5-130 A 47-GHz loop-stabilized VCO with phase-noise performance similar to that of Figure 5-129.

sation scheme indicates a phase noise of about −108 dBc/Hz at 100 kHz. This is consistent with reported results from other researchers who have used complicated phase-locked loops to achieve similar performance as seen in Figure 5-130.

5-9-6 Summary

A novel method of improving the phase noise of oscillators has been demonstrated. This approach works well for low-frequency oscillators such as 1000 MHz up to the millimeter-wave range. It only requires an additional wideband dc biasing circuit, which applies noise feedback. The correctness of the theory has been demonstrated by appropriate validation.

REFERENCES

1. Paul Danzer, ed., *The ARRL Handbook for Radio Amateurs,* 75th ed. Newington, CT: American Radio Relay League, 1997, Chap. 14, AC/RF Sources (Oscillators and Synthesizers).
2. Ulrich L. Rohde, Jerry C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*, 2nd ed. New York: McGraw-Hill, 1997, p. 380.
3. G. Vendelin, A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design: Using Linear and Nonlinear Techniques.* New York: Wiley, New York, 1990.
4. Samuel Y. Liao, *Microwave Devices and Circuits,* 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1985.
5. S. C. Peterson, "900-MHz μ-Strip," Philips Semiconductors application document.
6. H. Keller, "Electronic UHF Tuning in TV Receivers," *Radio-Fernsch-Phono-Praxis,* No. 3, 1967.
7. L. Micic, "The Tuner Diode," *Radio Mentor Electron.,* **32**(5), 404–405, 1966.

8. Kyung-Whan Yeom, "An Investigation of the High-Frequency Limit of a Miniaturized Commercial Voltage-Controlled Oscillator Used in 900-MHz Band Mobile-Communication Handset," *IEEE Trans. Microwave Theory Tech.,* **MTT-46**, 1165–1168, August 1998.
9. D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," *Proc. IEEE,* **54**, 329–330, 1966.
10. U. L. Rohde, C. R. Chang, and J. Gerber, "Parameter Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits Like Mixers and Oscillators," *Proceedings of the 23rd European Microwave Conference,* Madrid, Spain, September 6–9, 1993.
11. C. R. Chang, "Mixer Noise Analysis Using the Enhanced Microwave Harmonica," Compact Software *Transmission Line News*, **6**(2), 4–9, June 1992.
12. T. Antognetti and G. Massobrio, *Semiconductor Device Modeling with SPICE.* New York: McGraw-Hill, 1988, p. 91.
13. R. J. Hawkins, "Limitation of Nielsen's and Related Noise Equations Applied to Microwave Bipolar Transistors, and a New Expression for the Frequency and Current Dependent Noise Figure," *Solid State Electron.*, **20**, 191–196, 1977.
14. Tzu-Hwa Hus and Craig P. Snapp, "Low Noise Microwave Bipolar Transistor Sub-Hall-Micrometer Emitter Width," *IEEE Trans. Electron Devices*, **ED-25**, 723–730, June 1978.
15. R. A. Pucel and U. L. Rohde, "An Accurate Expression for the Noise Resistance R_n of a Bipolar Transistor for Use with the Hawkins Noise Model," *IEEE Microwave Guided Wave Lett.*, **3**(2), 35–37, February 1993.
16. See reference 3.
17. R. A. Pucel, W. Struble, Robert Hallgren, and U. L. Rohde, "A General Noise De-embedding Procedure for Packaged Two-Port Linear Active Devices," *IEEE Trans. Microwave Theory Tech.,* **MTT-40**(11), 2013–2024, November 1993.
18. U. L. Rohde, "Improved Noise Modeling of GaAs FETS: Using an Enhanced Equivalent Circuit Technique," *Microwave J.,* Part I: 87–101, November 1991; Part II: 87–95, December 1991.
19. V. Rizzoli, F. Mastri, and C. Cecchefti, "Computer-Aided Noise Analysis of MESFET and HEMT Mixers," *IEEE Trans. Microwave Theory Tech.,* **MTT-37**, 1401–1410, September 1989.
20. V. Rizzoli and A. Lippadni, "Computer-Aided Noise Analysis of Linear Multiport Networks of Arbitrary Topology," *IEEE Trans. Microwave Theory Tech.,* **MTT-33**, 1507–1512, December 1985.
21. V. Rizzoli, F. Mastri, and D. Masotti, "General-Purpose Noise Analysis of Forced Nonlinear Microwave Circuits," *Military Microwave,* 1992.
22. U. L. Rohde, "Key Components of Modern Receiver Design," *VHF Conference*, Weinheim, Germany, September 17–18, 1993.
23. C. R. Chang and U. L. Rohde, "The Accurate Simulation of Oscillator and PLL Phase Noise in RF Sources," *Wireless '94 Symposium,* Santa Clara, California, February 15–18, 1994.
24. U. L. Rohde, "Oscillator Design for Lowest Phase Noise," *Microwave Eng. Eur.*, 31–40, May 1994.
25. Dieter Scherer, "Design Principles and Test Methods for Low Phase Noise RF and Microwave Sources," *RF & Microwave Measurement Symposium and Exhibition*, Hewlett-Packard.
26. D. J. Healy III, "Flicker of Frequency and Phase and White Frequency and Phase Fluctuations in Frequency Sources," *Proceedings of the 26th Annual Symposium on Frequency Control*, Fort Monmouth, NJ, June 1972, pp. 43–49.
27. Michael M. Driscoll, "Low Noise Oscillator Design Using Acoustic and Other High *Q* Resonators," *44th Annual Symposium on Frequency Control*, Baltimore, MD, May 1990.
28. Ulrich L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design*. New York: Wiley, 1997.

INTERESTING PATENTS

Michel Fache and Claude Juliot, "Tunable UHF Oscillator with Harmonic Limitation," Lignes Telegraphoques et Telephoniques (Societe), Paris Cedex, France, U.S. Patent No. 4,146,850, March 27, 1979.

Thomas M. Higgins, Jr., "Negative Resistance Oscillator with Electronically Tunable Base Inductance," Hewlett-Packard Company, Palo Alto, California, U.S. Patent No. 5,373,264, December 13, 1994.

Osamu Ichiyoshi, "Voltage-Control Oscillator Which Suppresses Phase Noise Caused by Internal Noise of the Oscillator," NEC Corporation, Japan, U.S. Patent No. 5,351,014, September 27, 1994.

Shankar R. Joshi, Ulrich L. Rohde, and Klaus Eichel, "Phase-Locked Loop Circuits and Voltage Controlled Oscillator Circuits," Synergy Microwave Corp., Paterson, New Jersey, U.S. Patent No. 5,650,754, July 22, 1997.

Joseph H. Kiser, "Wide-Range Electronic Oscillator," Vari-L Co., Denver, Colorado, U.S. Patent No. 4,621,241, November 4, 1986.

Joseph H. Kiser, "Oscillator Voltage Regulator," Vari-L Co., Denver, Colorado, U.S. Patent No. 5,675,478, October 7, 1997.

Charles Lewis, "Low Phase Noise UHF and Microwave Oscillator," Z-Communications, Inc., San Diego, California, U.S. Patent No. 5,748,051, May 5, 1998.

Larry R. Lockwood, "Low-Noise Oscillator," Tektronix, Inc., Beaverton, Oregon, U.S. Patent No. 4,580,109, April 1, 1986.

Benedict J. Nardi, "Low Phase Noise Reference Oscillator," Watkins-Johnson Company, Palo Alto, California, U.S. Patent No. 5,341,110, August 23, 1994.

Leon Riebman, "Low-Noise Oscillator," AEL Defense Corp., Lansdale, Pennsylvania, U.S. Patent No. 5,166,647, November 24, 1992.

Takeshi Saitoh and Hiroshi Hatashita, "TV Tuner Oscillator with Feedback for More Low-Frequency Power," Hitachi Ltd., Tokyo, Japan, U.S. Patent No. 4,564,822, January 14, 1986.

FURTHER READING

Stanislaw Alechno, "Analysis Method Characterizes Microwave Oscillators," Part 1, *Microwaves & RF*, **36**(11), 82–86, November 1997.

P. Antognetti and G. Massobrio, *Semiconductor Device Modeling with SPICE.* New York: McGraw-Hill, 1993.

W. Anzill, F. X. Kärtner, and P. Russer, "Simulation of the Single-Sideband Phase Noise of Oscillators," *Second International Workshop of Integrated Nonlinear Microwave and Millimeterwave Circuits*, 1992.

Behagi, "Piece-wise Linear Modeling of Solid-State Varactor-Tuned Microwave Oscillators," *Proceedings of the IEEE Frequency Control Symposium*, 1992, pp. 415–519.

Noël Boutin, "RF Oscillator Analysis and Design by the Loop Gain Method," *Appl. Microwave & Wireless*, **11**(8), 32–48, August 1999.

P. Braun, B. Roth, and A. Beyer, "A Measurement Setup for the Analysis of Integrated Microwave Oscillators," *Second International Workshop of Integrated Nonlinear Microwave and Millimeterwave Circuits,* 1992.

C. R. Chang, M. B. Steer, S. Martin, and E. Reese, "Computer-Aided Analysis of Free-Running Microwave Oscillators," *IEEE Trans. Microwave Theory Tech.,* **MTT-39**, 1735–1745, October 1991.

Y. Cheng, K. Chen, K. Imai, and C. Hu, "A Unified MOSFET Channel Charge Model for Device Modeling in Circuit Simulation," *IEEE Trans. Computer-Aided Des.*, **17**, 641–644, August 1998.

J. Craninckx and M. Steyaert, "A 1.8-GHz Low-Phase-Noise CMOS VCO Using Optimized Hollow Spiral Inductors," *IEEE J. Solid-State Circuits*, **SC-32**(5), 736–744, May 1997.

P. Davis, P. Smith, E. Campbell, J. Lin, K. Gross, G. Bath, Y. Low, M. Lau, Y. Degani, J. Gregus, R. Frye, and K. Tai, "Silicon-on-Silicon Integration of a GSM Transceiver with VCO Resonator," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers,* pp. 248–249.

Alper Demir, Amit Mehrotra, and Jaijeet Roychowdhury, "Phase Noise and Timing Jitter in Oscillators," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 45–48.

P. J. Garner, M. H. Howes, and C. M. Snowden, "*Ka*-Band and MMIC pHEMT-Based VCOs with Low Phase-Noise Properties," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1531–1536, October 1998.

G. Gonzalez, *Microwave Transistor Amplifiers—Analysis and Design,* 2nd ed. New York: Prentice-Hall, 1997.

A. V. Grebennikov and V. V. Nikiforov, "An Analytic Method of Microwave Transistor Oscillator Design," *Int. J. Electron.,* **83**, 849–858, December 1997.

Andrey V. Grebennikov, "Microwave Transistor Oscillators: An Analytic Approach to Simplify Computer-Aided Design," *Microwave J.*, **42**(5), 292–300, May 1999.

Ali Hajimiri and Thomas H. Lee, "A General Theory of Phase Noise in Electrical Oscillators," *IEEE J. Solid-State Circuits*, **SC-33**, 179–194, February 1998.

Ali Hajimiri, Sotiriosw Limotyrakis, and Thomas H. Lee, "Phase Noise in Multi-gigahertz CMOS Ring Oscillators," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 49–52.

Frank Herzel and Behzad Razavi, "Oscillator Jitter Due to Supply and Substrate Noise," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 489–492.

Qiuting Huang, "On the Exact Design of RF Oscillators," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference*, pp. 41–44.

Kamran Iravani and Gary Miller, "VCOs with Very Low Sensitivity to Noise on the Power Supply," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 515–518.

K. M. Johnson, "Microwave Varactor-Tuned Transistor Oscillator Design," *IEEE Trans. Microwave Theory Tech.*, **MTT-14**(11), 564–572, September 1966.

K. M. Johnson, "Large-Signal GaAs MESFET Oscillator Design," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**(3), 217–226, March 1979.

W. E. Kamali, J. P. Grimm, R. Meierer, and C. Trirons, "New Design Approach for Wideband FET Voltage-Controlled Oscillators," *IEEE Trans. Microwave Theory Tech.,* **MTT-34**(10), 1059–1063, October 1986.

F. X. Kärtner, "Noise in Oscillating Systems," *Second International Workshop of Integrated Nonlinear Microwave and Millimeterwave Circuits*, 1992.

T. Kashiwa, T. Ishida, T. Katho, H. Kurusu, H. Hoshi, and Y. Mitsui, "V-Band High-Power Low Phase-Noise Monolithic Oscillators and Investigation of Low Phase-Noise Performance at High Drain Bias," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1559–1565, October 1998.

Peter Kinget, "A Fully Integrated 2.7V 0.35 μm CMOS VCO for 5GHz Wireless Applications," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 226–228.

J. Kitchen, "Octave Bandwidth Varactor-Tuned Oscillators," *Microwave J.*, **30**(5), 347–353, May 1987.

A. Kral, F. Behbahani, and A. A. Abidi, "RF-CMOS Oscillators with Switched Tuning," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference*, pp. 555–558.

Vênceslav F. Kroupa, ed., *Direct Digital Frequency Synthesizers.* New York: IEEE Press, 1999.

K. Kurokawa, "Some Basic Characteristics of Broadband Negative Resistance Oscillator Circuits," *Bell Syst. Tech. J.*, 1937–1955, July 1969.

Fred Labaar, "New Discriminator Boosts Phase-Noise Testing," *Microwaves*, 65–69, March 1982.

A. L. Lacaita and C. Samori, "Phase-Noise Performance of Crystal-Like *LC* Tanks," *IEEE Trans. Circuits Syst. Analog Digital Signal Processing*, **45**, 898–900, 1998.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

M. Madihian and T. Noguchi, "An Analytical Approach to Microwave GaAsFET Oscillator Design," *NEC Res. Dev.*, No. 82, 49–56, July 1986.

J. O. McSpadden, L. Fan, and K. Chang, "High-Efficiency *Ku*-Band Oscillators," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1566–1571, October 1998.

Mini-Circuits Engineering, "Designers' Considerations for VCOs Used in Phase-Locked Loops," *Microwave Prod. Digest*, 24, 34, 36, 37, 40, February 1999.

T. Nakagawa and T. Ohira, "A Novel Phase Noise Reduction Technique for MMIC Frequency Synthesizers that Uses Newly Developed Pulse Generator LSI," NTT Radio Communication Systems Laboratories, 1-2356 Take, Yokosuka-shi, Kanagawa 238-03, Japan.

Mohamed K. Nezami, "Evaluate the Impact of Phase Noise on Receiver Performance," *Microwaves & RF*, **41**(5), 165–172, May 1998.

E. C. Niehenke and R. D. Hess, "A Microstrip Low Noise X-Band Voltage-Controlled Oscillator," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**(12), 1075–1079, October 1986.

Chris O'Connor, "Develop Trimless Voltage-Controlled Oscillators," *Microwaves & RF*, **38**(7), 68–78, July 1999.

T. Ohira et al., "A Compact Full MMIC Module for Ku-Band Phase-Locked Oscillators," *IEEE Trans. Microwave Theory Tech.*, **MTT-37**(4), 723–728, April 1989.

D. P. Owen, "Measurement of Phase Noise in Signal Generators," *Marconi Instrum.,* **15**(6), 117–122, 19xx.

D. F. Page, "A Design Basis for Junction Transistor Oscillator Circuits," *Proc. IRE,* **46**, 1271–1280, November 1958.

Byeong-Ha Park, "1 GHz, Low-Power, Low-Phase-Noise, CMOS Voltage-Controlled Oscillator with an Integrated LC Resonator," private communication with the author.

D. F. Paterson, "Varactor Properties for Wideband Linear-Tuning Microwave VCOs," *IEEE Trans. Microwave Theory Tech.*, **MTT-28**(2), 110–119, February 1980.

Matthew Pope, Russell Crouch, and Jeff Gudewicz, "Dual Band Voltage Controlled Oscillators for Personal Communication," *Microwave Prod. Digest,* 52, 84, 85, May 1999.

Nenad Popovic, "Review of Some Types of Varactor Tuned DROs," *Appl. Microwave & Wireless,* **11**(8), 62–70, August 1999.

R. A. Pucel and J. Curtis, "Near-Carrier Noise in FET Oscillators," *IEEE MTT-S International Microwave Symposium Digest*, 1983, pp. 282–284.

S. Qi, K. Wu, and Z. Ou, "Hybrid Integrated HEMT Oscillator with a Multiple-Ring Nonradiative Dielectric (NRD) Resonator Feedback Circuit," *IEEE Trans. Microwave Theory Tech.,* **MTT-46**, 1552–1558, October 1998.

M. Regis, O. Llopis, and J. Graffeuil, "Nonlinear Modeling and Design of Bipolar Transistors Ultra-Low Phase-Noise Dielectric-Resonator Oscillators," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1589–1593, October 1998.

V. Rizzoli, F. Mastri, and D. Masotti, "A General Purpose Harmonic Balance Approach to the Computation of Near-Carrier Noise in Free-Running Microwave Oscillators," *MTT-S,* **x**, 309–312, 1993.

M. Runkel, "Analyse und Vergleich Spannungsgesteuerter Oszillatoren für Mobilfunkanwendungen," Master Thesis, Gerhard-Mercator-Universistät Gesamthochschule Duisburg, Chaired by Prof. Dr. Ingo Wolff, August 1994.

Carlo Samori, Andrea L. Lacaita, Francesco Vill, and Franco Zappa, "Spectrum Folding and Phase Noise in *LC* Tuned Oscillators," *IEEE Trans. Circuits Syst. Analog Digital Signal Processing,* **45**, 781–790, 1998.

Sarafian and B. Z. Kaplan, "A New Approach to the Modeling of the Dynamics of RF VCOs and Some of Its Practical Implications," *IEEE Trans. Circuits Syst. Fundam. Theory Applic.,* **40**(12), 895–901, December 1993.

G. Sauvage, "Phase Noise in Oscillators: A Mathematical Analysis of Leeson's Model," *IEEE Trans. Instrum. Meas.*, **IM-26**(4), x–x, December 1977.

Jwo-Shiun Sun, "Design and Analysis of Microwave Varactor-Tuned Oscillators," *Microwave J.,* 302–310, May 1999.

R. J. Trew, "Octave-Band GaAs FET YIG-Tuned Oscillators," *Electron. Lett.,* **13**(21), October 1977.

R. J. Trew, "MESFET Models for Microwave Computer-Aided Design," *Microwave J.,* **33**(5) 115–128, May 1990.

T. Yamasaki and E. Nagata, *Microwave Synthesizer for Terrestrial and Satellite Communications,* NEC Corporation, June 1989.

N. L. Wang and W. J. Ho, "X-Band HBT VCO with High-Efficiency CB Buffer Amplifier," *IEEE J. Solid-State Circuits,* **SC-27**(10), 1439–1443, October 1992.

Shing-Tak Yan and Howard C. Luong, "A 3-V 1.3-to-1.8-GHz CMOS Voltage-Controlled Oscillator with 0.3-ps Jitter," *IEEE Trans. Circuits Syst. Analog Digital Signal Processing,* **45**, 876–880, 1998.

D. A Warren, J. M. Golio, and W. L. Seely, "Large and Small Signal Oscillator Analysis," *Microwave J.*, **32**(5), 229–246, May 1989.

R. Winch, "Wideband Varactor-Tuned Oscillators," *IEEE J. Solid-State Circuits,* **SC-17**, 1214–1219, December 19, 1982.

Markus Zannoth, Bernd Kolb, Josef Fenk, and Robert Weigel, "A Fully Integrated VCO at 2 GHz," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers,* pp. 224–225.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

5

RF/WIRELESS OSCILLATORS

5-1 INTRODUCTION TO FREQUENCY CONTROL

Practically all modern telecommunication and test equipment uses frequency-control techniques based on *frequency synthesis*, the production of the many frequencies involved in a radiocommunication system's modulation, transmission, reception, and demodulation functions through the combination and mathematical manipulation of a very few input frequencies. Although the frequency synthesis techniques now ascendant in wireless systems—phase-locked loops and, less commonly, direct digital synthesizers—are fundamentally different, all are ultimately based on RF oscillators. This chapter covers oscillator theory, evaluation, and design.

Two types of oscillators are needed in a phase-locked-loop system (Figure 5-1). One, typically a *crystal oscillator*, generates the synthesizer's reference signal. As Figure 5-1 reflects, the reference oscillator may be built into the synthesizer IC in highly integrated systems. The other oscillator, a *voltage-controlled oscillator* (VCO), is varied in frequency by the system to produce the synthesizer's output signal. Although designing good oscillators remains somewhat like black magic or a special art, we will show that mathematics and CAD tools, applied in conjunction with practical experience, can keep the oscillator design process well under control.

5-2 BACKGROUND

The oscillator was probably first discovered by people who wanted to build an amplifier back in the vacuum-tube days. Its modern equivalent is shown in Figure 5-2. Here we see an active device, a bipolar transistor or field-effect transistor, with a tuned input and output circuit. Under ideal circumstances, the voltage at the input is 180° out of phase and the amplifier is "stable." Because of the unavoidable feedback capacitance, a certain amount of energy is transferred from the output to the input at about 90° out of phase. If one of the tuned circuits is detuned to the point where the phase shift is –180°, oscillation will occur. This means that the small energy from the output will be amplified, brought back to the input in phase and

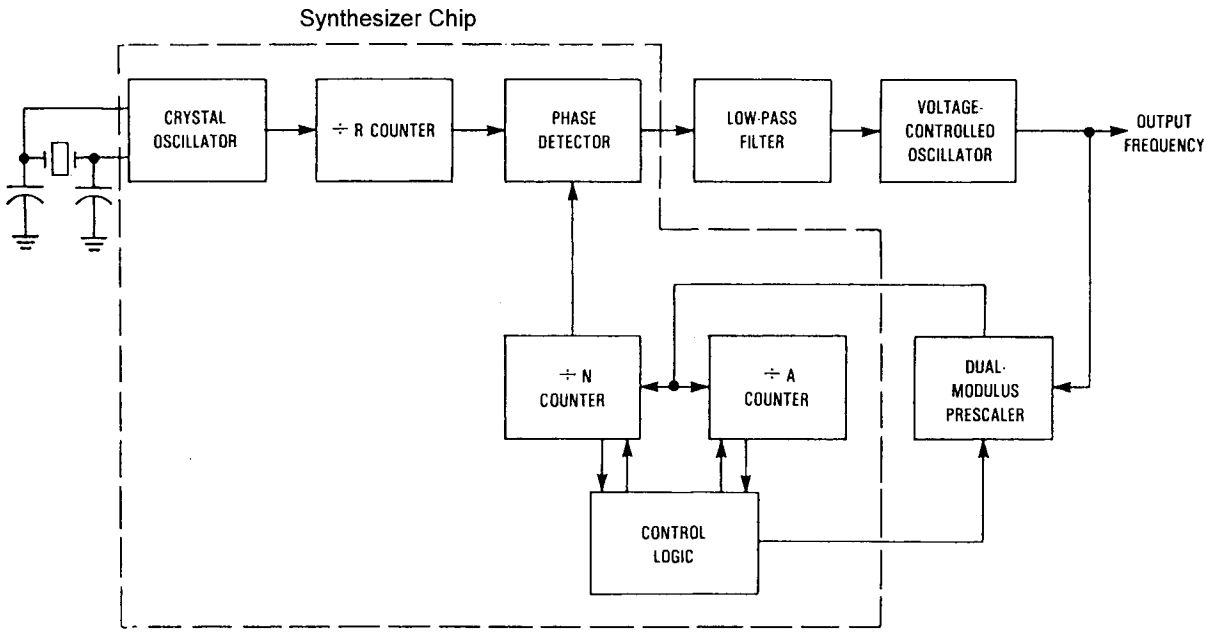


Figure 5-1 Block diagram of a modern, integrated frequency synthesizer. In this case, the designer has control over the VCO and the loop filter; the reference oscillator is part of the chip. In most cases (up to 2.5 GHz), the dual-modulus prescaler is inside the chip.

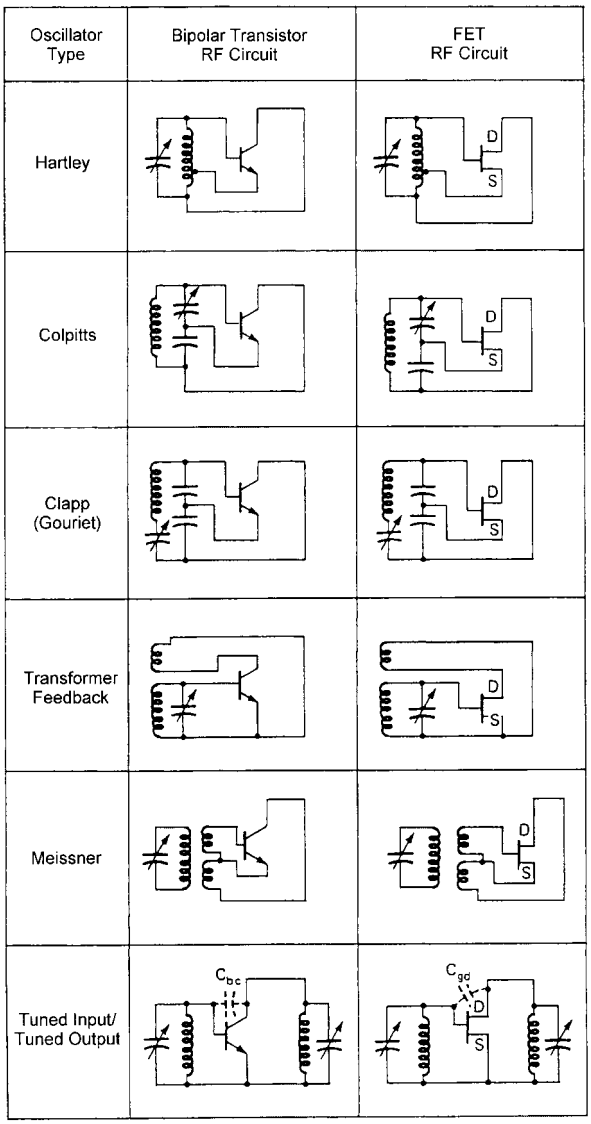


Figure 5-2 Schematic diagrams of RF connections for common oscillator circuits (dc and biasing circuits not shown) [1].

further amplified, causing the stage to "take off" and oscillate. Unless there is some mechanism to limit the amplitude of the oscillation at the input or output, its amplitude can theoretically increase to a level sufficient to destroy the device through breakdown effects or thermal runaway.

The criterion for oscillation was first described by Barkhausen [see Eq. (5-6) in the next section]. The conditions for oscillation exist when the feedback gain is high enough to cancel all losses and the difference between forward gain and reverse gain is less than zero. This implies also that oscillator circuits (Figure 5-2) provide a negative resistance, which is responsible for oscillation. The following mathematical treatment explains this.

5-3 OSCILLATOR DESIGN

A phase-locked loop generally has two oscillators: the oscillator at the output frequency and the reference oscillator. The reference oscillator at times can be another loop that is being mixed in, and the voltage-controlled oscillator (VCO) is controlled by either the reference or the oscillator loop. The VCO is one of the most important parts in a phase-locked-loop system, the performance of which is determined only by the loop filter at offsets inside the loop bandwidth, and only by the quality of the VCO design at offsets outside the loop bandwidth.

To some designers, VCO design appears to be magic. Shortly, we will go through the mathematics of the oscillator and some of its design criteria, but the results have only limited meaning. This is due to component tolerances, stray effects, and, most of all, nonlinear performance of the oscillator device, which can be modeled with only a certain degree of accuracy. However, after building oscillators for awhile, a certain feeling will be acquired for how to do this, and certain performance behavior will be predicted on a rule-of-thumb basis rather than on precise mathematical effort. For reasons of understanding, we will deal with the necessary mathematical equations, but we consider it essential to explain that these are only approximations.

5-3-1 Basics of Oscillators

An electronic oscillator is a device that converts dc power to a periodic output signal (ac power). If the output waveform is approximately sinusoidal, the oscillator is referred to as *sinusoidal.* There are many other oscillator types normally referred to as *relaxation* oscillators. For application in frequency synthesizers, we will explore only sinusoidal oscillators for reasons of purity and noise-sideband performance.

All oscillators are inherently nonlinear. Although the nonlinearity results in some distortion of the signal, linear analysis techniques can normally be used for the analysis and design of oscillators. Figure 5-3 shows, in block diagram form, the necessary components of an oscillator. It contains an amplifier with frequency-dependent forward loop gain $G(j\omega)$ and a frequency-dependent feedback network $H(j\omega)$.

The output voltage is given by

$$V_o = \frac{V_{\text{in}} G(j\omega)}{1 + G(j\omega)H(j\omega)} \tag{5-1}$$

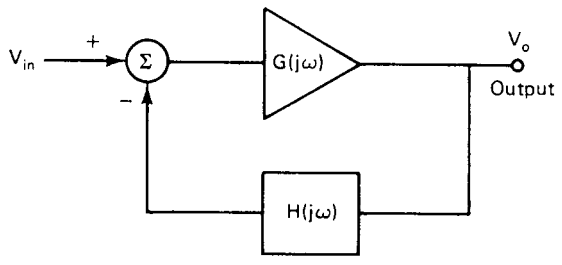


Figure 5-3 Block diagram of an oscillator showing forward and feedback loop components.

For an oscillator, the output V_o is nonzero even if the input signal $V_i = 0$. This can only be possible if the forward loop gain is infinite (which is not practical), or if the denominator

$$1 + G(j\omega)H(j\omega) = 0 \tag{5-2}$$

at some frequency ω_o. This leads to the well-known condition for oscillation (the *Nyquist criterion*), where at some frequency ω_o

$$G(j\omega_o)H(j\omega_o) = -1 \tag{5-3}$$

That is, the magnitude of the open-loop transfer function is equal to 1:

$$|G(j\omega_o)H(j\omega_o)| = 1 \tag{5-4}$$

and the phase shift is 180°:

$$\arg[G(j\omega_o)H(j\omega_o)] = 180° \tag{5-5}$$

This can be more simply expressed as follows: If in a negative-feedback system, the open-loop gain has a total phase shift of 180° at some frequency ω_o, the system will oscillate at that frequency provided that the open-loop gain is unity. If the gain is less than unity at the frequency where the phase shift is 180°, the system will be stable, whereas if the gain is greater than unity, the system will be unstable.

This statement is not correct for some complicated systems, but it is correct for those transfer functions normally encountered in oscillator design. The conditions for stability are also known as the *Barkhausen criteria*, which state that if the closed-loop transfer function is

$$\frac{V_o}{V_i} = \frac{\mu}{1 - \mu\beta} \tag{5-6}$$

the system will oscillate provided that $\mu\beta = 1$. This is equivalent to the Nyquist criterion, the difference being that the transfer function is written for a loop with positive feedback. Both versions state that the total phase shift around the loop must be 360° at the frequency of oscillation and the magnitude of the open-loop gain must be unity at that frequency.

The following analysis of the relatively simple oscillator shown in Figure 5-4 illustrates the design method. The linearized (and simplified) equivalent circuit of Figure 5-4 is given in Figure 5-5. h_{rb} has been neglected, and $1/h_{ob}$ has been assumed to be much greater than the load resistance R_L and is also ignored.

Note that the transistor is connected in the common-base configuration, which has no voltage phase inversion (the feedback is positive), so the conditions for oscillation are

$$|G(j\omega_o)H(j\omega_o) = 1| \tag{5-7}$$

and

$$\arg[G(j\omega_o)H(j\omega_o)] = 0° \tag{5-8}$$

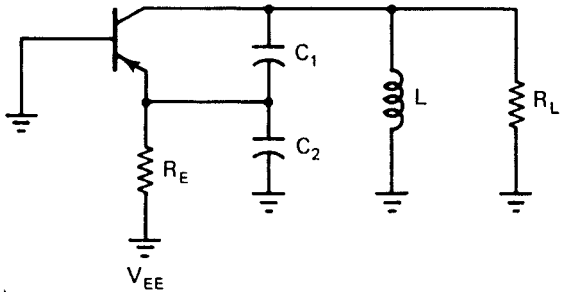


Figure 5-4 Oscillator with capacitive voltage divider.

The circuit analysis can greatly be simplified by assuming that

$$\frac{1}{\omega(C_2 + C_1)} << \frac{h_{ib}R_E}{h_{ib} + R_E} \tag{5-9}$$

and also that the Q of the load impedance is high. In this case the circuit reduces to that of Figure 5-6, where

$$V = \frac{V_o C_1}{C_1 + C_2} \tag{5-10}$$

and

$$R_{\text{eq}} = \frac{h_{ib}R_E}{h_{ib} + R_E}\left(\frac{C_1 + C_2}{C_1}\right)^2 \tag{5-11}$$

Figure 5-7 shows how the input impedance of an oscillator circuit—in this case, a BJT Colpitts oscillator minus its resonator—satisfies this condition for oscillation.

Then the forward gain

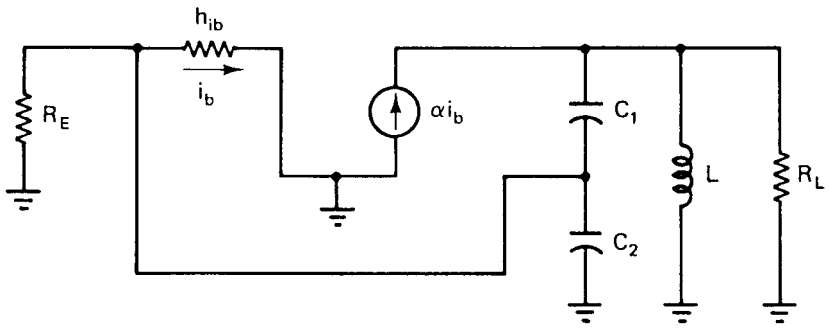


Figure 5-5 Linearized and simplified equivalent circuit of Figure 5-4.

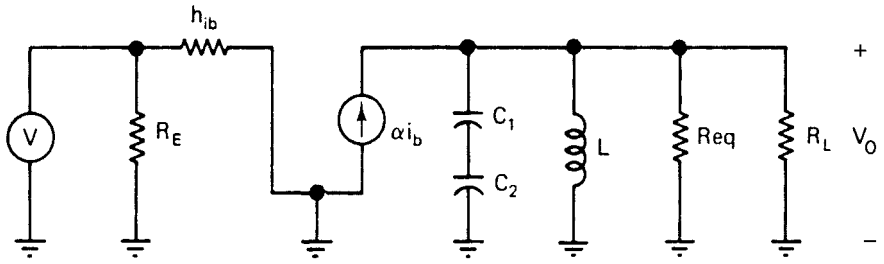


Figure 5-6 Further simplification of Figure 5-4, assuming high-impedance loads.

$$G(j\omega) = \frac{h_{fb}}{h_{ib}} Z_L = \frac{\alpha}{h_{ib}} Z_L \tag{5-12}$$

and

$$H(j\omega) = \frac{C_1}{C_1 + C_2} \tag{5-13}$$

where

$$Y_L = \frac{1}{Z_L} = \frac{1}{j\omega L} + \frac{1}{R_{\text{eq}}} + \frac{1}{R_L} + \frac{1}{j\omega C} \tag{5-14}$$

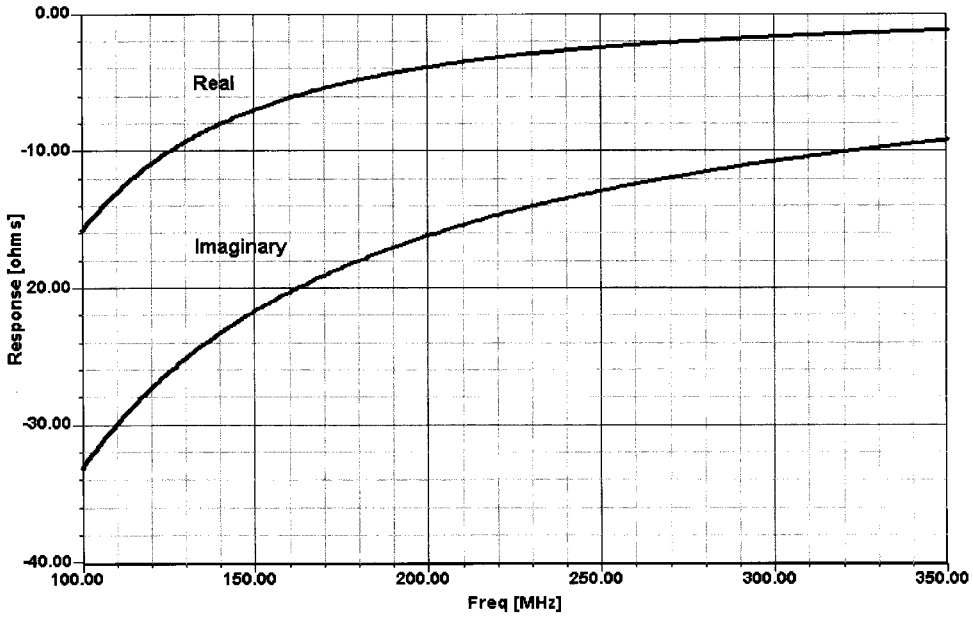


Figure 5-7 Calculated input impedance of a 200-MHz BJT Colpitts oscillator with its resonator removed. (Figure 5-46 shows an equivalent graph with the resonator present.)

A necessary condition for oscillation is that

$$\arg[G(j\omega)H(j\omega)] = 0^\circ \tag{5-15}$$

Since H does not depend on frequency in this example, if arg(GH) is zero, the phase shift of the load impedance Z_L must also be zero. This occurs only at the resonant frequency of the circuit,

$$\omega_o = \frac{1}{\sqrt{L[C_1C_2/(C_1 + C_2)]}} \tag{5-16}$$

At this frequency

$$Z_L = \frac{R_{\text{eq}}R_L}{R_{\text{eq}} + R_L} \tag{5-17}$$

and

$$GH = \frac{h_{fb}}{h_{ib}}\left(\frac{R_{\text{eq}}R_L}{R_{\text{eq}} + R_L}\right)\frac{C_1}{C_1 + C_2} \tag{5-18}$$

The other condition for oscillation is the magnitude constraint that

$$G(j\omega)H(j\omega) = \frac{\alpha}{h_{ib}}\left(\frac{R_{\text{eq}}R_L}{R_{\text{eq}} + R_L}\right)\frac{C_1}{C_1 + C_2} = 1 \tag{5-19}$$

Three-Reactance Oscillators. Although the block-diagram formulation of the stability criteria is an easy way to express stability mathematically, it is frequently not the easiest to apply since it is often difficult to identify the forward loop gain $G(j\omega)$ and the feedback ratio $H(j\omega)$ in electronic systems. A direct analysis of the circuit equations is frequently simpler than the block diagram interpretation (particularly for single-stage amplifiers). Figure 5-8 shows a generalized circuit for an electronic amplifier.

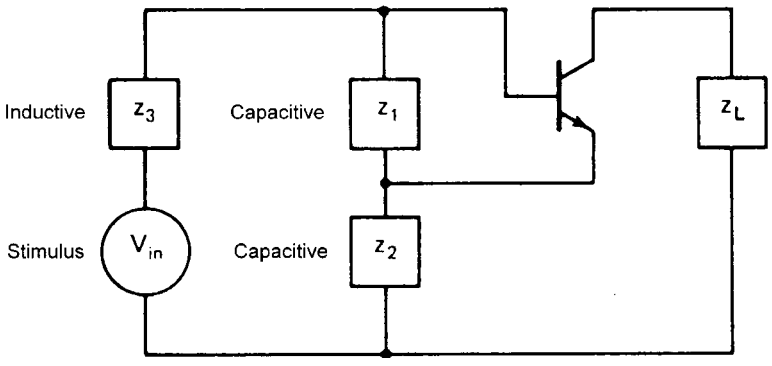


Figure 5-8 Generalized circuit for an oscillator using an amplifier model. Z_3 is inductive even with a capacitor in series with it.

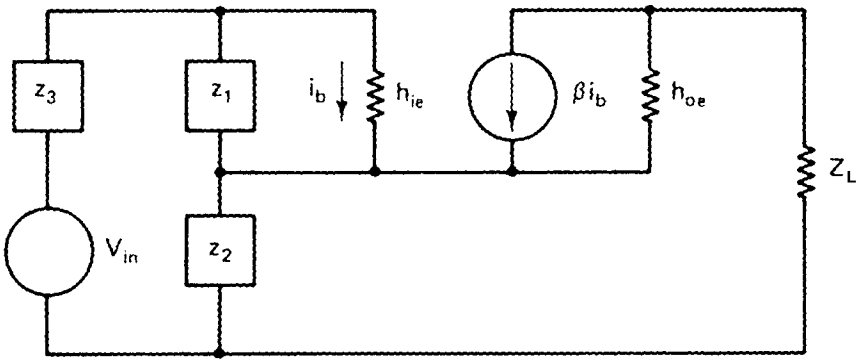


Figure 5-9 Small-signal equivalent circuit of Figure 5-8.

By inspecting Figure 5-8, one can see the three necessary reactances and the load. All other known circuits are derivatives of this by rotation. The small-signal equivalent circuit is given in Figure 5-9 (where h_{re} has been neglected).

Normally, h_{oe} can also be assumed sufficiently small and can be neglected. The loop equations are then

$$V_{\text{in}} = I_1(Z_3 + Z_1 + Z_2) - I_b Z_1 + \beta I_b Z_2 \tag{5-20}$$

$$0 = -I_1 Z_1 + I_b(h_{ie} + Z_1) \tag{5-21}$$

For the amplifier to oscillate, the currents I_b and I_1 must be nonzero even when $V_1 = 0$. This is only possible if the system determinant

$$\Delta = \begin{vmatrix} Z_3 + Z_1 + Z_2 & \beta Z_2 - Z_1 \\ -Z & h_{ie} + Z_1 \end{vmatrix} \tag{5-22}$$

is equal to 0. That is,

$$(Z_3 + Z_1 + Z_2)(h_{ie} + Z_1) - Z_1^2 + \beta Z_1 Z_2 = 0 \tag{5-23}$$

which reduces to

$$(Z_1 + Z_2 + Z_3)h_{fe} + Z_1 Z_2 \beta + Z_1(Z_2 + Z_3) = 0 \tag{5-24}$$

Assume for the moment that Z_1, Z_2, and Z_3 are purely reactive impedances. [It is easily seen that Eq. (5-24) does not have a solution if all three impedances are real.] Since both the real and imaginary parts must be zero, Eq. (5-24) is equivalent to the following equations if

$$h_{ie}(Z_1 + Z_2 + Z_3) = 0 \tag{5-25}$$

and

$$Z_1[(1 + \beta)Z_2 + Z_3] = 0 \tag{5-26}$$

Since β is real and positive, Z_2 and Z_3 must be of opposite sign for Eq. (5-26) to hold. That is,

$$(1 + \beta)Z_2 = -Z_3 \tag{5-27}$$

Therefore, since h_{ie} is nonzero, Eq. (5-25) reduces to

$$Z_1 + Z_2 - (1 + \beta)Z_2 = 0 \tag{5-28}$$

or

$$Z_1 = \beta Z_2 \qquad [C_1 < C_2; L_1 > L_2] \tag{5-29}$$

Thus, since β is positive, Z_1 and Z_2 will be reactances of the same kind. If Z_1 and Z_2 are capacitors, Z_3 is an inductor and the circuit is as shown in Figure 5-10. It is referred to as a *Colpitts* oscillator, named after the person who first described it.

If Z_1 and Z_2 are inductors and Z_3 is a capacitor as illustrated in Figure 5-11, the circuit is called a *Hartley* oscillator.

Example 1. Design a Colpitts circuit to oscillate at 200 MHz, using a transistor (operating at I_C = 6 mA dc) that has an input impedance

$$\text{Mag}(Z_{11}) = \frac{26 \text{ mV}}{I_C} \beta = 282\ \Omega (h_{ie} \triangleq Z_{11})$$

and

$$\beta_{RF} = 65; \quad \beta_{OSC} = 10 \quad \text{(large-signal condition)}$$

The oscillating transistor operates under large-signal conditions, generating harmonics in addition to the fundamental. Most of the transistor's current gain will come from dc biasing. β_{OSC}, its large-signal beta at the fundamental frequency of oscillation, must therefore be considerably less than β_{RF}, the small-signal, single-tone value, as the

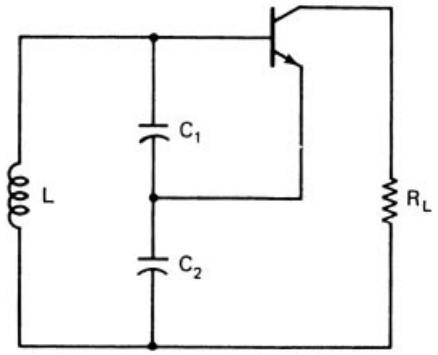


Figure 5-10 Colpitts oscillator.

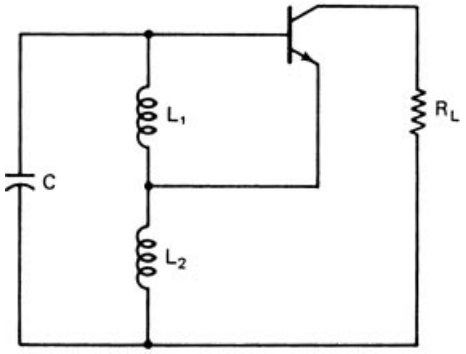


Figure 5-11 Hartley oscillator.

transistor's available current gain, β_0, is distributed among the dc, fundamental, and harmonic components present ($\beta_0 = \beta_{dc} + \beta_{OSC} + \beta_{OSC,F2} + \beta_{OSC,F3}\ldots$).

Solution: For the Colpitts circuit, Z_1 and Z_2 are capacitive reactances and Z_3 is an inductive reactance. Let $Z_2 = -j1.6\ \Omega$; then [Eq. (5-27)]

$$Z_3 = -(1 + \beta)Z_2 = -(1 + 10) \times (-1.6) \approx j17\ \Omega$$

and [Eq. (5-25)]

$$Z_1 = -(Z_2 + Z_3) \approx -j17\ \Omega$$

At 200 MHz, these impedances correspond to component values of $C_1 = 50$ pF, $C_2 = 500$ pF, and $L = 30$ nH. The oscillator will probably work with C_2 values larger than 500 pF, but at the loss of safety margin for tolerances in production, beta, temperature, and phase-noise performance. The completed circuit, except for biasing, is shown in Figure 5-12.

If Z_1, Z_2, Z_3, or h_{ie} is complex, the preceding analysis is more complicated, but the conditions for oscillation can still be obtained from Eq. (5-24). For example, if, in the Colpitts

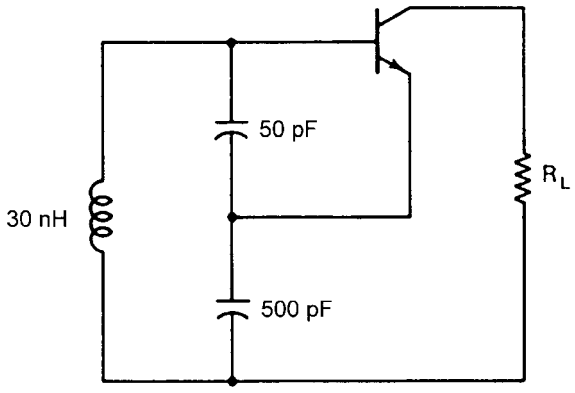


Figure 5-12 Design example of a Colpitts oscillator.

circuit, there is a resistor R in series with $L(Z_3 = R + j\omega L)$, Eq. (5-24) reduces to two equations:

$$h_{ie}\left(\omega L - \frac{1}{\omega C_1} - \frac{1}{\omega C_2}\right) - \frac{R}{\omega C_1} = 0 \tag{5-30}$$

and

$$h_{ie}R - \frac{1+\beta}{\omega^2 C_1 C_2} + \frac{1}{\omega C_1}\omega L = 0 \tag{5-31}$$

Define

$$C_1' = \frac{C_1}{1 + R/h_{ie}} \tag{5-32}$$

The resonant frequency at which oscillations will occur is found from Eq. (5-30) to be

$$\omega_o = \frac{1}{\sqrt{L[C_1' C_2/(C_1' + C_2)]}} \tag{5-33}$$

and for oscillations to occur $R_e(h_{ie}) = R_e(Z_{11})$ must be less than or equal to

$$R_e(h_{ie}) \leq \frac{1+\beta}{\omega_o^2 C_1 C_2} - \frac{L}{C_1} \tag{5-34}$$

If R becomes too large, Eq. (5-32) cannot be satisfied and oscillations will stop. In general, it is advantageous to have

$$X_{C_1} X_{C_2} = \frac{1}{\omega^2 C_1 C_2} \tag{5-35}$$

with $X \triangleq 1/\omega c$ as large as possible, since R can be large. However, if C_1 and C_2 are too small (large and X_{C_1} and X_{C_2}), the input and output capacitances of the transistor, which shunt C_1 and C_2, respectively, become important. A good, stable design will always have C_1 and C_2 much larger than the transistor capacitances they shunt.

Example 2. In Example 1, will the circuit still oscillate if the inductor now has a $Q_u = 100$? If the transistor input capacitance is 5 pF, what effect will this have on the system?

Solution: In Example 1, $X_L = 17\ \Omega$, $C_1 = 50$ pF, and $C_2 = 500$ pF. Since $C_1 = 50$ pF, adding 5 pF in parallel will change the equivalent C to 55 pF. As the inductor $Q_u = 100$, the equivalent resistance R in series with the lossless inductor is

$$R = \frac{17}{100} = 0.17\ \Omega$$

The new resonant frequency can be determined from Eqs. (5-32) and (5-33):

$$C_1' = \frac{C_1}{1 + R/h_{ie}} = \frac{55}{1 + (0.17/280)} \approx \frac{55}{1.0006} = 54.96 \text{ pF}$$

and

$$f_o = \frac{1}{2\pi L[C_1' C_2/(C_1' + C_2)]^{1/2}} \cong 191.56 \text{ MHz}$$

The effect of the finite inductor Q causes a negligible change in the oscillating frequency compared to the effect of the transistor input capacitance, which reduces the resonant frequency 4.2%. Because the large-signal beta, β_{OSC}, is 10, there is enough loop gain to maintain oscillation.

Two-Port Oscillator. Although Eqs. (5-25) and (5-26) can be used to determine the exact expressions for oscillation, they are often difficult to use and add little insight into the design process. An alternative interpretation, although not as accurate, will now be presented. It is based on the fact that an ideal tuned circuit (infinite Q), once excited, will oscillate infinitely because there is no resistance element present to dissipate the energy. In the actual case where the inductor Q is finite, the oscillations die out because energy is dissipated in the resistance. It is the function of the amplifier to maintain oscillations by supplying an amount of energy equal to that dissipated. This source of energy can be interpreted as a negative resistor in series with the tuned circuit. If the total resistance is positive, the oscillations will die out, while the oscillation amplitude will increase if the total resistance is negative. To maintain oscillations, the two resistors must be of equal magnitude. To see how a negative resistance is realized, the input impedance of the circuit in Figure 5-13 will be derived.

If h_{oe} is sufficiently small ($h_{oe} << 1/R_L$), the equivalent circuit is as shown in Figure 5-13. The steady-state loop equations are

$$V_{\text{in}} = I_{\text{in}}(X_{C_1} + X_{C_2}) - I_b(X_{C_1} - \beta X_{C_2}) \tag{5-36}$$

$$0 = -I_{\text{in}}(X_{C_1}) + I_b(X_{C_1} + h_{ie}) \tag{5-37}$$

After I_b is eliminated from these two equations, Z_{in} is obtained as

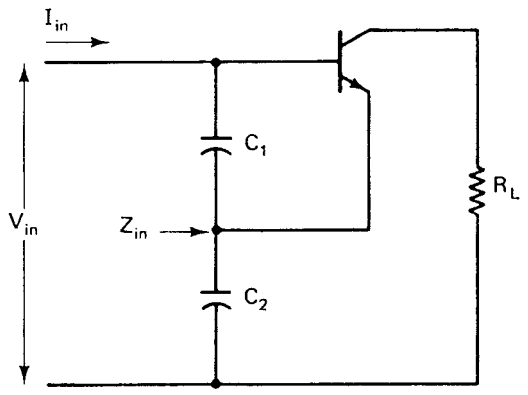


Figure 5-13 Calculation of input impedance of the negative-resistance oscillator.

$$Z_{\text{in}} = \frac{V_{\text{in}}}{I_{\text{in}}} = \frac{(1+\beta)X_{C_1}X_{C_2} + h_{ie}(X_{C_1} + X_{C_2})}{X_{C_1} + h_{ie}} \tag{5-38}$$

If $X_{C_1} << h_{ie}$, the input impedance is approximately equal to

$$Z_{\text{in}} \approx \frac{1+\beta}{h_{ie}} X_{C_1}X_{C_2} + (X_{C_1} + X_{C_2}) \tag{5-39}$$

$$Z_{\text{in}} \approx \frac{-g_m}{\omega^2 C_1 C_2} + \frac{1}{j\omega[C_1C_2/(C_1 + C_2)]} \tag{5-40}$$

That is, the input impedance of the circuit shown in Figure 5-14 is a negative resistor,

$$R = \frac{-g_m}{\omega^2 C_1 C_2} \tag{5-41}$$

in series with a capacitor,

$$C_{\text{in}} = \frac{C_1 C_2}{C_1 + C_2} \tag{5-42}$$

which is the series combination of the two capacitors.

With an inductor L (with the series resistance R_S) connected across the input, it is clear that the condition for sustained oscillation is

$$R_S = \frac{g_m}{\omega^2 C_1 C_2} \tag{5-43}$$

and the frequency of oscillation is

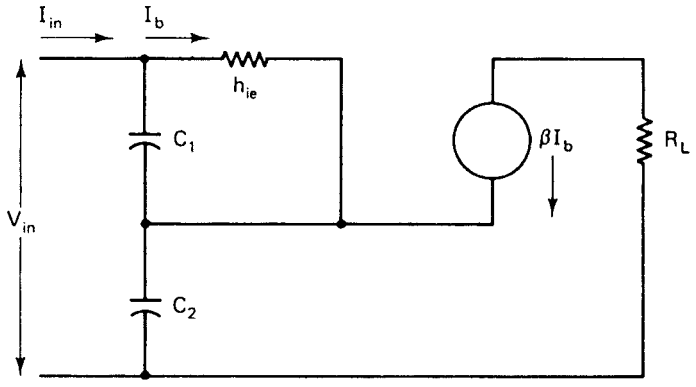


Figure 5-14 Equivalent small-signal circuit of Figure 5-13.

$$f_o = \frac{1}{2\pi\sqrt{L[C_1C_2/(C_1 + C_2)]}} \tag{5-44}$$

This interpretation of the oscillator readily provides several guidelines that can be used in the design. First, C_1 should be as large as possible so that

$$X_{C_1} << h_{ie} \tag{5-45}$$

and C_2 is to be large so that

$$X_{C_2} << \frac{1}{h_{oe}} \tag{5-46}$$

When these two capacitors are large, the transistor base-to-emitter and collector-to-emitter capacitances will have a negligible effect on the circuit's performance. However, Eq. (5-43) limits the maximum value of the capacitances since

$$r \leq \frac{g_m}{\omega^2 C_1 C_2} \leq \frac{G}{\omega^2 C_1 C_2} \tag{5-47}$$

where G is the maximum value of g_m. For a given product of C_1 and C_2, the series capacitance is at maximum when $C_1 = C_2 = C_m$. Thus Eq. (5-47) can be written

$$\frac{1}{\omega C_m} > \sqrt{\frac{r}{G}} \tag{5-48}$$

This equation is important in that it shows that for oscillations to be maintained, the minimum permissible reactance $1/\omega C_m$ is a function of the resistance of the inductor and the transistor's mutual conductance, g_m.

An oscillator circuit known as the *Clapp circuit* or *Clapp–Gouriet circuit* is shown in Figure 5-15. This oscillator is equivalent to the one just discussed, but it has the practical advantage of being able to provide another degree of design freedom by making C_o much smaller than C_1 and C_2.

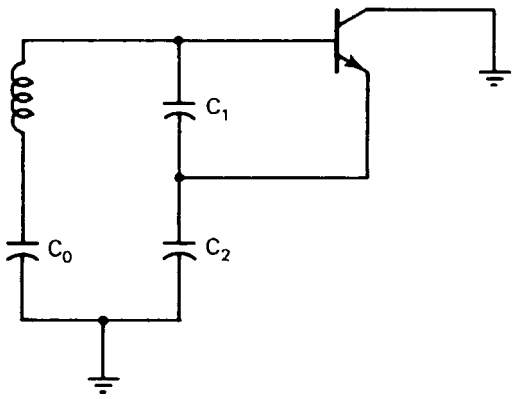


Figure 5-15 Circuit of a Clapp oscillator.

It is possible to use C_1 and C_2 to satisfy the condition of Eq. (5-47) and then adjust C_o for the desired frequency of oscillation ω_o, which is determined from

$$\omega_o L - \frac{1}{\omega_o C_o} - \frac{1}{\omega_o C_1} - \frac{1}{\omega_o C_2} = 0 \tag{5-49}$$

Amplitude Stability. Linearized analysis of the oscillator is convenient for determining the frequency but not the amplitude of the oscillation. The Nyquist stability criterion defines the frequency of oscillation as the frequency at which the loop phase shift is 360°, but it says nothing about the oscillation amplitude. If no provisions are taken to control the amplitude, it is susceptible to appreciable drift. Two frequently used methods for controlling the amplitude are operating the transistor in the nonlinear region or using a second stage for amplitude limiting. For the single-stage oscillator, amplitude limiting is accomplished by designing an unstable oscillator; that is, the loop gain is made greater than 1 at the frequency at which the phase shift is 180°. As the amplitude increases, the β of the transistor decreases, causing the loop gain to decrease until the amplitude stabilizes. This is a self-limiting oscillator. There are nonlinear analysis techniques predicting the amplitude of oscillation, but their results are approximate except in idealized cases, forcing the designer to resort to an empirical approach.

An example of a two-stage emitter-coupled oscillator is shown in Figure 5-16. In this circuit, amplitude stabilization occurs as a result of current limiting in the second stage. This circuit has the additional advantage that it has output terminals that are isolated from the feedback path. The emitter signal of Q_2, having a rich harmonic content, is normally used as output. Harmonics of the fundamental frequency can be extracted at the emitter of Q_2 by using an appropriately tuned circuit. Note that the collector of Q_2 is isolated from the feedback path.

Phase Stability. An oscillator has a frequency or phase stability that can be considered in two separate parts. First, there is *long-term* stability, in which the frequency changes over a period of minutes, hours, days, weeks, or even years. This frequency stability is normally limited by the circuit components' temperature coefficients and aging rates. The other part,

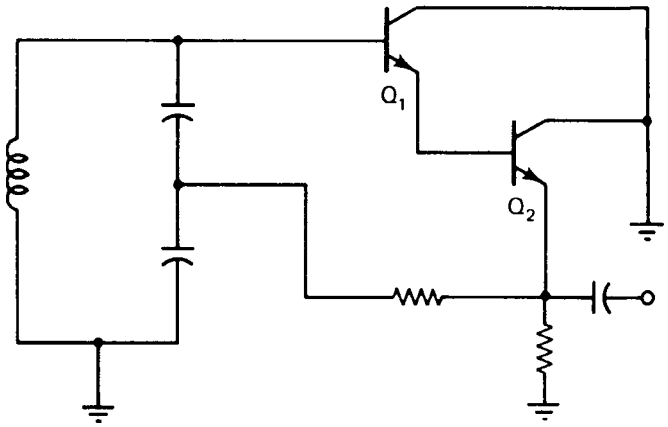


Figure 5-16 Two-stage, emitter-coupled oscillator.

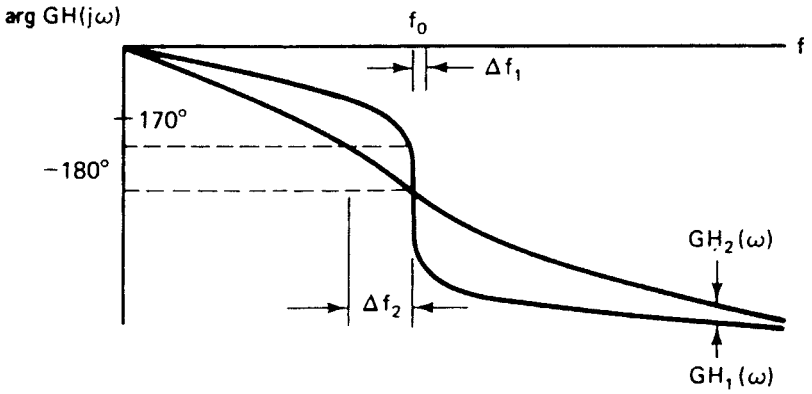


Figure 5-17 Phase plot of two open-loop systems with different resonator *Q* values.

short-term frequency stability, is measured in terms of seconds. One form of short-term instability is due to changes in phase of the system; here the term *phase stability* is used synonymously with frequency stability. It refers to how the frequency of oscillation reacts to small changes in phase shift of the open-loop system. It can be assumed that the system with the largest rate of change of phase versus frequency $d\phi/df$ will be the most stable in terms of frequency stability. Figure 5-17 shows the phase plots of two open-loop systems used in oscillators. At the system crossover frequency, the phase shift is −180°. If some external influence causes a shift in phase—say, it adds 10° of phase lag—the frequency will change until the total phase shift is again 0°. In this case, the frequency will decrease to the point where the open-loop phase shift is 170°. Figure 5-17 shows that Δf_2, the change in frequency associated with the 10° change in phase of GH_2, is greater than the change in frequency Δf_1, associated with the open-loop system GH_1, whose phase is changing more rapidly near the open-loop crossover frequency.

This qualitative discussion illustrates that $d\phi/df$ at $f = f_o$ is a measure of an oscillator's phase stability. It provides a good means of quantitatively comparing the phase stability of two oscillators. Consider the simple parallel tuned circuit shown in Figure 5-18.

For this circuit, the two-port is

$$\frac{V_o(j\omega)}{I(j\omega)} = \frac{R}{1 + jQ[(\omega/\omega_o) - (\omega_o/\omega)]} \tag{5-50}$$

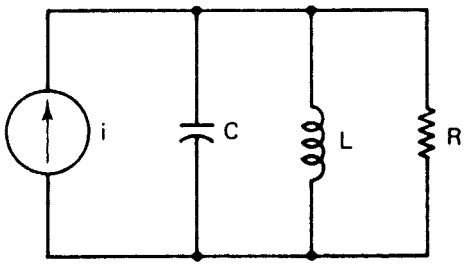


Figure 5-18 Parallel tuned circuit for phase-shift analysis.

where

$$\omega_o = \frac{1}{\sqrt{LC}} \quad \text{and} \quad Q = \frac{R}{\omega_o L} \tag{5-51}$$

The circuit phase shift is

$$\arg \frac{V_o}{I} = \theta = \tan^{-1} Q\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right) \tag{5-52}$$

and

$$\frac{d\theta}{d\omega} = \frac{1/Q}{1/Q^2 + [(\omega^2 - \omega_o^2)/\omega_o\omega]^2} \frac{\omega^2 + \omega_o^2}{(\omega_o\omega)^2} \tag{5-53}$$

At the resonant frequency ω_o,

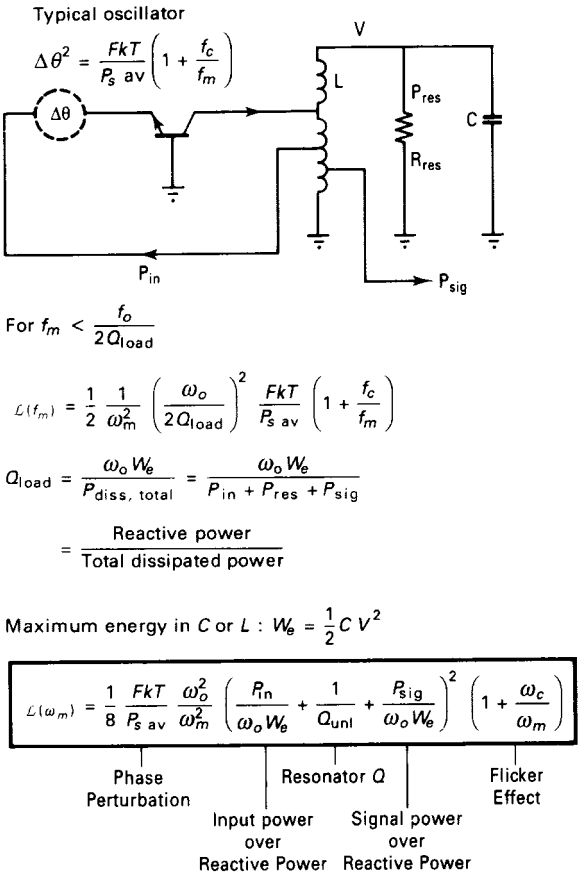


Figure 5-19 Diagram for a feedback oscillator illustrating the principles involved, and showing the key components considered in the phase noise calculation and its contribution.

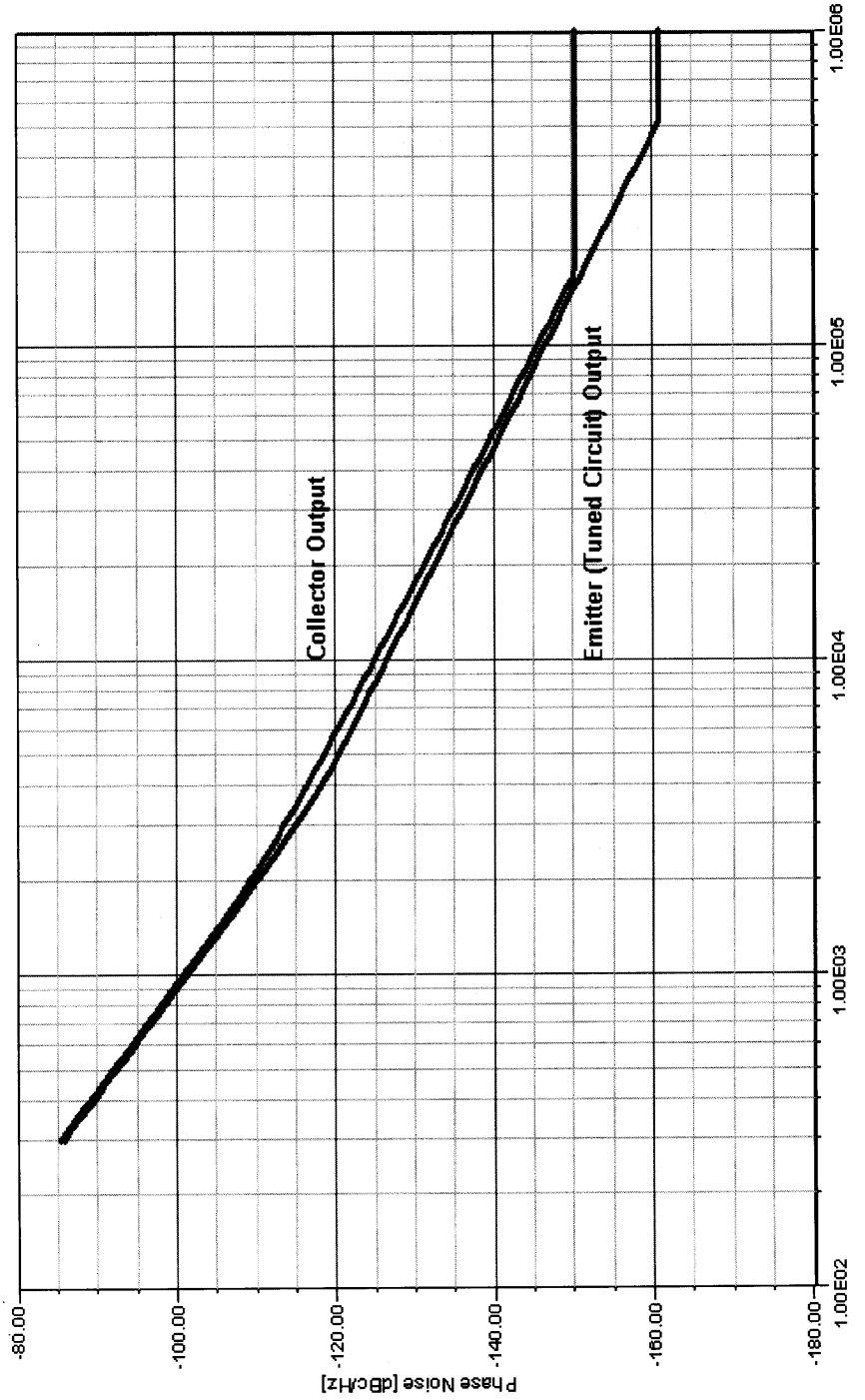


Figure 5-20 Phase noise with oscillator output taken from collector (upper trace) versus emitter (lower trace) of a BJT oscillator.

$$\left.\frac{dQ}{d\omega}\right|_{\omega=\omega_o} = \frac{2Q}{\omega_o} \tag{5-54}$$

The frequency stability factor S_F is defined as the change in phase $d\phi/d\omega$ divided by the normalized change in frequency $\Delta\omega/\omega_o$. That is,

$$S_F = 2Q \tag{5-55}$$

S_F is a measure of the short-term stability of an oscillator. Equation (5-54) indicates that the higher the circuit Q, the higher the stability factor. This is one reason for using high-Q circuits in oscillator circuits. Another reason is the ability of the tuned circuit to filter out undesired harmonics and noise.

The two oscillator types shown so far can be considered two-port oscillators because one port is responsible for oscillation and the other port supplies the output power. This type of output supposedly has better isolation (although this is not true at higher frequencies) but has a higher noise floor. A better way of extracting energy is shown in Figure 5-19.

These types of oscillators are essentially what is referred to as *two-port* oscillators. The name comes from the fact that the oscillating device has two ports, with the output energy taken from the output port, typically the collector or drain, while the tuned circuit is at the input port. For years, it was widely held that this type of design provided better isolation between the tuned circuit and termination; unfortunately, because of the Miller effect, this is not true. At the higher frequencies, changes in output-port loading affect the input port, resulting in *pulling*, a detuning effect that increases with frequency. In addition, taking the output from the collector or the drain results in a poorer signal-to-noise ratio than that achievable by extracting energy from the tuned circuit itself. The only real advantages the collector output provides are more power and higher efficiency, but definitely at the expense of phase noise, which is often referred to as *SSB phase noise* because it is expressed in terms of the decibel ratio of the noise power in a single (the upper or lower) noise sideband, in a 1-Hz bandwidth centered at a specified frequency offset from the oscillator carrier, to the carrier power. Figure 5-20 compares the phase noise of an oscillator with the energy taken off the collector and the tuned circuit; the difference is highly visible.

5-4 OSCILLATOR CIRCUITS

5-4-1 Hartley

Figure 5-21 shows the Hartley oscillator configuration, in which feedback is obtained by tapping the resonator. In practice, the resonator can consist of a single tapped inductor or two separate, magnetically uncoupled inductors, with feedback obtained at their junction.

5-4-2 Colpitts

Figure 5-22 shows the Colpitts oscillator configuration, in which feedback is obtained via a capacitive voltage divider.

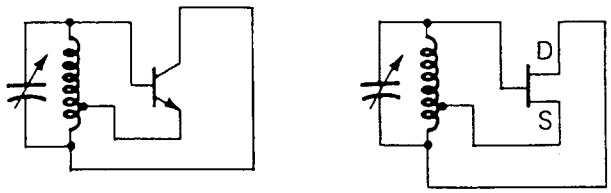


Figure 5-21 Hartley oscillator.

5-4-3 Clapp–Gouriet

Figure 5-23 shows the Clapp–Gouriet oscillator. Like the Colpitts, the Clapp–Gouriet obtains its feedback via a capacitive voltage divider; unlike the Colpitts, an additional capacitor series-tunes the resonator. The Pierce oscillator, a configuration used only with crystals, is a rotation of the Clapp–Gouriet oscillator in which the emitter is at RF ground.

5-5 DESIGN OF RF OSCILLATORS

5-5-1 General Thoughts on Transistor Oscillators

An estimate of the noise performance of an oscillator is

$$\mathcal{L}(\omega_m) = \frac{1}{8}\frac{FkT}{P_{s,\text{av}}}\frac{\omega_0^2}{\omega_m^2}\left(\frac{P_{\text{in}}}{\omega_0 W_e} + \frac{1}{Q_{\text{unl}}} + \frac{P_{\text{sig}}}{\omega_0 W_e}\right)^2\left(1 + \frac{\omega_c}{\omega_m}\right) \tag{5-56}$$

Phase perturbation

Resonator Q

Flicker effect

Input power/Reactive power ratio

Signal power/Reactive power ratio

We just state this here without presenting oscillator noise theory itself. (Also see Figure 5-19.) This equation is based on work done by Dieter Scherer of Hewlett-Packard about 1978. He was the first to introduce the flicker effect to the Leeson equation by adding the AM-to-PM

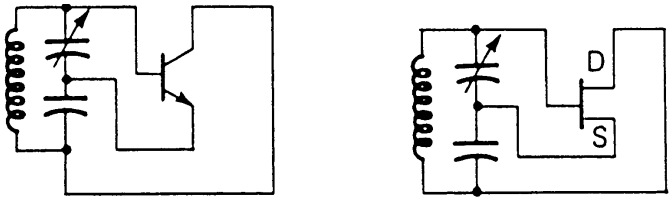


Figure 5-22 Colpitts oscillator.

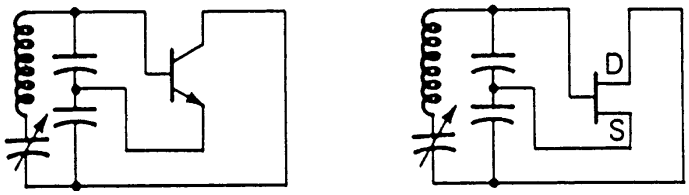


Figure 5-23 Clapp–Gouriet oscillator.

conversion effect, which is caused by the nonlinear capacitance of the active devices. This equation must be further expanded:

$$\pounds(f_m) = 10 \log\left\{\left[1 + \frac{f_0^2}{(2f_m Q_{\text{load}})^2}\right]\left(1 + \frac{f_c}{f_m}\right)\frac{FkT}{2P_{\text{sav}}} + \frac{2kTRK_0^2}{f_m^2}\right\} \tag{5-57}$$

where $\pounds(f_m)$ = ratio of sideband power in 1-Hz bandwidth at f_m to total power in dB
f_m = frequency offset
f_0 = center frequency
f_c = flicker frequency
Q_{load} = loaded Q of the tuned circuit
F = noise factor
$kT = 4.1 \times 10^{-21}$ at 300 K (room temperature)
P_{sav} = average power at oscillator output
R = equivalent noise resistance of tuning diode (typically 200 Ω to 10 kΩ)
K = oscillator voltage gain

The following table (from Motorola) shows the flicker corner frequency f_C as a function if I_C for a typical small-signal microwave BJT. $I_{C(\text{max})}$ of this transistor is about 10 mA.

I_C (mA)	f_C (kHz)
0.25	1
0.5	2.74
1	4.3
2	6.27
5	9.3

Note that f_C, which is defined by AF and KF in the SPICE model, increases with I_C. This gives us a clue about how f_C changes when a transistor oscillates. As a result of the bias-point shift that occurs during oscillation (see Example 1 of Section 5-3-1), an oscillating BJT's average I_C is higher than its small-signal I_C. KF is therefore higher for a given BJT operating as an oscillator than for the same transistor operating as a small-signal amplifier. This must be kept in mind when considering published f_C data, which are usually determined under small-signal conditions without being qualified as such. Determining a transistor's oscillating f_C is best done through measurement: Operate the device as a high-Q UHF oscillator (we

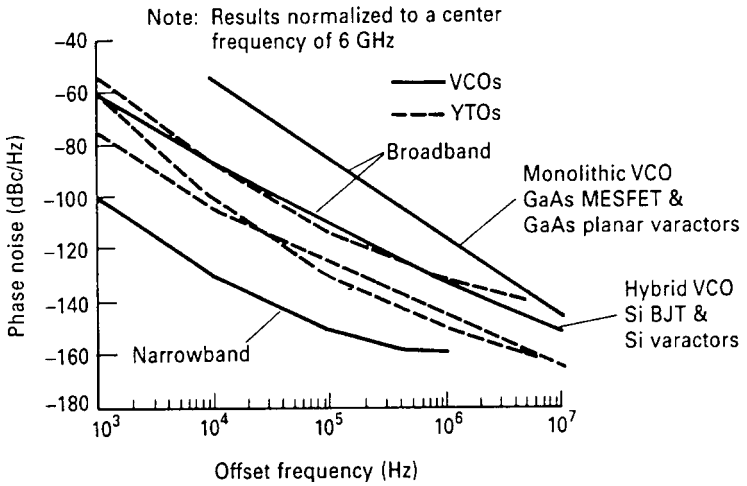


Figure 5-24 Phase noise of oscillators using different semiconductors and resonators.

suggest using a ceramic-resonator-based tank in the vicinity of 1 GHz), and measure its close-in (10 Hz to 10 kHz) phase noise versus offset from the carrier. f_C will correspond to a slight decrease in the slope of the phase noise versus offset curve. Generally, f_C varies with device type as follows: silicon JFETs, 50 Hz and higher; microwave RF BJTs, 1–10 kHz (as above); MOSFETs, 10–100 kHz; and GaAsFETs, 10–100 MHz. Figure 5-24 shows the phase noise of oscillators using different semiconductors and resonators.

Equation (5-57) is based on Rohde et al. [2]. The additional term introduces a distinction between a conventional oscillator and a VCO. Whether the voltage- or current-dependent capacitance is internal or external makes no difference; it simply affects the frequency.

For a more complete expression for a resonator oscillator's phase noise spectrum we can write

$$s_\phi(f_m) = \frac{\alpha_R F_0^4 + \alpha_E (F_0/2Q_L)^2}{f_m^3} + \frac{(2GFkT/P_0)\,(F_0/2Q_L)^2}{f_m^2} + \frac{2\alpha_R Q_L F_0^3}{f_m^2} + \frac{\alpha_E}{f_m} + \frac{2GFkT}{P_0} \tag{5-58}$$

where G = compressed power gain of the loop amplifier
F = noise factor of the loop amplifier
k = Boltzmann's constant
T = temperature in kelvins

P_0 = carrier power level (in watts) at the output of the loop amplifier
F_0 = carrier frequency in hertz
f_m = carrier offset frequency in hertz
$Q_L (= \pi F_0 \tau_g)$ = loaded Q of the resonator in the feedback loop
α_R, α_E = flicker-noise constants for the resonator and loop amplifier, respectively.

In frequency synthesizers, we have no use for *LC* oscillators that do not include a tuning diode or diodes, but it may still be of interest to analyze the low-noise, fixed-tuned *LC* oscillator first and later make both elements, inductor and capacitor, variable.

Later, we will show the performance changes if we utilize the two possible ways of getting coarse and fine tuning in oscillators:

1. Use of tuning diodes.
2. Use of switching diodes.

We will spend some time looking at the effects that switching and tuning diodes have in a circuit because they will ultimately influence the noise performance more strongly than the transistor itself.

The reason is that the noise generated in tuning diodes will be superimposed on the noise generated in the circuit, while switching diodes have losses that cause a reduction of circuit *Q*. The selection of the proper tuning and switching diodes is important, as is the proper way of connecting them. As both types are modifications of the basic *LC* oscillator, we start with the *LC* oscillator itself.

Early signal generators as were offered by several companies (namely, Rohde & Schwarz, Hewlett-Packard, Boonton Electronics, or Marconi), if they are not synthesized, use an air-dielectric variable capacitor or, as in the case of one particular Hewlett-Packard generator, the Model HP8640, a tuned cavity.

Tuning here is accomplished by changing the value of an air-dielectric variable capacitor or changing the mechanical length of a quarter-wave resonator.

Using the equations shown previously, it is fairly easy to calculate the performance of oscillators and understand how they work, but this does not necessarily optimize their design. For *LC* oscillator applications in which noise performance is crucial, the oscillator shown in Figure 5-25, used in the famous Rohde & Schwarz SMDU, is the state of the art. Its noise performance is equivalent to the noise found in the cavity-tuned oscillator used in Hewlett-Packard's popular HP8640 signal generator, and because of the unique way a tuning diode is coupled to the circuit, its modulation capabilities are substantially superior to any of the signal generators currently offered. To develop such a circuit from design equations is not possible. This circuit is the result of many years of experience and research and looks fairly simple. The grounded-gate FET circuit provides the best performance because it fulfills the important requirements of Eq. (5-56).

The tuned circuit is not connected directly to the drain, but the drain is put on a tap of the oscillator section. Therefore, the actual voltage across the tuning capacitor is higher than the supply voltage, and thus the energy stored in the capacitor is much higher than in a circuit connected between the gate electrode and ground, as it is in a normal Colpitts oscillator. In addition, the high output impedance of the FET does not load the circuit, which also provides a reduced noise contribution. Since this oscillator is optimized for best frequency-modulation

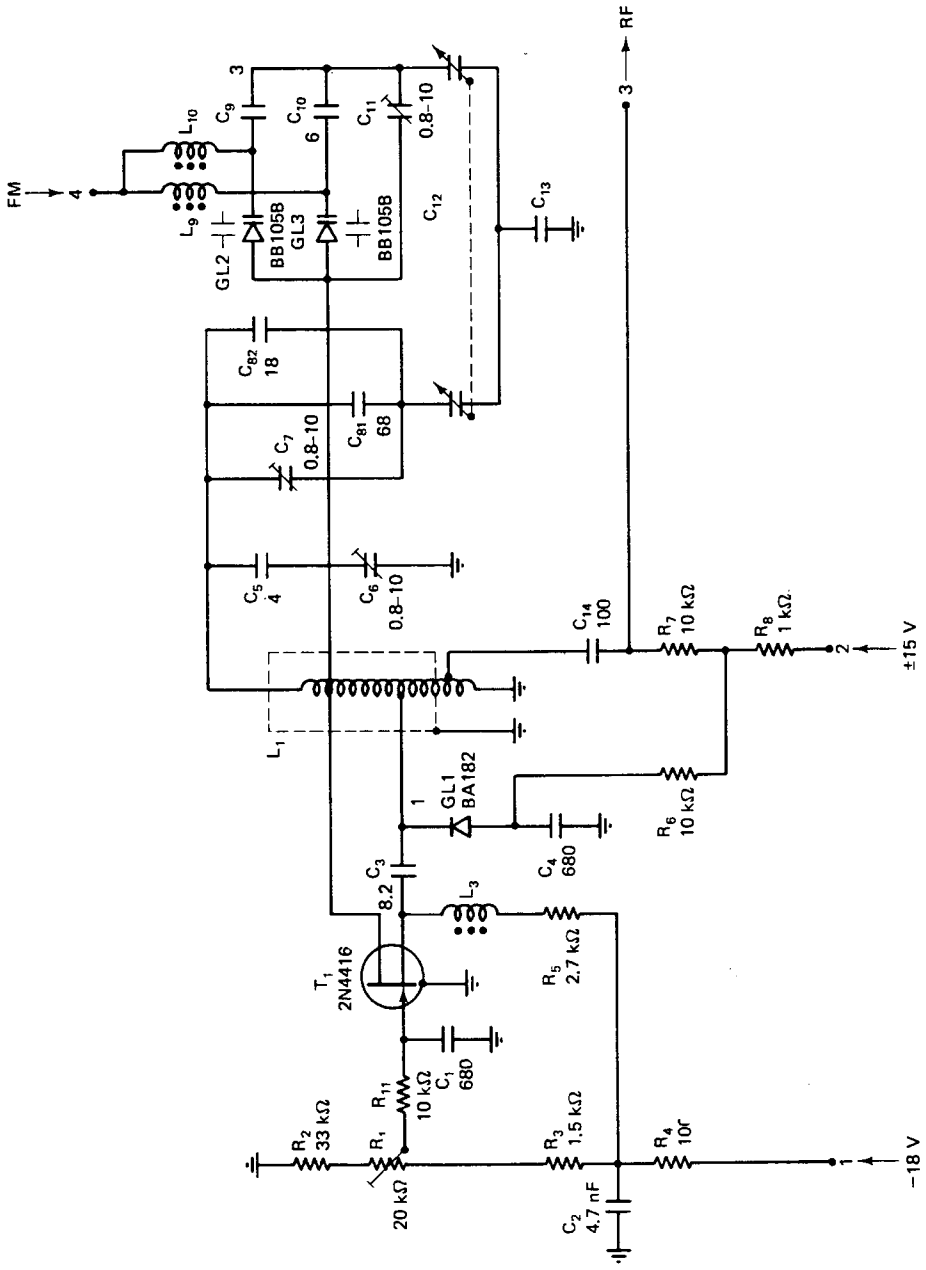


Figure 5-25 A 118–198-MHz oscillator from the Rohde & Schwarz SMDU signal generator.

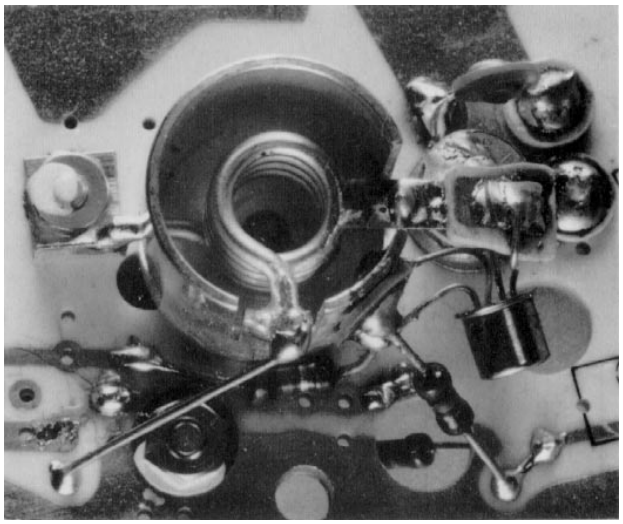


Figure 5-26 Photograph of the helical-resonator system from the Rohde & Schwarz SMDU signal generator.

performance in the FM frequency range, it becomes apparent that it fits the requirements of low-distortion stereo modulation.

For extremely critical blocking measurements in the 2-m region (140–160 MHz), the noise specification at 20 kHz off the carrier is of high importance, while the peak modulation typically does not exceed 5 kHz. Figure 5-25 shows the schematic of the oscillator section optimized for this frequency range.

Because of f_T limitations, it has been found experimentally that these *LC* oscillators, if based on silicon JFETs, should not be used above about 500 MHz; rather, a doubler stage should be employed. Analyzing the signal generators currently on the market, we find that their highest base band typically ranges from 200 to 500 MHz, with frequency doubling used to 1000 MHz. As can be seen in Figure 5-26, the mechanical layout of such an oscillator is extremely compact.

5-5-2 Two-Port Microwave/RF Oscillator Design

A common method for designing oscillators is to resonate the input port with a passive high-*Q* circuit at the desired frequency of resonance. It will be shown that if this is achieved with a load connected on the output port, the transistor is oscillating at both ports and is thus delivering power to the load port. The oscillator may be considered a two-port structure, with M_3 being the lossless resonating port and M_4 providing lossless matching such that all the external RF power is delivered to the load. See Figure 5-27. The resonating network has been described. Nominally, only parasitic resistance is present at the resonating port, since a high-*Q* resonance is desirable for minimizing oscillator noise. It is possible to have loads at both the input and output ports if such an application occurs, since the oscillator is oscillating at both ports simultaneously.

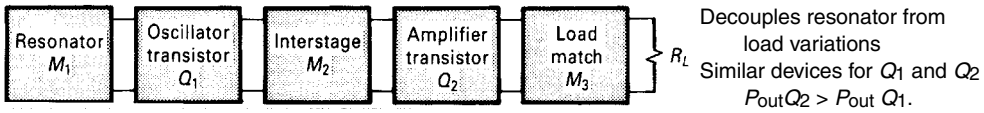


Figure 5-27 Buffered oscillator design.

The simultaneous oscillation condition is proved as follows. Assume that the oscillation condition is satisfied at port 1:

$$1/S'_{11} = \Gamma_G \tag{5-59}$$

Thus,

$$S'_{11} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L} = \frac{S_{11} - D\Gamma_L}{1 - S_{22}\Gamma_L} \tag{5-60}$$

$$\frac{1}{S'_{11}} = \frac{1 - S_{22}\Gamma_L}{S_{11} - D\Gamma_L} = \Gamma_G \tag{5-61}$$

By expanding Eq. (5-61), we find

$$\begin{gathered} \Gamma_G S_{11} - D\Gamma_L\Gamma_G = 1 - S_{22}\Gamma_L \\ \Gamma_L(S_{22} - D\Gamma_G) = 1 - S_{11}\Gamma_G \\ \Gamma_L = \frac{1 - S_{11}\Gamma_G}{S_{22} - D\Gamma_G} \end{gathered} \tag{5-62}$$

Thus,

$$S'_{22} = S_{22} + \frac{S_{12}S_{21}\Gamma_G}{1 - S_{11}\Gamma_G} \tag{5-63}$$

$$\frac{1}{S'_{22}} = \frac{1 - S_{11}\Gamma_G}{S_{22} - D\Gamma_G} \tag{5-64}$$

Comparing Eqs. (5-62) and (5-64), we find

$$1/S'_{22} = \Gamma_L \tag{5-65}$$

which means that the oscillation condition is also satisfied at port 2; this completes the proof. Thus, if either port is oscillating, the other port must be oscillating as well. A load may appear at either or both ports, but normally the load is in Γ_L, the output termination. This result can be generalized to an n-port oscillator by showing that the oscillator is simultaneously oscillating at each port:

$$\Gamma_1 S'_{11} = \Gamma_2 S'_{22} = \Gamma_3 S'_{33} = \cdots = \Gamma_n S'_{nn} \tag{5-66}$$

Before concluding this section on two-port oscillator design, the buffered oscillator shown in Figure 5-27 must be considered. This design approach is used to provide the following:

1. A reduction in load-pulling, which is the change in oscillator frequency when the load reflection coefficient changes.
2. A load impedance that is more suitable to wideband applications ($k < 1$).
3. A higher output power from a working design, although the higher output power can also be achieved by using a larger oscillator transistor.
4. This type of oscillator is not optimized for low phase noise. However, buffered oscillator designs are quite common in wideband YIG applications, where changes in the load impedance must not change the generator frequency.

Two-port oscillator design may be summarized as follows:

1. Select a transistor with sufficient gain and output-power capability for the frequency of operation. This may be based on oscillator datasheets, amplifier performance, or *S*-parameter calculation.
2. Select a topology that gives $k < 1$ at the operating frequency. Add feedback if $k < 1$ has not been achieved.
3. Select an output load-matching circuit that gives $|S'_{11}| > 1$ over the desired frequency range. In the simplest case, this could be a 50-Ω load.
4. Resonate the input port with a lossless termination so that $\Gamma_G S'_{11} = 1$. The value of S'_{22} will be greater than unity with the input properly terminated.

In all cases, the transistor delivers power to a load and the input of the transistor. Practical considerations of realizability and dc biasing will determine the best design.

For both bipolar and FET oscillators, a common topology is the common-base or common-gate, since a common-lead inductance can be used to raise S_{22} to a large value, usually greater than unity even with a 50-Ω generator resistor. However, it is not necessary for the transistor S_{22} to be greater than unity, since the 50-Ω generator is not present in the oscillator design. The requirement for oscillation is $k < 1$; then resonating the input with a lossless termination will ensure that $|S'_{11}| > 1$.

A simple example will clarify the design procedure. A common-base bipolar transistor (HP2001) was selected to design a fixed-tuned oscillator at 2 GHz. The common-base *S*-parameters and stability factors are given in Table 5-1.

Using the load circuit in Figure 5-28, we see that the reflection coefficients are

Table 5-1 *S* parameters and stability factors (V_{CE} = 15 V, I_C = 25 mA)

$L_B = 0$	$L_B = 0.5$ nH
$S_{11} = 0.94 \angle 174°$	$1.04 \angle 173°$
$S_{21} = 1.90 \angle -28°$	$2.00 \angle -30°$
$S_{12} = 0.013 \angle 98°$	$0.043 \angle 153°$
$S_{22} = 1.01 \angle -17°$	$1.05 \angle -18°$
$k = -0.09$	-0.83

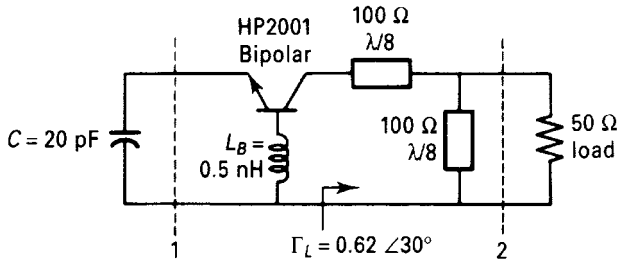


Figure 5-28 Oscillator example at 2 GHz.

$$\Gamma_L = 0.62\ \angle 30°$$

$$S'_{11} = 1.18\ \angle 173°$$

Thus, a resonating capacitance of G = 20 pF resonates the input port. In a YIG-tuned oscillator, this reactive element could be provided by the high-Q YIG element. For a dielectric-resonator oscillator (DRO), the puck would be placed to give $\Gamma_G \approx 1.0/\angle -173°$.

Another two-port design procedure is to resonate the Γ_G port and calculate S'_{22} until $|S'_{22}| > 1$, then design the load port to satisfy. This design procedure is summarized in Figure 5-29.

An example using this procedure at 4 GHz is given in Figure 5-30 using an AT41400 silicon bipolar chip in the common-base configuration with a convenient value of base and emitter

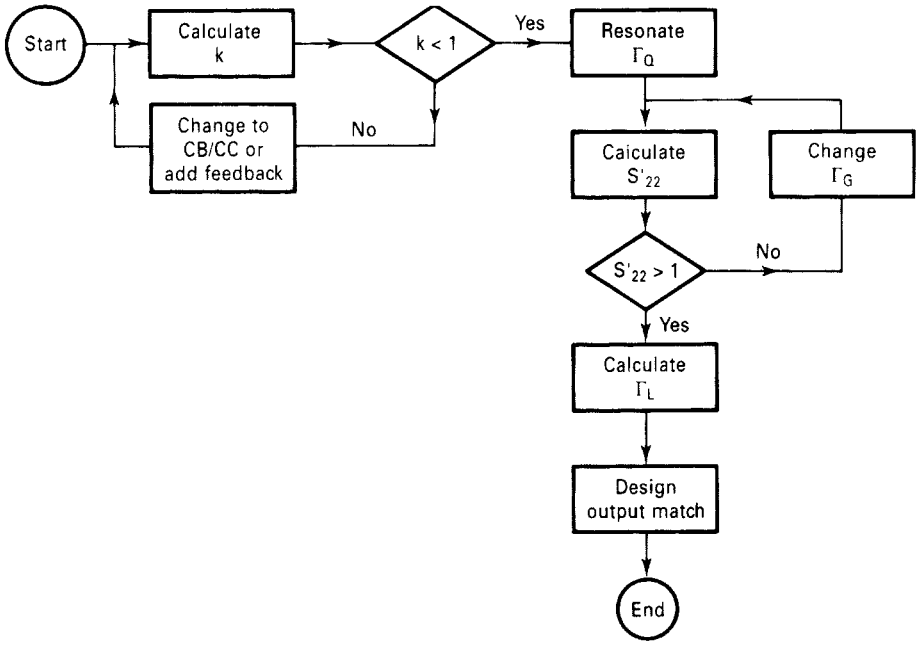


Figure 5-29 Oscillator design flowchart.

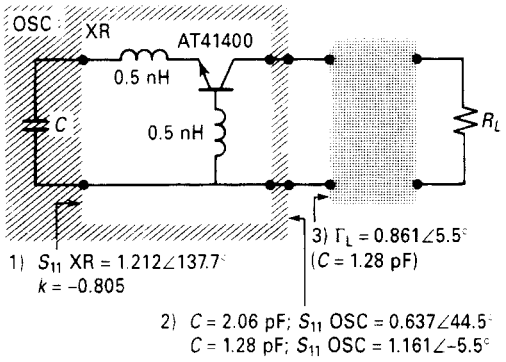


Figure 5-30 A 4-GHz lumped-resonator oscillator using the AT41400.

inductance of 0.5 nH. The feedback parameter is the base inductance, which can be varied if needed.

Since a lossless capacitor at 4 GHz of 2.06 pF gives $\Gamma_G = 1 - 0 \angle -137.7°$, this input termination is used to calculate S'_{22}, giving $S'_{22} = 0.637 \angle 44.5°$. This circuit will oscillate into any passive load. Varying the emitter capacitor about 20° on the Smith Chart to 1.28 pF gives $S'_{22} = 1.16 \angle -5.5°$, which will oscillate into a load of $\Gamma_L = 0.861 \angle 5.5°$. The completed lumped-element design is given in Figure 5-31.

5-5-3 Ceramic-Resonator Oscillators

An important application for a new class of resonators called ceramic resonators (CRs) has emerged for wireless applications. The CRs are similar to rigid coaxial cable, where the center conductor is connected at the end to the outside of the cable. These resonators are generally operating in quarter-wavelength mode and their characteristic impedance is approximately 10 Ω. Because their coaxial assemblies are made from a high-ε low-loss material with good silver plating throughout, the electromagnetic field is internally contained and therefore results in very little radiation. These resonators are therefore ideally suited for high-Q, high-density oscillators. The typical application for this resonator is VCOs ranging from not much more than 200 MHz up to about 3 or 4 GHz. At these high frequencies, the mechanical dimensions of the resonator become too tiny to offer any advantage. One of the

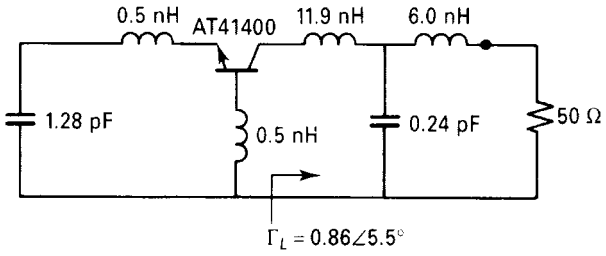


Figure 5-31 Completed lumped-resonator oscillator (LRO).

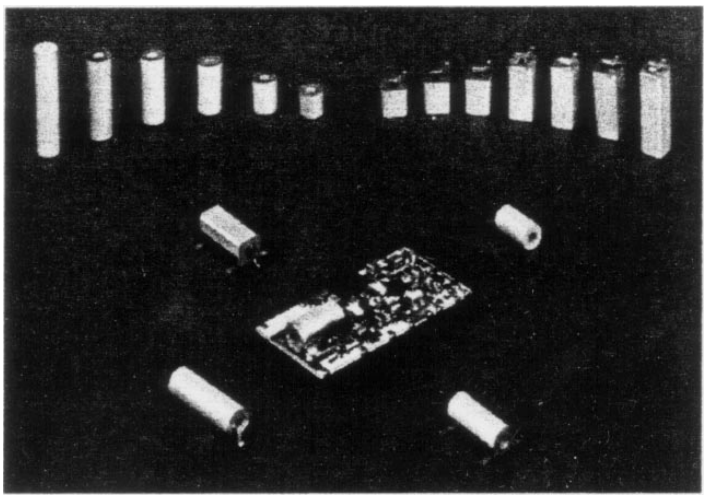


Figure 5-32 Ceramic resonators.

principal requirements is that the physical length is considerably larger than the diameter. If the frequency increases, this requirement can no longer be met.

Manufacturers supply ceramic resonators on a prefabricated basis. Figures 5-32 and 5-33 show the standard round/square packaging available and the typical dimensions for a ceramic resonator. Figure 5-34 shows a ceramic-resonator oscillator built on a ceramic substrate. The available material has a dielectric constant of 88 and is recommended for use in the 400–1500-MHz range. The next higher frequency range (800 MHz to 2.5 GHz) uses an ε of 38, while the top range (1–4.5 GHz) uses an ε of 21. Given the fact that ceramic resonators are prefabricated and have standard outside dimensions, the following quick calculation applies:

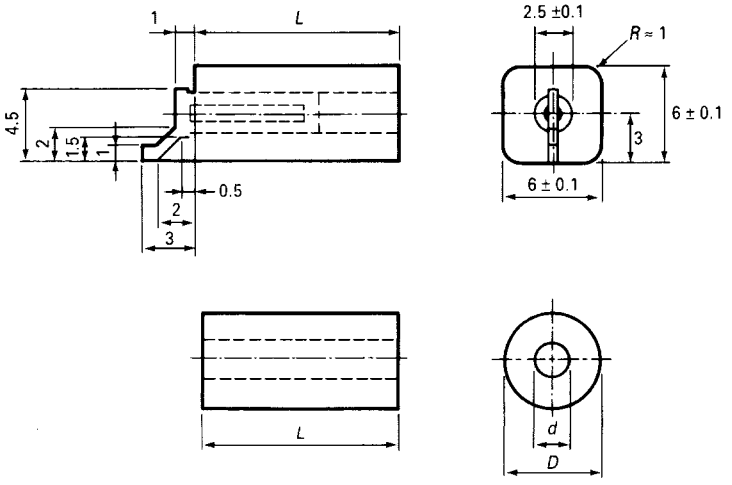


Figure 5-33 Standard round/square ceramic-resonator packaging and dimensions.

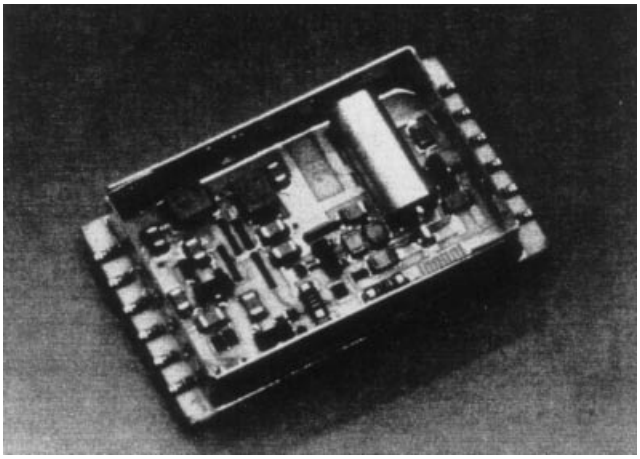


Figure 5-34 Modern ceramic-resonator oscillator for handheld cellular phones.

Relative dielectric constant of resonator material	$\varepsilon_r = 21$	$\varepsilon_r = 38$	$\varepsilon_r = 88$
Resonator length in millimeters	$l = \frac{16.6}{f}$	$l = \frac{12.6}{f}$	$l = \frac{8.2}{f}$
Temperature coefficient (ppm/°C)	10	6.5	8.5
Available temperature coefficients	–3 to +12	–3 to +12	–3 to +12
Typical resonator Q	800	500	400
Frequency range	1–4.5 GHz	800–2500 MHz	400–1500 MHz

Calculation of Equivalent Circuit. The equivalent parallel-resonant circuit afforded by a ceramic resonator has a resistance at resonance of

$$R_p = \frac{2(Z_0)^2}{R^* l} \tag{5-67}$$

where Z_0 = characteristic impedance of the resonator
l = mechanical length of the resonator
R^* = equivalent resistor due to metallization and other losses

As an example, one can calculate

$$C^* = \frac{2\pi\varepsilon_0\varepsilon_r}{\log_e(D/d)} = 55.61 \times 10^{-12}\,\frac{\varepsilon_r}{\log_e(D/d)} \tag{5-68}$$

and

$$L^* = \frac{\mu_r \mu_0}{2\pi} = \log_e\left(\frac{D}{d}\right) = 2 \times 10^{-7} \log_e\left(\frac{D}{d}\right) \tag{5-69}$$

$$Z_0 = 60\ \Omega \frac{1}{\sqrt{\varepsilon_r}} \log_e\left(\frac{D}{d}\right)$$

$$= 60 \frac{1}{\sqrt{88}} \log_e\left(\frac{6}{2.5}\right)$$

$$= 5.6\ \Omega \tag{5-70}$$

A practical example for $\varepsilon_r = 88$ and 450 MHz is

$$C_p = \frac{C^* l}{2} = 49.7\ \text{pF} \tag{5-71}$$

$$L_p = 8L^* l = 2.52\ \text{nH} \tag{5-72}$$

$$R_p = 2.5\ \text{k}\Omega \tag{5-73}$$

Figure 5-35 shows the schematic of a ceramic-resonator-based oscillator. Figures 5-36 and 5-37 show the simulated and measured phase noise of the oscillator.

By using ceramic-resonator-based oscillators in conjunction with miniature synthesizer chips, it is possible to build extremely small phase-locked-loop (PLL) systems for cellular telephone operation. Figure 5-38 shows one of the smallest currently available PLL-based synthesizers manufactured by Synergy Microwave Corporation. Because of the high-Q resonator, these types of oscillators exhibit extremely low phase noise. Values of better than –150 dBc/Hz at 1 MHz off the carrier are achievable. The ceramic resonator reduces the oscillator's sensitivity to microphonic effects and proximity effects caused by other components.

5-5-4 Using a Microstrip Inductor as the Oscillator Resonator*

A high-Q microstrip resonator [3–5] can be used to improve an oscillator's phase-noise performance. The following section, based on information in the Philips Semiconductors Application Note AN1777, describes the application of such a resonator in a BJT differential oscillator.

It can be seen from Eq. (5-56) that the phase noise is proportional to the term

$$\frac{1}{4Q_L^2} \tag{5-74}$$

This means that for any 10% increase in Q, we get something like 20% improvement in phase noise, which makes it most desirable to increase the Q. One way to reduce the loading of the

*Based on portions of the Philips Semiconductors/Signetics RF Communications Products Application Note AN1777, "Low-Voltage Front End Circuits: SA601, SA620," August 20, 1997. Used with permission.

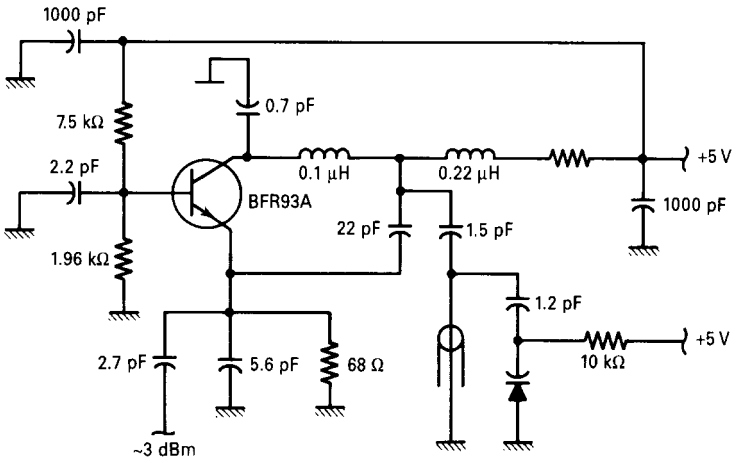


Figure 5-35 Schematic of a high-performance ceramic-resonator-based oscillator that can be used from 500 MHz to 2 GHz.

resonator is to tap the device down on the resonator. The underlying oscillator circuit is based on the differential type of oscillator found in ICs.

Increasing Loaded Q. Figure 5-39 shows the oscillator section of the Philips SA620 low-voltage front-end IC with a conventional second-order parallel-tuned tank circuit configured as the external resonator.

From the basic equations for parallel resonance, the resonator's loaded Q is given by

$$Q_L = R_P \left(\frac{C_T}{L_T} \right)^{1/2} \tag{5-75}$$

where

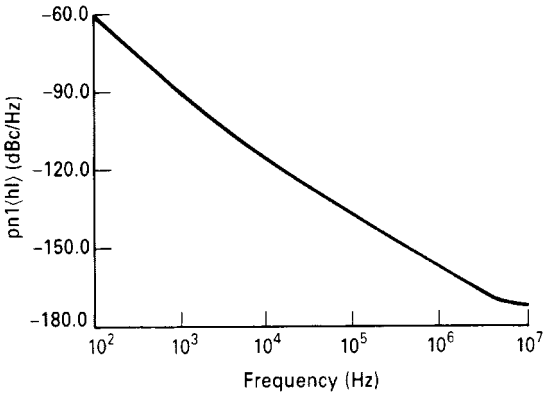


Figure 5-36 Simulated phase noise of an *npn* bipolar 1-GHz ceramic-resonator-based oscillator.

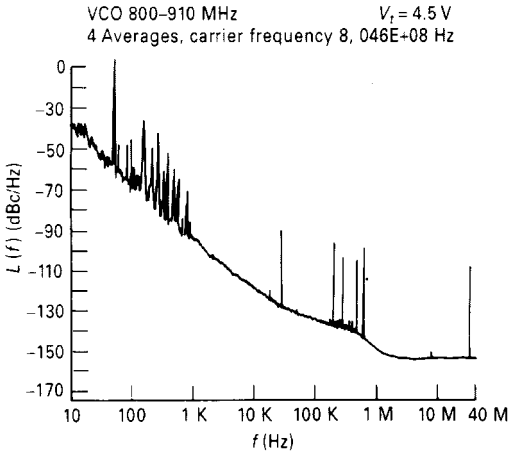


Figure 5-37 Measured phase noise of a ceramic-resonator-based oscillator.

$$R_P = R_T \| R_C = \left[\left(\frac{1}{R_T}\right) + \left(\frac{1}{R_C}\right)\right]^{-1} \tag{5-76}$$

represents the net shunt tank resistance appearing across the network at resonance. R_T represents losses in the inductance L_T and capacitance C_T (almost always dominated by inductor losses), and R_C is the active circuit's load impedance at resonance. Improving the quality of the tank components L_T and C_T (i.e., improving their Q) will increase R_T, but since R_C is low in the case of the SA620, the resulting increase in R_P is relatively small. We can

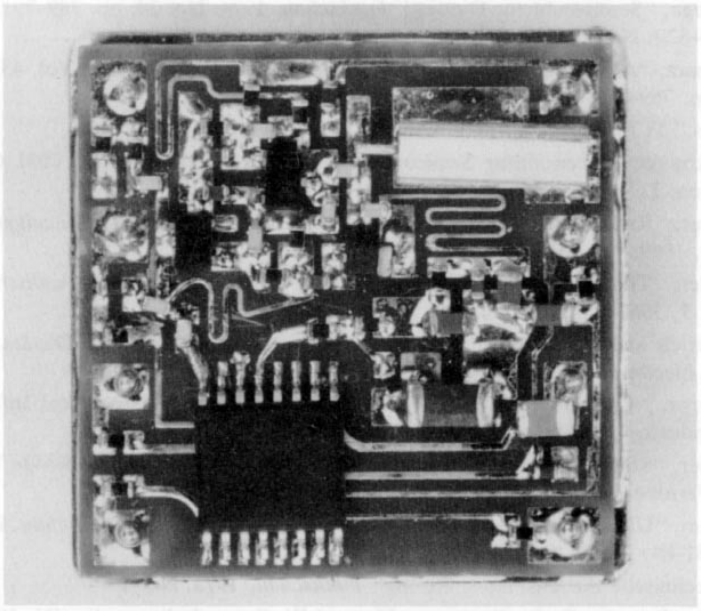


Figure 5-38 Miniature PLL-based synthesizer manufactured by Synergy Microwave Corporation.

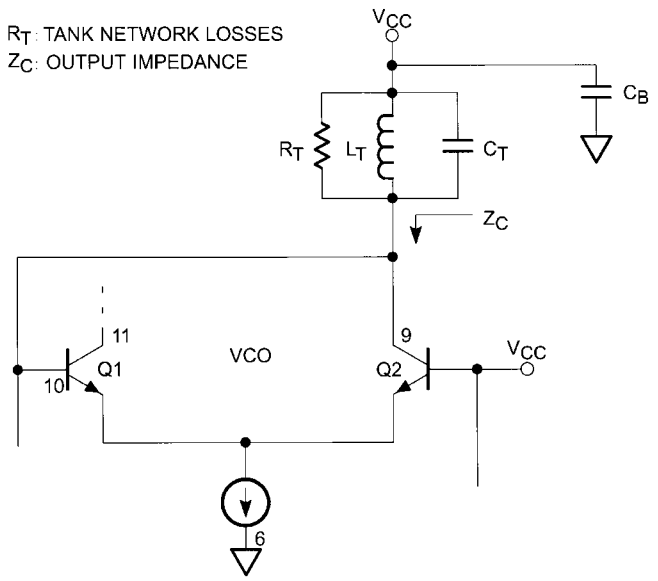


Figure 5-39 Oscillator portion of Philips SA620 front-end IC with external resonator.

increase R_P by decoupling Z_C from the tank circuit (note that R_C is the real part of Z_C at resonance). Decoupling this impedance by using either tapped-L or tapped-C tank configurations is possible. Inspection of the circuit shows that dc biasing is necessary for pins 9 and 10 (osc1 and osc2, respectively), so a tapped-C approach would require shunt-feeding these pins to V_{CC}. Thus, the most practical way is to employ a tapped-L network.

Figure 5-40 shows the basic tapped-L circuit. The effective multiplied shunt impedance at resonance across C_T is given by

$$\hat{R}_P = R_P\left[\left(\frac{L_{T2}}{L_{T1}}\right)+1\right]^{1/2} \tag{5-77}$$

with a corresponding increase in loaded Q:

$$Q_L = Q_L\left[\left(\frac{L_{T2}}{L_{T1}}\right)+1\right]^{1/2} \tag{5-78}$$

Feedback is caused by an RF voltage appearing across L_T, coupled through bypass capacitor C_B and its ground return to the cold end of the current source at pin 6. The significance of this path increases with frequency. At 900 MHz, it becomes very critical in obtaining stable oscillations and must be kept as short as possible. Attention to layout detail is pivotal here!

High-Q Microstrip Inductor. Realizing a stable tapped-L network using lumped surface-mount inductors is impractical due to their low unloaded Q and finite physical size. Another approach uses a high-Q short microstrip inductor. The conventional approach to microstrip resonators treats them as a length of transmission line terminated with a short that

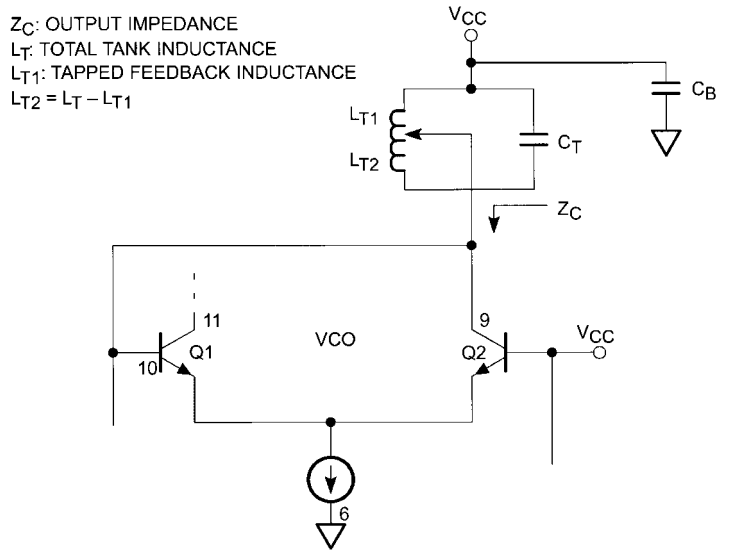


Figure 5-40 Oscillator portion of SA620 IC with tapped-microstrip resonator.

reflects back an inductance impedance that simulates a lumped inductor. The approach presented here deviates from this technique by specifically exploiting the very high unloaded Q attainable with short microstrip inductors, where we design for a large C/L ratio by making the microstrip a specific, but short, length. Resonance is achieved by replacing the short with whatever capacitance is needed for proper resonance. One equation for the inductance of a strip of metal having a length ℓ (mils) and width w (mils) is

$$L = [5.08\text{E} - 3]\,\ell\left[\ln\left(\frac{\ell}{w}\right) + 1.193\left(\frac{w}{\ell}\right)\right] \qquad \text{[3] (5-79)}$$

where L is in nH/mil. This technique results in a much shorter strip of metal for a given inductance. One intriguing property of microstrip inductors is that they are capable of very high unloaded Q values. An equation describing the quality factor for a "wide" microstrip is

$$Q_C = 0.63h[\sigma f_{\text{GHz}}]^{1/2} \qquad \text{[4] (5-80)}$$

where h is the dielectric thickness in centimeters, σ is the conductivity in S/m, and f is frequency in GHz. This equation predicts an unloaded Q exceeding 700 when silver-coated copper is used. Also, wide microstrip lines are defined as those whose strip width w to height h ratio is approximately greater than 1; that is, $w/h > 1$. The parallel capacitor formed by the metal strip over the ground plane should also be included and may be calculated by the classic formula and included in the total needed to resonate with the inductance result found from Eq. (5-79). This "strip" capacitance is given by

$$C_S = K\varepsilon_0\varepsilon_r\left(\frac{w\ell}{h}\right) \qquad \text{(5-81)}$$

where C_S is in farads, $\varepsilon_0 = 8.86\text{E} -12$ (F/cm) is the permittivity of free space, and ε_r is the relative dielectric constant of the substrate. A "fringe factor," K, is included to account for necessary fringing (estimated to be 2–15%). These equations are meant to provide insight into circuit behavior and should, therefore, be applied cautiously to specific applications.

UHF VCO Using the Tapped-Inductor Differential Oscillator at 900 MHz. The basic electrical circuit of this configuration is shown in Figure 5-41. Feedback occurs when sufficient RF voltage develops across the tap inductances L_{T1} and L_{T2} (due to tap 1 and tap 2, respectively); note that inductance is reckoned from the cold end of the tank at node A. Taps 1 and 2 should be as close as possible to pins 9 and 10, respectively. Also node A (cold end of tank) and node B (pin 6) must be as physically close together and exhibit as low an impedance as possible to discourage parasitic oscillations. This "inner loop," composed of L_{T2}, the inductance of connections between taps 1 and 2 and pins 9 and 10, respectively, and "stray" low-Q inductance between nodes A and B creates a parasitic loop that will support oscillation that no longer depends on the full tapped-L tank circuit. Oscillation based on this low-Q loop typically occurs well above 1.2 GHz and has been observed as high as 1.6 GHz, and exhibits very poor phase noise performance.

Another factor also affects this unwanted parasitic oscillation. As the full tank circuit's unloaded Q decreases, the circuit becomes conditionally stable and will eventually favor the parasitic loop exclusively. This occurs because the loaded Q of the tapped microstrip affects the magnitude of the RF voltage appearing across the entire tank. Thus, feedback at taps 1 and 2 is reduced as the Q decreases. This increases the likelihood of the parasitic inner loop controlling the oscillator, since its feedback voltage is largely independent of the microstrip tank Q.

Equation (5-80) shows that microstrip inductors are capable of very high unloaded Q values. This depends on a number of physical factors: frequency, dielectric thickness and

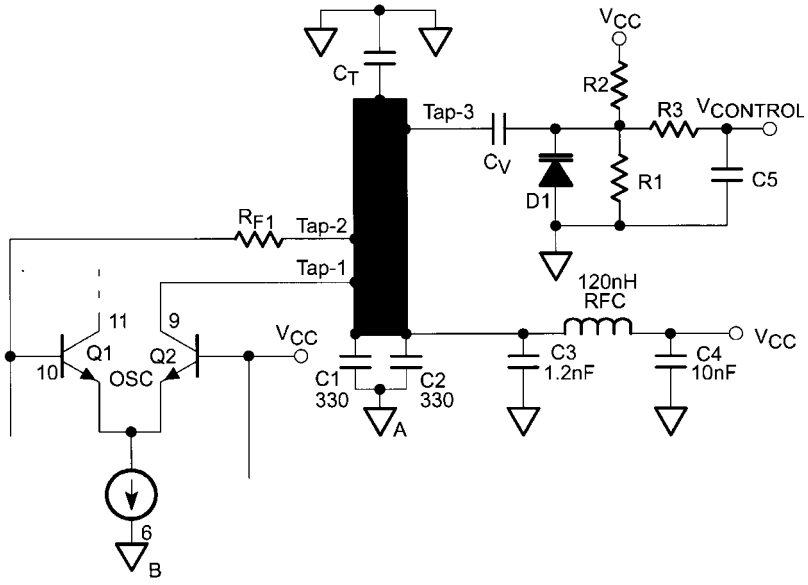


Figure 5-41 Differential oscillator with tapped-microstrip resonator.

quality, and the skin-depth conductance of the metal strip itself. For example, at 900 MHz, the unloaded Q of a silver-coated microstrip line is in the vicinity of 750. Coating the same line with lead (sloppy soldering can do it!) reduces the Q to less than 230. If the unloaded Q is too low, the loaded Q also drops. The circuit becomes conditionally stable and now may oscillate either at the higher parasitic-loop frequency or the desired lower frequency at which the entire microstrip tank resonates. Components should be connected by narrow traces closely connected to component leads to avoid soldering on the microstrip proper.

R_{F1} introduces necessary losses to control the inner-loop parasitic oscillation. Its size should be made as small as possible consistent with stability, startup, and good phase-noise performance. Since R_{F1} also decreases feedback, the oscillator's output falls as R_{F1} increases. Typically, experimental values from 4 to 30 Ω have proved sufficient.

Stability as a function of V_{CC} is a good way to assess conditional stability. Provided the microstrip tank has a high loaded Q, the oscillator's stability should be independent of V_{CC} down to less than 2.5 V or so. When loaded Q decreases sufficiently to favor the inner-loop parasitic, conditional stability will occur. This can readily be observed by increasing V_{CC} from about 1.0 V and noting whether the VCO frequency jumps back and forth unpredictably. Mixer and LO-port loading can also affect this condition and should be considered.

Dimensions for the microstrip resonator are based on several criteria. The strip must be "wide": Its width-to-height ratio must be on the order of 1 or greater. Minimum strip length seems to be about 200 mils for 62-mil-thick board. L/C ratios somewhere about 500 have yielded quite good results. Note that we usually specify the "L/C ratio" because it is always greater than 1, even though in a parallel-resonant circuit it appears in the Q equation as C/L. Thus, decreasing the "L/C ratio" by making the microstrip shorter helps the inductor's loaded Q, and the circuit's loaded Q_L, increase [see Eq. (5-75)]. However, the effect of the lower Q of C_T, even when using a high-Q surface-mount type, decreased the expected larger increase in the circuit's net loaded Q. Thus, the expected increase in Q_L may not be fully realizable. Assuming a 62-mil-thick board with FR5 dielectric, a strip 50 mils by 300 mils gave very good experimental results, where C_T was about 4.3 pF, with oscillation occurring from about 950 to 1000 MHz with various test boards. A shorter inductor designed to increase the C/L ratio requiring 7.5 pF experimentally resulted in only about 1–2-dB phase-noise improvement at 950 MHz. The reason for this is that as the microstrip inductor degenerates to a plain inductor, all of its advantages are lost.

Tap 1 may be anywhere between the cold end of the tank (where C_1 and C_2 are located) and tap 2. Tap 2 yields good results at about one-third the strip length. Making it too close to the cold end in an effort to increase Q_L will result in loss of control over the inner-loop parasitic. To keep Q_L as high as possible, C_T should be a high-Q surface-mount type. Differences greater than 5 dB in SSB phase noise have been observed between a generic NP0 surface-mount capacitor and a high-Q NP0 surface-mount capacitor.

Tap 3 can be at the end of the microstrip when C_T is connected. However, the tuning diode inevitably will cause a decrease in overall loaded Q, typically resulting in a phase-noise increase of as much as 5 dB. Moving it some distance down from the high end of the tank will decrease this effect and yield better results. A good starting point is about one-quarter down, or three-quarters of the length from the cold end of the tank. C_1 and C_2 are paralleled lower-value capacitors to yield a better low-impedance ground return to pin 6 (node B) since relatively large RF currents flow through them. Note that as Q_L increases, the peak circulating RF current will also increase, as will the RF voltage at the hot end of the microstrip. At 900 MHz, good results have been obtained when both are about 300 pF. The RF choke was chosen to be approximately series-resonant around 900 MHz, and constitutes the series feed path

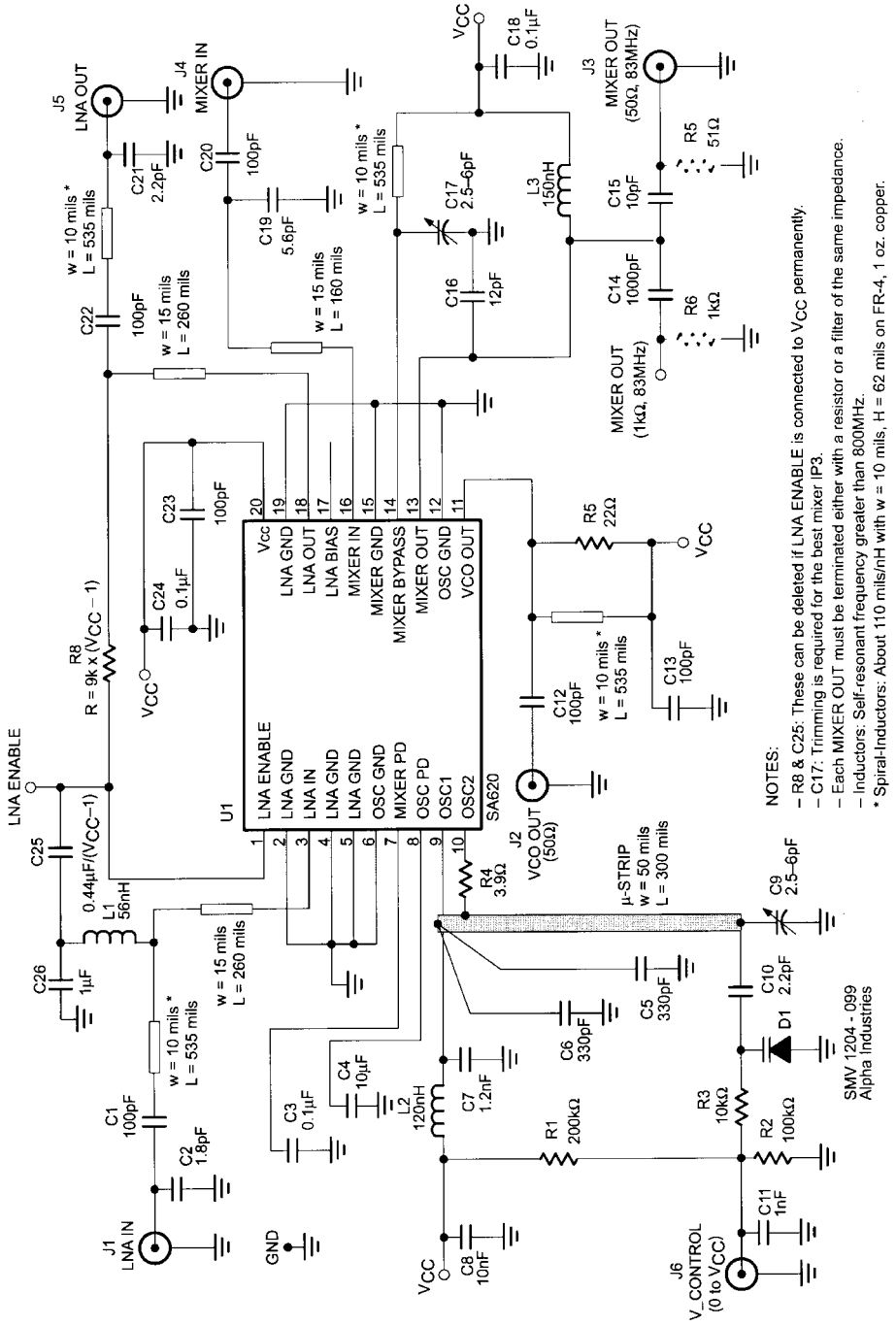


Figure 5-42 Philips SA620 application board schematic showing the tapped-microstrip resonator.

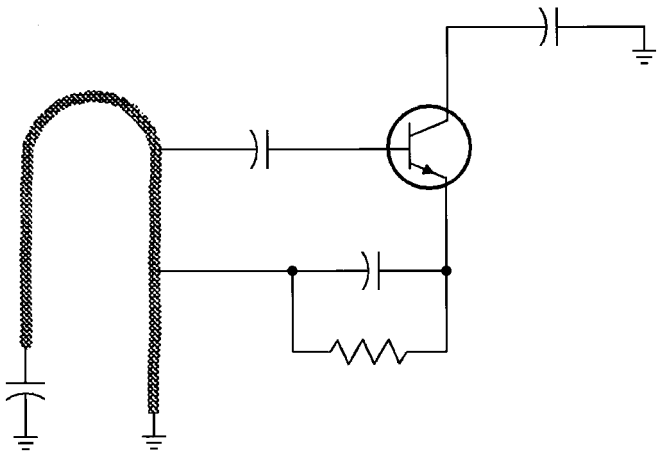


Figure 5-43 "Walking stick" microstrip Hartley resonator as used by Plextek in the 800–1500-MHz range. Below 800 MHz, the resonator's physical size becomes impractical; above 1500 MHz, the transistor's parasitic reactances make the Hartley configuration problematic.

for dc biasing. It may be possible to neglect this component entirely, provided that the PC-board connection to the cold end of the tank is very close to RF ground. Finally, note that shielding of the entire microstrip may be necessary to meet FCC Part 15 emission limitations (where applicable). Figure 5-42 shows one application of the tapped-microstrip oscillator with the Philips SA620 front-end IC.

5-5-5 Hartley Microstrip Resonator Oscillator

G. Smithson and S. Fitz of Plextek Communications Technology Consultants have reported on the construction of Hartley microstrip-resonator oscillators using thick-film techniques on ceramic (alumina) substrates for high-volume wireless products. The resonator in one such oscillator (Figure 5-43) exhibited an unloaded Q of approximately 80. Table 5-2 summarizes its phase noise versus tuning range with different tuning-diode configurations.

5-5-6 Crystal Oscillators

Most wireless applications do not require ultra-high-frequency stability that mandates the use of a primary standard, such as one based on cesium, or a secondary standard, such as one based on rubidium. For very stable oscillators at fixed or only slightly variable frequencies up to a few hundred megahertz, oscillators with piezoelectric resonators are generally used. A piezoelectric material transduces mechanical stress (deformation) to

Table 5-2 Achieved oscillator phase noise versus tuning range

Tuning Range, 0.5–3 V	Tuning Range, 1–3 V	Phase Noise at 20-kHz Offset
23%	16%	–100 dBc/Hz
16%	11%	–103 dBc/Hz
10%	8%	–106 dBc/Hz

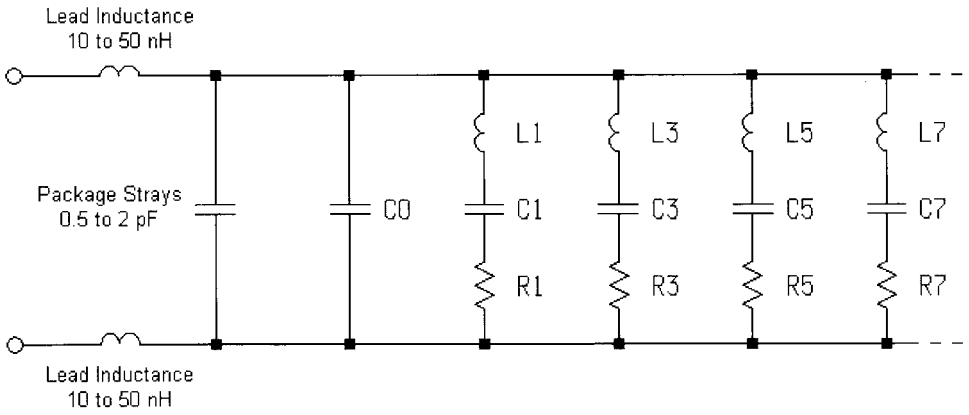


Figure 5-44 Electrical equivalent of a piezoelectric crystal resonator, showing fundamental (1) and overtone (3, 5, 7, . . .) resonances. (Courtesy of Roger Clark, Vectron International.)

Table 5-3 Approximate crystal parameters

Crystal Q Approximations ($F_{\text{resonance}}$ in MHz)

Nonprecision Fundamental Crystals:

Worst: $Q = \dfrac{2 \times 10^6}{F}$ Better: $Q = \dfrac{5 \times 10^6}{F}$

Precision Crystals Processed for High Q:

SC: $Q = \dfrac{12 \times 10^6}{F}$ AT Cut: $Q = \dfrac{10 \times 10^6}{F}$

Crystal Resistance (r) for 10-MHz, Third-Overtone Example Using Precision Crystal Approximation (F in MHz and C_1 in pF)

$$r = \frac{2\pi F C_1}{Q}$$

$$r = 6.964$$

Typical Crystal Parameters

10-MHz Fundamental: $C_1 = 0.02$ pF, $C_0 = 3.6$ pF, $Q = 350{,}000$, $r = 2.5\ \Omega$, $L = 0.012665$ H
50-MHz Third-Overtone: $C_1 = 0.0026$ pF, $C_0 = 4.212$ pF, $Q = 100{,}000$, $r = 13\ \Omega$, $L = 0.003897$ H
100-MHz Fifth-Overtone: $C_1 = 0.0005$ pF, $C_0 = 2.25$ pF, $Q = 80{,}000$, $r = 40\ \Omega$, $L = 0.005066$ H
155-MHz High-Frequency Fundamental: $C_1 = 0.005$ pF, $C_0 = 1.5$ pF, $Q = 20{,}000$, $r = 12\ \Omega$, $L = 0.002108$ H

Source: Roger Clark, Vectron International.

electrical stress (voltage), and vice versa. In noncritical applications up to 10 MHz, a piezoelectric ceramic may be used. The Q and temperature stability of such materials are relatively poor, however, so quartz is the resonator material of choice for temperature-stable, low-noise piezoelectric oscillators.

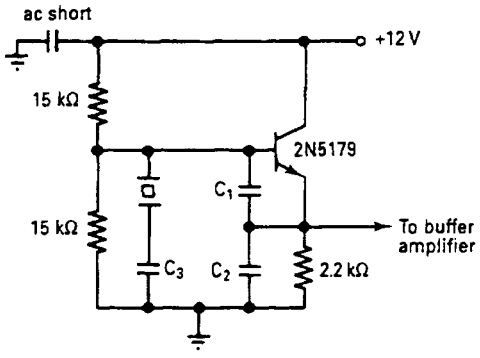
Just as an electrical resonator consisting of distributed L and C exhibits multiple electrical resonances, a physical piece of piezoelectric material, depending on its shape, exhibits multiple mechanical resonances. *Fundamental* resonance occurs at the frequency at which the crystal is one-half wavelength long; *overtone* resonances occur near odd multiples (3, 5, 7, etc.) of the fundamental. Figure 5-44 shows the electrical equivalent of a crystal. Table 5-3 presents examples of approximate parameters for several crystal types. Appropriate shaping and placement of the crystal's connecting electrodes, combined with suitable oscillator circuit design and adjustment, allow the selection or ehancement of a particular resonance for excitation.

Figure 5-45 shows a typical crystal oscillator circuit configuration, and the associated table supplies recommended values for the elements. This is in line with the chip manufacturer's recommendations for built-in oscillator circuits. Figure 5-46 shows how this oscillator's input impedance varies around resonance; Figure 5-47 shows its simulated phase noise. Placing a small resistance in series with the crystal (Figure 5-48) allows us to use the crystal as a high-Q filter, substantially improving the oscillator's spectral purity relative to that obtainable at the transistor emitter (Figure 5-49).

Some applications demand ultra-low-phase-noise crystal oscillators. Figures 5-50 and 5-51 show an appropriate circuit and its phase noise with and without the current-supplying hot carrier diode (HCD) used for noise reduction.

5-5-7 Voltage-Controlled Oscillators

A schematic of a ceramic-resonator-based VCO (Figure 5-35) has already been shown in the section on ceramic-resonator oscillators. To tune the oscillator within the required range,



	3 MHz	6 MHz	10 MHz	20 MHz	30 MHz
C_1 (pF)	330	270	180	82	43
C_2 (pF)	430	360	220	120	68
C_3 (pF)	39	43	43	36	32
Crystal C_L (pF)	32	32	30	20	15

C_3 is the approximate value needed to set oscillator on frequency using indicated crystal C_L.

Figure 5-45 A 3–30-MHz Colpitts crystal oscillator.

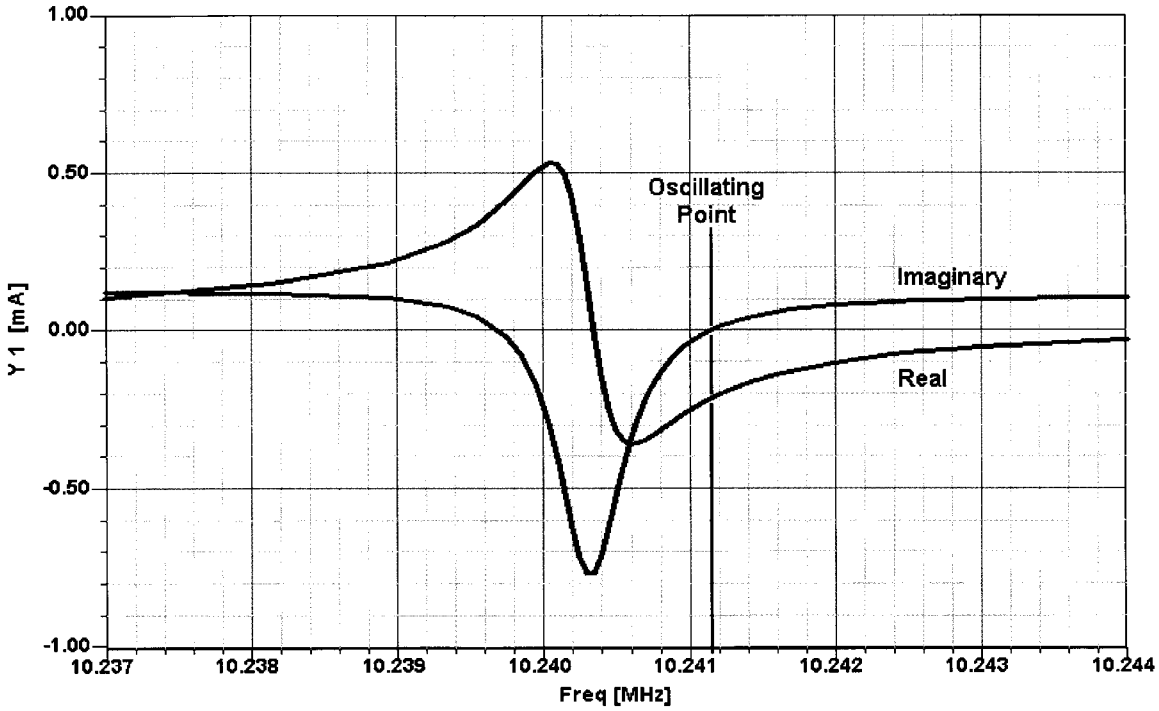


Figure 5-46 As with every oscillator, the crystal oscillator's input impedance is negative and real at resonance. This graph plots the real and imaginary components of the test voltage injected by the Oscillator Design Aid tool in Ansoft Serenade 8.0. Compare this graph with Figure 5-7, which plots an oscillator's input impedance with the resonator absent.

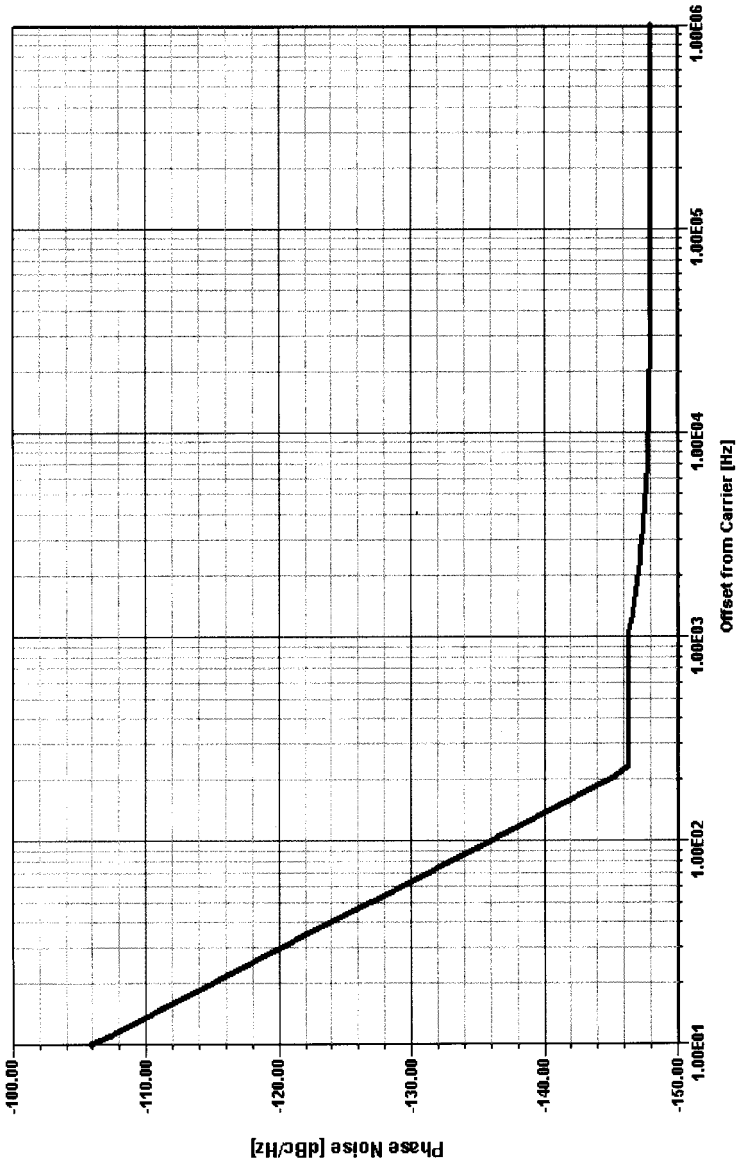


Figure 5-47 Simulated phase noise of the crystal oscillator.

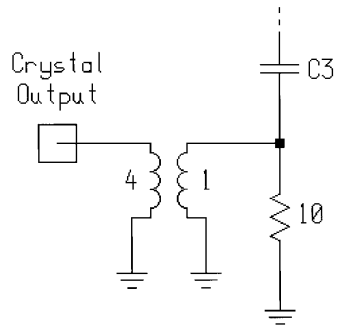


Figure 5-48 This alternative output tap uses the crystal's high selectivity for improved harmonic reduction.

tuning diodes are used. These diodes are often called *varactors* or *voltage-sensitive diodes.* By way of approximation, we can use the equation

$$C = \frac{K}{(V_R + V_D)^n} \tag{5-82}$$

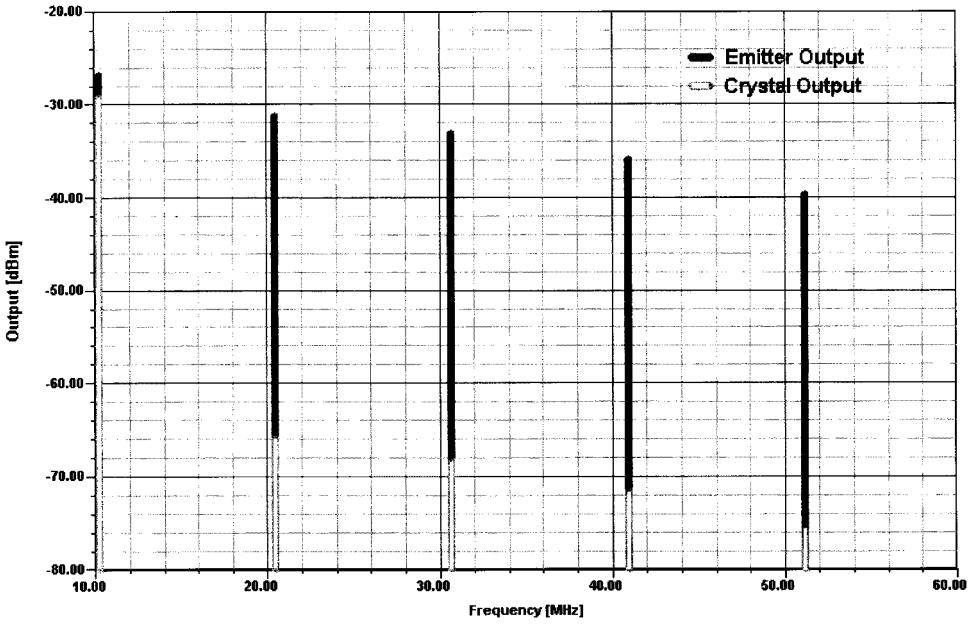


Figure 5-49 Comparison of the crystal oscillator's simulated output spectrum as obtained at the transistor emitter (dark gray, fundamental –26.7 dBm) and crystal (light gray, fundamental –29.1 dBm). At the emitter, the second harmonic is down only 4.3 dB relative to the carrier; at the crystal, the second harmonic is down 36 dB relative to the carrier.

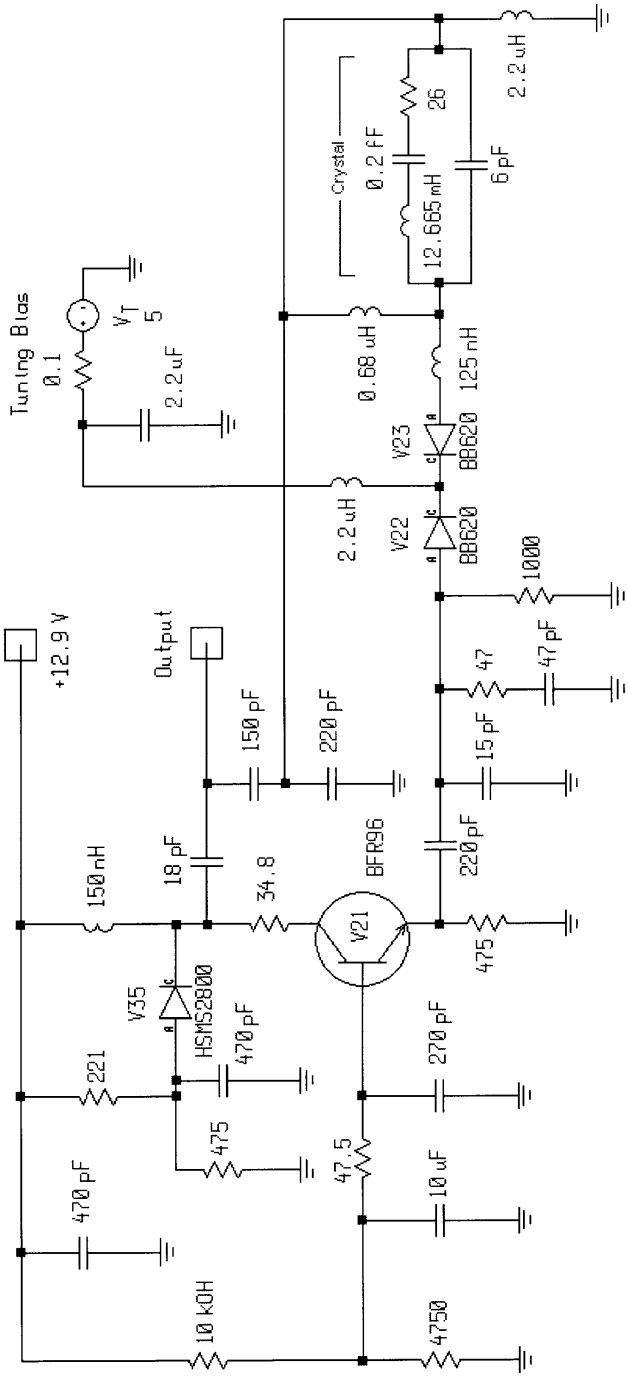


Figure 5-50 A 100-MHz VCXO suitable for ultra-low-phase-noise applications. The HSMS2800 hot-carrier diode (HCD) provides noise reduction.

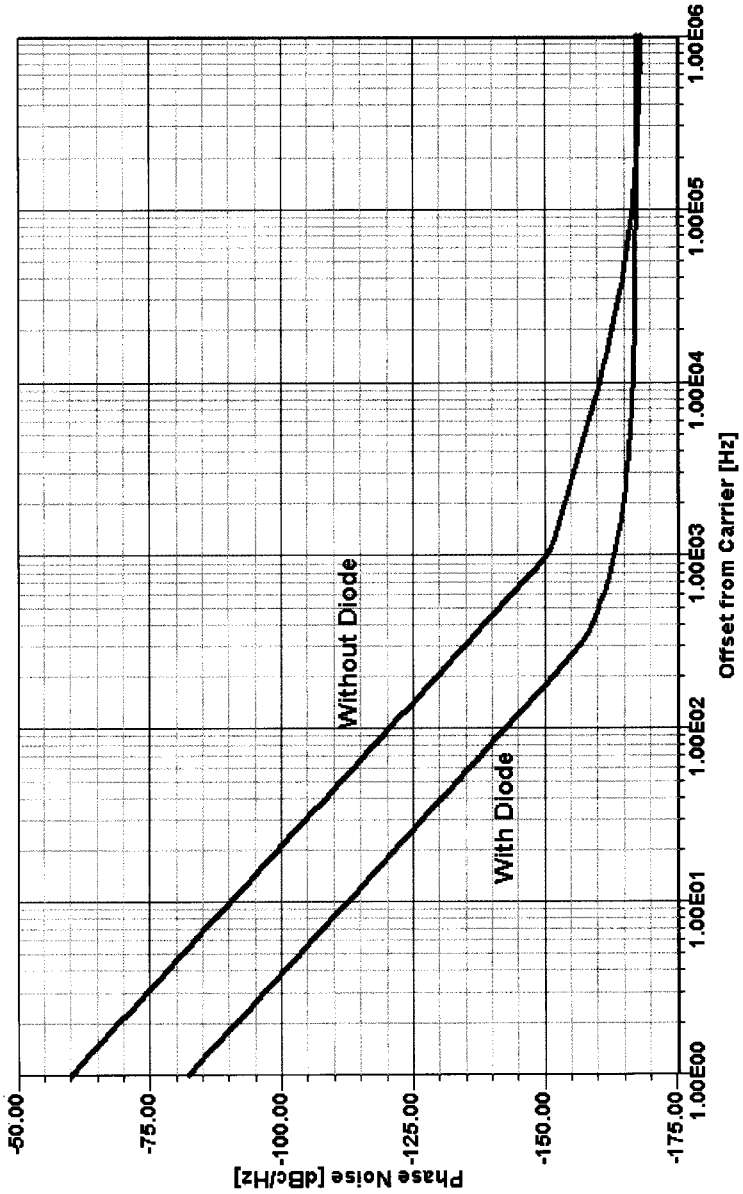


Figure 5-51 Phase noise of the VCXO with and without the noise-reduction diode.

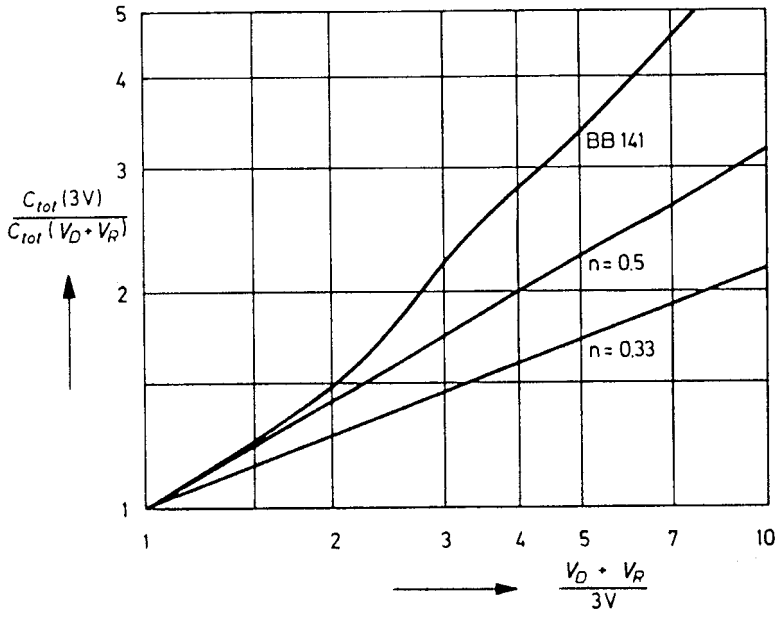


Figure 5-52 Capacitance–voltage characteristic for an alloyed capacitance diode ($n = 0.33$), a diffused capacitance diode ($n = 0.5$), and a wide-range tuning diode (BB141).

wherein all constants and all parameters determined by the manufacturing process are contained in K. The exponent is a measure of the slope of the capacitance/voltage characteristic and is 0.5 for alloyed diodes, 0.33 for single-diffused diodes, and (on average) 0.75 for tuner diodes with a hyperabrupt *pn* junction [6, 7]. Figure 5-52 shows the capacitance–voltage characteristics of an alloyed, a diffused, and a tuner diode.

Recently, an equation was developed that, although purely formal, describes the practical characteristic better than Eq. (5-82):

$$C = C_0 \left(\frac{A}{A + V_R} \right)^m \tag{5-83}$$

where C_0 is the capacitance at $V_R = 0$, and A is a constant whose dimension is the volt. The exponent m is much less dependent on voltage than the exponent n in Eq. (5-81).

The operating range of a capacitance diode or its useful capacitance ratio,

$$\frac{C_{\max}}{C_{\min}} = \frac{C_{\text{tot}}(V_{R,\min})}{C_{\text{tot}}(V_{R,\max})} \tag{5-84}$$

is limited by the fact that the diode must not be driven by the alternating voltage superimposed on the tuning voltage, either into the forward mode or the breakdown mode. Otherwise, rectification would take place, shifting the bias of the diode and considerably affecting its figure of merit.

There are several manufacturers of tuning diodes. Motorola is a typical supplier in this country; Siemens of Germany and Philips also provide good diodes. Table 5-4 and Figure 5-53 contain information for tuning diodes useful for our applications.

5-5-8 Diode-Tuned Resonant Circuits

Tuner Diode in Parallel-Resonant Circuit. Figures 5-54 and 5-55 illustrate two basic circuits for the tuning of parallel resonant circuits by means of capacitance diodes. In the circuit diagram of Figure 5-54, the tuning voltage is applied to the tuning diode via the bias resistor, R_B. Series-connected from the tuning diode to the top of the tank circuit is capacitor C_S, which completes the circuit for alternating current but isolates the cathode of the tuner diode from the coil and thus from the negative terminal of the tuning voltage. Moreover, a fixed parallel capacitance C_P is provided. The decoupling resistor preceding the bias resistor is large enough to be disregarded in the following discussion. Since for high-frequency purposes the biasing resistor is connected in parallel with the series capacitor, it is transformed into the circuit as an additional equivalent shunt resistance R_c. For the parallel loss resistance transformed into the circuit, we have the expression

$$R_c = R_B \left(1 + \frac{C_{\text{tot}}}{C_S}\right)^2 \tag{5-85}$$

Table 5-4 Example tuning diodes (Siemens)

	Maximum Ratings		Characteristics (T_A = 25 °C)								
Type	V_R (V)	I_F (mA)	C_T at (pF)	V_R (V)	C_T (pF)	V_R (V)	C_{ratio}	I_R (nA)	V_R (V)	Package	Chip Code
BB439	30	20	29	3	5.0	25	5.8	≤20	28	SOD-323	Q
BB535	30	20	18.7	1	2.1	28	8.9	≤10	30	SOD-323	A
BB545	30	20	20	1	2.0	28	10.0	≤10	30	SOD-323	C
BB639	30	20	38.3	1	2.65	28	14.5	≤10	30	SOD-323	K
BB639C	30	20	39	1	2.55	28	15.3	≤10	30	SOD-323	B
BB640	30	20	69	1	3.05	28	22.6	≤10	30	SOD-323	N
BB804 (dual)	20	50	44.75	2	26.15	8	1.71	≤20	16	SOT-23	E
BB814 (dual)	20	50	44.75	2	20.8	8	2.15	≤20	16	SOT-23	O
BB833	30	20	9.3	1	0.75	28	12.4	≤20	30	SOD-323	F
BB835	30	20	9.1	1	0.62	28	14.7	≤10	30	SOD-323	G
BB914 (dual)	20	50	43.75	2	18.7	8	2.34	≤20	16	SOT-23	P
BBY51 (dual)	7	20	5.30	1	3.10	4	1.75	≤10	6	SOT-23	H
BBY51-03W	7	20	5.30	1	3.10	4	1.75	≤10	6	SOD-323	H
BBY51-07 (dual)	7	20	5.30	1	3.10	4	1.75	≤10	6	SOT-143	H
BBY52 (dual)	7	20	1.75	1	1.25	4	1.40	≤10	6	SOT-23	I
BBY52-03W	7	20	1.75	1	1.25	4	1.40	≤10	6	SOD-323	I
BBY53 (dual)	6	20	5.30	1	2.40	3	2.20	≤10	4	SOT-23	L
BBY53-03W	6	20	5.30	1	2.40	3	2.20	≤10	4	SOD-323	L

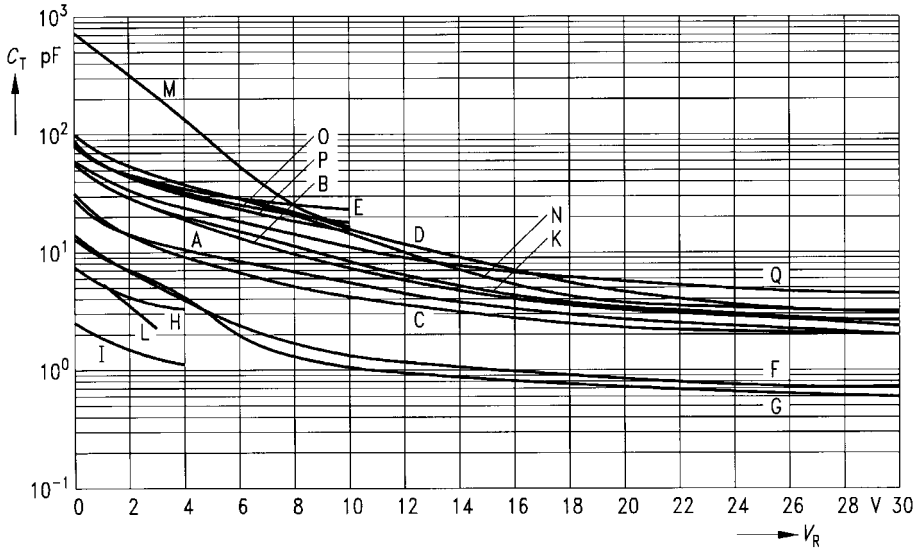


Figure 5-53 Diode capacitance C_T versus reverse voltage V_R for the diodes shown in Table 5-4. The curve labels correspond to the chip codes called out in the table.

If in this equation the diode capacitance is replaced by the resonator circuit frequency ω, we obtain

$$R_c = R_B \left(\frac{\omega^2 L C_S}{\omega^2 L (C_S + C_P) - 1} \right)^2 \tag{5-86}$$

In the circuit of Figure 5-55, the resonant circuit is tuned by two tuning diodes, which are connected in parallel via the coil for tuning purposes, but series-connected in opposition for high-frequency signals. This arrangement has the advantage that the capacitance shift caused by the ac modulation acts in opposite directions in these diodes and, therefore, cancels itself. The bias resistor R_B, which applies the tuning voltage to the tuning diodes, is transformed into the circuit at a constant ratio across the entire tuning range. Given two identical, loss-free tuning diodes, we obtain the expression

$$R_c = 4R_B \tag{5-87}$$

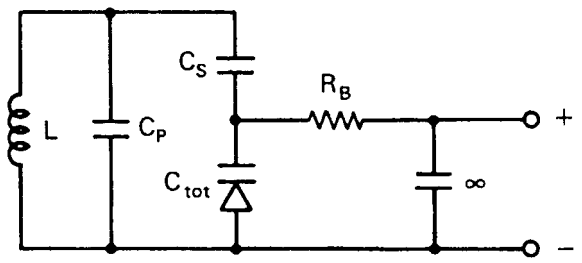


Figure 5-54 Parallel-resonant circuit with tuner diode, and bias resistor in parallel to the diode.

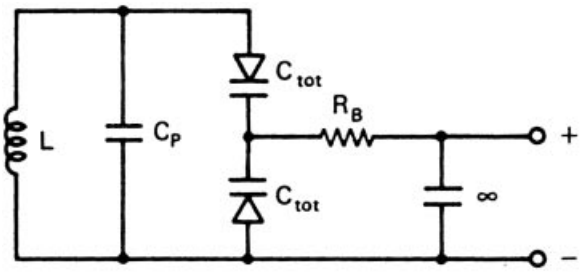


Figure 5-55 Parallel-resonant circuit with two tuning diodes.

Capacitances Connected in Parallel or in Series with the Tuning Diode. Figure 5-54 shows that a capacitor is usually in series with the tuner diode, in order to close the circuit for alternating current and, at the same time, to isolate one terminal of the tuning diode from the rest of the circuit with respect to direct current, so as to enable the tuning voltage to be applied to the diode. If possible, the value of the series resistor C_S will be chosen such that the effective capacitance variation is not restricted. However, in some cases—for example, in the oscillator circuit of receivers whose intermediate frequency is of the order of magnitude of the reception frequency—this is not possible, and the influence of the series capacitance will then have to be taken into account. By connecting the capacitor C_S, assumed to be lossless, in series with the diode capacitance C_{tot}, the tuning capacitance is reduced to the value

$$C^* = C_{tot} \frac{1}{1 + C_{tot}/C_S} \tag{5-88}$$

The Q of the effective tuning capacitance, taking into account the Q of the tuning diode, increases to

$$Q^* = Q\left(1 + \frac{C_{tot}}{C_S}\right) \tag{5-89}$$

The useful capacitance ratio is reduced to the value

$$\frac{C^*_{max}}{C^*_{min}} = \frac{C_{max}}{C_{min}} \frac{1 + C_{min}/C_S}{1 + C_{max}/C_S} \tag{5-90}$$

where C_{max} and C_{min} are the maximum and minimum capacitances, respectively, of the tuner diode.

On the other hand, the advantage is gained that, due to the capacitive potential division, the amplitude of the alternating voltage applied to the tuner diode is reduced to

$$\hat{v}^* = \hat{v} \frac{1}{1 + C_{tot}/C} \tag{5-91}$$

so that the lower value of the tuning voltage can be smaller, and this results in a higher maximum capacitance C_{max} of the tuning diode and a higher useful capacitance ratio. The

influence exerted by the series capacitor, then, can actually be kept lower than Eq. (5-89) would suggest.

The parallel capacitance C_P that appears in Figures 5-54 and 5-55 is always present, since wiring capacitances are inevitable and every coil has its self-capacitance. By treating the capacitance C_P, assumed to be lossless, as a shunt capacitance, the total tuning capacitance rises in value and, if C_S is assumed to be large enough to be disregarded, we obtain

$$C^* = C_{\text{tot}}\left(1 + \frac{C_P}{C_{\text{tot}}}\right) \tag{5-92}$$

The Q of the effective tuning capacitance, derived from the Q of the tuning diode, is

$$Q^* = Q\left(1 + \frac{C_P}{C_{\text{tot}}}\right) \tag{5-93}$$

or, in other words, it rises with the magnitude of the parallel capacitance. The useful capacitance is reduced:

$$\frac{C^*_{\text{max}}}{C^*_{\text{min}}} = \frac{C_{\text{max}}}{C_{\text{min}}}\,\frac{1 + C_P/C_{\text{max}}}{1 + C_P/C_{\text{min}}} \tag{5-94}$$

In view of the fact that even a comparatively small shunt capacitance reduces the capacitance ratio considerably, it is necessary to ensure low wiring and coil capacitances in the layout stage of the circuit design.

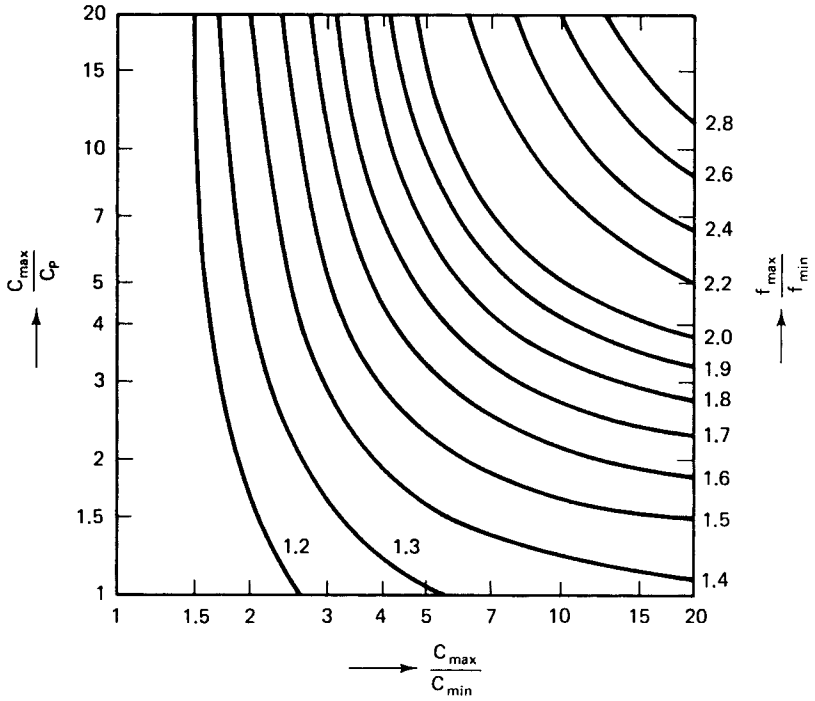


Figure 5-56 Diagram for determining capacitance ratio and maximum capacitance.

Tuning Range. The frequency range over which a parallel-resonant circuit (according to Figure 5-54) can be tuned by means of the tuning diode depends on the useful capacitance ratio of the diode and on the parallel and series capacitances present in the circuit. The ratio is

$$\frac{f_{\max}}{f_{\min}} = \left(\frac{1 + \dfrac{C_{\max}}{C_P(1 + C_{\max}/C_S)}}{1 + \dfrac{C_{\max}}{C_P(C_{\max}/C_{\min} + C_{\max}/C_S)}} \right)^{1/2} \tag{5-95}$$

In many cases, the value of the series capacitor can be chosen to be large enough for its effect to be negligible. In that case, Eq. (5-95) is simplified as follows:

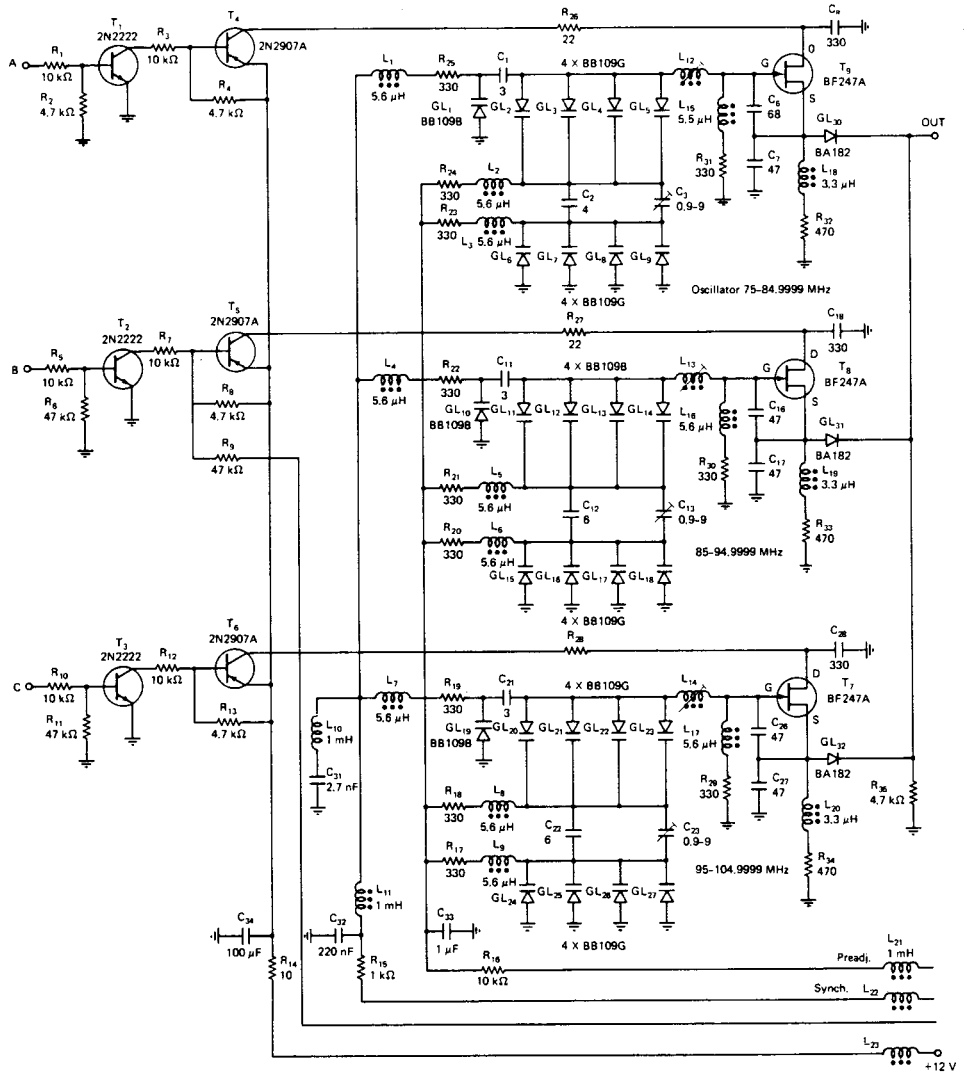


Figure 5-57 Oscillator and switching section of the Rohde & Schwarz ESH2/ESH3 test receiver.

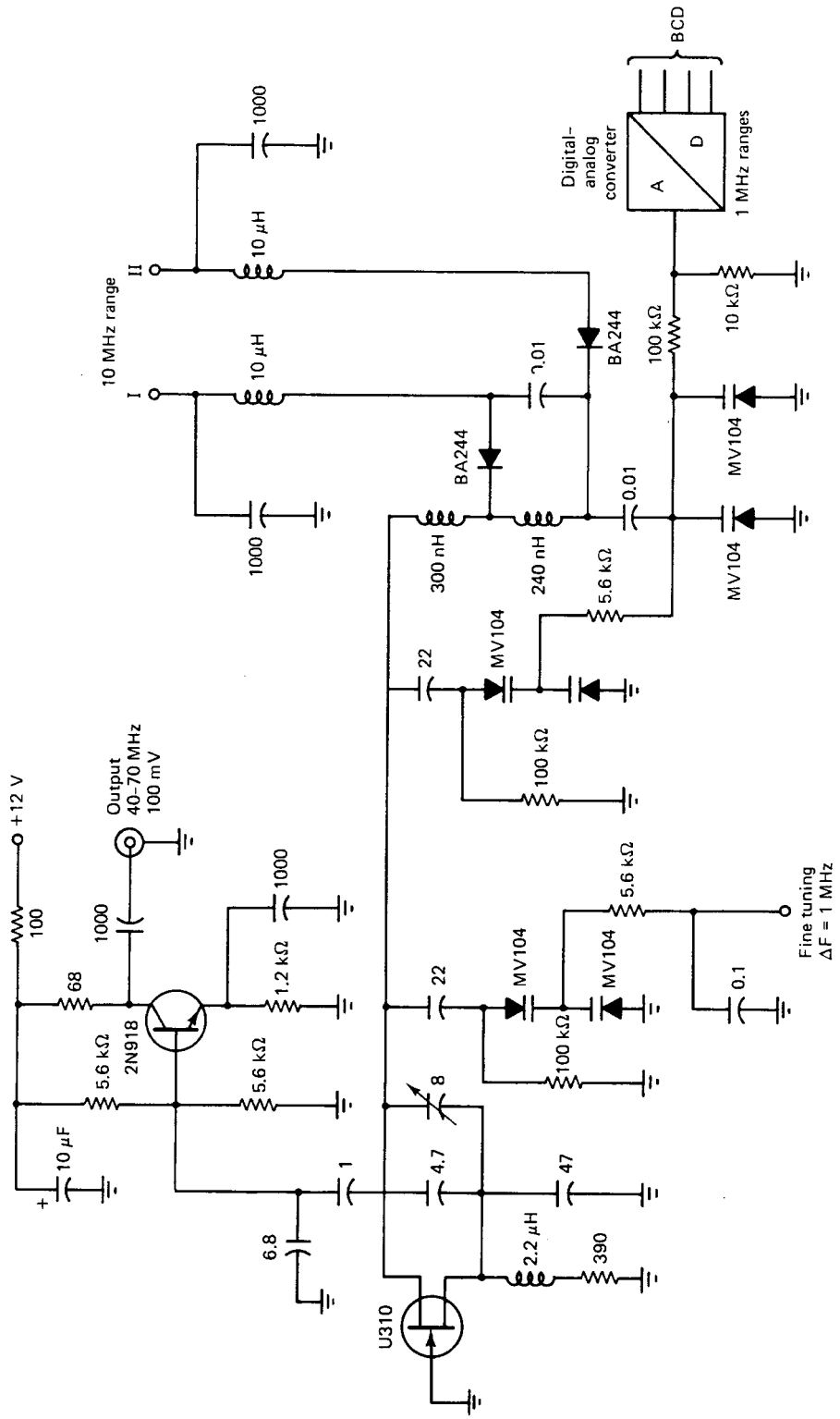


Figure 5-58 A 40–70-MHz VCO with two coarse-steering ranges and a fine-tuning range of 1 MHz.

$$\frac{f_{\max}}{f_{\min}} = \left(\frac{1 + C_{\max}/C_P}{1 + C_{\min}/C_P} \right)^{1/2} \tag{5-96}$$

From this equation, the diagram shown in Figure 5-56 is computed. With the aid of this diagram, the tuning diode parameters required for tuning a resonant circuit over a stipulated frequency range (i.e., the maximum capacitance and the capacitance ratio) can be determined. Whenever the series capacitance C_S cannot be disregarded, the effective capacitance ratio is reduced accorded to Eq. (5-90).

Tracking. When several tuned circuits are used on the same frequency, diodes must be selected for perfect tracking.

5-5-9 Practical Circuits

Oscillators with Coarse and Fine Tuning. After so much theory, it may be nice to take a look at some practical circuits, such as the one shown in Figure 5-57. This oscillator is being used in the Rohde & Schwarz ESN/ESVN40 field-strength meter, and in the HF1030 receiver produced by Cubic Communications, San Diego. This circuit combines all the various techniques shown previously. A single diode is being used for fine-tuning a narrow range of less than 1 MHz; coarse tuning is achieved with the antiseries diodes.

Several unusual properties of this circuit are apparent:

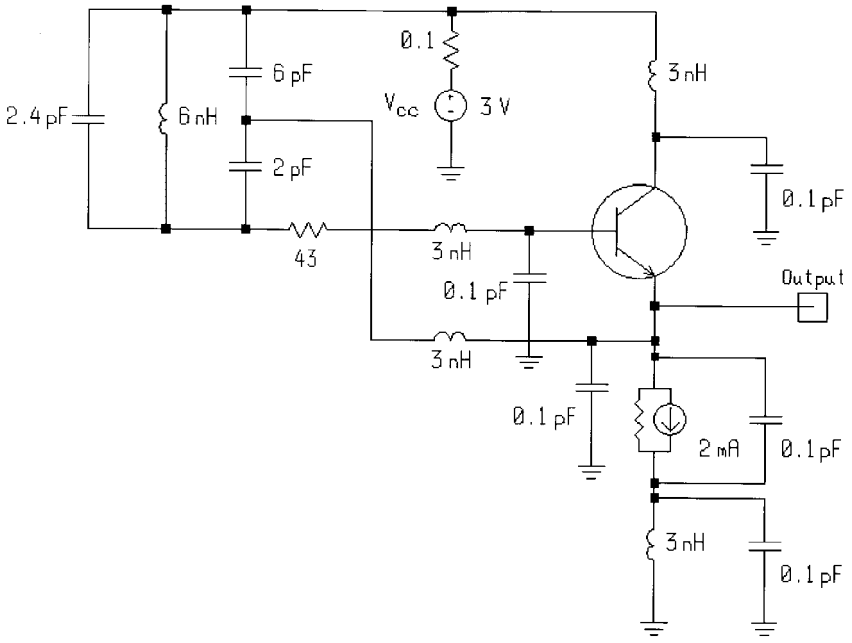


Figure 5-59 Schematic of the dc-coupled oscillator.

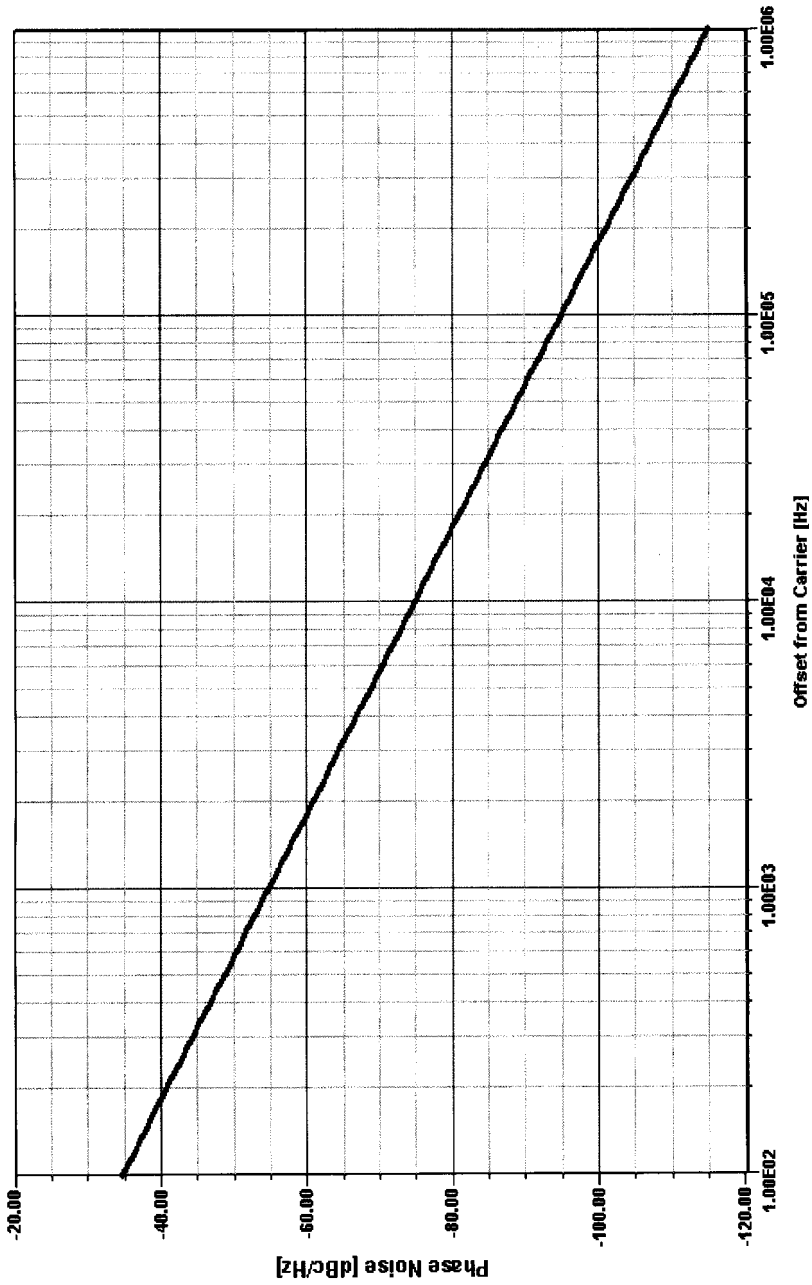


Figure 5-60 Phase noise of the dc-coupled oscillator operating at 897 MHz.

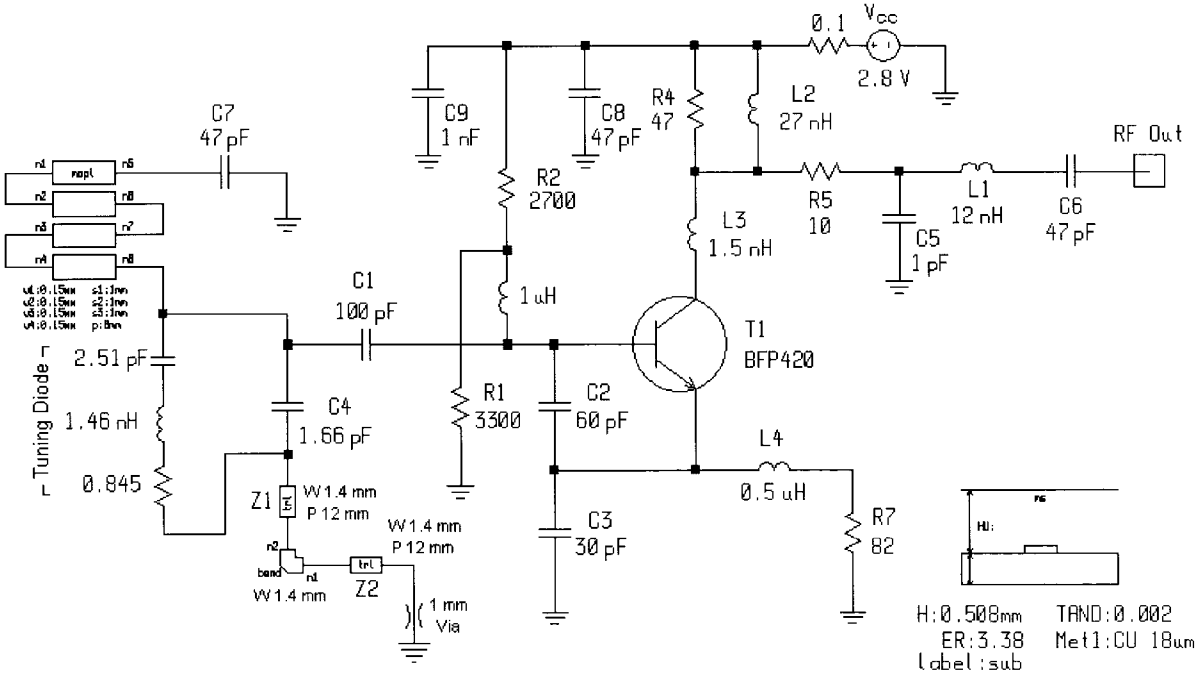


Figure 5-61 Schematic of the Siemens VCO. This oscillator (simulation) works with such high voltage-divider capacitances (C_2 and C_3) only because the circuit's large-signal transistor model, which was obtained under contract, exhibits unrealistically high gain at the frequency of oscillation. Actual capacitors with values on the order of those shown for C_2 and C_3 typically exhibit parasitic resistance and inductance of 0.3 Ω and 0.2 nH, respectively; the resulting series resonance would make any actual such capacitor act as an inductor at UHF. Obtaining practical "UHF useful" capacitances with values on the order of those shown for C_2 and C_3 would therefore necessitate the use of multiple, smaller-value capacitors in parallel—4 × 15 pF at C_2, for example.

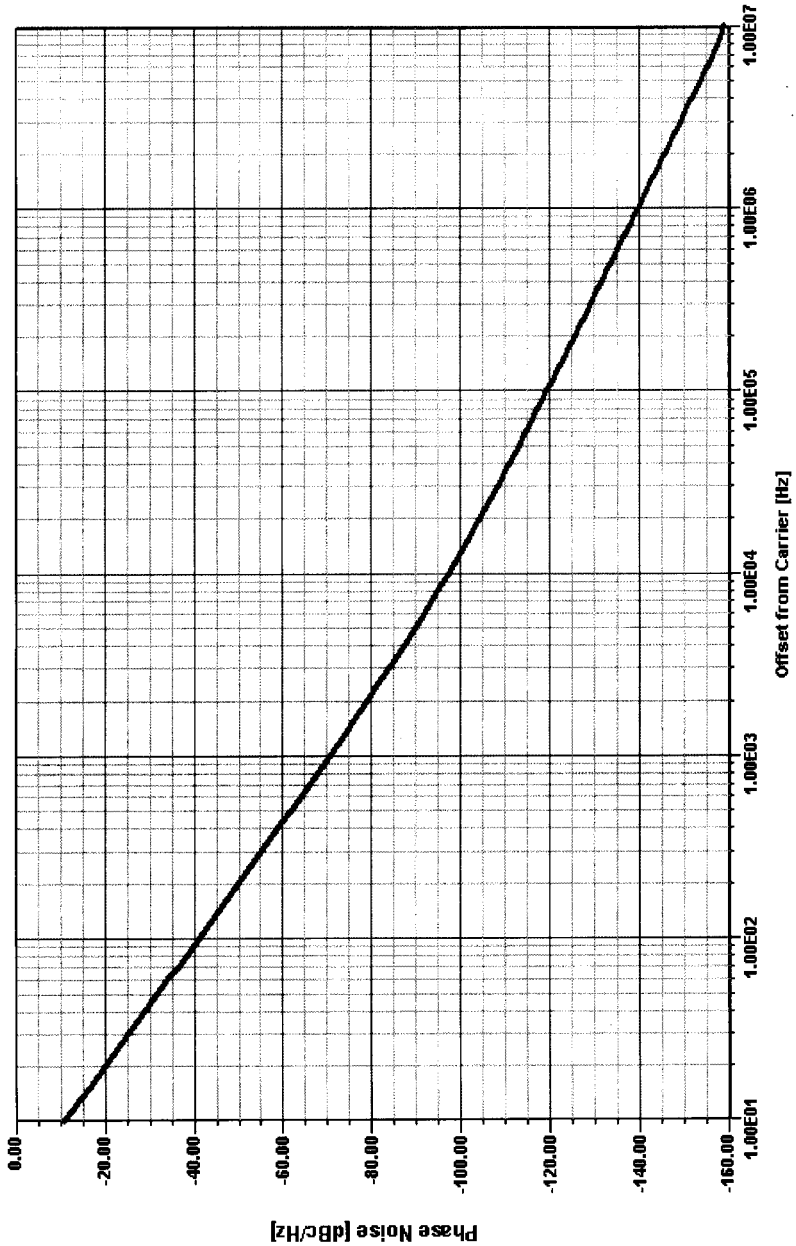


Figure 5-62 Simulated phase noise of the VCO. The frequency of oscillation is 1.052 GHz.

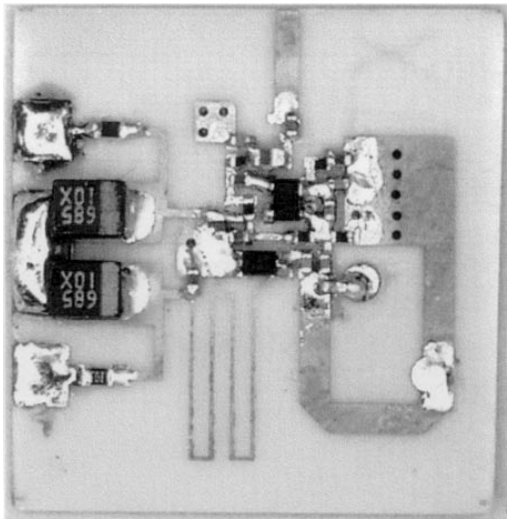


Figure 5-63 Implementation of the Siemens VCO.

1. The fine tuning is achieved with a tuning diode that has a much larger capacitance than that of the coupling capacitor in the circuit. The advantage of this technique is that the fixed capacitor and the tuning diode form a voltage divider, whereby the voltage across the tuning diode decreases as the capacitance increases. For larger values of the capacitance of the tuning diode, the Q changes and the gain K_0 increases. Because of the voltage division, the noise contribution and loading effect of the diode are reduced.
2. In the coarse-tuning circuit, several tuning diodes are used in parallel. The advantage of this circuit is a change in L/C ratio by using a higher C and storing more energy in the tuned circuit. There are no high-Q diodes available with such large capacitance values, and therefore preference is given to using several diodes in parallel rather than one tuning diode with a large capacitance, normally used only for AM (medium-frequency broadcast) tuner circuits.

We mentioned previously that, despite this, the coarse-tuning circuit will introduce noise outside the loop bandwidth, where it cannot be corrected. It is therefore preferable to incorporate switching diodes for segmenting ranges at the expense of switching current drain.

Figure 5-58 shows a circuit using a combined technique of tuning diodes for fine- and medium-resolution tuning, and coarse tuning with switching diodes.

dc-Coupled Oscillator. An interesting circuit that was cut out of an IC was supplied to us by David Lovelace of Motorola. Figure 5-59 shows the circuit and Figure 5-60 shows its phase noise. As a result of the constraints of integration, this BJT oscillator has no base bias circuitry in the conventional form. A fairly noisy performer, it ultimately was not used in production.

Siemens Colpitts Oscillator. A more elaborate approach that is essentially a standard Colpitts oscillator with a lot of detailed modeling is the oscillator of Figure 5-61, supplied

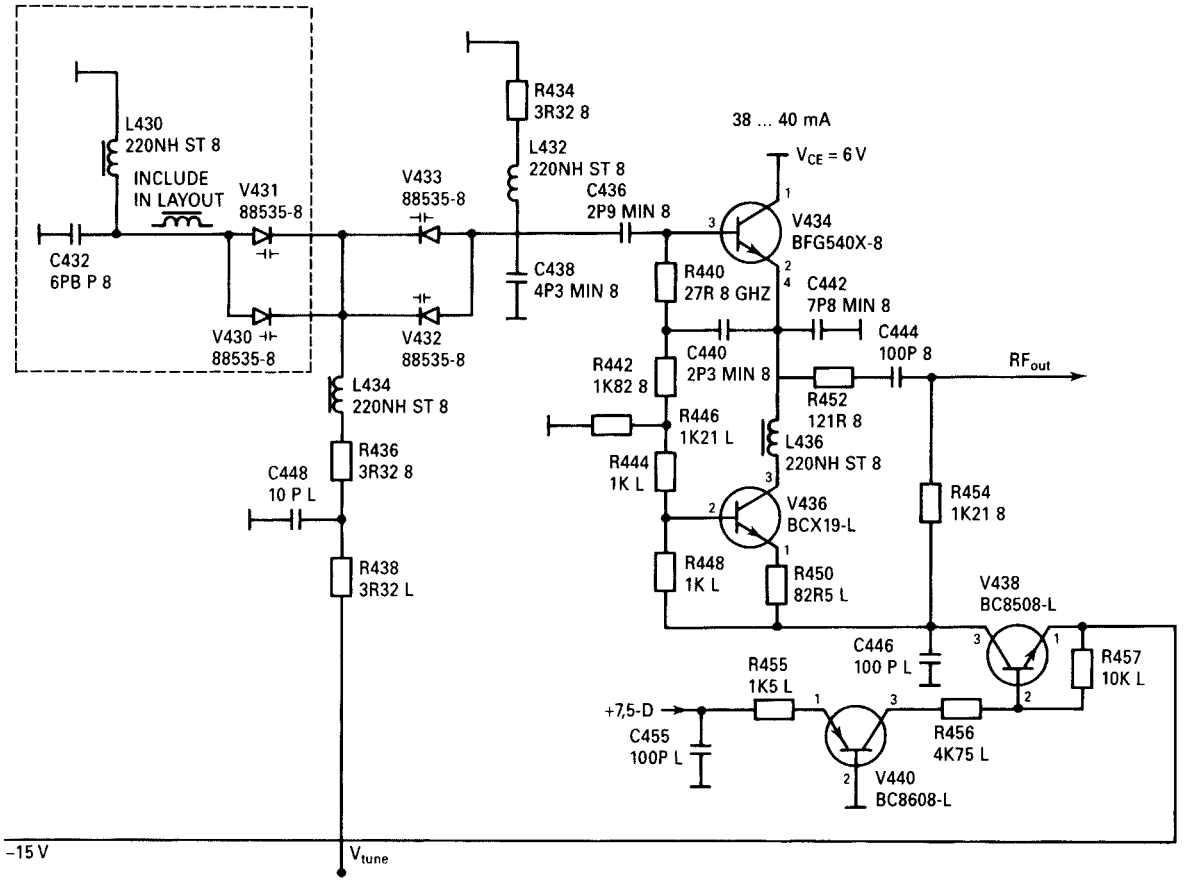


Figure 5-64 A recommended very-low-phase-noise VCO that operates in the 1-GHz region. The V436 stage is a constant-current generator responsible for improved phase-noise performance.

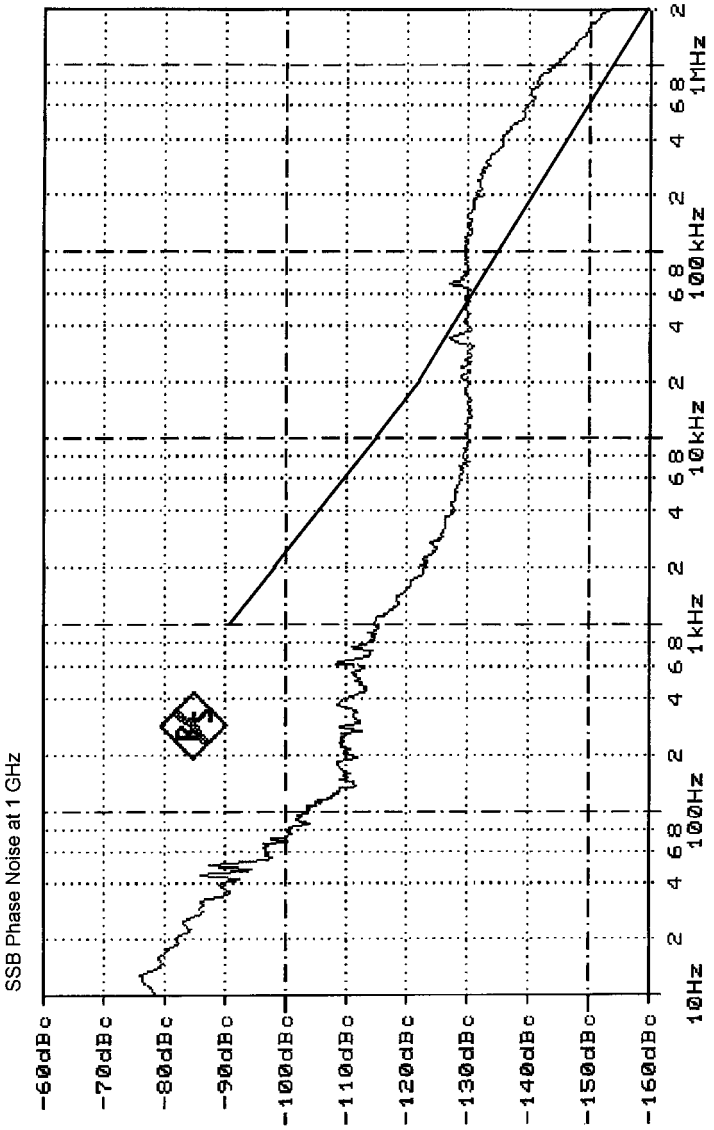


Figure 5-65 Phase noise of the low-noise VCO. The "ledges" from 100 Hz to 1 kHz and from 10 kHz to 100 kHz reflect the noise-cleanup effects of a dual-loop synthesizer; the superimposed line from 1 kHz to 2 MHz shows the oscillator's noise characteristic with both loops unlocked. The closed-loop response is noisier than the open-loop response at offsets above 60 kHz because a wide loop bandwidth is used to achieve a faster switching speed.

to us by Siemens. Considering its simplicity, its noise performance is good, as shown in Figure 5-62. Modeling this circuit involved the consideration of via holes and coupled lines. Figure 5-63 shows an actual implementation of this oscillator. The value of external emitter resistance (R7 in Figure 5-61) plays a critical role in determining a common-emitter Colpitts oscillator's upper frequency limit as described in Yeom [8].

dc-Stabilized Oscillator. For high-performance synthesizers, such as the Rohde & Schwarz series SMG/SMH, a dc-stabilized oscillator with fairly wide tuning range may be used. Figure 5-64 shows such an oscillator, the measured phase noise of which is shown in Figure 5-65. It should be noted that the output loop is magnetically coupled and therefore takes advantage of the inherent selectivity of the tuned circuit.

5-6 NOISE IN OSCILLATORS

In transmitters, oscillator noise can result in adjacent-channel interference and modulation errors; in receivers, oscillator noise can result in demodulation errors and degraded sensitivity and dynamic range. The specification, calculation, and reduction of oscillator noise are therefore of great importance in wireless system design.

5-6-1 Linear Approach to the Calculation of Oscillator Phase Noise

Since an oscillator can be viewed as an amplifier with feedback (Figure 5-66), it is helpful to examine the phase noise added to an amplifier that has a noise figure F. F is defined by [9]

$$F = \frac{(S/N)_{\text{in}}}{(S/N)_{\text{out}}} = \frac{N_{\text{out}}}{N_{\text{in}}G} = \frac{N_{\text{out}}}{GkTB} \tag{5-97}$$

$$N_{\text{out}} = FGkTB \tag{5-98}$$

$$N_{\text{in}} = kTB \tag{5-99}$$

where N_{in} is the total input noise power to a noise-free amplifier. The input phase noise in a 1-Hz bandwidth at any frequency $f_0 + f_m$ from the carrier produces a phase deviation given by (Figure 5-67)

$$\Delta\theta_{\text{peak}} = \frac{V_{n\,\text{RMS1}}}{V_{s,\text{av RMS}}} = \sqrt{\frac{fKT}{P_{s,\text{av}}}} \tag{5-100}$$

$$\Delta\theta_{\text{1RMS}} = \frac{1}{\sqrt{2}}\sqrt{\frac{FkT}{P_{s,\text{av}}}} \tag{5-101}$$

Since a correlated random phase-noise relation exists at $f_0 - f_m$, the total phase deviation becomes

$$\Delta\theta_{\text{RMS total}} = \sqrt{FkT/P_{s,\text{av}}} \tag{5-102}$$

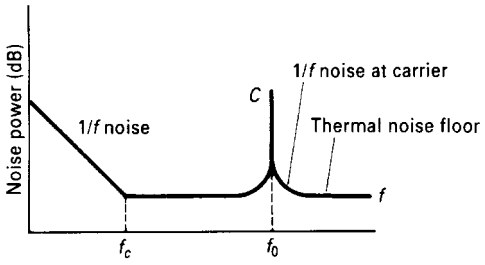


Figure 5-66 Noise power versus frequency of a transistor amplifier with an input signal applied.

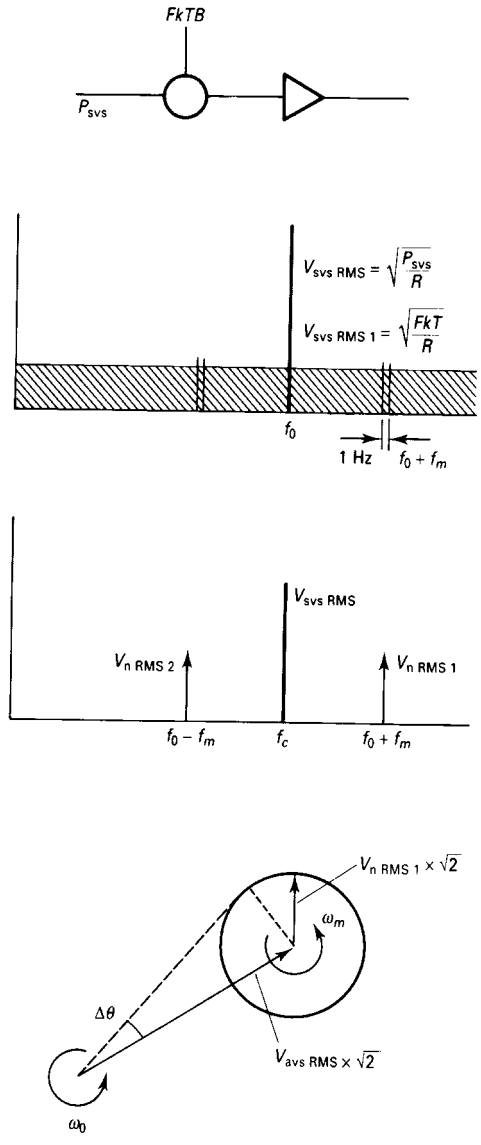


Figure 5-67 Phase noise added to carrier.

The spectral density of phase noise becomes

$$S_\theta(f_m) = \Delta\theta^2_{\text{RMS}} = FkTB/P_{s,\text{av}} \tag{5-103}$$

where $B = 1$ for a 1-Hz bandwidth. Using

$$kTB = -174 \text{ dBm/Hz} \qquad (B = 1) \tag{5-104}$$

allows a calculation of the spectral density of phase noise that is far removed from the carrier (i.e., at large values of f_m). This noise is the theoretical noise floor of the amplifier. For example, an amplifier with +10-dBm power at the input and a noise figure of 6 dB gives

$$S_\theta(f_m > f_c) = -174 \text{ dBm} + 6 \text{ dB} - 10 \text{ dBm} = -178 \text{ dBm} \tag{5-105}$$

Only if P_{OUT} is >0 dBm can we expect $\mathcal{L}$ (signal-to-noise ratio) to be greater than 174 dBc/Hz (1-Hz bandwidth). For a modulation frequency close to the carrier, $S_\theta(f_m)$ shows a flicker or $1/f$ component, which is empirically described by the corner frequency f_c. The phase noise can be modeled by a noise-free amplifier and a phase modulator at the input as shown in Figure 5-68.

The purity of the signal is degraded by the flicker noise at frequencies close to the carrier. The spectral phase noise can be described by

$$S_\theta(f_m) = \frac{FkTB}{P_{s,\text{av}}}\left(1 + \frac{f_c}{f_m}\right) \quad (B = 1) \tag{5-106}$$

No AM-to-PM conversion is considered in this equation. The oscillator may be modeled as an amplifier with feedback as shown in Figure 5-69. The phase noise at the input of the amplifier is affected by the bandwidth of the resonator in the oscillator circuit in the following way. The tank circuit or bandpass resonator has a low-pass transfer function

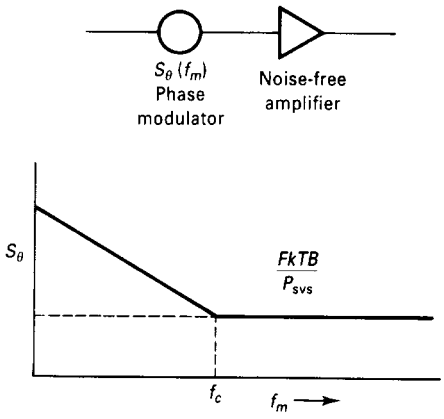


Figure 5-68 Phase noise modeled by a noise-free amplifier and phase modulator.

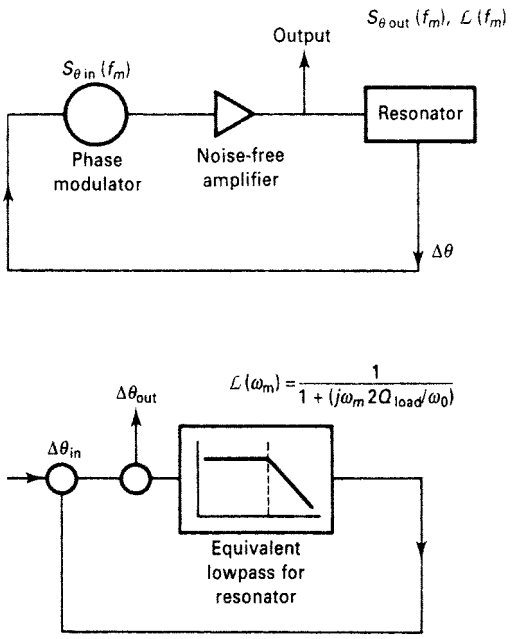


Figure 5-69 Equivalent feedback models of oscillator phase noise.

$$L(\omega_m) = \frac{1}{1 + j(2Q_L\omega_m/\omega_0)} \tag{5-107}$$

where

$$\omega_0/2Q_L = B/2 \tag{5-108}$$

is the half-bandwidth of the resonator. These equations describe the amplitude response of the bandpass resonator; the phase noise is transferred unattenuated through the resonator up to the half-bandwidth.

The closed-loop response of the phase feedback loop is given by

$$\Delta\theta_{\text{out}}(f_m) = \left(1 + \frac{\omega_0}{j2Q_L\omega_m}\right)\Delta\theta_{\text{in}}(f_m) \tag{5-109}$$

The power transfer becomes the phase spectral density

$$S_{\theta,\text{out}}(f_m) = \left[1 + \frac{1}{f_m^2}\left(\frac{f_0}{2Q_L}\right)^2\right] S_{\theta,\text{in}}(f_m) \tag{5-110}$$

where $S_{\theta,\text{in}}$ was given by Eq. (5-106). Finally, $\mathcal{L}(f_m)$ is

$$\mathcal{L}(f_m) = \frac{1}{2}\left[1 + \frac{1}{f_m^2}\left(\frac{f_0}{2Q_L}\right)^2\right] S_{\theta,\text{in}}(f_m) \tag{5-111}$$

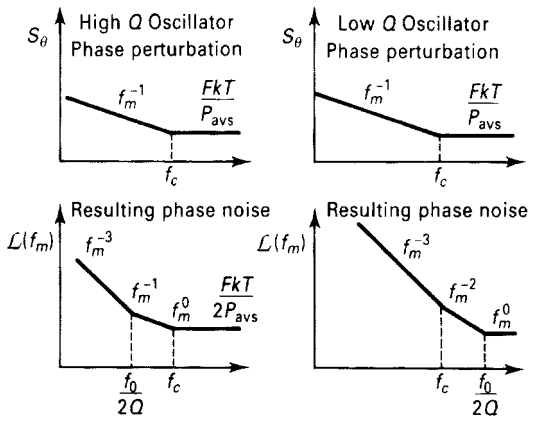


Figure 5-70 Equivalent feedback models of oscillator phase noise.

This equation describes the phase noise at the output of the amplifier (flicker corner frequency and AM-to-PM conversion are not considered). The phase perturbation $S_{\theta,\text{in}}$ at the input of the amplifier is enhanced by the positive phase feedback within the half-bandwidth of the resonator, $f_0/2Q_L$.

Depending on the relation between f_c and $f_0/2Q_L$, there are two cases of interest, as shown in Figure 5-70. For the low-Q case, the spectral phase noise is unaffected by the Q of the resonator, but the $\mathcal{L}(f_m)$ spectral density will show a $1/f^3$ and $1/f^2$ dependence close to the carrier. For the high-Q case, a region of $1/f^3$ and $1/f$ should be observed near the carrier. Substituting Eq. (5-106) into Eq. (5-111) gives an overall noise of

$$\mathcal{L}(f_m) = \frac{1}{2}\left[1 + \frac{1}{f_m^2}\left(\frac{f}{2Q_L}\right)^2\right]\frac{FkT}{P_{s,\text{av}}}\left(1 + \frac{f_c}{f_m}\right)$$

$$= \frac{FkTB}{2P_{s,\text{av}}}\left[\frac{1}{f_m^3}\frac{f^2 f_c}{4Q_L^2} + \frac{1}{f_m^2}\left(\frac{f}{2Q_L}\right)^2 + \left(1 + \frac{f_c}{f_m}\right)\right] \text{dBc/Hz} \tag{5-112}$$

Examining Eq. (5-112) gives the four major causes of oscillator noise: the upconverted $1/f$ noise or flicker FM noise, the thermal FM noise, the flicker phase noise, and the thermal noise floor, respectively.

Q_L (loaded Q) can be expressed as

$$Q_L = \frac{\omega_o W_e}{P_{\text{diss,total}}} = \frac{\omega_o W_e}{P_{\text{in}} + P_{\text{res}} + P_{\text{sig}}} = \frac{\text{Reactive power}}{\text{Total dissipated power}} \tag{5-113}$$

where W_e is the reactive energy stored in L and C,

$$W_e = \frac{1}{2} CV^2 \tag{5-114}$$

$$P_{\text{res}} = \frac{\omega_o W_e}{Q_{\text{unl}}} \tag{5-115}$$

$$\mathcal{L}(f_m) = \frac{1}{2}\left[1 + \frac{\omega_o^2}{4\omega_m^2}\left(\frac{P_{\text{in}}}{\omega_o W_e} + \frac{1}{Q_{\text{unl}}} + \frac{P_{\text{sig}}}{\omega_o W_e}\right)^2\right]\left(1 + \frac{\omega_c}{\omega_m}\right)\frac{FkT_o}{P_{\text{sav}}} \tag{5-116}$$

Input power/ Reactive power ratio

Resonator Q

Signal power/ Reactive power ratio

Flicker effect

Phase perturbation

This equation is extremely significant because it covers most of the causes of phase noise in oscillators. [AM-to-PM conversion must be added; see Eq. (5-56).] To minimize the phase noise, the following design rules apply:

1. Maximize the unloaded Q.
2. Maximize the reactive energy by means of a high RF voltage across the resonator and obtain a low L/C ratio. The limits are set by breakdown voltages of the active devices and the tuning diodes, and the forward-bias condition of the tuning diodes. If L becomes too high, the circuit degenerates into a squegging oscillator. Using lower L allows larger capacitance, which allows us to swamp the oscillating device's internal, nonlinear capacitance with external, linear capacitance, reducing the phase-noise degradation caused by the nonlinear capacitance. The ceramic-resonator oscillator best illustrates this approach, which may be hard to accomplish in discrete circuits because high-Q inductors with values above 1 nH cannot be built unless high-Q transmission lines or resonators are used.
3. Avoid saturation at all costs, and try to have either limiting or automatic gain control (AGC) without degradation of Q. Isolate the tuned circuit from the limiter or AGC circuit. Use tuning diodes in antiseries configurations to avoid forward bias.
4. Choose an active device with the lowest possible noise figure and flicker corner frequency. The noise figure of interest is the noise figure obtained with the actual impedance at which the device is operated. Using FETs rather than BJTs, it is preferable to deal with the equivalent noise voltage and noise currents rather than with the noise figure, since they are independent of source impedance. The noise figure improves as the ratio between source impedance and equivalent noise resistance increases. In addition, in a tuned circuit, the source impedance changes drastically as a function of the offset frequency, and this effect has to be considered. For low phase-noise operation, use a medium-power transistor. If you need your output power to be achieved at 6–9 mA, select a transistor with an $I_{C,\max}$ of 60–90 mA. Also avoid an f_T greater than 3× to 5× the operating frequency. To make transistors that are stable across their full frequency ranges, manufacturers add circuitry that makes the flicker corner frequency higher as f_T increases.

   The following transistors have the lowest noise figure:

   - BFG65/67 by Philips
   - BFR106/92; BFP405/420 by Siemens
   - 2SC3358/3356 by NEC
   - HXTR4105 by HP

- AT2100 (chip) by Avantek

Among the lowest-noise JFETs are the U310, 2N4416/17, and 2N5397.

In designing MMICs, one must resort to MESFETs, which are also referred to as GaAsFETs. They have a fairly high flicker frequency, which is typically around 10 MHz but can go up to as high as 100 MHz.

Table 5-5 shows equivalent circuits and measurement data for the Avantek AT2100 20-GHz *npn* silicon bipolar oscillator transistor. (Siemens of Germany is about to come out with 45-GHz f_T transistors, which will compete with SiGe transistors in the future.)

5. Phase perturbation can be minimized by using high-impedance devices such as FETs, where the signal-to-noise ratio of the signal voltage relative to the equivalent noise voltage can be made very high. This also indicates that in the case of a limiter, the limited voltage should be as high as possible.
6. Choose an active device with low flicker noise. The effect of flicker noise can be reduced through RF feedback. An unbypassed emitter resistor of 10–30 Ω in a bipolar circuit can improve the flicker noise by as much as 40 dB.

   The proper bias point of the active device is important, and precautions should be taken to prevent modulation of the input and output dynamic capacitances of the active device, which will cause amplitude-to-phase conversion and therefore introduce noise.
7. The energy should be coupled from the resonator rather than from another portion of the active device so that the resonator limits the bandwidth.
8. For microwave applications, the lowest-noise GaAsFETs are pseudomorphic HJ FETs. A reproduction of the datasheet of the NEC NE42484A, which is generally the state of the art today, follows; it is used with permission.

Equation (5-116) assumes that the phase perturbation and the flicker effect are the limiting factors, as practical use of such oscillators requires that an isolation amplifier must be used.

In the event that energy is taken directly from the resonator and the oscillator power can be increased, the signal-to-noise ratio can be increased above the theoretical limit of 174 dBc/Hz, due to the low-pass-filter effect of the tuned resonator. However, since this is mainly a theoretical assumption and does not represent a real-world system, this noise performance cannot be obtained. In an oscillator stage, even a total noise floor of –170 dBc/Hz is rarely achieved.

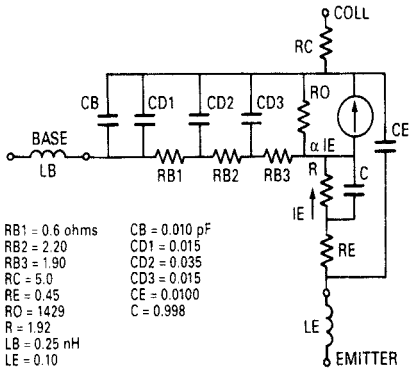
What other influences do we have that cause an oscillator's noise performance to degrade? So far, we have assumed that the Q of the tuned circuit is really determined only by the LC network and the loading effect of the transistor. In synthesizer applications, however, we find it necessary to add a tuning diode. The tuning diode has a substantially lower Q than that of a mica capacitor or even a ceramic capacitor. As a result of this, the noise sidebands change as a function of the additional loss. This is best expressed by adjusting the value for the loaded Q in Eq. (5-113).

There seems to be no precise mathematical way of predetermining the noise influence of a tuning diode, but the following approximation seems to give proper results:

$$\frac{1}{Q_{T,\text{load}}} = \frac{1}{Q_{\text{load}}} + \frac{1}{Q_{\text{diode}}} \tag{5-117}$$

Table 5-5 Linear and SPICE equivalent circuits for the Avantek AT21400 chip

LINEAR EQUIVALENT CIRCUIT[a]
V_{CE} = 8 V, I_C = 14 mA

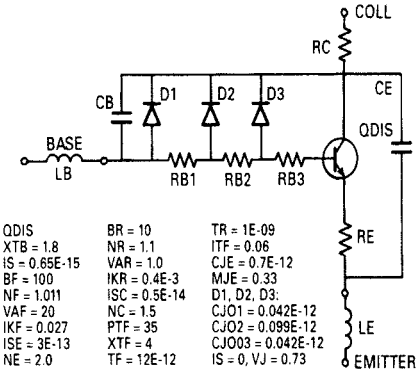


RB1 = 0.6 ohms
RB2 = 2.20
RB3 = 1.90
RC = 5.0
RE = 0.45
RO = 1429
R = 1.92
LB = 0.25 nH
LE = 0.10

CB = 0.010 pF
CD1 = 0.015
CD2 = 0.035
CD3 = 0.015
CE = 0.0100
C = 0.998

CCCS: M = .99, A = 0
R1 = R, R2 = 1E8, F = 22.7, T = 7.4

SPICE EQUIVALENT CIRCUIT[a]



QDIS
XTB = 1.8
IS = 0.65E-15
BF = 100
NF = 1.011
VAF = 20
IKF = 0.027
ISE = 3E-13
NE = 2.0

BR = 10
NR = 1.1
VAR = 1.0
IKR = 0.4E-3
ISC = 0.5E-14
NC = 1.5
PTF = 35
XTF = 4
TF = 12E-12

TR = 1E-09
ITF = 0.06
CJE = 0.7E-12
MJE = 0.33
D1, D2, D3:
CJO1 = 0.042E-12
CJO2 = 0.099E-12
CJO03 = 0.042E-12
IS = 0, VJ = 0.73

Modeled Scattering Parameters, Common Emitter[b]

T_A = 25°C, V_{CE} = 8 V, I_C = 14 mA

Freq.	S_{11}		S_{21}			S_{12}			S_{22}	
GHz	Mag	Ang	dB	Mag	Ang	dB	Mag	Ang	Mag	Ang
2	.67	−168	15.8	6.19	90	−34.0	.02	52	.65	−9
3	.67	−180	12.4	4.17	81	−32.0	.025	59	.65	−10
4	.67	172	10.0	3.15	74	−30.5	.03	64	.65	−11
5	.68	165	8.1	2.53	68	−28.0	.04	68	.65	−12
6	.68	160	6.5	2.12	62	−28.0	.04	70	.65	−14
7	.69	154	5.2	1.82	56	−26.0	.05	72	.65	−16
8	.70	149	4.1	1.60	51	−24.4	.06	73	.65	−18
9	.70	144	3.1	1.43	46	−24.4	.06	74	.65	−20
10	.71	140	2.2	1.29	41	−23.1	.07	74	.65	−22
11	.72	135	1.4	1.18	36	−21.9	.08	74	.65	−24
12	.73	131	0.7	1.08	31	−21.9	.08	75	.66	−26
13	.74	127	0	1.00	27	−20.9	.09	75	.66	−28
14	.75	123	−0.6	0.93	23	−20.0	.10	75	.66	−30
15	.76	120	−1.3	0.86	18	−19.2	.11	74	.67	−32
16	.77	116	−1.9	0.80	14	−19.2	.11	74	.67	−34
18	.79	109	−3.1	0.70	7	−17.7	.13	74	.68	−39
20	.80	103	−4.2	0.62	−1	−17.1	.14	73	.69	−44
22	.82	97	−5.4	0.54	−7	−15.9	.16	72	.70	−48
24	.84	92	−6.6	0.47	−14	−14.9	.18	71	.71	−53
26	.85	87	−8.0	0.40	−20	−14.0	.20	70	.72	−58

[a]These equivalent circuits are provided only as first-order design aids. Their accuracy for critical designs at very high frequencies has not been validated.

[b]*S* parameters are from the linear equivalent circuit. Below 10 GHz, they have been fit to measurements of die on a standard carrier with one bond wire to the base and four wires to the emitter.

NEC

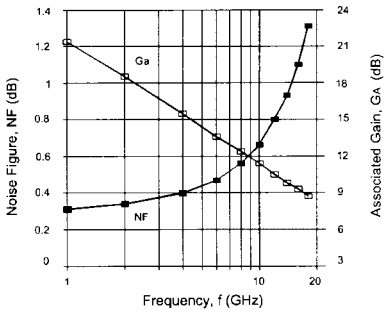
SUPER LOW NOISE PSEUDOMORPHIC HJ FET

NE42484A

FEATURES

- **VERY LOW NOISE FIGURE:** 0.8 dB TYP at 12 GHz
- **HIGH ASSOCIATED GAIN:** 10.5 dB TYP at 12 GHz
- **L_G = 0.35 µm, W_G = 200 µm**
- **LOW COST METAL CERAMIC PACKAGE**
- **TAPE & REEL PACKAGING OPTION AVAILABLE**

NOISE FIGURE & ASSOCIATED GAIN vs. FREQUENCY
V_{DS} = 2 V, I_{DS} = 10 mA



DESCRIPTION

The NE42484A is a pseudomorphic Hetero-Junction FET that uses the junction between Si-doped AlGaAs and undoped InGaAs to create very high mobility electrons. The device features mushroom shaped TiAl gates for decreased gate resistance and improved power handling capabilities. The mushroom gate also results in lower noise figure and high associated gain. This device is housed in an epoxy-sealed, metal/ceramic package and is intended for high volume consumer and industrial applications.

NEC's stringent quality assurance and test procedures ensure the highest reliability and performance.

ELECTRICAL CHARACTERISTICS (T_A = 25°C)

PART NUMBER PACKAGE OUTLINE			NE42484A 84AS		
SYMBOLS	**PARAMETERS AND CONDITIONS**	**UNITS**	**MIN**	**TYP**	**MAX**
NF_{OPT}[1]	Optimum Noise Figure, V_{DS} = 2.0 V, I_{DS} = 10 mA, f = 12 GHz	dB		0.8	1.2
G_A[1]	Associated Gain, V_{DS} = 2.0 V, I_{DS} = 10 mA, f = 12 GHz	dB	9.0	10.5	
P_{1dB}	Output Power at 1 dB Gain Compression Point, f = 12 GHz V_{DS} = 2.0 V, I_{DS} = 10 mA V_{DS} = 2.0 V, I_{DS} = 20 mA	 dBm dBm		 9.7 10.2	
G_{1dB}	Gain at P_{1dB}, f = 12 GHz V_{DS} = 2.0 V, I_{DS} = 10 mA V_{DS} = 2.0 V, I_{DS} = 20 mA	 dB dB		 10.3 10.5	
I_{DSS}	Saturated Drain Current, V_{DS} = 2.0 V, V_{GS} = 0 V	mA	15	40	70
V_P	Pinch-off Voltage, V_{DS} = 2.0 V, I_{DS} = 0.1 mA	V	-2.0	-0.8	-0.2
g_m	Transconductance, V_{DS} = 2.0 V, I_D = 10 mA	mS	45	60	
I_{GSO}	Gate to Source Leakage Current, V_{GS} = -3.0 V	µA		0.5	10.0
$R_{TH\ (CH-A)}$	Thermal Resistance (Channel to Ambient)	°C/W		750	
$R_{TH(CH-C)}$[2]	Thermal Resistance (Channel to Case)	°C/W			350

Notes:
1. Typical values of noise figures and associated gain are those obtained when 50% of the devices from a large number of lots were individually measured in a circuit with the input individually tuned to obtain the minimum value. Maximum values are criteria established on the production line as a "go-no-go" screening tuned for the "generic" type but not for each specimen.
2. R_{TH} (channel to case) for package mounted on an infinite heat sink.

NE42484A

ABSOLUTE MAXIMUM RATINGS[1] (T_A = 25°C)

SYMBOLS	PARAMETERS	UNITS	RATINGS
V_{DS}	Drain to Source Voltage	V	4.0
V_{GS}	Gate to Source Voltage	V	-3.0
I_{DS}	Drain Current	mA	I_{DSS}
I_{GRF}	Gate Current (RF Drive)	µA	200
P_{IN}	RF Input (CW)	dBm	15
T_{CH}	Channel Temperature	°C	150
T_{STG}	Storage Temperature	°C	-65 to +150
P_T	Total Power Dissipation	mW	165

Note:
1. Operation in excess of any one of these parameters may result in permanent damage.

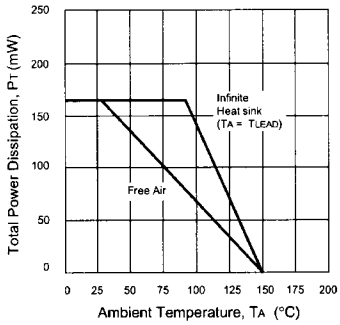
TYPICAL NOISE PARAMETERS (T_A = 25°C)

V_{DS} = 2 V, I_{DS} = 10 mA

FREQ. (GHz)	NF_{OPT} (dB)	G_A (dB)	Γ_{OPT} MAG	Γ_{OPT} ANG	Rn/50
1.0	0.31	21.4	0.78	10	0.43
2.0	0.34	18.5	0.76	28	0.38
4.0	0.40	15.5	0.72	58	0.28
6.0	0.47	13.6	0.65	84	0.21
8.0	0.56	12.4	0.57	113	0.15
10.0	0.66	11.4	0.50	141	0.10
12.0	0.80	10.5	0.44	173	0.09
14.0	0.93	9.8	0.39	-157	0.08
16.0	1.10	9.3	0.36	-125	0.08
18.0	1.31	8.7	0.35	-90	0.08

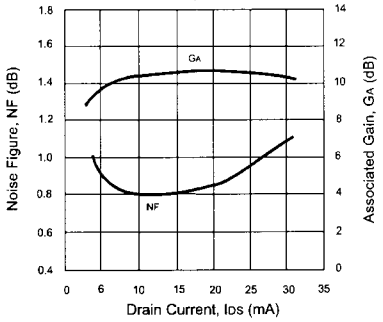
TYPICAL PERFORMANCE CURVES (T_A = 25°C)

TOTAL POWER DISSIPATION vs. AMBIENT TEMPERATURE[1]

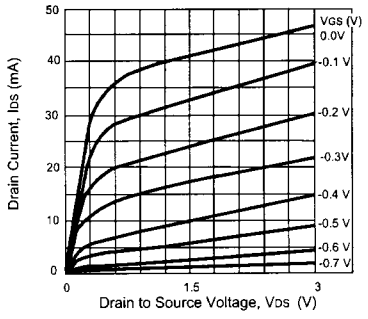


Note
1. If P_T exceeds the Free Air Value, reliable operation can be assured by measuring the worst-case temperature, $T_{(LEAD)}$, at the lead where heat flow is maximum (usually the source lead) and limiting T_A, P_T or $R_{TH\ (CKT)}$

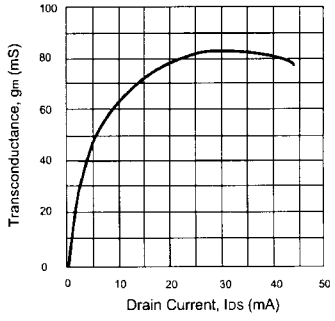
NOISE FIGURE AND ASSOCIATED GAIN vs. DRAIN CURRENT V_{DS} = 2 V, f = 12 GHz



DRAIN CURRENT vs. DRAIN TO SOURCE VOLTAGE



TRANSCONDUCTANCE vs. DRAIN CURRENT V_{DS} = 2.0 V



The tuning diode is specified to have a cutoff frequency $f_{\max}$, which is determined from the loss resistor R_S and the value of the junction capacitance as a function of voltage (e.g., measured at 3 V). This means that the voltage determines the Q and, consequently, the noise bandwidth.

5-6-2 AM-to-PM Conversion

We went into more detail in dealing with the mechanism and influence of tuning diodes in the voltage-controlled oscillators section, where we evaluated the various methods of building voltage-tunable oscillators using tuning diodes. In this section, we limit ourselves to practical results.

The loading effect of the tuning diode is due to losses, and these losses can be described by a resistor in parallel with the tuned circuit.

It is possible to define an equivalent noise $R_{a,\text{eq}}$ that, inserted in Nyquist's equation,

$$V_n = \sqrt{4kT_0R\,\Delta f} \tag{5-118}$$

where $kT_0 = 4.2 \times 10^{-21}$ at about 300 K, R is the equivalent noise resistor, and Δf is the bandwidth, determines an open-circuit noise voltage across the tuning diode. Practical values of $R_{a,\text{eq}}$ for carefully selected tuning diodes are in the vicinity of 200 Ω to 50 kΩ. If we now determine the noise voltage $V_n = \sqrt{4 \times 4.2 \times 10^{-21} \times 10{,}000}$, the resulting voltage value is 1.296×10^{-8} V $\sqrt{\text{Hz}}$.

This noise voltage generated from the tuning diode is now multiplied with the VCO gain K_0, resulting in the rms frequency deviation

$$(\Delta f_{\text{rms}}) = K_0 \times (1.296 \times 10^{-8}\ \text{V}) \text{ in 1-Hz bandwidth} \tag{5-119}$$

To translate this into an equivalent peak phase deviation,

$$\theta_d = \frac{K_0\sqrt{2}}{f_m}(1.296 \times 10^{-8}) \text{ rad in 1-Hz bandwidth} \tag{5-120}$$

or for a typical oscillator gain of 100 kHz/V,

$$\theta_d = \frac{0.00183}{f_m} \text{ rad in 1-Hz bandwidth} \tag{5-121}$$

For f_m = 25 kHz (typical spacing for adjacent-channel measurements for FM mobile radios), $\theta_c = 7.32 \times 10^{-8}$. This can be converted now into the SSB signal-to-noise ratio:

$$\mathcal{L}(f_m) = 20 \log_{10}\left(\frac{\theta_c}{2}\right) = -149 \text{ dBc/Hz} \tag{5-122}$$

For the typical oscillator gain of 10 MHz/V found in wireless applications, the resulting phase noise will be 20 dB worse [10 log (10 MHz ÷ 100 kHz)]. However, the best tuning diodes, like the BB104, have an R_n of 200 Ω instead of 10 kΩ, which again changes the picture. According to Eq. (5-118), with $kT_0 = 4.2 \times 10^{-21}$ the resulting noise voltage will be

$$V_n = \sqrt{4 \times 4.2 \times 10^{-21} \times 200}$$

$$= 1.833 \times 10^{-9} \text{ V}\sqrt{\text{Hz}} \tag{5-123}$$

From Eq. (5-114), the equivalent peak phase deviation for a gain of 10 MHz/V in a 1-Hz bandwidth is then

$$\theta_d = \frac{1 \times 10^7 \sqrt{2}}{f_m} (1.833 \times 10^{-9}) \text{ rad} \tag{5-124}$$

or

$$\theta_d = \frac{0.026}{f_m} \text{ rad in 1-Hz bandwidth} \tag{5-125}$$

With $f_m = 25$ kHz, $\theta_c = 1.04 \times 10^{-6}$. Expressing this as phase noise:

$$\mathcal{L}(f_m) = 20 \log_{10}\left(\frac{\theta_c}{2}\right) = -126 \text{ dBc/Hz} \tag{5-126}$$

The phase-noise value in Eq. (5-122) is that typically achieved in the Rohde & Schwarz SMDU or with the Hewlett-Packard 8640 signal generator and considered state-of-the-art for a free-running oscillator. It should be noted that both signal generators use a slightly different tuned circuit; the Rohde & Schwarz generator uses a helical resonator, whereas the Hewlett-Packard generator uses an electrically shortened quarter-wavelength cavity. Both generators are mechanically pretuned, and a tuning diode with a gain of about 100 kHz/V is used for AFC and frequency-modulation purposes. It is apparent that, because of the nonlinearity of the tuning diode, the gain is different for low dc voltages than for high dc voltages. The impact of this is that the noise varies with the tuning range.

If this oscillator had to be used for a synthesizer, its 1-MHz tuning range would be insufficient; therefore, a way had to be found to segment the band into the necessary ranges. In VCOs, this is typically done with switching diodes that allow the proper frequency bands to be selected. These switching diodes insert in parallel or series, depending on the circuit, with additional inductors or capacitors, depending on the design.

In low-energy-consuming circuits, the VCO frequently is divided into a coarse-tuning section using tuning diodes and a fine-tuning section with a tuning diode. In the coarse-tuning range, this results in very high gains, such as 1–10 MHz/V, for the diodes. The noise contribution of those diodes is therefore very high and can hardly be compensated by the loop. For low-noise applications, which automatically mean higher power consumption, it is unavoidable to use switching diodes.

Let us now examine some test results. Referring back to Eq. (5-106), Figure 5-71 shows the noise sideband performance as a function of Q, whereby the top curve with $Q_L = 100$ represents a somewhat poor oscillator and the lowest curve with $Q_L = 100{,}000$ probably represents a crystal oscillator where the unloaded Q of the crystal was in the vicinity of 3×10^6. Figure 5-72 shows the influence of flicker noise.

Corner frequencies of 10 Hz to 10 kHz have been selected, and it becomes apparent that around 1 kHz the influence is fairly dramatic, whereas the influence at 20 kHz off the carrier

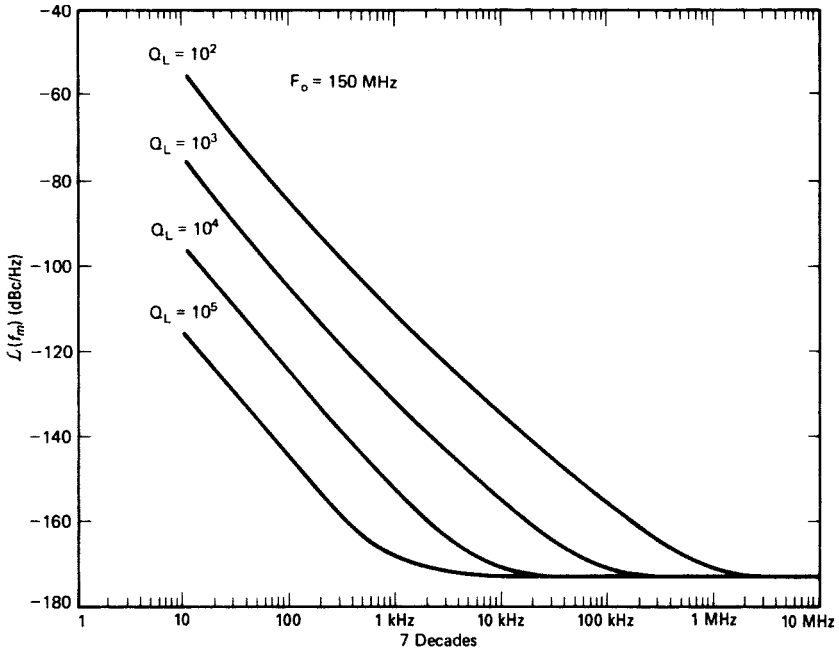


Figure 5-71 Noise sideband of an oscillator at 150 MHz as a function of the loaded Q of the resonator.

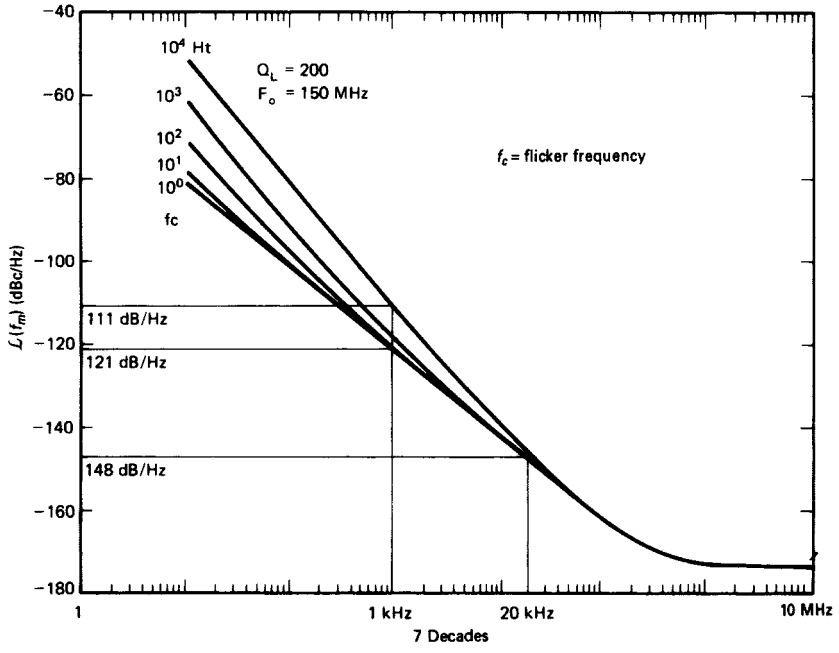


Figure 5-72 Noise-sideband performance as a function of the flicker frequency f_c varying from 10 Hz to 10 kHz.

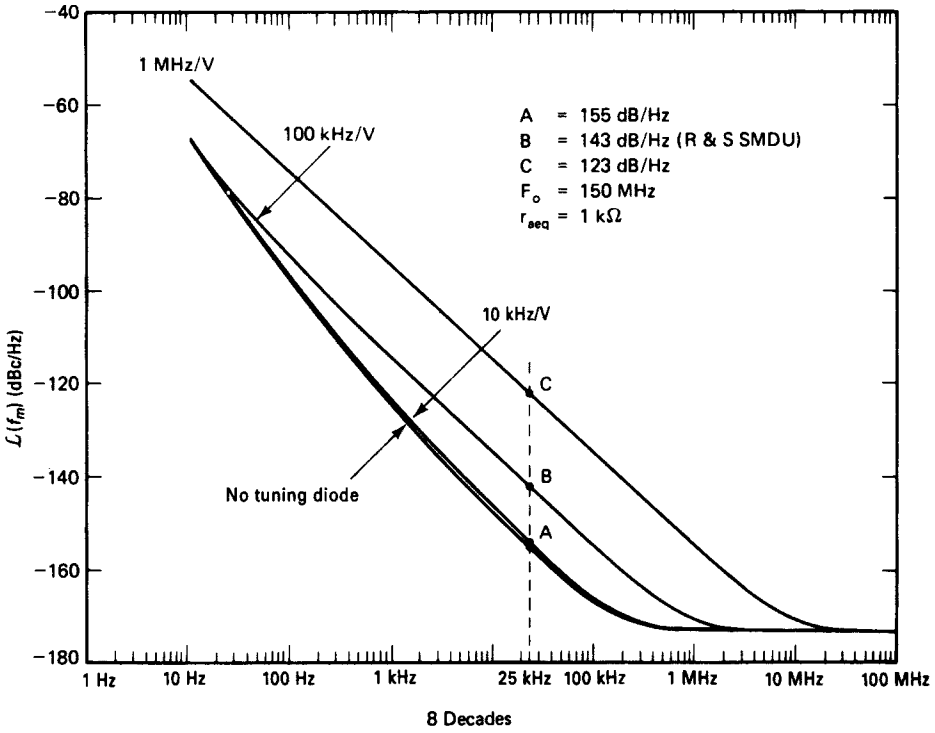


Figure 5-73 Noise-sideband performance of an oscillator at 150 MHz, showing the influence of various tuning diodes.

is not significant. Figures 5-73 and 5-74 show the influence of the tuning diodes on a high-Q oscillator.

Curve A in Figure 5-73 uses a lightly coupled tuning diode with a K_0 of 10 kHz/V; the lower curve is the noise performance without any diode. As a result, the two curves are almost identical, which can be seen from the somewhat smeared form of the graph. Curve B shows the influence of a tuning diode at 100 kHz/V and represents a value of −143 dBc/Hz from −155 dBc/Hz, already some deterioration. Curve C shows the noise if the tuning diode results in a 1-MHz/V VCO gain, and the noise sideband at 25 kHz has now deteriorated to −123 dBc/Hz. These curves speak for themselves.

Figure 5-74 shows how the phase noise of a 147-MHz VCO varies with the bias applied to its tuning diode. Because a tuning diode's Q increases with the reverse bias applied to it, we expect a VCO's phase-noise performance to improve, even if only slightly, as its tuning voltage is increased. Yet Figure 5-74 illustrates the reverse of this relationship. Inspecting this VCO's circuit (Figure 5-75), we see that the diode is in series with the circuit's 220-nH resonator. *All* of the RF current flowing through the resonator must flow through the relatively low-Q diode, and even small decreases in the diode's capacitance will significantly modify the tank's L/C ratio.

Graphing the diode's ac load line (Figure 5-76), we see that as the tuning voltage increases (i.e., as the tuning diode capacitance decreases), the RF voltage across the diode becomes disproportionately high. We note also that at every value of tuning voltage, the RF voltage across the diode goes positive (relative to the anode) during part of each cycle.

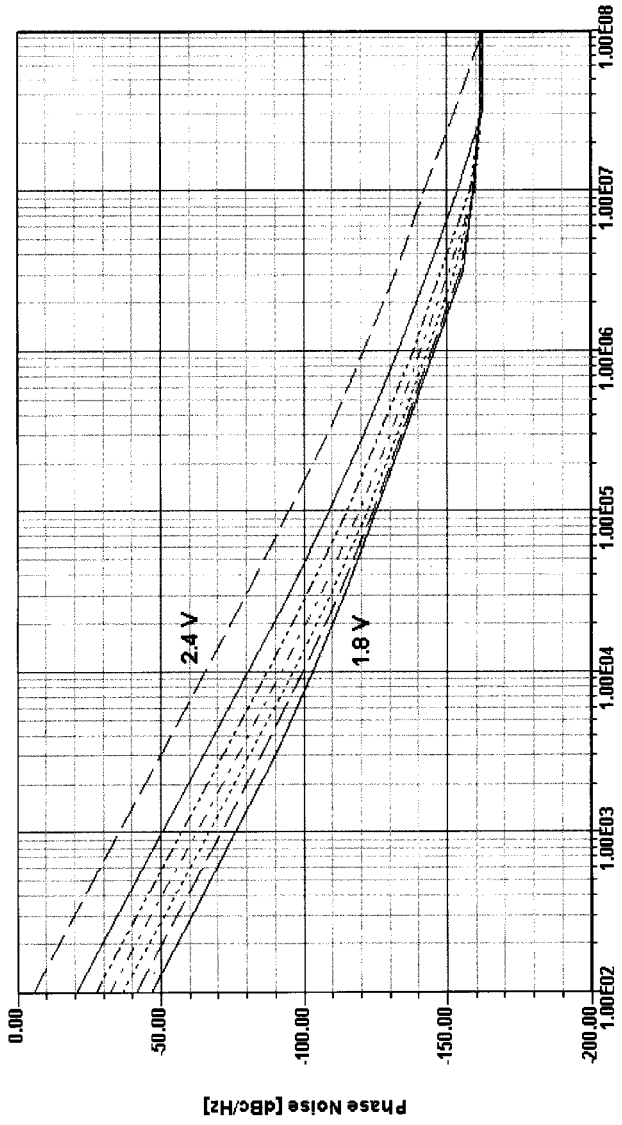


Figure 5-74 Noise-sideband performance of an oscillator at 147 MHz, showing the influence of a tuning diode operated at bias voltages from 1.8 to 2.4 V in 0.1-V increments.

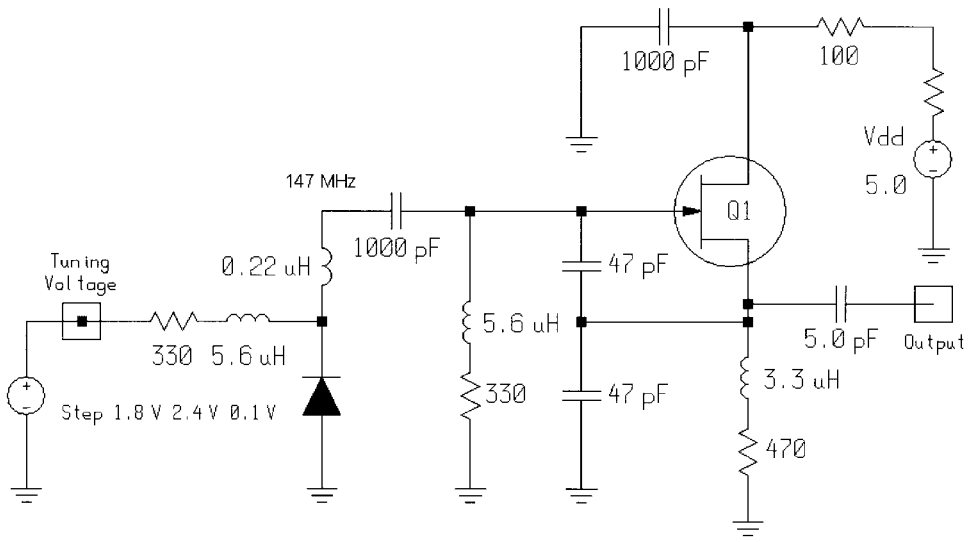


Figure 5-75 The VCO schematic. Because the tuning diode is in series with the resonator, it passes all of the tank's RF current and plays a disproportionately large role in determining the circuit's phase-noise performance.

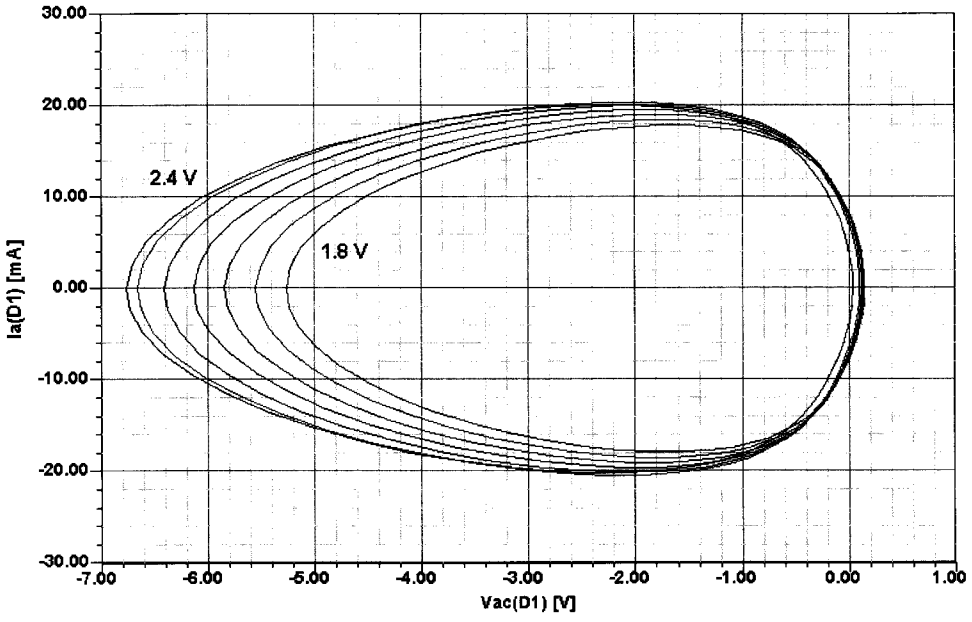


Figure 5-76 Graphing the tuning diode's ac load line reveals that increasing the diode's tuning bias (decreasing its capacitance) results in a disproportionally large increase in the RF voltage across the diode. At all tuning voltages, the RF voltage across the diode is slightly positive (relative to the anode) during part of each cycle.

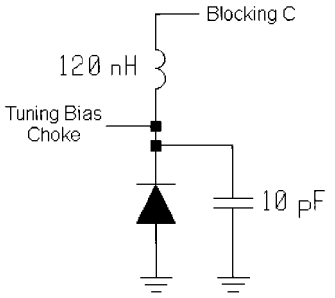


Figure 5-77 Adding capacitance across the tuning diode reduces the effect of its nonlinearities on phase-noise performance. We have also revalued the resonator to 120 nH to maintain oscillation near 147 MHz.

Modifying the oscillator tank to add capacitance in parallel with the tuning diode (Figure 5-77) reduces the tuning diode's effect on tank Q and L/C ratio, and therefore its contribution to phase noise. With this arrangement, we now see the expected relationship between tuning bias and phase noise: The higher the tuning bias, the lower the phase noise (Figure 5-78). Figure 5-79 shows that the diode's ac load line is now virtually unchanged across the tuning-voltage range. The RF voltage across the diode still goes positive (relative to the anode) during part of each cycle.

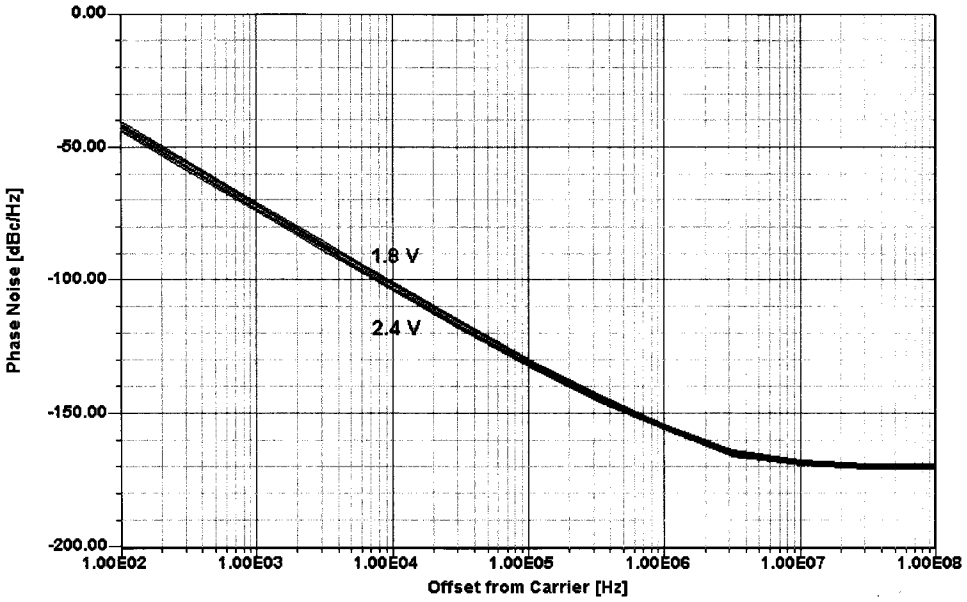


Figure 5-78 Now that the tuning diode passes only part of the tank's RF current, varying its capacitance does not result in the large change in phase-noise performance shown in Figure 5-74. Note also that the phase-noise-versus-tuning-bias relationship has reversed relative to that shown in Figure 5-74: The highest tuning voltage now corresponds to the lowest phase noise.

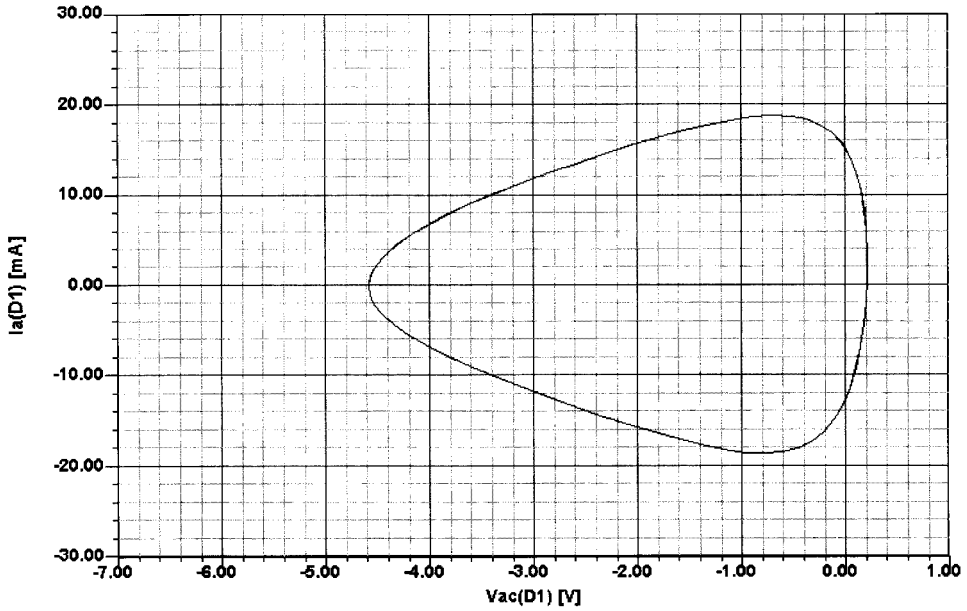


Figure 5-79 The ac load line for the tuning diode in Figure 5-77. Now the voltage swing across the diode is virtually the same at all tuning voltages. The RF voltage across the diode still goes positive (relative to the anode) during part of the cycle.

Figure 5-80 shows one more modification to the VCO tank. Now, the tuning diode is lightly coupled to the hot end of the tank, in parallel with the resonator. The result is phase-noise performance that is essentially unchanged across the 1.8–2.4-V tuning-voltage range. Figure 5-81 compares this circuit's phase-noise performance with the worst-case performance of the tank configurations in Figures 5-75 and 5-77. Figure 5-82 compares the diode ac load lines that correspond to the phase-noise curves in Figure 5-81. The best phase-noise performance corresponds to the diode ac load line that exhibits the narrowest voltage swing and never goes positive relative to the anode.

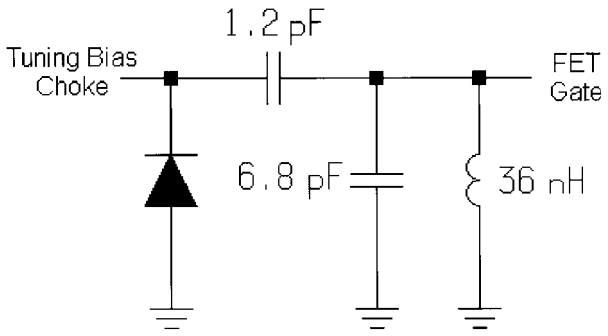


Figure 5-80 Tank circuit reconfigured as per Figure 5-54 to lightly couple the tuning diode in parallel with the resonator, with the resonator inductance again decreased to maintain oscillation near 147 MHz. The 1000-pF capacitor and gate choke of Figure 5-75 (5.6 μH in series with 330 Ω) are no longer needed.

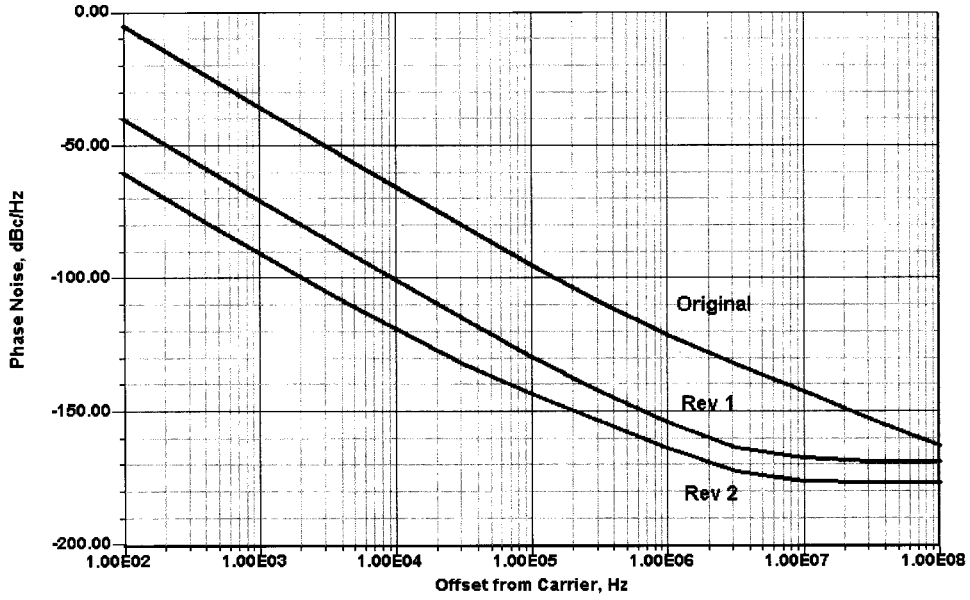


Figure 5-81 Comparison of worst-case phase-noise performance obtained with the tank configurations of Figures 5-75 (Original), 5-77 (Rev 1), and 5-80 (Rev 2).

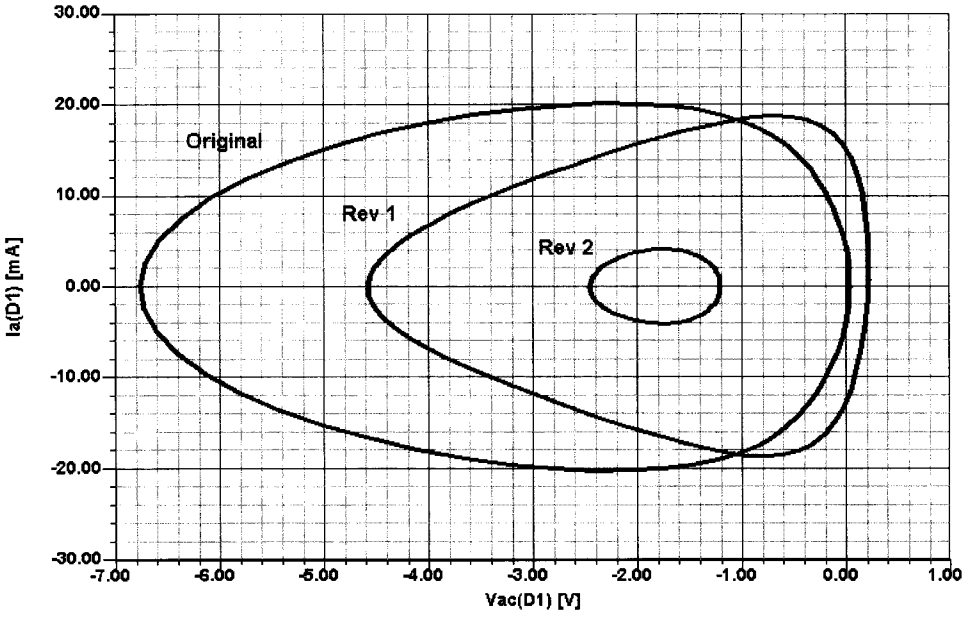


Figure 5-82 Comparison of the tuning diode's ac load line for the worst-case phase-noise curves of Figure 5-79.

It's important to keep in mind that we have not gained something for nothing in achieving the best phase-noise response in Figure 5-81. As we reduced the diode's phase-noise contribution, we also reduced the oscillator's tuning range—from 7.1% (Figure 5-75 circuit) to 1.8% (Figure 5-77 circuit) to 0.046% (Figure 5-80). Simultaneously optimizing phase-noise performance and tuning range requires a combination of techniques. First, we would maximize diode Q by using as high a tuning voltage as feasible. Second, we would use antiseries-connected diodes to minimize the RF voltage swing across them (Figure 5-55). Both of these techniques work against the achievement of wide tuning ranges by minimizing the capacitance swing that the diodes can contribute. As a result of this, and because hyperabrupt diodes are not available for the frequencies at which wireless VCOs operate, we would therefore use multiple antiseries diode pairs in parallel.

It is of interest to compare various oscillators. Figure 5-83 shows the performance of a 10-MHz crystal oscillator, a 40-MHz *LC* oscillator, the HP8640 cavity-tuned oscillator at 500 MHz, the 310–640-MHz switched-reactance oscillator of the HP8662 oscillator, and a 2–6-GHz YIG oscillator at 6 GHz.

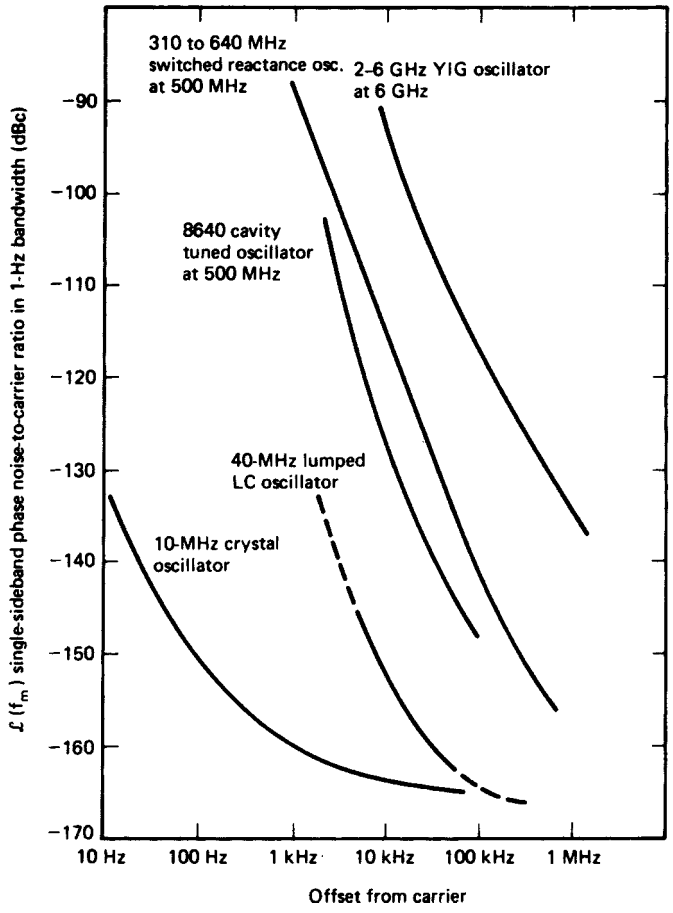


Figure 5-83 Comparison of noise-sideband performances of a crystal oscillator, *LC* oscillator, cavity-tuned oscillator, switched-reactance oscillator, and YIG oscillator.

5-6-3 Nonlinear Approach to the Calculation of Oscillator Phase Noise

The mechanism of noise generation in an oscillator combines the equivalent of frequency conversion (mixing) with the effect of AM-to-PM conversion. Therefore, to calculate oscillator phase noise, we must first be able to calculate the noise figure of a mixer. The following nonlinear approach is based on Ulrich L. Rohde, Chao-Ren Chang, and Jason Gerber, "Design and Optimization of Low-Noise Oscillators Using Nonlinear CAD Tools," *IEEE Frequency Control Symposium Proceedings*, 1994, pp. 548–554. This section presents the use of a novel algorithm for the computation of SSB carrier noise in free-running oscillators using the harmonic-balance (HB) nonlinear technique.

Traditional approaches relying on frequency conversion analysis are not sufficient to describe the complex physical behavior of a noisy oscillator. The accuracy of this nonlinear approach is based on the dynamic range of the harmonic-balance simulator and the quality of the parameter extraction for the active device. The algorithm described has also been verified with several examples up to millimeter wavelengths. This is the only algorithm that provides a complete and rigorous treatment of noise analysis for autonomous circuits.

Noise Generation in Oscillators. As shown earlier, the qualitative linearized picture of noise generation in oscillators is very well known [9]. The physical effects of random fluctuations taking place in the circuit are different depending on their spectral allocation with respect to the carrier:

- Noise components at low-frequency deviations result in frequency modulation of the carrier through mean square frequency fluctuation proportional to the available noise power.
- Noise components at high-frequency deviations result in phase modulation of the carrier through mean square phase fluctuation proportional to the available noise power.

We will demonstrate that the same conclusions can be derived quantitatively from the HB equations for an autonomous circuit.

Equivalent Representation of a Noisy Nonlinear Circuit. A general noisy nonlinear network can be described by the equivalent circuit shown in Figure 5-84. The circuit is divided into linear and nonlinear subnetworks as noise-free multiports. Noise generation is accounted for by connecting a set of noise voltage and noise current sources at the ports of the linear subnetwork [9–18].

Frequency Conversion Approach. The circuit supports a large-signal time-periodic steady state of fundamental angular frequency ω_0(carrier). Noise signals are small perturbations superimposed on the steady state, represented by families of pseudosinusoids located at the sidebands of the carrier harmonics. Therefore, the noise performance of the circuit is determined by the exchange of power among the sidebands of the unperturbed steady state through frequency conversion in the nonlinear subnetwork. Due to the perturbative assumption, the nonlinear subnetwork can be replaced with a multifrequency linear multiport described by a conversion matrix. The flow of noise signals can be computed by means of conventional linear circuit techniques.

The frequency conversion approach frequently used has limitations. The frequency-conversion approach is not sufficient to predict the noise performance of an autonomous

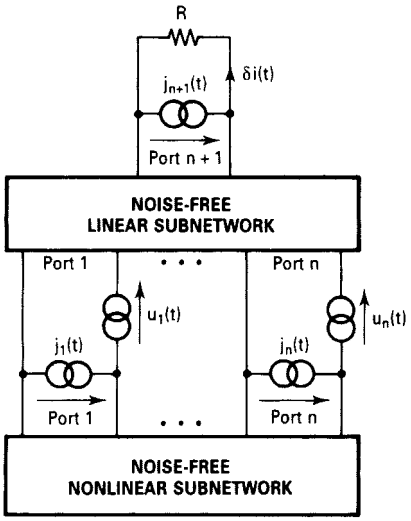


Figure 5-84 Equivalent circuit of a general noisy nonlinear network.

circuit. The spectral density of the output noise power and, consequently, the PM noise computed by the conversion analysis are proportional to the available power of the noise sources.

- In the presence of both thermal and flicker noise sources, PM noise increases as ω^{-1} for $\omega \to 0$ and tends to a finite limit for $\omega \to \infty$.
- Frequency conversion analysis correctly predicts the far-carrier noise behavior of an oscillator and, in particular, the oscillator noise floor; it does not provide results consistent with the physical observations at low deviations from the carrier.

This inconsistency can be removed by adding the modulation noise analysis. In order to determine the far-away noise using the autonomous circuit perturbation analysis, the following applies.

The circuit supports a large-signal time-periodic autonomous regime. The circuit is perturbed by a set of small sources located at the carrier harmonics and at the sidebands at a deviation ω from carrier harmonics. The perturbation of the circuit state (δX_B, δX_H) is given by the uncoupled sets of equations:

$$\frac{\partial E_B}{\partial E_B}\,\delta X_B = J_B(\omega) \tag{5-127}$$

$$\frac{\partial E_H}{\partial E_H}\,\delta X_H = J_H(\omega) \tag{5-128}$$

where E_B, E_H = vectors of HB errors
X_B, X_H = vectors of state-variable (SV) harmonics (since the circuit is autonomous, one of the entries X is replaced by the fundamental frequency ω_0)
J_B, J_H = vectors of forcing terms

The subscripts *B* and *H* denote sidebands and carrier harmonics, respectively.

For a spot noise analysis at a frequency ω, the noise sources can be interpreted in either of two ways:

- Pseudosinusoids with random amplitude and phase located at the sidebands. This noise generation mechanism is described by Eq. (5-127), which is essentially a frequency conversion equation relating the sideband harmonics of the state variables and of the noise sources. This description is exactly equivalent to the one provided by the frequency conversion approach. This mechanism is referred to as *conversion noise* [19–21].
- Sinusoids located at the carrier harmonics and randomly phase and amplitude modulated by pseudosinusoidal noise at frequency ω. This noise generation mechanism is described by Eq. (5-128), which describes noise-induced jitter of the circuit-state, represented by the vector δX_H. The modulated perturbing signals are represented by replacing the entries of J_H with the complex modulation laws. This mechanism is referred to as *modulation noise.* One of the entries of δX_H is $\delta\omega_0$, where $\delta\omega_0(\omega)$ is the phasor of the pseudosinusoidal components of the fundamental frequency fluctuations in a 1-Hz band at frequency ω. Equation (5-128) provides a frequency jitter with a mean square value proportional to the available noise power. In the presence of both thermal and flicker noise, PM noise increases as ω^{-3} for $\omega \to 0$ and tends to 0 for $\omega \to \infty$. Modulation noise analysis correctly describes the noise behavior of an oscillator at low deviations from the carrier and does not provide results consistent with physical observations at high deviations from the carrier.

The combination of both phenomena explains the noise in the oscillator shown in Figure 5-85, where the near-carrier noise dominates below ω_X and far-carrier noise dominates above ω_X. Figure 5-86 (itemized form) shows the noise sources as they are applied at the IF. We have arbitrarily defined the low oscillator output as IF. This applies to the conversion matrix calculation.

Figure 5-87 shows the contributions that have to be taken into consideration for calculation of the noise at the output. The accuracy of the calculation of the phase noise depends highly on the quality of the parameter extraction for the nonlinear device; in particular, high-frequency phenomena must be modeled properly. In addition, the flicker noise contribution is essential. This is also valid for mixer noise analysis.

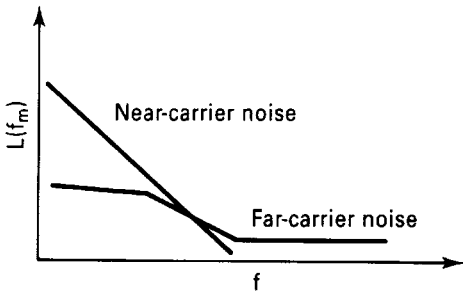


Figure 5-85 Oscillator noise components.

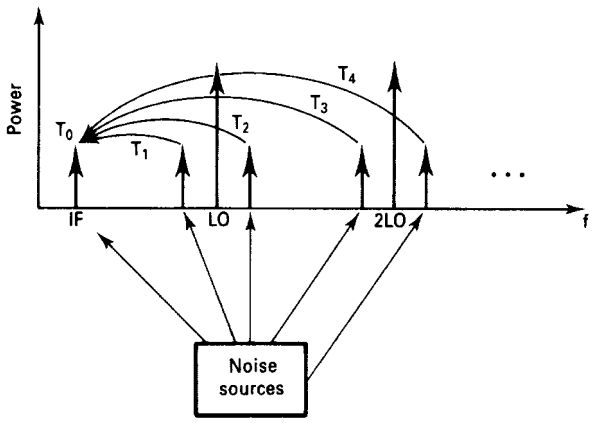


Figure 5-86 Noise sources where the noise at each sideband contributes to the output noise at the IF through frequency conversion.

Figure 5-88 shows the measured phase noise at 10 kHz of the carrier of an 800-MHz oscillator and the simulated result. The good agreement between measured and predicted data is worth noticing. At currents above 5 mA, the simulation becomes inaccurate as breakdown effects occur that were not part of the simulation.

Conversion Noise Analysis. The actual mathematics used to calculate the noise result (Ansoft Serenade 8.*x*) are as follows.

kth Harmonic PM Noise

$$\langle |\delta\Phi_k(\omega)|^2 \rangle = \frac{N_k(\omega) - N_{-k}(\omega) - 2\,\mathrm{Re}[C_k(\omega)]}{R|I_k^{SS}|^2} \tag{5-129}$$

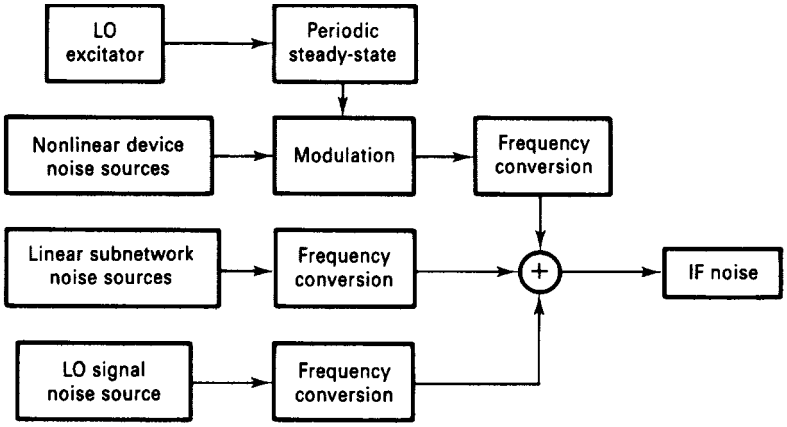


Figure 5-87 Noise mechanisms.

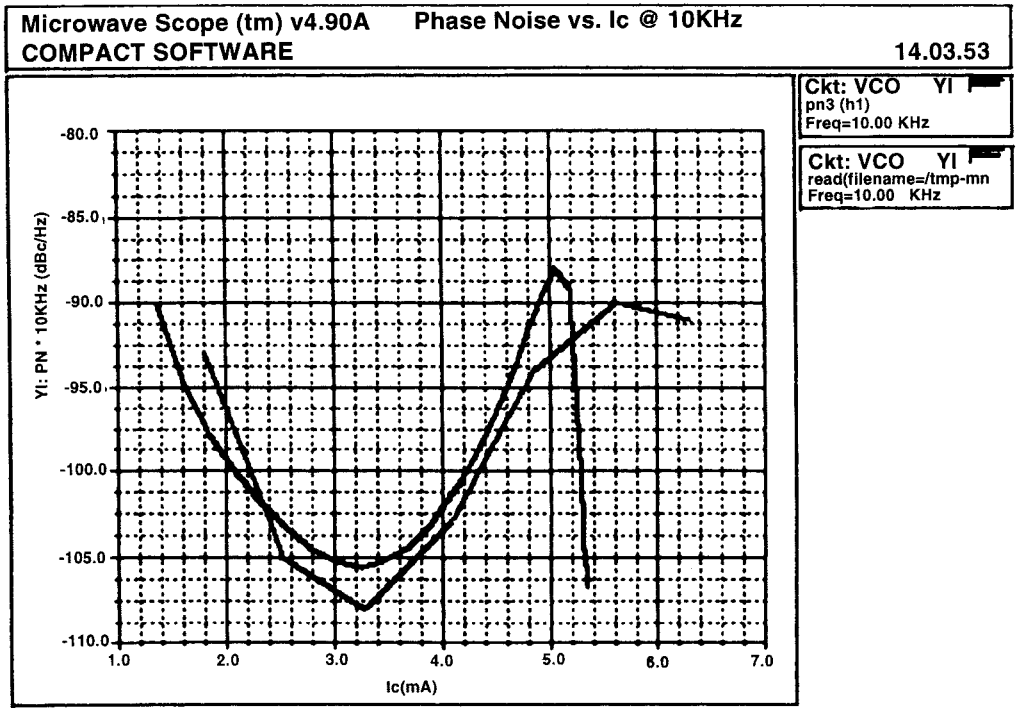


Figure 5-88 Bias-dependent measured and predicted phase noise at 10 kHz off the carrier of an 800-MHz oscillator.

kth Harmonic AM Noise

$$\langle |\delta A_k(\omega)|^2 \rangle = 2\,\frac{N_k(\omega) - N_{-k}(\omega) + 2\,\mathrm{Re}[C_k(\omega)]}{R|I_k^{SS}|^2} \tag{5-130}$$

kth Harmonic PM–AM Correlation Coefficient

$$C_k^{\mathrm{PMAM}}(\omega) = \langle \delta\Phi_k(\omega)\delta A_k(\omega)^* \rangle = -\sqrt{2}\,\frac{2\,\mathrm{Im}[C_k(\omega)] + j[N_k(\omega) - N_{-k}(\omega)]}{R|I_k^{SS|^2}} \tag{5-131}$$

where

$N_k(\omega)$, $N_{-k}(\omega)$ = noise power spectral densities at the upper and lower sidebands of the kth harmonic
$C_k(\omega)$ = normalized correlation coefficient of the upper and lower sidebands of the kth carrier harmonic
R = load resistance
I_k^{SS} = kth harmonic of the steady-state current through the load

Modulation Noise Analysis

kth Harmonic PM Noise

$$\langle |\delta\Phi_k(\omega)|^2 \rangle = \frac{k^2}{\omega^2}\,\mathrm{T}_F\,\langle \mathrm{J}_H(\omega)\mathrm{J}_H^t(\omega)\rangle \mathrm{T}_F^t \tag{5-132}$$

kth Harmonic AM Noise

$$\langle |\delta A_k(\omega)|^2 \rangle = \frac{2}{|I_k^{SS}|^2}\,\mathrm{T}_{Ak}\,\langle \mathrm{J}_H(\omega)\mathrm{J}_H^t(\omega)\rangle \mathrm{T}_{Ak}^t \tag{5-133}$$

kth Harmonic PM–AM Correlation Coefficient

$$C_k^{\mathrm{PMAM}}(\omega) = \langle \delta\Phi_k(\omega)\delta A_k(\omega)^* \rangle = \frac{k\sqrt{2}}{j\omega|I_K^{SS}|^2}\,\mathrm{T}_F\langle \mathrm{J}_H(\omega)\mathrm{J}_H^t(\omega)\rangle \mathrm{T}_{Ak}^t \tag{5-134}$$

where $\mathrm{J}_H(\omega)$ = vector of Norton equivalent of the noise sources
T_F = frequency transfer matrix
I_k^{SS} = kth harmonic of the steady-state current through the load

Experimental Variations. We will look at two examples where we compare predicted and measured data. Figure 5-89 shows the abbreviated circuit of a 10-MHz crystal oscillator. It uses a high-precision, high-Q crystal made by companies such as Bliley. Oscillators like this are intended for use as frequency and low-phase-noise standards. In this case, the circuit under consideration is part of the HP3048 phase-noise measurement system.

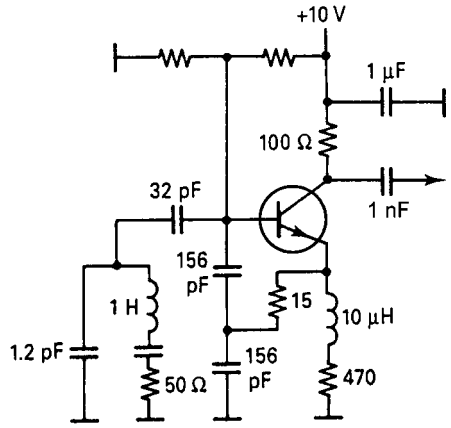


Figure 5-89 Abbreviated circuit of a 10-MHz crystal oscillator.

Figure 5-90 shows the measured phase noise of this HP frequency standard, and Figure 5-91 shows the phase noise predicted using the mathematical approach outlined above.

In cooperation with Motorola, we also analyzed an 800-MHz VCO. In this case, we also did the parameter extraction for the Motorola transistor. Figure 5-92 shows the circuit, a Colpitts oscillator that uses RF feedback in the form of a 15-Ω resistor and a capacitive

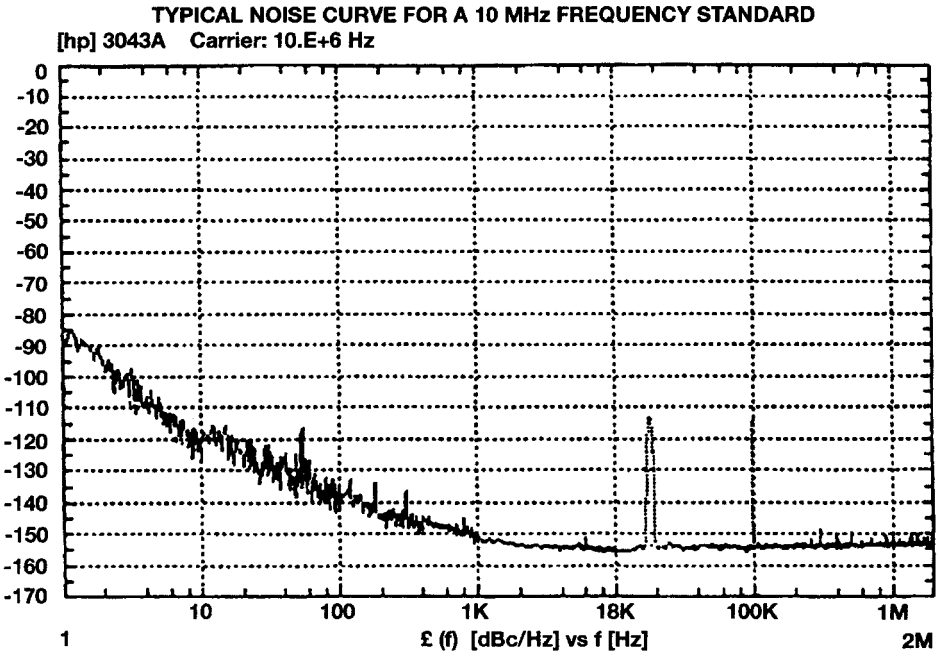


Figure 5-90 Measured phase noise for this frequency standard by HP.

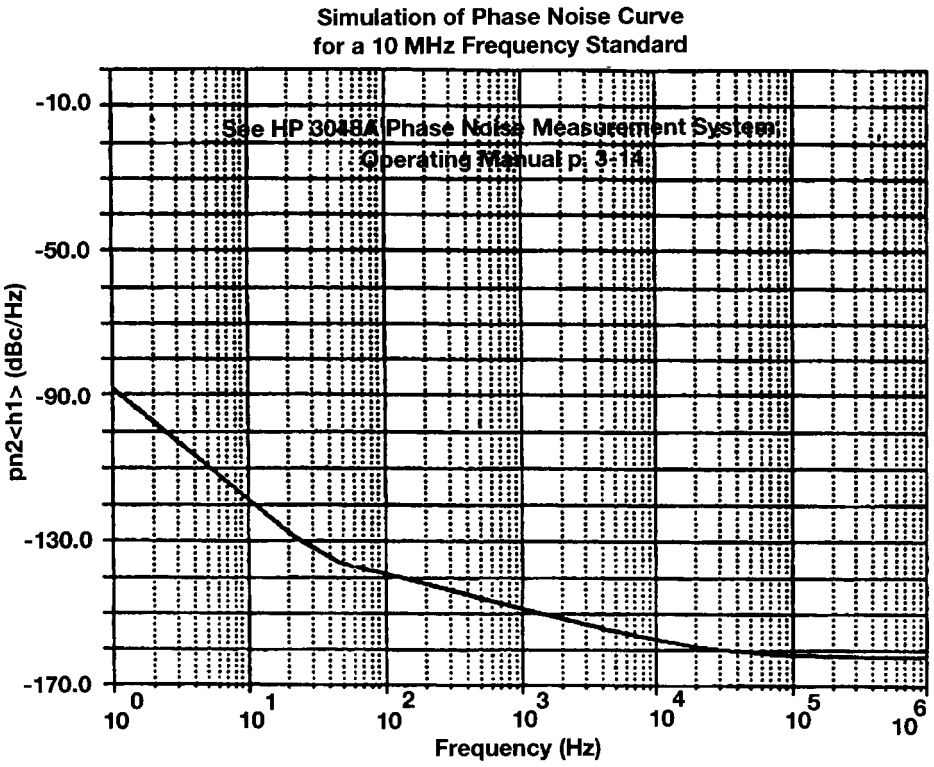


Figure 5-91 Simulated phase noise of the oscillator shown in Figure 5-89.

voltage divider consisting of 1 pF between the BJT's base and the feedback resistor, and 1 pF between the feedback resistor and common. Also, the tuned circuit is loosely coupled to this part of the transistor circuit. Figure 5-93 shows a comparison between predicted and measured phase noise for this oscillator. Figures 5-94, 5-95, and 5-96 show the oscillator's predicted output waveform, spectrum, and dc I–V and ac load line responses, respectively.

Phase Noise Optimization. By allowing the simulator's optimizer to vary the oscillator's feedback and dc operating point, phase noise can be improved. The values that were allowed to vary were the values of the capacitors in the feedback voltage divider, the value of a negative-feedback resistor (starting value, 15 Ω) that reduces AM–PM conversion, and the emitter resistor, which changes the transistor's dc bias and therefore affects its bias-dependent flicker noise. Figure 5-97 shows the improvement of the 800-MHz VCO as previously shown. While the close-in phase noise can be improved drastically by approximately 32 dB, the far-out noise at 20 MHz and beyond is deteriorated. This is due to the resistive feedback, which reduces AM-to-PM conversion [22–24].

In testing an oscillator's resistance to frequency pushing (frequency shift with supply-voltage changes), variations in phase noise can be observed. Phase noise is also bias dependent,

and this fact can be used for phase-noise optimization. Figure 5-98 shows the phase noise of an oscillator as its supply voltage is varied.

Summary. This section has demonstrated the power of harmonic-balance simulation in the design of low-phase-noise oscillators. The simulator can be used to accurately predict close-in and far-out phase noise based on the contribution of the transistor, the loaded Q of the resonator, and, in the case of a VCO, the contribution of the tuning diode.

This is the first time that a bias-dependent noise model for both FETs and bipolar transistors has been developed and used for this type of application, in which optimization is allowed to vary circuit components and the bias point for best performance.

Practical examples ranging from a 10-MHz crystal oscillator to a 1-GHz CR oscillator have been supplied. The theoretical background for the nonlinear oscillator and its validation have been presented.

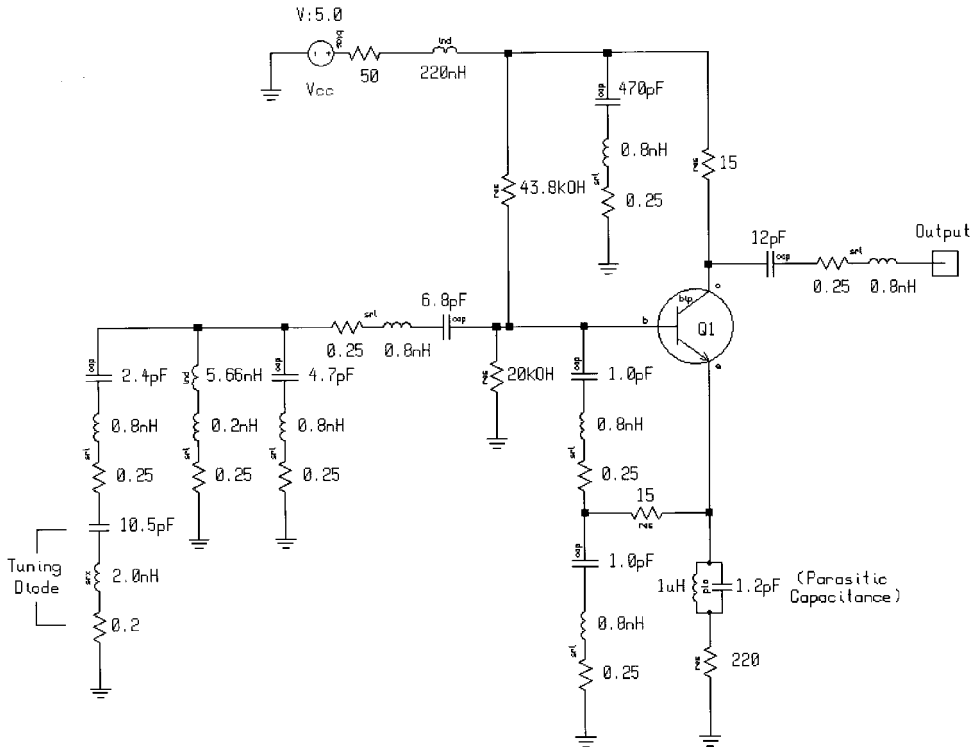


Figure 5-92 Colpitts oscillator that uses RF negative feedback between the emitter and capacitive voltage divider. To be realistic, we used actual components rather than ideal ones. The suppliers for the capacitors and inductors provide some typical values for the parasitics. The major changes are 0.8 nH and 0.25 Ω in series with the capacitors. The same thing applies for the main inductance, which has a parasitic connection inductance of 0.2 nH in series with a 0.25-Ω resistance. These types of parasitics are valid for a fairly large range of components assembled in surface-mount applications. Most engineers model the circuit only by assuming lossy devices, not adding these important parasitics. One of the side-effects we have noticed is that the output power is more realistic and, needless to say, the simulated phase noise agrees quite well with measured data. This circuit can also serve as an example for modeling amplifiers and mixers using surface-mount components.

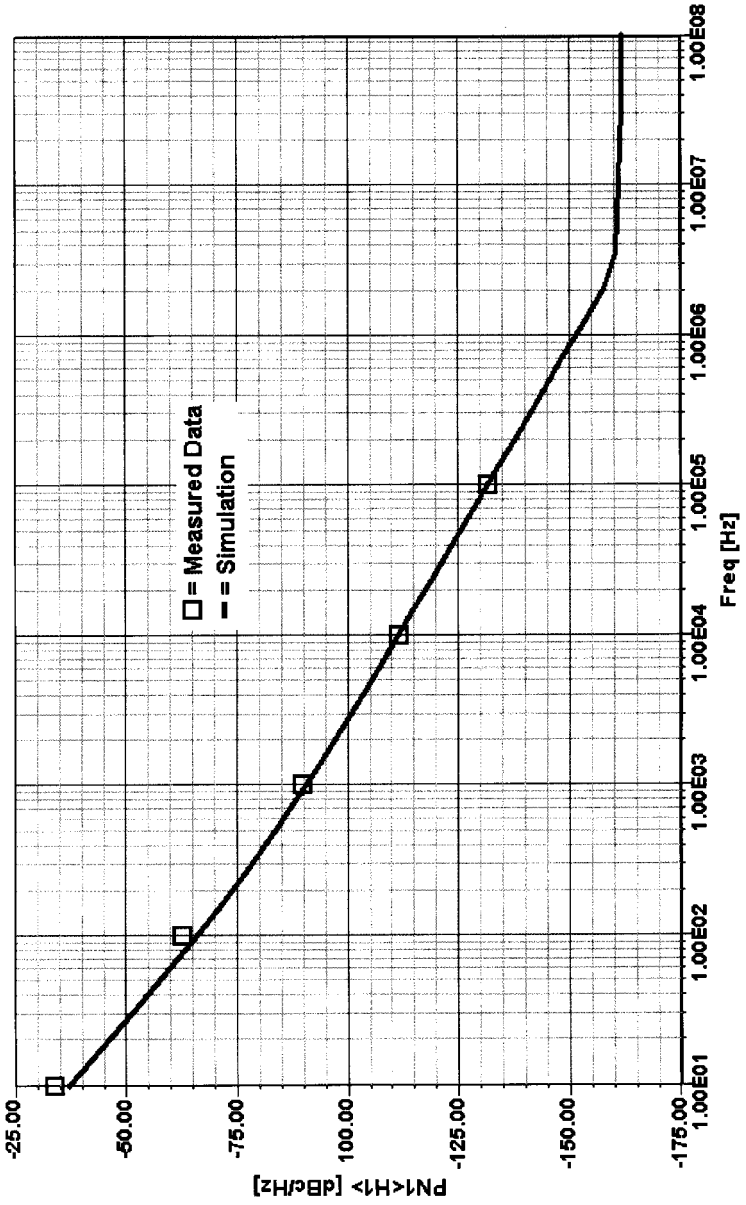


Figure 5-93 Comparison between predicted and measured phase noise for the oscillator shown in Figure 5-92.

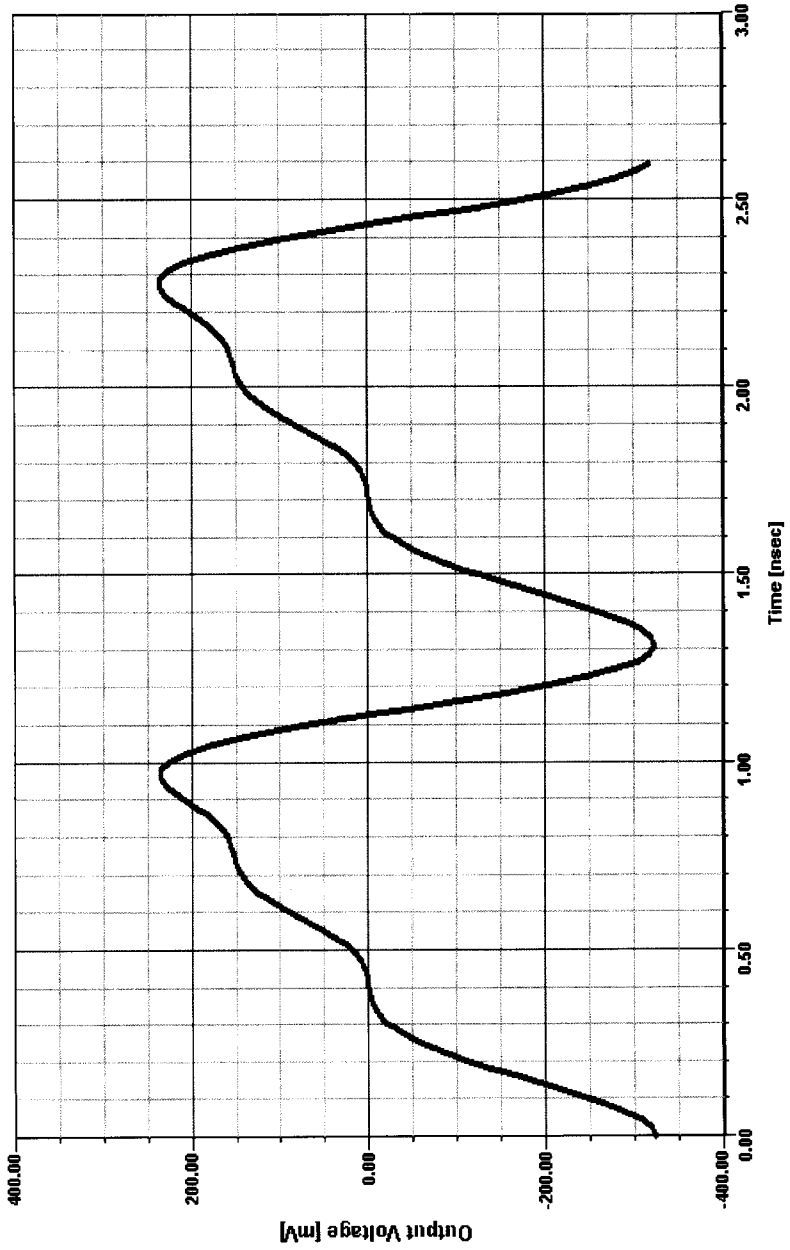


Figure 5-94 Predicted output waveform for the oscillator shown in Figure 5-92.

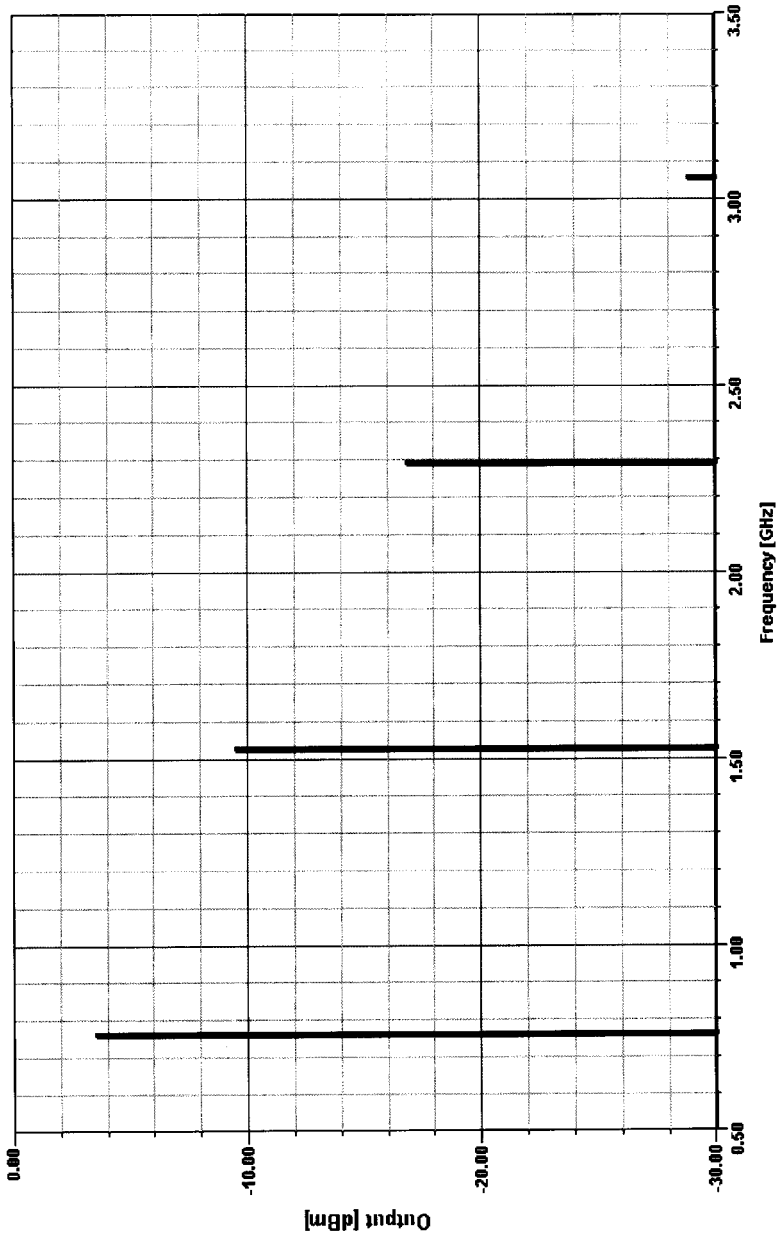


Figure 5-95 Predicted output spectrum for the oscillator shown in Figure 5-92.

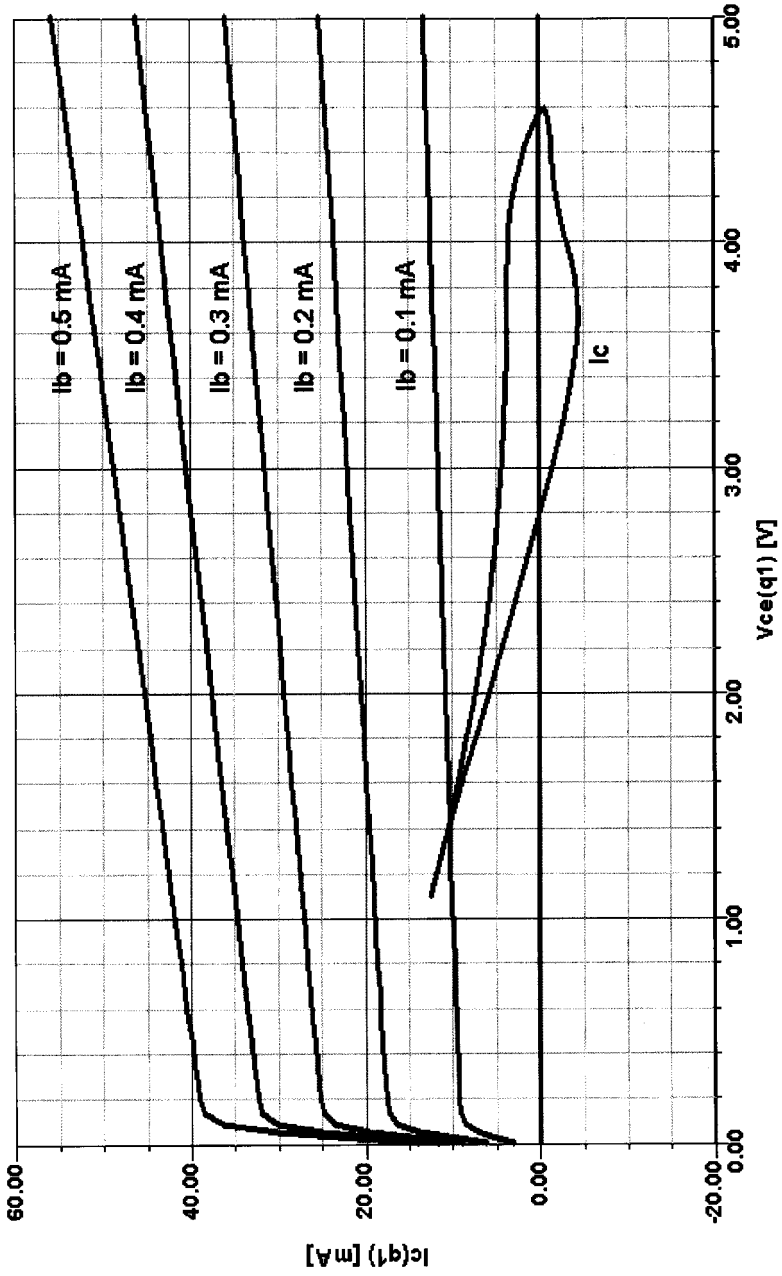


Figure 5-96 The ac load line and dc *I–V* curves for the transistor in the oscillator shown in Figure 5-92.

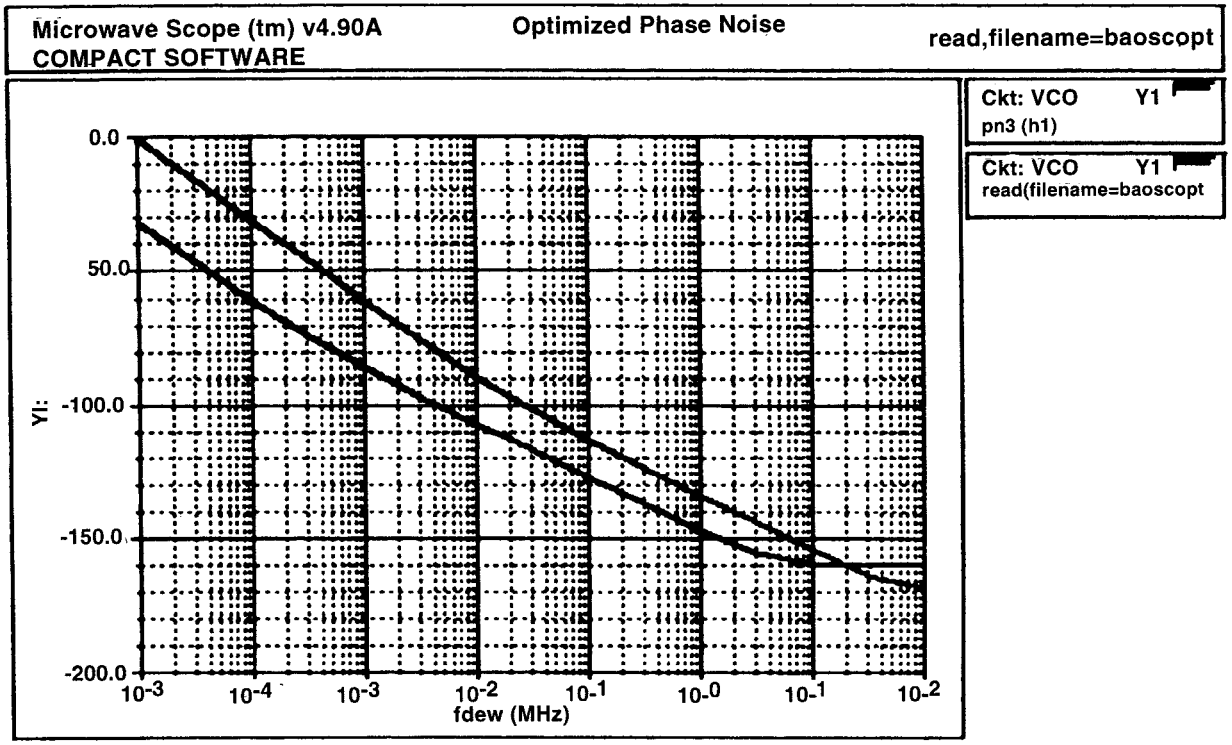


Figure 5-97 Phase noise of the 800-MHz VCO before and after optimization.

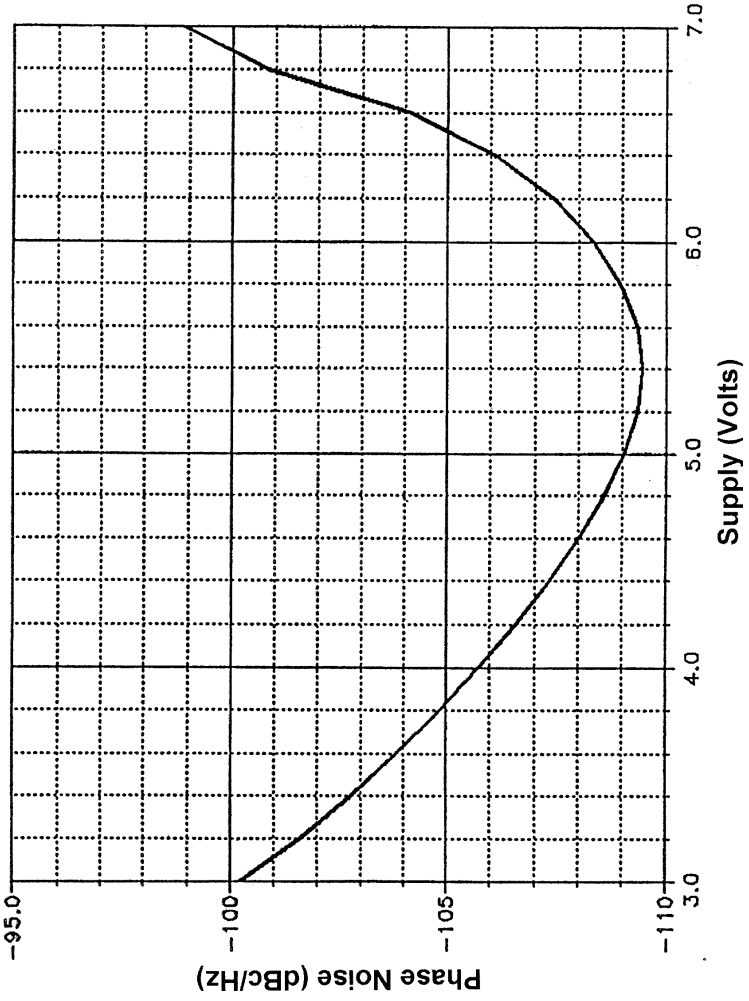


Figure 5-98 Calculated phase noise as a function of the supply voltage for a BJT oscillator. The graph shows a distinct minimum for a particular bias, which could be translated into a collector current change. Such phase-noise minimization could also be accomplished by changing just the voltage applied to the bias portion of the oscillator.

Acknowledgment. This research work for nonlinear operation in oscillators, partly sponsored by the U.S. Department of Defense MIMIC program, was performed in cooperation with Professor Vittorio Rizzoli's group of the Department of Electrical Engineering, Information and Systems Engineering at the University of Bologna, Italy, and the Motorola Portable Products Division, Florida, as well as the engineering team of Ansoft Corporation in New Jersey.

5-7 OSCILLATORS IN PRACTICE

5-7-1 Oscillator Specifications

Although we have tended to emphasize phase noise as a critical indicator of oscillator quality, phase noise is only one of a number of oscillator performance criteria that must be considered in wireless system design. In addition to phase noise, the following characteristics are used by commercial companies to describe oscillator performance:

Frequency Pushing. Frequency pushing characterizes the degree to which an oscillator's frequency is affected by its supply voltage. A sudden current surge caused by activating a transceiver's RF power amplifier may produce a spike on the VCO's dc power supply and a consequent frequency jump. Like tuning sensitivity (below), pushing is specified in frequency/voltage form and is tested by varying the VCO's dc supply voltage (typically ± 1 V) with its tuning voltage held constant.

Harmonic Output Power. The harmonic content is measured relative to the output power. Typical values are 20 dB or more suppression relative to the fundamental. This suppression can be improved by additional filtering.

Output Power. The output power of the oscillator, typically expressed in dBm, is measured into a 50-Ω load. The output power is always combined with a specification for flatness or variation. A typical specification would be 0 dBm ± 1 dB.

Output Power as a Function of Temperature. All active circuits vary in performance as a function of temperature. An oscillator's output power over a temperature range should vary less than a specified value, such as 1 dB.

Post-tuning Drift. After a voltage step is applied to the tuning diode input, the oscillator frequency may continue to change until it settles to a final value. This post-tuning drift is one of the parameters that limits the bandwidth of the VCO input.

Power Consumption. This characteristic conveys the dc power, usually specified in milliwatts and sometimes qualified by operating voltage, required by the oscillator.

Sensitivity to Load Changes. To keep costs down, many wireless applications use a VCO alone, without the buffering action of a high reverse-isolation amplifier stage. In such applications, *frequency pulling*, the change of frequency resulting from partially reactive loads, is an important oscillator characteristic. Pulling is commonly specified in terms of the frequency shift that occurs when the oscillator is connected to a load that exhibits a non-unity VSWR (such as 1.75, usually referenced to 50 Ω), compared to the frequency that results with a unity VSWR load (usually 50 Ω). Frequency pulling must be minimized, especially in cases where power stages are close to the VCO unit

and short pulses may affect the output frequency. Such feedback may make phase locking impossible.

Spurious Outputs. A VCO's spurious output specification, expressed in decibels, enumerates the strength of unwanted and nonharmonically related components relative to the oscillator fundamental. Because a stable, properly designed oscillator is inherently clean, such *spurs* are typically introduced only by external sources in the form of radiated or conducted interference. See *harmonic output power.*

Temperature Drift. Although the synthesizer is responsible for locking and maintaining the oscillator's frequency, the VCO's frequency change as a function of temperature is a critical parameter and must be specified. Its value varies between 10 kHz/°C to several hundred kHz/°C depending on the center frequency and tuning range.

Tuning Characteristic. This specification shows the relationship, depicted as a graph, between a VCO's operating frequency and the tuning voltage applied. Ideally, the correspondence between operating frequency and tuning voltage is linear. See *tuning linearity.*

Tuning Linearity. For stable synthesizers, a constant deviation of frequency versus tuning voltage is desirable. It is also important to make sure that there are no breaks in the tuning range—for example, that the oscillator does not stop operating with a tuning voltage of 0 V. See *tuning characteristic.*

Tuning Sensitivity, Tuning Performance. This characteristic, typically expressed in megahertz per volt (MHz/V), enumerates how much a VCO's frequency changes per unit of tuning-voltage change.

Tuning Speed. This characteristic is defined as the time necessary for the VCO to reach 90% of its final frequency on the application of a tuning-voltage step. Tuning speed depends on the internal components between the input pin and tuning diode—including, among other things, the capacitance present at the input port. The input port's parasitic elements determine the VCO's maximum possible modulation bandwidth.

5-7-2 More Practical Circuits

Because of the need for integration, we will present three IC oscillators of increasing complexity. The circuit in Figure 5-99 is found in early Siemens ICs. Figure 5-100 shows its phase-noise performance.

Figure 5-101 shows the basic push–pull oscillator topology underlying the circuit in Figure 5-99. The omission of the current source and active bias components improves the simpler circuit's phase-noise performance over that of the circuit shown in Figure 5-99. Figure 5-102 compares the phase noise of this basic circuit to the integrated version in Figure 5-99.

Figure 5-103 is for those who speculate on the possibility of replacing the BJTs in Figure 5-101 with LDMOS FETs or JFETs. Figure 5-104 compares its phase noise to that of the BJT case. The rule of thumb is that FETs do not have enough gain for high-Q oscillation in oscillators above 400 MHz, and bipolar transistors are a better choice. Because of their higher flicker noise contribution, GaAsFETs should be considered only at frequencies above 4 or 5 GHz.

Figure 5-105 shows a circuit based on two differential amplifiers and frequently found in Gilbert cell IC applications. Figure 5-106 shows its phase-noise performance.

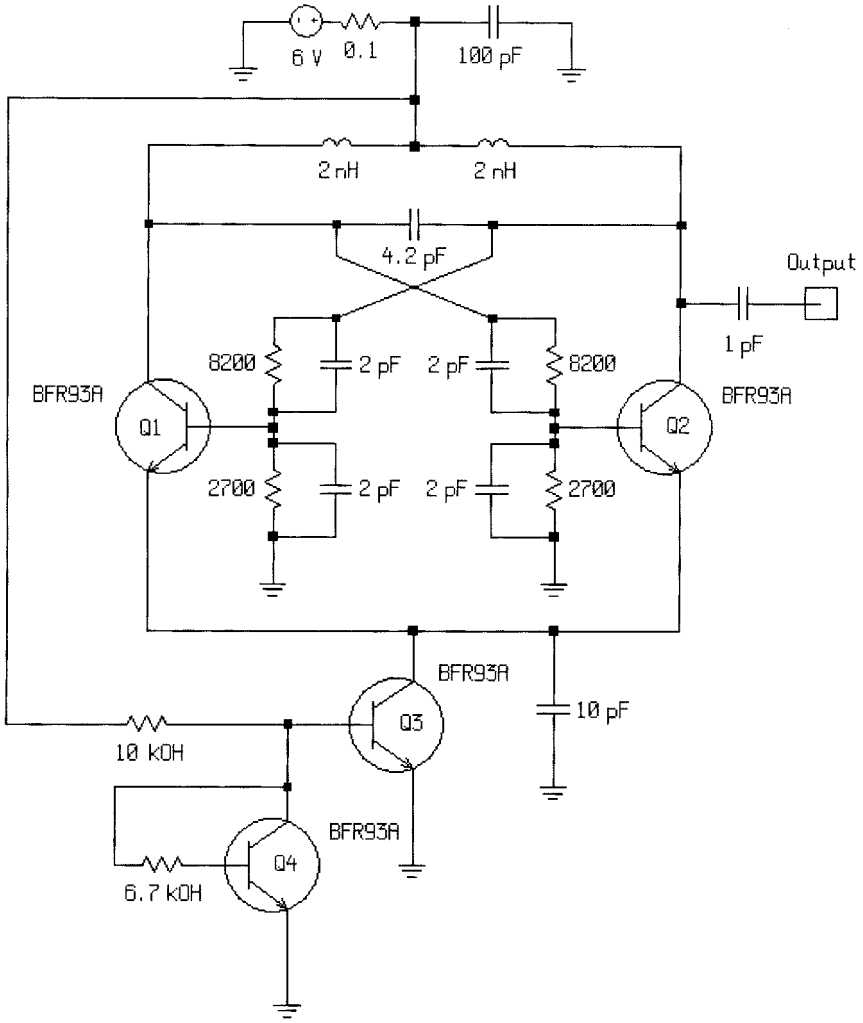


Figure 5-99 Schematic of the Siemens IC oscillator.

Figure 5-107 is a validation circuit that demonstrates a comparison of measured phase-noise performance versus its prediction by circuit simulation software. The circuit is based on the core of Motorola's MC12148 ECL oscillator IC, the topology of which is derived from the superseded MC1648. Motorola specifies the MC12148's typical phase-noise performance as –90 dBc/Hz at an offset of 25 kHz. The close agreement between the measured and simulated results highlights the value of state-of-the-art CAD tools in wireless oscillator design (Figure 5-108).

Figure 5-108 presents the results of *two* simulations: one with ideal L and C components, and another with realistic values specified for these reactances. In using a circuit

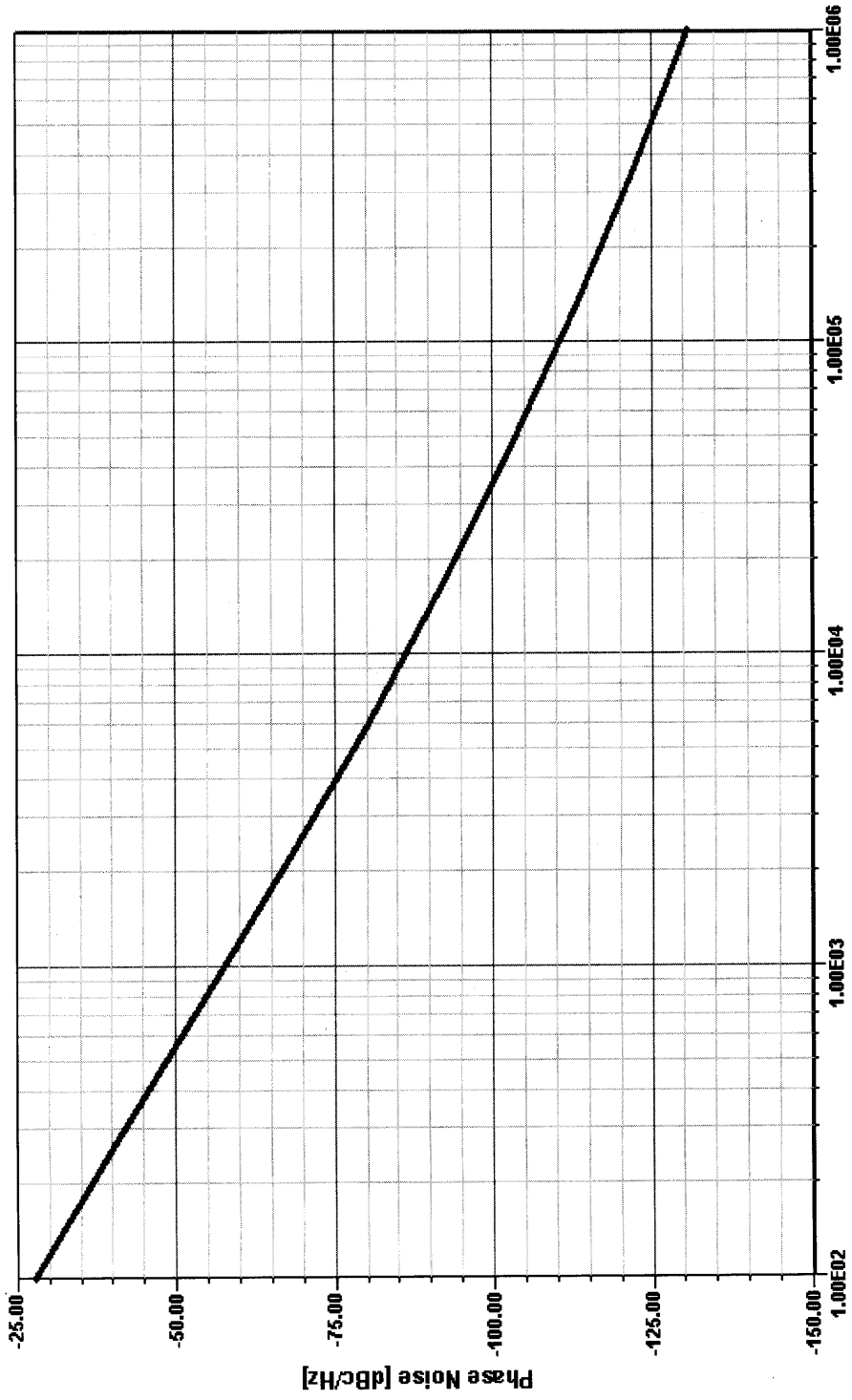


Figure 5-100 Phase noise of the Siemens IC oscillator. The frequency of oscillation is 1.051 GHz.

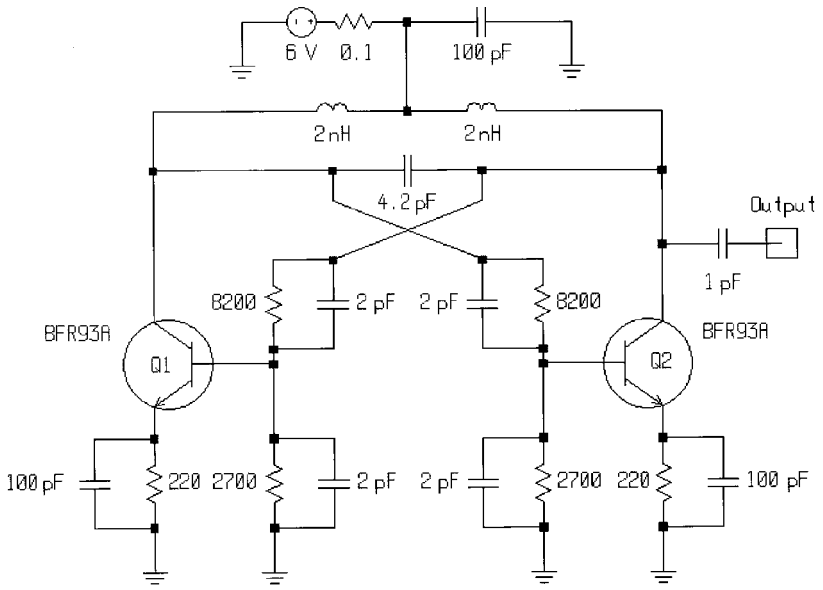


Figure 5-101 Basic push–pull oscillator underlying the circuit shown in Figure 5-99.

simulator to validate and compare oscillator topologies, it's important to start with ideal components because we are interested in comparing the inherent phase-noise performance of various circuits—particularly the role played by the nonlinear reactances in active devices. Because the Q values of practical components are so low in the microwave range (very generally, 200–300 for capacitors and 60–70 for inductors), the inherent merits of one topology, or one device, over another would be masked by loading effects if we conducted our investigations using nonideal reactances. Once initial evaluation is completed, realistic Q values should be specified for increased accuracy in phase-noise, output-power, and output-spectrum simulations.

Finally, as evidence of how moving from lumped to distributed techniques can improve oscillator performance at frequencies where *LC* tanks become problematic, Figure 5-109 compares, for a simulated BJT Colpitts oscillator operating at 2.3 GHz, the difference in phase-noise performance obtainable with a resonator consisting of an ideal 2-nH inductor and a $\frac{1}{4}\lambda$ transmission line (11 Ω, 90° long at 2.6 GHz, attenuation 0.1 dB/m) with the transistor biased by constant-current and constant-voltage sources.

Silicon/GaAs-Based Integrated VCOs and Possible Difficulties. Due to process variations, the center frequency of the VCO cannot be repeated without trimming. Trimming can be avoided by making the VCO gain large enough to compensate for tank circuit variations, but this degrades phase-noise performance and increases the VCO susceptibility to unwanted spurs. For example, a VCO with a gain of 10 MHz/V has a phase noise of –80 dBc/Hz at 10 kHz, while a VCO with a gain of 75 MHz/V has a phase noise of only –60 dBc/Hz at 10 kHz. All but the tank inductors are integrated.

Tank inductors are generally not integrated for reasons of size and Q. The area of a 1-nH inductor is approximately 42,000 mm^2, which is ~3.5% of a typical IC; the area of a 10-nH

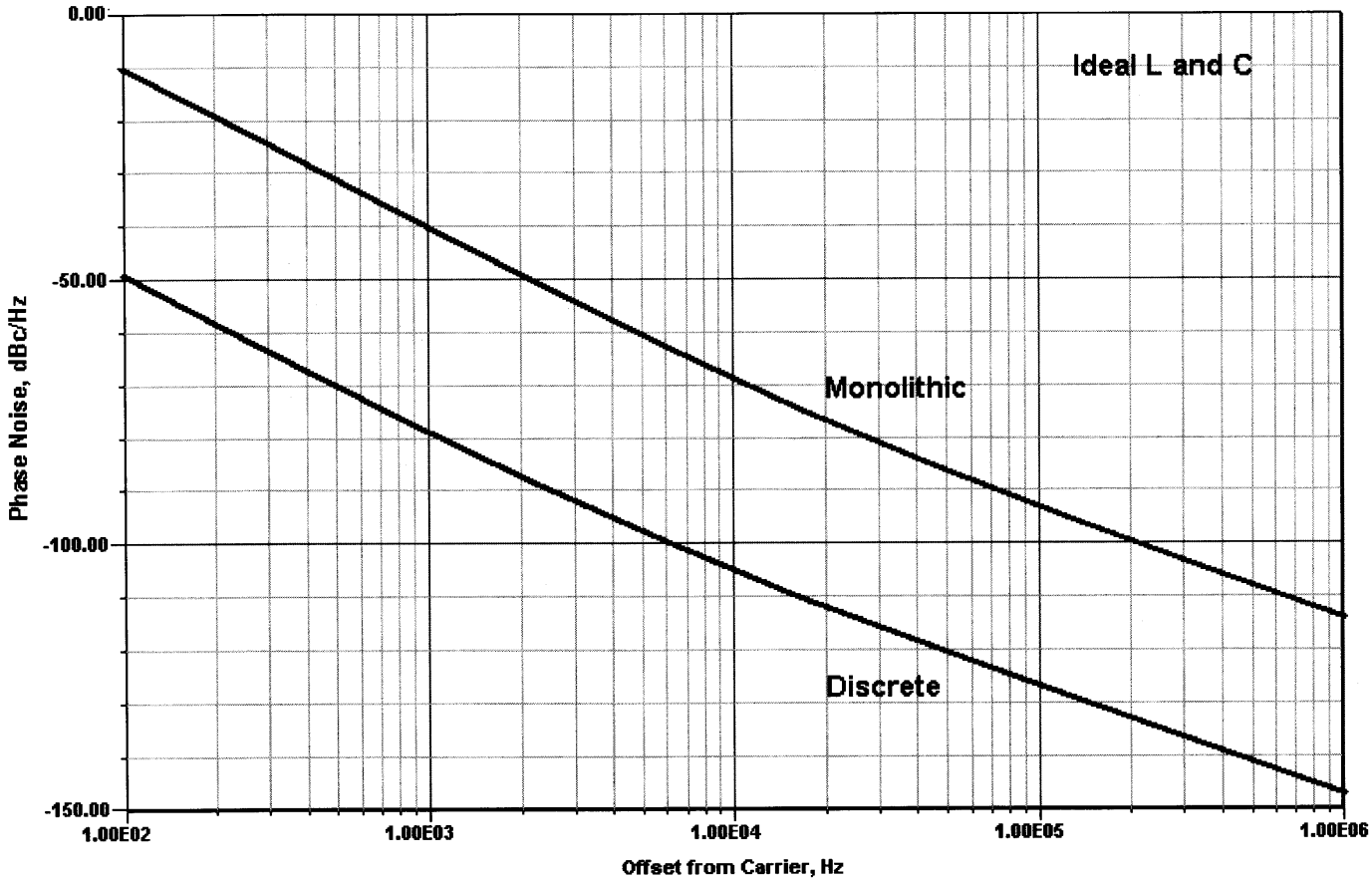


Figure 5-102 The basic version of the oscillator (Discrete curve) is a better phase-noise performer than the integrated version (Monolithic curve). The frequency of oscillation is 1.039 GHz.

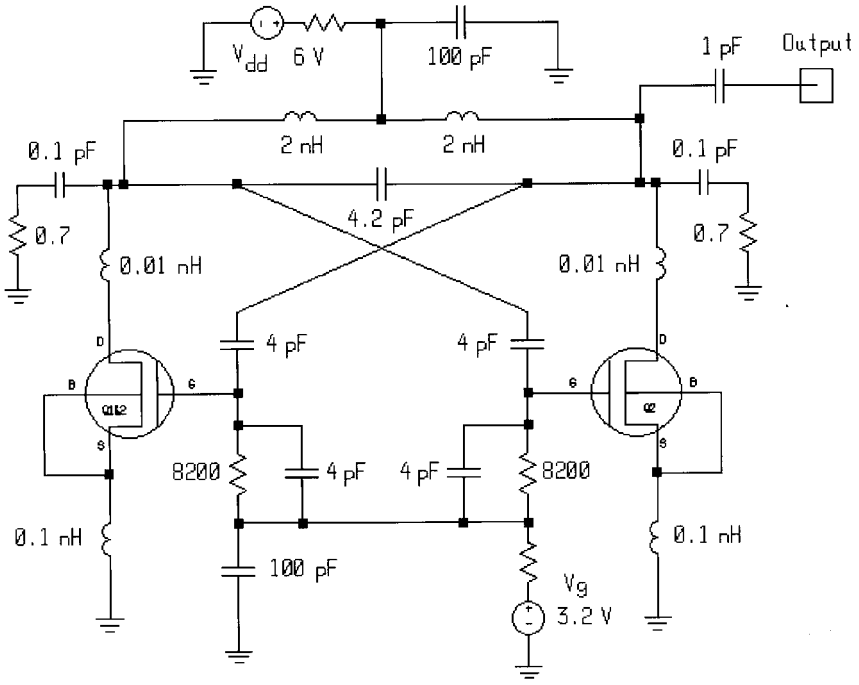


Figure 5-103 A 1.04-GHz push–pull oscillator using LDMOS FETs instead of BJTs. (LDMOS FET parameters courtesy of David K. Lovelace, Motorola, Inc.)

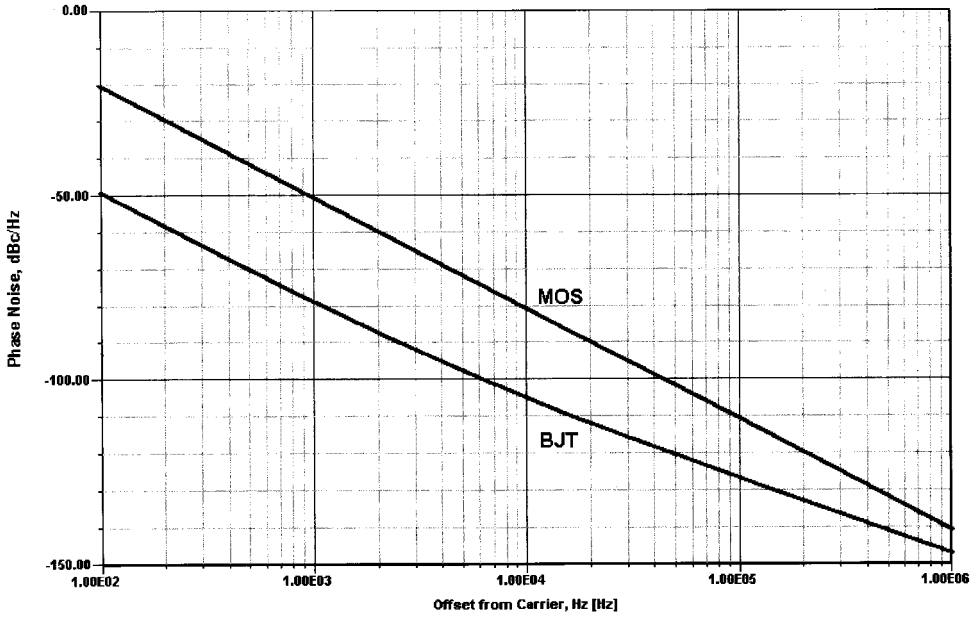


Figure 5-104 Phase-noise comparison of the BJT and MOSFET oscillators.

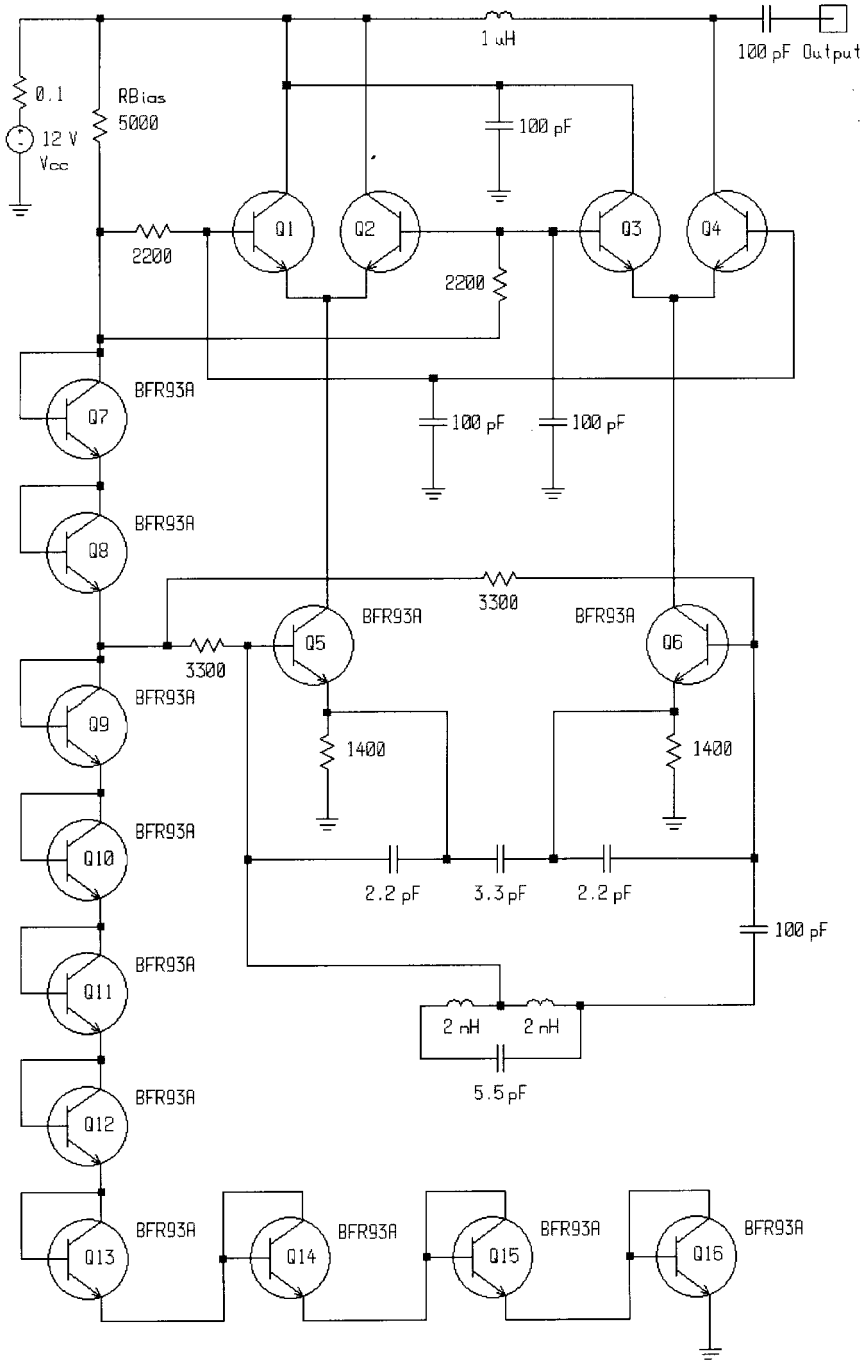


Figure 5-105 A 1.022-GHz IC oscillator based on two differential amplifiers. With an RF signal fed to the bases of Q_1–Q_4 and Q_2–Q_3 in push–pull, and IF output extracted from the collectors of Q_1–Q_3 and Q_2–Q_4 in push–pull, the circuit would act as a Gilbert cell converter. In the circuit shown, we have modified the differential amplifiers' collector circuitry to give single-ended output and avoid cancellation of the oscillator signal.

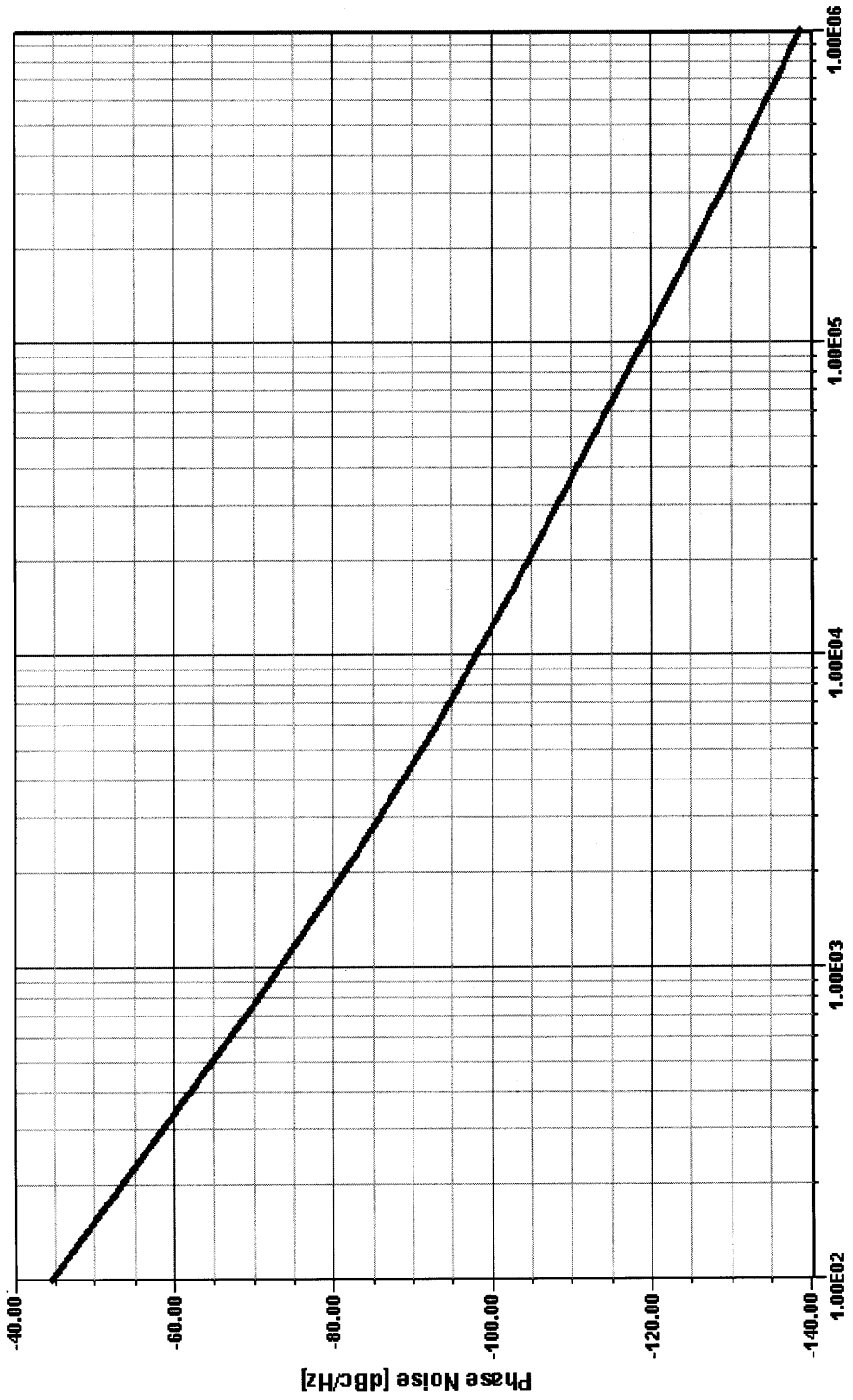


Figure 5-106 Phase noise of the two-differential-amplifier oscillator circuit.

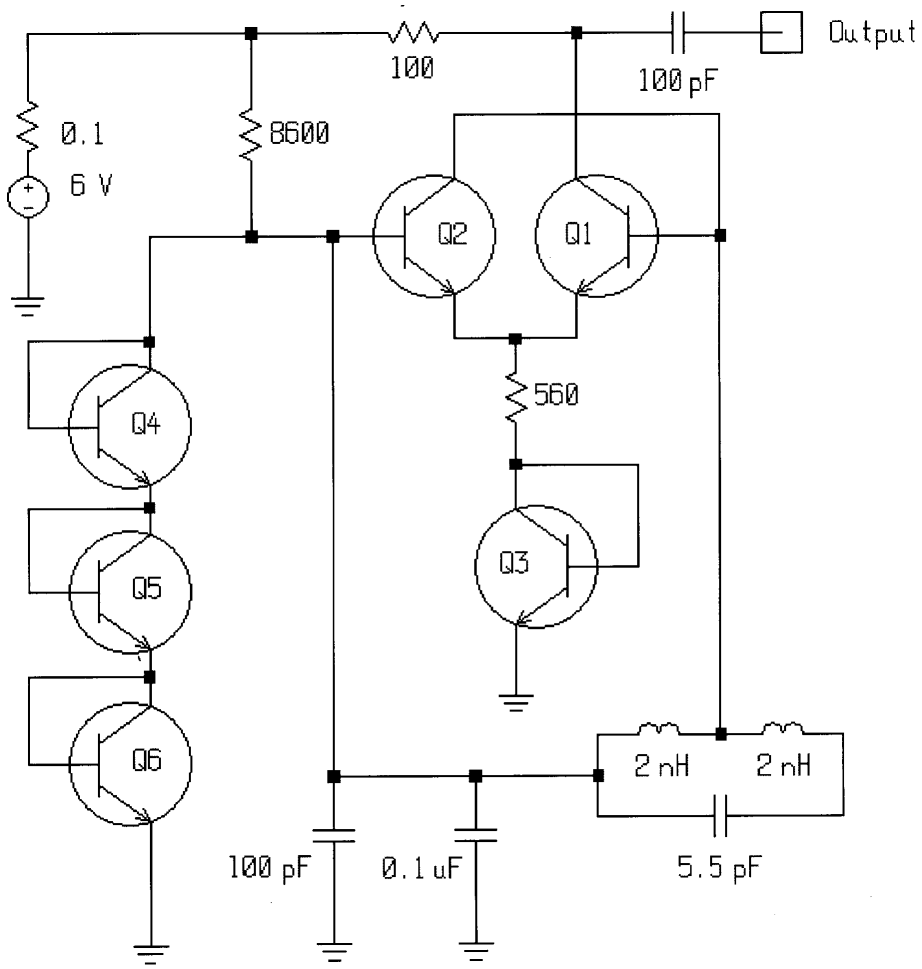


Figure 5-107 This validation circuit models the oscillator topology at the core of Motorola's MC12148 ECL oscillator IC.

inductor is approximately 167,000 mm^2, which is ~14% of a typical IC. For GaAs substrate, the resulting Q would be about 5; for silicon, as we already saw, 20 is already a very high value.

The customer can use PCB traces for the required inductance instead of external surface-mount inductors (cost: $0.03 each, most expensive passive component!). It must be pointed out that even for a nonintegrated approach, the higher VCO gain is noisier by definition. This was explained in Section 5-6-2. These problems are not unique to silicon; they apply to GaAs also. Figure 5-110 shows the size relationship of on-chip capacitors and inductors compared to the actual transistor; therefore, on-chip inductors should really be avoided for reasons of space and Q. Just imagine how the picture would change if the 1.1-nH inductor had to be 11 nH instead! It would dominate the IC area, and therefore the IC cost as well.

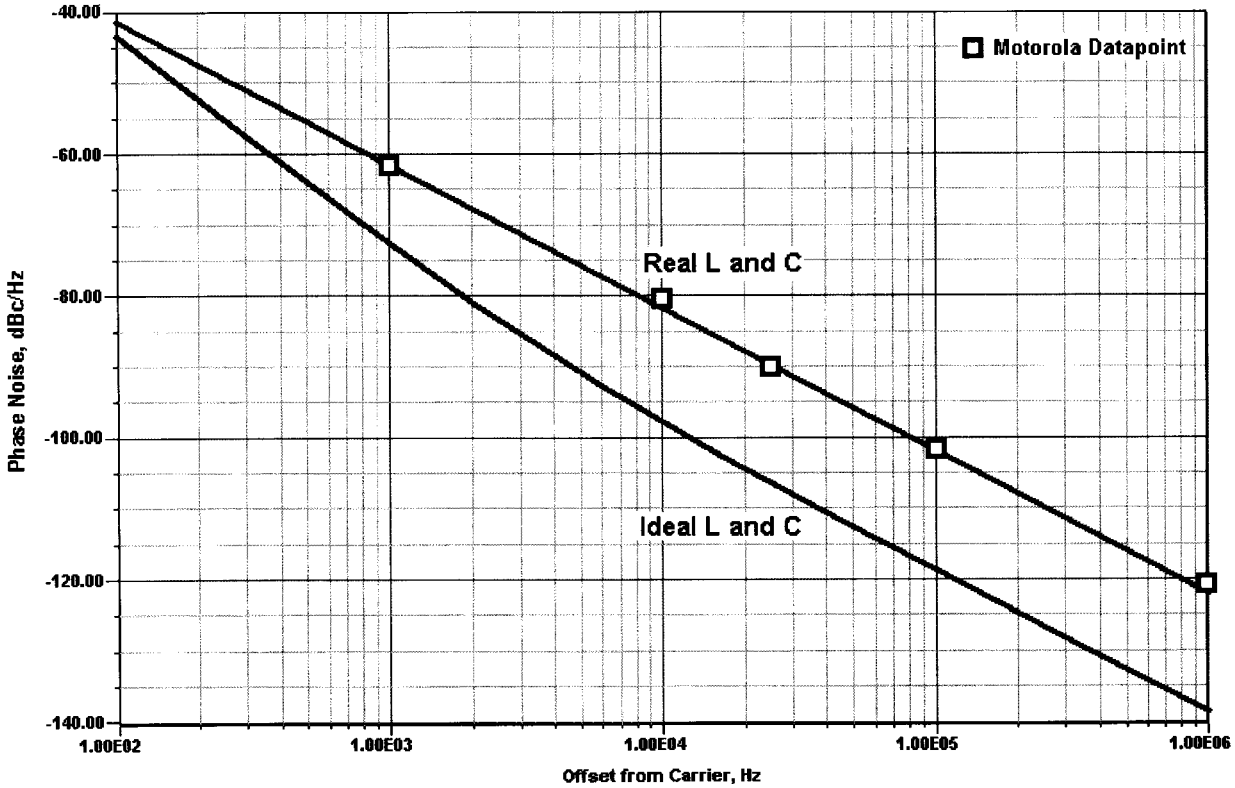


Figure 5-108 Comparison of measured versus simulated phase noise of the MC12148 differential oscillator. The simulation (1.016 GHz) predicts a phase noise of –89.54 dBc/Hz at an offset of 10 kHz (Real L and C curve), which agrees well with Motorola's specification of –89.54 dBc/Hz (typical at 930 MHz) at this offset. The Ideal L and C curve plots the simulation's phase-noise performance with ideal reactances, as opposed to the Real L and C curve, which graphs the results with realistic Q values specified for its reactive components ($Q = 250$ for its C values and $Q = 65$ for its L, all at 1 GHz).

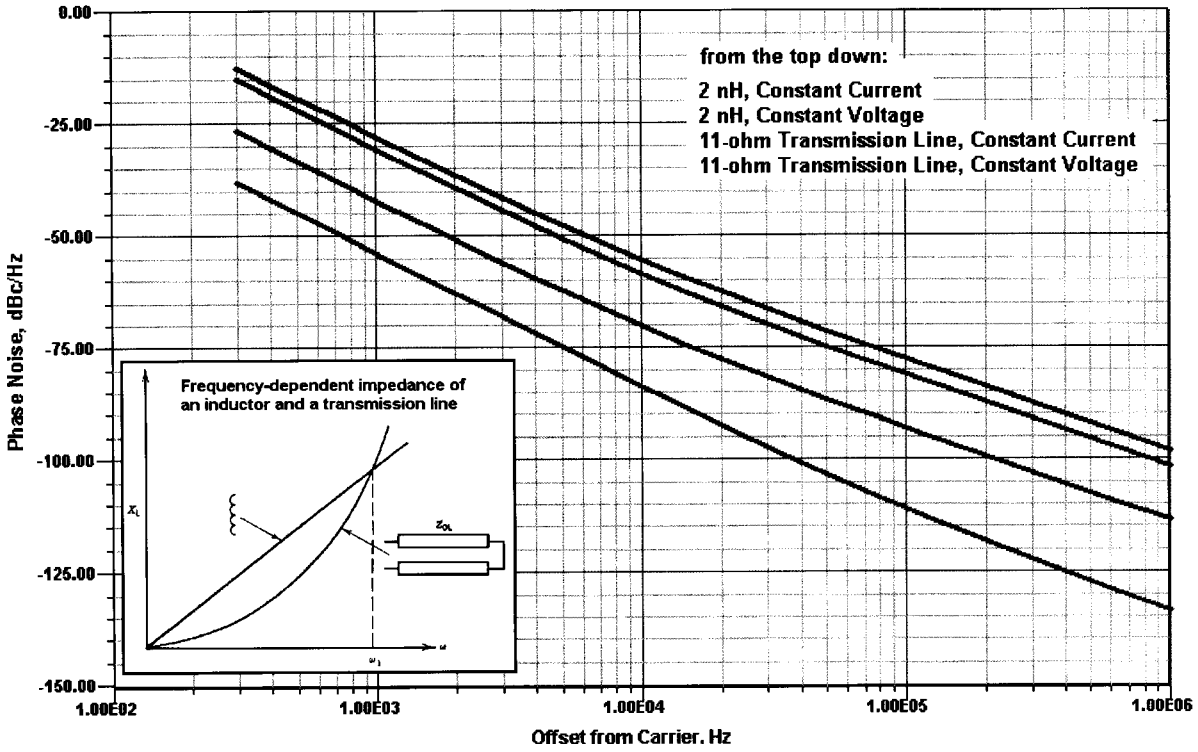


Figure 5-109 Phase-noise performance of a 2.3-GHz BJT oscillator with a resonator consisting of an inductor (2 nH) and a 1/4 λ transmission line (11 Ω, approximating the behavior of a dielectric resonator) with bias from a constant-current source and a low-impedance, resistive constant-voltage source.

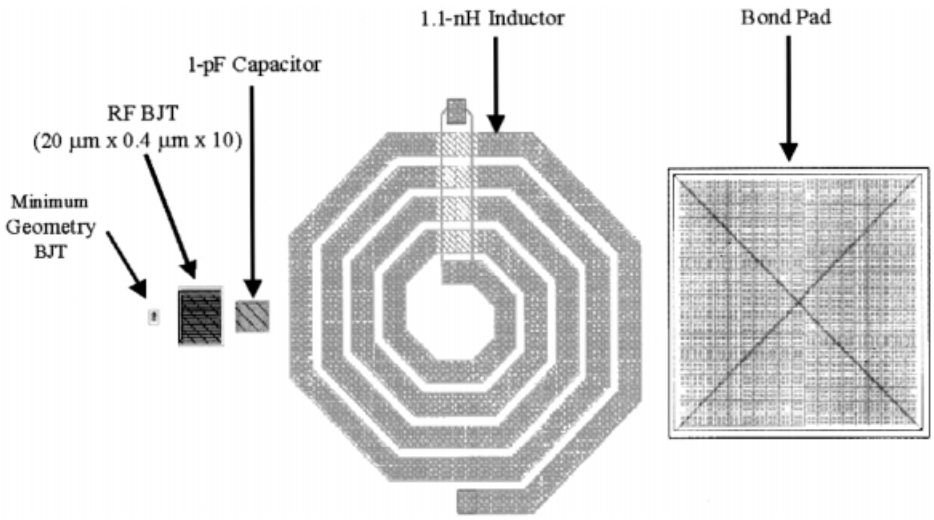


Figure 5-110 Same-scale comparison of space requirements for transistors, capacitors, inductors, and bond pads. For reference purposes, the coil diameter is 0.2 mm.

5-8 DESIGN OF RF OSCILLATORS USING CAD

5-8-1 Harmonic-Balance Simulation

HB analysis performs steady-state analysis of periodically excited circuits. The circuit to be analyzed is split into linear and nonlinear subcircuits. The linear subcircuit is calculated in the frequency domain. Features of this aspect of the HB process include:

- Use of distributed models in the spectral domain
- Matrix formulation that can enable reduction of internal nodes
- Major speed advantage
- Straightforward intermodulation and mixer analysis

The nonlinear subcircuit is calculated in the time domain. Features of this aspect of the HB process include:

- Nonlinear models derived directly from device physics
- Intuitive, easy, and logical circuit representation

Figure 5-111 diagrams this approach for a MESFET amplifier. Figure 5-112 charts a general-purpose nonlinear design algorithm that includes optimization. Modern analysis tools that must provide accurate phase-noise calculation should be based on the principle of harmonic balance. Figure 5-113 shows a BJT microwave oscillator entered into the schematic-capture module of a commercially available HB simulator (Ansoft Serenade 8.0); Figure 5-114 shows this oscillator's simulated phase noise. By the way, HB analysis can also handle mixers.

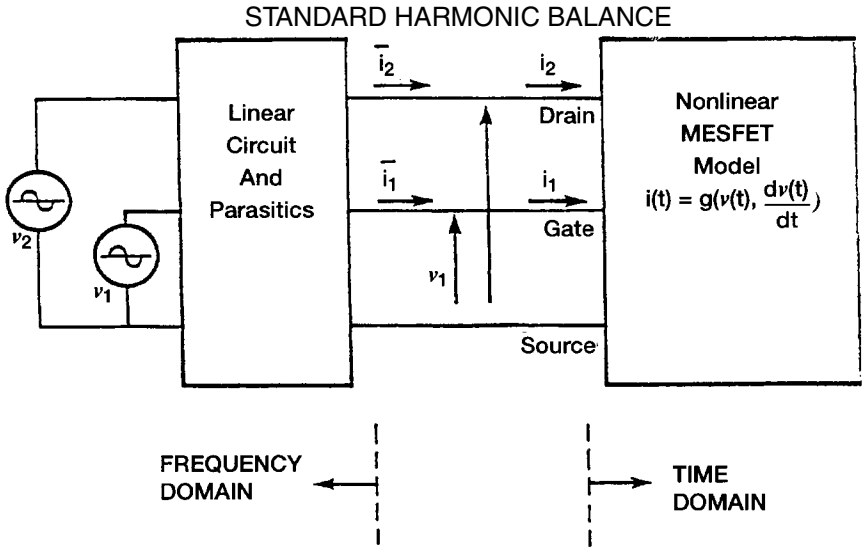


Figure 5-111 MESFET circuit partitioned into linear and nonlinear subcircuits for harmonic-balance analysis. Applied gate and drain voltages, and relevant terminal voltages and currents, are indicated.

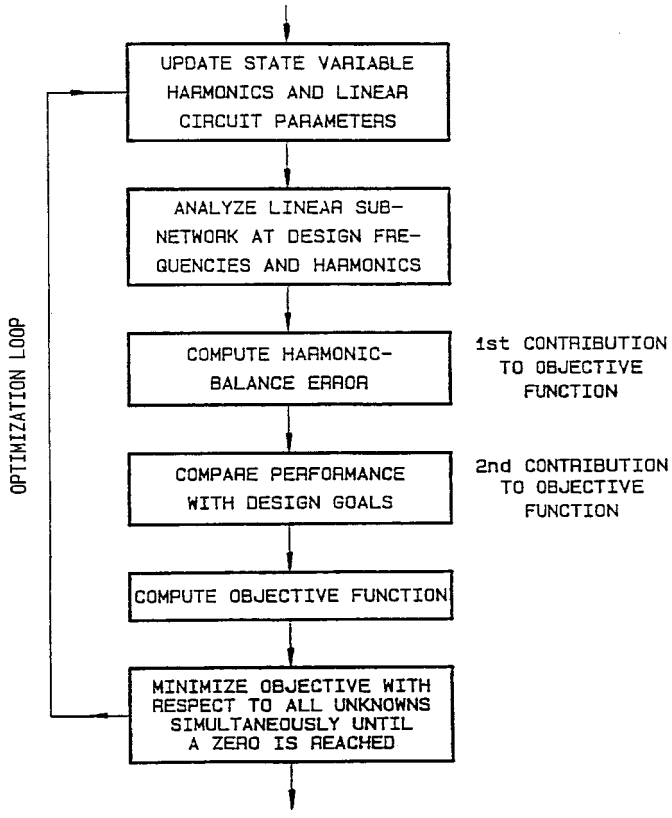


Figure 5-112 Flowchart of a general-purpose harmonic-balance design algorithm that includes optimization.

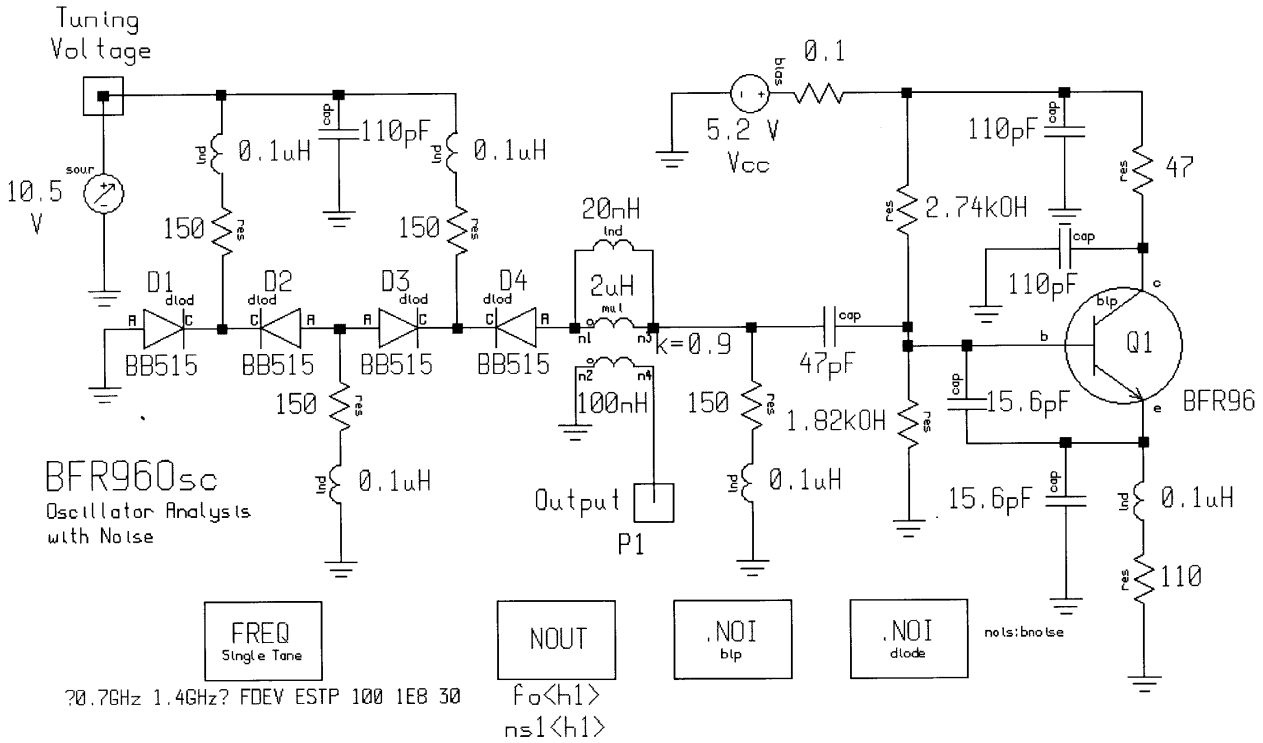


Figure 5-113 BJT microwave oscillator entered into the schematic-capture module of a commercial HB analysis program.

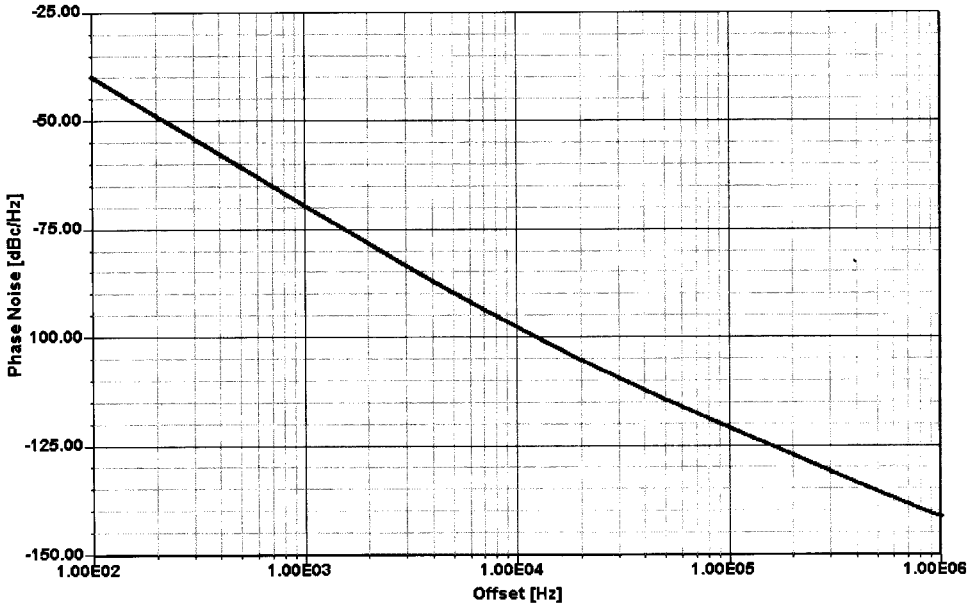


Figure 5-114 Calculated phase noise of the microwave BJT oscillator shown in Figure 5-113.

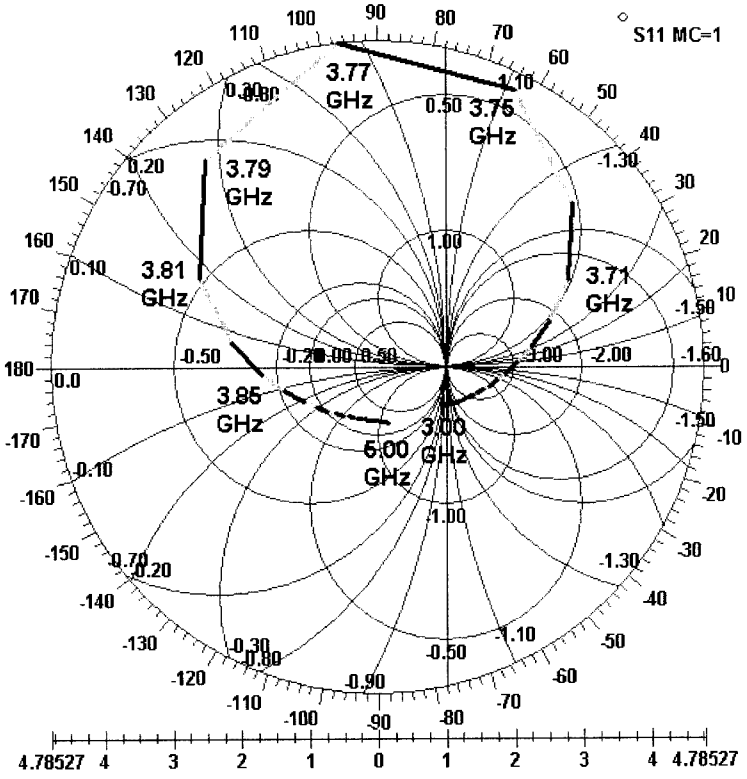


Figure 5-115 S_{11} response of a 3.8-GHz BJT DRO as modeled by Ansoft's Super-Spice time-domain simulator.

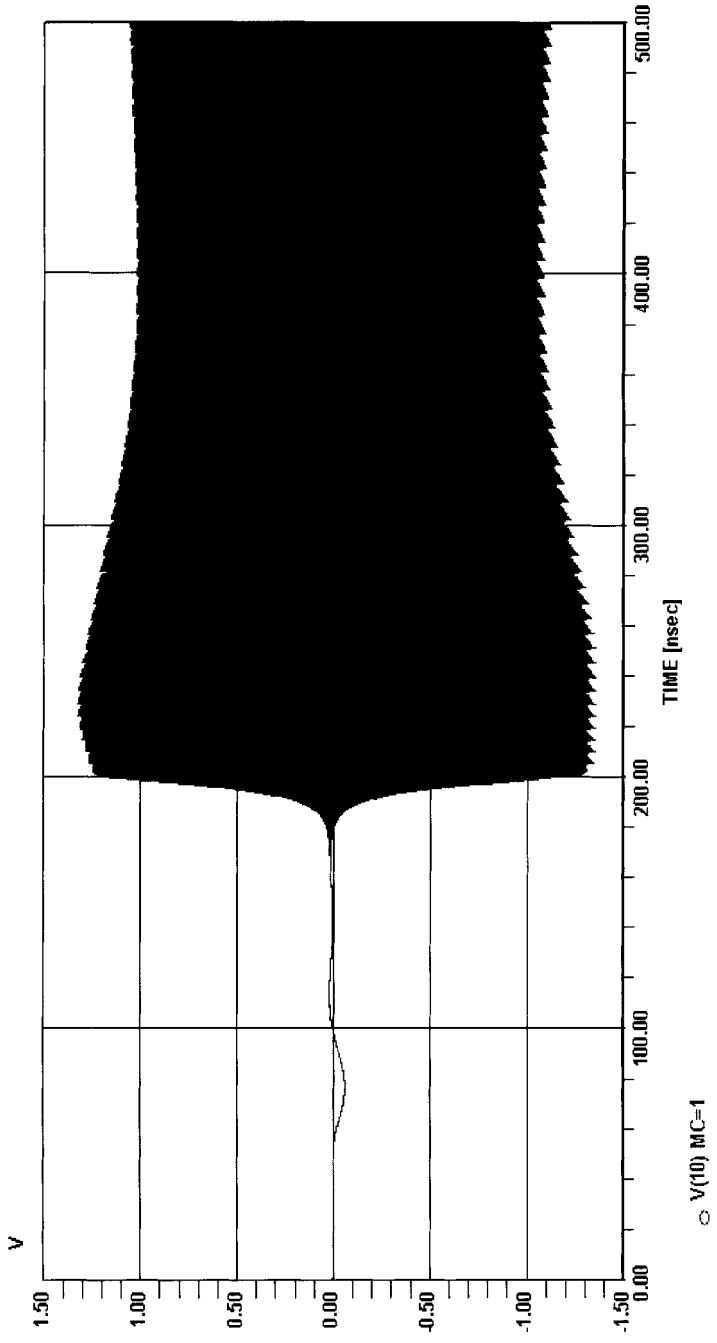


Figure 5-116 Output voltage of the BJT DRO as simulated by Super-Spice.

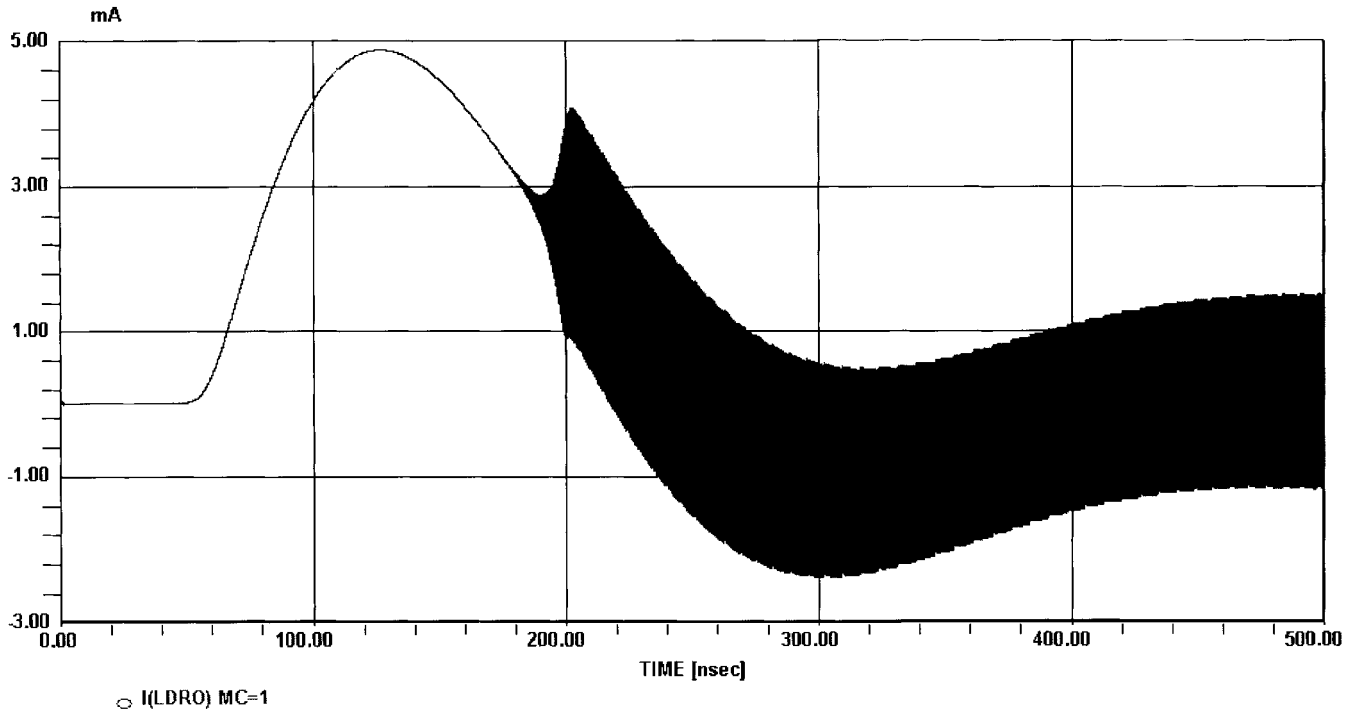


Figure 5-117 Current through the BJT DRO's equivalent resonator inductance from startup to amplitude stabilization.

5-8-2 Time-Domain Simulation

SPICE, a general-purpose simulation program capable of nonlinear dc, linear ac, and nonlinear transient analysis, is probably the most common of the time-domain methods used for nonlinear circuit analysis. Its ability to represent arbitrary time-domain waveforms and model time delays, level shifts, and the sinking and sourcing characteristics of such circuits have made it the simulator of choice for digital designers concerned with analog issues. SPICE's application in RF, microwave, and wireless circuit design is limited by its difficulties with simulating multitone excitations, and coupling and transmission-line effects—including losses, parasitics, and dispersion for distributed elements—and its inefficiency in simulating steady-state conditions.

One area in which SPICE's transient analysis capabilities can be particularly useful to RF and wireless designers is the simulation of oscillator startup. To illustrate this, we present a 3.8-GHz dielectric resonator oscillator modeled with Ansoft Corporation's Super-Spice simulator. Figure 5-115 shows how the oscillator's S_{11} response, positive throughout the simulation's frequency span of 3–5 GHz, peaks at 3.77 GHz. Figure 5-116 shows the oscillator's output voltage as the circuit starts oscillating and amplitude-stabilizes; Figure 5-117 shows the current through the equivalent inductance of the circuit's dielectric resonator across the same time span.

5-9 PHASE-NOISE IMPROVEMENTS OF INTEGRATED RF AND MILLIMETER-WAVE OSCILLATORS *

5-9-1 Introduction

The generation of microwave and millimeter-wave frequencies can be done by lower-frequency VCOs multiplied up in frequency, such as that provided by comb generators; by using hybrid GaAsFET-based oscillators with external resonators; or—for the best phase noise obtainable—YIG oscillators whose physical size and cost do not always provide a practical solution. A significant improvement in oscillator noise performance can be obtained with a novel feedback circuit that uses internally generated noise and cancels the close-in noise in a bandwidth of up to 1 MHz. This results in a phase noise improvement of more than 15 dB at microwave frequencies for the same topology, and the feedback system can be made part of the biasing circuit. This feedback circuit can be used over a wide frequency range and works well at VHF/UHF frequencies and in the millimeter-wave area.

5-9-2 Review of Noise Analysis

Section 5-6 of this chapter covered noise in oscillators in detail, so we will merely review oscillator noise issues here. The first linear model for a basic oscillator, without considering semiconductor noise, was developed by Leeson in 1966 [9] and has been quoted in numerous applications.

The phase noise of a VCO is completely determined by

*This section is based on a paper given at the Fifth International Workshop on Integrated Nonlinear Microwave and Millimeterwave Circuits, sponsored by IEEE at Gerhard-Mercator University, Duisburg, Germany, October 1–2, 1998.

$$£(f_m) = 10 \log\left\{\left[1 + \frac{f_0^2}{(2f_m Q_{\text{load}})^2}\right]\left(1 + \frac{f_c}{f_m}\right)\frac{FkT}{2P_{s,\text{av}}} + \frac{2kTRK_0^2}{f_m^2}\right\} \tag{5-135}$$

where $£(f_m)$ = ratio of sideband power in a 1-Hz bandwidth at f_m to total power in dB
f_m = frequency offset
f_0 = center frequency
f_c = flicker frequency
Q_{load} = loaded Q of the tuned circuit
F = noise factor
$kT = 4.1 \times 10^{-21}$ at 300 K (room temperature)
$P_{s,\text{av}}$ = average power at oscillator output
R = equivalent noise resistance of tuning diode (typically 200 Ω to 10 kΩ)
K_0 = oscillator voltage gain

When adding an isolating amplifier the noise of an *LC* oscillator is determined by

$$S_\phi(f_m) = [a_R F_0^4 + a_E(F_0/2Q_L)^2]/f_m^3$$
$$+ [(2GFkT/P_0)(F_0/2Q_L)^2]/f_m^2$$
$$+ 2a_R Q_L F_0^3/f_m^2$$
$$+ a_E/f_m + 2GFkT/P_0$$

where G = compressed power gain of the loop amplifier
F = noise factor of the loop amplifier
k = Boltzmann's constant
T = temperature in kelvins
P_0 = carrier power level (in watts) at the output of the loop amplifier
F_0 = carrier frequency in hertz
f_m = carrier offset frequency in hertz
$Q_L (= \pi F_0 \tau_g)$ = loaded Q of the resonator in the feedback loop
a_R, a_E = flicker noise constants for the resonator and loop amplifier, respectively

In 1978, Dieter Scherer from Hewlett-Packard added the $(1 + f_c/f_m)$ term, which addresses the 1/*f* noise or flicker corner frequency [25]. The 1/*f* frequency phenomenon is based on the surface effects inside the semiconductor and varies from 50 Hz for silicon FETs, to 5 kHz and more for silicon microwave bipolar transistors, to 1 MHz and higher for GaAsFETs. As a rule of thumb, one can state that the higher the frequency of operation, the higher the f_T for the active device, the higher the flicker corner frequency will be. Figure 5-118 shows the spectral distribution as a function of offset frequency. This linearization was first discovered by individuals from NIST, formerly the National Bureau of Standards, in Boulder, Colorado. The left corner is dominated by the flicker corner frequency and the oscillator itself shows at the output a combination of modulation and conversion

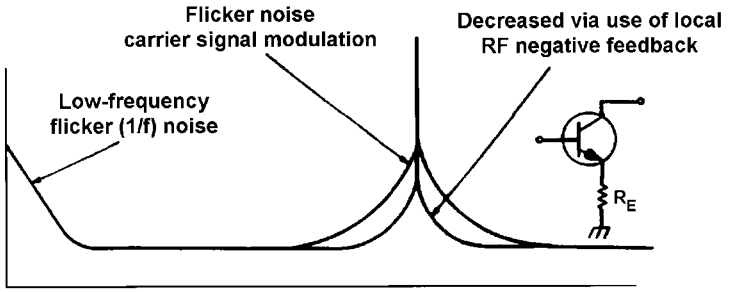


Figure 5-118 Reduction of transistor flicker-of-phase noise via use of local negative feedback (emitter resistance), showing, for a BJT, the flicker noise or corner frequency, as well as the phase noise that results from the flicker noise. The negative feedback from the unbypassed emitter resistor reduces the flicker noise by up to 40 dB [26, 27].

noise. In addition to the $1/f$ noise, the real-life oscillator exhibits a wideband noise, which is due to nonlinearities—specifically, the AM-to-PM conversion resulting from nonlinear capacitance and transconductance of the transistor. This complete linear "noise equation" was first published by Rohde et al. [2]. Therefore, the noise model for the transistor had to be enhanced by describing this mechanism. Essentially, it is equal to a tuning diode coupled to the resonator, which is modulated by thermal noise, and therefore results in a higher phase-noise contribution than the oscillator itself. This noise equation also has Q playing a major role. The phase noise, therefore, ignoring other frequencies, is proportional to $1/Q^2$. In monolithic circuits, the loss of transmission lines, which determines the Q, is related to the substrate material. One has little choice of material, particularly in microwave and millimeter-wave oscillators. Printed resonators sadly provide low Q at these frequencies. A similar problem related to the material is the choices of tuning diodes. In many, if not all, cases, one uses a GaAsFET with the gate connected to the drain or to the source as a tuning diode, and the resulting Q is also disappointing.

The example in Figure 5-118 takes advantage of a microwave bipolar transistor and also is applicable for heterojunction bipolar transistors (HBTs). The transconductance of the bipolar transistor, which is approximately 39 mA/V, increases to values up to 20 times larger than exhibited by GaAsFETs. This method of using negative feedback reduces the gain of the GaAsFET too much and is, therefore, not applicable.

5-9-3 Workarounds

When looking at the modes of operation of test sets for phase-noise measurement, the one method that requires only one signal generator is most desirable. Based on the frequency discriminator method, it uses a cable of appropriate length that acts as a delay line. The oscillator under test drives a double-balanced mixer [one port directly (LO port)] and drives the RF port "delayed" via the delay line cable, and the mixer is therefore driven by the same frequency with a different phase. This setup has been used for phase-noise measurements very successfully. One can easily connect a low-noise FFT-based audio spectrum analyzer to the mixer output and measure the signal-to-noise ratio as a function of offset. With modern FFT analyzers available up to 10 MHz and higher, a fast and reliable measurement can be done. The length of the cable determines the frequency range over which an accurate

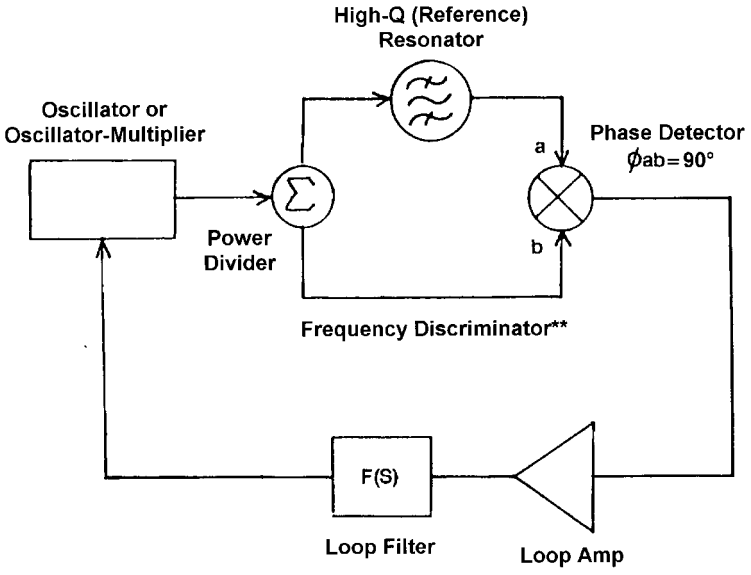
measurement is possible. This phase-noise information (dc and ac voltage) at the output can be used to modulate the VCO and improve the phase noise in the frequency range over which this method is applicable. Fluke and other companies have built signal generators based on this principle to improve the phase noise. The drawback of this method is that the delay line acts like a band-limiting filter, and, therefore, one cannot compensate over a large range of frequencies. The delay line in question, responsible for the phase shift, can also be substituted by a high-Q resonator. By closing the loop, one can "clean up" the phase noise of the device under test (DUT). The difficulty with this is that it requires an external high-Q resonator. One might as well use the same resonator for the oscillator. Therefore, this method is not suitable for MMICs. See Figure 5-119 for more details.

As far as the actual measurement is concerned, one has to observe certain limitations of the test setup as explained in Figure 5-120. When using the delay line instead of a high-Q resonator (low loss is assumed for coaxial cable), one needs to determine the physical length of the cable. The physical length of the cable determines the delay and, therefore, the discriminator noise floor as a function of the offset frequency. Typically, 8 feet of RG-8/U cable results in a delay of 12 ns. As can be seen from Figure 5-121, it is advisable to have several lengths of cable depending on the range of frequency for which one wants to do the measurements under different lengths of cable [28].

Since we are primarily concerned about the flicker noise, here is another, similar, approach to "cleaning up" the flicker noise, as shown in Figure 5-122. The bandwidth over which this technique works depends on the bandwidth of the phase detector system, which can be made fairly wide. A derivative of this will provide a useful solution.

5-9-4 Reduction of Flicker Noise

From inspecting a device datasheet, one is always pleasantly surprised about very low spot noise figures, while the 1/f corner frequency is rarely specified at all. The mechanism that



**Resonator may be operated in reflection mode as well as transmission mode.

Figure 5-119 AFC stabilization of an external oscillator.

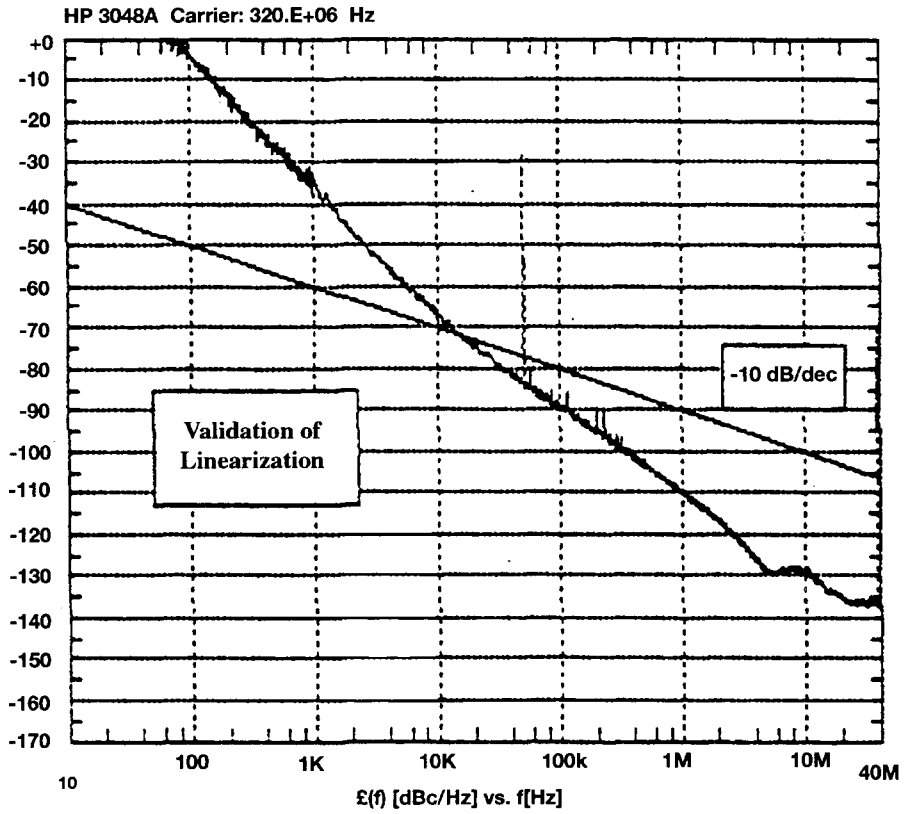


Figure 5-120 Display of a typical phase-noise measurement using the delay line principle. This method is only applicable where $x = \sin(x)$. The measurements above the solid line violate this relationship and therefore are not valid.

transforms the noise, which is generated by the $1/f$ contribution into what we referred to as modulation noise, is the AM-to-PM conversion. If one could "linearize" the transistor, the flicker noise effect would drastically be reduced.

Figure 5-123 shows two variations on an approach that is consistent with the previous method. In Figure 5-123a, the noise across the 56-Ω emitter resistor of oscillator transistor Q1 is sampled, amplified and inverted in Q2, and used to modulate Q1's base voltage. A low-frequency, general-purpose *npn* transistor (BC848C) is used at Q2. A high-frequency, higher-power transistor (BFR93A) is used at Q1; an HBT could also be used. Figure 5-123b shows the same approach using a variation of the *pnp* active biasing circuit shown in Figure 3-192. Figure 5-124 shows the phase-noise performance for the circuit of Figure 5-123a.

5-9-5 Applications to Integrated Oscillators

Figure 5-125 shows a silicon-based BiCMOS technology wireless oscillator. The three big inductors can be seen easily. This design by Motorola is a good candidate for using this innovative technology. Applying this feedback technology, we measure the phase noise shown in Figure 5-126.

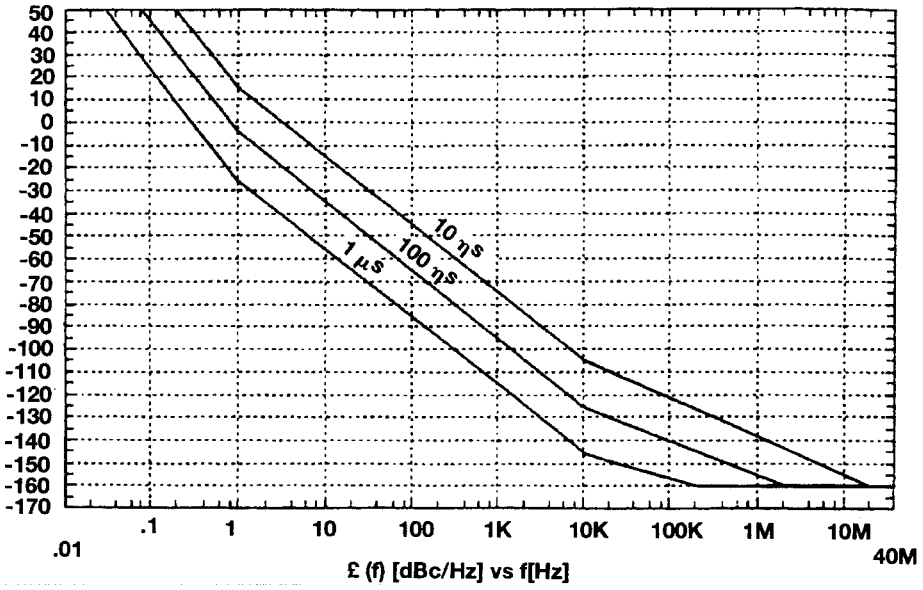


Figure 5-121 Dynamic range as a function of cable delay. A delay of 1 μs is ideal for microwave frequencies.

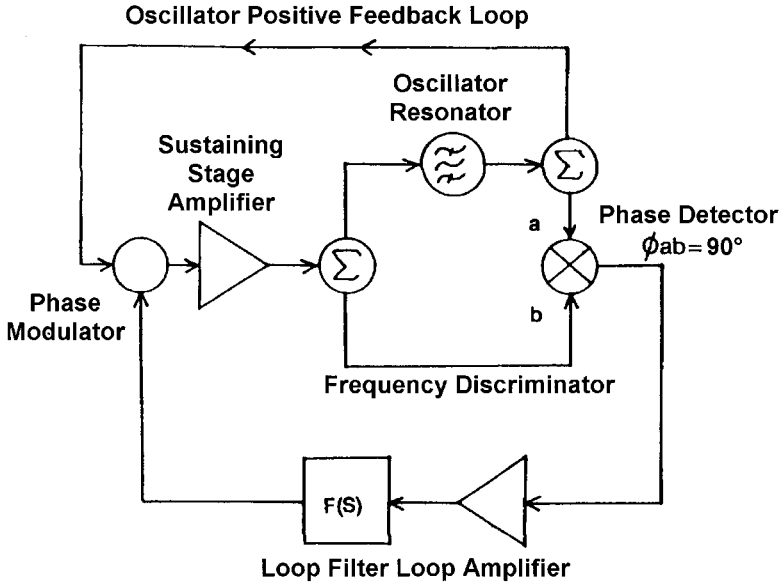
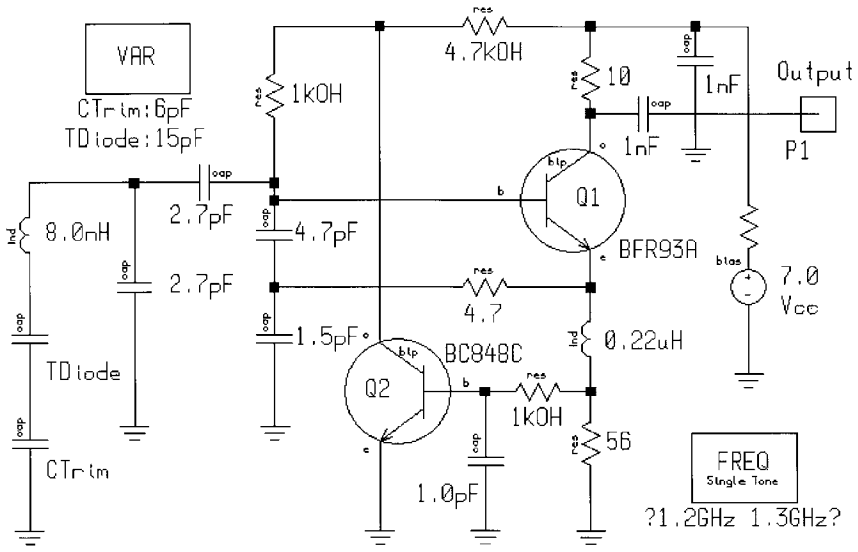
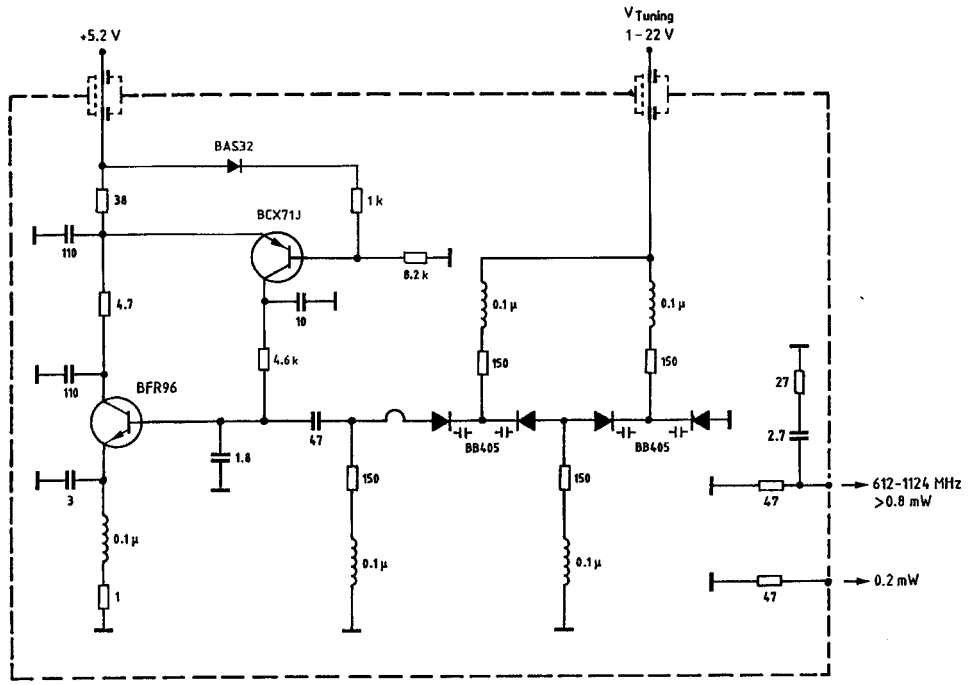


Figure 5-122 Noise reduction in an oscillator whose signal flicker-of-frequency noise is primarily due to sustaining stage flicker-of-phase noise. A phase perturbation in the sustaining stage produces a frequency change in the oscillator, which produces a change in the signal phase shift through the resonator detected as the phase detector. The resonator may be operated simultaneously in transmission line and reflection modes in the oscillator in the oscillator and discriminator portions of the circuit, respectively.



(a)



(b)

Figure 5-123 BJT-based oscillators with noise feedback. (a) The noise sampling is done in the transistor emitter; (b) the noise sampling is done in the collector using a variation of the active biasing circuit shown in Figure 5-124.

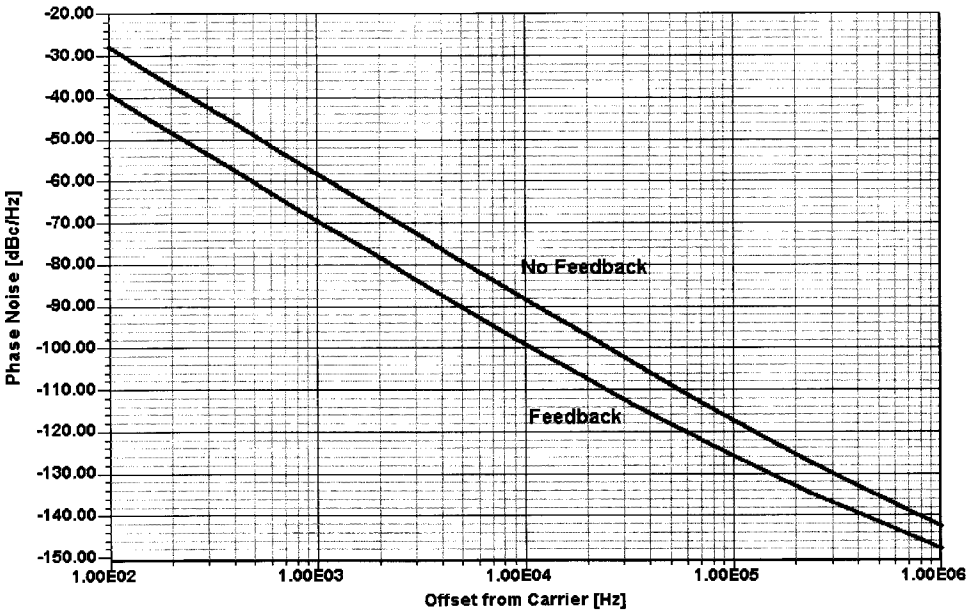


Figure 5-124 Phase noise of the oscillator in Figure 5-123a with and without noise feedback.

Moving up into the millimeter-wave area, one of the target oscillators is the Ka-band MMIC voltage-controlled oscillator designed for Ka-band smart munitions applications (Figure 5-127). This voltage-controlled oscillator MMIC employs 0.25-μm gate length, double-heterojunction, pseudomorphic, high-electron-mobility transistor (PHEMT) technology. This is a custom chip developed by Martin Marietta under the U.S. Government MMIC program. Ansoft Serenade tools were used in this design.

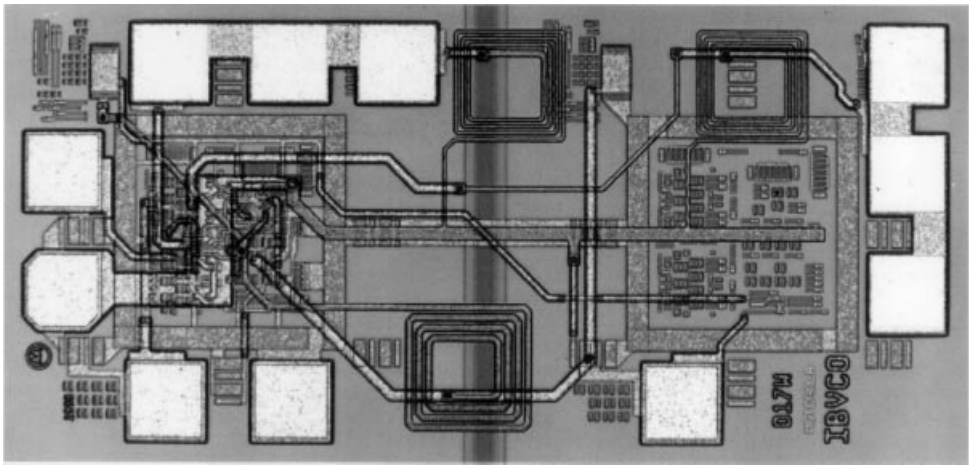


Figure 5-125 Layout of the Motorola 800-MHz monolithic differential oscillator IC in Si technology.

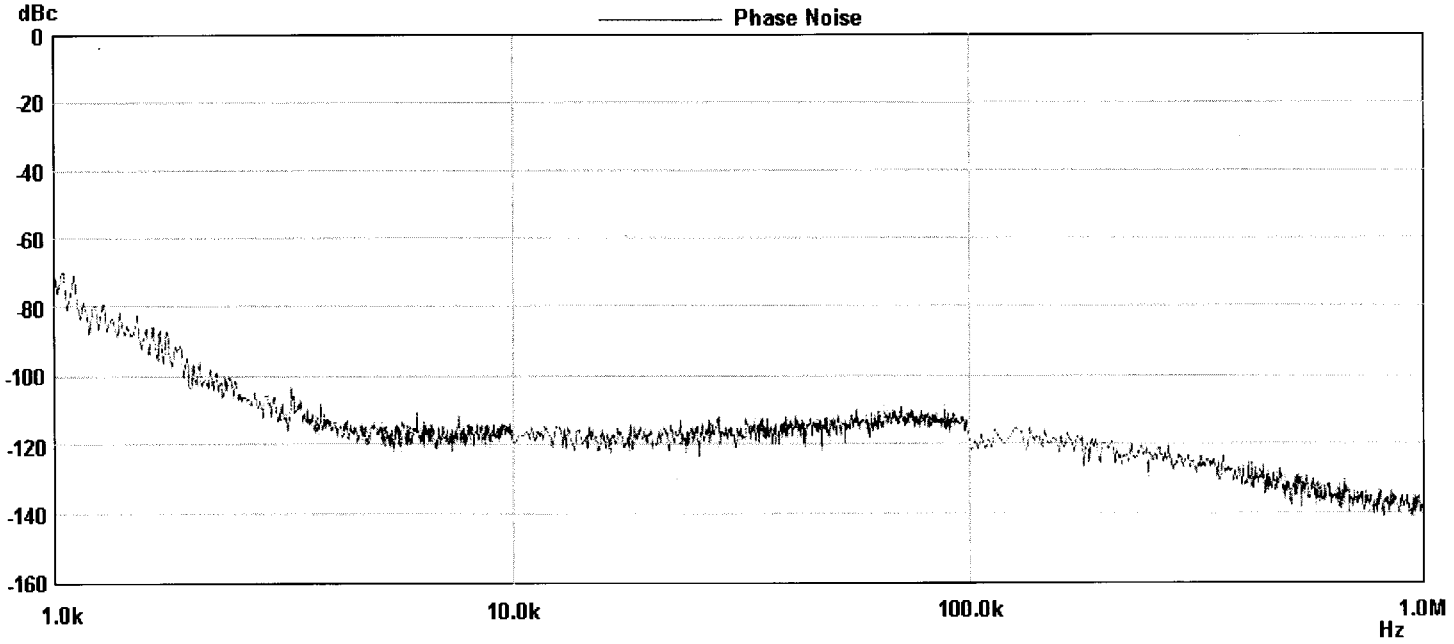


Figure 5-126 Phase noise of the oscillator of Figure 5-125. The noise floor is limited by the on-board isolation amplifier.

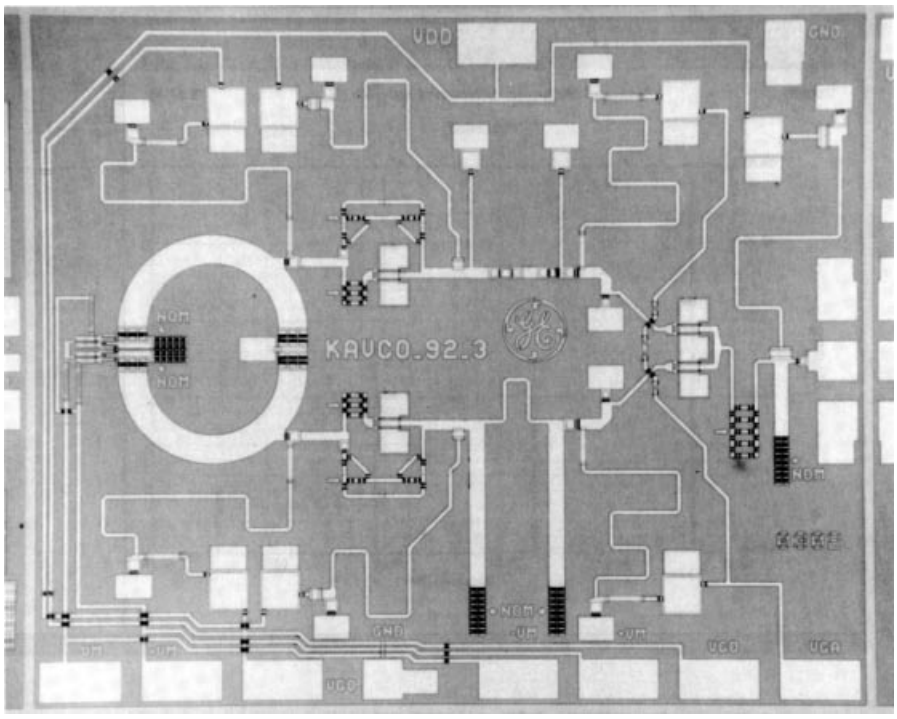


Figure 5-127 Layout of the Ka-band voltage-controlled oscillator.

Features

- Fundamental-mode, differential-ring VCO
- Electronically tunable
- Output power greater than 16 dBm
- Power-added efficiency (PAE) of 13%
- Compact size for easy integration with power amplifier
- 0.25-μm pseudomorphic HEMTs

Specifications

- Frequency range: Ka band
- Output power: 16 dBm minimum

When applying the same technology as shown in the previous bipolar example, the phase noise can be reduced significantly. Since field-effect transistors require a negative gate voltage, it is possible to sample the noise in the drain current using a resistor, amplify the noise with an appropriate loop amplifier, and modulate the gate voltage within a 1-MHz bandwidth.

A spectrum analyzer measurement of this oscillator type, as shown in Figure 5-128, indicates a phase noise of −78.6 dBc/Hz at an offset frequency of 100 kHz at a center frequency of 38 GHz, while the closed-loop phase noise of Figure 5-129 using the compen-

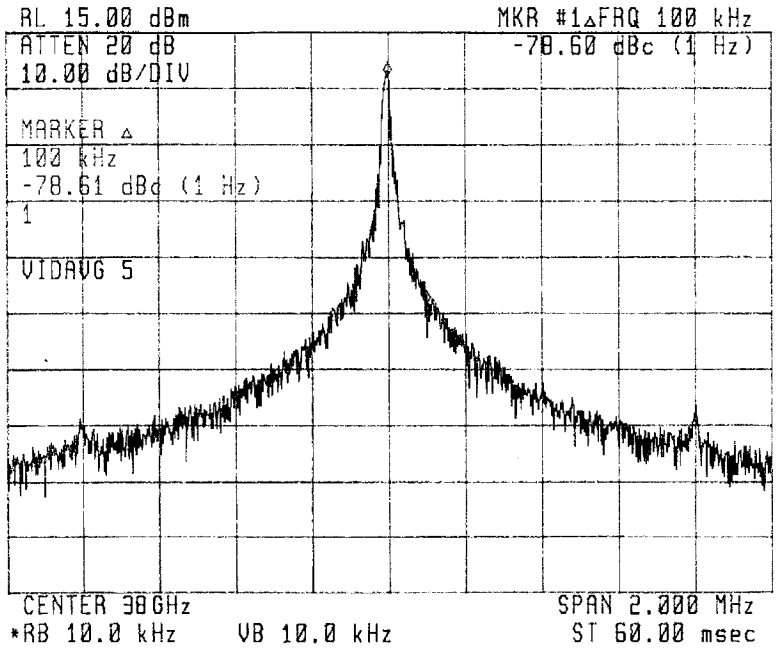


Figure 5-128 Spectrogram of the 38-GHz free-running push–pull VCO shown in Figure 5-127.

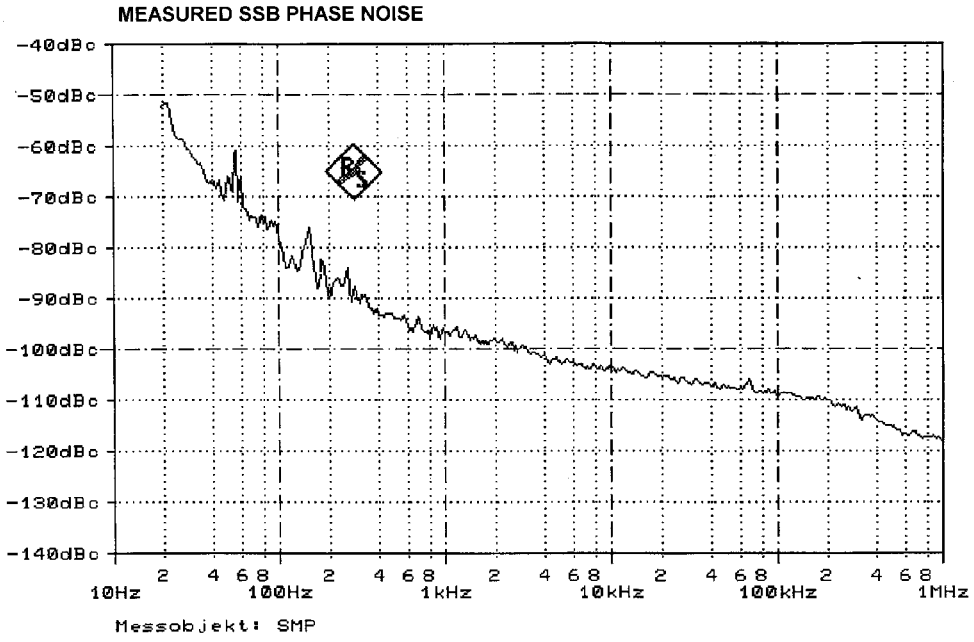


Figure 5-129 Phase-noise performance of a 38-GHz oscillator with phase-noise “clean-up” system.

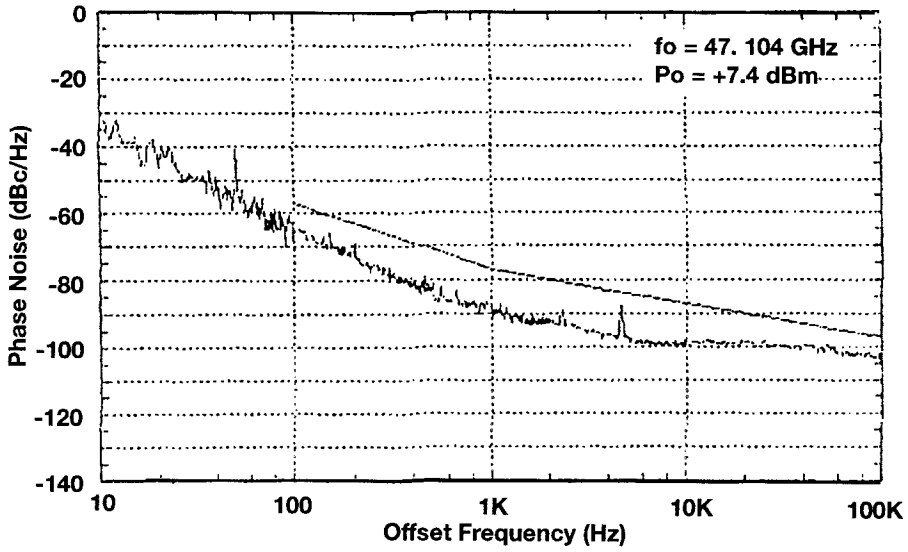


Figure 5-130 A 47-GHz loop-stabilized VCO with phase-noise performance similar to that of Figure 5-129.

sation scheme indicates a phase noise of about −108 dBc/Hz at 100 kHz. This is consistent with reported results from other researchers who have used complicated phase-locked loops to achieve similar performance as seen in Figure 5-130.

5-9-6 Summary

A novel method of improving the phase noise of oscillators has been demonstrated. This approach works well for low-frequency oscillators such as 1000 MHz up to the millimeter-wave range. It only requires an additional wideband dc biasing circuit, which applies noise feedback. The correctness of the theory has been demonstrated by appropriate validation.

REFERENCES

1. Paul Danzer, ed., *The ARRL Handbook for Radio Amateurs,* 75th ed. Newington, CT: American Radio Relay League, 1997, Chap. 14, AC/RF Sources (Oscillators and Synthesizers).
2. Ulrich L. Rohde, Jerry C. Whitaker, and T. T. N. Bucher, *Communications Receivers: Principles and Design*, 2nd ed. New York: McGraw-Hill, 1997, p. 380.
3. G. Vendelin, A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design: Using Linear and Nonlinear Techniques.* New York: Wiley, New York, 1990.
4. Samuel Y. Liao, *Microwave Devices and Circuits,* 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1985.
5. S. C. Peterson, "900-MHz μ-Strip," Philips Semiconductors application document.
6. H. Keller, "Electronic UHF Tuning in TV Receivers," *Radio-Fernsch-Phono-Praxis,* No. 3, 1967.
7. L. Micic, "The Tuner Diode," *Radio Mentor Electron.,* **32**(5), 404–405, 1966.

8. Kyung-Whan Yeom, "An Investigation of the High-Frequency Limit of a Miniaturized Commercial Voltage-Controlled Oscillator Used in 900-MHz Band Mobile-Communication Handset," *IEEE Trans. Microwave Theory Tech.,* **MTT-46**, 1165–1168, August 1998.
9. D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," *Proc. IEEE,* **54**, 329–330, 1966.
10. U. L. Rohde, C. R. Chang, and J. Gerber, "Parameter Extraction for Large Signal Noise Models and Simulation of Noise in Large Signal Circuits Like Mixers and Oscillators," *Proceedings of the 23rd European Microwave Conference,* Madrid, Spain, September 6–9, 1993.
11. C. R. Chang, "Mixer Noise Analysis Using the Enhanced Microwave Harmonica," Compact Software *Transmission Line News*, **6**(2), 4–9, June 1992.
12. T. Antognetti and G. Massobrio, *Semiconductor Device Modeling with SPICE.* New York: McGraw-Hill, 1988, p. 91.
13. R. J. Hawkins, "Limitation of Nielsen's and Related Noise Equations Applied to Microwave Bipolar Transistors, and a New Expression for the Frequency and Current Dependent Noise Figure," *Solid State Electron.*, **20**, 191–196, 1977.
14. Tzu-Hwa Hus and Craig P. Snapp, "Low Noise Microwave Bipolar Transistor Sub-Hall-Micrometer Emitter Width," *IEEE Trans. Electron Devices*, **ED-25**, 723–730, June 1978.
15. R. A. Pucel and U. L. Rohde, "An Accurate Expression for the Noise Resistance R_n of a Bipolar Transistor for Use with the Hawkins Noise Model," *IEEE Microwave Guided Wave Lett.*, **3**(2), 35–37, February 1993.
16. See reference 3.
17. R. A. Pucel, W. Struble, Robert Hallgren, and U. L. Rohde, "A General Noise De-embedding Procedure for Packaged Two-Port Linear Active Devices," *IEEE Trans. Microwave Theory Tech.,* **MTT-40**(11), 2013–2024, November 1993.
18. U. L. Rohde, "Improved Noise Modeling of GaAs FETS: Using an Enhanced Equivalent Circuit Technique," *Microwave J.,* Part I: 87–101, November 1991; Part II: 87–95, December 1991.
19. V. Rizzoli, F. Mastri, and C. Cecchefti, "Computer-Aided Noise Analysis of MESFET and HEMT Mixers," *IEEE Trans. Microwave Theory Tech.,* **MTT-37**, 1401–1410, September 1989.
20. V. Rizzoli and A. Lippadni, "Computer-Aided Noise Analysis of Linear Multiport Networks of Arbitrary Topology," *IEEE Trans. Microwave Theory Tech.,* **MTT-33**, 1507–1512, December 1985.
21. V. Rizzoli, F. Mastri, and D. Masotti, "General-Purpose Noise Analysis of Forced Nonlinear Microwave Circuits," *Military Microwave,* 1992.
22. U. L. Rohde, "Key Components of Modern Receiver Design," *VHF Conference*, Weinheim, Germany, September 17–18, 1993.
23. C. R. Chang and U. L. Rohde, "The Accurate Simulation of Oscillator and PLL Phase Noise in RF Sources," *Wireless '94 Symposium,* Santa Clara, California, February 15–18, 1994.
24. U. L. Rohde, "Oscillator Design for Lowest Phase Noise," *Microwave Eng. Eur.*, 31–40, May 1994.
25. Dieter Scherer, "Design Principles and Test Methods for Low Phase Noise RF and Microwave Sources," *RF & Microwave Measurement Symposium and Exhibition*, Hewlett-Packard.
26. D. J. Healy III, "Flicker of Frequency and Phase and White Frequency and Phase Fluctuations in Frequency Sources," *Proceedings of the 26th Annual Symposium on Frequency Control*, Fort Monmouth, NJ, June 1972, pp. 43–49.
27. Michael M. Driscoll, "Low Noise Oscillator Design Using Acoustic and Other High *Q* Resonators," *44th Annual Symposium on Frequency Control*, Baltimore, MD, May 1990.
28. Ulrich L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design*. New York: Wiley, 1997.

INTERESTING PATENTS

Michel Fache and Claude Juliot, "Tunable UHF Oscillator with Harmonic Limitation," Lignes Telegraphoques et Telephoniques (Societe), Paris Cedex, France, U.S. Patent No. 4,146,850, March 27, 1979.

Thomas M. Higgins, Jr., "Negative Resistance Oscillator with Electronically Tunable Base Inductance," Hewlett-Packard Company, Palo Alto, California, U.S. Patent No. 5,373,264, December 13, 1994.

Osamu Ichiyoshi, "Voltage-Control Oscillator Which Suppresses Phase Noise Caused by Internal Noise of the Oscillator," NEC Corporation, Japan, U.S. Patent No. 5,351,014, September 27, 1994.

Shankar R. Joshi, Ulrich L. Rohde, and Klaus Eichel, "Phase-Locked Loop Circuits and Voltage Controlled Oscillator Circuits," Synergy Microwave Corp., Paterson, New Jersey, U.S. Patent No. 5,650,754, July 22, 1997.

Joseph H. Kiser, "Wide-Range Electronic Oscillator," Vari-L Co., Denver, Colorado, U.S. Patent No. 4,621,241, November 4, 1986.

Joseph H. Kiser, "Oscillator Voltage Regulator," Vari-L Co., Denver, Colorado, U.S. Patent No. 5,675,478, October 7, 1997.

Charles Lewis, "Low Phase Noise UHF and Microwave Oscillator," Z-Communications, Inc., San Diego, California, U.S. Patent No. 5,748,051, May 5, 1998.

Larry R. Lockwood, "Low-Noise Oscillator," Tektronix, Inc., Beaverton, Oregon, U.S. Patent No. 4,580,109, April 1, 1986.

Benedict J. Nardi, "Low Phase Noise Reference Oscillator," Watkins-Johnson Company, Palo Alto, California, U.S. Patent No. 5,341,110, August 23, 1994.

Leon Riebman, "Low-Noise Oscillator," AEL Defense Corp., Lansdale, Pennsylvania, U.S. Patent No. 5,166,647, November 24, 1992.

Takeshi Saitoh and Hiroshi Hatashita, "TV Tuner Oscillator with Feedback for More Low-Frequency Power," Hitachi Ltd., Tokyo, Japan, U.S. Patent No. 4,564,822, January 14, 1986.

FURTHER READING

Stanislaw Alechno, "Analysis Method Characterizes Microwave Oscillators," Part 1, *Microwaves & RF*, **36**(11), 82–86, November 1997.

P. Antognetti and G. Massobrio, *Semiconductor Device Modeling with SPICE.* New York: McGraw-Hill, 1993.

W. Anzill, F. X. Kärtner, and P. Russer, "Simulation of the Single-Sideband Phase Noise of Oscillators," *Second International Workshop of Integrated Nonlinear Microwave and Millimeterwave Circuits*, 1992.

Behagi, "Piece-wise Linear Modeling of Solid-State Varactor-Tuned Microwave Oscillators," *Proceedings of the IEEE Frequency Control Symposium*, 1992, pp. 415–519.

Noël Boutin, "RF Oscillator Analysis and Design by the Loop Gain Method," *Appl. Microwave & Wireless*, **11**(8), 32–48, August 1999.

P. Braun, B. Roth, and A. Beyer, "A Measurement Setup for the Analysis of Integrated Microwave Oscillators," *Second International Workshop of Integrated Nonlinear Microwave and Millimeterwave Circuits,* 1992.

C. R. Chang, M. B. Steer, S. Martin, and E. Reese, "Computer-Aided Analysis of Free-Running Microwave Oscillators," *IEEE Trans. Microwave Theory Tech.,* **MTT-39**, 1735–1745, October 1991.

Y. Cheng, K. Chen, K. Imai, and C. Hu, "A Unified MOSFET Channel Charge Model for Device Modeling in Circuit Simulation," *IEEE Trans. Computer-Aided Des.*, **17**, 641–644, August 1998.

J. Craninckx and M. Steyaert, "A 1.8-GHz Low-Phase-Noise CMOS VCO Using Optimized Hollow Spiral Inductors," *IEEE J. Solid-State Circuits*, **SC-32**(5), 736–744, May 1997.

P. Davis, P. Smith, E. Campbell, J. Lin, K. Gross, G. Bath, Y. Low, M. Lau, Y. Degani, J. Gregus, R. Frye, and K. Tai, "Silicon-on-Silicon Integration of a GSM Transceiver with VCO Resonator," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers,* pp. 248–249.

Alper Demir, Amit Mehrotra, and Jaijeet Roychowdhury, "Phase Noise and Timing Jitter in Oscillators," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 45–48.

P. J. Garner, M. H. Howes, and C. M. Snowden, "*Ka*-Band and MMIC pHEMT-Based VCOs with Low Phase-Noise Properties," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1531–1536, October 1998.

G. Gonzalez, *Microwave Transistor Amplifiers—Analysis and Design,* 2nd ed. New York: Prentice-Hall, 1997.

A. V. Grebennikov and V. V. Nikiforov, "An Analytic Method of Microwave Transistor Oscillator Design," *Int. J. Electron.,* **83**, 849–858, December 1997.

Andrey V. Grebennikov, "Microwave Transistor Oscillators: An Analytic Approach to Simplify Computer-Aided Design," *Microwave J.*, **42**(5), 292–300, May 1999.

Ali Hajimiri and Thomas H. Lee, "A General Theory of Phase Noise in Electrical Oscillators," *IEEE J. Solid-State Circuits*, **SC-33**, 179–194, February 1998.

Ali Hajimiri, Sotiriosw Limotyrakis, and Thomas H. Lee, "Phase Noise in Multi-gigahertz CMOS Ring Oscillators," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 49–52.

Frank Herzel and Behzad Razavi, "Oscillator Jitter Due to Supply and Substrate Noise," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 489–492.

Qiuting Huang, "On the Exact Design of RF Oscillators," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference*, pp. 41–44.

Kamran Iravani and Gary Miller, "VCOs with Very Low Sensitivity to Noise on the Power Supply," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference,* pp. 515–518.

K. M. Johnson, "Microwave Varactor-Tuned Transistor Oscillator Design," *IEEE Trans. Microwave Theory Tech.*, **MTT-14**(11), 564–572, September 1966.

K. M. Johnson, "Large-Signal GaAs MESFET Oscillator Design," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**(3), 217–226, March 1979.

W. E. Kamali, J. P. Grimm, R. Meierer, and C. Trirons, "New Design Approach for Wideband FET Voltage-Controlled Oscillators," *IEEE Trans. Microwave Theory Tech.,* **MTT-34**(10), 1059–1063, October 1986.

F. X. Kärtner, "Noise in Oscillating Systems," *Second International Workshop of Integrated Nonlinear Microwave and Millimeterwave Circuits*, 1992.

T. Kashiwa, T. Ishida, T. Katho, H. Kurusu, H. Hoshi, and Y. Mitsui, "V-Band High-Power Low Phase-Noise Monolithic Oscillators and Investigation of Low Phase-Noise Performance at High Drain Bias," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1559–1565, October 1998.

Peter Kinget, "A Fully Integrated 2.7V 0.35 μm CMOS VCO for 5GHz Wireless Applications," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 226–228.

J. Kitchen, "Octave Bandwidth Varactor-Tuned Oscillators," *Microwave J.*, **30**(5), 347–353, May 1987.

A. Kral, F. Behbahani, and A. A. Abidi, "RF-CMOS Oscillators with Switched Tuning," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference*, pp. 555–558.

Vênceslav F. Kroupa, ed., *Direct Digital Frequency Synthesizers.* New York: IEEE Press, 1999.

K. Kurokawa, "Some Basic Characteristics of Broadband Negative Resistance Oscillator Circuits," *Bell Syst. Tech. J.*, 1937–1955, July 1969.

Fred Labaar, "New Discriminator Boosts Phase-Noise Testing," *Microwaves*, 65–69, March 1982.

A. L. Lacaita and C. Samori, "Phase-Noise Performance of Crystal-Like *LC* Tanks," *IEEE Trans. Circuits Syst. Analog Digital Signal Processing*, **45**, 898–900, 1998.

Sen-You Liu and Huey-Ru Chuang, "A 2.4-GHz Transceiver RF Front-End for ISM-Band Digital Wireless Communications," *Appl. Microwave & Wireless*, **10**(5), 32–48, June 1998.

M. Madihian and T. Noguchi, "An Analytical Approach to Microwave GaAsFET Oscillator Design," *NEC Res. Dev.*, No. 82, 49–56, July 1986.

J. O. McSpadden, L. Fan, and K. Chang, "High-Efficiency *Ku*-Band Oscillators," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1566–1571, October 1998.

Mini-Circuits Engineering, "Designers' Considerations for VCOs Used in Phase-Locked Loops," *Microwave Prod. Digest*, 24, 34, 36, 37, 40, February 1999.

T. Nakagawa and T. Ohira, "A Novel Phase Noise Reduction Technique for MMIC Frequency Synthesizers that Uses Newly Developed Pulse Generator LSI," NTT Radio Communication Systems Laboratories, 1-2356 Take, Yokosuka-shi, Kanagawa 238-03, Japan.

Mohamed K. Nezami, "Evaluate the Impact of Phase Noise on Receiver Performance," *Microwaves & RF*, **41**(5), 165–172, May 1998.

E. C. Niehenke and R. D. Hess, "A Microstrip Low Noise X-Band Voltage-Controlled Oscillator," *IEEE Trans. Microwave Theory Tech.*, **MTT-27**(12), 1075–1079, October 1986.

Chris O'Connor, "Develop Trimless Voltage-Controlled Oscillators," *Microwaves & RF*, **38**(7), 68–78, July 1999.

T. Ohira et al., "A Compact Full MMIC Module for Ku-Band Phase-Locked Oscillators," *IEEE Trans. Microwave Theory Tech.*, **MTT-37**(4), 723–728, April 1989.

D. P. Owen, "Measurement of Phase Noise in Signal Generators," *Marconi Instrum.,* **15**(6), 117–122, 19xx.

D. F. Page, "A Design Basis for Junction Transistor Oscillator Circuits," *Proc. IRE,* **46**, 1271–1280, November 1958.

Byeong-Ha Park, "1 GHz, Low-Power, Low-Phase-Noise, CMOS Voltage-Controlled Oscillator with an Integrated LC Resonator," private communication with the author.

D. F. Paterson, "Varactor Properties for Wideband Linear-Tuning Microwave VCOs," *IEEE Trans. Microwave Theory Tech.*, **MTT-28**(2), 110–119, February 1980.

Matthew Pope, Russell Crouch, and Jeff Gudewicz, "Dual Band Voltage Controlled Oscillators for Personal Communication," *Microwave Prod. Digest,* 52, 84, 85, May 1999.

Nenad Popovic, "Review of Some Types of Varactor Tuned DROs," *Appl. Microwave & Wireless,* **11**(8), 62–70, August 1999.

R. A. Pucel and J. Curtis, "Near-Carrier Noise in FET Oscillators," *IEEE MTT-S International Microwave Symposium Digest*, 1983, pp. 282–284.

S. Qi, K. Wu, and Z. Ou, "Hybrid Integrated HEMT Oscillator with a Multiple-Ring Nonradiative Dielectric (NRD) Resonator Feedback Circuit," *IEEE Trans. Microwave Theory Tech.,* **MTT-46**, 1552–1558, October 1998.

M. Regis, O. Llopis, and J. Graffeuil, "Nonlinear Modeling and Design of Bipolar Transistors Ultra-Low Phase-Noise Dielectric-Resonator Oscillators," *IEEE Trans. Microwave Theory Tech.*, **MTT-46**, 1589–1593, October 1998.

V. Rizzoli, F. Mastri, and D. Masotti, "A General Purpose Harmonic Balance Approach to the Computation of Near-Carrier Noise in Free-Running Microwave Oscillators," *MTT-S,* **x**, 309–312, 1993.

M. Runkel, "Analyse und Vergleich Spannungsgesteuerter Oszillatoren für Mobilfunkanwendungen," Master Thesis, Gerhard-Mercator-Universistät Gesamthochschule Duisburg, Chaired by Prof. Dr. Ingo Wolff, August 1994.

Carlo Samori, Andrea L. Lacaita, Francesco Vill, and Franco Zappa, "Spectrum Folding and Phase Noise in *LC* Tuned Oscillators," *IEEE Trans. Circuits Syst. Analog Digital Signal Processing,* **45**, 781–790, 1998.

Sarafian and B. Z. Kaplan, "A New Approach to the Modeling of the Dynamics of RF VCOs and Some of Its Practical Implications," *IEEE Trans. Circuits Syst. Fundam. Theory Applic.,* **40**(12), 895–901, December 1993.

G. Sauvage, "Phase Noise in Oscillators: A Mathematical Analysis of Leeson's Model," *IEEE Trans. Instrum. Meas.*, **IM-26**(4), x–x, December 1977.

Jwo-Shiun Sun, "Design and Analysis of Microwave Varactor-Tuned Oscillators," *Microwave J.,* 302–310, May 1999.

R. J. Trew, "Octave-Band GaAs FET YIG-Tuned Oscillators," *Electron. Lett.,* **13**(21), October 1977.

R. J. Trew, "MESFET Models for Microwave Computer-Aided Design," *Microwave J.,* **33**(5) 115–128, May 1990.

T. Yamasaki and E. Nagata, *Microwave Synthesizer for Terrestrial and Satellite Communications,* NEC Corporation, June 1989.

N. L. Wang and W. J. Ho, "X-Band HBT VCO with High-Efficiency CB Buffer Amplifier," *IEEE J. Solid-State Circuits,* **SC-27**(10), 1439–1443, October 1992.

Shing-Tak Yan and Howard C. Luong, "A 3-V 1.3-to-1.8-GHz CMOS Voltage-Controlled Oscillator with 0.3-ps Jitter," *IEEE Trans. Circuits Syst. Analog Digital Signal Processing,* **45**, 876–880, 1998.

D. A Warren, J. M. Golio, and W. L. Seely, "Large and Small Signal Oscillator Analysis," *Microwave J.*, **32**(5), 229–246, May 1989.

R. Winch, "Wideband Varactor-Tuned Oscillators," *IEEE J. Solid-State Circuits,* **SC-17**, 1214–1219, December 19, 1982.

Markus Zannoth, Bernd Kolb, Josef Fenk, and Robert Weigel, "A Fully Integrated VCO at 2 GHz," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers,* pp. 224–225.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

6

WIRELESS SYNTHESIZERS

6-1 INTRODUCTION

Increasing integration has drastically narrowed the range of component choices open to wireless synthesizer designers. The design of a high-performance synthesizer is largely reduced to selecting the most advanced synthesizer integrated circuit (IC) and, if the synthesizer uses a phase-locked loop (PLL), designing or obtaining the best voltage-controlled oscillator (VCO). This chapter covers synthesizer theory, evaluation, and design, including PLL and direct digital synthesis (DDS) techniques. Going into all the details of frequency synthesizers would be beyond the scope of this book. For depth and background on this subject, we recommend the book by Rohde [1].

6-2 PHASE-LOCKED LOOPS

6-2-1 PLL Basics

Figure 6-1 shows a complete PLL synthesizer block diagram indicating the areas over which the designer actually has control. Once the VCO signal has been translated from analog to quasidigital (square-wave) form in circuitry similar to a line receiver, the synthesizer IC takes over. The VCO receives an analog control signal that results from the integration of digital pulses from a phase/frequency discriminator. Modern phase/frequency discriminators, which are part of the PLL IC, use edge-triggered loop locks and generate correcting pulses of either positive or negative sign. The output portion of such a phase detector is frequently referred to as a *charge pump* because the resulting dc control voltage, which ultimately determines the oscillator frequency, is processed by a loop filter containing at least one large capacitor that is charged or discharged as necessary to maintain the control-voltage level necessary for loop lock.

Figure 6-2 shows the block diagram of a single-loop synthesizer. Unless special techniques are used, such as the *fractional-N-division* principle, the step size or channel frequency spacing is equal to the reference frequency. When describing frequency synthesizers mathe-

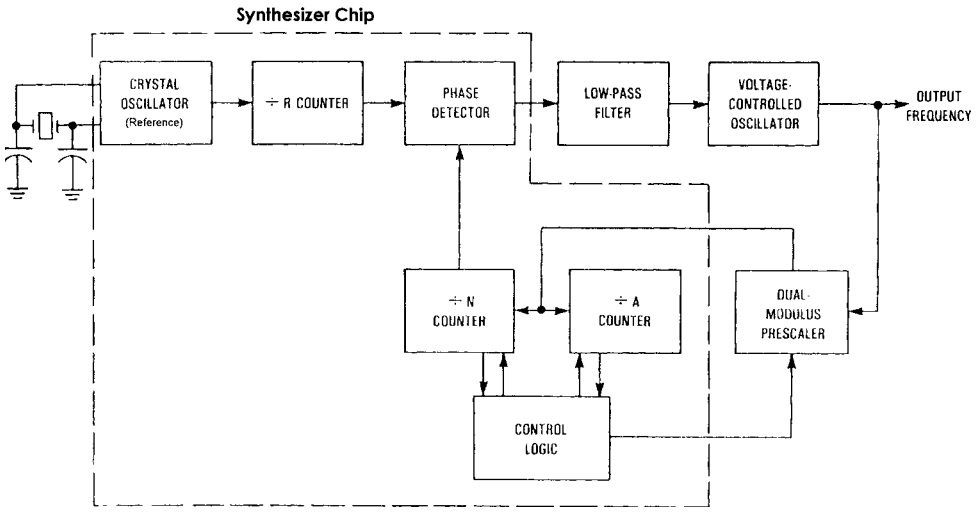


Figure 6-1 Block diagram of an integrated frequency synthesizer. In this case, the designer has control over the VCO and the loop filter; the reference oscillator is part of the chip. In most cases (up to 2.5 GHz), the dual-modulus prescaler is also inside the chip.

matically, we usually use a linearized model. Because most effects occurring in the phase detector are highly nonlinear, only the so-called piecewise-linear treatment allows adequate approximation.

We assume that the VCO shown in Figure 6-2 is tunable over the frequency range from 410 to 510 MHz. Its output is divided to the reference frequency in a programmable divider (i.e., divide by N) whose output is fed to one of the input of the phase/frequency detector and compared with the reference frequency supplied to the other input. The loop filter at the output of the phase detector suppresses reference-frequency components while also serving as an integrator. The dc control voltage at the output of the loop filter tunes the VCO until the divided frequency and phase equal those of the reference. In this simple example with the divider set to 45,000 and the reference set to 1 kHz, the VCO is controlled to a frequency of 450 MHz. A fixed division of the frequency standard output produces the reference frequency of the appropriate step size. Frequency standards are typically operated at 1, 5, or 10 MHz to take advantage of high crystal stability. A 5-MHz frequency standard would be

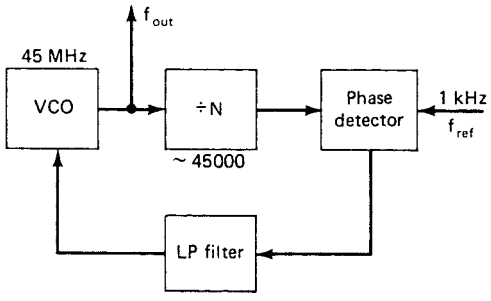


Figure 6-2 Block diagram of a single-loop synthesizer.

divided by 500 in the example. The operating range of the PLL is determined by the maximum operating frequency of the programmable divider, by its division-range ratio, and by the tuning range of the VCO.

The PLL is nonlinear because the phase detector is nonlinear. However, it can be accurately approximated by a linear model when the loop is in lock. The response, when the loop is closed, may be expressed as

$$\frac{\theta_c(s)}{\theta_r(s)} \equiv B(s) = \frac{\text{Forward gain}}{1 + \text{Open-loop gain}}$$

$$= \frac{G(s)}{1 + G(s)/N} \tag{6-1}$$

where $G(s) = G_1(s)G_2(s)F(s)/s$, and θ_c and θ_r are the phases of the controlled oscillator and the reference, respectively.

When the loop is locked, it is assumed that the phase-detector output voltage is proportional to the difference in phase between its inputs; that is,

$$V_\theta = K_\theta(\theta_r - \theta_i) \tag{6-2}$$

where V_θ is the output voltage of the phase detector, and θ_r and θ_i are the phases of the reference signal and the divided VCO signal, respectively. K_θ is the phase-detector gain factor and has the dimensions of volts per radian. It is also assumed that the VCO can be modeled as a linear device whose output frequency differs from its free-running frequency by an increment of frequency

$$2\pi\delta f = K_0 V_c \tag{6-3}$$

where V_c is the voltage of the output of the low-pass filter, and K_0 is the VCO gain factor with the dimensions of radians per second per volt. Because frequency is the time derivative of phase, the VCO operation can be described as

$$2\pi\delta f \equiv \frac{d\theta_c}{dt} = K_0 V_c \tag{6-4}$$

With these assumptions, the PLL may be represented by the linear model shown in Figure 6-3.

The linear transfer function relating $\theta_c(s)$ and $\theta_r(s)$ is

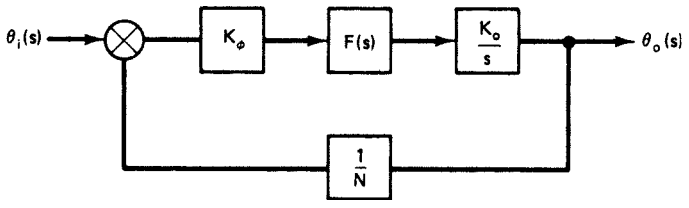


Figure 6-3 Block diagram of a linearized model of a PLL.

$$B(s) = \frac{\theta_c(s)}{\theta_r(s)} = \frac{K_\theta K_0 F(s)/s}{1 + K_\theta K_0 F(s)/Ns} \tag{6-5}$$

The forward gain is

$$G(s) = \frac{K_\theta K_0 F(s)}{s} \tag{6-6}$$

and the open-loop gain is

$$G(s)H(s) = \frac{K_\theta K_0 F(s)}{Ns} \tag{6-7}$$

which leads to the transfer formula of Eq. (6-1).

There are various choices of filter response $F(s)$. Because the VCO by itself is an integrator, we can use a simple *RC* filter following the phase detector. This arrangement is called a *Type 1* filter. Because the components used, together with feedthrough capacitors and other stray effects, can cause excess phase shift, it is necessary to ensure that stability criteria are satisfied. If the gain of a passive loop is too small to provide adequate drift stability of the output phase, especially if a high division ratio is used, the best solution to this problem is the use of an amplifier as an integrator. In most frequency synthesizers, the active-filter-integrator approach is preferred to the passive one. Some frequency synthesizer chips have a single-ended output. In such cases, the use of an additional integrator requires some precautions.

6-2-2 Phase/Frequency Comparators

The phase/frequency comparator can be divided into two types:

1. Phase detectors.
2. Phase-frequency comparators.

This means that the phase comparator has limited means to compare two signals and only accepts phase, not frequency, information. In this case, particular measures have to be taken to pull the VCO into the locking range. The phase comparators require special locking help. Here we are analyzing only the performance.

Diode Rings. The diode ring is normally driven with two signals with sinusoidal waveform and also is some sort of mixer. Here it will suffice to derive the gain characteristic K_θ of the device. If the input signal is $\theta_i = A_i \sin \omega_o t$, and the reference signal is $\theta_r = A_r \sin(\omega_o t + \phi)$, where ϕ is the phase difference between the two signals, the output signal θ_e is

$$\theta_e = \theta_i \theta_r = \frac{A_i A_r}{2} K \cos \phi - \frac{A_i A_r}{2} K \cos(2\omega_o t + \phi) \tag{6-8}$$

where K is the mixer gain. One of the primary functions of the low-pass filter is to eliminate the second-harmonic term before it reaches the VCO. The second harmonic will be assumed to be filtered out and only the first term will be considered, so

$$\theta_e = \frac{A_t A_r}{2} K \cos \phi \tag{6-9}$$

When the error signal is zero, $\phi = \pi/2$. Thus, the error signal is proportional to phase differences from 90°. For small changes in phase $\Delta\phi$,

$$\theta_e \cong \frac{\pi}{2} + \Delta\phi = \frac{A_t A_r}{2} K \left[\cos\left(\frac{\pi}{2} + \Delta\phi \right) \right]$$

$$= \frac{A_t A_r}{2} K \sin \Delta\phi \tag{6-10}$$

For a small phase perturbation $\Delta\phi$,

$$\theta_e \cong \frac{A_t A_r K}{2} \Delta\phi \tag{6-11}$$

The phase-detector output was assumed to be

$$\theta_e = K(\theta_i - \theta_o) \tag{6-12}$$

and the phase detector scale factor K_θ is given by

$$K_\theta = \frac{A_t A_r K}{2} \tag{6-13}$$

The phase-detector scale factor K_θ depends on the input signal amplitudes; the device can be considered linear only for constant-amplitude input signals and for small deviations in phase. For larger deviations in phase,

$$\theta_e = K_\theta \sin \Delta\phi \tag{6-14}$$

which describes a nonlinear relation between θ_e and ϕ.

In frequency synthesizers, the reference is typically generated from a reference oscillator and is lower than the VCO frequency, which is divided by a programmable divider. Both signals, therefore, are square waves rather than sine waves, and theoretically, a diode ring can be driven by those two signals.

A drawback to the diode-ring phase detector is that its output voltage is very small—several hundred millivolts at most. A postdetector dc amplifier, which will unavoidably introduce noise, is therefore required.

Edge-Triggered JK Master–Slave Flip-Flops. The fundamental idea of the sequential phase comparator we will be dealing with is that there are two outputs available, one to charge and one to discharge a capacitor. Output 1 then is high if the Signal 1 frequency is

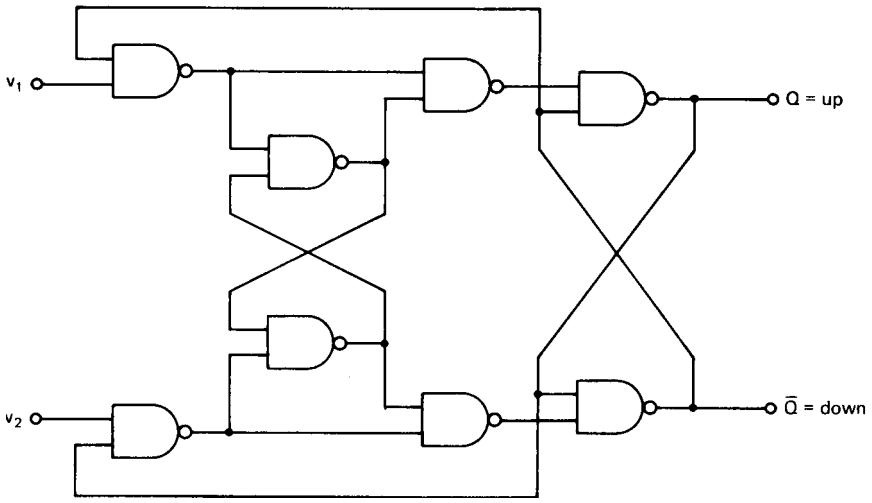


Figure 6-4 Edge-triggered JK master–slave flip-flop.

greater than the Signal 2 frequency; or if the two frequencies are equal, if Signal 1 leads Signal 2 in phase.

Output 2 is high if the frequency of Signal 2 is greater than that of Signal 1, or if the signal frequencies are the same and Signal 2 leads Signal 1 in phase.

Figure 6-4 shows the minimum configuration to build such a phase comparator. It can be operated from -2π to $+2\pi$, and an active amplifier is recommended as a charge pump. The Q output of the JK master–slave flip-flop is set to one by the negative edge of Signal 1, while the negative edge of Signal 2 resets it to zero. Therefore, the output $\overline{Q}$ is the complement of Q. The output voltage $\overline{V}$ is defined as the weighted duty cycle of Q and $\overline{Q}$ This means that a positive contribution is made when $Q = 1$ and a negative contribution (discharge) is made when $Q = 0$. The averaging and filtering of the unwanted ac component is done by a subsequent integrator. The integrator then is called a *charge pump*, as the loop capacitor is being charged and discharged depending on whether Q is high or low.

If the system using the JK flip-flop is not in lock, and there is a large difference between frequencies f_1 and f_2 at the output, the output will not be zero, but instead will be positive or negative relative to one-half the supply voltage. This is an advantage and indicates that this system is frequency sensitive. We therefore call it a *phase/frequency comparator* because it is capable of detecting both phase and frequency offsets. In its locking and pull-in performance, it is similar to an exclusive-OR gate.

For better understanding, let us look at a few cases where the system is in lock. It should be noted that whereas the exclusive-OR gate is sensitive to the duty cycle of the input signals, the JK flip-flop responds only to the edges, and therefore the phase/frequency comparator can be used for asymmetrical waveforms. Let us assume first that the input Signals 1 and 2 have the same frequency. Figure 6-5 shows what happens if the phase error is about 0, π, and 2π. In those cases, the duty cycle at the output is about 0%, 50%, or 100%, respectively. The narrow output pulses may cause spikes on the power-supply line and lines in the vicinity, and precautions must be taken to filter them.

The output voltage $\overline{V}$ is the average of the signal Q and is a linear function of the phase error.

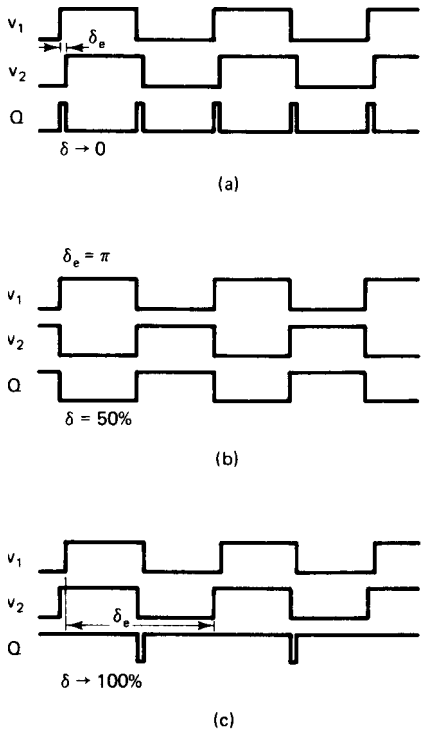


Figure 6-5 Performance of the JK phase/frequency comparator for different input signals.

Now let us take a look at several cases where the system is not in lock. Figure 6-6 shows the case where f_1 is substantially higher than f_2. As a result, the output duty cycle is close to 100% and the VCO frequency is being pulled up to higher frequencies. If the frequency at Input 2 is much higher than that at Input 1, the opposite is true. This proves that the device is sensitive to frequency changes.

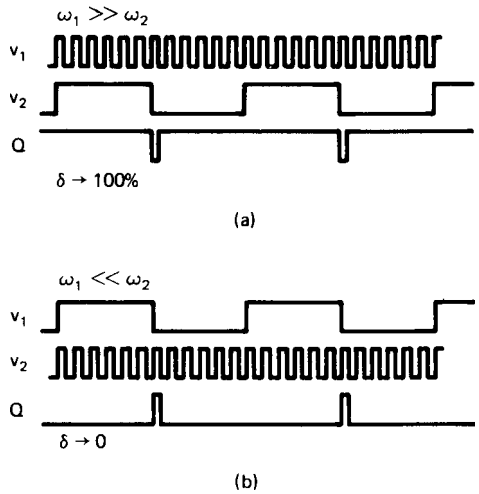


Figure 6-6 Phase-detector output for two input frequencies that are substantially different.

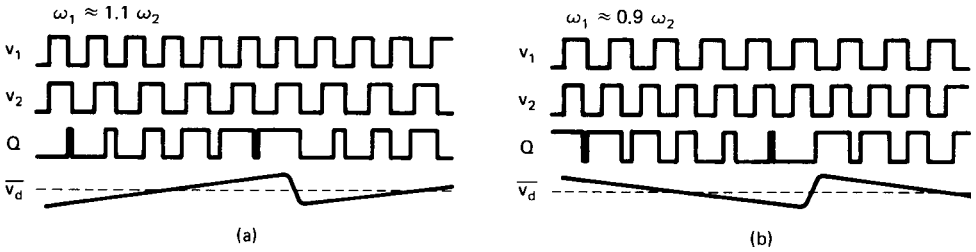


Figure 6-7 Performance of the phase detector for small frequency differences.

In cases where both frequencies are about the same, as shown in Figure 6-7, the crossover area is not clearly defined. The first picture shows the case where f_2 is 10% higher than f_1 and the duty cycle is changing periodically between 0% and 100%. Therefore, the ac voltages look like a sawtooth, with a rate equal to the difference of both frequencies. The same holds true if the two inputs are reversed. In the case where both frequencies are identical, the JK flip-flop behaves the same as an exclusive-OR gate. From this discussion, it can be concluded that, while this phase/frequency comparator was included to explain how it works, it is not a very desirable device for practical purposes because of the uncertainty of its behavior close to lock.

Digital Tristate Comparators. The digital tristate phase/frequency comparator is probably the most universally used and most important next to the sample/hold comparator. Although the diode-ring and exclusive-OR gate phase detectors have some applications, the tristate phase/frequency comparator can be used widely. Even in cases where a sample/hold comparator theoretically could be used, it may be inferior as far as reference attenuation or noise is concerned, but it is generally well-behaved. Unfortunately, the tristate system is very complex and shows a number of unusual phenomena. Such a digital tristate comparator is shown in Figure 6-8 using two D flip-flops and a NAND gate. The Q_2 output signal is filtered with the low-pass filter. The operation of this logic circuit is readily analyzed using the state-transition diagram, as shown in Figure 6-9. The D flip-flop outputs go high on the leading edge of their respective clock inputs and remain high until they are reset. The reset

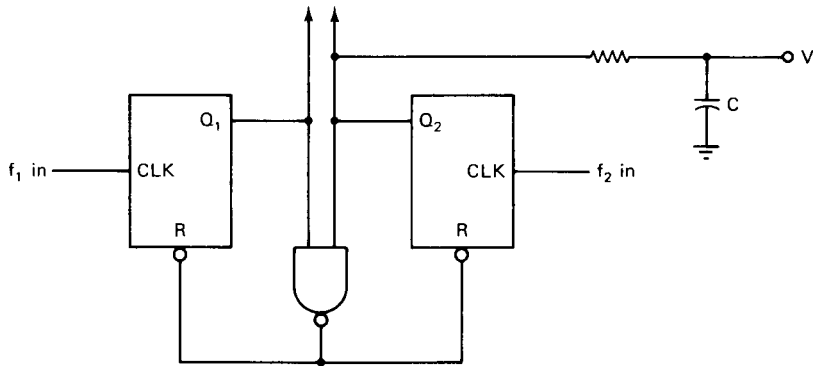


Figure 6-8 Phase detector with two D flip-flops and a NAND gate. In this book, this type of phase detector will be called a tristate comparator.

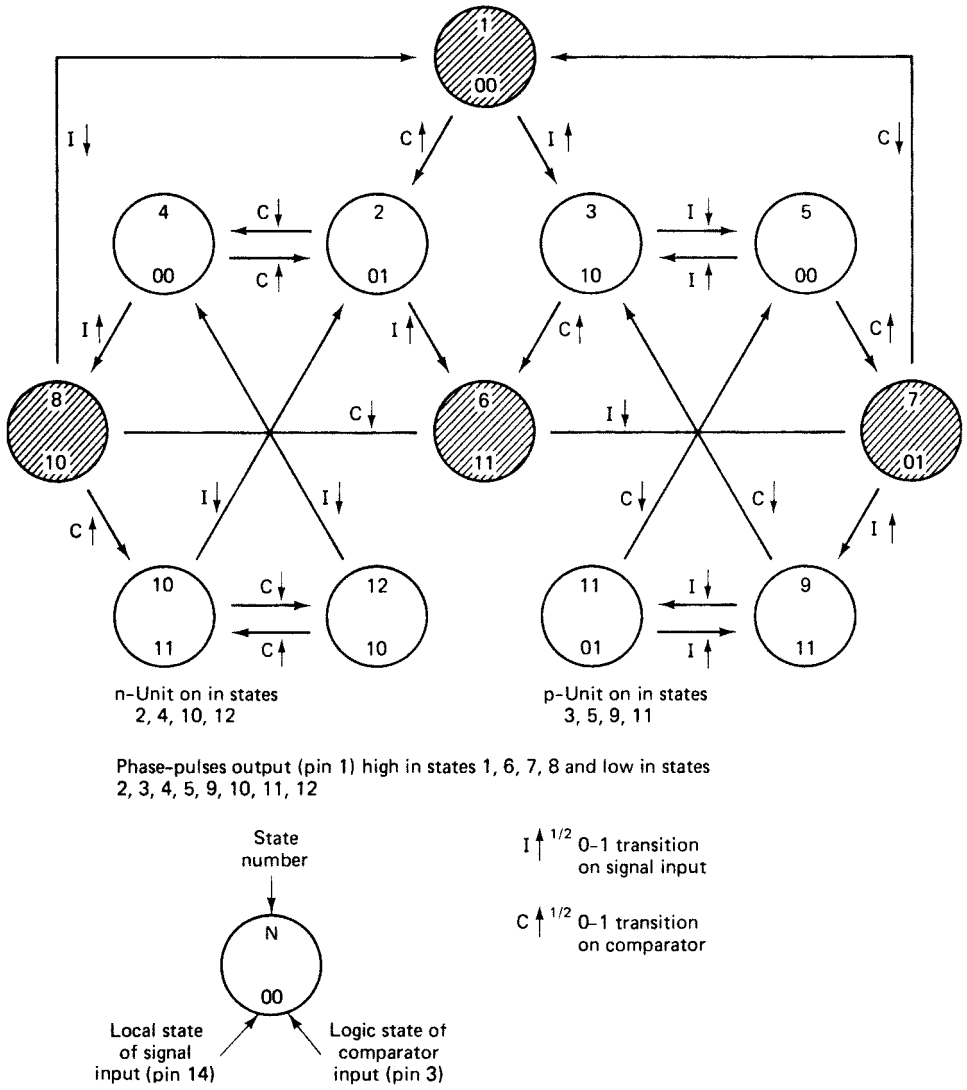


Figure 6-9 Logic diagram of the tristate detector.

signal occurs when both inputs are high. When both signals are in phase and of the same frequency, both outputs will remain low, and no signal will be applied to the operational amplifier. When the two frequencies are the same, the dc output voltage transfer characteristic will be as shown in Figure 6-10. If the two signal frequencies are not the same, the output voltage will depend on both the relative frequency difference and the phase difference. The timing diagram of Figure 6-11 illustrates the case in which $f_2 = 3f_1$. In part (a) of the figure, the leading edge of f_1 occurs just after that of f_2, so that Q_2 is high 50% of the time, and the average value of the PD output is 50%. In Figure 6-11b, the leading edge of f_1 occurs just before that of f_2, so Q_2 is high almost all the time and the average output voltage is approximately V.

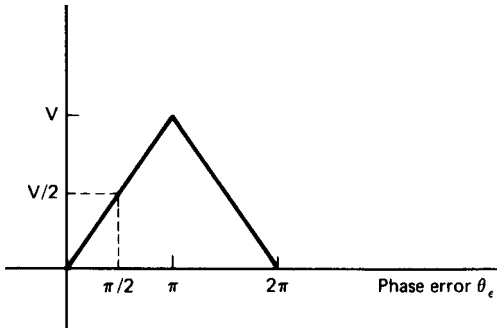


Figure 6-10 Transfer characteristic of the tristate phase/frequency comparator.

The output voltage averaged over all of the phase differences is then 67% for $f_2 = 3f_1$. In general, it can be said that the average output (averaged over all the phase differences) is given by

$$V_{\text{ave}} = 1 - \frac{f_1}{f_2} V \tag{6-15}$$

provided that f_2 is greater than f_1. This expression is plotted in Figure 6-12 together with the cases in which f_1 is greater than f_2.

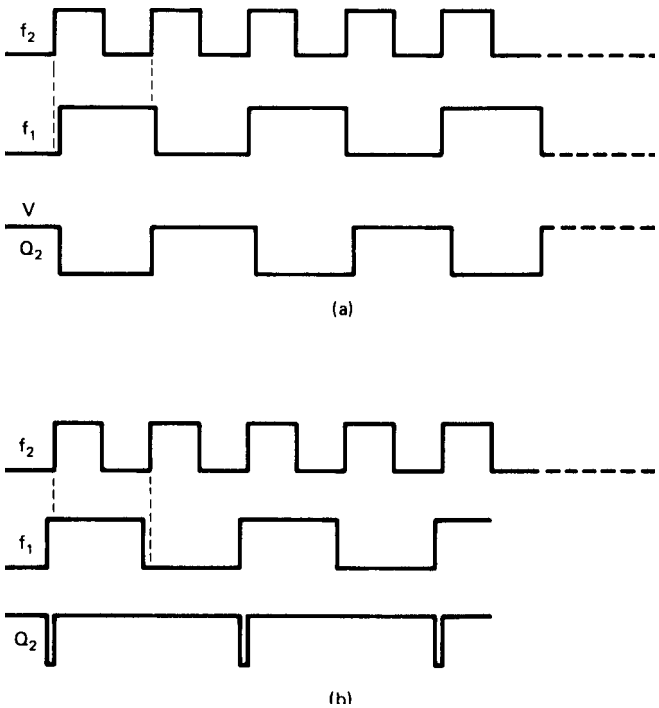


Figure 6-11 Output waveform of the tristate frequency comparator for different input frequencies.

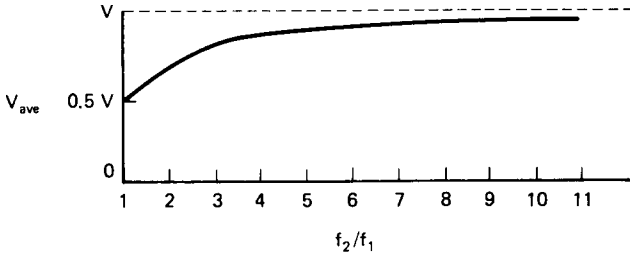


Figure 6-12 Average output voltage as a function of frequency ratio.

The digital network used in this realization is only one of a large number of logic circuits that could be used. Many IC manufacturers now produce a *quad-D* circuit that functions much like the dual-D flip-flop; the main difference is that when the frequency of one signal is more than twice that of the other signal, the corresponding output will be high all of the time. Therefore, a larger voltage is applied to the VCO and the loop response is faster. An example of the quad-D circuit is shown in Figure 6-13.

The most popular digital tristate phase/frequency comparator on the market is the one used in the CD4046 PLL IC, shown in Figure 6-14. It contains an additional phase comparator, an exclusive-OR gate that can be used as a lock indicator. In addition, two FETs are used to sum the two outputs. A slightly faster version in TTL technique is the Motorola MC4044. The fastest version in ECL is the MC12040, also made by Motorola, shown in Figure 6-15. Sometimes it is convenient to build the phase/frequency comparator in discrete technique to add additional features. Figure 6-16 shows an example.

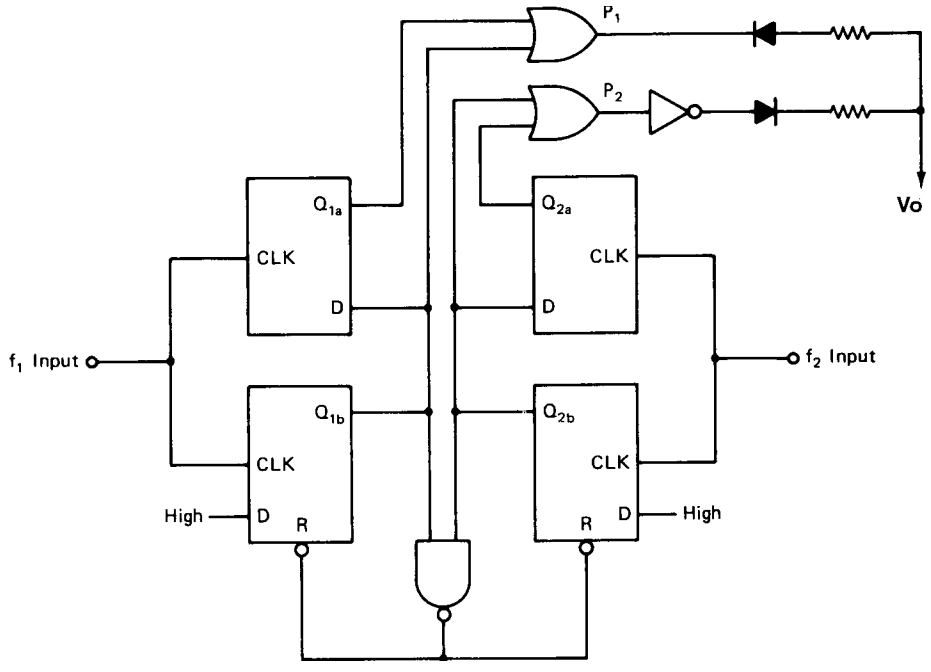


Figure 6-13 Example of a quad-D circuit.

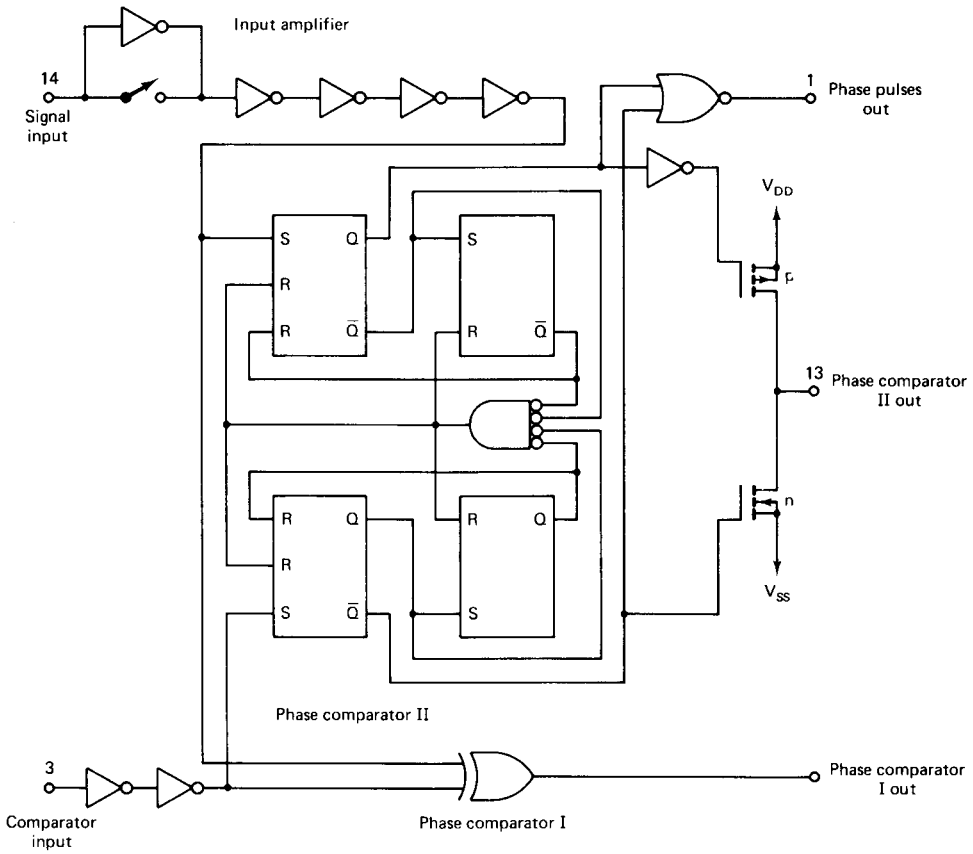


Figure 6-14 Block diagram of CD4046 phase/frequency comparator. (Courtesy of Motorola Semiconductor Products, Inc.)

This particular tristate phase/frequency comparator has a peculiarity that was first mentioned by Egan and Clark [2]. When actually building a phase-locked loop with this phase/frequency comparator, or that of the CD4046 type, by going through the normal mathematical design routine, it becomes apparent that the expected performance and the actual results differ as follows:

1. The reference suppression will be better than expected.
2. The phase error or tracking will be worse than expected.
3. The phase margin will differ and the system may not lock despite the fact that the calculation is correct.

The reason for all this is due to two effects:

1. The flip-flops are not absolutely alike, and as a result of this, the output in the crossover region is not zero.
2. If there is no or very little correction voltage required, the gain of the phase detector will drop substantially.

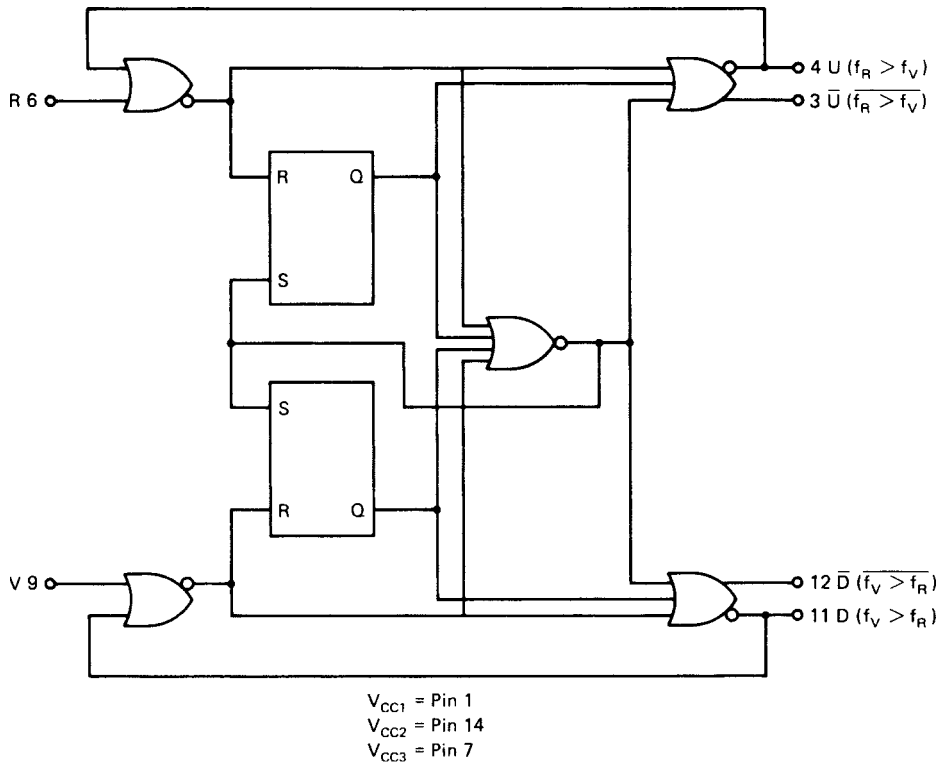


Figure 6-15 Block diagram of Motorola MC12040 phase/frequency comparator. (Courtesy of Motorola Semiconductor Products, Inc.)

Let us assume the ideal situation where the output of the phase/frequency comparator feeding the charge pump does not have to correct any error, the system is drift free, and there are no leakage currents. The holding capacitor of the charge pump would maintain constant voltage and, as there is no drift, no correction voltage would be necessary.

The flip-flops, however, introduce a certain amount of jitter, and a certain amount of jitter is also introduced by the frequency dividers, both the reference divider and the programmable divider. This jitter results in an uncertainty regarding the zero crossings, and extremely narrow pulses will appear at the output of the summation amplifier used in the CD4046.

Under the ideal assumption that there are no corrections required and those pulses would not exist, the reference suppression would be infinite, as there is no output. Therefore, the reference suppression—disregarding the effect of the loop filter—depends only on how well this condition is met.

The change of gain seems somewhat surprising, but as we think of it, if there is no correction and no update, there is also no gain. It is impossible to meet this condition, which is fortunate, but with regard to the temperature stability and aging characteristics of some devices, predicting actual performance may be difficult.

There are several remedies to this problem. A simple version is to introduce a controlled amount of leakage. While the electrolytic capacitor required in the charge pump will have some leakage, it is better to set a leakage current that is independent of temperature and

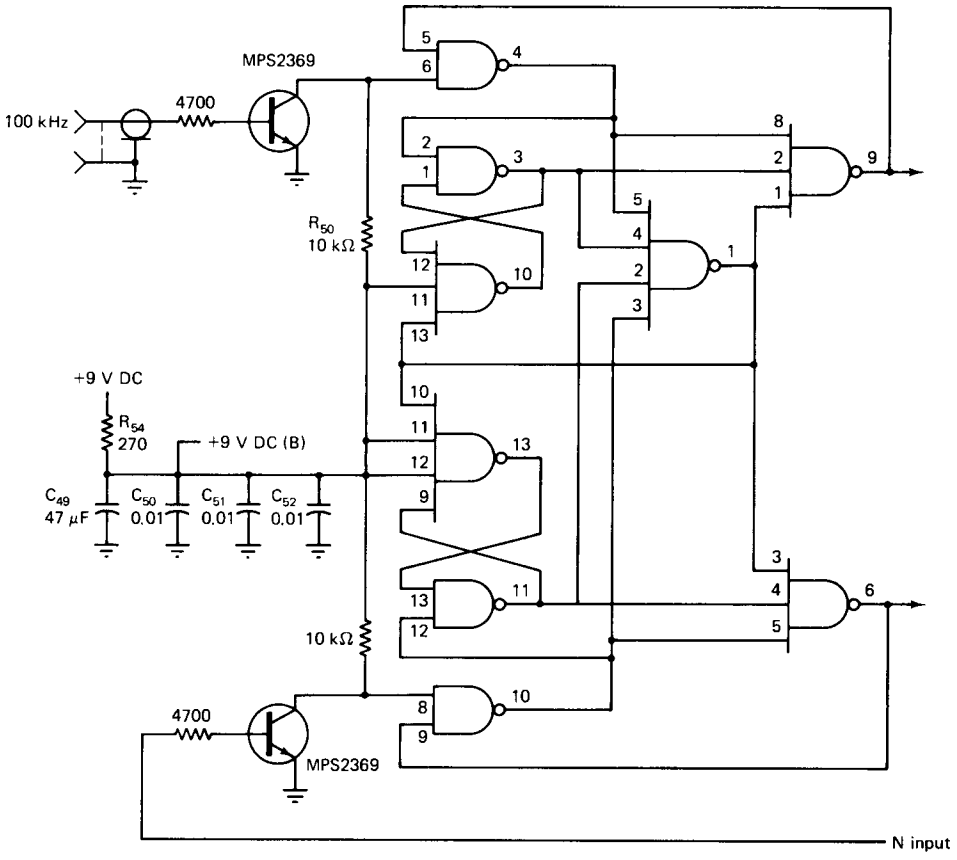


Figure 6-16 Possible version of tristate phase/frequency comparator.

aging. This can be accomplished by putting a 1-MΩ resistor from the output of the CD4046 to ground. The phase/frequency comparator then has to deliver an output current, and this output current is determined by a resistor that can be independent of temperature and other effects. As a result of this, the duty cycle of the output pulses of the phase/frequency comparator will change and the pulses will become wider. The wider the pulse, the more energy it contains; therefore, the wider the pulse, the more the reference suppression degrades.

It is theoretically possible to connect one side of the 1-MΩ resistor, instead of to ground, to the wiper of a potentiometer, and set the voltage in such a manner that this offset is compensated; but again, since the phase will shift theoretically, one must adjust the potentiometer according to the actual phase error. This is not a very convenient arrangement.

A somewhat better method was proposed by Fairchild several years ago, and the hardware was possibly realized in newer ICs. It was proposed to insert a gate in one of the output arms of the phase/frequency comparator before the signal is fed to the summation amplifier and a periodic current disturbance introduced. This disturbance has the same rate as the reference frequency and is of extremely short duration, such that the output contains only fairly high harmonics of the reference, which is easily filtered as it contains very little energy. This periodic disturbance upsets the output of the phase/frequency comparator and has an effect

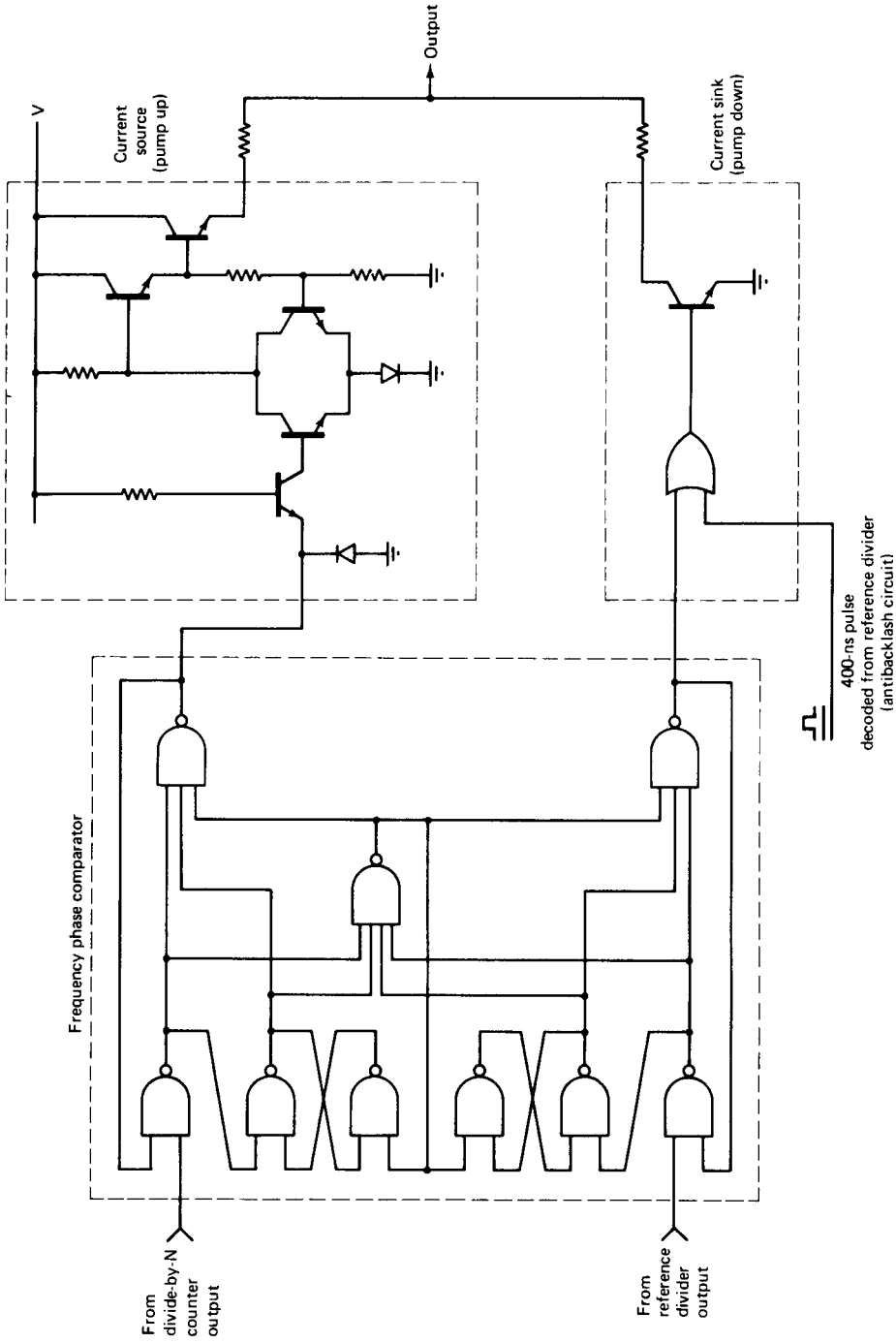


Figure 6-17 Tristate detector with antibacklash circuit included.

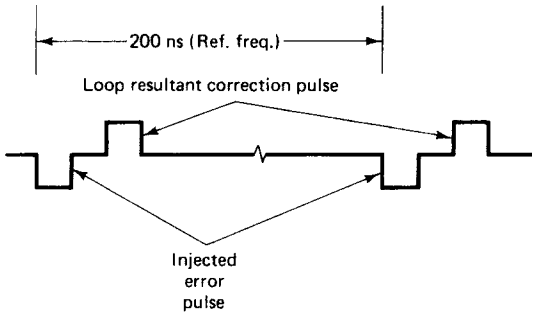


Figure 6-18 Output of phase/frequency detector with antibacklash circuit.

similar to that of a leakage resistor. The advantage of this method, however, is that this is done at a fairly high frequency and does not introduce low-frequency noise, which the 1-MΩ resistor does.

Figure 6-17 shows the circuit that accomplishes this, and Figure 6-18 shows the effect on the output pulses. The charge pump output exhibits a short negative-going pulse followed immediately by a short positive-going pulse. This can also be called an *antibacklash* feature, and it prevents operating in the dead zone. (This zone is not really a dead zone because of leakage currents in the tuning diode.) The duration and proximity of these pulses are such that they cause no net change in the charge of the integrator. Figure 6-19 shows the response of a phase/frequency detector near loop lock, including the dead zone; this may not be true for ECL.

6-2-3 Filters for Phase Detectors Providing Voltage Output

Figure 6-20 shows the passive *RC* filter for the second-order loop typically used in PLL synthesizers. The transfer characteristic of the filter is

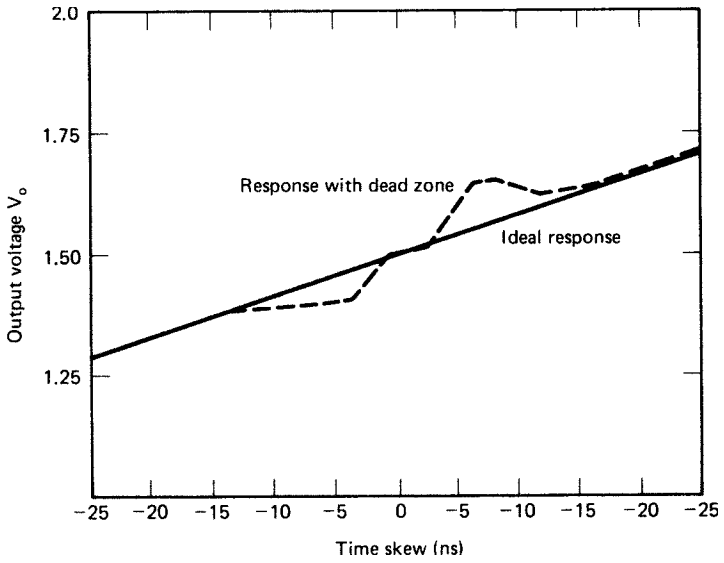


Figure 6-19 Response of phase/frequency detector near loop lock, resulting in a dead zone.

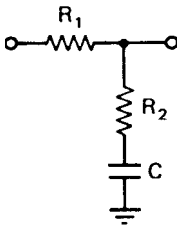


Figure 6-20 Schematic diagram of a typical passive *RC* filter.

$$\frac{V_0(s)}{V_I(s)} = \frac{1 + s\tau_2}{1 + s(\tau_1 + \tau_2)} \tag{6-16}$$

where $\tau_1 = R_1C$ and $\tau_2 = R_2C$.

Figure 6-21 shows the schematic for the active filter for the second-order loop. Its transfer characteristic is

$$\frac{V_0(s)}{V_I(s)} = \frac{1 + s\tau_2}{s\tau_1} \tag{6-17}$$

where $\tau_1 = R_1C$ and $\tau_2 = R_2C$.

If only one active integrator is used, we have a *Type 1* PLL. If two integrators are used, as in building an active filter, we have a *Type 2 second-order* loop. Here, *second-order* refers to the denominator polynomial of the transfer function. If we insert a simple low-pass filter such as the one shown in Figure 6-20, but with $R_2 = 0$, we obtain

$$F(s) = \frac{1}{1 + s\tau} \tag{6-18}$$

If we let $K = K_0K_\theta/N$, the transfer function $B(s)$ becomes

$$B(s) = \frac{N}{s^2/\omega_n^2 + 2\zeta s/\omega_n + 1} \tag{6-19}$$

where $\omega_n = \sqrt{K/\tau}$ and $2\zeta = \omega_n/K = \sqrt{1/K\tau}$. Here, ζ is the damping factor of the loop and ω_n is the natural frequency.

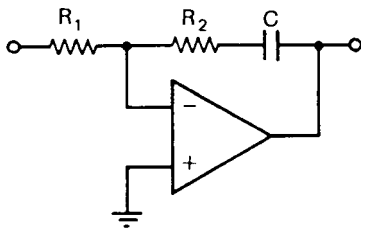


Figure 6-21 Schematic diagram of an active filter for a second-order loop.

The frequency response of the second-order transfer function is determined by ζ. For $\zeta = 0.707$, the transfer function becomes the second-order maximally flat, or Butterworth, response. For values of $\zeta < 0.707$, the gain exhibits peaking in the frequency domain. The maximum value of the frequency response can be found by setting the derivative of its maximum to zero. The frequency at which the maximum occurs is

$$\omega_p = \omega_n \sqrt{1 - 2\zeta^2} \tag{6-20}$$

The 3-dB bandwidth B is found to be

$$B = f_n[1 - 2\zeta^2 + (2 - 4\zeta^2 + 4\zeta^4)^{1/2}]^{1/2} \tag{6-21}$$

where $f_n = \omega_n/2\pi$.

The time required for the output to rise from 10% to 90% of its final value is the rise time t_r. It is approximately related to the system bandwidth by the relation

$$t_r = \frac{2.2}{B} \tag{6-22}$$

The *RC* time constant of this simple filter determines both the natural loop frequency and the damping factor ζ. To improve the performance of the filter, we need more flexibility. When the series resistor R_2 is not zero, we obtain the original *RC* filter of Figure 6-20. The transfer function of this filter is

$$B(s) = \frac{N[s\omega_n(2\zeta - \omega_n/K) + \omega_n^2]}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{6-23}$$

where $\omega_n = \sqrt{K/\tau}$ and $2\zeta = (1 + K\tau_2)/\sqrt{K\tau}$ and τ is written for $\tau_1 + \tau_2$.

The determination of the 3-dB bandwidth for this general Type 1 second-order loop is somewhat more complex than the earlier computation, but after calculation, we obtain

$$B = f_n[a + (a^2 + 1)^{1/2}]^{1/2} \tag{6-24}$$

where we have written

$$a = 2\zeta^2 + 1 - \frac{\omega_n(4\zeta - \omega_n/K)}{K} \tag{6-25}$$

The noise bandwidth of the Type 1 second-order loop is

$$B_n = \pi f_n \left(\zeta + \frac{1}{4\zeta}\right) \tag{6-26}$$

In the case of the active filter, where we have two integrators, the closed-loop transfer function of the Type 2 second-order PLL with a perfect integrator is

$$B(s) = \frac{N(2\zeta\omega_n s + \omega_n^2)}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{6-27}$$

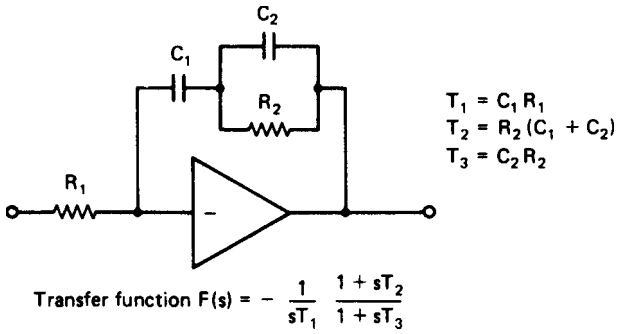


Figure 6-22 Schematic diagram of an active filter for a third-order loop.

where $\omega_n = (KR_2/\tau_2 R_1)^{1/2}$, $2\zeta = (K\tau_2 R_2/R_1)^{1/2}$, and $K = K_\theta K_0/N$, as usual. The 3-dB bandwidth of the Type 2 second-order filter is

$$B = f_n \left\{ 2\zeta^2 + 1 + \left[(2\zeta^2 + 1)^2 + 1 \right]^{1/2} \right\}^{1/2} \tag{6-28}$$

and the noise bandwidth is

$$B_n = \frac{(KR_2/R_1) + (1/\tau_2)}{4} \tag{6-29}$$

The Type 2 third-order loop is defined by the active integrator, as shown in Figure 6-22. The additional capacitor across the second resistor increases suppression of the reference frequency. The advantage of the higher-order loop is that, for the same loop bandwidth, it offers more reference-frequency suppression than the second-order loop. Conversely, for the same suppression it offers a faster lock-in time. More details are given in Rohde [1].

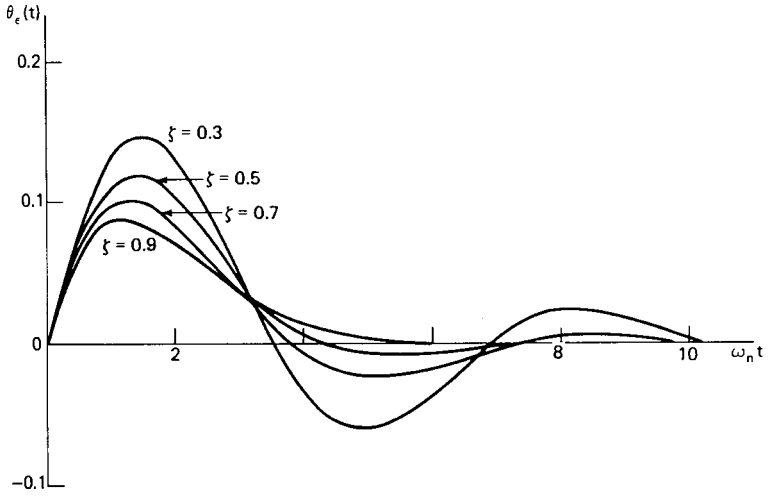


Figure 6-23 The error response of a Type 1 second-order PLL to unit-step change in frequency for various damping ratios ζ with K constant. The steady-state error $2\zeta/\omega = 1/K$.

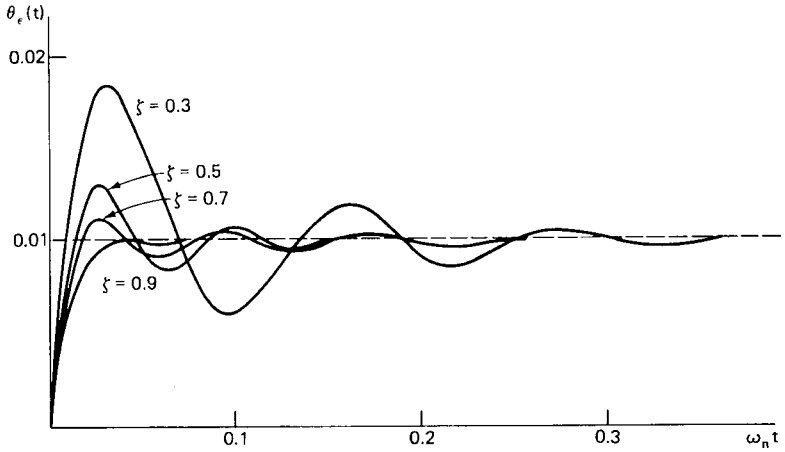


Figure 6-24 The error response of a Type 2 second-order PLL to unit-step change in frequency for various damping ratios ζ with ω_n constant. The steady-state error is zero.

Transient Response. The Laplace transform can be used to calculate the response of the PLL to a change in frequency. Figure 6-23 shows the normalized output response of the Type 1 second-order loop, and Figure 6-24 shows the normalized output response of the Type 2 second-order loop. We determine from both functions that a damping ratio of 0.707 will produce a peak overshoot of less than 10% for the Type 1 second-order loop and of less than 20% for the Type 2 second-order loop when $\omega_n t \geq 4.5$. The settling time is therefore determined to be $t_s = 4.5/\omega_n$. For more details on the actual design of synthesizer loops, the reader should refer to Rohde [1].

The Philips UMA1018M is an example of the synthesizer implementations now available in IC form. Designed for use in portable radiotelephones, the UMA1018M contains two frequency synthesizers. One, intended for first-LO use, can operate at input frequencies from 50 to 1.25 GHz; the second, intended for use as the second or "IF" LO in a double-conversion system, can operate at input frequencies from 20 to 300 MHz. Both loops use the same reference signal (3–40 MHz), which must be supplied by an external crystal oscillator. Figure 6-25 shows the UMA1018M's block diagram; Figure 6-26 shows the block diagram of a sample UMA1018M application; and Figure 6-27 shows a UMA1018M application schematic. Finally, Figures 6-28 and 6-29 show a UMA1018M-based system's close-in phase noise and reference suppression, respectively.

6-2-4 Charge-Pump-Based Phase-Locked Loops*

The basic drawback of the conventional phase/frequency detector, expressed in V/rad, is its dead-zone phenomenon. The loop gain is really determined by the linearity of the phase detector. Its can drop to zero when no correction is needed. One way around this had been to provide an external resistance, connected between the phase-detector output and common, that draws current but reduces the reference suppression. A more modern way to overcome

*Portions of this section are based on information contained in the National Semiconductor datasheet "LMX1501A/LMX1511 PLLatinum™ 1.1 GHz Frequency Synthesizer for RF Personal Communications," November 1995. Used with permission.

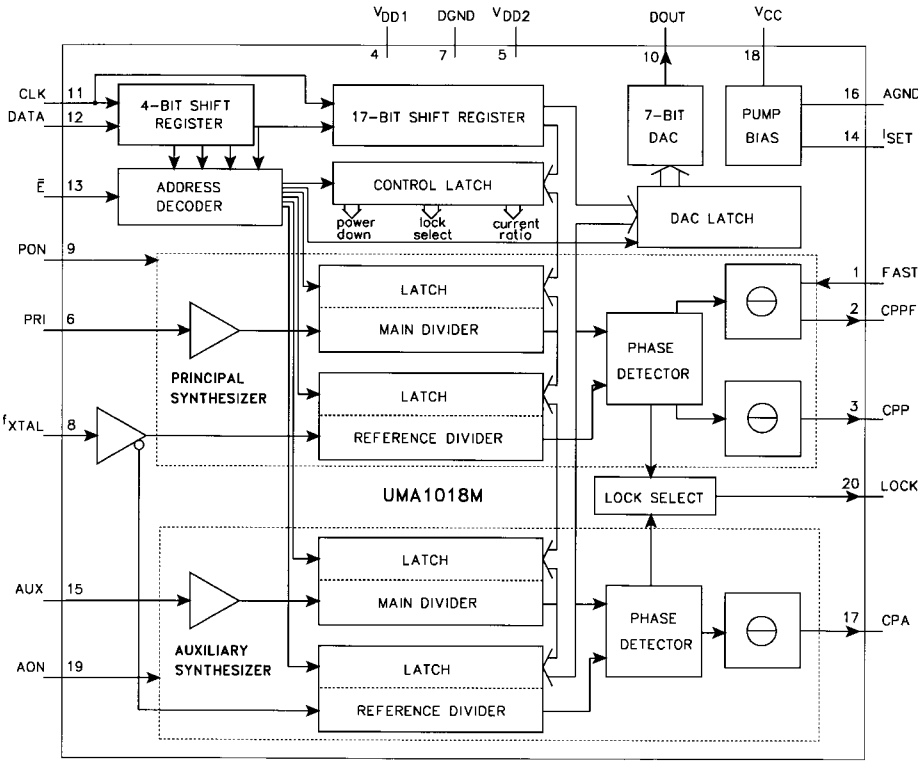


Figure 6-25 Block diagram of the Philips UMA1018M dual-synthesizer chip.

this problem is to resort to a charge pump. Figure 6-30 uses a CMOS-based charge pump, with the resistor R_L limiting the available current.

If more control over the actual current is needed, here is a recommendation by National Semiconductor that improves the flexibility.

External Charge Pump. Figure 6-31 shows one possible architecture for an external charge-pump current source. The signals ϕ_p and ϕ_r in the diagram correspond to the phase-detector outputs of the National LMX1501/1511 frequency synthesizers. These logic signals are converted into current pulses, using the circuitry shown in Figure 6-31, to enable either charging or discharging of the loop filter components to control the PLL's output frequency.

Referring to Figure 6-31, the design goal is to generate a 5-mA current that is relatively constant to within 5 V of the power-supply rail. To accomplish this, it is important to establish as large a voltage drop across R_5 and R_8 as possible without saturating Q_2 and Q_4. Approximately 300 mV provides a good compromise, allowing the current source reference generated to be relatively repeatable in the absence of good Q_1–Q_2, Q_3–Q_4 matching (matched transistor pairs are recommended.) The ϕ_p and ϕ_r outputs are rated for a maximum load current of 1 mA, while 5-mA current sources are desired. The voltages developed across R_4 and R_9 will consequently be approximately 258 mV, or 42 mV < R_8–R_5, due to the current density differences [0.026 × ln(5 mA/1 mA)] through the Q_1–Q_2, Q_3–Q_4 pairs.

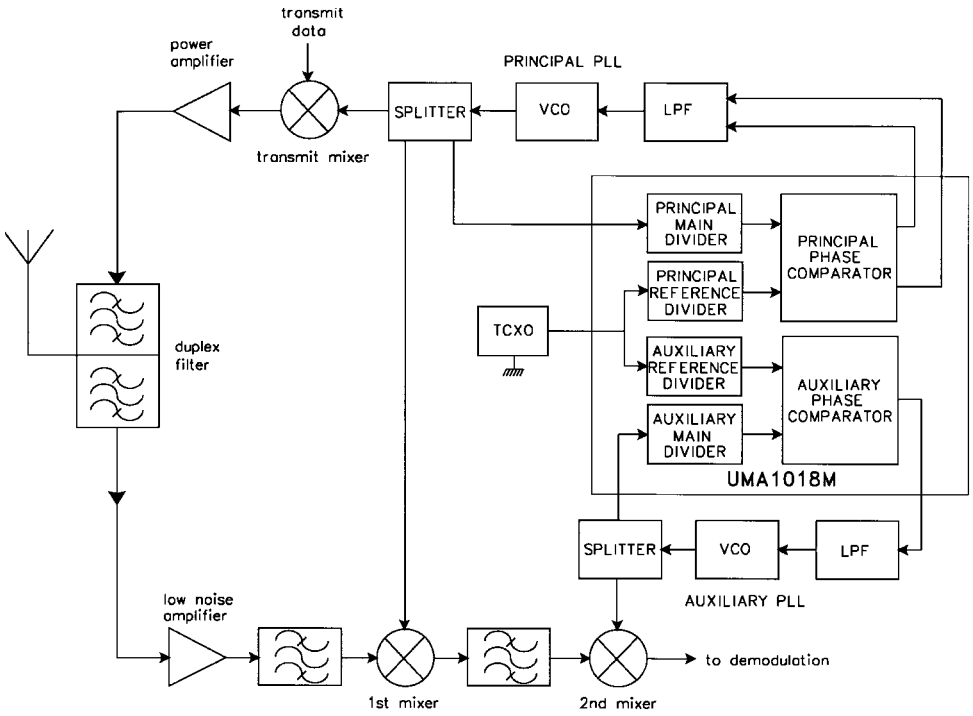


Figure 6-26 Block diagram of a typical UMA1018M application.

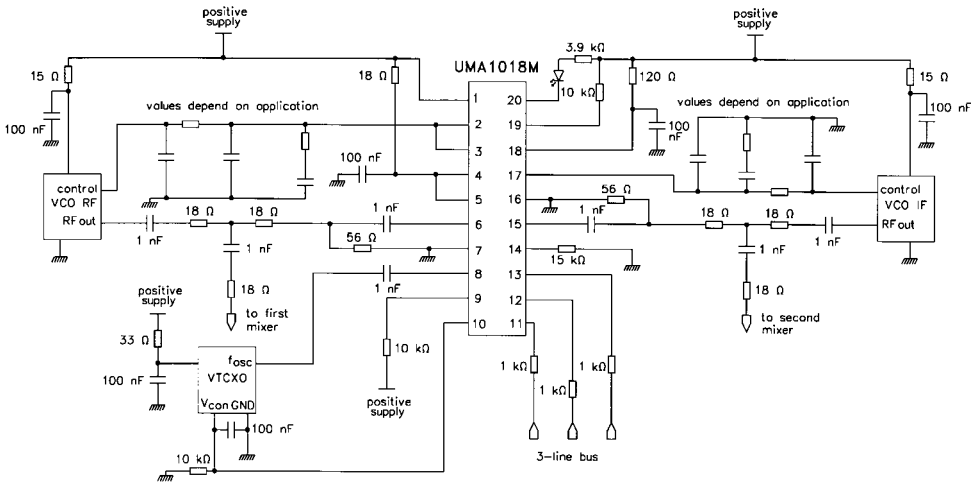


Figure 6-27 Schematic of a typical UMA1018M application.

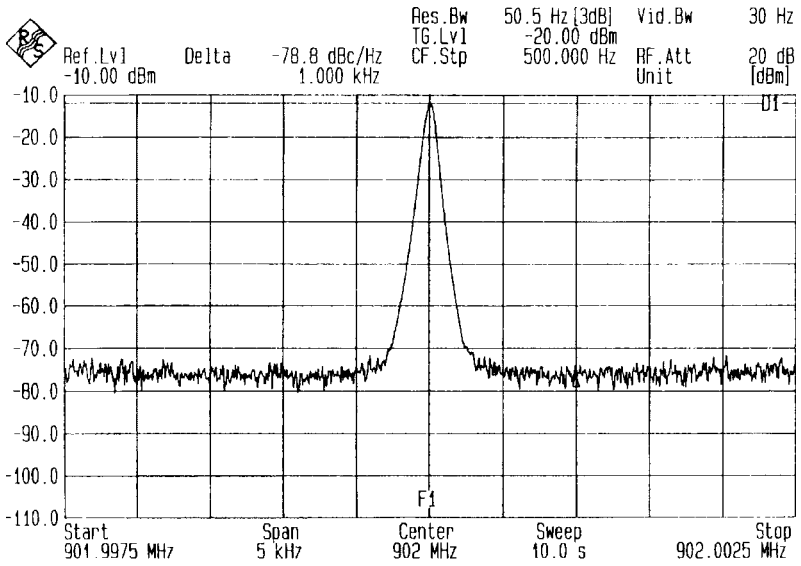


Figure 6-28 Close-in noise of a UMA1018M-based system (principal synthesizer).

To calculate the value of R_7, it is necessary to first estimate the forward base-to-emitter voltage drop (V_{fn}, V_{fp}) of the transistors used, the V_{OL} drop of ϕ_p, and the V_{OH} drop of ϕ_r under 1-mA loads (ϕ_p's $V_{OL} < 0.1$V and ϕ_r's $V_{OH} < 0.1$V).

Knowing these parameters along with the desired current allows us to design a simple external charge pump. Separating the pump-up and pump-down circuits facilitates the nodal analysis and gives the following equations:

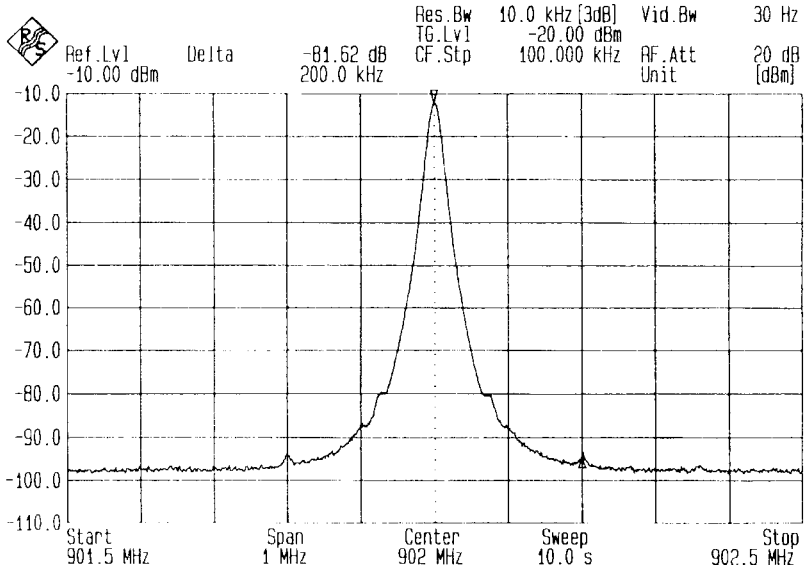


Figure 6-29 Wideband output spectrum of a UMA1018M-based system (principal synthesizer) showing reference-frequency breakthrough.

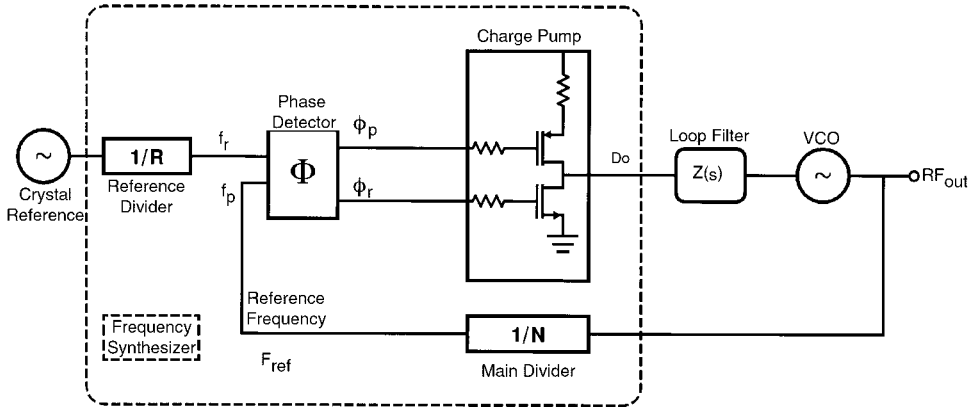


Figure 6-30 A basic charge-pump PLL.

$$R_4 = \frac{V_{R5} - V_T \times \ln(i_{\text{source}}/i_{p\,\max})}{i_{\text{source}}} \tag{6-30}$$

$$R_9 = \frac{V_{R8} - V_T \times \ln(i_{\text{sink}}/i_{n\,\max})}{i_{\text{sink}}} \tag{6-31}$$

$$R_5 = \frac{V_{R5} \times (\beta_p + 1)}{i_{p\,\max} \times (\beta_p + 1) - i_{\text{source}}} \tag{6-32}$$

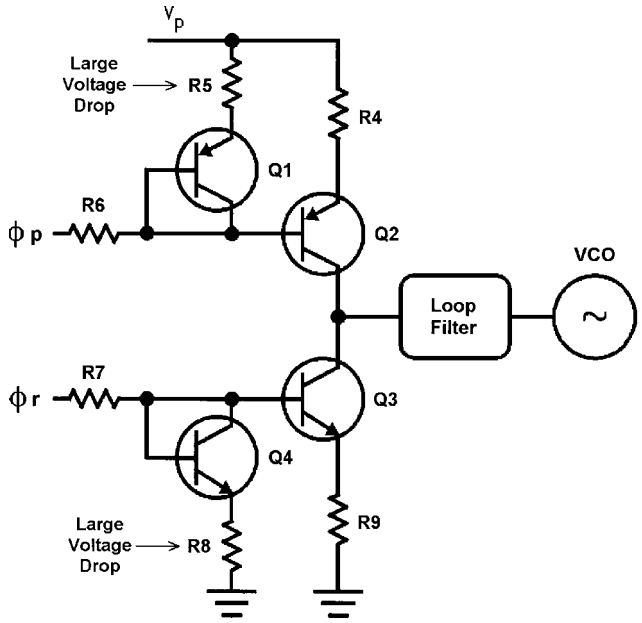


Figure 6-31 External charge-pump current source.

$$R_8 = \frac{V_{R8} \times (\beta_n + 1)}{i_{r\,\max} \times (\beta_n + 1) - i_{\text{source}}} \tag{6-33}$$

$$R_6 = \frac{(V_p - V_{VOL\phi_p}) - (V_{R5} + V_{fp})}{i_{p\,\max}} \tag{6-34}$$

$$R_7 = \frac{(V_p - V_{VOH\phi_r}) - (V_{R8} + V_{fn})}{i_{\max}} \tag{6-35}$$

Example

Typical Device Parameters: $\beta_n = 100$, $\beta_p = 50$.
Typical System Parameters: $V_p = 5.0$ V; $V_{cntl} = 0.5$ V – 4.5 V; $V_{\phi p} = 0.0$ V; $V_{\phi r} = 5.0$ V.
Design Parameters: $I_{\text{sink}} = I_{\text{source}} = 5.0$ mA; $V_{fn} = V_{fp} = 0.8$ V; $I_{r\,\max} = I_{p\,\max} = 1$mA; $V_{R8} = V_{R5} = 0.3$ V; $V_{OL\phi_p} = V_{OH\phi_r} = 100$mV.

Therefore, select

$$R_4 = R_9 = \frac{0.3\text{ V} - 0.026 \times \ln(5.0\text{ mA} \div 1.0\text{ mA})}{5\text{ mA}} = 51.6\ \Omega \tag{6-36}$$

$$R_5 = \frac{0.3\text{ V} \times (50 + 1)}{1.0\text{ mA} \times (50 + 1) - 5.0\text{ mA}} = 332\ \Omega \tag{6-37}$$

$$R_8 = \frac{0.3\text{ V} \times (100 - 1)}{1.0\text{ mA} \times (100 + 1) - 5.0\text{ mA}} = 315.6\ \Omega \tag{6-38}$$

$$R_6 = R_7 = \frac{(5\text{ V} - 0.1\text{ V}) - (0.3\text{ V} + 0.8\text{ V})}{1.0\text{ mA}} = 3.8\ \text{k}\Omega \tag{6-39}$$

Design Example: A Passive PLL Filter. Figure 6-32 shows the loop filter topology we will use for this example. The advantage of a passive filter as opposed to an active filter is that the filter itself introduces no noise, while the active filter using an op-amp can frequently cause more harm than good because of its noise contribution.

The time constants that determine the pole and zero frequencies of the filter shown in Figure 6-32 are

$$\tau_1 = R_2 \frac{C_1 C_2}{C_1 + C_2} \tag{6-40}$$

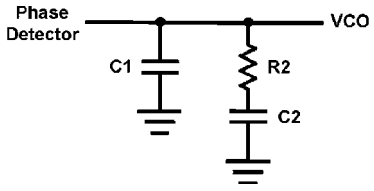


Figure 6-32 Schematic of the second-order loop filter.

and

$$\tau_2 = R_2 C_2 \tag{6-41}$$

The required values for τ_1 and τ_2 depend on the values we specify for the loop bandwidth, ω_p (equal to the frequency at which the loop's gain falls to 0 dB, or unity), and phase margin, ϕ_p, which is defined as the difference between 180° and the loop's phase shift at ω_p. For the best loop performance, we require a phase margin of 45°. The loop bandwidth must carefully be chosen with regard to lock time, phase noise, stability, and the reference-energy suppression. In this example, we will design our loop based on 10% of the reference frequency (f_{ref}, 200 kHz), so $\omega_p = 2\pi \times 20$ kHz, or 125.6 kHz. The required values for τ_1 and τ_2 can be determined from

$$\tau_1 = \frac{\sec \phi_p - \tan \phi_p}{\omega_p} \tag{6-42}$$

and

$$\tau_2 = \frac{1}{\omega_p^2 \tau_1} \tag{6-43}$$

From the time constants τ_1 and τ_2 and the loop bandwidth ω_p, we obtain the values for C_1, C_2, and R_2 as follows:

$$C_1 = \frac{\tau_1}{\tau_2} \frac{K_\phi K_{\text{VCO}}}{\omega_p^2 \tau_1} \sqrt{\frac{1 + (\omega_p \tau_2)^2}{1 + (\omega_p \tau_1)^2}} \tag{6-44}$$

$$C_2 = C_1 \left(\frac{\tau_2}{\tau_1} - 1 \right) \tag{6-45}$$

$$R_2 = \frac{\tau_2}{C_2} \tag{6-46}$$

where K_{VCO} = VCO tuning voltage constant (frequency versus voltage ratio) in MHz/V

K_ϕ = phase-detector/charge-pump gain constant (its ratio of current output to input phase differential) in mA

N = main divider ratio

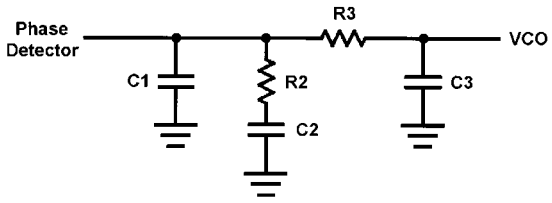


Figure 6-33 Third-order loop filter for greater reference-energy suppression.

If additional suppression of the loop's reference frequency is needed, a low-pass filter section can be added, resulting in the third-order loop filter shown in Figure 6-33.

The additional reference-frequency attenuation provided by R_3C_3 can be found from

$$\text{Atten} = 20 \log[(2\pi f_{\text{ref}} R_3 C_3)^2 + 1] \tag{6-47}$$

The low-pass section contributes an additional pole, which must be low enough to provide significant additional reference attenuation and high enough ($\geq 5\ \omega_p$) not to compromise the loop's stability. The time constant τ_3 of the added low-pass section can be found from

$$\tau_3 = R_3 C_3 \tag{6-48}$$

We can find τ_3 for a given value of additional attenuation from

$$\tau_3 = \frac{\sqrt{10^{(\text{Atten}/20)} - 1}}{2\pi f_{\text{ref}}} \tag{6-49}$$

To compensate for the added low-pass section, we recalculate the filter component values using the new loop bandwidth, ω_c, which can be found from

$$\omega_c = \frac{(\tan \phi)(\tau_1 + \tau_3)}{(\tau_1 + \tau_3)^2 + \tau_1 \tau_3} \left(\sqrt{1 + \frac{(\tau_1 + \tau_3)^2 + \tau_1 \tau_3}{[(\tan \phi)(\tau_1 + \tau_3)]^2}} - 1 \right) \tag{6-50}$$

We then reduce the phase-margin degradation caused by R_3C_3 by increasing C_1 and C_2 while slightly decreasing R_2. Calculating new values for C_1, C_2, and R_2 requires that we first determine the value for τ_2 as modified by R_3C_3:

$$\tau_2' = \frac{1}{\omega_c^2(\tau_1 + \tau_3)} \tag{6-51}$$

We can calculate C_1 as

$$C_1 = \frac{\tau_1}{\tau_2'} \frac{K_\phi K_{\text{VCO}}}{\omega_c^2 N} \left(\frac{1 + \omega_c^2 \tau_2'^2}{(1 + \omega_c^2 \tau_1^2)(1 + \omega_c^2 \tau_3^2)} \right)^{1/2} \tag{6-52}$$

As with the original second-order filter, C_2 and R_2 are calculated by means of Eqs. (6-45) and (6-46), respectively.

The values of R_3 and C_3 are somewhat arbitrary. The following rules of thumb apply: The value of C_3 should be less than that of C_1 and C_2, and preferably $\leq C_1/10$ to keep τ_3 from interacting with τ_1 and τ_2. Also, R_3 should be $\geq 2R_2$. Any capacitance already present on the VCO tuning line, including the input capacitance of the tuning diode(s), must be allowed for in selecting the value actually used for C_3 in the constructed loop filter.

The conversion from the voltage gain of a VCO to the charge current as needed here can be obtained from the formula [3]

$$K_\phi = \frac{I_{\text{charge}}}{2\pi f_R C_R} \tag{6-53}$$

which can be solved for the charge current, I_{charge}, as

$$I_{\text{charge}} = K_\phi 2\pi f_R C_R \tag{6-54}$$

where K_ϕ is the phase-detector/charge-pump gain factor in V/rad, f_R is the reference frequency, and C_R is the memory ("ramp") capacitance, typically 0.1 μF. The loop filters of Figures 6-32 and 6-33 include C_R as C_1.

Example. Design a third-order loop filter for a 900-MHz synthesizer with a 200-kHz reference based on the following parameters:

$K_{\text{VCO}} = 20$ MHz/V
$K_\phi = 5$ mA
$N = 4500$
$\omega_p = 2\pi \times 20$ kHz = 1.256E5
$\phi_p = 45°$
Atten = 20 dB

$$\tau_1 = \frac{\sec \phi_p - \tan \phi_p}{\omega_p} = 3.29\text{E}{-6} \tag{6-55}$$

$$\tau_3 = \frac{\sqrt{10^{(20/20)}} - 1}{(2\pi \times 200\text{E}3)^2} = 2.387\text{E}{-6} \tag{6-56}$$

$$\omega_c = \frac{3.29\text{E}{-6} + 2.387\text{E}{-6}}{[(3.29\text{E}{-6} + 2.387\text{E}{-6})^2 + 3.29\text{E}{-6} \times 2.387\text{E}{-6})]} \times \left(1 + \sqrt{\frac{3.29\text{E}{-6} + 2.387\text{E}{-6})^2 + 3.29\text{E}{-6} \times 2.387\text{E}{-6}}{[(3.29\text{E}{-6} + 2.387\text{E}{-6})]^2}}\right) - 1 \tag{6-57}$$

$$= 7.045\text{E}4 \tag{6-58}$$

$$\tau_2 = \frac{1}{(7.045\text{E}4)^2 \times (3.29\text{E}{-6} + 2.387\text{E}{-6})} = 3.549\text{E}{-5}$$

$$C_1 = \left(\frac{3.29\text{E}{-6}\ (5.0\text{E}{-3}) \times 20\text{E}6}{3.549\text{E}{-5}\ (7.045\text{E}4)^2 \times 4500}\right) \times \left(\frac{1 + (7.0454\text{E}4)^2 \times (3.549\text{E}{-5})^2}{[1 + (7.045\text{E}4)^2 \times (3.29\text{E}{-6})^2][1 + (7.045\text{E}4)^2 \times (2.39\text{E}{-6})^2]}\right)^{1/2} \tag{6.59}$$

$$= 1.085 \text{ nF}$$

$$C_2 = 1.085 \text{ nF} \times \left(\frac{3.55\text{E}{-5}}{3.29\text{E}{-6}} - 1 \right) = 10.6 \text{ nF} \tag{6-60}$$

$$R_2 = \frac{3.55\text{E}{-5}}{10.6\text{E}{-9}} = 3.35 \text{ k}\Omega \tag{6-61}$$

If we choose $R_3 = 22$ kΩ, then

$$C_3 = \frac{2.34\text{E}{-6}}{22\text{E}3} = 106 \text{ pF} \tag{6-62}$$

Selecting the nearest standard value for each component gives $C_1 = 1000$ pF, $R_2 = 3.3$ kΩ, $C_2 = 10$ nF, $R_3 = 22$ kΩ, and $C_3 = 100$ pF.

6-2-5 How to Do a Practical PLL Design Using CAD

1. Have a modern CAD tool give you a synthesized oscillator, such as a quarter-wave microstrip oscillator (Figure 6-34). Here we have used Compact Software's PLL Design Kit. (There are other similar programs on the market.) The loop synthesis program will ask you to specify the oscillator transistor's dc current and so on, and it will calculate the circuit's output power.

2. Ask your design tool for the oscillator's open- and closed-loop phase noise as a function of other noise sources (Figure 6-35). The program must consider the tuning diode as a major noise contributor. Here the loop bandwidth has been made too wide and makes the actual

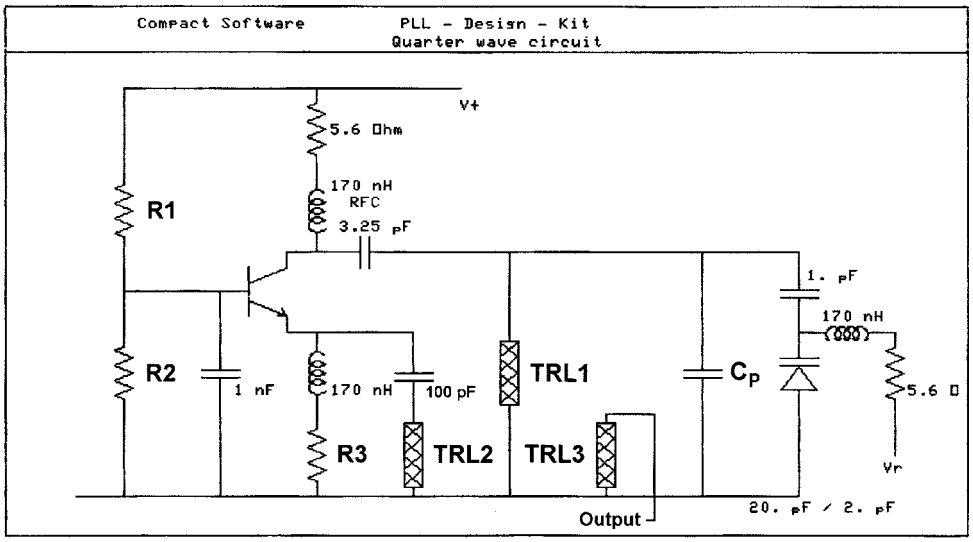


Figure 6-34 To design a PLL, begin with a suitable CAD-generated oscillator. This 900-MHz circuit is similar to that of the ceramic-resonator oscillator shown in Figure 5-33. TRL_1, TRL_2, and TRL_3 are coupled transmission lines; the resonator, TRL_1, is slightly less than λ/4 at the operating frequency. TRL_2 provides feedback, and TRL_3 provides output coupling. R_1, R_2, and R_3 must be specified; C_P must be determined according to the desired tuned range.

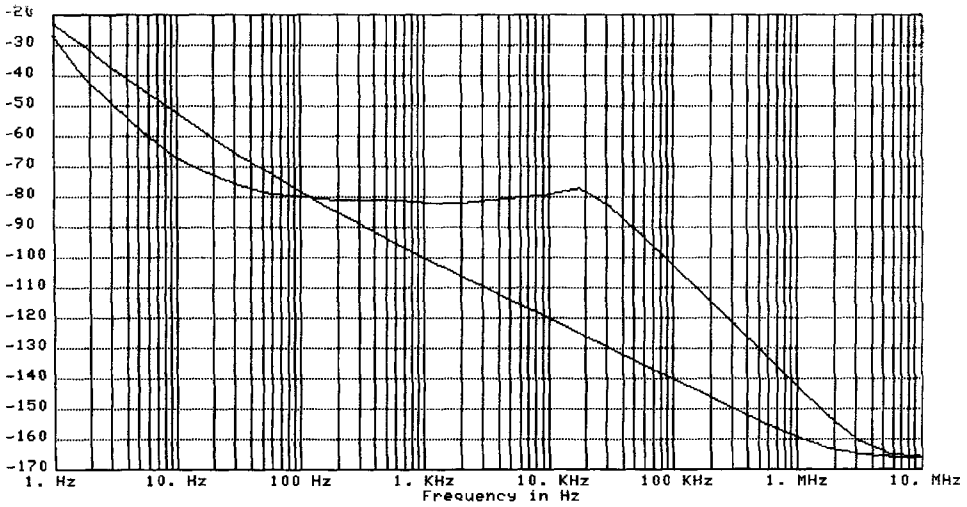


Figure 6-35 Predicted open- and closed-loop phase noise for the oscillator shown in Figure 6-34. The portion of the closed-loop curve up to 20 kHz represents the phase noise of the loop's crystal reference oscillator multiplied up. The VCO operates at 900 MHz; the reference frequency is 200 kHz. For references of 2 and 20 MHz, the phase noise would drop by 20 and 40 dB, respectively. See Table 6-1 and Figure 6-36.

phase noise worse. This plot is valid only if the dividers are the only noise contributors. The phase/frequency detector also adds noise following the equation

$$\mathcal{L} = \mathcal{L}_0 + 10 \log(F_r)$$

where $\mathcal{L}_0$ is a constant that is equivalent to the phase/frequency-detector noise with F_r = 1 Hz. $\mathcal{L}$ as a function of F_r is given below for standard PLL chips:

Table 6-1 Phase noise v. reference frequency derived from a high-performance 50-MHz crystal oscillator and multiplied to 880 MHz

Offset from Carrier	Phase Noise of 50-MHz Standard Oscillator (dBc/Hz)	Phase Noise in dBc/Hz After Dividing the Standard Down to: and Multiplying the Result to 880 MHz:		
		25 kHz	1.25 MHz	50 MHz
10 Hz	–80	–55	–55	–55
100 Hz	–110	–69	–85	–85
1 kHz	–140	–69	–103	–115
10 kHz	–160	–69	–103	–135
100 kHz	–160	–69	–103	–135
1 MHz	–160	–69	–103	–135

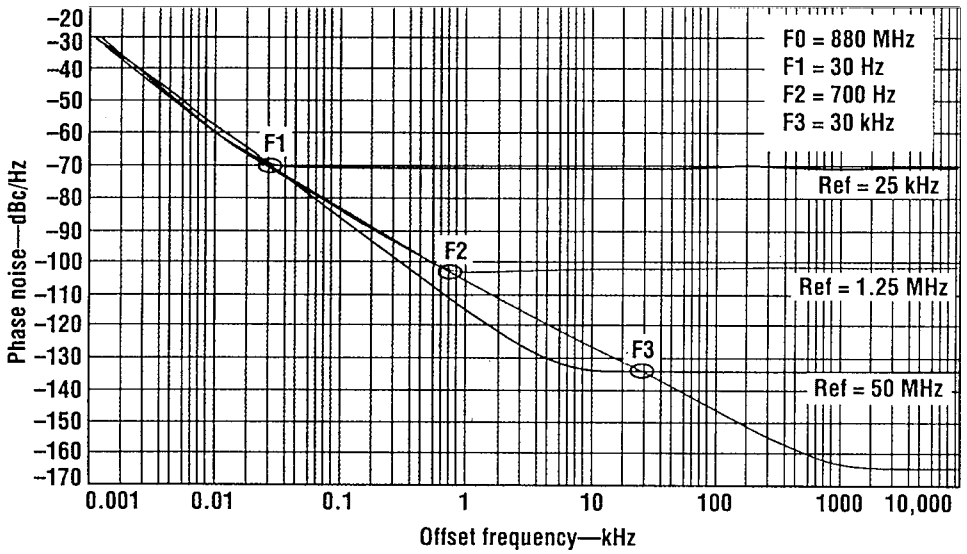


Figure 6-36 Result of dividing a 50-MHz standard oscillator down to different reference frequencies and then multiplying these frequencies to 880 MHz. The continuing line below breakpoint F3 shows the performance of the free-running 50-MHz oscillator.

$\mathcal{L}$ (dBc/Hz)	F_r (Hz)
−168 to −170	10k
−164 to −168	30k
−155 to −160	200k
−150 to −155	1M
−145	10M

3. Ask your synthesis program to design a loop filter for best performance, meaning a phase margin of 45° (Figure 6-37). Here we have arbitrarily included an additional low-pass filter with a cutoff frequency of 30 kHz to illustrate its effect on the loop performance.

4. Generate a Bode plot for the loop design and check the loop's phase margin at f_0, the frequency at which the loop gain is 0 dB. In this example, the plot (Figure 6-38) shows that because of the 30-kHz low-pass filter, we achieve only about 12° phase margin instead of the desired 45°. We can also see that the slope of the gain response through f_0 is not maintained at 6 dB/octave (20 dB/decade) over a sufficiently wide gain span. Maintaining the 6-dB/octave slope for loop-gain values of +10 to −10 dB is essential.

We can determine from Figure 6-38 that the loop gain is (60 dB at 200 kHz, the loop's reference frequency (f_{ref}). Adding this to the phase-detector's reference suppression (at least 40 dB) would appear to net us an overall reference suppression of 100 dB, but in any real implementation, crosstalk on the PLL PC board would reduce this value to between 80 and 90 dB.

5. As a result of the insufficient phase margin, the loop rings (Figure 6-39), locking in 322 μs instead of a possible 80 μs—a lock time four times as long as intended. Achievement of a phase margin of 45°, sometimes characterized merely as a rule of thumb, minimizes

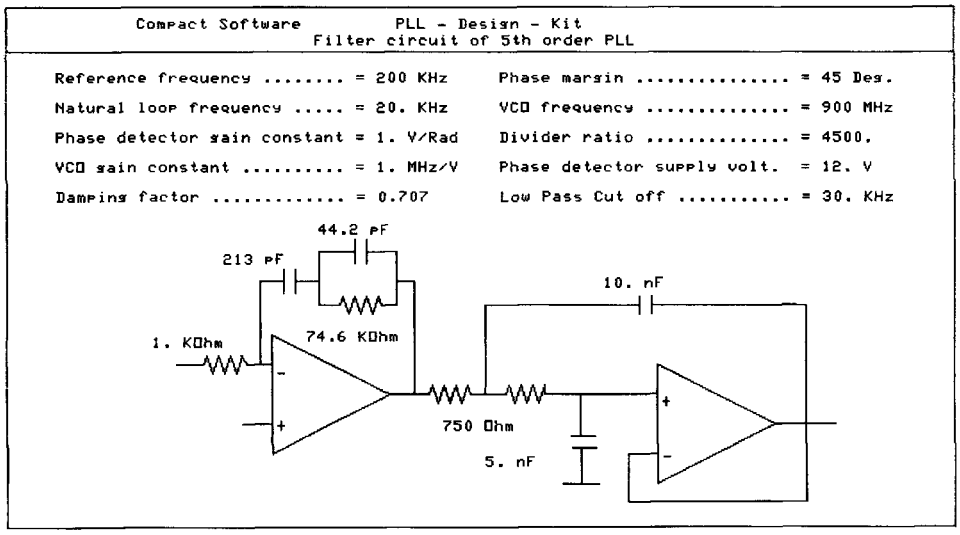


Figure 6-37 CAD-generated loop filter design based on a reference frequency of 200 kHz, a loop bandwidth of 20 kHz, and the optimum phase margin of 45°. We have arbitrarily included a 30-kHz low-pass filter to illustrate its effects on loop performance.

locking time and overshoot. Values less than 45° result in excessive overshoot and ringing, as shown here, and values greater than 45° result in an overdamped loop that crawls into lock.

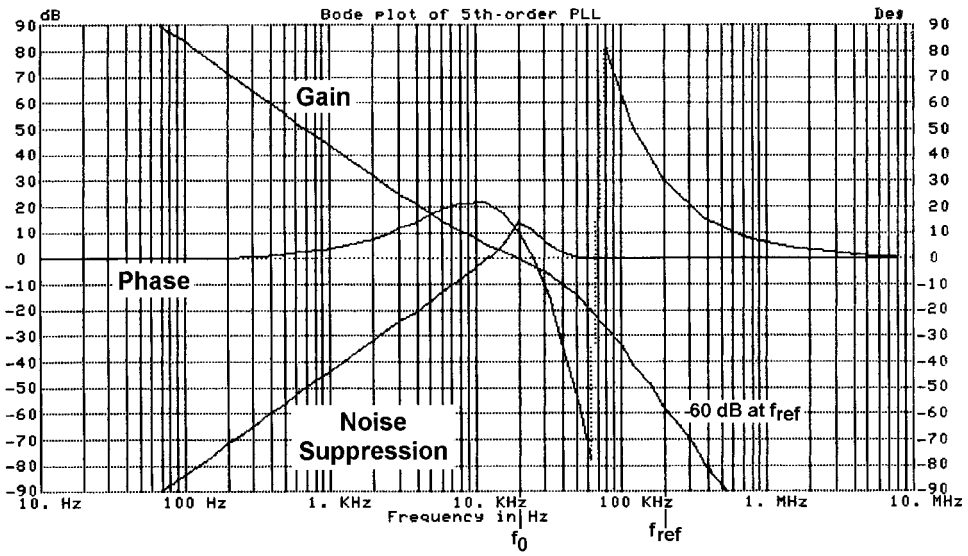


Figure 6-38 Bode plot for the PLL. The 30-kHz low-pass filter disallows a phase margin of 45° at f_0; instead, the margin is only about 20°.

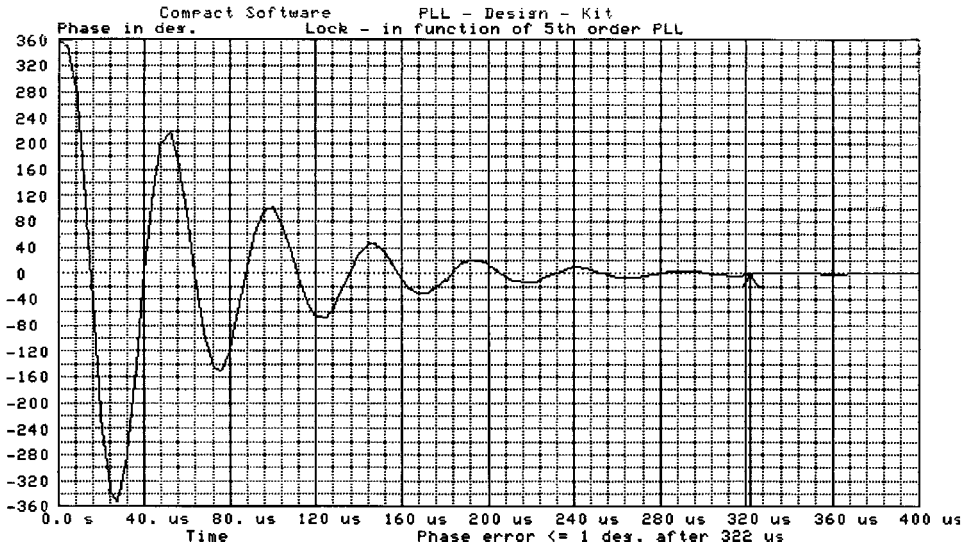


Figure 6-39 Lock-in response of the PLL. As a result of the insufficient phase margin, the loop is underdamped and takes 322 μs to achieve lock.

6-3 FRACTIONAL-*N*-DIVISION PLL SYNTHESIS

6-3-1 The Fractional-*N* Principle

The principle of the fractional-*N* PLL synthesizer has been around for awhile. In the past, implementation of this has been done in an analog system. It would be ideal to be able to build a single-loop synthesizer with a 1.25- or 50-MHz reference and yet obtain the desired step-size resolution, such as 25 kHz. This would lead to the much smaller division ratio and much better phase-noise performance.

An alternative would be for N to take on fractional values. The output frequency could then be changed in fractional increments of the reference frequency. Although a digital divider cannot provide a fractional division ratio, ways can be found to accomplish the same task effectively.

The most frequently used method is to divide the output frequency by $N + 1$ every M cycles and to divide by N the rest of the time. The effective division ratio is then $N + 1/M$, and the average output frequency is given by

$$f_o = \left(N + \frac{1}{M}\right) f_r \tag{6-63}$$

This expression shows that f_o can be varied in fractional increments of the reference frequency by varying M. The technique is equivalent to constructing a fractional divider, but the fractional part of the division is actually implemented using a phase accumulator. The phase accumulator approach is illustrated by the following example. This method can be expanded to frequencies much higher than 6 GHz using the appropriate synchronous dividers.

Example. Consider the problem of generating 899.8 MHz using a fractional-*N* loop with a 50-MHz reference frequency,

$$899.8 \text{ MHz} = 50 \text{ MHz}\left(N + \frac{K}{F}\right)$$

The integral part of the division *N* has to be set to 17 and the fractional part *K*/*F* needs to be 996/1000 (the fractional part *K*/*F* is not a integer); the VCO output has to be divided by 996 every 1000 cycles. This can easily be implemented by adding the number 0.996 to the contents of an accumulator every cycle. Every time the accumulator overflows, the divider divides by 18 rather than by 17. Only the fractional value of the addition is retained in the phase accumulator. If we move to the lower band or try to generate 850.2 MHz, *N* remains 17 and *K*/*F* becomes 4/1000. This method of using fractional division was first introduced by using analog implementation and noise cancellation, but today it is implemented totally as a digital approach. The necessary resolution is obtained from the dual-modulus prescaling, which allows for a well-established method of achieving a high-performance frequency synthesizer operating at UHF and higher frequencies. Dual-modulus prescaling avoids the loss of resolution in a system compared to a simple prescaler; it allows a VCO step equal to the value of the reference frequency to be obtained. This method needs an additional counter and the dual-modulus prescaler then divides one or two values depending on the state of its control. The only drawback of prescalers is the minimum division ratio of the prescaler for approximately N^2. The dual-modulus divider is the key to implementing the fractional-*N* synthesizer principle. Although the fractional-*N* technique appears to have a good potential of solving the resolution limitation, it also has its own complications. Typically, an overflow from the phase accumulator, which is the adder with the output feedback to the input after being latched, is used to change the instantaneous division ratio. Each overflow produces a jitter at the output frequency, caused by the fractional division, and is limited to the fractional portion of the desired division ratio.

In our case, we had chosen a step size of 200 kHz, and yet the discrete sidebands vary from 200 kHz for *K*/*F* = 4/1000 to 49.8 MHz for *K*/*F* = 996/1000. It will become the task of the loop filter to remove those discrete spurious frequencies. While in the past the removal of the discrete spurs has been accomplished by using analog techniques, various digital methods are now available. The microprocessor has to solve the following equation:

$$N^* = \left(N + \frac{K}{F}\right) = [N(F - K) + (N + 1)K] \tag{6-64}$$

Example. For F_o = 850.2 MHz, we obtain

$$N^* = \frac{850.2 \text{ MHz}}{50 \text{ } MHz} = 17.004$$

Following the formula above,

$$N^* = \left(N + \frac{K}{F}\right) = \frac{[17(1000 - 4) + (17 + 1) \times 4]}{1000}$$

$$= \frac{16932 + 72}{1000} = 17.004$$

$$F_o = 50 \text{ MHz} \times \frac{(16932 + 72)}{1000}$$

$$= 846.6 \text{ MHz} + 3.6 \text{ MHz}$$

$$= 850.2 \text{ MHz}$$

By increasing the number of accumulators, frequency resolution much below 1-Hz step size is possible with the same switching speed.

There is an interesting, generic problem associated with *all* fractional-*N* synthesizers. Assume for a moment that we use our 50-MHz reference and generate a 550-MHz output frequency. This means our division factor is 11. Aside from reference-frequency sidebands (±50 MHz) and harmonics, there will be no unwanted spurious frequencies. Of course, the reference sidebands will be suppressed by the loop filter by more than 90 dB. For reasons of phase noise and switching speed, a loop bandwidth of 100 kHz has been considered. Now, taking advantage of the fractional-*N* principle, say we want to operate at an offset of 30 kHz (550.03 MHz). With this new output frequency, the inherent spurious-signal reduction mechanism in the fractional-*N* chip limits the reduction to about 55 dB. Part of the reason why the spurious-signal suppression is less in this case is that the phase/frequency detector acts as a mixer, collecting both the 50-MHz reference (and its harmonics) and 550.03 MHz. Mixing the 11th reference harmonic (550 MHz) and the output frequency (550.03 MHz) results in output at 30 kHz; since the loop bandwidth is 100 kHz, it adds nothing to the suppression of this signal. To solve this, we could consider narrowing the loop bandwidth to 10% of the offset. A 30-kHz offset would equate to a loop bandwidth of 3 kHz, at which the loop speed might still be acceptable, but for a 1-kHz offset, the necessary loop bandwidth of 100 Hz would make the loop too slow. A better way is to use a different reference frequency—one that would place the resulting spurious product considerably outside the 100-kHz loop filter window. If, for instance, we used a 49-MHz reference, multiplication by 11 would result in 539 MHz. Mixing this with 550.03 MHz would result in spurious signals at ±11.03 MHz, a frequency so far outside the loop bandwidth that it would essentially disappear. Starting with a VHF, low-phase-noise crystal oscillator, such as 130 MHz, one can implement an intelligent reference-frequency selection to avoid these discrete spurious signals. An additional method of reducing the spurious contents is maintaining a division ratio greater than 12 in all cases. Actual tests have shown that these reference-based spurious frequencies can be repeatably suppressed by 80–90 dB.

6-3-2 Spur-Suppression Techniques

While several methods have been proposed in the literature [4–9], the method of reducing the noise by using a sigma-delta modulator has shown the most promise. The concept is to get rid of the low-frequency phase error by rapidly switching the division ratio to eliminate the gradual phase error at the discriminatory input. By changing the division ratio rapidly between different values, the phase errors occur in both polarities, positive as well as negative, and at an accelerated rate that explains the phenomenon of high-frequency noise push-up. This noise, which is converted to a voltage by the phase/frequency discriminator and loop filter, is filtered out by the low-pass filter. The main problem associated with this noise shaping technique is that the noise power rises rapidly with frequency. Figure 6-40 shows noise contributions with such a sigma-delta modulator in place.

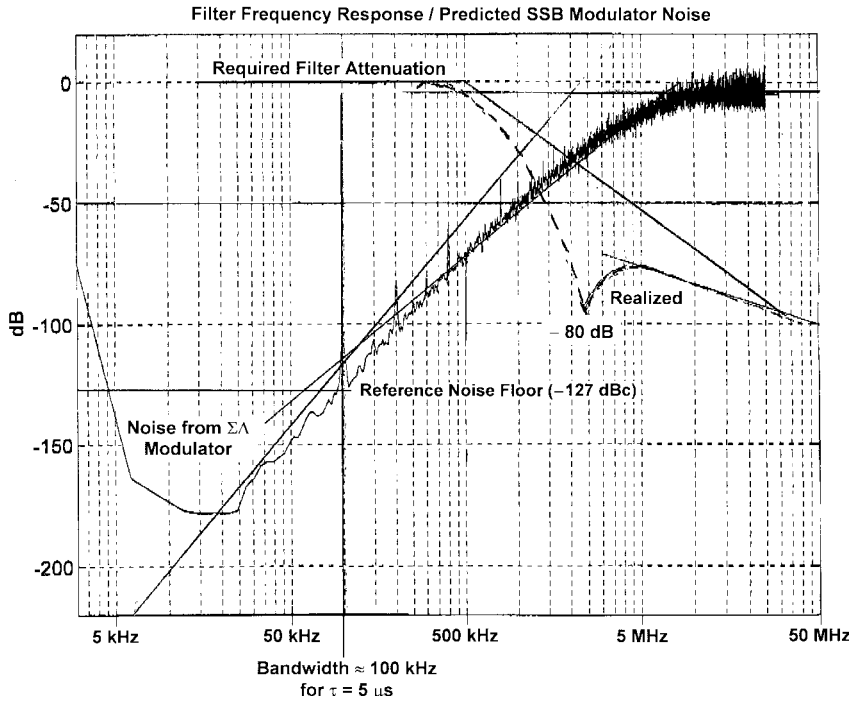


Figure 6-40 The filter frequency response/phase-noise analysis graph shows the required attenuation for the reference frequency of 50 MHz and the noise generated by the sigma-delta converter (three steps) as a function of the offset frequency. It becomes apparent that the sigma-delta converter noise dominates above 80 kHz unless attenuated.

On the other hand, we can now, for the first time, build a single-loop synthesizer with switching times as fast as 6 μs and very little phase-noise deterioration inside the loop bandwidth, as seen in Figure 6-40. Since this system maintains the good phase noise of the ceramic-resonator-based oscillator, the resulting performance is significantly better than the phase noise expected from high-end signal generators. However, this method does not allow us to increase the loop bandwidth beyond the 100-kHz limit, where the noise contribution of the sigma-delta modulator takes over.

Table 6-2 shows some of the modern spur-suppression methods. These three-stage sigma-delta methods with larger accumulators have the most potential [4–9].

Table 6-2 Modern spur-suppression methods

Technique	Feature	Problem
DAC phase estimation	Cancel spur by DAC	Analog mismatch
Pulse generation	Insert pulses	Interpolation jitter
Phase interpolation	Inherent fractional divider	Interpolation jitter
Random jittering	Randomize divider	Frequency jitter
Sigma-delta modulation	Modulate division ratio	Quantization noise

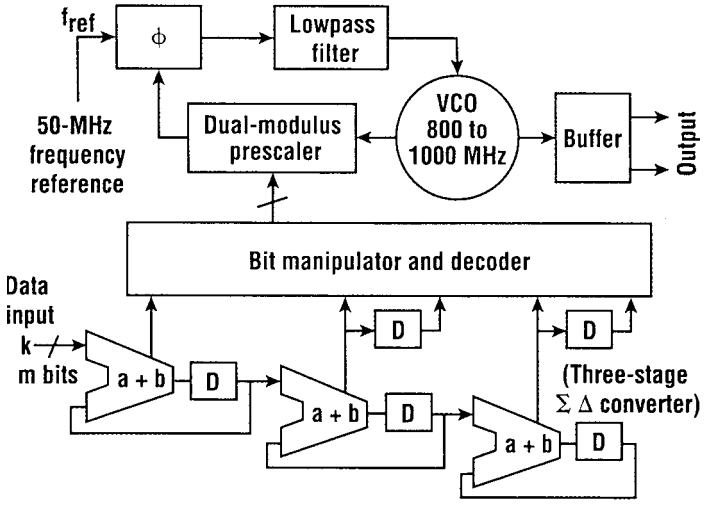


Figure 6-41 Block diagram of the fractional-*N* synthesizer built using a custom IC capable of operation at reference frequencies up to 150 MHz. The frequency is extensible up to 3 GHz using binary (÷2, ÷4, ÷8, etc.) and fixed-division counters.

The power spectral response of the phase noise for the three-stage sigma-delta modulator is calculated from

$$L(f) = \frac{(2\pi)^2}{12 f_{\text{ref}}} [2\sin(\pi f / f_{\text{ref}})]^{2(n-1)} \text{ rad}^2/\text{Hz} \tag{6-65}$$

where n is the number of the stage of the cascaded sigma-delta modulator [10]. Equation (6-65) shows that the phase noise resulting from the fractional controller is attenuated to negligible levels close to the center frequency, and further from the center frequency, the phase noise is increased rapidly and must be filtered out prior to the tuning input of the VCO to prevent unacceptable degradation of spectral purity. A loop filter must be used to filter the noise in the PLL loop. Figure 6-40 shows the plot of the phase noise versus the offset frequency from the center frequency. A fractional-*N* synthesizer with a three-stage sigma-delta modulator as shown in Figure 6-41 has been built. The synthesizer consists of a phase/frequency detector, an active low-pass filter (LPF), a voltage-controlled oscillator (VCO), a dual-modulus prescaler, a three-stage sigma-delta modulator, and a buffer. Figure 6-42 shows the inner workings of the chip in greater detail.

After designing, building, and predicting the phase-noise performance of this synthesizer, it becomes clear that measuring the phase noise of such a system becomes tricky. Standard measurement techniques that use a reference synthesizer would not provide enough resolution because there are no synthesized signal generators on the market good enough to measure such low values of phase noise. Therefore, we had to build a comb generator that would take the output of the oscillator and multiply this up 10–20 times.

Passive phase-noise measurement systems, based on delay lines, are not selective, and the comb generator confuses them. However, the Rohde & Schwarz FSEM spectrum analyzer with the K-4 option has sufficient resolution to be used for phase-noise measurements. All of the Rohde & Schwarz FSE-series spectrum analyzers use a somewhat more discrete

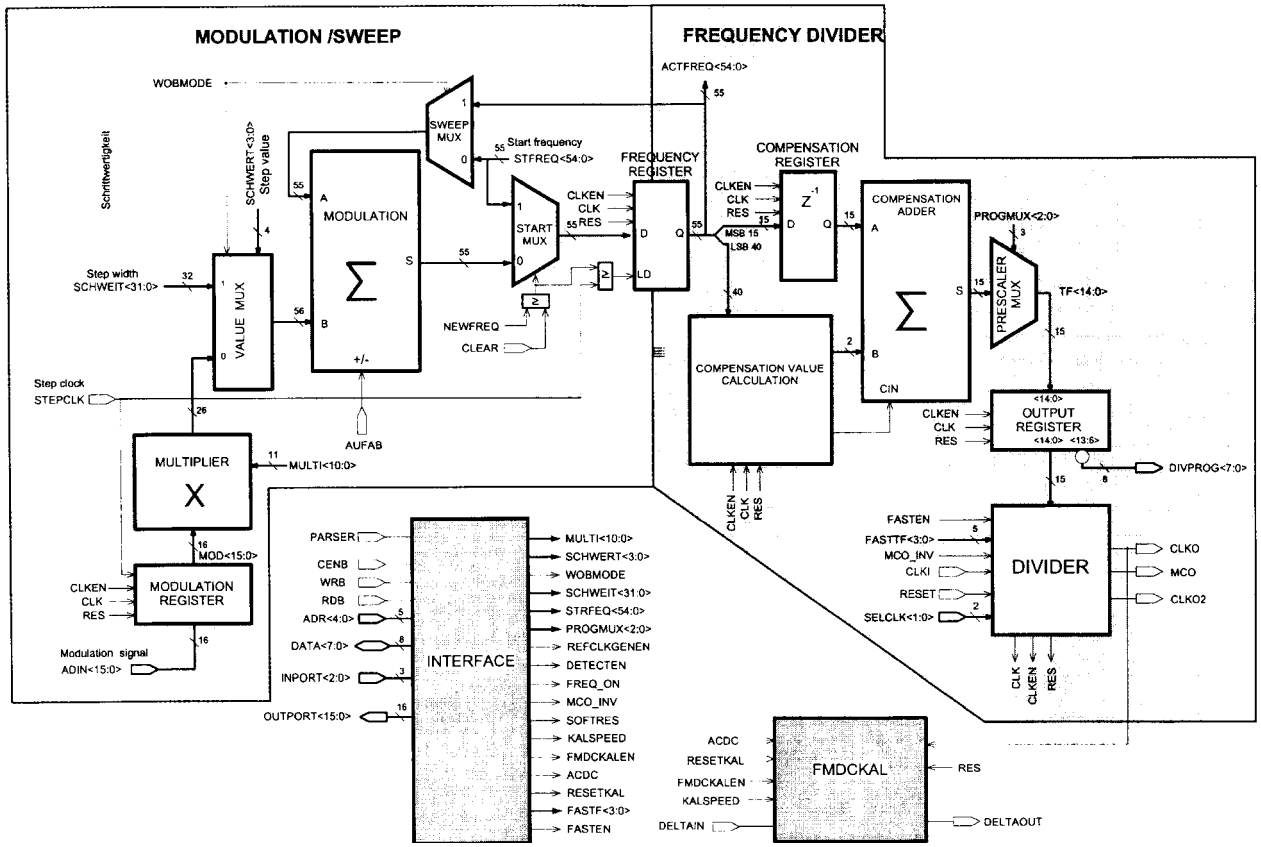


Figure 6-42 Detailed block diagram of the inner workings of the fractional-*N*-division synthesizer chip.

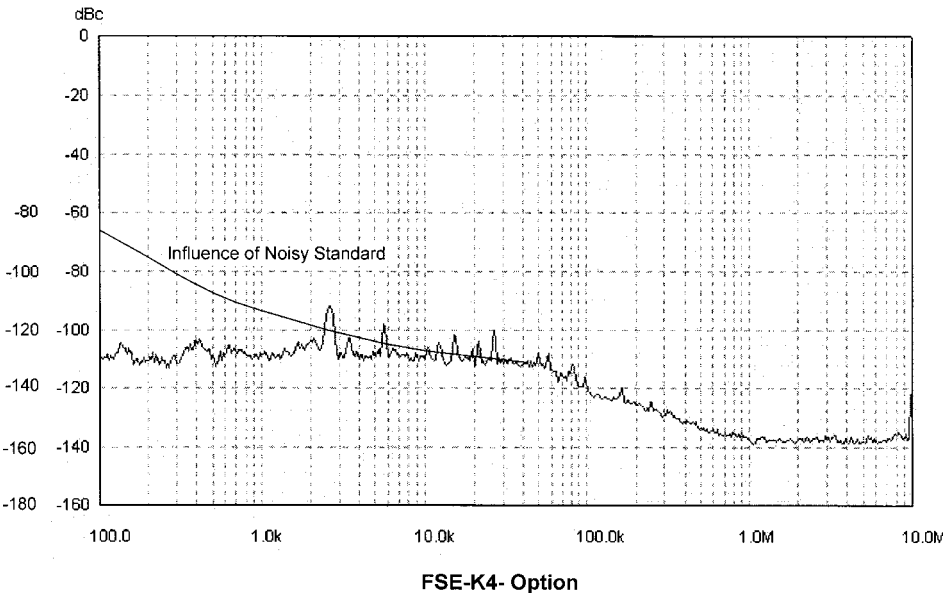


Figure 6-43 Measured phase noise of the fractional-*N*-division synthesizer using a custom-built, high-performance 50-MHz crystal oscillator as reference, with the calculated degradation due to a noisy reference plotted for comparison. Both synthesizer and spectrum analyzer use the same reference.

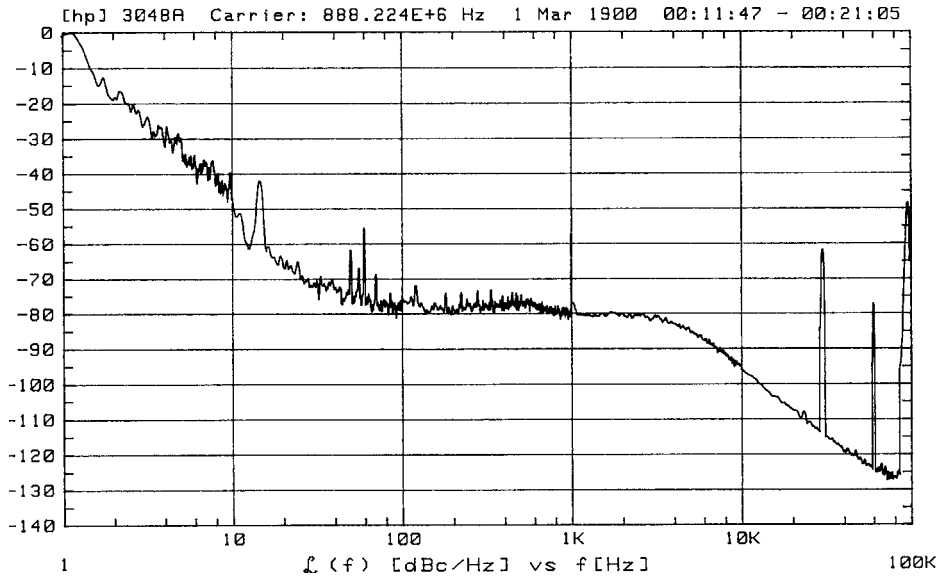


Figure 6-44 Measured phase noise of an 880-MHz synthesizer using a conventional synthesizer chip. Comparing this to Figure 6-43, we see that a large improvement is possible using fractional-*N*-division synthesizers, as shown in this product description.

fractional-division synthesizer with a 100-MHz reference. Based on the multiplication factor of 10, it turns out that there is enough dynamic range in the FSEM analyzer with the K-4 option to be used for phase-noise measurement. The useful frequency range off the carrier for the system is 100 Hz to 10 MHz—perfect for this measurement.

Figure 6-43 shows the measured phase noise of the final frequency synthesizer.

During the measurements, we determined that we needed a 50-MHz crystal oscillator with better phase noise. Upon examination of the measured phase noise shown in Figure 6-43, it can be seen that the oscillator used as the reference was significantly better. Otherwise, this phase noise would not have been possible. Also, the loop filter cutoff frequency of about 100 kHz can be recognized by the roll-off in Figure 6-43. This fractional-*N*-division synthesizer with a high-performance VCO has a significantly better phase noise than other example systems in this frequency range. To demonstrate this improvement, phase-noise measurements were made on standard systems, using typical synthesizer chips. Although the phase

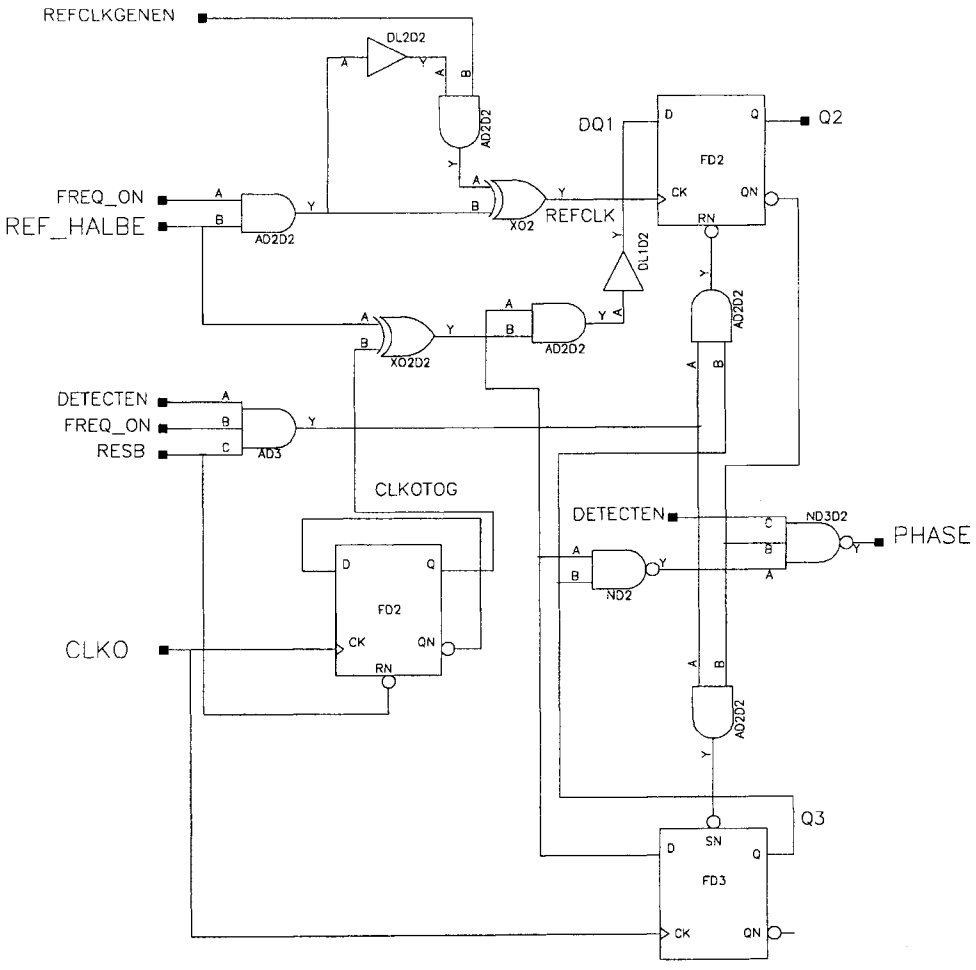


Figure 6-45 Custom-built phase detector with a noise floor of better than –168 dB. This phase detector shows extremely low phase jitter.

noise by itself and the synthesizer design are quite good, they are no match for this new approach, as can be seen in Figure 6-44 [11].

By combining the very best available technologies, such as using high-end VCOs with ceramic-resonator-based tuned circuits or high-*Q LC* arrangements including microstrips on Teflon material and modern fractional-*N*-division synthesizer blocks that can operate with a 50-MHz reference, it is possible to build an extremely high-quality system.

We have also learned that the limits of this technology are determined by the reference crystal oscillator and possibly the phase detector used; these must be specially designed to match the synthesizer's performance. (Figure 6-45 shows a block diagram of the phase detector.) Due to the very large bandwidth, switching speeds at 6 μs were made possible. The resolution of the synthesizer itself depends on the accumulator size. Step sizes of 25 kHz up to several megahertz were successfully tested. For wideband applications, some of the critical points are at 10 kHz where –120 dBc/Hz is desired, at 800 kHz, better than –153 dBc/Hz, and at 3 MHz, better than –155 dBc/Hz.

As validation, three types of synthesizers have been built: one covering 75–105 MHz, with 1-Hz resolution, for an HF transceiver; another covering 700–2000 MHz; and a third covering 2700–3500 MHz, also with better than 1-Hz resolution.

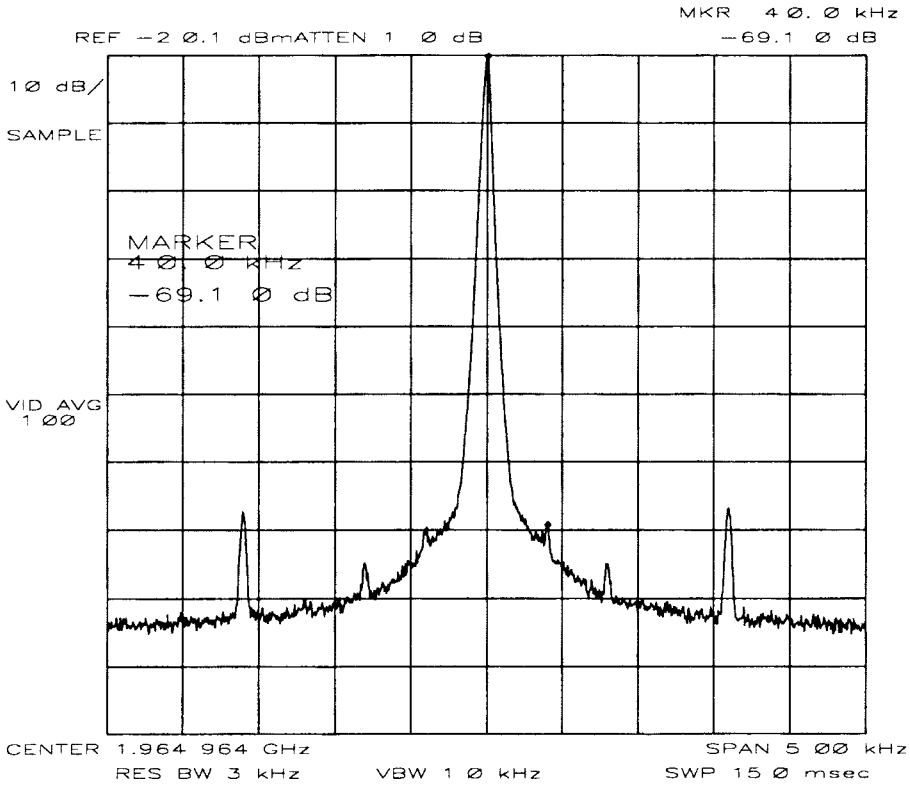


Figure 6-46 Result of fractional spur measurement for an LMX2350-based synthesizer. The spurs at ±40 kHz are –69.1 dBc (fractional compensation = 4/16); the center frequency is 1.964964 GHz. (Reprinted with permission of National Semiconductor Corporation.)

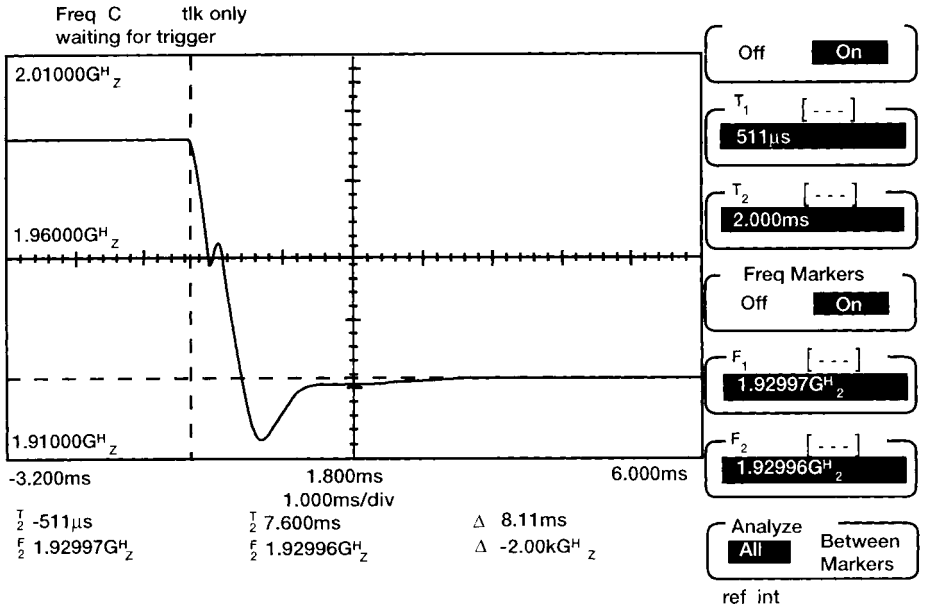


Figure 6-47 Result of lock time measurement for an LMX2350 based synthesizers (Reprinted with permission of National Semiconductor Corporation.)

The fractional-*N* PLL implementations described so far are intended mainly for high-performance, base-station applications. Simplified, highly integrated versions are available for mass-market wireless applications. One solution, National Semiconductor's LMX235*x* family, implements a fractional-*N* RF synthesizer and an integer-*N* IF synthesizer in one IC. The LMX2350's RF synthesizer can operate at input frequency up to 2.5 GHz; the LMX2352's, up to 1.2 GHz. The IF synthesizer in both parts can operate from 10 to 550 MHz. Figure 6-46 shows the close-in output spectrum of an LMX2350-based synthesizer at 1.965 GHz, and Figure 6-47 shows its lock time. A sample LMX2350 application is shown in Figure 6-48.

Fractional-*N* synthesizer chips are also available from Philips. Examples of Philips fractional-*N* parts include the SA7016DH (1.3 GHz), SA7025DK (main synthesizer, 1.0 GHz; auxiliary synthesizer, 150 MHz), SA7026DK (1.3 GHz/550 MHz), SA8016DH (2.5 GHz), SA8025ADK (1.8 GHz/150 MHz), and SA8026DK (2.5 GHz/550 MHz).

6-4 DIRECT DIGITAL SYNTHESIS

Direct digital frequency synthesis (DDFS), also referred to as *direct digital synthesis* (DDS), consists of generating a digital representation of the desired signal and then using a D/A converter to convert the digital representation to an analog waveform. Recent advances in high-speed microelectronics, particularly the microprocessor, make DDS practical at frequencies in and below the high-frequency band (as of this writing). Systems can be compact, use low power, and provide very fine frequency resolution with virtually instantaneous frequency changes. DDS is finding increasing application, particularly in conjunction with PLL synthesizers.

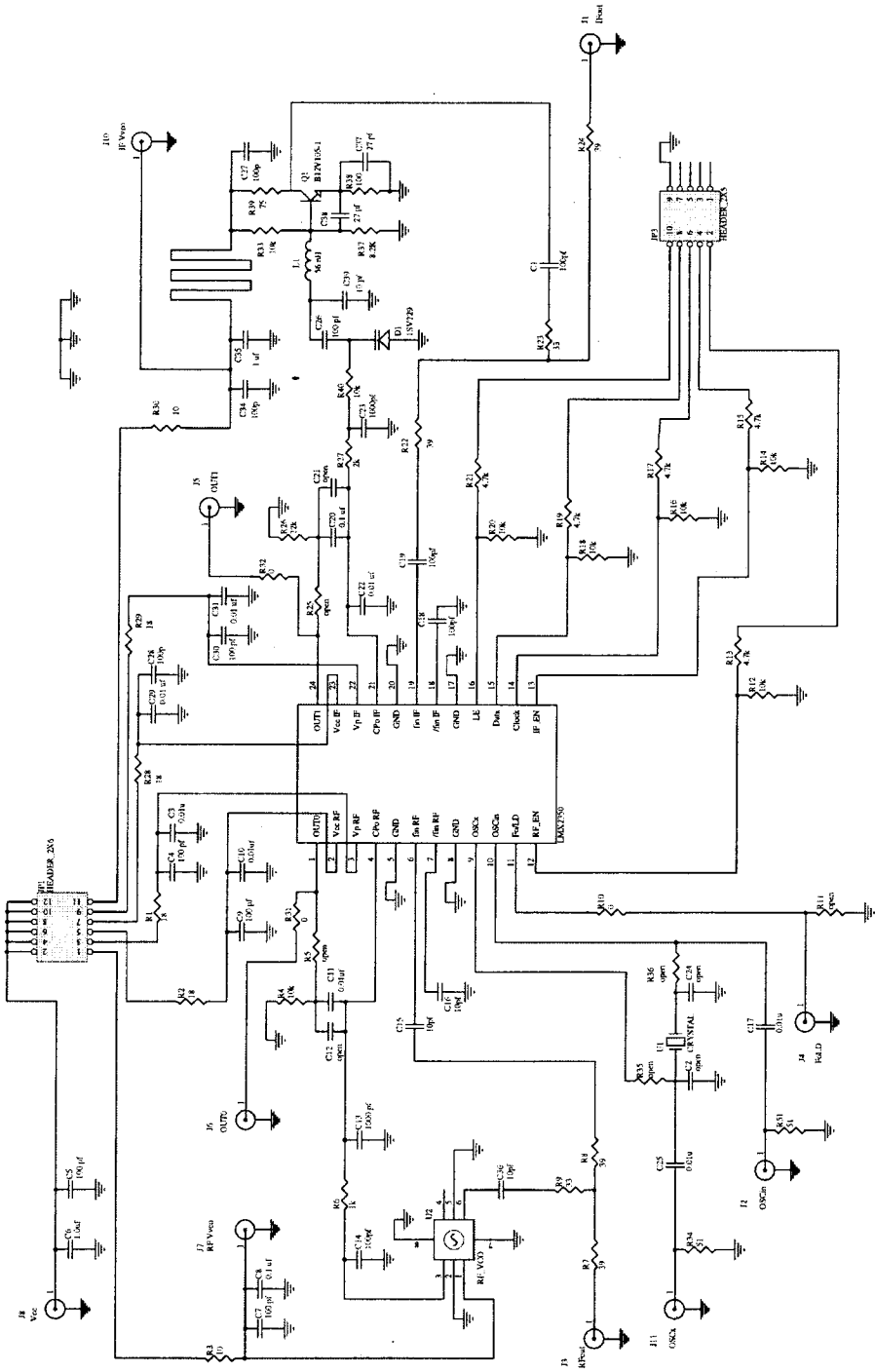


Figure 6-48 LMX2350 sample application schematic. (Reprinted with permission of National Semiconductor Corporation.)

DDS uses a single-frequency source (clock) as a time reference. One method of digitally generating the values of a sine wave is to solve the digital recursion relation as follows:

$$Y_n = [2\cos(2\pi ft)]Y_{n-1} - Y_{n-2} \tag{6-66}$$

This is solved by $Y_n = \cos(2\pi fnt)$. There are at least two problems with this method, however. The noise can increase until a limit cycle (nonlinear oscillation) occurs. Also, the finite word length used to represent $\cos(2\pi ft)$ places a limitation on the frequency resolution. Another method of DDS, direct table lookup, consists of storing the sinusoidal amplitude coefficients for successive phase increments in memory. Advances in miniaturization and the lowering cost of ROM make this the most frequently used technique.

One method of direct table lookup outputs the same N points for each cycle of the sine wave and changes the output frequency by adjusting the rate at which the points are computed. It is relatively difficult to obtain fine frequency resolution with this approach, so a modified table-lookup method is generally used. It is this method that we describe here. The function $\cos(2\pi ft)$ is approximated by outputting the function $\cos(2\pi fnT)$ for $n = 1, 2, 3, \ldots$, where T is the interval between conversions of digital words in the D/A converter and n represents the successive sample numbers. The sampling frequency, or rate, of the system is $1/T$. The lowest output frequency waveform contains N distinct points in its waveform, as illustrated in Figure 6-49. A waveform of twice the frequency can be generated, using the same sampling rate, but outputting every other data point. A waveform k times as fast is obtained by outputting every kth point at the same rate $1/T$. The frequency resolution, then, is the same as the lowest frequency, f_L.

The maximum output frequency is selected so that it is an integral multiple of f_L; that is, $f_U = kf_L$. If P points are used in the waveform of the highest frequency, N ($=kP$) points are used in the lowest frequency waveform. The number N is limited by the available memory size. The minimum value that P can assume is usually taken to be four. With this small value of P, the output contains many harmonics of the desired frequency. These can be removed by the use of low-pass filtering in the D/A output. For $P = 4$, the period of the highest frequency is $4T$, resulting in $f_U = 4f_L$. Thus, the highest attainable frequency is determined by the fastest sampling rate possible.

In the design of this type of DDS, the following guidelines apply:

- The desired frequency resolution determines the lowest output frequency f_L.
- The number of D/A conversions used to generate f_L is $N = 4k = 4f_U/f_L$ provided that four conversions are used to generate f_U ($P = 4$).
- The maximum output frequency f_U is limited by the maximum sampling rate of the DDS, $f_U \leq 1/4T$. Conversely, $T \leq 1/4f_U$.

The architecture of the complete DDS is shown in Figure 6-50. To generate nf_L, the integer n addresses the register, and each clock cycle kn is added to the content of the accumulator so that the content of the memory address register is increased by kn. Each knth point of the memory is addressed, and the content of this memory location is transferred to the D/A converter to produce the output sampled waveform.

To complete the DDS, the memory size and length (number of bits) of the memory word must be determined. The word length is determined by system noise requirements. The amplitude of the D/A output is that of an exact sinusoid corrupted with the deterministic noise due to truncation caused by the finite length of the digital words (quantization noise).

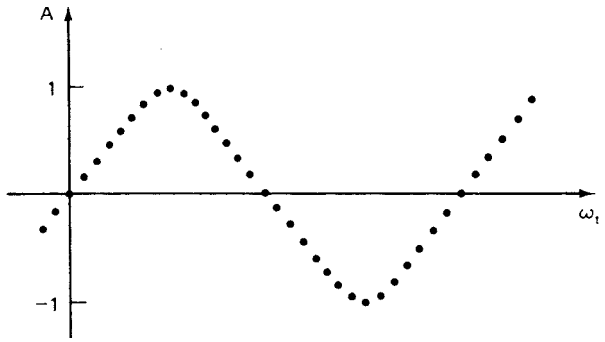


Figure 6-49 Synthesized waveform generated by direct digital synthesis.

If an $(n + 1)$-bit word length (including one sign bit) is used and the output of the A/D converter varies between ±1, the mean noise from the quantization will be

$$\rho^2 = \frac{1}{12}\left(\frac{1}{2}\right)^{2n} = \frac{1}{3}\left(\frac{1}{2}\right)^{2(n+1)} \tag{6-67}$$

The mean noise is averaged over all possible waveforms. For a worst-case waveform, the noise is a square wave with amplitude and $\frac{1}{2}(\frac{1}{2})^n$ and$\rho^2 = \frac{1}{4}(\frac{1}{2})^{2n}$. For each bit added to the word length, the spectral purity improves by 6 dB.

The main drawback of the low-power DDS is that it is limited to relatively low frequencies. The upper frequency is directly related to the maximum usable clock frequency; today, the limit is about 1 GHz. DDS tends to be noisier than other methods, but adequate spectral purity can be obtained if sufficient low-pass filtering is used at the output. DDS systems are easily constructed using readily available microprocessors. The combination of DDS for fine frequency resolution plus other synthesis techniques to

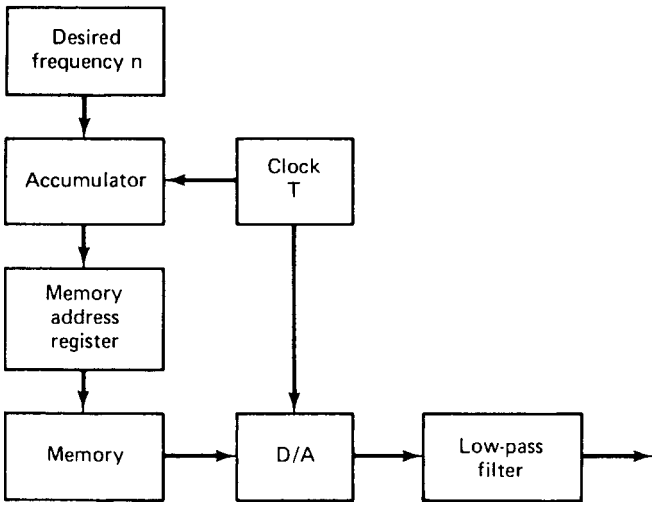


Figure 6-50 Block diagram of a direct digital frequency synthesizer.

obtain higher-frequency output can provide high resolution with very rapid setting time after a frequency change. This is especially valuable for frequency-hopping spread-spectrum systems.

Figure 6-51 shows the functional block diagram of a DDS system. In analyzing both the resolution and signal-to-noise ratio (or rather signal-to-spurious performance) of the DDS, one has to know the resolution and input frequencies. As an example, if the input frequency is approximately 35 MHz and the implementation is for a 32-bit device, the frequency resolution compared to the input frequency is 35E6 ÷ 2^{32} = 35E6 ÷ 4.294967296E9 or 0.00815 Hz ≈ 0.01 Hz. Given the fact that modern shortwave radios with a first IF of about 75 MHz will have an oscillator between 75 and 105 MHz, the resolution at the output range is more than adequate. In practice, one would use the microprocessor to round it to the next increment of 1 Hz relative to the output frequency.

As to the spurious response, the worst-case spurious response is approximately 20 log 2^R, where R is the resolution of the digital/analog converter. For an 8-bit A/D converter, this would mean approximately 48 dB down (worst case), as the output loop would have an analog filter to suppress close-in spurious noise. In our application, we will use an 8-bit external D/A converter. However, devices such as the Analog Devices AD7008 DDS modulator have a 10-bit resolution, as shown in Figure 6-52. Ten bits of resolution can translate into 20 log 2^{10}, or 60 dB of suppression. The actual spurious response would be much better. The current production designs for communication applications, such as shortwave transceivers, despite the fact that they are resorting to a combination of PLLs and DDSs, still end up somewhat complicated. By using 10 MHz from the DDS and using a single-loop PLL system, one can easily extend the operation to above 1 GHz but with higher complexity and power consumption.

Figure 6-53 shows the necessary components of a single-PLL system. Some communication equipment uses this approach.

Figure 6-54 shows the combination of a standard PLL and DDS, as implemented in the ICOM IC-736 HF/6-meter transceiver. This approach uses DDS in a frequency range between 500 kHz and 1 MHz. This frequency is upconverted either to 60 MHz for the shortwave band or 90 MHz for the 6-meter ham band. The resulting frequency is used as an auxiliary LO to convert the frequency of the first LO (69.0415 to 102.0115 MHz) down to the synthesizer IF between 8.5 and 41.5 MHz. There is an additional divide-by-2 stage in the loop, which therefore requires a reference frequency of 250 kHz instead of 500 kHz. This is done to extend the operating range of the synthesizer chip, including its prescaler's capability of operating at much higher frequencies, although it does not have such a hybrid DDS approach incorporated (Figure 6-55).

While this approach obtains a fairly small division ratio, it is still a four-loop synthesizer. One loop is the DDS itself. The second is the translator loop that mixes the DDS up to 60 MHz. The third is the main loop responsible for the desired output frequency. The fourth loop, so to speak, is the generation of the auxiliary LO frequencies at 60 and 90 MHz, which are derived from the 30-MHz frequency standard. For reasons of good phase noise, it employs a total of five VCOs. Because this loop's division ratio varies between 80 and 17, its gain is subject to considerable variation.

The 10.7-MHz signal from the crystal filter goes to a single-chip PLL (U5, a Motorola MC145170), which contains all the necessary dividers and the phase/frequency discriminator. The operational amplifier (U6, an OPA27) is driven from a 28-V source, and the negative supply of the OPA27 is connected as a voltage doubler, which receives its ac voltage from the synthesizer IC. This trick allows extension of the VCO control voltage. The output from

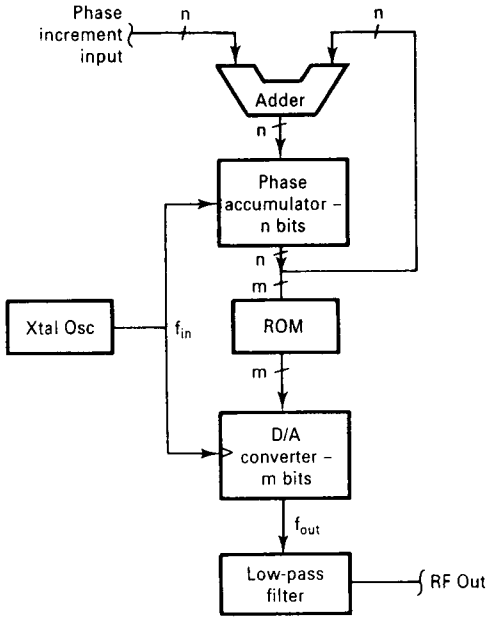


Figure 6-51 Block diagram of a DDS system.

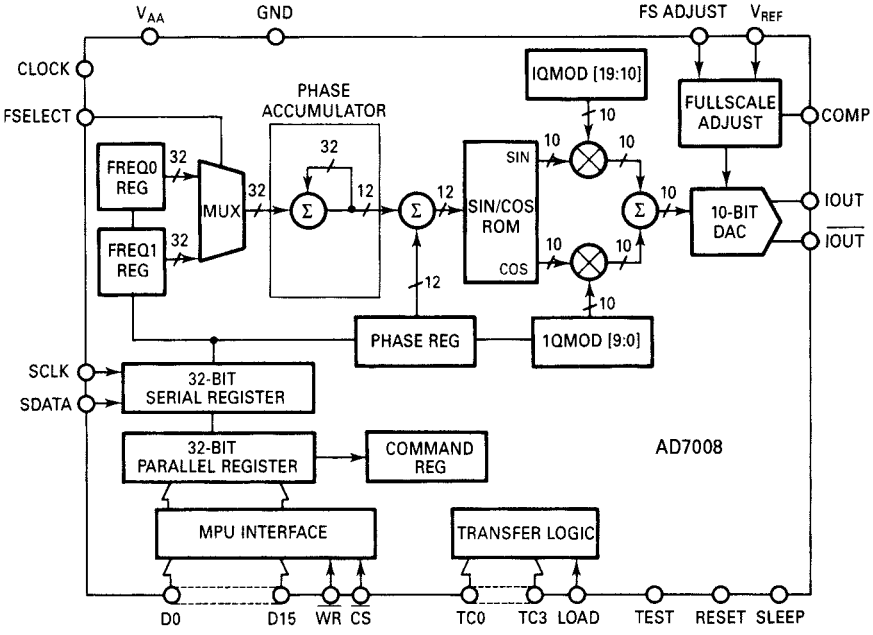


Figure 6-52 Functional block diagram of the Analog Devices AD7008 DDS modulator.

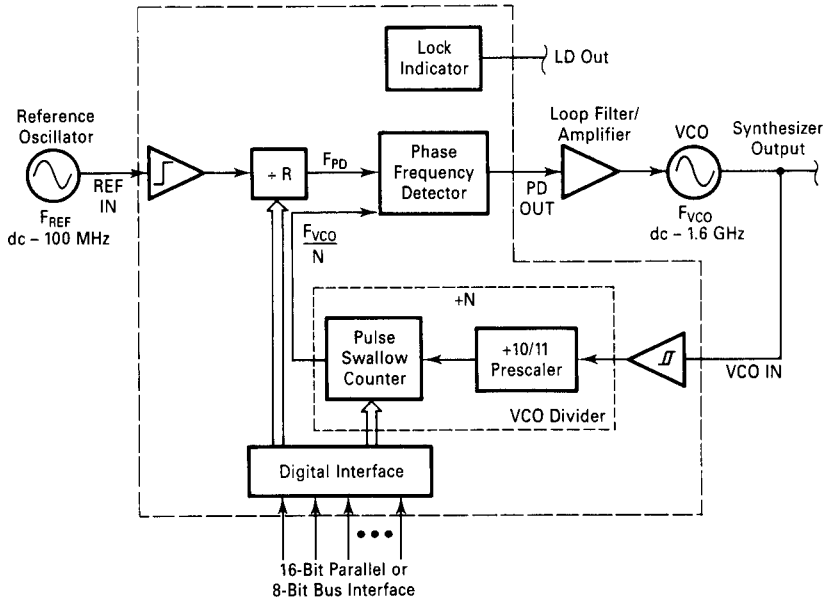


Figure 6-53 Block diagram of a single-loop PLL synthesizer showing all the necessary components for microwave and RF applications.

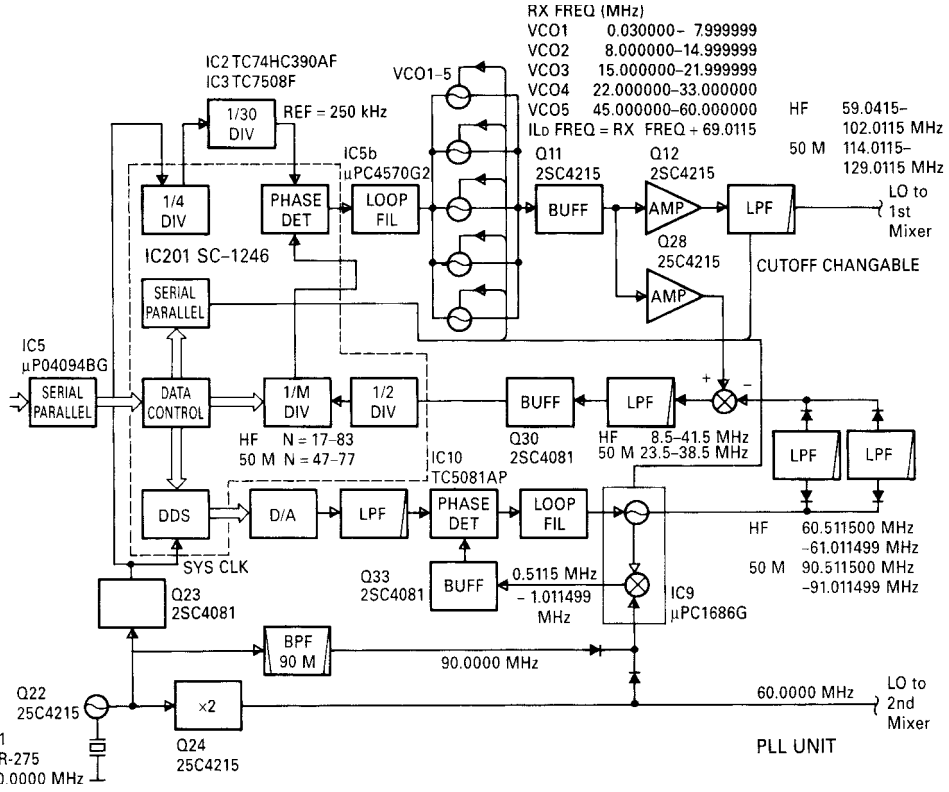


Figure 6-54 Synthesizer used in the ICOM IC-736 HF/6-meter transceiver. The IC-736 combines the DDS and PLL approaches.

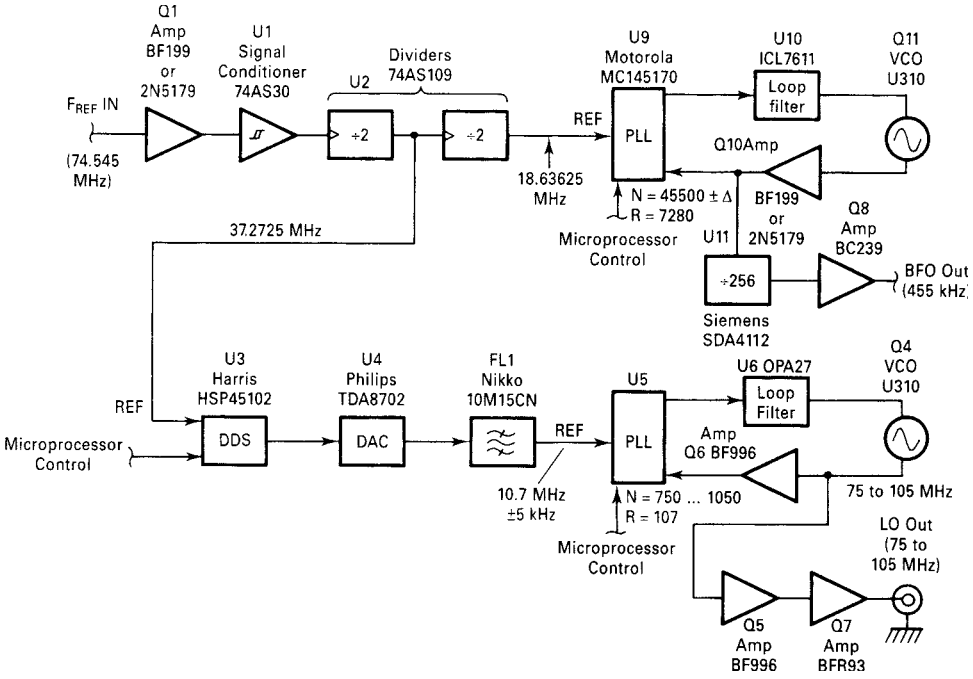


Figure 6-55 Hybrid synthesizer that provides output at about 455 kHz, and from 75 to 105 MHz at approximately 0.01-Hz resolution. This synthesizer uses a combination of a standard PLL and DDS.

the VCO is applied to a distribution-amplifier system (Q5, Q6, Q7). Q5, a dual-gate MOSFET, drives the PLL IC's f_{in} input; Q6, a dual-gate MOSFET, and Q7, a BJT, supply +17 dBm LO drive for the first mixer.

REFERENCES

1. Ulrich L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design* . New York: John Wiley, 1997.
2. W. Egan and E. Clark, "Test Your Charge-Pump Phase Detectors," *Electron. Des.* **26**(12), 134–137, June 7, 1978.
3. Motorola Application Note AN969, "Operation of the MC145145 PLL Frequency Synthesizer with Analog Phase Detector."
4. Alexander W. Hietala, Cary, Illinois; Duane C. Rabe, Rolling Meadows, Illinois Motorola,Inc., Shaumburg, Illinois, "Latched Accumulator Fractional-*N* Synthesis with Residual Error Reduction," U.S. Patent No. 5,093,632, March 3, 1992.
5. Thomas A. D. Riley, Osgoode, Canada, Carleton University, Ottawa, Canada, "Frequency Synthesizers Having Dividing Ratio Controlled Sigma-Delta Modulator," U.S. Patent No. 4,965,531, October 23, 1990.
6. Nigel J. R. King, Wokingham, England, Racal Communications Equipment Limited, England, "Phase Locked Loop Variable Frequency Generator," U.S. Patent No. 4,204,174, May 20, 1980.
7. John Norman Wells, St. Albans, Hertfordshire (GB), Marconi Instruments, St. Albans, Hertfordshire (GB), "Frequency Synthesizers," European Patent No. 012579OB2, July 5, 1995. See

also John N. Wells, "Frequency Synthesizers," Marconi Instruments Ltd., St. Albans, Hertfordshire, Great Britain, U.S. Patent No. 4,609,881, September 2, 1986.

8. Thomas Jackson, Twickenham, Middlesex (GB), Plessey Overseas Ltd., Ilford, Essex (GB), "Improvement In or Relating to Synthesizers," European Patent No. 0214217Bl, June 6, 1996.
9. Thomas Jackson, Twickenham, Middlesex (GB), Plessey Overseas Ltd., Ilford, Essex (GB), "Improvement In or Relating to Synthesizers," European Patent No. W086/05046, August 28, 1996.
10. Byeong-Ha Park, Georgia Institute of Technology, "A Low Voltage, Low Power CMOS 900 MHz Frequency Synthesizer," Ph.D. dissertation, December 1997.
11. Cris E. Hill, "All Digital Fractional-*N* Synthesizer for High Resolution Phase Locked Loops," *Appl. Microwave & Wireless*, **9**(6), 62–69, November/December 1997, **10**(1), 38–42, January/February 1998.

INTERESTING PATENTS

Donald A. Bradley, "Ultra-Low-Noise Microwave Synthesizer," Wiltron Company, Morgan Hill, California, U.S. Patent No. 5,374,902, December 20, 1994.

Soenke Mehrgardt, "Frequency Divider Circuit," Deutsche ITT Industries GmbH, Freiburg, Germany, U.S. Patent No. 4,694,475.

Esteban G. Najle and Robert M. Buckley, "Phase Noise Measurements Utilizing a Frequency Down-Conversion/Multiplier, Direct Spectrum Measurement Technique," Harris Corporation, Melbourne, Florida, and Advanced Testing Technologies, Commack, New York, U.S. Patent No. 5,337,014.

Fractional-*N* Synthesizers

Alain Albarello, Andre Roullet, and Alberto Pimentel, "Fractional-Division Frequency Synthesizer for Digital Angle-Modulation," Thomson-CSF, Paris, France, U.S. Patent No. 4,492,936, January 8, 1985.

Colin Attenborough, "Fractional-*N* frequency Synthesizer with Modulation Compensation," Plessey Overseas Ltd., Ilford, Essex, Great Britain, U.S. Patent No. 4,686,488, August 11, 1987.

Robert J. Bosselaers, "PLL Including an Arithmetic Unit," U.S. Patent No. 3,913,928, October 1975.

R. G. Cox, "Frequency Synthesizer," Hewlett-Packard Company, Palo Alto, California, U.S. Patent No. 2,976,945.

James A. Crawford, "Enhanced Analog Phase Interpolation for Fractional-*N* Frequency Synthesis," Hughes Aircraft Company, Los Angeles, California, U.S. Patent No. 4,586,005, April 29, 1986.

Albert T. Crowley, "PLL Frequency Synthesizer Including Fractional Digital Frequency Divider," RCA Corporation, New York, New York, U.S. Patent No. 4,468,632, August 28, 1984.

Allen P. Edwards, "Low Phase Noise RF Synthesizer," Hewlett-Packard Company, Palo Alto, California, U.S. Patent No. 4,763,083, August 9, 1988

Bar-Giora Goldberg, "Digital Frequency Synthesizer," Sciteq Electronics, Inc., San Diego, California, U.S. Patent No. 4,752,902, June 21, 1988.

Bar-Giora Goldberg, "Digital Frequency Synthesizer Having Multiple Processing Paths," U.S. Patent No. 4,958,310, September 18, 1990.

Bar-Giora Goldberg, "Programmable Fractional-*N* Frequency Synthesizer," Sciteq Electronics, Inc., San Diego, California, U.S. Patent No. 5,224,132, June 29, 1993.

W. G. Greken, "Digital Frequency Synthesizer," General Dynamics, U.S. Patent No. 3,882,403.

Alexander W. Hietala and Duane C. Rabe, "Latched Accumulator Fractional*N* Synthesis with Residual Error Reduction," Motorola, Inc., Schaumburg, Illinois, U.S. Patent No. 5,093,632, March 3, 1992.

Leland Jackson, "Digital Frequency Synthesizer," U.S. Patent No. 3,734,269, May 1973.

Thomas Jackson, "Fractional-*N* Synthesizer," Plessey Overseas Ltd., Ilford, Essex, Great Britain, U.S. Patent No. 4,758,802, July 19, 1988.

Thomas Jackson, "Frequency Synthesizer of the Fractional Type," Plessey Overseas Ltd., Ilford, Essex, Great Britain, U.S. Patent No. 4,800,342, January 24, 1989.

Thomas Jackson, "Improvement In or Relating to Synthesizers," Plessey Overseas Ltd., Ilford, Essex, Great Britain, European Patent No. 0214217B1, June 6, 1996.

Thomas Jackson, "Improvement In or Relating to Synthesizers," Plessey Overseas Ltd., Ilford, Essex, Great Britain, European Patent No. WO86/05046, August 28, 1996.

Nigel J. R. King, "Phase Locked Loop Variable Frequency Generator," Racal Communications Equipment Ltd., Wokingham, England, U.S. Patent No. 4,204,174, May 20, 1980.

Nicholas G. Kingsbury, "Frequency Synthesizer with Fractional Division Ratio and Jitter Compensation," Marconi Company Ltd., Chelmsford, Great Britain, U.S. Patent No. 4,179,670, December 18, 1979.

Charles A. Kingsford-Smith, "Device for Synthesizing Frequencies Which Are Rational Multiples of a Fundamental Frequency," Hewlett-Packard Company, Palo Alto, California, U.S. Patent No. 3,928,813, December 23, 1975.

Frederick L. Martin, "Frequency Synthesizer with Spur Compensation," Motorola, Inc., Schaumburg, Illinois, U.S. Patent No. 4,918,403, April 17, 1990.

Frederick L. Martin, "Frequency Synthesizer with Spur Compensation," Motorola, Inc., Schaumberg, Illinois, U.S. Patent No. 4,816,774, March 28, 1989.

Kenneth D. McCann, "Frequency Synthesizer Having Jitter Compensation," U.S. Philips Corp., New York, New York, U.S. Patent No. 4,599,579, July 8, 1986.

Brian M. Miller, "Multiple-Modulator Fractional-*N* Divider," Hewlett-Packard Company, Palo Alto, California, U.S. Patent No. 5,038,117, August 6, 1991.

Edward J. Nossen, "Digitized Frequency Synthesizer," RCA Corp., New York, New York, U.S. Patent No. 4,206,425, June 3, 1980.

Victor S. Reinhardt and Iradj Shahriary, "Spurless Fractional Divider Direct Digital Frequency Synthesizer and Method," Hughes Aircraft Company, Los Angeles, California, U.S. Patent No. 4,815,018, March 21, 1989.

Joel Remy, "Frequency Synthesizer Including a Fractional Multiplier," Adret Electronique, Paris, France, U.S. Patent No. 4,458,329, July 3, 1984.

Thomas A. D. Riley, "Frequency Synthesizers Having Dividing Ratio Controlled Sigma-Delta Modulator," Carleton University, Ottawa, Canada, U.S. Patent No. 4,965,531, October 23, 1990.

Wayne P. Shepherd, Darell E. Davis, and Wan F. Tay, "Fractional-*N* Synthesizer Having Modulation Spur Compensation," Motorola, Inc., Schaumburg, Illinois, U.S. Patent No. 5,021,754, June 4, 1991.

William J. Tanis, "Frequency Synthesizer Having Fractional-*N* Frequency Divider in PLL," Engelman Microwave, U.S. Patent No. 3,959,737, May 25, 1976.

John Norman Wells, "Frequency Synthesizers," Marconi Instruments, St. Albans, Hertfordshire, Great Britain, European Patent No. 0125790B2, July 5, 1995.

John N. Wells, "Frequency Synthesizers," Marconi Instruments Ltd., St. Albans, Hertfordshire, Great Britain, U.S. Patent No. 4,609,881, September 2, 1986.

Charles E. Wheatley III, "Digital Frequency Synthesizer with Random Jittering for Reducing Discrete Spectral Spurs," Rockwell International Corporation, El Segundo, California, U.S. Patent No. 4,410,954, October 18, 1983.

Mark A. Wheatley, Leslie A. Lepper, and Nigel K. Webb, "Frequency Modulated Phase Locked Loop with Fractional Divider and Jitter Compensation," Racal-Dana Instruments Ltd., Berkshire, Great Britain, U.S. Patent No. 5,038,120, August 6, 1991.

Richard O. Yaeger, "Fractional Frequency Divider," RCA Corp., Princeton, New Jersey, U.S. Patent No. 4,573,176, February 25, 1986.

FURTHER READING

Wei-Zen Chen and Jieh-Tsorng Wu, "A 2 V 1.6 GHz BJT Phase-Locked Loop," *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference*, pp. 563–566.

J. Craninckx and M. Steyaert, "A Fully Integrated CMOS DCS-1800 Frequency Synthesizer," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 372–373.

Bart De Smedt and Georges Gielen, "Nonlinear Behavioral Modeling and Phase Noise Evaluation in Phase Locked Loops," *Proceedings of the 1998 IEEE Custom Integrated Circuits Conference*, pp. 53–56.

Norman M. Filiol, Thomas A. D. Riley, Calvin Plett, and Miles Copeland, "An Agile ISM Band Frequency Synthesizer with Built-in GMSK Data Modulation," *IEEE J. Solid State Circuits*, **SC-33**(7), 998–1007, July 1998.

Fujitsu Microelectronics, Inc., *Super PLL Application Guide*, 1998.

Bar-Giora Goldberg, "Analog and Digital Fractional-*n* PLL Frequency Synthesis: A Survey and Update," *Appl. Microwave & Wireless*, **42**(6), 32–42, June 1999.

New Technique for Analyzing Phase-Locked Loops, June 1975.

Vênceslav F. Kroupa, ed., *Direct Digital Frequency Synthesizers*. New York: IEEE Press, 1999.

Steve Lo, Christian Olgaard, and Dennis Rose, "A 1.8V/3.5mA 1.1GHz/300MHz CMOS Dual PLL Frequency Synthesizer IC for RF Communications," *Proceedings of the 1998 IEEE Custom Integrated Circuits Conference*, pp. 571–574.

G. Palmisano et al., "Noise in Fully-Integrated PLLs," *Proceedings of the Sixth AACD 97*, Como, Italy, 1997.

Byeong-Ha Park and Phillip E. Allen, "A 1GHz, Low-Phase-Noise CMOS Frequency Synthesizer with Integrated LC VCO for Wireless Communications," *Proceedings of the 1998 IEEE Custom Integrated Circuits Conference*, pp. 567–570.

Orlando Peña, "SPICE Tools Provide Behavioral Modeling of PLLs," *Microwaves & RF*, **40**(11), 71–80, November 1997.

Benjamin Sam, "Hybrid Frequency Synthesizer Combines Octave Tuning Range and Millihertz Steps," *Appl. Microwave & Wireless*, **42**(5), 76–84, May 1999

Morris Smith, "Phase Noise Measurement Using the Phase Lock Technique," Wireless Subscriber Systems Group (WSSG) AN1639, Motorola.

V. von Kaenel, D. Aebischer, R. van Dongen, and C. Piguet, "A 600MHz CMOS PLL Microprocessor Clock Generator with a 1.2GHz VCO," *1998 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 396–397.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

APPENDIX A

HBT HIGH-FREQUENCY MODELING AND INTEGRATED PARAMETER EXTRACTION*

This appendix presents, for the first time, a novel nonlinear model for accurate dc, small-signal, and noise characterization of AlGaAs/GaAs heterojunction bipolar transistors (HBTs). A new set of equations are introduced to take into account the bias, temperature, and frequency dependencies in noise calculations. This model provides significant improvement in predicting small-signal noise for HBT-based circuits. We also present an integrated method for accurate HBT model parameter extraction by fitting the dc, multibias *S*-parameter, and noise measurements simultaneously. The extracted model provides accurate small-signal, dc, and noise characteristics. This technique is general and can be used for parameter extraction of other microwave devices such as MESFETs and HEMTs. Our new HBT model is validated using devices from different foundries. An integrated parameter extraction technique is demonstrated for a foundry HBT, and excellent results are obtained.

A-1 INTRODUCTION

With the increasing demand of high-speed and high signal-to-noise ratio (SNR) in analog and digital communication, high-frequency and low-noise circuits and systems are becoming more important. This leads to immense interest in the application of heterojunction bipolar transistors (HBTs) for high-speed digital and microwave circuits. HBTs, combining high transconductance, output resistance, and power density with high breakdown voltage [1], are rapidly becoming viable candidates for low-noise amplifier (LNA) applications across the entire microwave frequency spectrum and well into the millimeter bands. A particularly important application is in monolithic microwave integrated circuits (MMICs), where highly

*©1997 IEEE. Reprinted, with permission, from *IEEE Trans. Microwave Theory Tech.*, Vol. 45, No. 12, December 1997, pp. 2493–2502.

accurate bias- and temperature-dependent noise models are critical for first-pass design success.

Accurate device models are critical to the success of the design of nonlinear microwave circuits such as amplifiers, oscillators, mixers, receivers, and synthesizers [2–5], particularly in MMIC design [6]. Many HBT models have been developed during the last two decades using Monte Carlo methods [7, 8] and models based on the Gummel–Poon (GP) bipolar junction transistor (BJT) model [9, 10]. However, these models suffer from either computationally intensive or inadequate device performance prediction, especially in noise characterization.

Low-noise microwave circuit design engineers are often frustrated that the noise responses of the fabricated circuits are quite different from the results obtained using CAD programs. Several design passes may be necessary to meet specifications. Even in mixer designs, the noise of the fabricated circuits may be significantly higher than that predicted by the CAD tool. The discrepancy arises primarily from the poor estimation of the device noise figure.

Parameter extraction by fitting the model responses to measurements is the primary method of obtaining the model parameter values of equivalent circuit models. Conventionally, parameter extraction is based on dc, *S*-parameter, and large-signal measurements (e.g., [11–13]). The models extracted are suitable for dc, small-signal, and large-signal analysis. However, noise analysis with these models often contains substantial errors.

In this appendix, a novel HBT model for high-frequency and low-noise applications is presented. This model is based on a new formulation that takes into account the bias, temperature, and frequency dependencies of the device noise characteristics to provide accurate performance prediction. Two types of noise equivalent circuits, a linearized T model and a hybrid-π model, are derived from the modified GP model described in the literature [14–17]. Self-heating effects are also taken into consideration using the thermal equivalent circuits.

We also present an integrated parameter extraction approach by fitting the model responses to the dc, multibias *S*-parameter, and multibias noise measurements simultaneously. The models extracted not only provide precise characterization of device small-signal and dc responses but also give accurate prediction of device noise performance. The models can be used in a variety of applications for small-signal and large-signal analysis as well as noise analysis. The advantages of our approach become more significant in the design of low-noise microwave circuits.

Our new HBT model is presented in Section A-2. The integrated parameter extraction technique is addressed in Section A-3. HBT model validation using devices from various foundries is given in Section A-4. Section A-5 demonstrates the HBT model parameter extraction for a foundry device. Two models are extracted: *Model 1* is based on dc and *S*-parameter measurements, while *Model 2* is based on dc, *S*-parameter, and noise measurements. Model 1 shows large errors in noise performance even though the dc and *S*-parameter fits are good. Model 2 provides accurate prediction in dc, *S*-parameter, and noise responses.

A-2 HIGH-FREQUENCY HBT MODELING

The equivalent circuit of our HBT model consists of three parts: intrinsic circuit, device parasitics, and package parasitics as shown in Figure A-1. The device and package parasitics play a very important role in high-frequency modeling. Without considering these parasitics, the model would not be able to predict the ac performance accurately. A typical equivalent

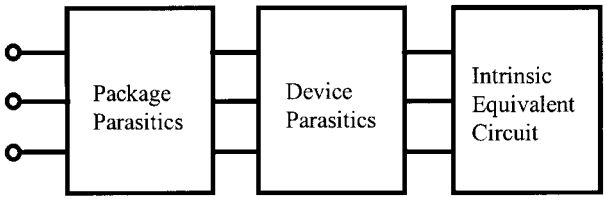


Figure A-1 The three parts of the equivalent circuit for the HBT model.

circuit representing package parasitics is shown in Figure A-2. The schematic of the intrinsic, extrinsic, and thermal circuits for our HBT model is depicted in Figure A-3. This circuit is a modified GP model implemented in SPICE [18]. The base resistor R_{bb} in the original GP model is split into R_{b1} and R_{b2}, and a base capacitor C_{bx} is added between these two resistors. This modification is necessary for accurate HBT characterization at high frequencies. Additional base resistors and capacitors have been added to the model in [19] for the SPICE equivalent circuits. However, these components are external to the device and noise correlation is not able to be taken into account in SPICE simulators. In our work, R_{b1}, R_{b2}, and C_{bx} are inside the device model incorporated into the harmonic-balance simulator [20], which takes into account the noise correlation and provides a much more accurate noise prediction. The detailed dc, capacitance, and transient-time equations of the model are available in [20], which is derived from [10,15].

A-2-1 dc and Small-Signal Model

Using the intrinsic base–emitter voltage V_{be} and collector–emitter voltage V_{ce} shown in Figure A-3 as state variables, we can write the HBT model equations as

$$I_B = I_B(\phi, V_{be}, V_{ce}, T_j(\phi, V_{be}, V_{ce})) \tag{A-1a}$$

$$I_C = I_C(\phi, V_{be}, V_{ce}, T_j(\phi, V_{be}, V_{ce})) \tag{A-1b}$$

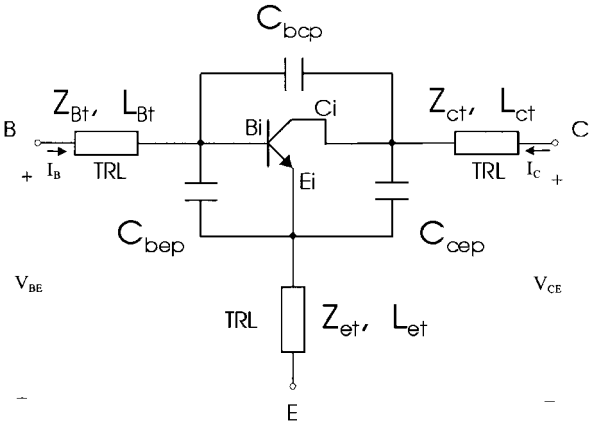


Figure A-2 A typical equivalent circuit representing package parasitics of HBTs.

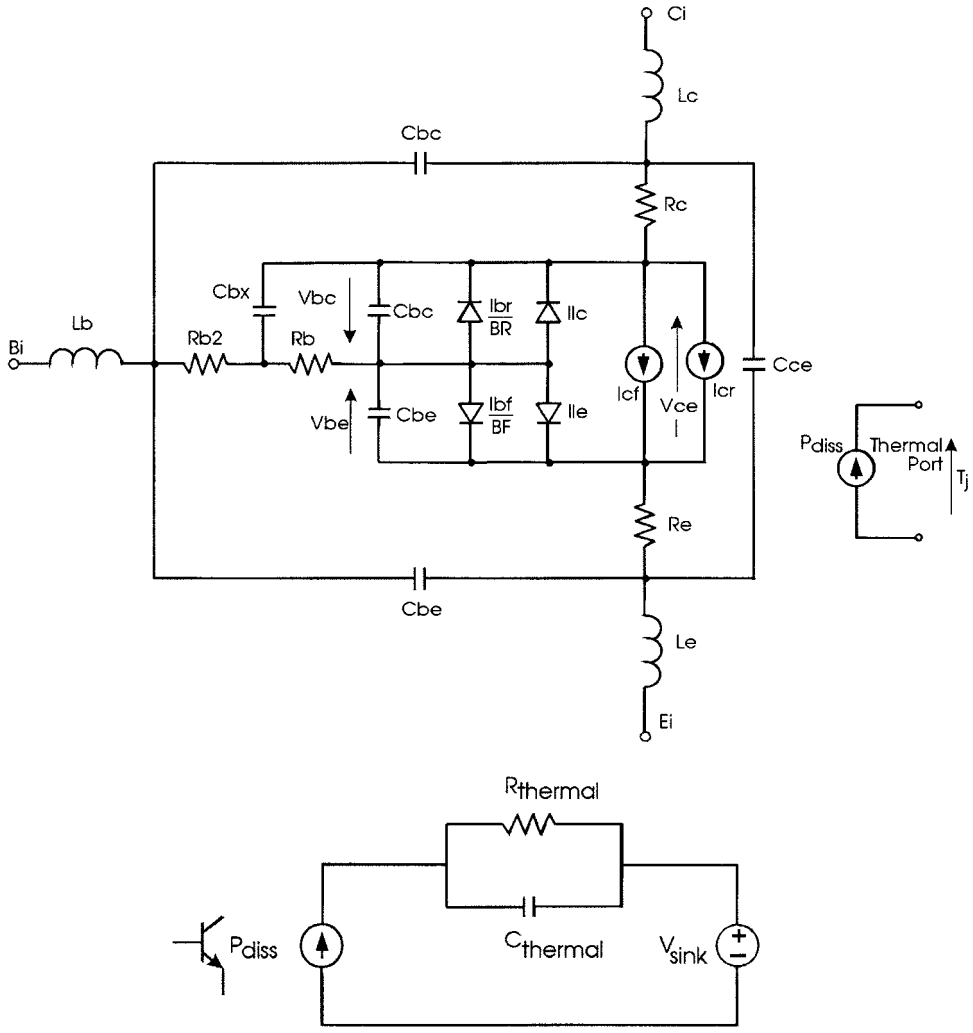


Figure A-3 Schematic of the intrinsic, extrinsic, and thermal circuits for the HBT model.

$$V_{BE} = V_{BE}(\phi, V_{be}, V_{ce}, T_j(\phi, V_{be}, V_{ce})) \tag{A-1c}$$

$$V_{CE} = V_{CE}(\phi, V_{be}, V_{ce}, T_j(\phi, V_{be}, V_{ce})) \tag{A-1d}$$

which indicate that the extrinsic base current I_B, collector current I_C, base–emitter voltage V_{BE}, and the collector–emitter voltage V_{CE}, as shown in Figure A-2, are functions of model parameters ϕ, state variables V_{be} and V_{ce}, and temperature T_j. The device temperature T_j is also dependent on the model parameters and state variables. Any two equations in Eqs. (A1a)–(A1d) can be used to solve for the state variables. We use Eqs. (A-1a) and (A-1d) with a Newton method in our dc simulation to obtain the dc operating points and produce constant-I_B versus V_{CE} curves. For a fixed set of parameters ϕ, the Newton iteration is written as

$$\begin{bmatrix} V_{be} \\ V_{ce} \end{bmatrix}^{k+1} = \begin{bmatrix} V_{be} \\ V_{ce} \end{bmatrix}^{k} - \left(\begin{bmatrix} \dfrac{\partial I_B}{\partial V_{be}} & \dfrac{\partial V_{CE}}{\partial V_{be}} \\ \dfrac{\partial I_B}{\partial V_{ce}} & \dfrac{\partial V_{CE}}{\partial V_{ce}} \end{bmatrix}^{k} \right)^{-1} \begin{bmatrix} I_B - I_B^M \\ V_{CE} - V_{CE}^M \end{bmatrix}^{k} \tag{A-2}$$

where I_B^M and V_{CE}^M are the measured or applied base bias current and collector–emitter voltage. The entries to the Jacobian matrix can be obtained by deriving Eqs. (A1a) and (A1d) with respect to the corresponding state variable; for instance,

$$\frac{\partial I_B}{\partial V_{be}} = \frac{\partial I_B}{\partial V_{be}} + \frac{\partial I_B}{\partial T_j}\frac{\partial T_j}{\partial V_{be}} \tag{A-3}$$

In our implementation, these entries are calculated analytically, which significantly reduces the computational time compared to the perturbation method used in other implementations (e.g., [10, 15]).

After the dc bias point is obtained, a small-signal equivalent circuit is then established by linearizing the model at the bias point, and this is used for *S*-parameter and noise calculations. Two types of small-signal models are discussed in the following sections. As we will see, the bias, temperature, and frequency dependencies are taken into account in model evaluations for small-signal and noise simulation.

A-2-2 Linearized T Model

The equivalent circuit for the linearized T model is sketched in Figure A-4. We add a collector–emitter resistor R_{ce} into the conventional T model [18] to provide more accurate device performance prediction. R_{ce} is also critical in parameter extraction for a good match between modeled and measured responses. The intrinsic linearized parameters A_0, C_i, R_c, R_e, C_e, and R_{ce} are functions of ϕ, V_{be}, V_{ce}, and T_j obtained in the dc simulation. For example,

$$A_0 = A_0(\phi, V_{be}, V_{ce}, T_j(\phi, V_{be}, V_{ce})) = \left. \frac{\dfrac{\partial I_c'}{\partial V_{be}}}{\dfrac{\partial I_b'}{\partial V_{be}} + \dfrac{\partial I_c'}{\partial V_{be}}} \right|_{\phi, V_{be}, V_{ce}, T_j = \text{constant}} \tag{A-4}$$

where

$$I_c' = I_{cf} - I_{cr}$$

$$I_b' = \frac{I_{bf}}{BF} + I_{le}$$

I_{cf}, I_{cr}, I_{bf}, and I_{le} are illustrated in Figure A-3, and BF is the ideal forward current gain.

For the noise responses, the minimum noise figure $F_{\min}$, the optimal noise reflection coefficient Γ_{opt} and the equivalent normalized noise resistance R_N are calculated using the following equations [16, 17]:

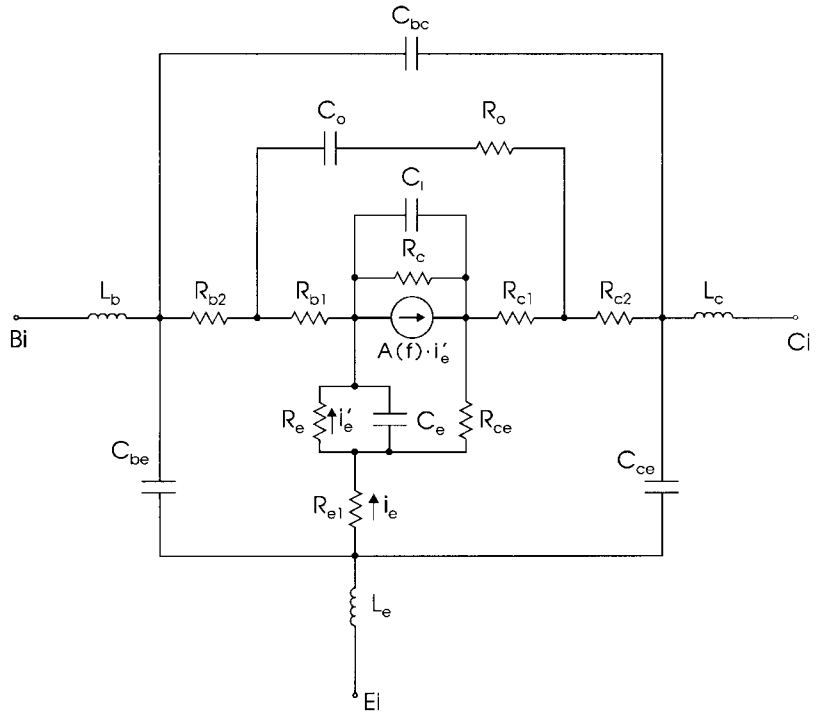


Figure A-4 The linearized T model for small-signal and noise calculations.

$$F_{\min} = a\,\frac{R_{b1} + R_{\text{opt}}}{R_e} + \left(1 + \frac{f^2}{f_b^2}\right)\frac{1}{A_0} + \frac{F_C}{1+f} \tag{A-5}$$

The optimum source resistance is calculated by

$$R_{\text{opt}} = \left[R_{b1}^2 - X_{\text{opt}}^2 + \left(1 + \frac{f^2}{f_b^2}\right)\frac{R_e(2R_{b1} + R_e)}{A_0 a}\right]^{1/2} \tag{A-6}$$

the optimum source reactance is

$$X_{\text{opt}} = \left(1 + \frac{f^2}{f_b^2}\right)\frac{2\pi f C_e R_e^2}{A_0 a} \tag{A-7}$$

and the normalized noise resistance is

$$R_N = R_{b1}\left(\alpha + 1 - \frac{1}{A_0}\right) + \frac{R_e}{2}\left\{\alpha + \frac{R_{b1}^2}{R_e^2}\left[1 - A_0 + \frac{f^2}{f_b^2} + \frac{f^2}{f_e^2} + \left(\frac{1}{A_0} - 1 - \frac{f^2}{f_b f_e}\right)^2\right]\right\} \tag{A-8}$$

where f is the operating frequency, f_b is the cutoff frequency of the ac common-base current gain $A(f)$, A_0 is the dc current gain, and F_C is the flicker noise corner frequency. The flicker

noise used for small-signal noise calculation is different from the one used for large-signal noise evaluation [20]. The coefficients α, a, and f_e are calculated as follows:

$$\alpha = \frac{1 + (f/f_b)^2}{A_0^2} \tag{A-9}$$

$$a = \left[\left(1 + \frac{f^2}{f_b^2}\right)\left(1 + \frac{f^2}{f_b^2}\right) - A_0\right]\frac{1}{A_0} \tag{A-10}$$

$$f_e = \frac{1}{2\pi N_{\text{fac}} C_e R_e^2} \tag{A-11}$$

where N_{fac} is a noise factor.

In addition to the above calculation, the effect of the device and package parasitic circuits is added to obtain the overall noise response using a common two-port noise evaluation method.

A-2-3 Linearized Hybrid-π Model

The equivalent circuit for the linearized hybrid-π model is shown in Figure A-5. The intrinsic linearized parameters Y_{be}, Y_{bc}, G_{m0}, and G_0 are functions of model parameters ϕ, V_{be}, V_{ce}, and T_j obtained in the dc simulation. For example,

$$G_{m0} = G_{m0}\left(\phi, V_{be}, V_{ce}, T_j(\phi, V_{be}, V_{ce})\right) = \left.\frac{\partial I_c'}{\partial V_{be}}\right|_{\phi, V_{be}, V_{ce}, T_j = \text{constant}} \tag{A-12}$$

The intrinsic noise model of this hybrid-π circuit is developed by incorporating the correlation of the noise sources and the frequency dependency of the noise sources. The normalized noise correlation matrix can be written

$$C = \frac{1}{4KT_jB}\begin{bmatrix} \overline{i_b i_b^*} & \overline{i_b i_c^*} \\ \overline{i_c i_b^*} & \overline{i_c i_c^*} \end{bmatrix} \tag{A-13}$$

where

$$\overline{i_b i_b^*} = 4KT_j\text{Re}[Y_{11}]B - 2qI_bB \tag{A-14a}$$

$$\overline{i_c i_b^*} = 2KT_j(G_m - 2\text{Re}[Y_{12}])B - 2qI_cB \tag{A-14b}$$

$$\overline{i_c i_c^*} = 4KT_j\text{Re}[Y_{22}]B + 2qI_cB \tag{A-14c}$$

where K is the Boltzmann constant, B is the bandwidth, q is the electron charge, and I_b and I_c are the intrinsic base and collector currents, respectively. Y_{11}, Y_{12}, and Y_{22} correspond to the intrinsic Y-matrix of the intrinsic model, as shown in Figure A-5 (region within the heavy box), and $\text{Re}[Y_{ij}]$ represents the real part of Y_{ij}.

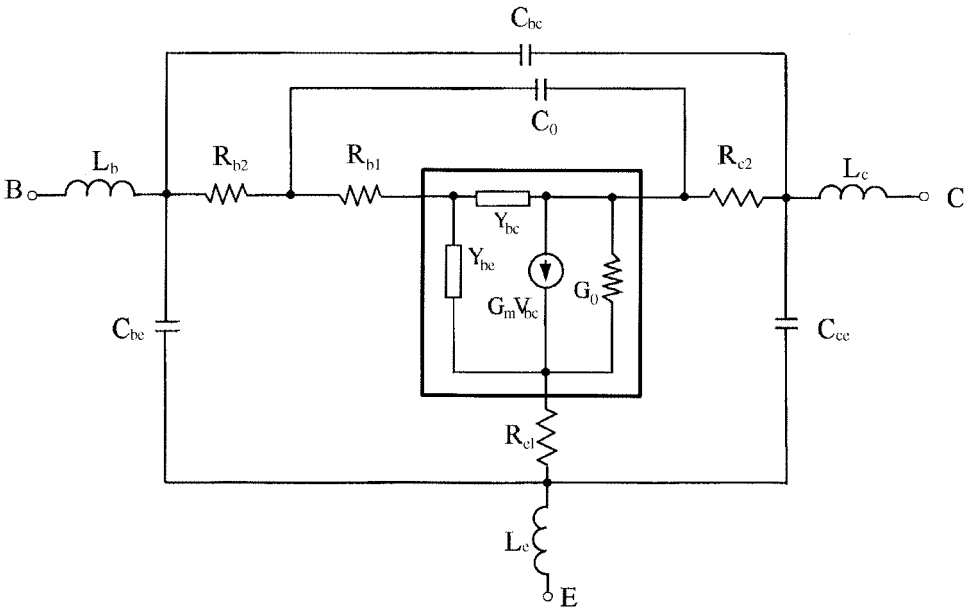


Figure A-5 The linearized hybrid-π model for small-signal and noise calculations.

At zero frequency (dc), the noise power terms $\overline{i_b i_b^*}$ and $\overline{i_c i_c^*}$ of the intrinsic device reduce to the SPICE shot noise model. Also, the cross-correlation terms reduce to zero, indicating that there is no noise component common to both the collector and base currents.

The intrinsic noise parameters (R_N, $F_{\min}$, Γ_{opt}) are derived from the correlation matrix C of Eq. (A-13) as

$$R_N = C_{11} \tag{A-15}$$

$$G_{\text{opt}} = \sqrt{\frac{C_{22}}{C_{11}} - \left(\frac{\text{Im}(C_{12})}{C_{11}}\right)^2} \tag{A-16}$$

$$B_{\text{opt}} = \frac{\text{Im}(C_{12})}{C_{11}} \tag{A-17}$$

$$F_{\min} = 1 + 2\sqrt{C_{11}C_{22} - \text{Im}(C_{12})^2 + \text{Re}(C_{12})} \tag{A-18}$$

where $\text{Re}(C_{12})$ and $\text{Im}(C_{12})$ are the real part and imaginary part of C_{12}, respectively.

A-3 INTEGRATED PARAMETER EXTRACTION

In this section we present integrated parameter extraction by fitting the dc, multibias *S*-parameter, and noise measurements simultaneously.

A-3-1 Formulation of Integrated Parameter Extraction

The objective function for optimization in our integrated parameter extraction includes dc, *S*-parameter, and noise measurements. It is formulated as

$$E_{\text{total}}(\phi) = E_{\text{dc}}(\phi) + E_S(\phi) + E_N(\phi) \tag{A-19}$$

where ϕ are the model parameters to be extracted, E_{total} is the total error, and

$$E_{\text{dc}}(\phi) = \sum_{i=1}^{N_{\text{dc}}} \sum_{k,l}^{N_P} \| W_{\text{dc}_{ikl}} [R_{\text{dc}_{ikl}}(\phi) - M_{\text{dc}_{ikl}}] \| \tag{A-20a}$$

$$E_S(\phi) = \sum_{i=1}^{N_{S\text{dc}}} \sum_{j=1}^{N_{SF}} \sum_{k,l}^{N_P} \| W_{S_{ijkl}} [R_{S_{ijkl}}(\phi) - M_{S_{ijkl}}] \| \tag{A-20b}$$

$$E_N(\phi) = \sum_{i=1}^{N_{N\text{dc}}} \sum_{j=1}^{N_{NF}} \sum_{k}^{N_{NP}} \| W_{N_{ijk}} [R_{N_{ijk}}(\phi) - M_{N_{ijk}}] \| \tag{A-20c}$$

are the dc error, *S*-parameter error, and noise error, respectively. N_P is the number of ports of the device. N_{dc} is the number of dc measurement points. $N_{S\text{dc}}$ and N_{SF} are, respectively, the number of bias points and frequencies at which the *S*-parameter measurements are taken. N_{NP} is the number of noise parameters and $N_{N\text{dc}}$ and N_{NF} are, respectively, the number of bias points and frequencies at which the noise measurements are taken. The *R*'s and *M*'s are the model responses and the corresponding measurements. The *W*'s are the weighting factors applied to the corresponding errors. The $\|\bullet\|$ denotes the least-squares norm. A Levenberg–Marquardt optimization process was applied to

$$\underset{\phi}{\text{minimize}} \quad E_{\text{total}}(\phi) \tag{A-21}$$

The dc, multibias *S*-parameter, and multibias noise measurements are simultaneously matched to the model responses during optimization. A direct parameter extraction and systematic optimization strategy [21] can be used for efficient model extraction.

A-3-2 Model Optimization

The optimization scheme consists of two iteration loops. The first iteration loop is for dc simulation to obtain the dc operation point as described in Section A-2. The second loop is used for adjusting the model parameters ϕ to match the model responses to the measurements. This procedure can be described by the following algorithm.

Step 1: Initialize the model parameter ϕ.
Step 2: Initialize the state variables V_{be} and V_{ce}.
Step 3: Perform dc simulation at each bias point to obtain the solution for state variables using the Newton iteration of Eq. (A-2).
Step 4: Obtain the linearized small-signal circuit and calculate the model responses including the *S*-parameters and noise parameters at each bias point and frequency.

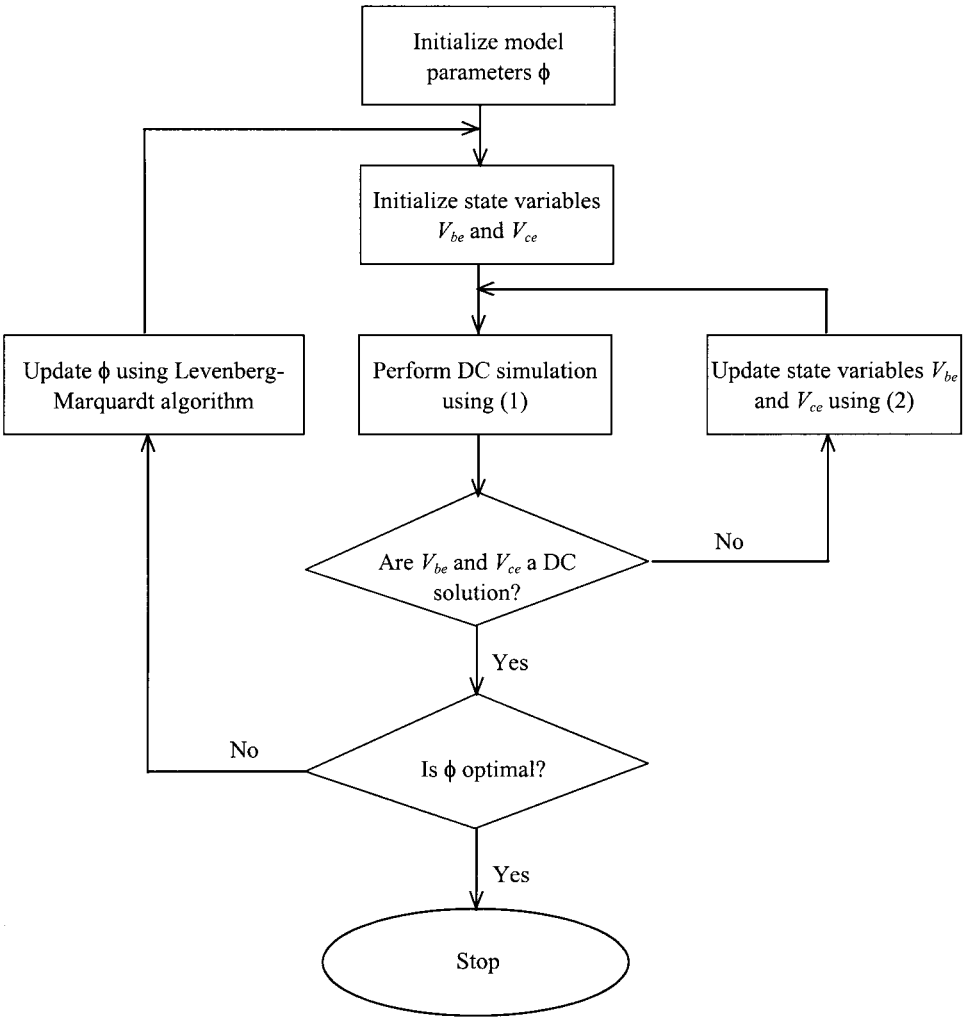


Figure A-6 Flowchart for model optimization.

Step 5: Calculate the objective function for optimization using Eqs. (A-19) and (A-20).
Step 6: If ϕ is optimal, stop. Otherwise, update ϕ according to the Levenberg–Marquardt optimization algorithm and go to *Step 2*.

This algorithm is illustrated by the flowchart in Figure A-6.

A-4 NOISE MODEL VALIDATION

We use the linearized hybrid-π model for noise model validation in this section and leave the linearized T model for the integrated parameter extraction example, which will be discussed in Section A-5.

To validate the model, on-wafer bias- and temperature-dependent noise parameter measurements are performed using the ATN-NP5 system for various foundries. Parameter extraction is performed using the procedure and model described above.

Figure A-7 compares the measured and modeled noise parameters for a foundry HBT biased at V_{CE} = 2 V and I_B = 428 μA. Two models are compared with the measured result: our model (referred to as Mod-PI) and the SPICE shot noise model (referred as MOD-SPICE), which is used in SPICE and harmonic-balance simulators. From the results it is evident that at lower frequencies both models show the same predictions; however, for the reasons explained earlier, at higher frequencies the SPICE noise model deviates rapidly from the measurements, while our model is very consistent with the measurements.

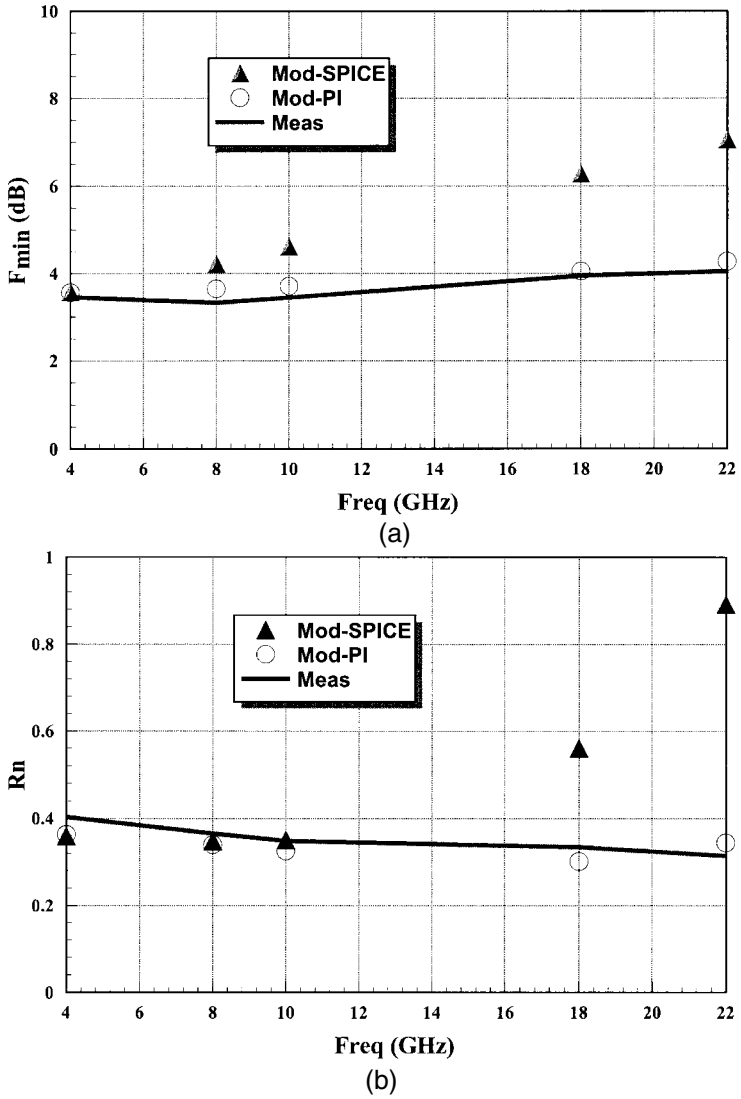
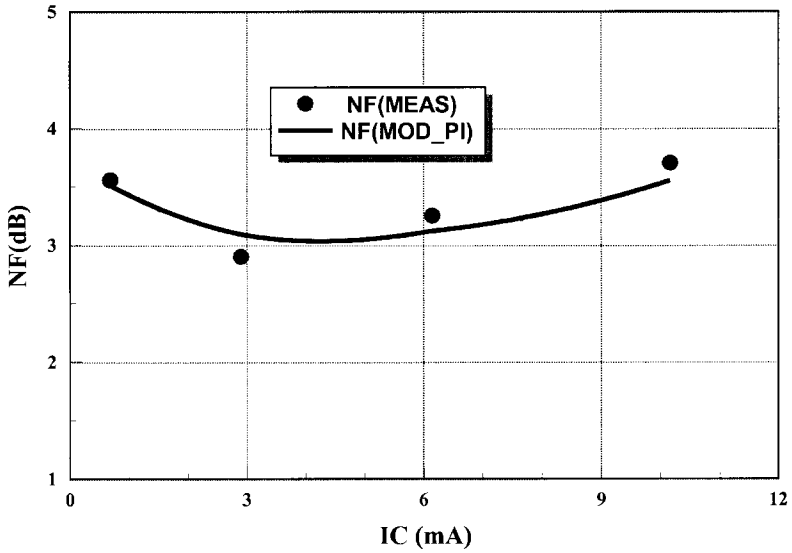
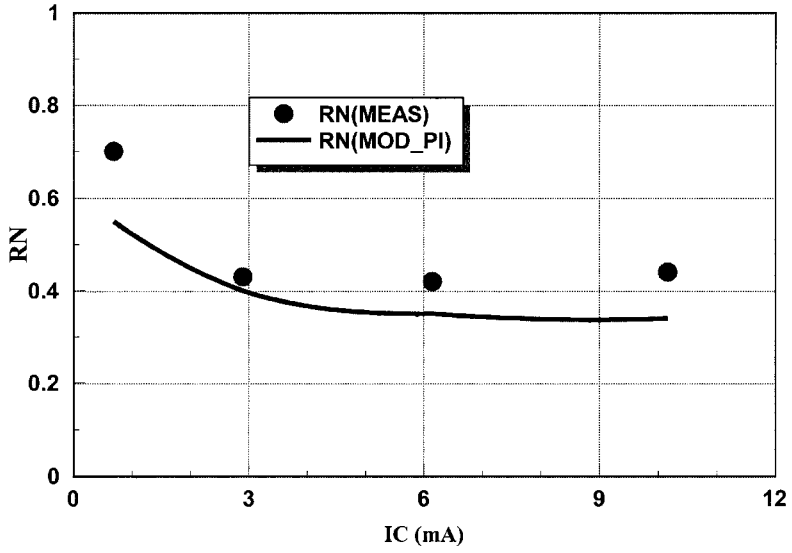


Figure A-7 Comparison of measured and modeled (Ansoft model and SPICE) noise parameters: (a) $F_{\min}$ and (b) R_n as a function of frequency for V_{CE} = 2.0 V and I_B = 428 μA.

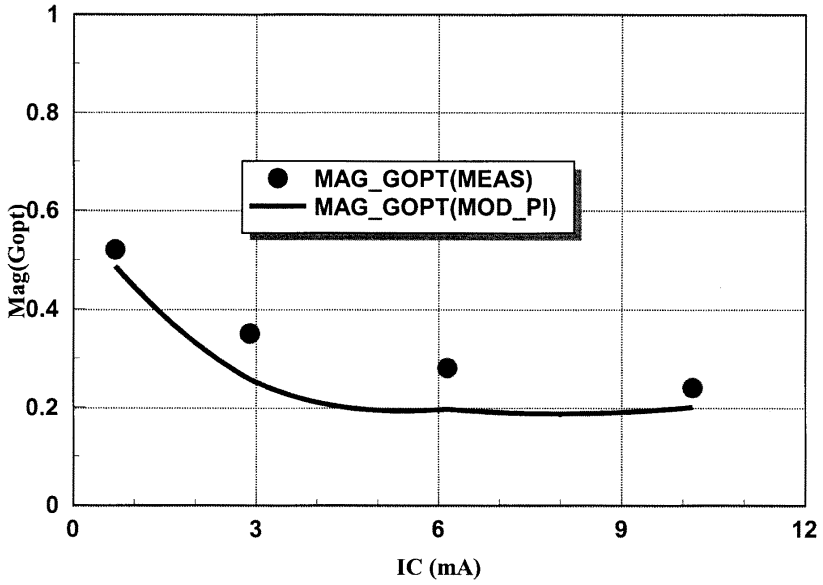


(a)

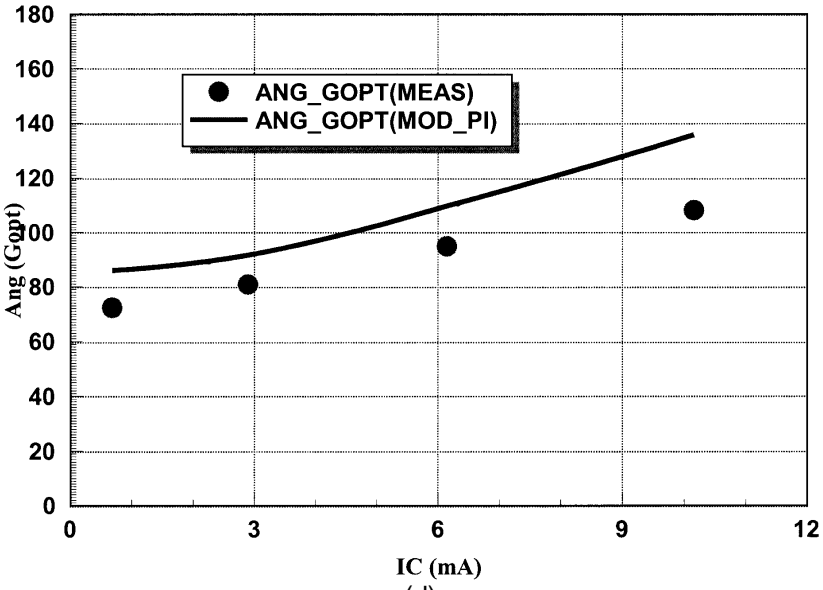


(b)

Figure A-8 Bias dependency of the noise parameters versus collector current at 10 GHz, with bias at V_{CE} = 2 V and I_B varied from 50 to 200 μA. The graphs are (a) noise figure, (b) R_N, (c) $|\Gamma_{opt}|$, and (d) $\angle|\Gamma_{opt}|$.



(c)



(d)

Figure A-8 (*Continued*)

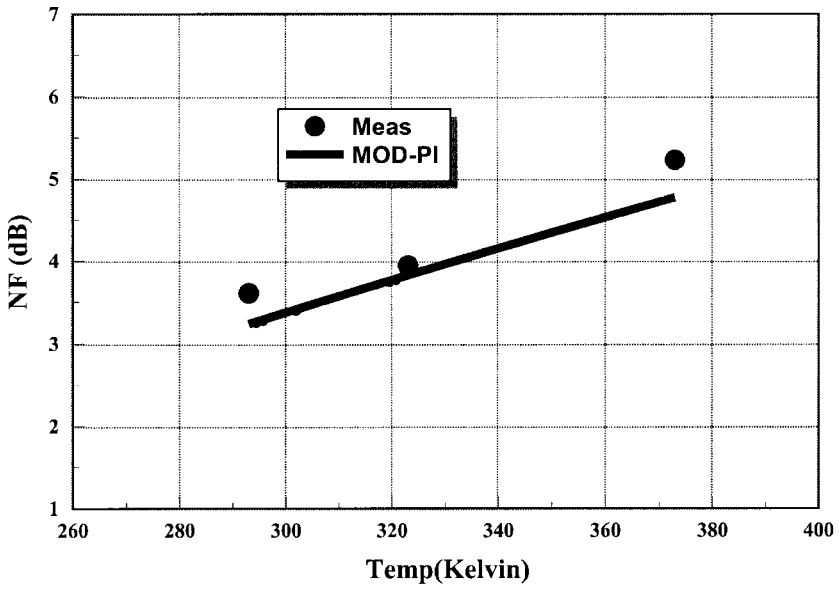


Figure A-9 Temperature dependency of the noise figure at 10 GHz, with the device biased at V_{CE} = 1 V and I_B = 100 μA. Temperature was varied from 20 to 100 °C.

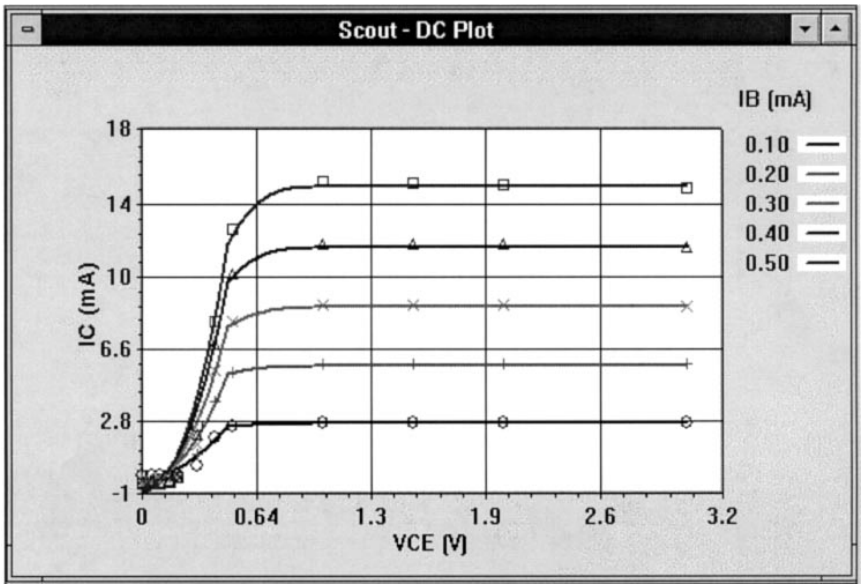
The next validation process is the model's accuracy with bias. For this case, the base current is varied from 50 to 200 μA, and V_{CE} is set to 2 V. The frequency is 10 GHz. Figure A-8 illustrates the bias-dependent noise characteristics for a foundry HBT. Very good correlation is obtained between the measured and modeled noise parameters.

Finally, the temperature-dependent model is validated as shown in Figure A-9. The noise figure of the HBT is plotted versus temperature in kelvins. In this case, the device is biased at V_{CE} =1 V and I_B = 100 μA and the temperature varied from 293 to 373 K (20 to 100 °C). We can see a very good match between the measured and modeled results.

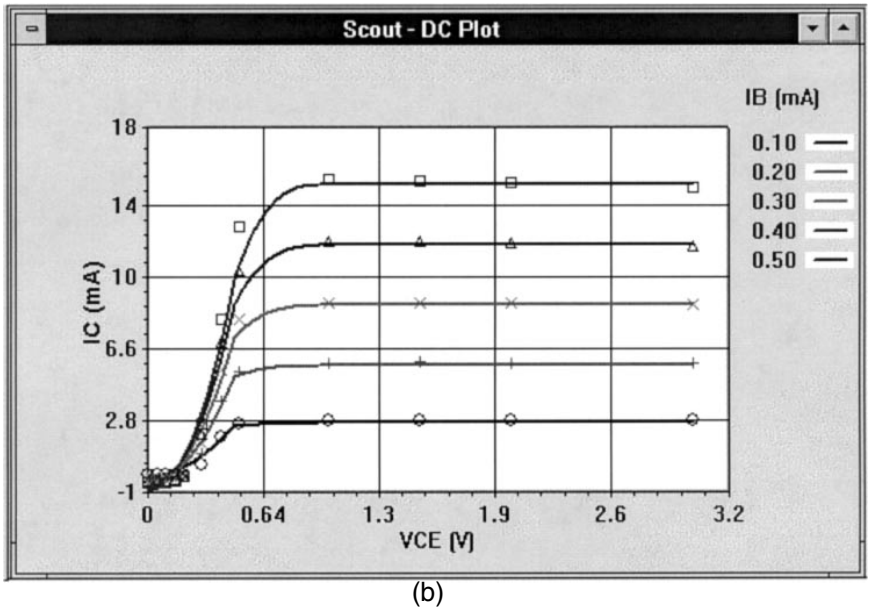
A-5 PARAMETER EXTRACTION OF AN HBT MODEL

Parameter extraction of an HBT model is carried out using the linearized T model illustrated in Section A-2.

The data used for parameter extraction include dc measurements at 60 points, and *S*-parameter measurements and noise measurements at 9 bias points and 17 frequencies. Two models are extracted: Model 1 using dc and *S*-parameter measurements and Model 2 using dc, *S*-parameter, and noise measurements. The modeled and measured dc responses are plotted in Figure A-10. The modeled and measured *S* parameters at three bias points are shown in Figures A-11 through A-13. The modeled and measured minimum noise figures at three bias points are shown in Figures A-14 through A-16. From Figures A-11 through A-13 we can see that the dc responses and *S* parameters of both models match the measurements very well. However, the noise responses of Model 2 are much closer to the measurements than those of Model 1, which are illustrated in Figures A-14 through A-16. If both models are used for noise analysis in circuit design, Model 2 will provide much more accurate results than Model 1.

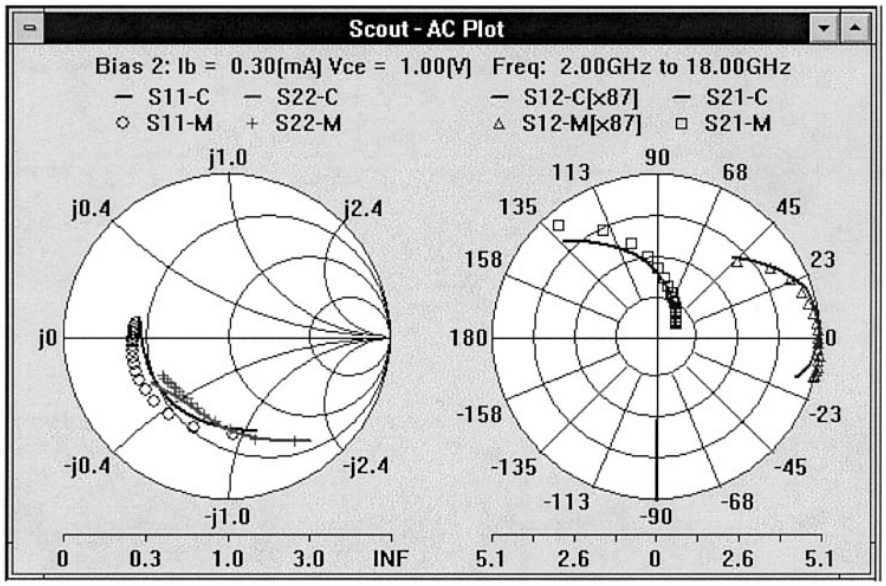


(a)

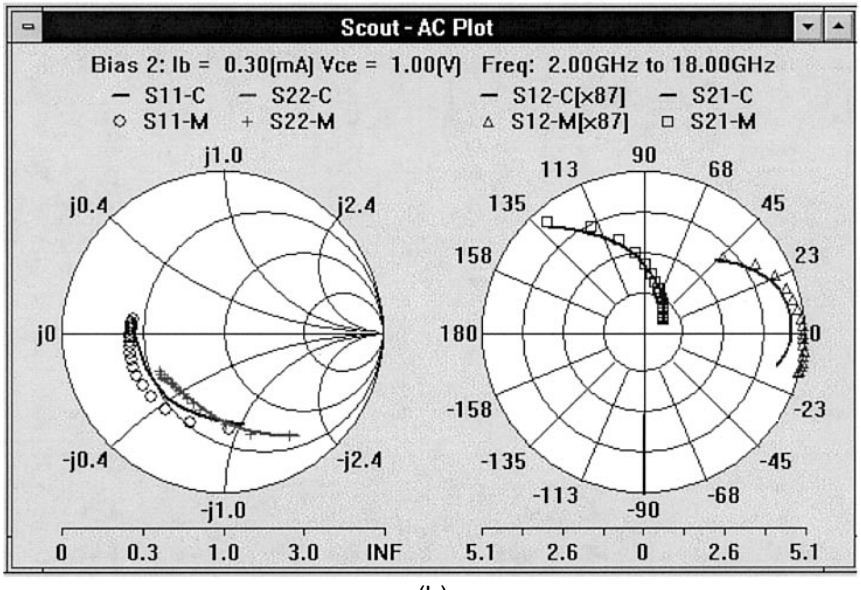


(b)

Figure A-10 Measured (discrete points) and modeled (solid lines) dc *I–V* curves of (a) Model 1 and (b) Model 2.

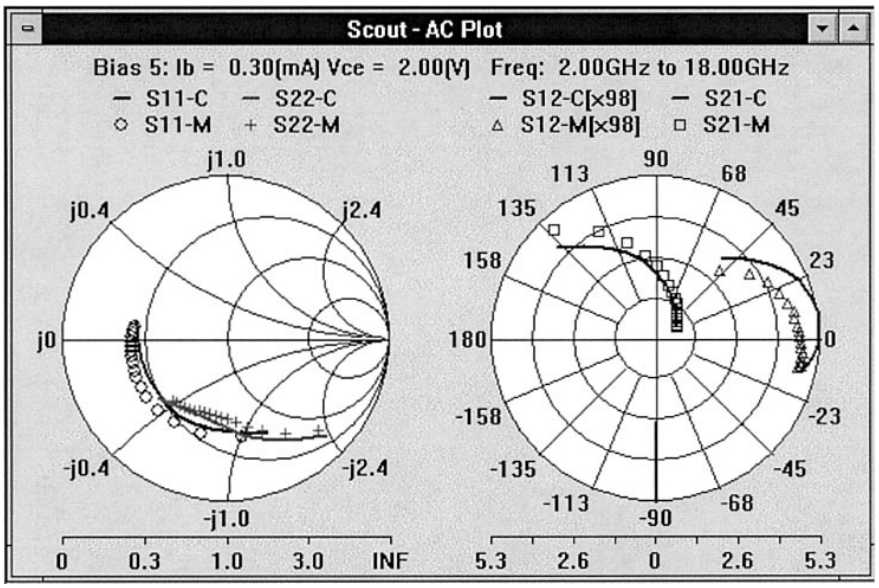


(a)

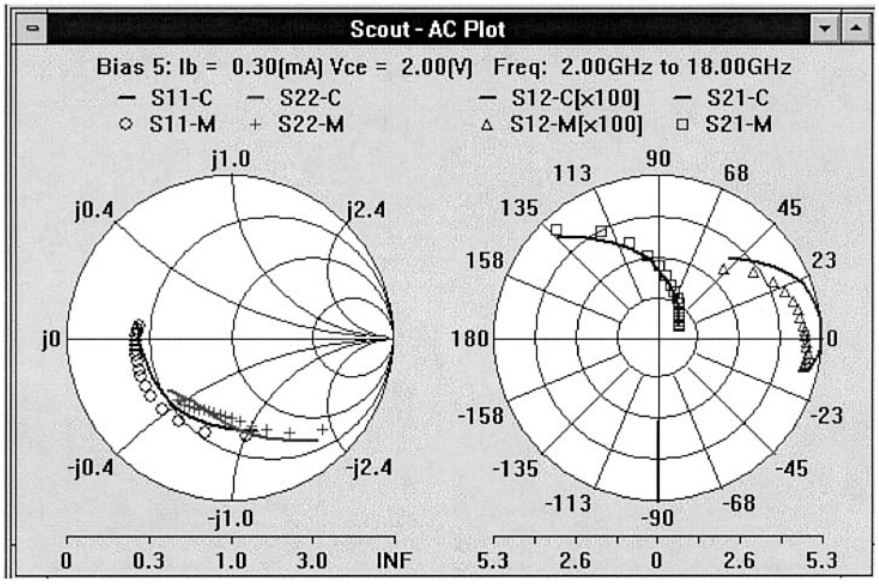


(b)

Figure A-11 Measured (discrete points) and modeled (solid lines) *S* parameters of (a) Model 1 and (b) Model 2 at bias point $I_B = 0.3$ mA and $V_{CE} = 1$ V.



(a)



(b)

Figure A-12 Measured (discrete points) and modeled (solid lines) *S* parameters of (a) Model 1 and (b) Model 2 at bias point $I_B = 0.3$ mA and $V_{CE} = 2$ V.

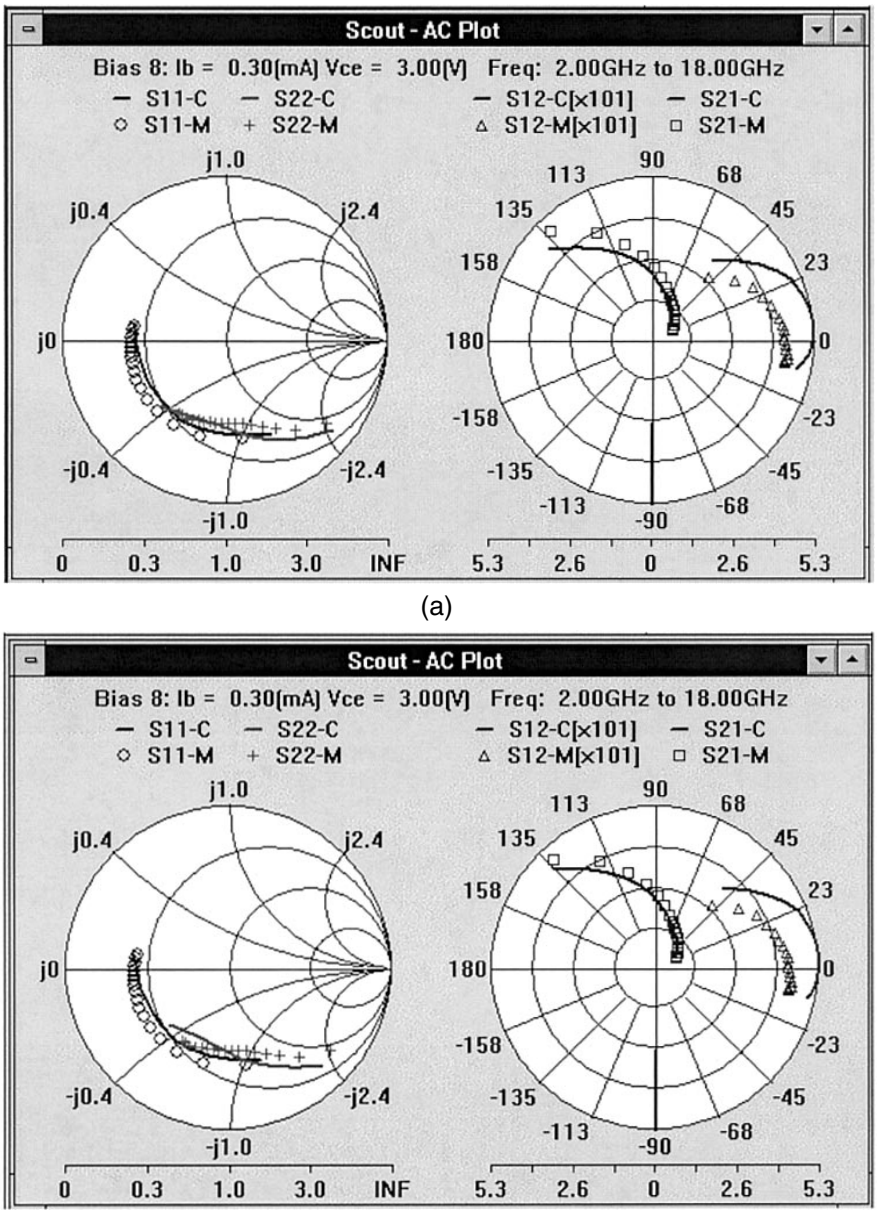
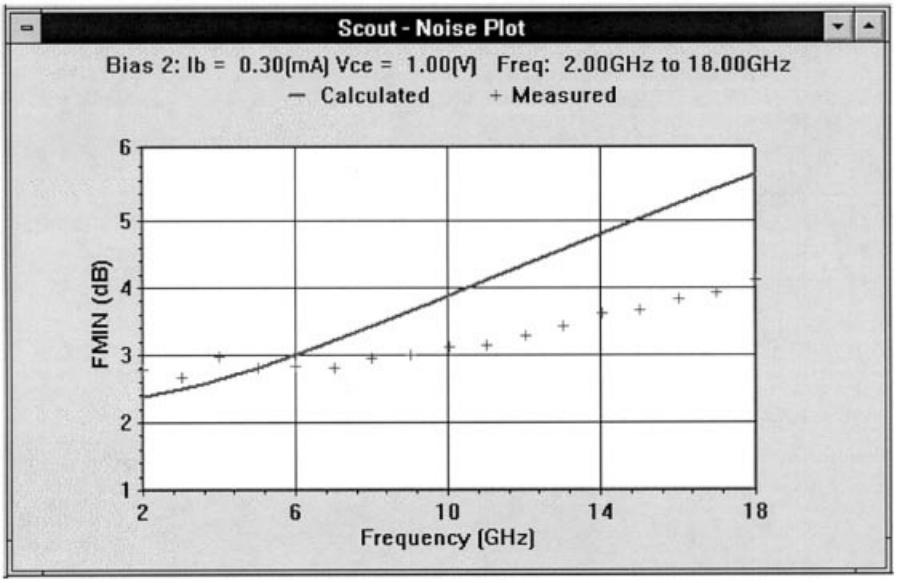
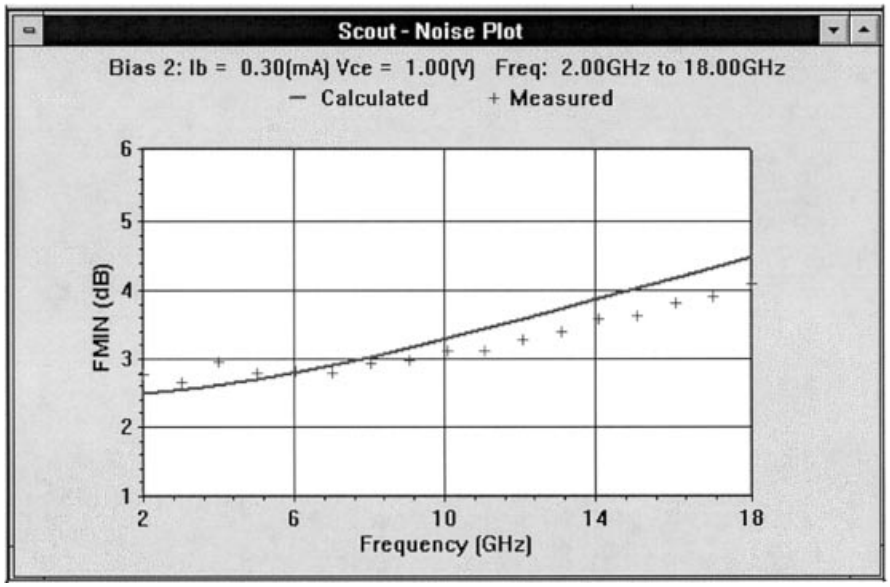


Figure A-13 Measured (discrete points) and modeled (solid lines) *S* parameters of (a) Model 1 and (b) Model 2 at bias point $I_B = 0.3$ mA and $V_{CE} = 3$ V.

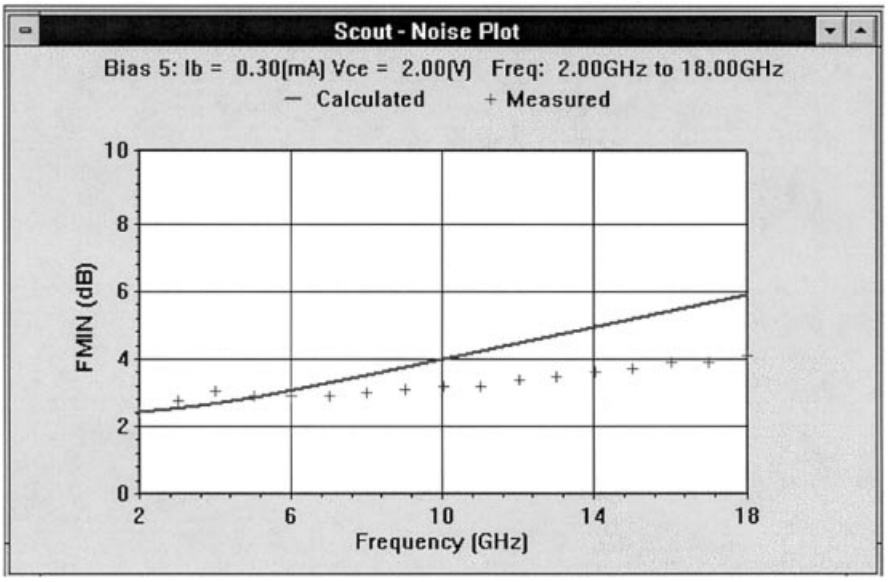


(a)

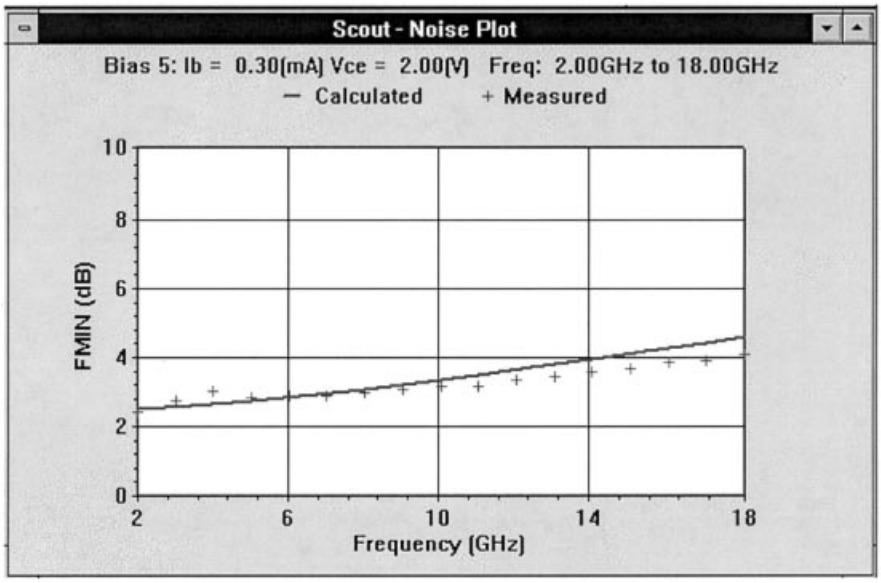


(b)

Figure A-14 Measured (discrete points) and modeled (solid lines) minimum noise figures of (a) Model 1 and (b) Model 2 at bias point $I_B = 0.3$ mA and $V_{CE} = 1$ V.

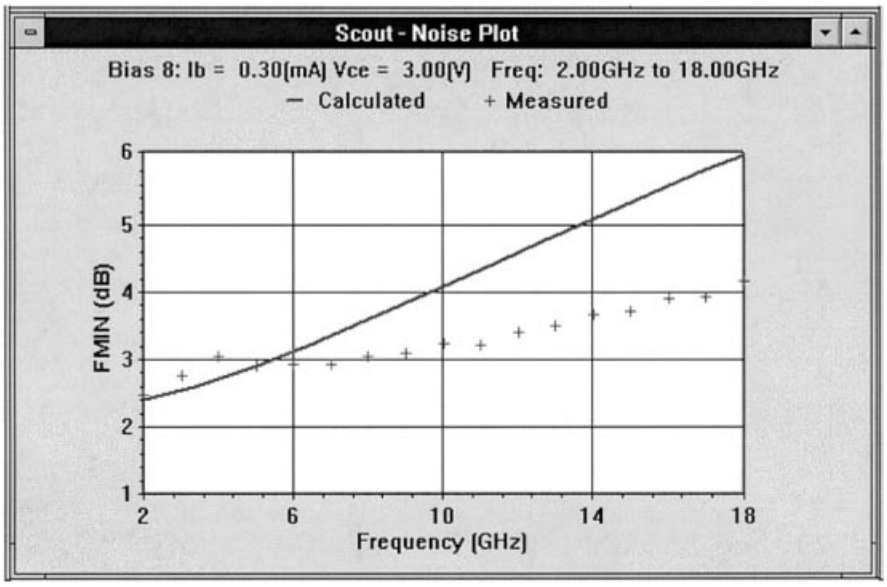


(a)

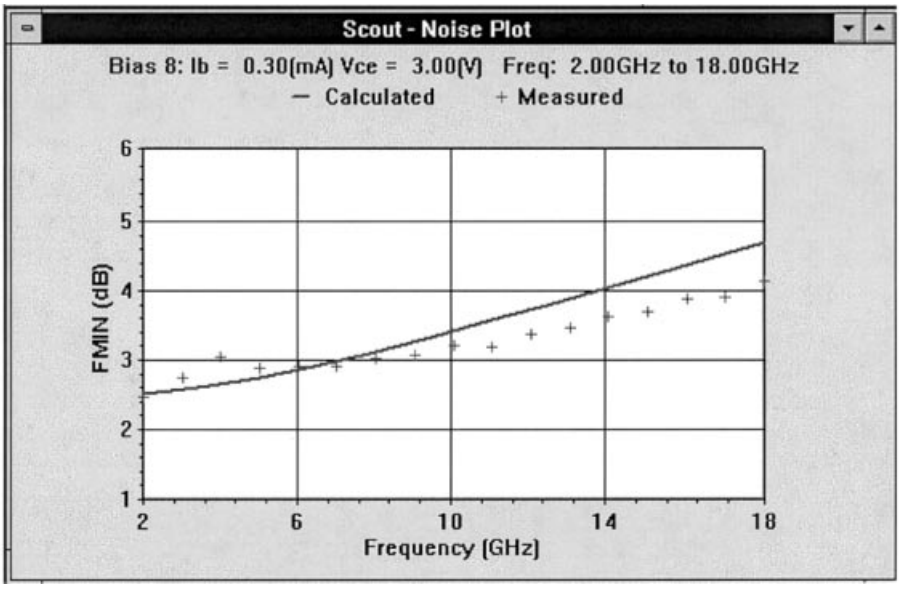


(b)

Figure A-15 Measured (discrete points) and modeled (solid lines) minimum noise figures of (a) Model 1 and (b) Model 2 at bias point $I_B = 0.3$ mA and $V_{CE} = 2$ V.



(a)



(b)

Figure A-16 Measured (discrete points) and modeled (solid lines) minimum noise figures of (a) Model 1 and (b) Model 2 at bias point $I_B = 0.3$ mA and $V_{CE} = 3$ V.

A-6 CONCLUSIONS

We have presented a novel nonlinear model for accurate dc, small-signal, and noise characterization of AlGaAs/GaAs HBTs. Two linearized model topologies have been described. Our model takes into account the bias, temperature, and frequency dependencies for small-signal and noise calculations and provides accurate device performance prediction. The model has been validated using devices from different foundries. This model can be implemented into SPICE or harmonic-balance simulators, thereby facilitating the design of microwave integrated circuits that incorporate HBTs. Our innovative noise calculation method can similarly be applied to the noise calculations of the FET device family.

We have exploited an integrated parameter extraction method fitting dc, multibias *S*-parameter, and multibias noise measurements simultaneously. The models extracted using this approach are capable of predicting the device small-signal responses and dc characteristics as well as noise performances accurately and can be used in a wide range of applications. Microwave circuit design engineers will benefit from such models, which allow the attainment of first-pass circuit design and thus reduce development cost. This creative technique is general and can be used for parameter extraction of other devices.

ACKNOWLEDGMENT

The work on which this appendix is based was sponsored, in part, by the Air Force Phase II SBIR Program Contract Number F33615-95-C-1707 on bias-dependent noise modeling of HBTs. The authors gratefully acknowledge discussions with Dr. Bob Pucel and thank Bob Neidhard of WL/ELM, Wright Patterson Air Force Base, for providing measurements. Participation of the two foundries toward model validation is greatly appreciated.

REFERENCES

1. R. Anholt, *Electrical and Thermal Characterization of MESFETs, HEMTs, and HBTs* . Norwood, MA: Artech House, 1995.
2. G. Vendelin, A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques*. New York: Wiley, 1990.
3. E. L. Kollberg, ed., *Microwave and Millimeter-Wave Mixers* . New York: IEEE Press, 1984.
4. U. L. Rohde, J. Whitaker, and T. N. N. Bucher, *Communications Receivers*, 2nd ed., New York: McGraw-Hill, 1997.
5. U. L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design* . New York: Wiley, 1997.
6. R. A. Pucel, ed., *Monolithic Microwave Integrated Circuits* . New York: IEEE Press, 1985.
7. K. Tomizawa, Y. Awano, and N. Hashizume, "Monte Carlo Simulation of AlGaAs/GaAs Heterojunction Bipolar Transistors," *IEEE Electron Device Lett.,* **EDL-5**, 362, 1984.
8. J. Hu, K. Tomizawa, and D. Pavlidis, "Monte Carlo Approach to Transient Analysis of HBTs with Different Collector Designs," *IEEE Electron Device Lett.,* **EDL-10**, 55–57, 1989.
9. A. A. Grinberg, M. Shur, R. Fisher, and H. Morkoc, "Investigation of the Effects of Graded Layer and Tunneling on the Performance of AlGaAs/GaAs Heterojunction Bipolar Transistors," *IEEE Trans. Electron Devices*, **ED-31**, 1758–1765, 1984.

10. R. Anholt, J. Gerber, R. Tayrani, and J. Pence, "HBT Model Parameter Extractor for SPICE and Harmonic Balance Simulators," *IEEE MTT-S International Microwave Symposium Digest* , pp. 1257–1260, 1994.
11. J. W. Bandler, R. M. Biernacki, Q. Cai, S. H. Chen, S. Ye, and Q. J. Zhang, "Integrated Physics-Oriented Statistical Modeling, Simulation and Optimization," *IEEE Trans. Microwave Theory Tech.,* **MTT-40**, 1374–1400, 1992.
12. S. G. Skaggs, J. Gerber, G. Bilbro, and M. B. Steer, "Parameter Extraction of Microwave Transistors Using a Hybrid Gradient Descent and Tree Annealing Approach," *IEEE Trans. Microwave Theory Tech.,* **MTT-41**, 726–729, 1993.
13. T. Fernández, Y. Newport, J. M. Zamanillo, A. Tazón, and A. Mediavilla, "Extracting a Bias-Dependent Large-Signal MESFET Model from Pulsed I/V Measurements," *IEEE Trans. Microwave Theory Tech.,* **MTT-44**, 372–378, 1996.
14. M. E. Hafizi, C. R. Crowell, and M. E. Grupen, "The dc Characteristics of AlGaAs/GaAs Heterojunction Bipolar Transistors with Application to Device Modeling," *IEEE Trans. Electron Devices*, **ED-37**, 2121–2129, 1990.
15. R. Anholt, J. Gerber, R. Tayrani, and J. Pence, "A Self-heating HBT Model for Harmonic Balance Simulators with Parameter Extraction," *1994 Asia-Pacific Microwave Conference* , pp. 1029–1032, 1994.
16. U. L. Rohde and C. R. Chang, "Parameter Extraction for Large-Signal Noise Models and Simulation of Noise in Large-Signal Circuits," *Microwave J.*, **36**(5), 222–239, May 1993.
17. R. A. Pucel and U. L. Rohde, "An Accurate Expression for the Noise Resistance R_n of a Bipolar Transistor for Use with the Hawkins Noise Model," *IEEE Microwave Guided Wave Lett.* , **3**, 35–37, 1993.
18. P. Antognetti and G. Massobrio, eds., *Semiconductor Modeling with SPICE*. New York: McGraw-Hill, 1988.
19. *Communications Components Designer's Catalog, GaAs and Silicon Products* , Hewlett-Packard data book, 1993.
20. *Microwave Harmonica Reference Manual* , Ansoft Corp., Compact Software Division, Elmwood Park, NJ.
21. *Compact Scout Reference Manual* , Ansoft Corp., Compact Software Division, Elmwood Park, NJ.

FURTHER READING

Alphonse Harter, "SiGe Technology Enhances Front-End Radio Performance," *Wireless Systems Design,* **4**(10), 11–14, October 1999.

RF/Microwave Circuit Design for Wireless Applications. Ulrich L. Rohde, David P. Newkirk
Copyright © 2000 John Wiley & Sons, Inc.
ISBNs: 0-471-29818-2 (Hardback); 0-471-22413-8 (Electronic)

APPENDIX B

NONLINEAR MICROWAVE CIRCUIT DESIGN USING MULTIHARMONIC LOAD-PULL SIMULATION TECHNIQUE*

This appendix presents practical applications of the multiharmonic load-pull simulation technique to the design of nonlinear microwave circuits. A systematic scheme using multiharmonic load-pull simulation explores the effects of each harmonic termination on the circuit performance and finds the optimal load at each harmonic. The circuit performance can significantly be improved following this systematic design procedure. This approach provides an efficient means of nonlinear microwave circuit design. Its advantages are illustrated for the design of two power amplifiers and a frequency doubler.

B-1 INTRODUCTION

With the rapid development of wireless communication in recent years, the quality of the nonlinear circuits, such as power amplifiers, oscillators, and frequency doublers, is becoming an increasingly critical component of the entire communication system. A good design of such nonlinear circuits can significantly improve the performance of the system and reduce its cost.

The harmonic-balance (HB) technique has widely been used for nonlinear microwave circuit design since its implementation in commercial CAD software. Many new applications have been developed for the HB method such as stability analysis [1, 2] and load-pull simulation [3]. Load-pull techniques and harmonic load tuning have been used successfully for large-signal device characterization and nonlinear circuit design (e.g., [4–7]). However,

*This appendix is based on Qian Cai, Jason Gerber, Chao-Ren Chang, and Ulrich L. Rohde, "Nonlinear Microwave Circuit Design Using Multi-harmonic Load-Pull Simulation Technique," *Int. J. Microwave Millimeter-Wave Computer-Aided Eng.*, 1999.

the design procedure has never been described systematically. This prevents or restricts the practical usage of the load-pull techniques.

Here we present a systematic scheme to use the load-pull technique efficiently in nonlinear microwave circuit design. Multiharmonic load-pull simulation using the HB method is used as a vehicle for our presentation. Response contours are simulated by sampling the corresponding harmonic impedance of the selected tuner connected to the input or output port. Optimal harmonic impedances are located from the response contours. A step by step design procedure is described.

Our process can be classified into two major steps: finding the optimal loading at each harmonic and checking the power levels of higher harmonics. By checking the power levels of higher harmonics, we can readily see the effects of higher harmonic loading on the circuit performance. Further steps are carried out based on the investigation of the higher harmonic loading effects. It is a very effective way for microwave engineers to achieve good, if not the best, results in nonlinear circuit design.

The multiharmonic load-pull simulation is presented in Section B-2. Applications of multiharmonic load-pull simulation to nonlinear circuit design are addressed in Section B-3. The circuit designs of two power amplifiers and a frequency doubler are demonstrated.

B-2 MULTIHARMONIC LOAD-PULL SIMULATION USING HARMONIC BALANCE

B-2-1 Formulation of Multiharmonic Load-Pull Simulation

The multiharmonic load-pull simulation is implemented within the nonlinear simulator, Microwave Harmonica [8], which uses an efficient HB technique [9]. The circuit topology for the multiharmonic load-pull simulation can be sketched generally as Figure B-1, where M tuners are placed at the M external ports considered. The formulation can generally be expressed as

$$E(X, Z) = 0 \tag{B-1}$$

where E is the vector of HB errors, X is the set of all harmonic state variables, and Z is the set of harmonic loads on all external ports. The kth subvector of E can be written

$$E_k(X, Z) = A(k\omega_0, Z)\Phi(X, Z) + B(k\omega_0, Z)\Psi(X, Z) + D(k\omega_0, Z) \tag{B-2}$$

where $0 \le k \le N_H$ (N_H is the number of harmonics used in the simulation), A and B are circuit matrices, D is a set of driving functions, and Φ and Ψ are, respectively, the harmonic vectors of instantaneous voltages v and currents i at the nonlinear subnetwork ports.

The harmonic loads Z can be written in the following matrix form:

$$Z = \begin{bmatrix} Z_1(0\omega_0) & Z_1(1\omega_0) & \cdots & Z_1(N_H\omega_0) \\ Z_2(0\omega_0) & Z_2(1\omega_0) & \cdots & Z_2(N_H\omega_0) \\ \vdots & \vdots & \vdots & \vdots \\ Z_M(0\omega_0) & Z_M(1\omega_0) & \cdots & Z_M(N_H\omega_0) \end{bmatrix} \tag{B-3}$$

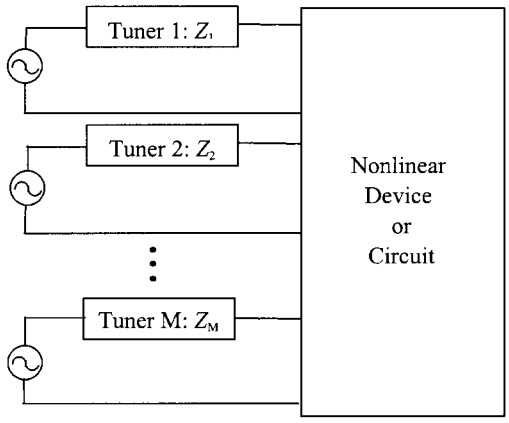


Figure B-1 Circuit topology for multiharmonic load-pull simulation.

where M is the number of external ports considered and $Z_i(k\omega_0)$ is the load at the kth harmonic and the ith port with $(0\omega_0)$ representing the dc component. The purpose of multiharmonic load-pull simulation is to find the optimal harmonic loading with respect to the design specifications. An impedance sampling method is used in our implementation of multiharmonic load-pull simulation. Only one component of Z is sampled at the defined impedance plane while the others are kept constant at each step. The HB simulation is performed at each sampling point to solve for the circuit responses specified. After the simulations are finished at all sampling points, load-pull contours are plotted on the Smith Chart and then the optimal point is located.

B-2-2 Systematic Design Procedure

Without losing generality, and for easy illustration, we consider the two-port circuit shown in Figure B-2, where tuners are placed at the source and load ports. The impedance sampling is achieved by adjusting the tuner parameters, which can be described as follows.

$R_i(k\omega_0)$—The resistance at the ith tuner and the kth harmonic, $i = 1, 2$, and $k = 1, 2, \ldots, N_H$.
$X_i(k\omega_0)$—The reactance at the ith tuner and the kth harmonic, $i = 1, 2$, and $k = 1, 2, \ldots, N_H$.

In order to obtain a uniform sampling of points, the tuner impedance is mapped to a reflection coefficient using a reference impedance Z_r and the mapping equation

$$\Gamma_i(k\omega_0) = \frac{Z_i(k\omega_0) - Z_r}{Z_i(k\omega_0) + Z_r} \tag{B-4}$$

The impedance sampling is transferred to two sweeps: the magnitude of $\Gamma_i(k\omega_0)$ $[|\Gamma_i(k\omega_0)|]$ from 0 to 1 and the angle of $\Gamma_i(k\omega_0)$ $[\angle\Gamma_i(k\omega_0)]$ from 0° to 360°. Only the impedance at one selected tuner i and one harmonic k, that is, $\Gamma_i(k\omega_0)$, is allowed to be tuned at a time and the other harmonic impedances are fixed at any meaningful values. In the following presentation we select $Z_r = 50\ \Omega$.

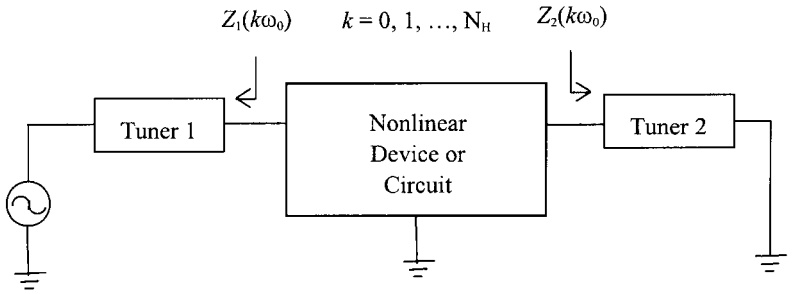


Figure B-2 Circuit schematic of a typical two-port nonlinear circuit.

The systematic design procedure can be described as follows:

Step 1: Start load-pull simulation by sampling the impedance at the fundamental frequency $Z_i(\omega_0)$ and find the optimal load $Z_{io}(\omega_0)$.

Step 2: Fix $Z_i(\omega_0)$ at $Z_{io}(\omega_0)$ and check the output spectrum to see the effects of higher harmonic loads on the circuit responses. If the effects of higher harmonic loads are significant, let $k = 2$, and go on to *Step 3*. Otherwise stop.

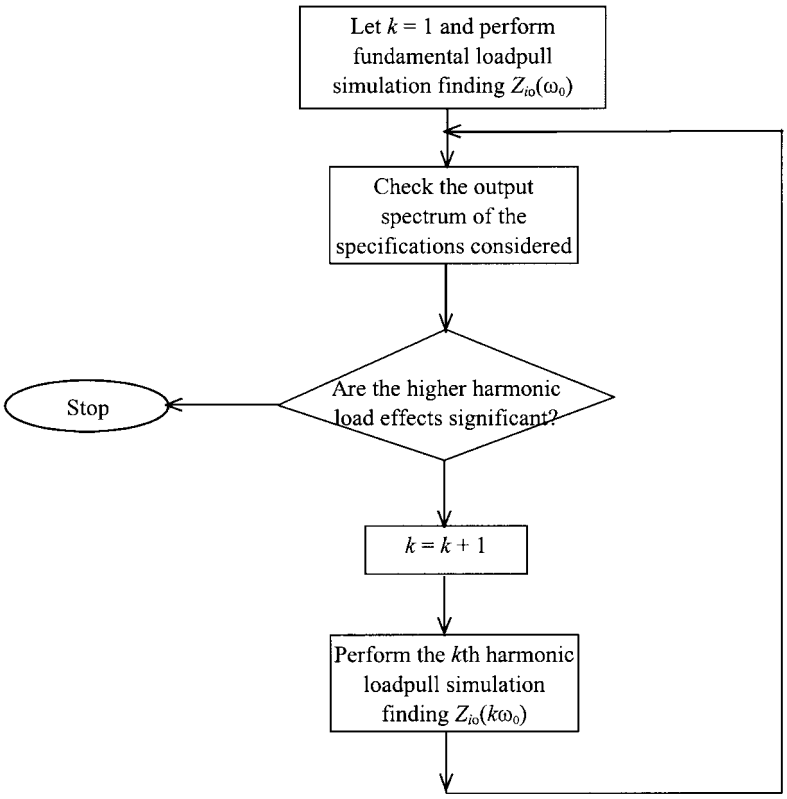


Figure B-3 Flowchart of design procedure using multiharmonic load-pull simulation.

Step 3: Perform load-pull simulation by sampling the kth harmonic impedance $Z_i(k\omega_0)$ and find the optimal load $Z_{io}(k\omega_0)$.

Step 4: Fix $Z_i(k\omega_0)$ at $Z_{io}(k\omega_0)$ and check the output spectrum to see the effects of higher harmonic loads on the circuit responses. If the effects of higher harmonic loads are significant, let $k = k + 1$, and go to *Step 3*. Otherwise stop.

This procedure can be repeated for all the tuners to achieve the optimal solution. It can be summed up as a load-pull simulation and spectrum-checking process. The flowchart in Figure B-3 illustrates this procedure.

B-3 APPLICATION OF MULTIHARMONIC LOAD-PULL SIMULATION

B-3-1 Narrowband Power Amplifier Design

Two power amplifiers, Amplifier 1 and Amplifier 2, are considered. The amplifiers are designed to operate at 0.5 GHz. The circuit schematic used for amplifier design is shown in Figure B-4. A Siemens power MESFET CLY10 [10] is used in our design. The MESFET is modeled by the modified Materka model [11] implemented in Microwave Harmonica [8]. Without losing generality, and for easy illustration, we fix all harmonic impedances of Tuner 1 (at the source port) at 50 Ω and tune Tuner 2 (at the load port) for both examples. Six harmonics are used in the HB simulation.

In Amplifier 1, the MESFET is biased at $V_{gs} = -1.3$ V and $V_{ds} = 5$ V, and the amplifier is designed as a normal Class A type with an input power of 10 dBm. Following the procedure described above, we perform load-pull simulation by sampling $Z_2(\omega_0)$ of Tuner 2 while other harmonic impedances of Tuner 2 are fixed at 50 Ω. The load-pull contours of power gain and power-added efficiency (PAE) are plotted in Figure B-5. The optimal load $Z_{2o}(\omega_0)$ is found to be $9.07 + j12.99$, at which the power gain is 18.54 dB and PAE is about 20%. The output power spectrum at this point is plotted in Figure B-6. By checking the output power spectrum of Figure B-6, we can see that the power levels at all higher harmonics are very small. Therefore, the higher harmonic loads will not have significant effects on the amplifier performance. To verify this conclusion, we perform load-pull simulation by sampling $Z_2(2\omega_0)$ of Tuner 2 while $Z_2(\omega_0)$ is fixed at $Z_{2o}(\omega_0)$. It is found that the best performance with respect to power gain and PAE is obtained for purely reactive second-harmonic loads, which is consistent with the results obtained by Berini et al. [5]. The output power versus the phase of $\Gamma_2(2\omega_0)$ is shown in Figure B-7, from which we can see that the influence of $Z_2(2\omega_0)$ is very small and our conclusion is justified.

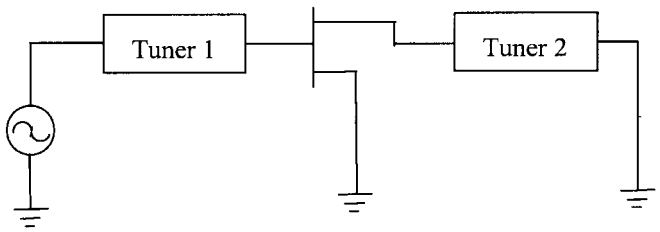
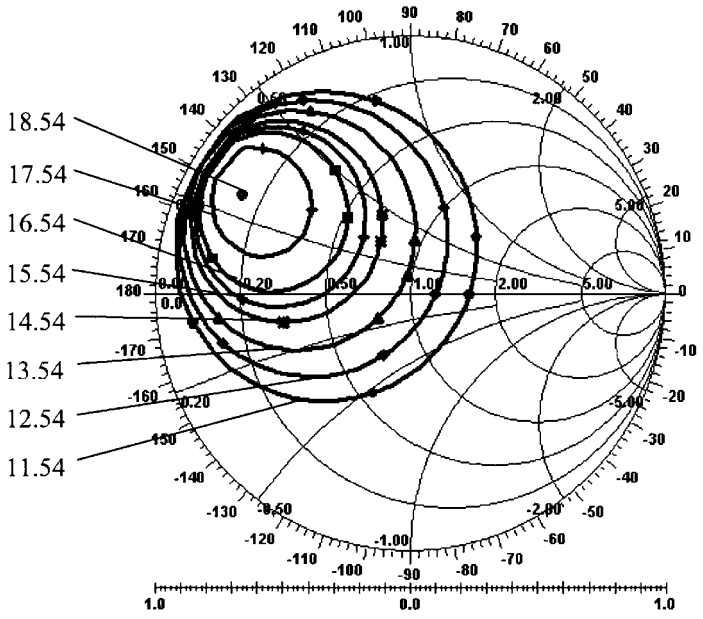
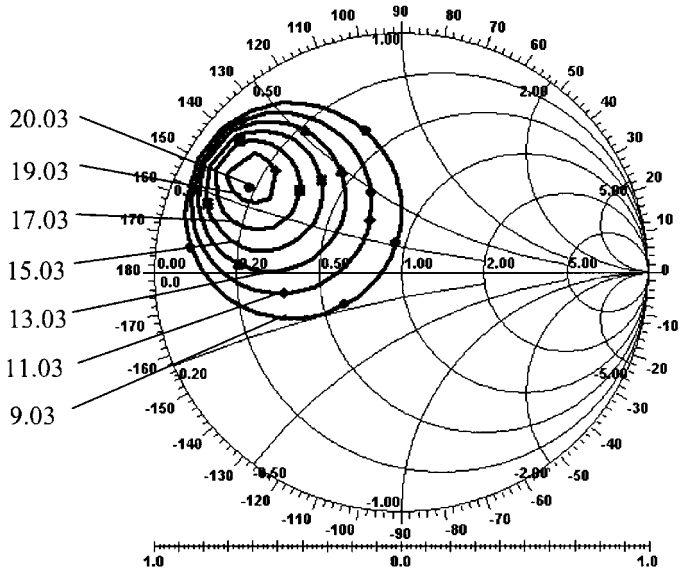


Figure B-4 Circuit topology for the amplifier design using multiharmonic load-pull simulation.



(a)



(b)

Figure B-5 Load-pull contours of (a) power gain (dB) and (b) PAE (%) of Amplifier 1 obtained from fundamental load-pull simulation.

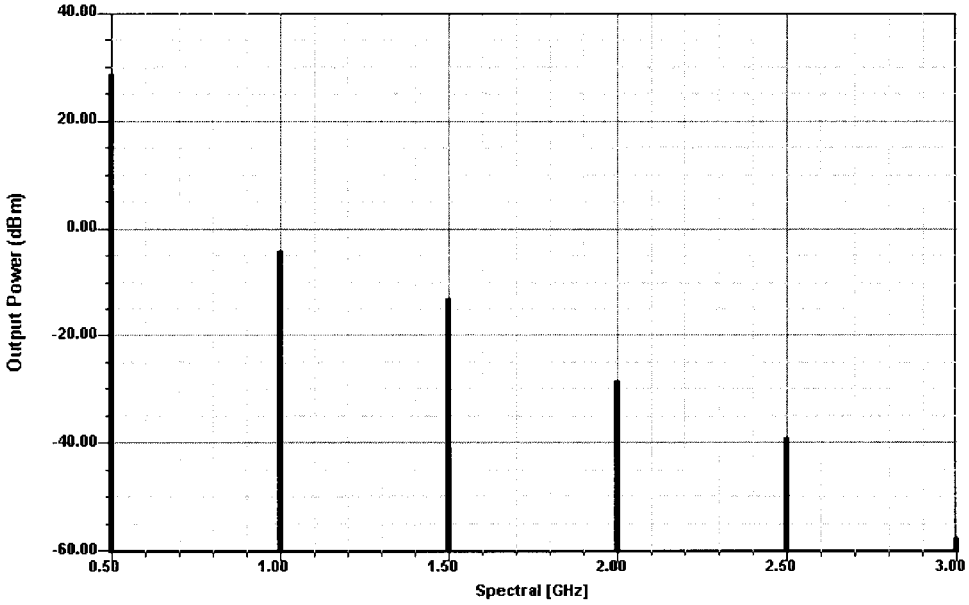


Figure B-6 Output power spectrum of Amplifier 1 when $Z_2(\omega_0) = Z_{2o}(\omega_0)$.

In order to illustrate the effects of higher harmonic loading on the amplifier performance, the MESFET is biased at $V_{gs} = -2.0$ V and $V_{ds} = 5$ V, and the input power is increased to 20 dBm in Amplifier 2. Figure B-8 shows the contours of power gain and PAE at the fundamental load-pull simulation. The optimal load $Z_{2o}(\omega_0)$ is $4.31 + j13.30$ for power gain (12.66 dB)

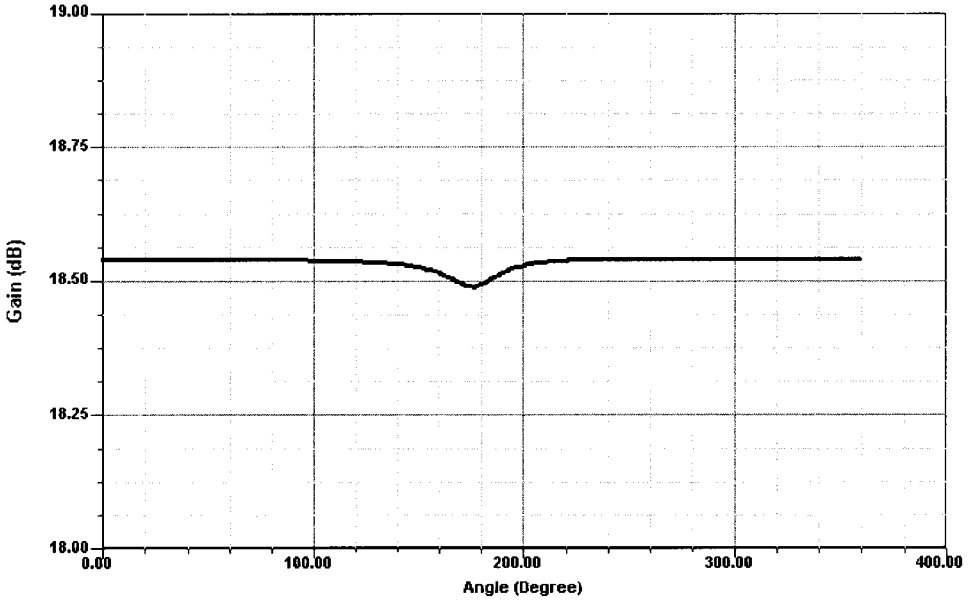
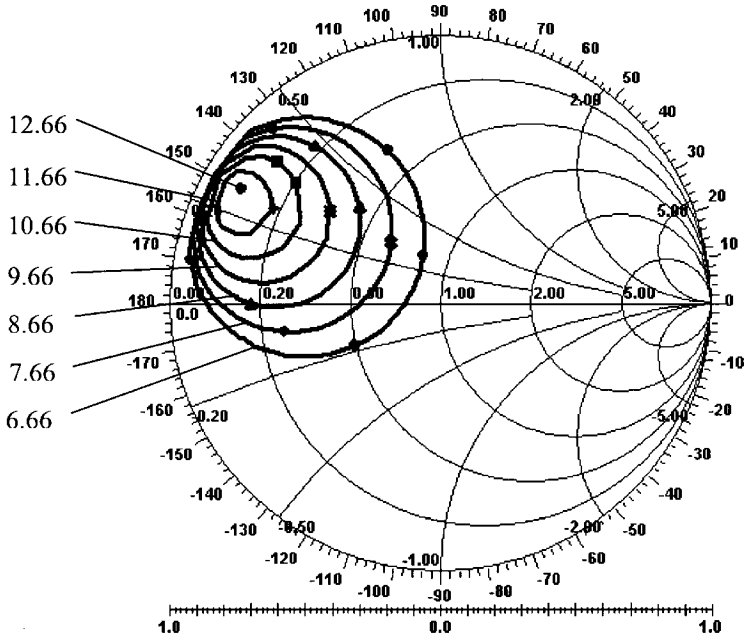
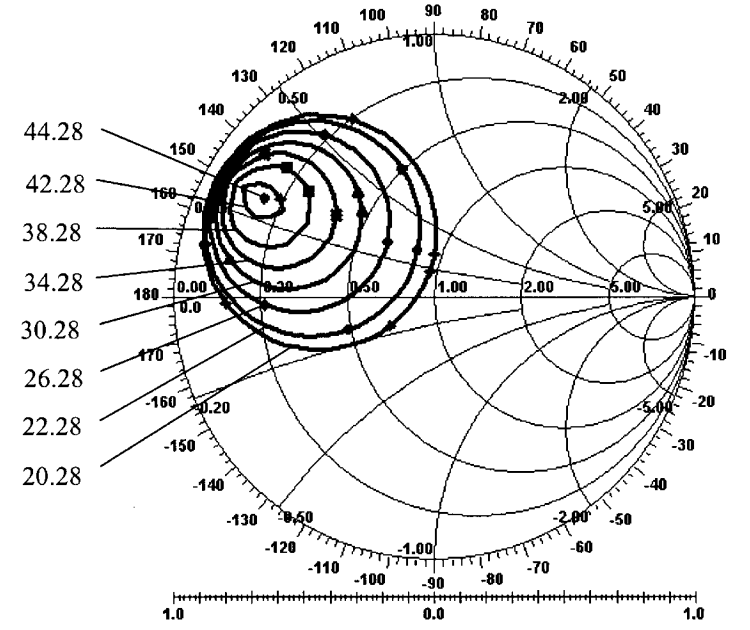


Figure B-7 Power gain versus the phase of $\Gamma_2(2\omega_0)$ for Amplifier 1.



(a)



(b)

Figure B-8 Load-pull contours of (a) power gain (dB) and (b) PAE (%) of Amplifier 2 obtained from fundamental load-pull simulation.

and 7.61 + j13.11 for PAE (44.28%). Though a compromise between the power gain and PAE can be obtained, we select 7.61 + j13.11 as our $Z_{2o}(\omega_0)$ and consider the PAE as the primary specification throughout the following process. The output spectrum at this point is shown in Figure B-9, which indicates that the output power levels at higher harmonics are significant, and thus the higher harmonic loads are critical to the amplifier performance.

By fixing $Z_2(\omega_0)$ at 7.61 + j13.11 and sampling $Z_2(2\omega_0)$, we perform the second-harmonic load-pull simulation. The contours of PAE are shown in Figure B-10, which also indicates that the best point for the second-harmonic load will be pure reactance near the short-circuit point. The effect of the phase of $\Gamma_2(2\omega_0)$ on PAE is shown in Figure B-11. The optimal value of $Z_{2o}(2\omega_0)$ is –j6.58, where the phase of $\Gamma_2(2\omega_0)$ is 195°. The PAE is improved from 44.28% to 47.18%. The output power spectrum at $Z_2(\omega_0)$ = 7.61 + j13.11 and at $Z_2(2\omega_0)$ = –j6.58 is shown in Figure B-12. From Figure B-12, we can see that the amplifier performance can be further improved by a proper third-harmonic load. The third-harmonic load-pull simulation is carried out by sampling $Z_2(3\omega_0)$ while $Z_2(\omega_0)$ and $Z_2(2\omega_0)$ are maintained at their optimal values. $Z_{2o}(3\omega_0)$ is found to be –j65.16 at which the value of PAE is 47.66%. Following the same procedure, we perform load-pull simulations up to the sixth harmonic. The results are as follows:

Harmonic Load	Optimal Value
$Z_2(\omega_0)$	7.61 + j13.11
$Z_2(2\omega_0)$	–j6.58
$Z_2(3\omega_0)$	–j65.16
$Z_2(4\omega_0)$	–j2.18
$Z_2(5\omega_0)$	–j28.81
$Z_2(6\omega_0)$	–j23.32

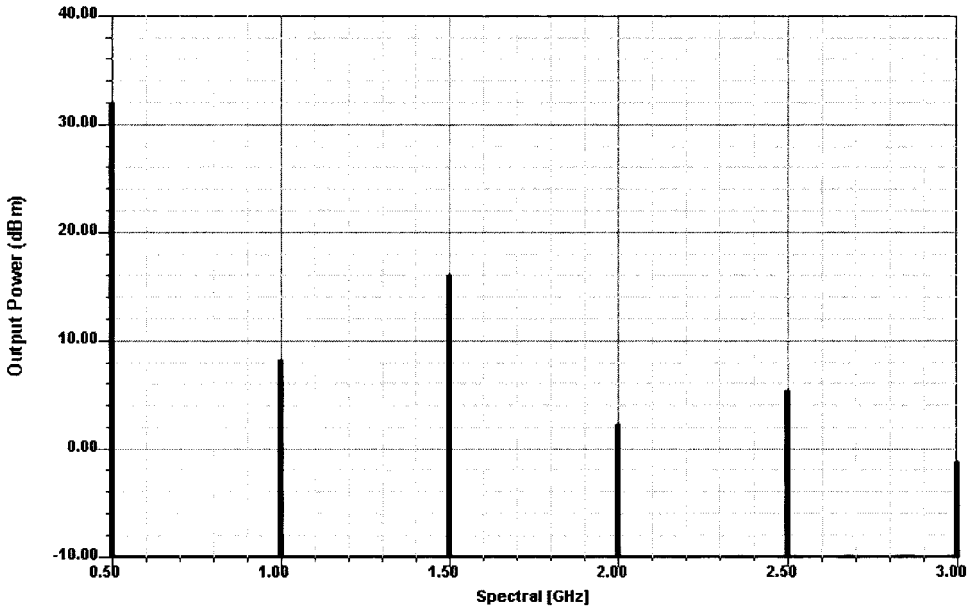


Figure B-9 Output power spectrum of Amplifier 2 when $Z_2(\omega_0) = Z_{2o}(\omega_0)$.

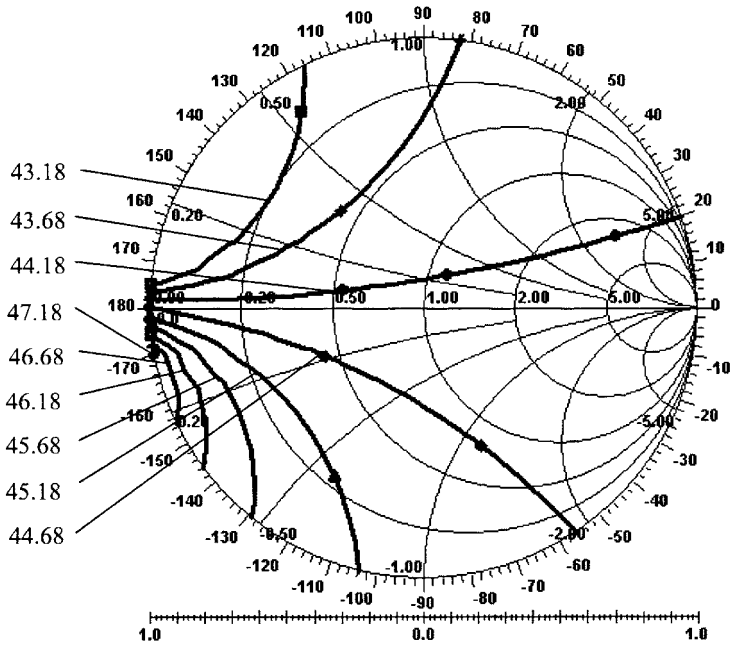


Figure B-10 Load-pull contours of PAE (%) of Amplifier 2 obtained from the second-harmonic load-pull simulation.

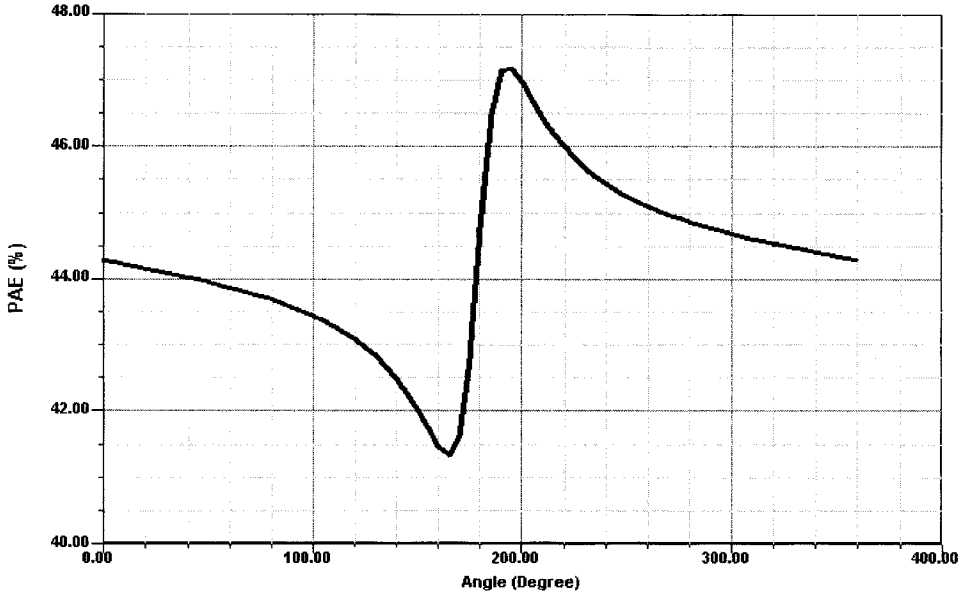


Figure B-11 PAE versus the phase of $\Gamma_2(2\omega_0)$ for Amplifier 2.

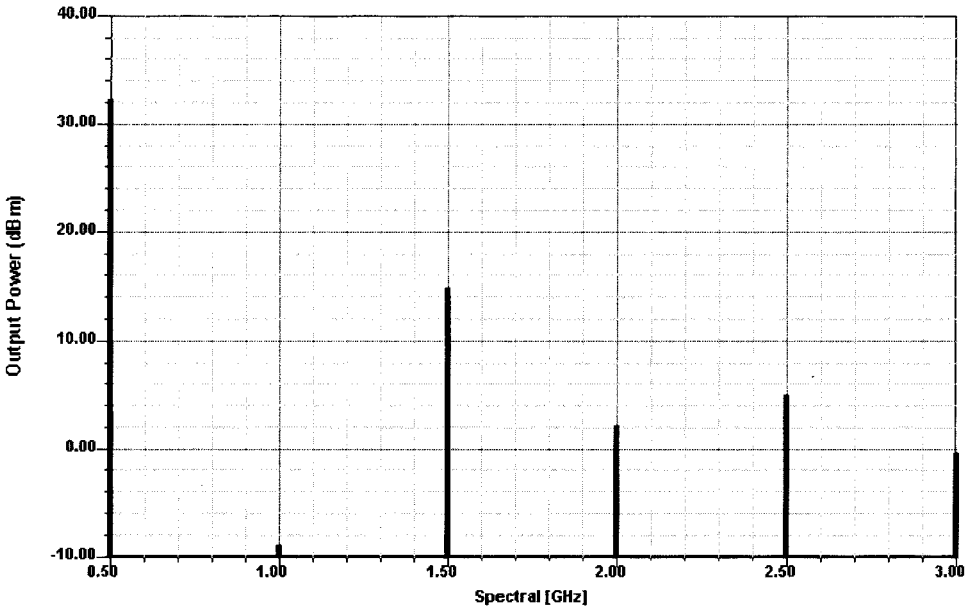


Figure B-12 Output power spectrum of Amplifier 2 when $Z_2(\omega_0) = Z_{2o}(\omega_0)$ and $Z_2(2\omega_0) = Z_{2o}(2\omega_0)$.

The PAE is 48.13% and power gain is 12.38 dB at the final design. The output power spectrum of the final design is shown in Figure B-13, which indicates that the output power levels at the second and higher harmonics have been suppressed significantly compared to the ones shown in Figure B-9.

B-3-2 Frequency Doubler Design

The circuit schematic shown in Figure B-14 is used for the frequency doubler design. The MESFET is biased at pinchoff, with $V_{gs} = -1.9$ V and $V_{ds} = 5$ V. The frequency of the input waveform is 5 GHz and the input power is 3 dBm. Six harmonics are considered in the HB simulation. Both Z_S and Z_L are 50 Ω. The conversion gain (CG) and the spectral purity (SP), are considered as the design specifications, calculated, respectively, by

$$CG = \frac{P_L(2\omega_0)}{P_{avs}(\omega_0)} \tag{B-5}$$

and

$$SP = \frac{P_L(2\omega_0)}{\sum_{k=1,\, k\neq 2}^{N_H} P_L(k\omega_0)} \tag{B-6}$$

where P_L is the power delivered to the load Z_L and P_{avs} is the available source power. In addition to the load-pull simulation, we also perform source-pull simulation in the design.

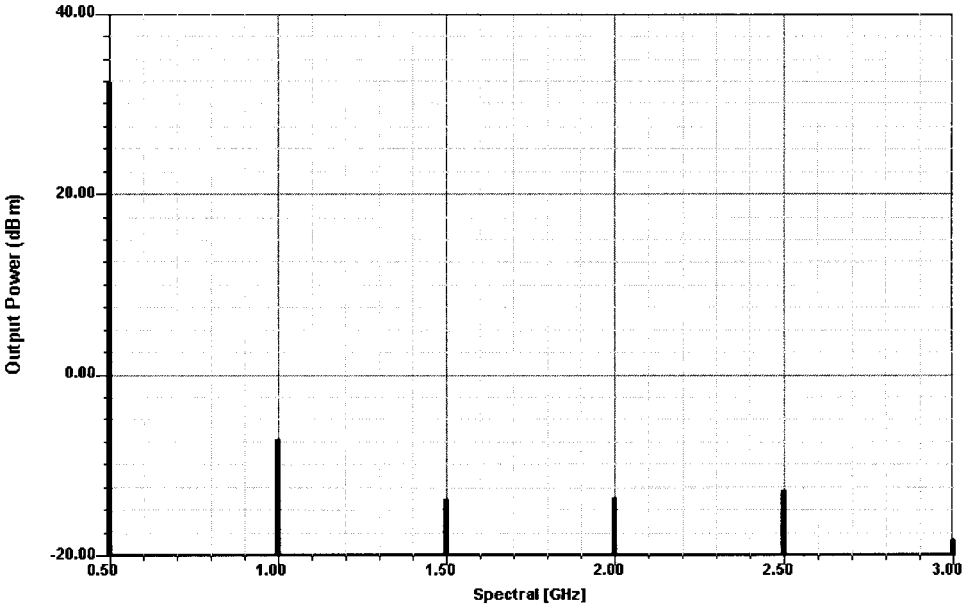


Figure B-13 Output power spectrum of Amplifier 2 at the final design.

Note that the harmonic impedance $Z_i(k\omega_0)$ is the tuner impedance without the load Z_L or the source impedance Z_S, and $\Gamma_i(k\omega_0)$ is the reflection coefficient of the tuner only in this design example.

First, the fundamental load-pull simulation is performed by sampling $\Gamma_2(\omega_0)$ at Tuner 2 while the impedances of all harmonics at Tuner 1 and other harmonic impedances of Tuner 2 are fixed at 0 Ω. The contours of SP with respect to $\Gamma_2(\omega_0)$ are plotted in Figure B-15, from which we can see that the optimal tuner impedance is infinity, which means that the fundamental harmonic component of the output waveform should be completely reflected by an open circuit. We select $1 \angle 0$ as the optimal point $[\Gamma_{2o}(\omega_0)]$, at which the SP is 13.21 dB and the CG is –9.75 dB. The second-harmonic load-pull simulation is performed by

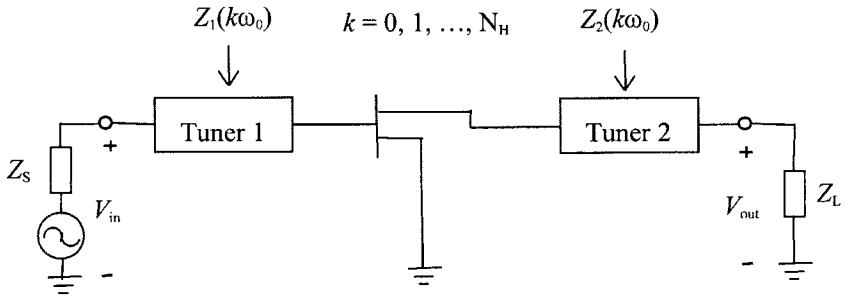


Figure B-14 Circuit topology for frequency doubler design using multiharmonic load-pull simulation.

sampling $\Gamma_2(2\omega_0)$ with $\Gamma_2(\omega_0)$ being fixed at $\Gamma_{2o}(\omega_0)$. The contours of SP with respect to $\Gamma_2(2\omega_0)$ are plotted in Figure B-16. The $\Gamma_{2o}(2\omega_0)$ is found to be $1 \angle 50$, at which the SP is improved from 13.21 to 18.14 dB and the CG from –9.75 to –5.89 dB. Following the same procedure, the load-pull simulation is performed up to the sixth harmonic. It is found that all the optimal harmonic impedances of Tuner 2 are infinity, except for the second harmonic, at which the optimal impedance is found to be $j127.23$. At the end of the load-pull simulation, the SP is 25.28 dB and the CG is –5.87 dB. The CG cannot be improved very much by tuning the harmonic impedances of Tuner 2.

The source-pull simulation is carried out by sampling the harmonic impedances of Tuner 1 using the same procedure as load-pull simulation with all the harmonic impedances of Tuner 2 being fixed at their optimal values. It is found that the most dominant harmonic impedance of Tuner 1 is the fundamental harmonic impedance. The optimal impedances at other harmonics of Tuner 1 turn out to be 0. The contours of CG with respect to $\Gamma_1(\omega_0)$ obtained by fundamental source-pull simulation with other harmonic impedances of Tuner 1 set to 0 Ω are plotted in Figure B-17. The $\Gamma_{1o}(\omega_0)$ is found to be $1 \angle 40$. The CG is significantly improved from –5.87 to 2.07 dB with a minor 0.5-dB degradation of spectral purity. The input and output voltage waveforms under the optimal termination conditions are shown in Figure B-18, from which we can see that the frequency of the output voltage is doubled to 10 GHz with a reasonable gain. The distortion of the input voltage waveform is due to the reflection of the higher-order harmonic voltages.

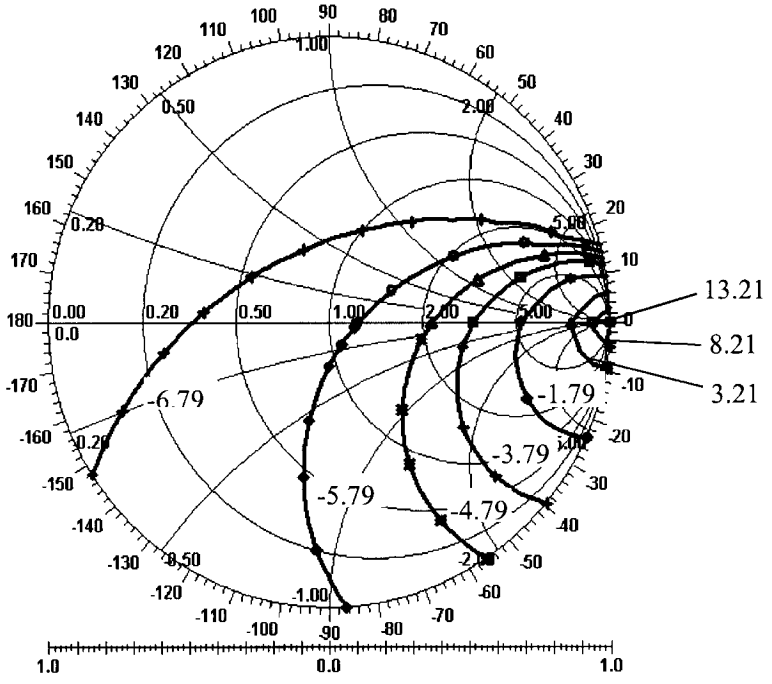


Figure B-15 Contours of spectral purity (dB) with respect to the fundamental impedance of Tuner 2 [$\Gamma_2(\omega_0)$] obtained from fundamental load-pull simulation of the frequency doubler.

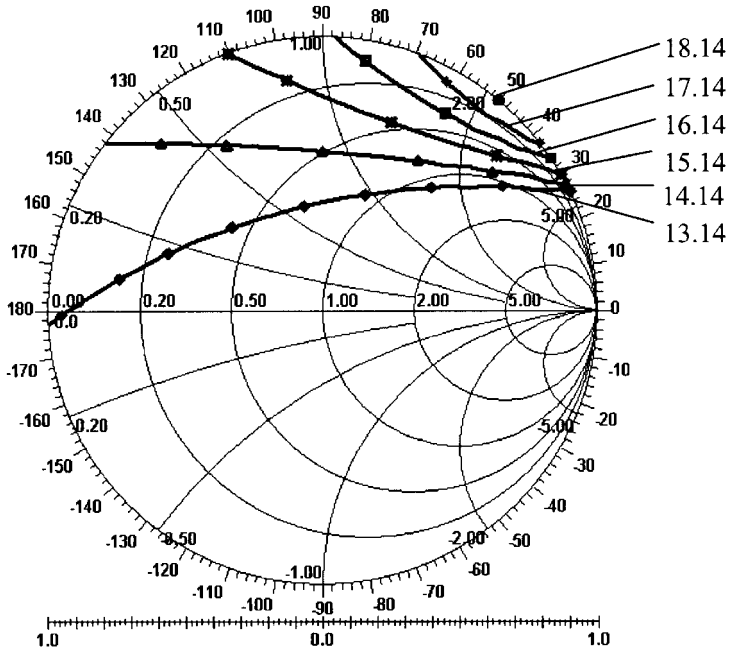


Figure B-16 Contours of spectral purity (dB) with respect to the second-harmonic impedance of Tuner 2 [$\Gamma_2(2\omega_0)$) obtained from the second-harmonic load-pull simulation of the frequency doubler.

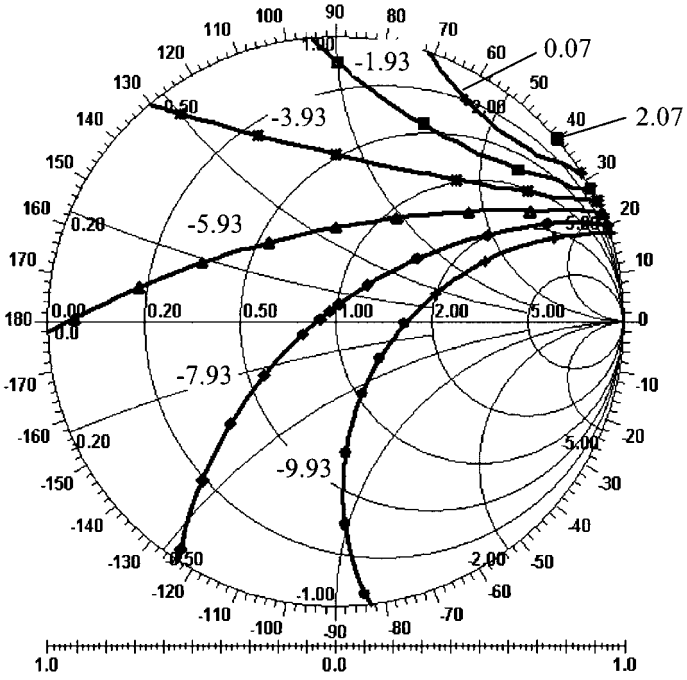


Figure B-17 Contours of conversion gain (dB) with respect to the fundamental impedance of Tuner 1 [$\Gamma_1(\omega 0)$] obtained from fundamental source-pull simulation of the frequency doubler.

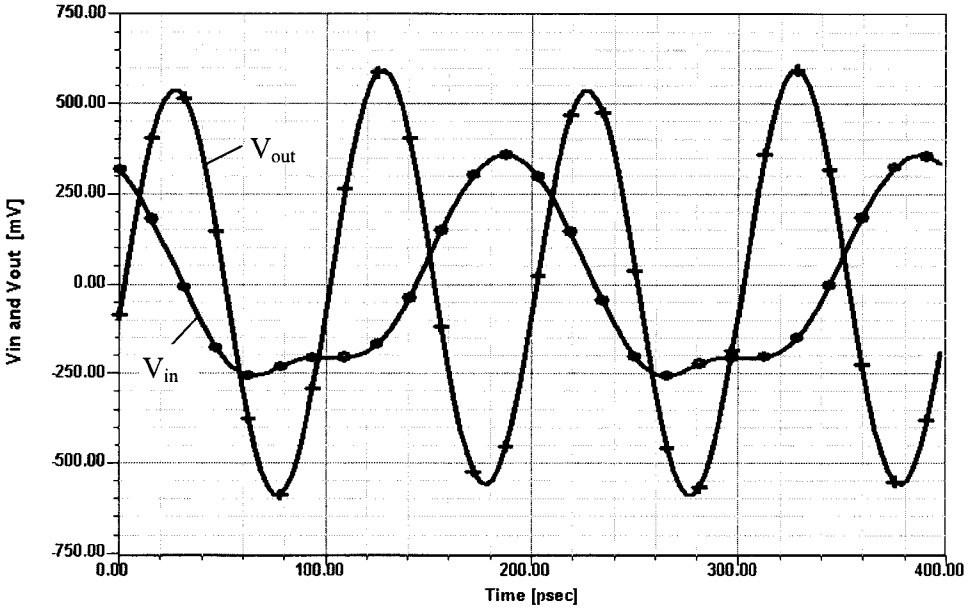


Figure B-18 Input and output voltage waveforms of the frequency doubler under the optimal termination conditions.

B-4 CONCLUSIONS

We have presented practical applications of multiharmonic load-pull simulation to the design of nonlinear microwave circuits. A systematic procedure has been addressed for effective circuit design. It has been demonstrated that the circuit performance can be improved significantly by a proper design of harmonic loads. This method can be applied to the design of nonlinear microwave circuits such as the power amplifier and frequency doubler to achieve the optimal solution and to utilize the maximum potential of the devices employed in the circuits.

B-5 NOTE ON THE PRACTICALITY OF LOAD-PULL-BASED DESIGN

From the standpoint that the wireless designer cannot have access to too many design aids, load-pull-based design techniques can serve as a valuable addition to the wireless designer's toolbox. We caution, however, that the current surge of interest in load-pull design techniques may owe as much to fashion and wishful thinking as it does to the practical achievability of the improved hardware performance that load-pull analysis may predict. Realizing the gain, PAE, and spectral-purity predictions of analyses like those presented here depends on the realization of optimum, controllable terminations at the signal fundamental *and its significant harmonics*—a nontrivial design challenge to say the least.

ACKNOWLEDGMENT

The work on which this appendix is based was sponsored under the DARPA MAFET Thrust 1 program, Agreement No. MDA972-96-3-0001. Partial contributions of Dr. Solti Peng of TI, Dallas, Texas, to some of the materials presented here are gratefully acknowledged.

REFERENCES

1. V. Rizzoli and A. Neri, "State of the Art and Present Trends in Nonlinear Microwave CAD Techniques," *IEEE Trans. Microwave Theory Tech*., **MTT-36**, 343–365, 1988.
2. J. Gerber and C. R. Chang, "Application of the Harmonic-Balance Method to the Stability Analysis of Oscillators," *IEEE MTT-S International Microwave Symposium Digest* , pp. 1587–1590, 1997.
3. Q. Cai, J. Gerber, and S. Peng, "A Systematic Scheme for Power Amplifier Design Using a Multi-harmonic Load-Pull Simulation Technique," *IEEE MTT-S International Microwave Symposium Digest*, 1998.
4. R. Hajji, F. Beauregard, and F. Ghannouchi, "Multitone Power and Intermodulation Load-Pull Characterization of Microwave Transistors Suitable for Linear SSPAs Design," *IEEE Trans. Microwave Theory Tech.*, **MTT-45**, 1093–1099, 1997.
5. P. Berini, M. Desgagn, F. Ghannouchi, and R. G. Bosisio, "An Experimental Study of the Effects of Harmonic Loading on Microwave MESFET Oscillators and Amplifiers," *IEEE Trans. Microwave Theory Tech.,* **MTT-42**, 943–950, 1994.
6. M. A. Khatibzadeh and H.Q. Tserng, "Harmonic Tuning of Power FETs at X-Band," *IEEE MTT-S International Microwave Symposium Digest* , pp. 989–992, 1990.
7. B. Kopp and D. D. Heston, "High-Efficiency 5-watt Power Amplifier with Harmonic Tuning," *IEEE MTT-S International Microwave Symposium Digest*, pp. 839–842, 1988.
8. *Microwave Harmonica Reference Manual* , Ansoft Corporation, Compact Software Division, Elmwood Park, NJ.
9. V. Rizzoli, C. Cecchetti, A. Lipparini, and F. Mastri, "General-Purpose Harmonic Balance Analysis of Nonlinear Microwave Circuits Under Multitone Excitation," *IEEE Trans. Microwave Theory Tech.,* **MTT-36**, 1650–1660, 1988.
10. *Einzelhalbleiter Discrete and RF Semiconductors* , Siemens Data Book 02.97, 1997.
11. A. Materka and T. Kacprzak, "Computer Calculation of Large-Signal GaAs FET Amplifier Characteristics," *IEEE Trans. Microwave Theory Tech.* , **MTT-33**, 129–135, 1985.