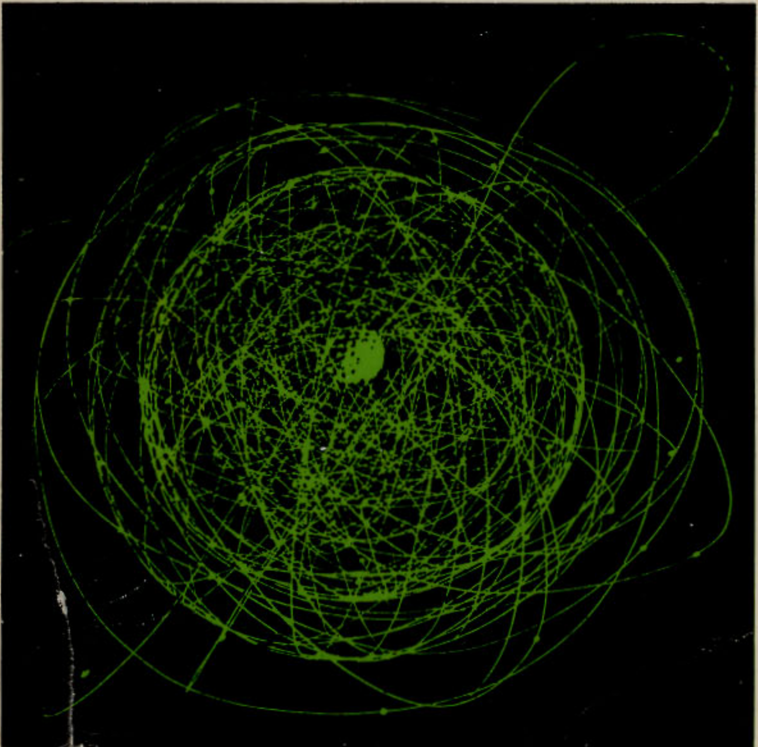
Advanced science
Readings

Modern chemistry

edited by J. G. Stark



Advanced science
Readings
Modern chemistry
edited by J. G. Stark

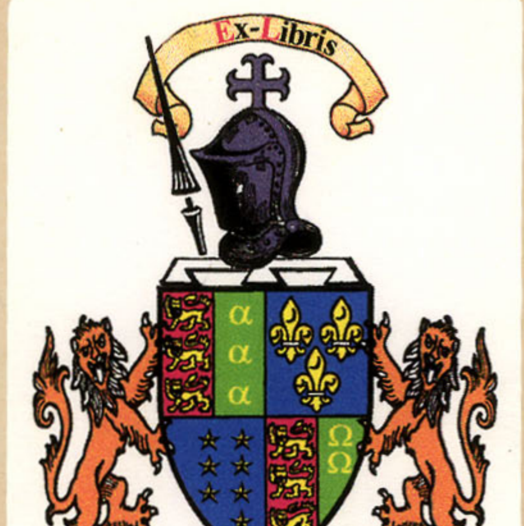
14 080161 8

Penguin Education

Advanced Science Readings

Modern Chemistry

Edited by J. G. Stark



Modern Chemistry

Selected Readings

Edited by J. G. Stark

Penguin Books

Penguin Books Ltd, Harmondsworth,
Middlesex, England
Penguin Books Inc., 7110 Ambassador Road,
Baltimore, Md 21207, U.S.A.
Penguin Books Australia Ltd, Ringwood,
Victoria, Australia

First published 1970

Made and printed in Great Britain by
Richard Clay (The Chaucer Press) Ltd, Bungay, Suffolk
Set in Monotype Times

Contents

Preface

The twentieth century has seen a series of rapid advances in the physical sciences, many of which have had a profound effect on chemical thought at all levels, and this book of readings in chemistry is intended to provide an up-to-date background for sixth-form and first-year university students. It consists of a series of articles which have been published during the last five years in the leading journals of chemical education in the United Kingdom and the United States. The articles are grouped under six main headings: atomic and molecular structure, stereochemistry, energetics and kinetics, acids and bases, the chemical elements and organic reaction mechanisms.

This book of readings is *not* a textbook and it does not aim to be comprehensive. Instead it offers a series of clearly written articles on some of the main themes of modern chemistry. A short list of suggestions for further reading is given at the end of each article and it is hoped that the reader who has access to a technical library will make the fullest use of the facilities it offers.

The editor wishes to thank the authors concerned, particularly Mr E. Cartmell and Professor J. A. Campbell, for permission to make minor modifications to their articles.

SI Units

International agreement has been reached between the standards authorities of most of the leading scientific countries to adopt a common system of metric units known as the SI system (Système Internationale d'Unités) and SI units will be increasingly used in universities, colleges and schools in the next few years. The most notable change which the adoption of the SI system will involve for the chemist is probably the replacement of the calorie as the unit of heat energy by the joule (J).

There are six basic SI units:

Basic physical quantity	*Name of unit*	*Symbol*
length	metre	m
mass	kilogramme	kg
time	second	s
electric current	ampere	A
thermodynamic temperature	kelvin	K
luminous intensity	candela	cd

The mole (mol) is a seventh unit, widely used by chemists, which may be introduced as a further SI unit. In addition there are *derived units* which are defined in terms of the basic units or other derived units, including the following:

Physical quantity	*Name of unit*	*Symbol and definition*
force	newton	N ($kg\,m\,s^{-2}$)
energy, heat	joule	J ($kg\,m^2\,s^{-2}$)
power	watt	W ($kg\,m^2\,s^{-3} = J\,s^{-1}$)
electric charge	coulomb	C (A s)
electric potential difference	volt	V ($kg\,m^2\,s^{-3}\,A^{-1} = J\,A^{-1}\,s^{-1}$)
electric resistance	ohm	Ω ($kg\,m^2\,s^{-3}\,A^{-2} = V\,A^{-1}$)
frequency	hertz	Hz (s^{-1})
customary temperature	degree Celsius	°C (K–273·15)

SI Units

Agreed prefixes and symbols are used to indicate multiples of units in powers of ten, including the following.

Multiple	*Prefix*	*Symbol*
10^{6}	mega	M
10^{3}	kilo	k
10^{-3}	milli	m
10^{-6}	micro	μ
10^{-9}	nano	n
10^{-12}	pico	p

The values of some physical constants in SI units and conversion factors are given in the following tables.

Quantity	*Value*
velocity of light c	$2{\cdot}998 \times 10^{8}\ \mathrm{m\ s^{-1}}$
Boltzmann constant k	$1{\cdot}380 \times 10^{-23}\ \mathrm{J\ K^{-1}}$
mass of proton	$1{\cdot}672 \times 10^{-27}\ \mathrm{kg}$
mass of neutron	$1{\cdot}675 \times 10^{-27}\ \mathrm{kg}$
mass of electron	$9{\cdot}109 \times 10^{-31}\ \mathrm{kg}$
charge of electron e	$1{\cdot}602 \times 10^{-19}\ \mathrm{C}$
Planck constant h	$6{\cdot}626 \times 10^{-34}\ \mathrm{J\ s}$
Avogadro constant N_A	$6{\cdot}022 \times 10^{23}\ \mathrm{mol^{-1}}$
molar gas volume at s.t.p.	$2{\cdot}241 \times 10^{-2}\ \mathrm{m^{3}\ mol^{-1}}$
gas constant R	$8{\cdot}314\ \mathrm{J\ K^{-1}\ mol^{-1}}$
Faraday constant F	$9{\cdot}649 \times 10^{4}\ \mathrm{C\ mol^{-1}}$
triple point of water	$273{\cdot}16\ \mathrm{K}$

1 calorie $= 4{\cdot}184$ J.
1 C $= 3 \times 10^{9}$ e.s.u.
1 Å $= 10^{-10}$ m $= 10^{-1}$ nm $= 10^{2}$ pm.
1 atmosphere $= 760$ mmHg $= 1{\cdot}013 \times 10^{5}\ \mathrm{N\ m^{-2}}$.

References

The International System (SI) Units (BS 3763), British Standards Institution, 1964.

The Use of SI Units (PD 5686), British Standards Institution, 1967 (revised edn, 1969).

M. L. McGLASHAN, *Physico-Chemical Quantities and Units*, Royal Institute of Chemistry, 1968.

Part One Atomic and Molecular Structure

The concept of structure is one of the main themes in modern chemistry and the attempt to correlate the structure of substances with their properties is often a fruitful exercise. E. Cartmell in his article 'Atomic structure' introduces modern ideas on the subject, dealing with such aspects of atomic structure as electronic configuration (by a consideration of ionization energies), quantum numbers, the periodic classification and atomic orbitals.

R. L. Livingston's article emphasizes the importance of crystal structure in modern chemistry courses and outlines an approach to the subject at an introductory level which will also be useful to the student as a preparation for more advanced work.

G. Allen summarizes the main experimental methods available for the determination of molecular structure, dealing in turn with spectroscopic techniques and diffraction methods. The type of information provided by the various techniques is assessed and the article contains a useful table in which the main methods for structure determination are put into the perspective of the preparative chemist.

1 E. Cartmell

Atomic Structure

E. Cartmell, 'Atomic structure', *Chemistry Today*, O.E.C.D., 1963, ch. 4, pp. 67–86.

The Traditional Approach

The traditional approach to atomic structure usually starts by pointing out that the theory that matter is made up from a very large number of very small, indivisible particles called atoms goes back to classical times (Leukippos, Demokritos), but stresses that this was a philosophical theory, no serious attempt being made to relate it to observations of natural phenomena.

The beginning of scientific inquiry into atomic structure can be traced in the works of the 'natural philosophers' of Renaissance and post-Renaissance times (e.g. Galileo, Newton, Boyle), but detailed discussion of atomic structure is based upon the work of John Dalton. It is shown first of all that the simple postulates of the Dalton theory explain the laws of stoichiometry – constancy of mass, constancy of composition, and combination in multiple and in reciprocal proportions. The next stage in the historical development of the subject is the determination of the relative weights of the atoms, which requires both a knowledge of 'combining weights' and also a knowledge of the 'combining powers' of the atoms themselves.

The description of the determination of atomic weights by chemical methods involves an account of the work of Gay-Lussac, Avogadro and Cannizzaro, and the story is one of the outstanding examples of the power and beauty of the scientific method. However, relative atomic weights are now always determined by the mass spectrograph, and the principles of this instrument can easily be understood once the electronic theory has been discussed. The modern theory of atomic structure also makes possible a description of 'combining power' in quite simple electronic terms.

Modern Atomic Theory

The atomic nature of electricity

Any discussion of atomic theory requires an introductory treatment of some of the properties of static and of current electricity. The important points are:

(i) that there are two kinds of electric charge which can for convenience be called *positive* and *negative*;
(ii) that like charges repel and unlike charges attract; and
(iii) that the force between two charges e_1 and e_2 separated by a distance r is, by the Coulomb law, $e_1 e_2/r^2$.

Conduction of electricity through gases

The well-known experiments on the properties of the cathode rays (Maltese cross, deflection experiments, measurement of charge, etc.) should, so far as is possible, be demonstrated experimentally. The principles of the Thomson experiment for the determination of e/m can then be understood. It is important to note that the same value for e/m ($1{\cdot}76 \times 10^8$ C g^{-1}) is obtained whatever the nature of the residual gas in the cathode ray tube.

Conduction of electricity through solutions

Here Faraday's laws must be quoted,[1] and comparatively simple experiments can be performed to show, for example, that one mole of silver is liberated from acidified silver nitrate solution by passing $9{\cdot}65 \times 10^4$ coulombs of electricity, whereas the liberation of one mole of copper from acidified copper sulphate solution requires $2 \times 9{\cdot}65 \times 10^4$ coulombs.

These and other similar facts can be explained if:

(i) chemical compounds produce both positively and negatively charged particles (ions) in conducting solutions;

---

[1] The quantity of a substance produced during electrolysis is proportional to the time and to the current; and the quantities of different substances produced by the passage of a certain quantity of electricity (i.e. number of electrons) are in the same proportion as the formula weight divided by the charge of the corresponding ion.

(ii) the charge on such ions is always a small whole number times the charge on the hydrogen ion; and

(iii) electricity is 'atomic'.

G. J. Stoney's introduction of the word 'electron' to describe the 'atom' of electricity, and Faraday's introduction of the words 'cation' (positive charge) and 'anion' (negative charge) can be mentioned at this stage.

Charge–mass ratios

The experiments on electrolysis give the following values for ionic charge to mass ratios:

charge–mass ratio for silver ions, Ag^+, $0{\cdot}0895 \times 10^4$ C g^{-1};
charge–mass ratio for copper ions, Cu^{2+}, $0{\cdot}304 \times 10^4$ C g^{-1};
charge–mass ratio for hydrogen ions, H^+, $9{\cdot}58 \times 10^4$ C g^{-1}.

The charge–mass ratio for the electron, $1{\cdot}76 \times 10^8$ C g^{-1}, is thus greater than the ratio for the hydrogen ion by the factor

$$\frac{1{\cdot}76 \times 10^8}{95\,800} = 1836.$$

Now since the electrolysis experiments show that ionic charges are small integral multiples of the charge on the hydrogen ion, the high charge–mass ratio for the electron must be ascribed to a very small value for the electronic mass m, and not to a very high value for the charge. We thus define the electron as a negatively charged particle of mass 1/1836 that of the proton (H^+).

Some further electron properties

Further examples of the properties of electrons may now be mentioned, for example:

(i) the Millikan experiment for the determination of electronic charge: here, the sudden changes in velocity of an oil drop moving in an electric field dramatically illustrate the arrival of one or more electrons on the drop;

(ii) the thermionic effect: here it should be noted that the charged particles emitted by heating different substances are always electrons, and that this effect is the foundation of the science of experimental electronics (although in recent years transistors, which work on a different principle, have been introduced);

(iii) the photoelectric effect: the particles ejected from metal surfaces illuminated by visible or ultraviolet light can again be shown to be electrons.

The Rutherford atom

Atoms are electrically neutral, so that if the negatively charged electrons are common constituents of all matter, atoms must consist of some arrangement of positive and negative charges. Early atomic structure theories pictured the atom as a cloud of positive charge with electrons as negative particles arranged either inside the positive charge cloud (Thomson) or outside it (Perrin). The historical approach to the Rutherford nuclear theory of the atom requires consideration of:

(i) the discovery of X-rays by W. K. Röntgen;
(ii) the discovery of radioactivity by H. Becquerel;
(iii) the discovery of alpha, beta and gamma radiation;
(iv) the discovery of scattering of alpha particles by matter.

These effects cannot easily be demonstrated although scintillations due to alpha particles can be observed through a microscope in a darkened room, for example in the luminous paint on the hands of a clock.

The scattering of alpha particles through large angles can be explained by assuming that the positive charge and almost all of the mass of the atom is concentrated in a small nucleus of diameter about 10^{-13} cm. Atomic sizes are of the order of 10^{-8} cm (i.e. about 10^5 times the size of the nucleus) and it is very important that the reader should appreciate how great is this difference. It is suggested that the relative sizes might be illustrated by using a football to represent the nucleus, and working out the size of the stadium in which this ball is located.

Measurement of nuclear mass and charge

Removal of one or more electrons from an atom produces a positively charged particle, whose charge–mass ratio can in principle be determined by deflection experiments using electric and magnetic fields. Very many mass spectrographs have been designed to measure this ratio. A recent version, the 'time-of-flight' mass spectrograph, can be described in simple terms. The time taken for a positively charged particle to pass from one electrode to another, maintained at a constant potential difference, will depend on the mass, heavier particles taking longer than light ones. The arrival of particles of different masses can be detected electronically and displayed on a cathode ray tube. If the instrument is calibrated using particles whose mass is defined as a standard reference value, nuclear masses can be measured very accurately. Until recently oxygen = 16 was taken as the standard, but now carbon = 12 has been adopted since more accurate mass measurements are possible if this latter standard is used.

If the particle whose mass has thus been determined carries an elementary positive charge, the corresponding atomic weight of the neutral atom can be obtained by adding the weight of one electron. In this way a table of atomic weights of very high accuracy can be obtained without recourse either to the determination of chemical equivalents or to assumptions about valency or 'combining power' of the atoms. The mass spectrometer shows that the charge–mass ratio for hydrogen ions, H^+, is 1836 times smaller than the corresponding ratio for the electron, so that if we assume the charge to have the same magnitude but opposite sign in the two cases, the mass of the proton must be 1836 times as great as the electron mass.

The results for helium, however, show that the charge–mass ratio corresponds to a doubly charged particle of mass about four times that of the proton. It was therefore assumed at one time that the nucleus of the helium atom contained four protons mixed in some way with two electrons to reduce the over-all positive charge to two units. Later work showed that this was not correct, and it is now thought that the nucleus of helium contains two protons and two *neutrons*, the latter being uncharged particles of mass almost the same as the proton mass.

The total number of particles in the nucleus, that is, protons + neutrons, is called the *mass number*. Hydrogen thus has a mass number of 1 and an atomic number of 1, whereas helium has a mass number of 4 and an atomic number of 2.

It will be seen later that the chemical properties of atoms are mainly determined by the number and the arrangement of the electrons around the nucleus. The number of these electrons will be equal to the atomic number of the atom, since there must be a sufficient number of electrons to neutralize the positive charge on the nucleus. An arrangement of atoms in order of atomic number thus provides a logical sequence in which to discuss the properties of the elements, and a chemical element itself may now be defined as a substance all of whose atoms have the same atomic number.

Nuclear charges can be detected and measured in the indirect, but very beautiful experiments of Moseley. He measured the frequencies of the 'characteristic' X-rays emitted when elements were bombarded with cathode rays of suitable energy. Then, by assigning an integer Z to each element, starting with $Z = 1$ for hydrogen, and increasing by unit steps in order of increasing atomic number, he was able to show experimentally that $\sqrt{\nu}$ is proportional to $Z-a$, where ν is the frequency and a is a constant. Z was identified as the number of units of positive charge (protons) in the nucleus.

The results of these experiments were important for a number of reasons, for example:

(i) the identification of the atomic number as the number of protons in the nucleus;
(ii) the correction of the order in which certain elements occur in the periodic classification – Ar, K; Co, Ni; Te, I;
(iii) the identification of elements by characteristic X-radiation.

The Bohr atom

The next step in the study of atomic structure is to discover how the electrons are arranged. Some important experimental evidence comes from the study of atomic spectra.

If an electric discharge is passed through a tube containing hydrogen gas at low pressure, a pink coloured glow is observed. This can be examined using either a simple spectroscope, or an inexpensive replica grating, and a few sharp, coloured lines can be observed. Each of these lines has a definite frequency, ν, which can be measured. Bohr explained the radiation of energies at a limited number of definite frequencies by adapting the quantum theory to the problem of atomic structure. This theory, developed by Max Planck in 1900 to account for the results of experiments on the radiation emitted by hot bodies, introduced the entirely new idea that energy was emitted in small 'packets' called *quanta*. These quanta have energy E given by $E = h\nu$, where ν is the frequency and h a universal constant called the Planck constant. The Bohr theory of the hydrogen atom assumes:

(i) that electrons move in closed circular paths or orbits around the nucleus;
(ii) that the energy of an electron in a given orbit is constant (this is in contradiction to classical electromagnetic theory); and
(iii) that radiation is emitted only when an electron undergoes a transition from one orbit to another orbit of different radius. The frequency of the radiation emitted is given by a very simple expression $\Delta E = h\nu$, where ΔE represents the energy difference between two levels (say $E_2 - E_1$), h is the Planck constant and ν the frequency. In this way the experimental observation of sharp lines in the hydrogen atomic spectrum can be linked with the energies of the electron moving in one or other of a number of discrete orbits.

Bohr was further able to calculate that the possible values for the electron energy E in the hydrogen atom were given by the expression:

$$E_n = -\frac{Rch}{n^2}$$

where R is constant (the Rydberg constant), c is the velocity of light, h is Planck's constant and n is an integer called the principal quantum number, taking values 1, 2, 3, 4, etc. (The minus sign appears since the energy zero is defined as the state in which the electron is at infinite separation from the nucleus.)

This expression can now be used to construct an *energy level diagram* of the type shown in Figure 1. Energy is plotted on a vertical scale, with the $E = 0$ value, corresponding to complete removal of the electron (ionization), at the top. The energy of the electron in a given orbit is shown by a horizontal line, and it is in fact customary to speak of the *energy levels* of an atom. Letters K, L, M, etc. are assigned to the various levels, as shown in the diagram. The electron in a hydrogen atom is normally in the K-orbit or K-level, and this is called the 'ground' state – or the state of lowest energy. When energy is fed into the system, as in the discharge tube, the electron is 'excited' into a higher energy state (e.g. *L* or *M*); it will eventually revert to a lower, more stable state, and in doing so radiation of a particular frequency will be emitted. Bohr showed that the lines observed in the visible region of the hydrogen spectrum arose from transitions ending on the *L*-level; lines of much greater frequency were later discovered by Lyman in the ultraviolet region of the spectrum, and these arise from transitions ending on the ground level.

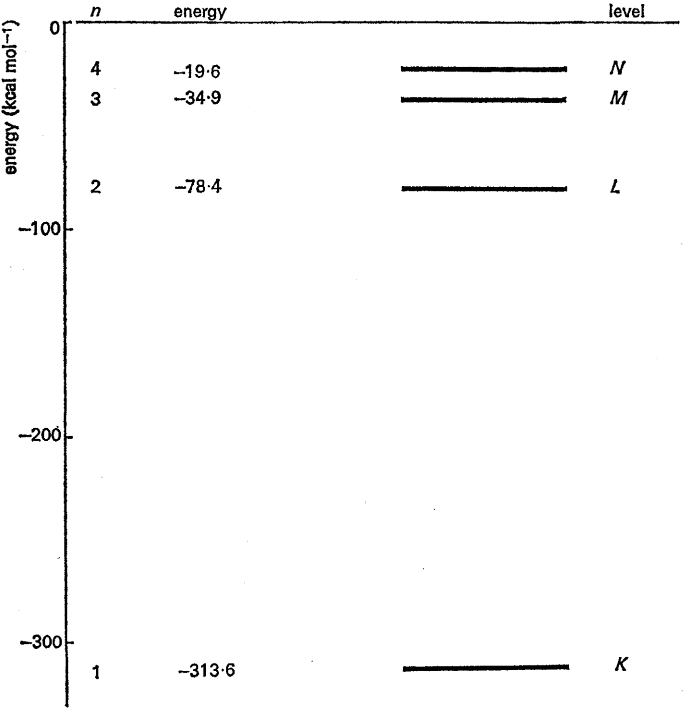


Figure 1 Energy level diagram for hydrogen

One of the first triumphs of the Bohr theory was the good agreement between the electron energies of the different orbits calculated by the theory and determined experimentally from the measured frequencies of the radiated energy. Now the energy of the electron in the K-level is $-313{\cdot}6$ kcal mol^{-1} (or 13·59 electron volts per atom)[1]; this means that 313·6 kcal mol^{-1} of energy must be supplied to remove the electron completely from the atom.

The Bohr theory accounted satisfactorily for the energy relations in the hydrogen atom but was difficult to apply to other atoms. As a result of further theoretical study, it was replaced some ten years later by a new quantum theory. We shall develop this new theory using ionization energies as the principal experimental tool.

Ionization energies

The problem of the arrangements of the electrons in atoms other than hydrogen can best be approached by a consideration of ionization energies. Figure 2 shows the first ionization energy of the elements from hydrogen to sodium inclusive. (N.B. The *first* ionization energy is the energy required to remove one electron from a neutral atom, the *second* ionization energy is the energy needed to remove an electron from a singly charged ion . . . and so on.) The sharp drop in ionization energy from helium to lithium can be explained if it is assumed that, whereas in helium two electrons occupy the K-level, in lithium there are two electrons in the K-level and one in the L-level, which represents an orbit of greater radius. The two electrons in helium are close to the nucleus, and it needs a lot of energy to drag one of them out of the atom. In lithium, however, the 'outer' electron in the L-level is further away from the nucleus, and it is screened from the full attractive force of the nucleus by the two 'inner' K-electrons. Further study of this and similar ionization energy diagrams (Figures 2 and 3) provides experimental evidence for the arrangement of the electrons in energy levels and sub-levels – the so-called 'electron configurations'.

Electron Configurations, the Pauli Principle and the 'Building-up' (*Aufbau*) Principle

The quantum numbers

Figure 3, which shows the successive ionization energies for potassium, can be explained if the nineteen electrons in potassium

---

[1] 1 eV atom^{-1} = 23·069 kcal mol^{-1}.

are arranged in a number of different energy levels. Thus the well-defined breaks in the ionization energy diagram after 1, 8, 8 and 2 electrons respectively have been removed, can be correlated with the 'principal quantum number *n*' which we have already introduced on page 19. This quantum number is related to the energy, and the ionization energy diagram for potassium tells us that in potassium there are 2 electrons in a level associated with principal quantum number 1, 8 in the level for which $n = 2$, 8 in the level for which $n = 3$, and 1 in the level for which $n = 4$. This electron configuration for potassium can be written therefore as 2, 8, 8, 1.

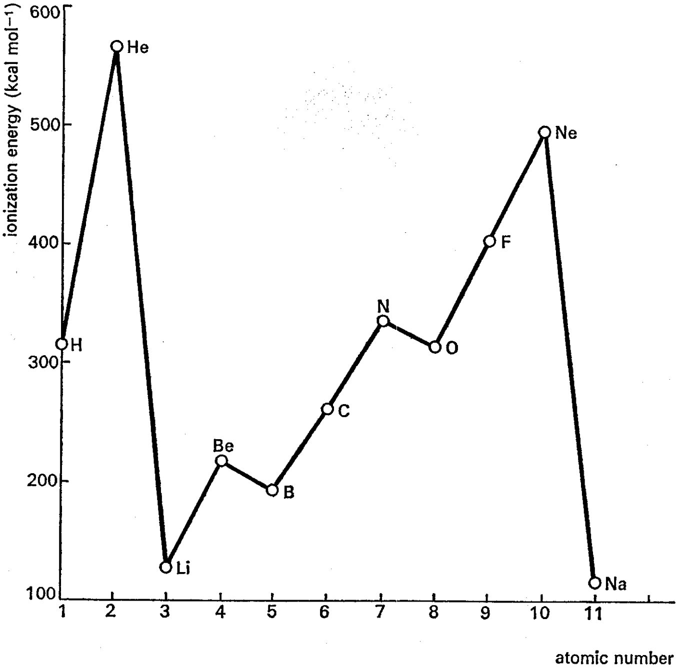


Figure 2 First ionization energies

A more detailed study of these diagrams shows that the levels containing 8 electrons are themselves sub-divided into groups of 2 and 6 electrons, so that the potassium configuration is now written 2; 2, 6; 2, 6; 1. At this stage the *four* quantum numbers must be introduced to describe electrons in the different levels and sub-levels. The procedures might appear arbitrary, but it can be pointed out that there are good theoretical reasons for them, which advanced students will eventually hear about. We therefore

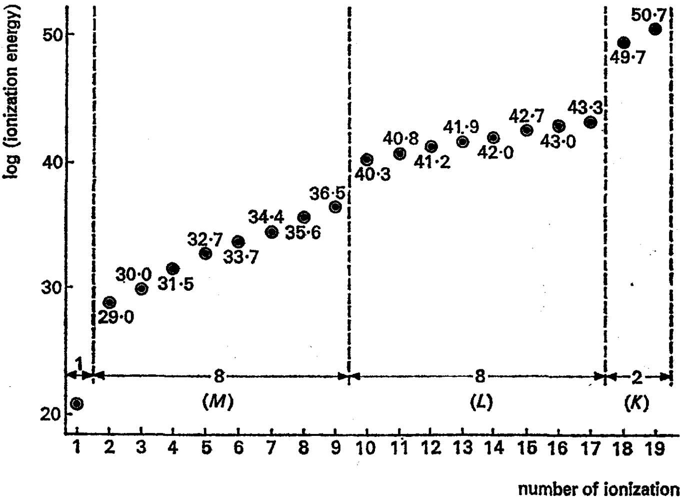


Figure 3 Successive ionization energies for potassium

state that four quantum numbers can be associated with an electron as follows:

n — the principal quantum number, which takes the values $1, 2, 3, \ldots, \infty$,

l — the azimuthal quantum number, which takes the values $0, 1, 2, \ldots, (n-1)$,

m — the magnetic quantum number, which takes the values $l, (l-1), \ldots, 0, \ldots, -(l-1), -l$,

s — the spin quantum number, which takes the values $\pm\frac{1}{2}$.

It is also convenient at this stage to introduce the description of electrons for which $l = 0, 1, 2, 3$, etc. as *s*-, *p*-, *d*- and *f*-electrons.

The Pauli principle

The Pauli exclusion principle states that in any atom no two electrons can be assigned the same set of four quantum numbers. No proof of this principle exists but so far nothing in nature has appeared to contradict it, and it brings a most beautiful order into the study of electron configurations, atomic structure and molecular structure. Table 1 shows the maximum number of electrons that can be accommodated in the different energy levels when the Pauli principle is taken into account.

Thus two electrons can occupy a level for which $n = 1$. These electrons must have 'opposite spins', that is, if the spin quantum number for electron 1 is $+\frac{1}{2}$, that for electron 2 must be $-\frac{1}{2}$. We can write the configuration of the helium atom in a number of ways, for example a 'shorthand' notation $1s^2$, where the first figure gives the principal quantum number n, s gives the l quantum number and the superscript gives the number of electrons in this particular level, or a more detailed description $\underset{1s}{\boxed{\uparrow\downarrow}}$ where the arrows represent electrons with opposite spins (such electrons are said to be 'paired').

When $n = 2$ there are two sub-levels, one ($l = 0$) containing a maximum of two electrons, and another ($l = 1$) containing six electrons. Thus the electronic configuration of the neon atom can be written:

$$1s^22s^22p^6 \text{ or } \underset{1s}{\boxed{\uparrow\downarrow}}\ \underset{2s}{\boxed{\uparrow\downarrow}}\ \underset{2p}{\boxed{\uparrow\downarrow\,|\,\uparrow\downarrow\,|\,\uparrow\downarrow}}$$

The Aufbau *or 'building-up' principle*

We can now take the atoms in order of increasing atomic number and work out the electron configurations. The *Aufbau* principle states that the electronic arrangement in any atom is obtained by assuming the arrangement is that in the atom immediately preceding the one in question, together with one more electron which goes into the level of lowest available energy – due regard being paid to the requirements of the Pauli principle. Thus, to get the arrangement in lithium, of atomic number 3, we first of all write down the arrangement in helium (atomic number 2) $1s^2$, and then put the third electron into the lowest available level – in this

Table 1
The Distribution of Electrons in Energy Levels According to the Pauli Principle

n	l	m	s	*Maximum no. of electrons*
1	0	0	$\pm\frac{1}{2}$	2
2	0	0	$\pm\frac{1}{2}$	2
	1	−1, 0, +1	$\pm\frac{1}{2}$	6
3	0	0	$\pm\frac{1}{2}$	2
	1	−1, 0, +1	$\pm\frac{1}{2}$	6
	2	−2, −1, 0, +1, +2	$\pm\frac{1}{2}$	10
4	0	0	$\pm\frac{1}{2}$	2
	1	−1, 0, +1	$\pm\frac{1}{2}$	6
	2	−2, −1, 0, +1. +2	$\pm\frac{1}{2}$	10
	3	−3, −2, −1, 0, +1, +2, +3	$\pm\frac{1}{2}$	14

case $2s^1$. We must therefore know the relative energies of the different levels and sub-levels. The information can be got experimentally from atomic spectra and ionization energy measurements and is displayed in Figure 4.

Beryllium will have the configuration $1s^2 2s^2$, but with boron the additional electron must go into a new level, since the $2s$ level is already fully occupied. Reference to Figure 4 shows that the next available level is the $2p$ one, so that boron is assigned the

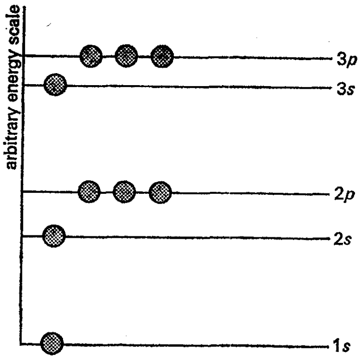


Figure 4 Energy levels (atomic number < 21)

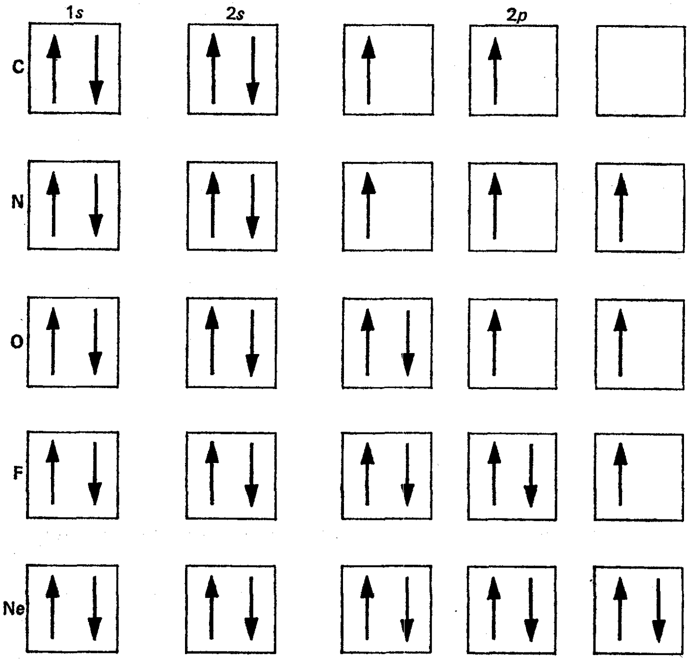


Figure 5 Electron configurations – carbon to neon

configuration $1s^2 2s^2 2p^1$. The filling up of the p-levels then continues for the elements carbon to neon, and we get the configurations:

C $1s^2 2s^2 2p^2$
N $1s^2 2s^2 2p^3$
O $1s^2 2s^2 2p^4$
F $1s^2 2s^2 2p^5$
Ne $1s^2 2s^2 2p^6$

A careful study of the atomic spectra of these elements, together with measurements of ionization energies, shows that when p- (or d- or f-) levels are being filled up, the electrons distribute themselves so that, as far as possible, they retain parallel spins. This is because electrons repel each other, and they will not pair together if levels of suitable energy are available which can be singly occupied. We can therefore write the configurations of the elements carbon to neon inclusive in the form shown in Figure 5. The configuration of nitrogen, with three unpaired electrons, is rather stable; this is clearly shown in Figure 2 where the ionization energy for oxygen is smaller than that for nitrogen.

The Theoretical Foundation of the Periodic System

It is possible to show experimentally that the elements can be arranged in groups possessing similar chemical properties. We can now write down the electron configurations of these various groups.

Group I. *The alkali metals*

Li $1s^2 2s^1$
Na $1s^2 2s^2 2p^6 3s^1$
K $1s^2 2s^2 2p^6 3s^2 3p^6 4s^1$
etc.

Here each atom has one 'outer' electron in an s-level, and all the other 'inner' electrons completely fill the s- and p-levels of smaller principal quantum number. The characteristic chemical properties of the group are evidently linked with the ease with which the outer electron can be removed (low ionization energy) to give an ion of charge $+1$ with a noble gas structure.

Group II. *The alkaline earth metals*

Be $1s^22s^2$
Mg $1s^22s^22p^63s^2$
Ca $1s^22s^22p^63s^23p^64s^2$
etc.

Here each member of the group has two 'outer' electrons, and removal of these will give ions with charge +2 and noble gas configurations. We need not give the configuration of each group in detail. The number of 'outer' electrons in the main group elements steadily increases as we go across the table, so that the halogens in Group VII have the configuration:

F $1s^22s^22p^5$
Cl $1s^22s^22p^63s^23p^5$
Br $1s^22s^22p^63s^23p^63d^{10}4s^24p^5$
etc.

Here the addition of one electron to each atom will produce halide ions with unit negative charge and noble gas structures. It is important to note that all the noble gas elements (except helium) have configurations with completely filled *s*- and *p*-levels:

Ne $1s^22s^22p^6$
Ar $1s^22s^22p^63s^23p^6$
Kr $1s^22s^22p^63s^23p^63d^{10}4s^24p^6$
etc.

There are eight electrons in the outer shell in each case, and these are often called the stable octet.

These examples show that the electron configuration of the main group elements is a periodic function of atomic number, and the arrangement arrived at on the basis of chemical properties is also obtained if one collects elements with analogous configurations into groups.

An interesting situation arises with element number 21, scandium. We write down, first of all, the configuration for element number 20, calcium. This is $1s^22s^22p^63s^23p^64s^2$ and we then look for the lowest available level into which we can put one more electron. Now for atoms of atomic number greater than 20, the order of energy levels is found experimentally to be that shown in Figure 6, that is the next electron goes into a 3*d* level, and not a 4*p* level, and we get a sequence of ten elements in which electrons are gradually filling up the five *d*-levels, each of which can contain 2 electrons. Thus, for example, Ti has outer electrons arranged to give [Ar] $3d^24s^2$, where [Ar] represents the electron configuration of argon. Mn has outer electrons arranged to give [Ar] $3d^5 4s^2$. The sequence of ten elements forms the so-called first transition series. Here again we can show that vertical groups of transition metals (see Table 2) have similar arrangements for their outer electrons. A difficulty very often arises at this point. In titanium, for example, the last two electrons to be added go into the 3*d* level, and it is natural to assume that in the formation of the Ti^{2+} ion, these two *d*-electrons will be removed. In fact this is not so; the relative order of the energy levels shown in Figure 6 may change when the atoms acquire positive charges. As the positive charge increases (for example in the series Ca, Sc^+, Ti^{2+}) the energy levels tend to arrange themselves in order of increasing quantum number (for example 3*s*, 3*p*, 3*d*, 4*s*, 4*p*, etc.). The

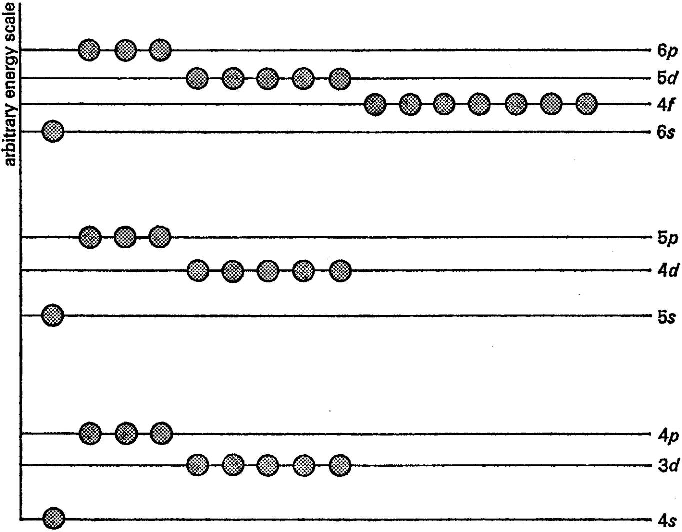


Figure 6 Energy levels (atomic number > 21)

Table 2

The Periodic Table

	s block		d block										p block					
Group	I	II	IIIA	IVA	VA	VIA	VIIA	VIII			IB	IIB	III	IV	V	VI	VII	
	Li	Be	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	B	C	N	O	F	Ne
	Na	Mg	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	Al	Si	P	S	Cl	Ar
	K	Ca	La[1]	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Ga	Ge	As	Se	Br	Kr
	Rb	Sr	Ac[2]										In	Sn	Sb	Te	I	Xe
	Cs	Ba											Tl	Pb	Bi	Po	At	Rn
	Fr	Ra																

f block													
[1]Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
[2]Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr

configuration of the outer electrons of Ti^{2+} is therefore $3d^2$, and not $4s^2$.

Element number 57, lanthanum, has a configuration $5d^1 6s^2$ for its outer electrons. Element number 58, cerium, however, does not have the configuration $5d^2 6s^2$ for its outer electrons, since at this stage the available level of lowest energy for the next electron to be added is $4f$. Thus, starting at cerium we get a sequence of 14 elements in which electrons are filling up the $4f$ level. These elements are usually classified together as the *lanthanide* series. They have very similar chemical properties, since the only difference between them is in the number of 'inner' $4f$ electrons. A similar sequence may start at element number 89, actinium. From thorium onwards electrons can fill up an inner $5f$ shell. However, in these heavy atoms the energy levels are very close together and in many of their reactions thorium and protactinium behave as though they had the electronic structure characteristic of a transitional series – that is, the electrons may occupy $6d$ levels.

If therefore we assume that chemical properties are determined by electronic configurations, then the Pauli principle and the *Aufbau* principle provide a theoretical foundation for the periodic system. The electronic configurations of the noble gases now become extremely important, since they evidently represent very stable arrangements. This is revealed not only by the 'inertness' of these gases, but also by their very high ionization energies, as shown in Figure 2.

The Charge Cloud Model

Charge clouds and electron density

The Bohr theory of atomic structure implies that the position and the momentum of an electron can in principle be defined precisely. There is a fundamental difficulty here, however, since no experiments can be devised to measure these quantities simultaneously; any attempt to locate the electron introduces an interaction that changes the momentum of the electron. This difficulty is usually referred to as the Heisenberg uncertainty principle, which states that it is impossible to measure, simultaneously and precisely, both the position and the momentum of an electron. We have, therefore, to rely on the 'probability' or the 'chance' of finding an electron at a particular point, and this leads to the so-called 'charge cloud' model of atomic structure.

This model pictures the electrons as a 'cloud' of negative charge surrounding the nucleus, the number of units of negative charge in this cloud being equal to the number of electron particles in the Rutherford or Bohr model of the same atom. Consider a hypothetical experiment: suppose that the position of an electron in a hydrogen atom could be photographed at any given instant. The electron is moving in an unknown path at high speed, and a fraction of a second later a new photograph would record the electron in a new position. The superposition of several million such photographs, taken at very small intervals of time, would produce a picture resembling a cloud made up from an enormous number of points, each representing a position of the electron. The density of such a cloud can be defined as the number of points in unit volume, and an important feature of the electron charge cloud in atoms is that the density is not uniform throughout the cloud; in hydrogen, for example, there is a large density per unit volume close to the nucleus, but the density decreases as the distance from the nucleus increases. Figure 7 shows a section through the hydrogen atom charge cloud; most of the electron charge is concentrated close around the nucleus, but there is a finite, if small, chance that the electron might be found quite a long distance from the nucleus. It is important to emphasize that this picture of the hydrogen atom represents a total charge of one electron unit. In theory there is no sharp boundary to the charge cloud, but in practice a boundary surface can be drawn to enclose

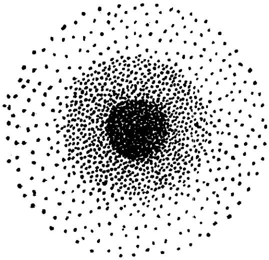


Figure 7 The hydrogen atom charge cloud

almost all of the charge, since the density decreases very rapidly as the distance from the nucleus increases.

Whereas in the Bohr theory electrons were said to move in orbits, electrons in the charge cloud occupy *orbitals*; electrons in *s*-, *p*-, *d*- and *f*-levels occupy *s*-, *p*-, *d*- and *f*-orbitals. The same energy level diagrams shown in Figures 4 and 6 are used to describe the energies of the orbitals. Thus when $n = 1$, there is one *s*-orbital, when $n = 2$ there are one *s*- and three *p*-orbitals, when $n = 3$ there are one *s*-, three *p*- and five *d*-orbitals and so on.

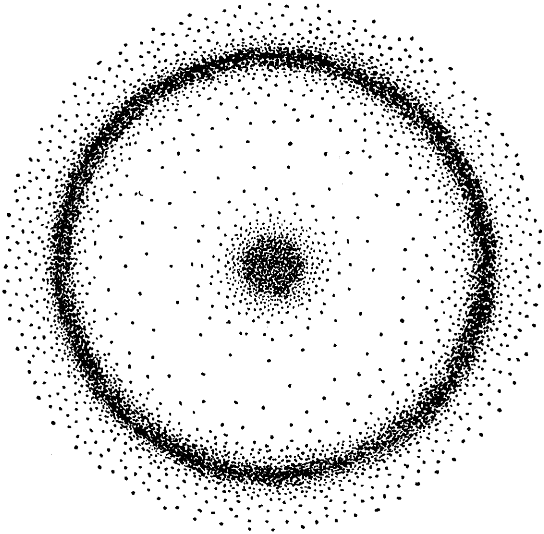


Figure 8 The lithium atom charge cloud

We can now consider electron orbitals in a number of atoms. The single electron in hydrogen forms a spherical charge cloud, as do the two electrons in helium. Figure 8 shows a section through the charge cloud of a lithium atom; the greater part of the charge cloud of two of the three electrons is mainly concentrated in a spherical cloud close to the nucleus, but the charge cloud of the third electron is mainly concentrated in a spherical shell some distance further out. Thus there is a distinction here between 'inner' and 'outer' orbitals just as there was in the Bohr theory.

The s- and p-orbitals

Hydrogen, helium, lithium and beryllium atoms are described by spherical charge clouds and, in general, all *s*-electrons (that is electrons for which the *l* quantum number is zero) have spherical charge clouds – they are in *s*-orbitals. The size of the orbital depends upon the principal quantum number – as *n* increases the radius of the orbital increases.

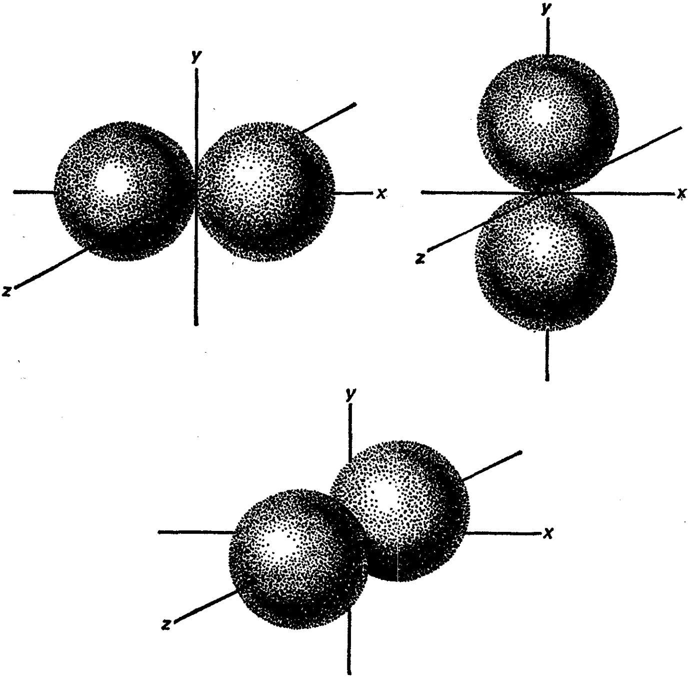


Figure 9 The *p*-orbitals

The *p*-electrons, however, do not have spherical orbitals. The shape of the charge cloud is shown in Figure 9; there are two 'pear-shaped' clouds or 'lobes' which have zero density at the nucleus, and maximum density on either side of the nucleus along collinear axes. Now Table 1 and Figure 4 show that for a given *n* there are three *p*-levels of equal energy. The charge cloud theory represents these as three *p*-orbitals of the same shape and of equal energies, but oriented along three Cartesian axes *x*, *y* and *z* as shown in Figure 9.

The d-orbitals

These are rather more complicated since unlike the *p*-orbitals they differ in shape as well as in orientation. Four of the five *d*-orbitals for a given principal quantum number can each be represented by the four lobes with mutually perpendicular axes shown in Figure 10(a); these axes may lie along two of the three Cartesian axes, or they may make an angle of 45° with selected pairs of Cartesian axes. The fifth *d*-orbital is shown in Figure 10(b); here there are two lobes aligned along the *z*-axis, with a 'collar' or 'tyre' of charge surrounding the nucleus.

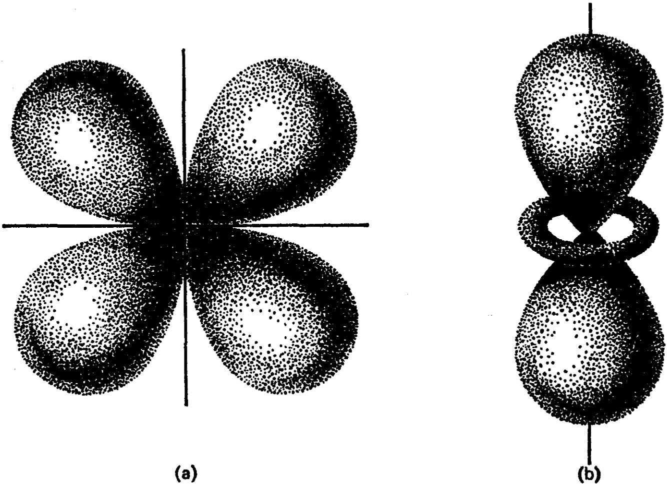


Figure 10 The *d*-orbitals

Hybrid Orbitals

The orbitals we have so far discussed are not necessarily those which are involved when atoms combine to form molecules. A simple principle – namely, that electrons repel each other and try to get as far apart as they can – is used to discuss the geometry of the orbitals linking atoms in molecules.

In general, these orbitals contain two (paired) electrons and these paired electrons also try to keep as far apart as possible. Thus when two electron pairs in the valence shell of an atom are involved in molecule formation, as for example beryllium in $BeCl_2$ (gaseous), the two orbitals are collinear, since this geometry puts them as far apart as possible (see Figure 11a). Figure 11(b) shows how three electron pairs are arranged in coplanar (trigonal) orbitals with an angle of 120° between adjacent axes (for example boron in BCl_3). Four electron pairs are arranged in orbitals pointing from the centroid to the vertices of a regular tetrahedron – for example carbon in CH_4 (Figure 11c). These are the so-called hybrid orbitals, since they can be described mathematically in terms of combinations of the functions used to describe *s*- and *p*-orbitals. Thus the collinear, coplanar and tetrahedral orbital arrangements are described as *sp*, sp^2 and sp^3 hybrid orbitals respectively.

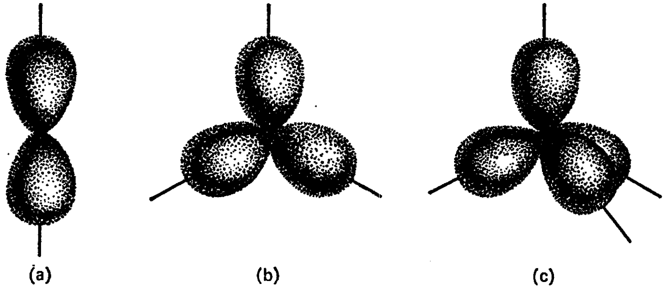


Figure 11 (a) collinear *sp* hybrid orbitals
(b) coplanar sp^2 hybrid orbitals
(c) tetrahedral sp^3 hybrid orbitals

For many purposes it is quite sufficient to say that charge clouds repel each other according to the Coulomb law, and that when there are around an atom, two, three or four electron pairs, the stable arrangement will be one in which the charge clouds are as far apart from each other as possible.

Bibliography

G. I. BROWN, *A New Guide to Modern Valency Theory*, Longmans, 1967.
E. CARTMELL and G. W. A. FOWLES, *Valency and Molecular Structure*, 3rd edn, Butterworths, 1966.[1]
J. E. FREY, 'Discovery of the noble gases and foundations of the theory of atomic structure', *J. Chem. Educ.*, vol. 43 (1966), no. 7, pp. 371–4.
L. PAULING, *The Chemical Bond*, Oxford University Press, 1967.
H. H. SISLER, *Electronic Structure, Properties and the Periodic Law*, Chapman & Hall, 1965.
J. E. SPICE, *Chemical Binding and Structure*, Pergamon, 1964.

[1] More advanced.

2 R. L. Livingston

Crystal Geometry

R. L. Livingston, 'The teaching of crystal geometry in the introductory course', *Journal of Chemical Education*, vol. 44 (1967), no. 7, pp. 376–82.

Crystal structure is becoming an increasingly important topic in chemistry courses. It is the purpose of this paper to outline an approach to crystal structure at the elementary level which will prepare the student for more advanced work in this field, or which could be used as the beginning in a more advanced course. Crystal geometry is stressed, giving a foundation for the study of the determination of atomic positions by X-ray diffraction. Coordinates for atoms within the unit cell are given for most structures in order to acquaint the student with the notation employed in the description of crystal structures in standard reference works.

Lattices and Unit Cells

In order to discuss the internal structure of a given crystal, the X-ray crystallographer selects a convenient coordinate system with which to designate the positions of the atoms (or ions) within the crystal. Various coordinate systems are used to define *lattices*; a lattice is merely a *regularly repeated set of points*.

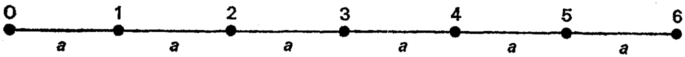


Figure 1 A one-dimensional lattice. If there are atoms at each lattice point, the coordinates of the atoms are 0, 1, 2, 3, . . ., and −1, −2, −3, . . .

For example, Figure 1 shows a one-dimensional lattice – a set of points regularly repeated along a straight line. The separation between adjacent points is always the same and is designated by *a*. The set of points may be imagined to continue an infinite distance in either direction.

It is important to understand that this set of lattice points may be generated from only one point provided we are told that it is a one-dimensional lattice and are given the value of the distance *a*. With this information, we may start at point O and move a distance *a* to the right, generating point 1. Moving one more unit distance *a* to the right generates point 2, etc. And, of course, we can start at point O and move a unit distance *a* to the left, then another unit distance *a* to the left, etc. In this manner we build up the entire set of lattice points by what we call unit translations. If there is an atom at each lattice point, we have a one-dimensional crystal.

Now consider a two-dimensional lattice such as is shown in Figure 2. In order to specify such a lattice, we need to give the angle γ between the directions of translation and specify the values *a* and *b* of the unit translations along these directions. Once these three quantities are known for a two-dimensional lattice, the entire lattice (set of points) can be generated by unit translations along the directions of the two axes. We would thus generate not only the few points shown in Figure 2, but would generate a set of lattice points extending to infinity in two dimensions. Since we are restricted to two dimensions, all of our points lie in one plane.

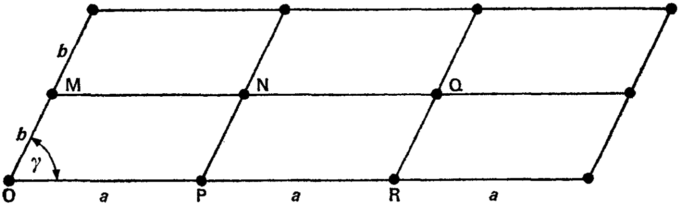


Figure 2 A two-dimensional lattice. If there are atoms at each lattice point, the coordinates of the atoms are given by all possible combinations of two integers

If point O is our arbitrarily chosen origin, then we can give coordinates for any of our lattice points. Thus point Q has coordinates (2, 1); this means it is away from the origin by 2 (positive) units of *a* and one (positive) unit of *b*. There is thus a lattice point corresponding to every pair of integers (positive and negative) which we might care to write down, for example, (5, 7), (−2, 23), (489, −35), etc. It should be clear that a and b are *vectors*; they have both length and direction.

Now imagine that each lattice point in Figure 2 represents an atom. We then have a (hypothetical) two-dimensional crystal with only a few atoms of the crystal being shown in the figure. This crystal could be described in either of the two following ways.

Method I. Give the magnitudes of a and b, the angle between a and b and state that there is an atom at the origin.

With this information, unit translations of the atom at the origin will generate all of the other atoms; there will be an atom at each of the lattice points. Coordinates of the atoms are (0, 0), (0, 1), (1, 0), (1, 1), etc.

Method II. In this case, a two-dimensional *model* is used. The model consists of the four atoms designated O, M, N and P; these form a parallelogram with sides of length a and b and with the specified angle γ between these sides. If we start with this model (containing four atoms) in its original position and move it by a distance a to the right, atom N generates atom Q, atom P generates atom R while atoms M and O take the positions originally occupied by atoms N and P, respectively. By repeating such unit translations among both the a and b directions, the entire (two-dimensional) crystal is built up.

While either of the above methods will suffice, the second method is often to be preferred since it gives a 'picture' of the crystal that is easier to visualize; this is especially true in three-dimensional cases.

The four atoms O, M, N and P in Figure 2 have a special significance. These four atoms, arranged as they are in a parallelogram, constitute a (two-dimensional) *unit cell.* The unit cell is frequently chosen as the 'model' of the crystal. It is usually chosen in as simple a way as possible – atoms O, M, N and P or atoms P, N, Q and R, for example; but it could be chosen in another way – atoms P, M, N and R, for example. Unit translations of any of these unit cells would generate the entire crystal.

In many patterns of wallpaper, there is a regular repetition scheme and it is possible to pick out a (two-dimensional) unit cell.

This is chosen so that unit translations of the cell will generate the entire pattern.

Returning to our description of Figure 2, we recall that in our first method we gave the coordinates of one atom but in our second method we showed four atoms. The question thus arises, how many atoms are there per cell? There is only one *atom per unit cell* in this case. It may be regarded as the atom at the origin (0, 0) or it may be computed from the model of the unit cell. Each of the four atoms (say O, M, N and P) is actually shared by four unit cells; this is most apparent for atom N but is also true for the other atoms. There is thus the equivalent of $\frac{1}{4}$ of an atom at each of the four positions O, M, N and P, and this makes a total of one atom per unit cell.

The above example has been given in some detail for a two-dimensional case since most students find it easier to visualize arrangements in two dimensions than in three dimensions. In order to discuss most crystals, we must extend our thinking to three dimensions.

To extend Figure 2 to three dimensions, we merely add another vector which we designate as c; this vector must not be parallel to either a or b. To specify the direction of this vector, we must give values for two angles: the angle β between a and c and then the angle α between b and c. In the most general case, the length of the vector c would be different from the length of either a or b. Thus to define our three-dimensional array of lattice points we need to specify six quantities: three lengths (a, b and c) and three angles (α, β and γ).

In practice, the structures of certain crystals are readily described in terms of this type of coordinate system. Many crystals possess some symmetry and hence may be described in a system in which two or all three of the angles are equal to one another and possibly to 90° and/or in which two or three of the unit translations are equal. The system with maximum symmetry is the *cubic system*. In this system $a = b = c$ and $\alpha = \beta = \gamma = 90°$. This is, no doubt, the three-dimensional coordinate system with which the student is most familiar, and the lattice points in the cubic system are easy to visualize. Note also that only one quantity (parameter) need be specified when we are dealing with an element which forms a cubic crystal; this is usually designated as a or a_0

and represents the edge of a unit cell as well as the length of our unit translations.

If the lengths of a, b and c are different from one another but the three angles are 90°, then our coordinate system (or crystal system) is said to be orthorhombic or merely *rhombic*. If $a = b \neq c$ and if the three angles are 90°, we have the *tetragonal* system.

Cubic Crystals

Simple cubic

A unit cell (or model) of a simple cubic crystal is shown in Figure 3. If we take atom O as the origin, then atom A has coordinates (1, 0, 0). (These are simply the x-, y- and z-coordinates in terms of a_0; that is, a coordinate of 1 means one unit of a_0.) Atom B has coordinates (1, 1, 0), atom E is at (1, 1, 1), etc. All atoms are assumed to be identical.

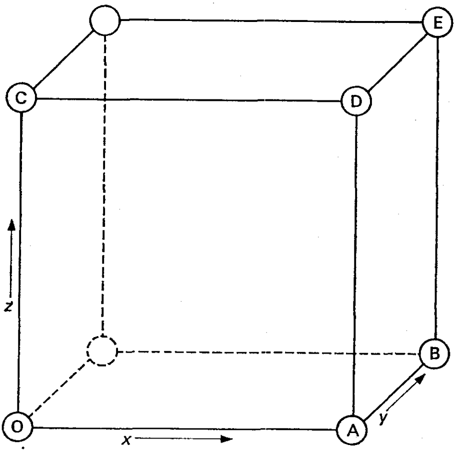


Figure 3 The unit cell for a simple cubic crystal. There are atoms at (0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0), etc.

The entire crystal is visualized as being built up by unit translations of this model; this process would generate atoms at all lattice points in our simple cubic three-dimensional array of points. There would thus be atoms occurring at all points in space whose coordinates (in terms of a_0) can be expressed by three integers, positive or negative.

We may thus describe a simple cubic crystal by Method I by saying that it is cubic and that there is an atom at the origin (0, 0, 0). The entire crystal is then built up by unit translations; we should also specify the length a_0 of these translations.

The drawing of the unit cell also indicates one atom per unit cell; each of the eight atoms at the corners is shared by eight unit cells so we have $\frac{8}{8}$ or one atom per unit cell.

In describing crystals, it is frequently desirable to give the number of nearest neighboring atoms or *coordination number.* For a simple cubic crystal, this number is six; that is, each atom is surrounded by six atoms all of which are the same distance away from it. Referring to Figure 3, consider atom A. There are three nearest neighbors shown in the figure – atoms O, B and D. The other three nearest neighbors are in adjacent unit cells and hence are not shown in the figure. The student should be able to visualize the positions of these other three atoms; it is suggested that he locate them and then check their coordinates. Their coordinates are (2, 0, 0), (1, 0, -1) and (1, -1, 0).

The simple cubic structure is useful as a reference point for visualizing other types of crystals and is an easy starting point for discussion of actual crystal structures. Only one substance (polonium) has so far been found to crystallize according to the simple cubic arrangement shown in Figure 3. This great scarcity of examples is very likely due to the fact that there is a large amount of empty space in this arrangement. Stated otherwise, it is an inefficient method of packing spheres. If the spheres in Figure 3 are enlarged in size so that they just touch, a simple calculation will show that only slightly more than half of the unit cell is actually occupied by the spheres.

Body-centered cubic

A body-centered cubic unit cell is shown in Figure 4. It consists of atoms of a given element at the corners of a cube plus an additional atom of the same element at the center of the cube.

Description by Method I involves stating that the crystal is cubic and that identical atoms are located at (0, 0, 0) and ($\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{2}$). Unit translations of these two atoms generate not only those shown in Figure 4 but also generate atoms at all points in space whose coordinates are given by three integral or by three half-integral numbers. (For example, there are atoms at (2, 10, -17), at ($3\frac{1}{2}$, $-5\frac{1}{2}$, $12\frac{1}{2}$), etc.) Since only two atoms are required to generate all other positions, there must be two atoms per unit cell. For a complete quantitative description, the length of the cell edge a_0 must also be given.

The model of the unit cell (Figure 4) shows nine atoms. The atom in the center of the cell belongs entirely to the unit cell but each of the atoms at the corners is shared by eight other cells; this also shows that there are $1+\frac{8}{8}$ or two atoms per unit cell.

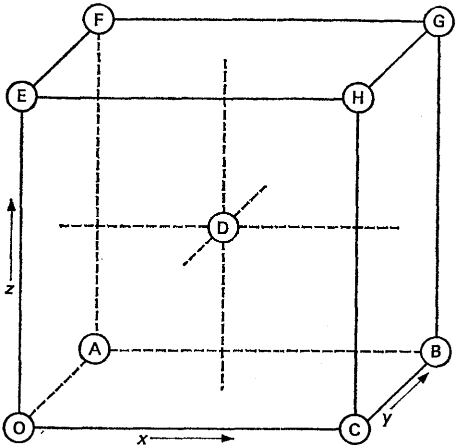


Figure 4 The unit cell for a body-centered cubic crystal. There are atoms of the same element at the corners and the center of the cube. These atoms have been drawn quite small; a more realistic picture is obtained by imagining each of them as being enlarged by an equal amount until the center atom touches the eight corner atoms. Unit translations of the atoms at (0, 0, 0) and ($\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{2}$) will generate the remaining atoms in the figure

The coordination number for this structure is eight. This is most readily visualized by considering the atom (designated D) at the center of the unit cell in Figure 4. The eight nearest neighbors, all at the same distance away, are the eight corner atoms. Now, what is the distance between the center atom and any corner atom? This may be calculated as follows: The distance from atom O to atom C is a_0; this is also the distance from atom O to atom A. By Pythagoras' theorem, the distance from atom C to atom A is $(a_0^2+a_0^2)^{\frac{1}{2}} = a_0\sqrt{2}$. Thus the diagonal on any face of a cubic unit cell (face diagonal or d_f) is $a_0\sqrt{2}$. Now consider the triangle formed by atoms A, C and H. The distance from atom A to atom H, again by the Pythagorean theorem, is $[(a_0\sqrt{2})^2+a_0^2]^{\frac{1}{2}}$ or $a_0\sqrt{3}$. This is the space diagonal or d_s. The distance from atom D to atom H (or from atom D to atom A) is just one half of this or $\frac{1}{2}a_0\sqrt{3}$. Thus the shortest distance between the centers of the atoms in a body-centered cubic structure is $\frac{1}{2}a_0\sqrt{3}$ or $0{\cdot}866a_0$.

It is important to realize that all atoms in this structure have identical surroundings. For example, atom C in Figure 4 also is surrounded by eight atoms at a distance of $0{\cdot}866a_0$ away. The student should visualize the positions of all these atoms by looking at Figure 4 and imagining the adjacent unit cells.

A number of elements crystallize in the body-centered cubic structure; these include sodium, potassium, vanadium, α-chromium, γ-chromium, α-tungsten, α-iron and δ-iron. The cell edges are known quite accurately for most of these; for tungsten, $a_0 = 3{\cdot}16469$ Å at 25°C and for sodium, $a_0 = 4{\cdot}2906$ Å at 25°C. Measurements of these cell edges at different temperatures by X-ray diffraction provide a convenient method of determining the coefficients of expansion of metal crystals.

The density of sodium is much less than the density of tungsten since the atoms of the former are lighter and are packed into a larger volume. The density of a crystal of known structure may be calculated as follows: With two atoms per unit cell, the mass per unit cell for sodium is[1]

$$\frac{2\times 22{\cdot}99}{6{\cdot}023\times 10^{23}}\ \text{g}$$

and the volume of a unit cell is $(4{\cdot}291\times 10^{-8})^3$ cm³. Thus the density is

$$\frac{2\times 22{\cdot}99}{6{\cdot}023\times 10^{23}\times(4{\cdot}291\times 10^{-8})^3} = 0{\cdot}9662\ \text{g cm}^{-3}.$$

A similar calculation may be made for tungsten for comparison with sodium.

[1] Atomic weight of sodium = 22·99, Avogadro constant = $6{\cdot}023\times 10^{23}$.

Face-centered cubic

The unit cell of the face-centered cubic arrangement is shown in Figure 5.

Description by Method I states that the crystal is cubic and lists the coordinates of four atoms of the same element, specifically those designated by O, C, A and D. The student should verify that these atoms have coordinates of (0, 0, 0), (0, $\frac{1}{2}$, $\frac{1}{2}$,), ($\frac{1}{2}$, $\frac{1}{2}$, 0) and ($\frac{1}{2}$, 0, $\frac{1}{2}$), respectively. (Again, remember that these coordinates are given in terms of a_0 which is the unit cell edge or distance from atom O to atom B.). It is important, also, to see that these four atoms, O, C, A and D, are sufficient to generate the entire structure by unit translations; such translations applied to atom O will produce all other atoms at the corners while atom C will generate atom E, A will generate H, and D will generate F. As before, continued unit translations produce atoms outside the cell and our model is visualized as extending essentially to infinity in all directions.

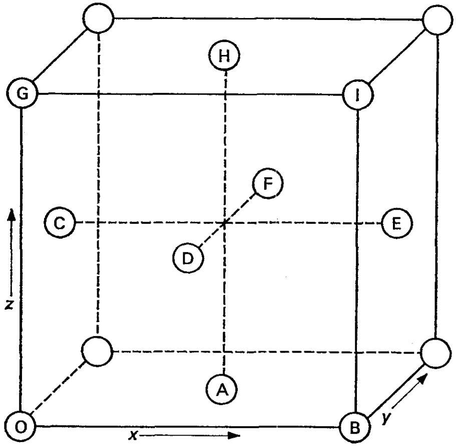


Figure 5 The face-centered cubic unit cell. There are atoms of the same element at the corners of the cube and at the center of each of the six faces of the cube. To obtain a more realistic picture, imagine each atom as being equally enlarged until the atoms 'touch' along a face diagonal; for example, atoms O and D would touch atom I. The coordinates of atoms O, A, C and D are sufficient to describe the unit cell

By reference to Figure 5, the student should convince himself that there are four atoms per unit cell even though the model shown contains fourteen atoms. As before, each of the eight corner atoms is shared by eight unit cells so that the corner atoms contribute one atom to the cell. Each of the six atoms at the centers of the faces is shared by two unit cells so that these atoms contribute $\frac{6}{2}$ or three atoms per unit cell. The total is thus $\frac{8}{8}$ (corners) plus $\frac{6}{2}$ (faces) giving four atoms per unit cell.

The coordination number for a face-centered cubic structure is 12. This is most readily seen by reference to Figure 6. This figure also shows a face-centered cubic unit cell but the origin has been moved by $\frac{1}{2}a_0$ from its position in Figure 5. The twelve atoms A, B, C, D, E, F, G, H, J, K, L and M are all the same distance from atom O; thus the coordination number is 12. The distance from atom O to any of the nearest neighbors is $\frac{1}{2}a_0\sqrt{2}$.

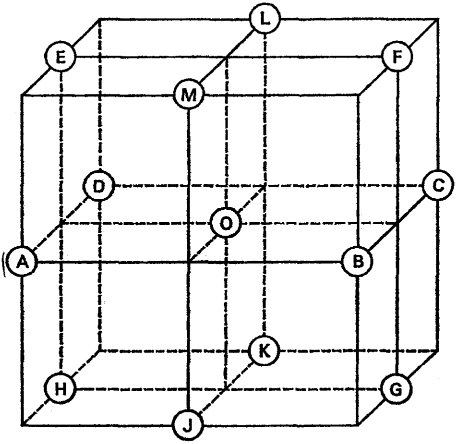


Figure 6 Alternative representation of the face-centered unit cell

It is also possible to visualize the coordination number of 12 by reference to Figure 5. The student is urged to do this, even though it is somewhat more difficult than in Figure 6. In Figure 5, atoms O, G, I and B are all the same distance from atom D; this distance is $\frac{1}{2}a_0\sqrt{2}$. Also at the same distance are atoms C, H, E and A. Finally, there are four atoms in an adjacent unit cell which are also at the same distance from atom D. (These last four atoms may be generated by unit translations of atoms C, H, E and A in the minus y direction.)

Many metals crystallize in a face-centered cubic arrangement. Examples are silver, gold, aluminium, calcium, β-nickel, copper, γ-iron and lead. At very low temperatures, neon, argon, krypton and xenon form such crystals. The dimensions of the unit cells are frequently known with high accuracy; for silver, $a_0 = 4{\cdot}0862$ Å; for gold, $a_0 = 4{\cdot}07864$ Å; for copper, $a_0 = 3{\cdot}61496$ Å; and for calcium, $a_0 = 5{\cdot}576$ Å, all at 25°C. If we regard atoms as rigid spheres, we may calculate the radius (or diameter) of any of these atoms. Thus, from the center of atom O to the center of atom I in Figure 5, the distance is $a_0\sqrt{2}$ and this is equal to the sum of four atomic radii. The student should verify that the radius of a calcium atom is thus given as 1·97 Å.

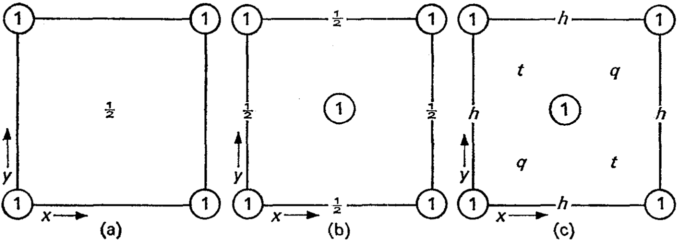


Figure 7 (a) Projection of the body-centered cubic unit cell on to the xy plane. The x- and y-coordinates of each atom are indicated in the conventional way. The z-coordinate of each atom is indicated by a number. Thus ○ means a z-coordinate of zero and 1 means a z-coordinate of 1; ① then indicates an atom in the plane of the paper (xy plane) *and* an atom above the plane of the paper by one unit of a_0. The designation $\frac{1}{2}$ indicates an atom above the plane of the paper by $\frac{1}{2}a_0$.
(b) Projection of the face-centered cubic unit cell on to the xy plane.
(c) Projection of the unit cell of diamond. The symbol ① has the same meaning as in (a) while q, h and t are used for z-coordinates of $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$, respectively

There is still another method of representing unit cells such as those shown in Figure 4 and Figure 5. This involves the projection of the atoms on to one plane and is thus convenient for representation on paper. This scheme is used in Figure 7(a) for a body-centered and in Figure 7(b) for a face-centered unit cell (in the cubic system) and is explained in the caption for Figure 7(a).

Diamond

The arrangement of the carbon atoms in diamond may be described as a face-centered cubic arrangement with additional carbon atoms halfway from the center of the cube to alternate corners. A projection of this structure on to the plane of the paper is shown in Figure 7(c). It should be clear that there are eight atoms per unit cell in this structure and that the coordination number of each carbon atom is four. Silicon, germanium and grey tin also crystallize with this structure.

Sodium chloride

Sodium chloride is our first example of the crystal structure of a compound and we now must concern ourselves with two different kinds of atoms. In the case of most salts, the crystals consist of ions rather than atoms.

The unit cell for sodium chloride is shown in Figure 8. The structure consists of two *interpenetrating face-centered cubic structures*. The student should convince himself that both the sodium ions and the chloride ions form face-centered cubic arrangements.

If we choose one of the chloride ions as the arbitrary origin, the coordinates of the ions in the unit cell are as follows:

Cl^- $(0, 0, 0)$, $(\frac{1}{2}, \frac{1}{2}, 0)$, $(\frac{1}{2}, 0, \frac{1}{2})$, $(0, \frac{1}{2}, \frac{1}{2})$
(Note that these are the same as in Figure 5.)
Na^+ $(\frac{1}{2}, 0, 0)$, $(0, \frac{1}{2}, 0)$, $(0, 0, \frac{1}{2})$, $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$

The origin of the lattice containing the sodium ions is displaced from the origin of the lattice containing the chloride ions by one half of the unit cell edge. (The coordinates of the sodium ions may be obtained by adding $\frac{1}{2}$ to the x-coordinate of each chloride ion; this gives $(\frac{1}{2}, 0, 0)$, $(1, \frac{1}{2}, 0)$, $(1, 0, \frac{1}{2})$ and $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. But note that an atom at $(1, \frac{1}{2}, 0)$ is, by the unit translation principle, equivalent to one at $(0, \frac{1}{2}, 0)$. In giving coordinates in the unit cell, it is customary to use numbers as small as possible; thus we use $(0, \frac{1}{2}, 0)$ rather than $(1, \frac{1}{2}, 0)$ etc.)

In sodium chloride, each sodium ion is surrounded by six equidistant chloride ions and each chloride ion is surrounded by six equidistant sodium ions. Thus the coordination number of each ion is six. If no distinction is made between the sodium ion and the chloride ion, the structure would be described as simple cubic. Since these two ions are different, such a description is incorrect but the coordination number is the same as for a simple cubic structure.

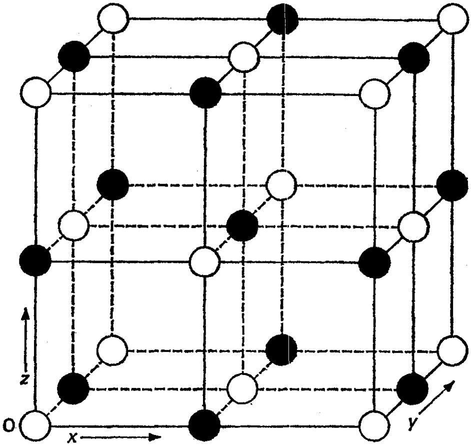


Figure 8 The unit cell for sodium chloride. Dark circles may represent sodium ions and light circles chloride ions or vice versa

There are four sodium ions and four chloride ions belonging to each unit cell. Sometimes it is stated that there are four 'molecules' of sodium chloride per unit cell but this designation is misleading since there are no molecules present. The formula NaCl indicates that sodium and chlorine must be present in a one-to-one ratio and this means that there must be the same number of ions of each element in a unit cell.

Zinc sulfide (*sphalerite or zinc blende*)

The structure of this compound may be visualized by reference to Figure 7(c). If the atoms designated *t* and *q* are sulfur atoms and if the remaining atoms are zinc atoms, then the structure is that of zinc blende. (We could have chosen the designation the other way around; that is, the atoms designated *t* and *q* could have represented zinc atoms with the remainder as sulfur atoms.) Note the close relationship between the diamond structure and the zinc blende structure.

The student should write out the coordinates for the atoms in this structure and also convince himself that the atoms designated *t* and *q* form a face-centered cubic lattice. Thus the zinc blende structure consists of two interpenetrating face-centered cubic structures. The zinc blende structure differs from the sodium chloride structure in the displacement of the origin of one of these structures with respect to the other. In zinc blende, one origin is displaced from the other by one quarter of the space diagonal. This can be visualized by adding a few atoms to Figure 7(c).

The coordination numbers of both zinc and sulfur are the same as for carbon in diamond.

Cesium chloride

The cesium chloride structure may be easily visualized by reference to Figure 4. If the center D represents a chloride ion and the remaining letters represent cesium ions, then we have the cesium chloride structure. (Again, we could have chosen the opposite designation.) In this case we have *two interpenetrating simple cubic arrangements* with the origin of one structure displaced by one half of the space diagonal from the origin of the other structure. The student should verify that the coordination number is eight for both ions and that the closest distance of approach of the ions is $\frac{1}{2}a_0\sqrt{3}$.

Some texts refer to cesium chloride as a body-centered structure. Such a description fails to recognize that there are two different elements present. Further, if body-centered were the correct description for cesium chloride, then in order to be consistent, the sodium chloride structure would be called simple cubic.

Each of the above structures for compounds is actually a prototype and reference is frequently made to the sodium chloride structure, the zinc sulfide (zinc blende) structure and the cesium chloride structure. Most of the alkali halides have the sodium chloride type structure but a few have the cesium chloride structure. Compounds having the zinc sulfide structure are gallium arsenide, zinc selenide and copper(I) bromide.

Copper(I) oxide

The unit cell for copper(I) oxide is shown in Figure 9. It will be noted that the copper atoms (represented by light circles) form a face-centered cubic structure. Not so obvious is the fact that the oxygen atoms form a body-centered structure, but this may be seen by visualizing additional oxygen atoms in adjacent unit cells. The coordinates of the copper atoms are the same as were those of the chloride ions in sodium chloride whereas the oxygen atoms have coordinates of ($\frac{1}{4}$, $\frac{1}{4}$, $\frac{1}{4}$) and ($\frac{3}{4}$, $\frac{3}{4}$, $\frac{3}{4}$). The coordination number is four for the oxygen atoms and two for the copper atoms.

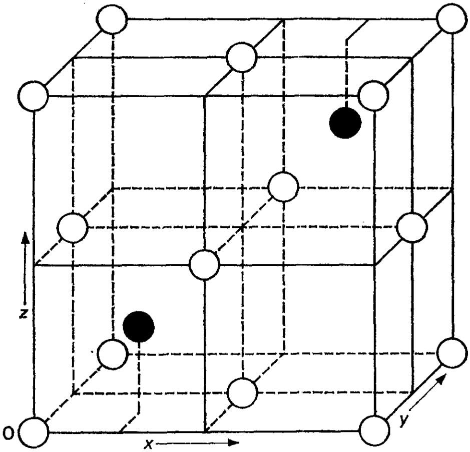


Figure 9 The unit cell for copper(I) oxide. The light circles represent copper and the dark circles represent oxygen

An alternative representation of the unit cell for copper(I) oxide is shown in Figure 10, where the light circles again represent copper. It is easy to see that the dark circles (oxygen) form a body-centered structure. The copper atoms have coordinates of ($\frac{1}{4}$, $\frac{1}{4}$, $\frac{1}{4}$), ($\frac{3}{4}$, $\frac{3}{4}$, $\frac{1}{4}$), ($\frac{1}{4}$, $\frac{3}{4}$, $\frac{3}{4}$) and ($\frac{3}{4}$, $\frac{1}{4}$, $\frac{3}{4}$), and thus have the same coordinates as the carbon atoms which are completely within the diamond unit cell; such an arrangement is face-centered cubic. Figure 10 may be generated from Figure 9 by displacing the origin by one quarter of the space diagonal. Thus, these two figures represent exactly the same structure even though their appearances are quite different. It should be noted that there are four atoms of copper and two atoms of oxygen per unit cell; these are in the correct ratio of two to one as they must be for the formula Cu_2O.

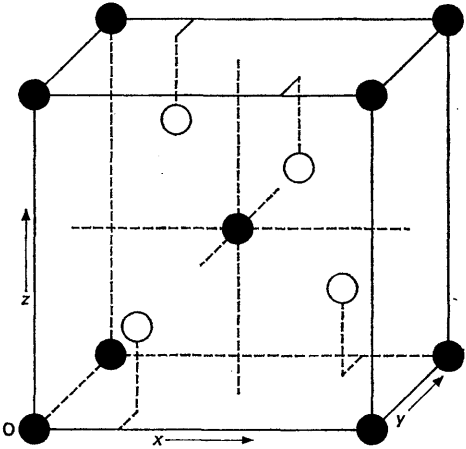


Figure 10 Alternative representation of the unit cell for copper(I) oxide. The light circles represent copper and the dark circles represent oxygen

Tetragonal Crystals

Crystals belonging to the tetragonal system are referred to a coordinate system with all angles equal to 90° and $a = b \neq c$. Usually we visualize the c-axis as running in a vertical direction and this means that we can describe a unit cell in the tetragonal system as a box with a square base and rectangular sides.

Indium

This element crystallizes in a body-centered tetragonal structure; that is, there is an atom at each corner and at the center of the box-shaped unit cell. The cell dimensions at 22°C are $a_0 = b_0 = 3{\cdot}244$ Å and $c_0 = 4{\cdot}938$ Å.

Lead monoxide (PbO)

Lead monoxide forms tetragonal crystals with the unit cell having a base of 3·975 Å on each side and a height of 5·023 Å at 27°C. Oxygen atoms are at the corners and at the centers of the top and bottom (not at the centers of the sides); this is an 'end-centered' arrangement. Lead atoms occur in a pair of opposite sides, 1·198 Å above the base and halfway between the vertical edges. Other lead atoms occur in the other pair of opposite sides, 3·825 Å above the base and halfway between the vertical edges. From this description, it should be possible for the student to construct a sketch of the unit cell in perspective or to show the projection of the unit cell onto the *xy* plane.

An alternative description of the unit cell is as follows: oxygen atoms occur at (0, 0, 0) and ($\frac{1}{2}$, $\frac{1}{2}$, 0) while lead atoms occur at (0, $\frac{1}{2}$, u) and at ($\frac{1}{2}$, 0, $\bar{u}$) where $u = 0{\cdot}2385$ and the designation $\bar{u}$ means $-u$. This indicates that the z-coordinates of two lead atoms are $0{\cdot}2385 \times 5{\cdot}023$ Å $= 1{\cdot}198$ Å. There is another pair of lead atoms whose z-coordinate is $-1{\cdot}198$ Å, and these atoms lie in an adjacent unit cell. Unit translations of these two atoms in the (positive) z-direction generate atoms on two faces of the unit cell under consideration and having z-coordinates of $1-u$ or 3·825 Å.

In calculating the distance between atoms (or ions) in a crystal, it is convenient to make use of Pythagoras' theorem in three directions. If x_1, y_1 and z_1 are the coordinates of atom 1, and x_2, y_2 and z_2 are the coordinates of atom 2, then the distance between these atoms is given by

$$d_{1-2} = \sqrt{\{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2\}}.$$

This formula holds for any crystal system in which the angles between the coordinates are 90°. In general, the coordinates in Ångströms should be used but in the cubic system calculations can be made in terms of a_0 and the result obtained multiplied by a_0 to obtain the distance in Ångströms.

The above discussion has assumed *perfect* crystals. Most actual crystals are imperfect and the various kinds of imperfections may cause interesting and important changes in the properties of crystals.

Finally, it should be pointed out that our discussion has considered only a few of the simplest types of crystals. The majority of known crystals are more complex than the ones discussed here. In particular, many substances of biological interest such as proteins, hormones, sugars, starch, penicillin and vitamins have quite complex crystal structures.

Bibliography

R. C. EVANS, *An Introduction to Crystal Chemistry*, 2nd edn, Cambridge University Press, 1964.

F. SEEL, *Atomic Structure and Chemical Bonding*, Methuen, 1963.

J. E. SPICE, *Chemical Binding and Structure*, Pergamon, 1964.

A. F. WELLS, *Structural Inorganic Chemistry*, 3rd edn, Oxford University Press, 1962.

3 G. Allen

Modern Methods for the Determination of Molecular Structure

G. Allen, 'Modern methods for the determination of molecular structure', *Chemistry Student*, vol. 2 (1968), no. 1, pp. 5–12.

The methods used by chemists to determine the structures of molecules have their origins in the pioneering work in atomic physics carried out at the turn of this century. New experimental techniques were introduced and soon they were used to study simple molecules as well as atoms. This was a most important event. It led to collaboration between physicists and chemists which resulted in the current unified views on molecular structure. Hitherto the chemist had based his ideas of molecular structure mainly on indirect evidence from the chemical composition and reactivity of substances. The physicist, too, had been concerned with the properties of matter in bulk but he had made useful progress by treating molecules as hard, featureless spheres. In the last two decades the new experimental techniques of the early twentieth century have become routine operations which are no longer the prerogative of the physicist. Today the chemist concerns himself more and more with these physical methods to determine molecular structures of ever-increasing complexity.

The physical methods available are necessarily indirect. Even large molecules, such as proteins having molecular weights of about one million, are much too small to be seen by the eye or any optical instrument because molecules are much smaller than the wavelength of visible light. The electron microscope seeks to overcome this fundamental difficulty by using a beam of electrons which has a range of wavelengths of the same order of magnitude as the distances between atoms in a molecule, that is, about 1 Å. Unfortunately it has no direct application to the determination of molecular structure. Even with the most modern instrument only the outlines of very large molecules can be seen.

The methods actually used fall into two main classes:

1. *Diffraction methods* in which molecular structures are deduced from the patterns observed when X-ray, electron or neutron beams having wavelengths of about 1 Å are scattered by the sample.
2. *Spectroscopic methods* which measure the discrete wavelengths at which the sample emits or absorbs electromagnetic radiation.

In practice information is obtained at two different levels. The diffraction methods are capable of giving a full structural analysis in which distances between nuclei are determined to between ±0·01 and ±0·05 Å. The methods are, however, time-consuming. About six months is required with modern equipment to solve a structure containing, say, twenty atoms. Spectroscopic techniques give a restricted amount of information about the shape of the molecule, but only for very simple molecules containing a few atoms do they give a detailed analysis. Nevertheless spectroscopic methods are widely used because they permit rapid identification of the chemical bonds or groups present in complex molecules.

There are, of course, many other physical methods of obtaining information about molecular structure. Measurement of dipole moments or optical activity are among the techniques which have contributed to our present knowledge. Such methods provide less detailed information, and must be regarded as secondary techniques in the light of advances made in diffraction and spectroscopic methods during the past decade. This article will not survey the secondary methods.

The Electromagnetic Spectrum

Both diffraction and spectroscopic methods involve the interaction of the sample being studied with electromagnetic radiation. The molecules scatter, emit or absorb the radiation, and the key to the nature of the molecular process involved lies in the energy quanta of the radiation. The quantum theory postulates that electromagnetic radiation consists of photons which travel at constant velocity $c = 3{\cdot}0 \times 10^{10}$ cm s^{-1}. The energy of each photon is related to the wavelength λ cm or the frequency ω hertz by:

$$E = h\omega = \frac{hc}{\lambda}\ \mathrm{J},$$

where $h = 6{\cdot}62 \times 10^{-34}$ J s is Planck's constant. If we now consider a situation in which each molecule scatters, emits or absorbs one photon then one mole of the substance interacts with radiation energy quanta amounting to:

$$E = \frac{Nhc}{J\lambda} = \frac{2{\cdot}86}{\lambda}\ \mathrm{cal\ mol^{-1}},$$

where N is the Avogadro constant and $J = 4{\cdot}18$ joules per calorie. The emission or absorption of electromagnetic radiation which occurs in spectroscopic methods of structure determination is accompanied by a change in the total energy of one molecule equal to the energy of the photon. In accordance with the quantum theory absorption or emission of radiation only occurs at specific wavelengths or frequencies corresponding to the separation between pairs of energy levels of the molecule as shown schematically in Figure 1. In Table 1 the electromagnetic spectrum is set out in terms of the units used to measure frequency and wavelength in different regions. Energy quanta and the molecular processes excited by the emission or absorption of radiation are also listed. It will be seen that long wavelength radiation (such as microwaves or radio waves) excites transitions involving very small energy changes and is unlikely to disrupt chemical

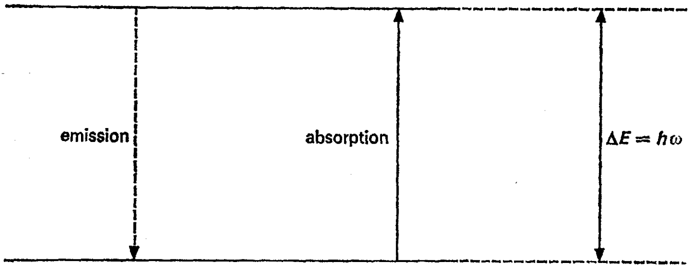


Figure 1 Transitions between two energy levels leading to emission or absorption of electromagnetic radiation

Table 1

The Electromagnetic Spectrum

Wavelength	1 Å	1000 Å	1 μm	10 μm	100 μm	1 mm	1 cm	10 cm	1 m	10 m	100 m	
Frequency	3×10^{18}		3×10^{14}		3×10^{12}		3×10^{10}		3×10^{8}		3×10^{6}	Hz
							30000		300	30	3	MHz
			10000	1000	100	10						cm^{-1}
	X-rays neutrons electrons	u.v.visible		infrared			microwave			radio wave		
Energy quantum	286000	286		2·86			$2{\cdot}86 \times 10^{-3}$			$2{\cdot}86 \times 10^{-6}$		kcal mol^{-1}
		12·4		$1{\cdot}24 \times 10^{-1}$			$1{\cdot}24 \times 10^{-4}$			$1{\cdot}24 \times 10^{-7}$		eV
Molecular phenomena	diffraction	electronic transitions		vibrational transitions			rotational transitions electron spin transitions			nuclear spin transitions		

bonds within the molecules as bond strengths are of the order of magnitude of 50–100 kcal mol^{-1}. Ultraviolet and X-rays, on the other hand, can excite energy changes much larger in magnitude than bond strengths and consequently they must be used with extreme care when molecular structures are investigated.

As we have already noted, the radiations employed in diffraction techniques have wavelengths of the order of 1 Å. Consequently very large energy quanta are involved. Fortunately in these techniques the radiation is scattered elastically by the atoms and so the energy of the molecule remains unchanged after the event. However, there is always a danger that the radiation will be absorbed in secondary processes not related to the technique for structure determination, which would lead to decomposition of the sample. Many readers will be aware that this danger accounts for the precautions taken in X-ray radiography in hospitals to shield the operator and the patient from the biochemical effects of X-radiation.

At this point we can summarize by noting that modern methods of structure determination can be divided into two main groups:

(i) spectroscopic methods which utilize electromagnetic radiation of longer wavelengths,
(ii) diffraction methods which employ short wavelength X-ray, electron or neutron beams.

Now we can review the individual methods and assess the information which they yield.

Spectroscopic Techniques

The origin of *atomic spectra* is well understood in terms of transitions between different energy levels. These levels correspond to different electronic structures of the atom. The spectra lie mainly in the ultraviolet and visible regions. Figure 2 shows part of the spectrum of the hydrogen atom and the transitions which give rise to the spectral lines.

To understand *molecular spectra* we must explore the energy levels available to a molecule. Let us begin with a simple molecule. The energy levels of a diatomic molecule such as H_2, HCl or CO

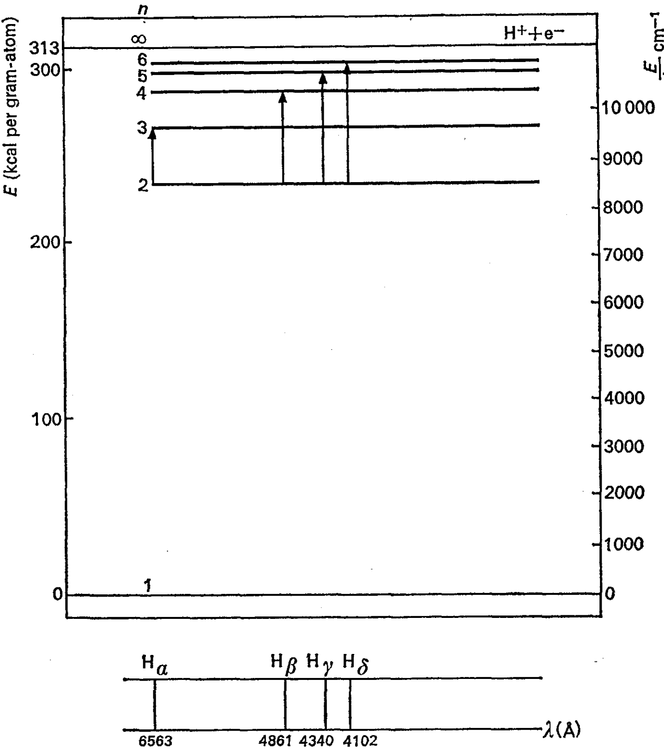


Figure 2 Energy levels and Balmer series of spectral lines in the hydrogen atom

are shown in Figure 3. In the H atom there is one main set of levels corresponding to the electron occupying the 1*s*, or 2*s*, or 2*p*, or 3*s*, . . ., and so on orbital. Molecules have three principal sets of levels which give rise, in turn, to three types of molecular spectroscopy.

1. *Electronic spectroscopy*

The widely spaced set of levels is comparable in spacing with the levels available to the hydrogen atom. They arise from the electronic structures that the molecule adopts when electrons are

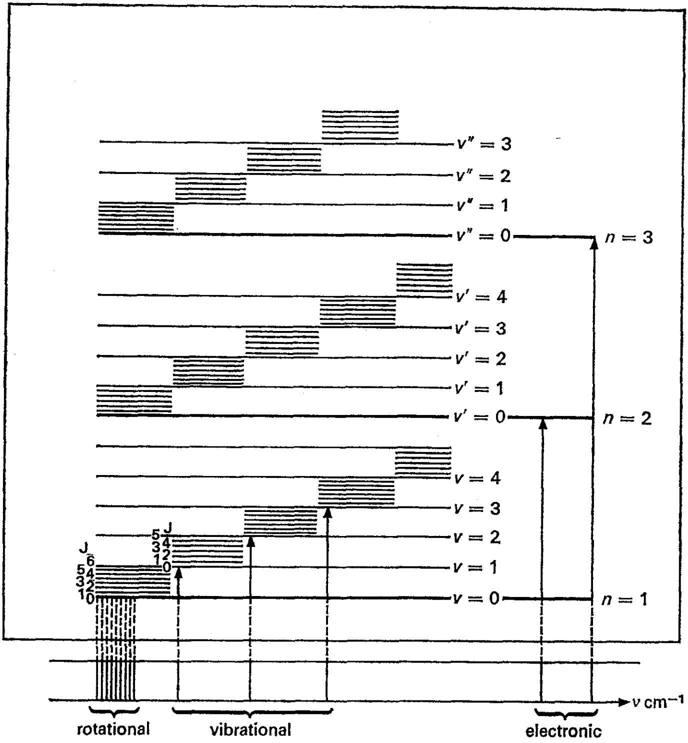


Figure 3 Energy levels of a typical diatomic molecule. J = quantum number of rotational energy levels, v = quantum number of vibrational energy levels, n = quantum number of electronic energy levels

promoted from the most stable electronic configuration into orbitals of higher energy. Transitions between these levels give spectral bands (rather than sharp lines) in the ultraviolet or visible regions. The spectrum is a characteristic feature of the molecule and can be used as a means of identification.

The spectra of polyatomic molecules are more complex but the positions of the bands can be correlated with the types of bonds present in the structure. Molecules which have conjugated bonds or double or triple bonds have characteristic absorption bands in the visible and near ultraviolet regions (λ = 2000–6000 Å), for example aromatic ring, conjugated ethylenic and carbonyl systems. On the other hand transitions associated with electrons in single bonds such as C—C, C—H and C—O lie at wavelengths less than 2200 Å. Thus the ultraviolet spectrum can be used as a 'finger-print' method for identifying the presence of specific groups. In addition inorganic chemists also make use of ultraviolet spectra to identify electron transitions of transition metal atoms in different chemical environments. In all these applications the sample is usually studied in the liquid or solid state, and with modern spectrometers the spectrum is obtained in a matter of minutes.

2. *Vibrational spectroscopy*

Associated with each electronic energy state of the molecule is a set of less widely spaced levels. These are the vibrational energy levels of the molecule.

A diatomic molecule behaves as though the two atoms were connected by a spring. Levels of higher energy correspond to increased amplitudes of vibration. Vibrational spectra lie in the infrared region, usually between 4000 and 200 cm^{-1}, and arise from transitions between the vibrational energy levels. However, only diatomic molecules which have permanent dipole moments have infrared absorption and emission spectra, because the molecules must have electric moments which undergo periodic changes in magnitude at the appropriate frequency if they are to emit or interact with electromagnetic radiation. Thus HCl, HI, CO, NO, IBr have infrared vibrational spectra whereas H_2, O_2, N_2, Cl_2, I_2 have not. The spectrum not only identifies the molecule but its existence shows that the molecular structure is unsymmetrical, that is, it contains two different atoms.

Raman spectroscopy provides an indirect way of observing the vibrational spectra of *all* diatomic molecules. In this form of molecular spectroscopy a parallel beam of monochromatic visible light of frequency ν_i is used to irradiate the sample. The spectrum of the light scattered at right angles to the incident beam consists mainly of radiation of unchanged frequency ν_i but there are also weak lines at $\nu_i \pm \nu_{\mathrm{vib}}$, where ν_{vib} is the frequency of the spectral line corresponding to the separation between vibrational energy levels. These weak lines arise because some of the photons gain or lose energy in the scattering process due to the molecules in the sample simultaneously undergoing vibrational transitions. Therefore, in Raman spectra, the frequency shifts from the central line correspond to the vibrational spectra of the molecules.

Vibrational spectra can be used to determine the symmetry of the structure of a diatomic molecule simply because symmetrical molecules such as H—H, O—O and N—N only have vibrational Raman spectra. Unsymmetrical molecules have both Raman and infrared vibrational spectra.

In polyatomic molecules the complexity of the vibrational spectrum increases as the number of atoms in the molecule increases. The geometry of the molecule is also important. If the molecule is symmetrical some vibrations do not produce a dipole moment change within the molecule. Just as for diatomic molecular vibrations, these motions only give rise to Raman vibrational spectral lines. Only those vibrational motions which do involve changes in dipole moment give infrared spectral lines. Thus if a molecule has different vibrational frequencies in the Raman spectrum from those observed in the infrared spectrum we can deduce the geometrical symmetry of the molecule. The use of vibrational spectra to deduce the geometrical shape of molecules is particularly important in inorganic chemistry.

Organic and inorganic chemists make use of vibrational spectra in a simpler way since the positions of the spectral lines correlate with molecular structure. Specific bonds and chemical groups have characteristic vibrational infrared or Raman frequencies, as shown in Table 2.

Table 2

Bond	*Approximate frequency* (cm^{-1})
O—H	3600
N—H	3400
C—H (aromatic)	3050–3100
C—H (aliphatic)	2800–3000
C≡N	2250
C=O	1700–1750
C—Cl	700–750

Infrared spectra are most commonly used for this purpose since they can be recorded for a few milligrams of sample. They act as 'finger-prints' for the molecular structure.

3. *Rotational spectroscopy*

The most closely spaced sets of energy levels in Figure 3 are associated with the rotational motion of the molecule in each of its vibrational states. The spacing of the rotational levels and hence the position of the spectral lines is determined by the moment of inertia of the molecule. The spectra lie in the far infrared and microwave regions. In practice moments of inertia can be calculated with great accuracy from the rotational spectra and hence bond lengths of diatomic molecules are also obtained since

$$I = \frac{m_1 m_2}{m_1 + m_2} r^2,$$

where I is the moment of inertia, r the bond length and m_1 and m_2 the atomic masses. As for vibrational spectra, only diatomic molecules which have permanent dipole moments have infrared and microwave rotational spectra, but *all* diatomic molecules have rotational Raman spectra. However, the precision with which bond lengths can be determined from rotational Raman spectra is not as high as that attained from microwave spectra.

In polyatomic molecules the rotational spectra are complicated by the fact that the molecules may be non-linear and thus have two or three different moments of inertia. Nevertheless analysis of microwave rotational spectra, in favourable cases, is still the most precise method available for determining moments of inertia and bond lengths.

The physical state of the sample is important in rotational spectroscopy. Rotational lines are only observed in the spectra of gases. Solids and liquids have broad, continuous spectra which do not give the required information for estimating moments of inertia and bond lengths.

Magnetic Resonance Spectra

So far we have considered the emission or absorption of radiation by samples which are simply placed in front of the spectrometer entrance slit. The resulting spectra lie in the ultraviolet, visible, infrared and microwave regions. If these samples are now placed in a uniform magnetic field of 3000–20000 oersted, two new series of spectra can be observed. They lie in the microwave and radio wave regions of the electromagnetic spectrum; the exact location of the spectral lines depends on the strength of the magnetic field.

Since new lines are observed, the sample has acquired new sets of energy levels when placed in the magnetic field. These levels arise from the fact that certain nuclei and all unpaired electrons behave like tiny bar magnets in the magnetic field. In chemical bonding we are familiar with the concept that two electrons can only share the same energy level (that is orbital) if they have spins of opposite sign. Let us consider the simplest possible case – an atom such as H, Na or K in which a single, unpaired electron is moving in the region of the nucleus. The statement that an electron has 'spin' means that it has spin angular momentum about an axis. In classical terms this means that the electron is behaving as though it were spinning about an axis. This motion is equivalent to electricity circulating around a loop of wire, consequently it generates a magnetic moment along the axis of spin as shown schematically in Figure 4.

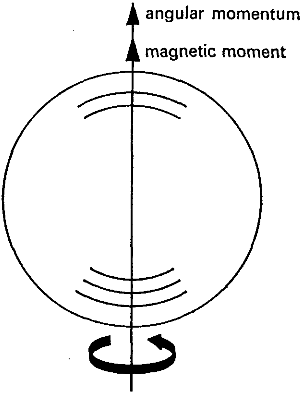


Figure 4 Magnetic moment generated by classical spinning charged particle

Therefore if we place an atom or molecule which has a free electron in a magnetic field the magnetic moment will, like a bar magnet, have one position in which it has lower energy (equivalent to opposite poles attracting ↓↑) and another of high energy (equivalent to an arrangement in which like poles repel ↑↑). Thus each energy level of the electron splits into two levels according to the orientation of the magnetic moment associated with electron spin (Figure 5).

The spin is characterized by a spin quantum number $s=\frac{1}{2}$ which, in the direction of the magnetic field, takes the values either $+\frac{1}{2}$ or $-\frac{1}{2}$. The new spectrum which is observed in the presence of a magnetic field corresponds to the absorption of radiation of energy ΔE by electrons having $s = -\frac{1}{2}$. The spin of the electron is thus changed to $+\frac{1}{2}$ as it undergoes a transition to the upper state. The spacing ΔE and hence the frequency depend on the magnetic field strength H_0 (Figure 6). For organic molecules (or radicals) with unpaired electrons the absorption line is at $\lambda = 3$ cm in a field of 3500 oersted and 0·8 cm at 12000 oersted. Such spectra are known as electron spin resonance (e.s.r.) spectra. They give information about the location of the unpaired electron in the molecule and must be counted among the methods for structure determination. However, e.s.r. spectra are observed only for molecules which have unpaired electrons, for example radicals or molecules containing transition metal atoms.

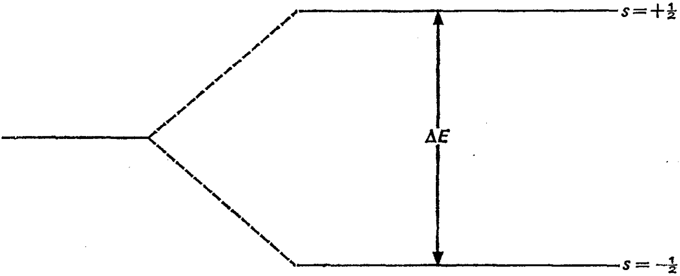


Figure 5 Splitting of electronic energy level in a magnetic field

Nuclear spin resonance is of much greater importance for the analysis of normal molecules. Nuclei such as ^{16}O, ^{12}C do not have 'spin' properties and so a magnetic field has no effect on their energy levels. ^1H, ^2H, ^{14}N, ^{15}N, ^{19}F, etc. do have nuclear spin and therefore they behave like magnets when placed in a magnetic field. However their magnetic moments are about a thousand times smaller than that of an unpaired electron and so, although nuclei have analogous magnetic spectra, the electromagnetic radiation is absorbed at correspondingly lower frequencies.

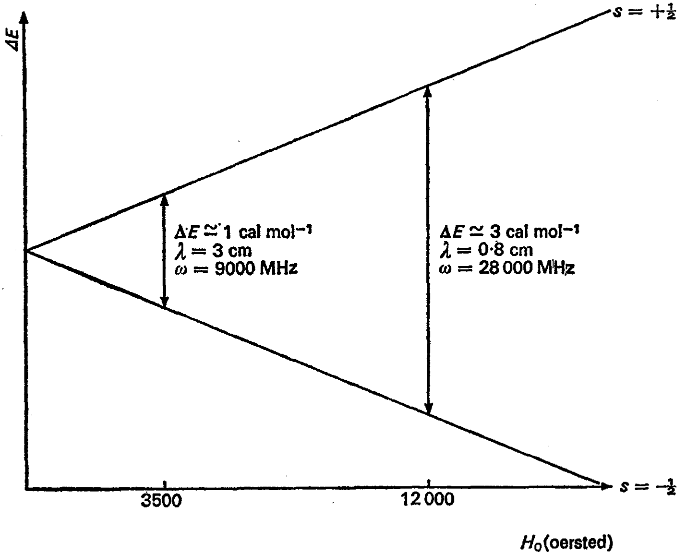


Figure 6 Origin of electron spin resonance spectra

Unlike electron spins, nuclear spins are not coupled and all molecules containing nuclei which have spin also have nuclear spin magnetic resonance (n.m.r.) spectra. The frequency at which the absorption spectrum is observed depends on the strength of the magnetic field and the nature of the nucleus as shown in Table 3.

The use of proton n.m.r. spectra is widespread in organic and inorganic chemistry. Applications depend on the fact that the spectral lines in liquids and solutions are very sharp. Thus it is found that protons in different chemical environments resonate at slightly different frequencies. From these very small shifts

(1 to 1000 Hz at 100 MHz) different chemical groups in a molecule can be identified and the relative number of protons in each group estimated. Also the number of protons in the groups *adjacent* to a given chemical group can be deduced from the structure of the spectrum of the protons in the group.

Table 3

	H_0 *(oersted)*	*Frequency* (MHz)
^1H	14 100	60
	23 500	100
^2H	23 500	14
^{19}F	23 500	94

Because of the detailed information which can be obtained on complex molecules, and because only a few milligrams of sample are required, proton n.m.r. spectroscopy is now one of the most widely used methods. Other nuclei with magnetic properties, for example ^{19}F, ^{31}P, ^{15}N, provide structural information of a similar nature.

The n.m.r. spectra of solids, too, are sometimes used in structure determinations since the distance between nuclei can be calculated from the shape of the very broad absorption lines which are observed for solid samples. In practice, only relatively simple structures can be solved, for example in $CaSO_4.H_2O$ crystals the H . . . . . H distance within the water molecule can be determined.

Mössbauer and nuclear quadrupole resonance spectroscopy are two other forms of absorption spectroscopy which involve transitions between energy levels generated by the magnetic properties of nuclei. These techniques also provide information concerning molecular structure, but their applications are limited and they are not likely to be as widely used as the other spectroscopic methods.

Mass Spectroscopy

When molecules are ionized by an electric discharge they tend to lose an electron, a proton or larger fragments. Conventional mass spectrometers separate ionized species according to their mass and charge. This form of spectroscopy is, of course, quite different from the other methods discussed above but in the past decade it has become one of the most useful techniques for analysing molecular structures. The development has occurred because very high resolution instruments which are capable of measuring mass ratios with a precision of a few parts per million are now commercially available. This increase in precision is important because the masses of atoms are not whole numbers. Suppose a sample is found to have a mass of 44 and the exact value is obtained to an accuracy of 1 part in 10^5 then the result can distinguish between the possible combinations of atoms which make up to this mass, for example

N_2O	44·0151	C_2H_4O	44·0403
CO_2	44·0038	C_2H_6N	44·0641
CH_2NO	44·0277	C_3H_8	44·0767

The structures of large molecules can be investigated by this technique and it has the advantage of requiring only a milligram of sample in either the solid, liquid or gaseous state.

Diffraction Techniques

Diffraction is a property of all forms of wave motion. In the diffraction techniques used to determine molecular structures, a parallel, monochromatic beam of X-rays, neutrons or electrons of wavelength about 1 Å is scattered by different atoms in the molecule. A typical apparatus is shown in Figure 7. The radiation scattered by atom 1 travels a slightly different distance from the radiation scattered by atom 2, to reach the photographic film at point A. When the path difference is exactly an integral number of wavelengths the two rays combine to produce a spot on the film but if their paths differ by exactly half the wavelength interference results in darkness and hence no spot at A.

If the sample is a single crystal it is mounted with one axis perpendicular to the incident beam of radiation. During the course of the exposure the crystal is slowly rotated so that successive planes of atoms in the crystal are brought into orientations which produce a series of spots on the photographic film.

From the position and relative intensities of the spots and a knowledge of the wavelength of the radiation, the relative positions of the atoms in the crystal can be determined. If the crystal is composed of molecules, for example benzene or naphthalene, rather than atoms or ions as in sodium chloride or metal crystals, this information allows the molecular structure to be specified in terms of bond lengths and bond angles.

When the sample is a powder rather than a single crystal, the incident beam in Figure 7 falls on a mass of crystallites in random orientations. Rotation of the sample is unnecessary since some of the crystallites will be orientated at the appropriate angle to produce a spot on the film. In this case, however, the crystallites do not have a particular axis orientated perpendicularly to the incident beam; instead the axes are orientated in all directions around the beam. Hence the scattered radiation forms a circle in the plane perpendicular to the incident beam instead of the series of discrete spots obtained from a single crystal.

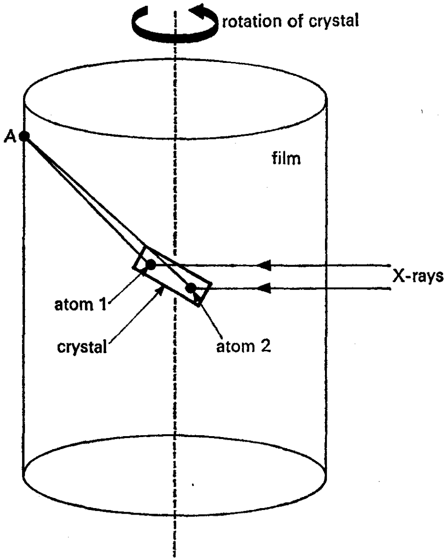


Figure 7 Schematic diagram of a diffraction apparatus

Unfortunately powder patterns are difficult to interpret but they are used quite extensively as 'finger-prints' to identify structures semi-quantitatively.

Gaseous samples also produce circular diffraction patterns but these are more readily interpreted because the only permanent three-dimensional structure present in the diffracting medium is that of the molecule. The molecules are randomly orientated with respect to each other and consequently there are no extended arrays of atoms to give additional reflections and so to complicate the diffraction pattern. As in the case of crystals, the molecular structure can be deduced from the pattern if the dimensions of the apparatus and the wavelength of the radiation are known.

X-ray Diffraction

This technique is mainly applied to the analysis of crystals. X-rays are scattered primarily by the electron clouds which surround each atomic nucleus. The scattering power of each atom increases regularly with increase in atomic number and so the contributions from heavy atoms tend to dominate the diffraction patterns. Consequently the heavier atoms can be located more precisely than hydrogen atoms in a molecular structure.

The accuracy with which structures can be determined is less than that attainable in rotational spectroscopy but the method is applicable to much larger and more complex molecules provided, of course, that they can be obtained in the form of single crystals. The use of computers enables large molecules containing many atoms to be tackled and has reduced the time required to resolve a structure involving, say, 10–20 atoms from years to weeks or a few months.

Neutron Diffraction

At the present time the principal use of neutron diffraction is to obtain more detailed information on crystal structures which have been partially resolved by X-ray methods. Sufficiently powerful sources of neutrons are available only in nuclear reactors and therefore instruments are not as numerous as X-ray diffractometers. The technique is particularly valuable, however,

Table 4

Comparison of Methods

Technique	*Region*	*Cost of instrument*[1] *(£)*	*Time scale of* *(a) experimental measurement*	*(b) analysis of results*	*Form of sample*	*Routine or specialist*	*Remarks*
Spectroscopic							
Electronic	Ultraviolet/ visible	1000–10000	minutes	minutes	solid, liquid or gas	R	'Finger-print' only, widely used
Vibrational (*a*)	Infrared	2000–2500	minutes	minutes	solid, liquid or gas	R	'Finger-print' only, widely used
(*b*)	Infrared Raman (visible)	2000–10000 5000–18000	minutes/ hours	days/ weeks	solid, liquid or gas	S	Full analyses only possible on small and/or highly symmetrical molecules
Rotational	Far infrared Microwave	3000–10000 n.c.a.	minutes/ hours	days/ weeks	gas	S	Very accurate moments of inertia for simple molecules
E.S.R.	Microwave	7000–20000	minutes	hours/ days	solid, liquid or gas	R	Only for molecules with unpaired electron spin
N.M.R.	Radio wave	10000–25000	minutes	minutes/ days	liquid	R	^{1}H, ^{19}F spectra very widely used
Mass Spectrum		10000–30000	minutes/ hours	minutes/ hours	solid, liquid or gas	R	Very widely used
Diffraction							
X-ray	1 Å	5000–35000	hours/days	weeks	single crystal	S	A most powerful method, automatic instruments increase speed and usefulness
Neutron	1 Å	n.c.a.	hours/days	weeks	single crystal	S	Useful for locating protons more precisely
Electron	1 Å	n.c.a.	hours	days/ weeks	gas	S	Extends the range covered by rotational spectroscopy

1. n.c.a. = no cost available.

because neutrons are scattered by atomic nuclei and, in contrast to X-ray scattering, the scattering power of the proton is considerably greater than the scattering powers of many heavier nuclei.

Electron Diffraction

Electron diffraction competes directly with microwave and far infrared rotational spectroscopy, as it is applied mainly to the determination of the molecular structure of gases. Electron diffraction measurements do not give such precise values for bond lengths and angles, but they are capable of dealing with more complex structures. Furthermore, a permanent dipole moment is not an essential condition as it is for rotational spectroscopy in the microwave or infrared regions, and so the molecular parameters of molecules such as CCl_4, $SiCl_4$ and diphenyl can be obtained from diffraction studies. The problem of calculating the molecular structure from the diffraction pattern is less arduous than in the case of single crystals, but in common with all diffraction techniques it is practised by specialists rather than as a routine operation.

Summary

An attempt is made in Table 4 to put some of the methods for structure determination into the perspective of the preparative chemist. This is a difficult exercise because, in general, the amount of information gathered in an experiment is related to the time spent in collecting the data. The various techniques can be classified, however, either as routine operations or as ones which require specialist operators. The form of the sample required is also a relevant factor. Readers will notice that the routine methods which produce a limited amount of information fairly quickly are the most popular.

The effective scientist is the man who can (a) choose the most appropriate method to obtain an answer quickly, and (b) recognize when the investigation has reached the point of diminishing returns. Consideration should also be given to the cost of the exercise – but perhaps that is too much to expect from a chemist caught up in some compelling investigation! In any event it is clear that modern methods for the determination of molecular structure form a substantial part of the budget of an industrial or university chemistry research laboratory.

Bibliography

G. M. BARROW, *The Structure of Molecules*, Benjamin, 1963.
W. S. BREY, *Physical Methods for Determining Molecular Geometry*, Chapman & Hall, 1966.
R. K. KULLNIG, C. M. MARTINI and F. C. NACHOD, 'Nuclear magnetic resonance spectroscopy', *Chemistry*, vol. 38 (1965), no. 5, pp. 6–13, and no. 7, pp. 17–23.
J. E. PAGE, 'Infrared spectroscopy', *Penguin Science News*, 1960, no. 54, pp. 7–26.
J. E. SPICE, *Chemical Binding and Structure*, Pergamon, 1964.
P. J. WHEATLEY, *The Determination of Molecular Structure*, 2nd edn, Oxford University Press, 1968.

Part Two **Stereochemistry**

E. Sherwin in his article on 'The shapes of simple molecules' shows how the electron-pair repulsion theory can be used as a *qualitative* method for predicting molecular shapes. The more sophisticated hybridization method is also briefly discussed.

In recent years great advances have been made in our understanding of the three-dimensional structure of organic molecules, and stereochemical concepts are of considerable importance in organic chemistry. G. Baddeley indicates the type of information on which our ideas concerning stereochemistry are based and also discusses the stereochemical aspects of addition reactions to ethylenic double bonds and of elimination reactions. The importance of stereochemistry as a tool in the elucidation of organic reaction mechanisms is explained by G. Illuminati in his article 'Organic reactions' on page 273.

4 E. Sherwin

The Shapes of Simple Molecules

E. Sherwin, 'The shapes of simple molecules', *Chemistry Student*, vol. 1 (1967), no. 4, pp. 90–94.

Introduction

Before starting on a discussion about the shapes of simple molecules it might be reasonable to define our terms and to say why we are interested in them.

A simple definition of a molecule often used is: the simplest particle of a substance capable of independent existence. We then sometimes go on to talk about substances such as sodium chloride and write molecular formulae such as NaCl. In elementary courses it is not always explained that it is only in the vapour phase – not often encountered for sodium chloride in such courses

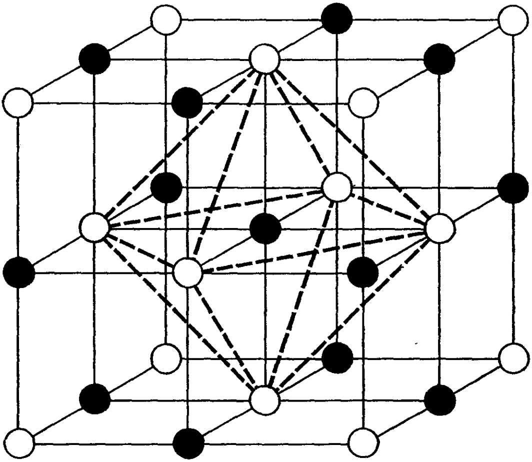


Figure 1 The unit cell of the sodium chloride crystal

– that such a thing as a simple molecule of sodium chloride exists. In the solid we have an aggregate of positive and negative ions but no particular pair of ions can be said to be a molecule (Figure 1). In solution the aggregate dissociates and the ions become surrounded by solvent molecules, that is, in water the ions are hydrated. Even covalent compounds may have molecules which are far from simple, for example we may write BN as the formula of boron nitride, but this gives no indication that both known forms of the compound are polymeric and that the molecules are extremely large. The two forms of this compound have structures similar to those of graphite and diamond, respectively. It would be hard to say exactly what a molecule is in these examples.

For the purposes of this article we shall regard simple molecules as those for which the size of the molecule is defined by the simple empirical formula or some simple multiple of it. Thus we can regard SF_6 and S_2Cl_2 as simple molecules.

Why do we Wish to Know the Shapes?

We shall see that the shape of a simple molecule is determined by the number of valence-shell electrons of the atoms concerned. In turn the physical and chemical properties are largely determined by the shape of the molecules.

One extremely important example will suffice to illustrate this point, namely water. In writing the formula for it by working to the octet rule, one is tempted to write a linear structure:

$$\mathrm{H{-}\underset{\cdot\cdot}{\overset{\cdot\cdot}{O}}{-}H}$$

The angular structure is well established and it is this that makes the molecule a dipole, that is the centres of positive and negative charge of the molecule do not coincide. This is caused in part by the different electronegativities of the two elements (such that the bonding electrons are not shared equally) and also by the two unshared pairs of electrons on the side remote from the hydrogen atoms.

The dipolar character of the molecules gives rise to a stronger attraction between them (hydrogen bonding) which results in the substance being far less volatile than one would anticipate from molecular size alone. While water is liquid at ordinary temperatures, the hydrides of other elements around oxygen in the periodic table are gases (NH_3, HF, H_2S). These strong attracting forces also result in higher surface tension and latent heats than for other liquids of similar molecular weight. The solvent action in forming hydrated ions and the hydrolytic reaction on many covalent compounds also result from the strong electron donor properties of the water molecules through their unshared pairs. In addition these directed forces due to the dipolar character result in a very highly ordered open structure in ice, such that ice has a lower density than water and there are anomalous changes in density just above the melting point.

Factors which Determine Shape

Single-bonded structures

We can now turn our attention to discussing what determines the shape of a molecule and the method by which we can deduce the shape.

The most important thing to remember is that both shared and unshared electrons in the valence shell, that is the outermost shell which is occupied, play a part in determining the distribution of bonds around an atom. It is a simple matter to add the number of valence-shell electrons of an atom to the number of shared electrons from the atoms bound on to it if we know how many valence-shell electrons the atoms have. The arrangement is determined by two opposing principles:

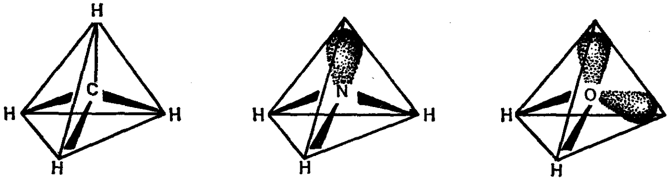
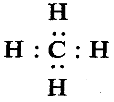


Figure 2 The arrangement of the electron pairs in methane (interbond angle 109° 28′), ammonia (106° 47′) and water (104° 27′)

1. Lewis's idea that electrons are associated in pairs still holds, the electrons in a pair being attracted to each other by something called spin when the spins are opposite.

2. Having got our electrons in pairs, we must remember another important piece of knowledge – like charges repel – which tells us that the pairs of electrons around an atom repel each other and will stay as far apart as possible.

On this basis we should expect a simple molecule like methane to be tetrahedral, that is the hydrogen atoms are at the corners of a regular tetrahedron with the carbon at the centre having the bonding pairs of electrons pointing towards the corners of the tetrahedron (Figure 2). We can arrive at this via the Lewis structure



$$\begin{array}{c} \mathrm{H} \\ \mathrm{H} : \overset{\cdot\cdot}{\underset{\cdot\cdot}{\mathrm{C}}} : \mathrm{H} \\ \mathrm{H} \end{array}$$

which will help us to see that there are four pairs of electrons around the carbon, four electrons from the carbon and one each from the hydrogen atoms. The repulsion between these four pairs results in their adopting the position in which the repulsive forces are at a minimum, and this is the tetrahedral arrangement. If the molecule were planar as is suggested by the Lewis structure, the repulsive forces would not be at a minimum since the pairs would then be closer together, that is 90° to each other, than in the tetrahedral configuration (109° 28′).

These two paragraphs contain the whole basis of working out the shapes of simple molecules. We can sum it up by saying: work out the number of pairs of electrons around an atom (from the number of valence-shell electrons and those provided by other atoms held by the bonds) and these will arrange themselves as far apart as possible.

In methane the four pairs are arranged symmetrically since all four are identical. Water and ammonia molecules also involve four pairs of electrons around the central atom and these will be expected to be tetrahedrally arranged. The shapes of molecules can be determined by various experimental techniques, for example diffraction methods (neutron or X-ray) and spectroscopic methods such as infrared. In the case of these two molecules the angle between bonding pairs is not quite the same as in methane (Figure 2), but steadily decreases as the number of bonding pairs decreases. Here we see the influence of another important principle in determining the finer details of the structures. The unshared pairs only belong to one atom and so produce a bigger electron density nearer to that atom than does a bonding pair which is shared by two atoms. The repulsive forces depend on the magnitude of the charge in addition to distance apart, and so there will be stronger repulsions between unshared pairs than between the bonding pairs. The repulsion between a bonding pair and an unshared pair of electrons will be intermediate between the two. The order of magnitude of the various repulsions will therefore be:

Unshared pair with unshared pair > unshared pair with bond pair > bond pair with bond pair.

These ideas were introduced in 1940 by Sidgwick and Powell. Perhaps because other events were uppermost in people's minds at that time their usefulness was not appreciated widely until the principles were restated by Gillespie and Nyholm in 1957. (See 'References' at the end of this article.)

These ideas of repulsions between pairs of electrons may be illustrated quite simply by the mechanical analogy of plastic modelling balloons.[1] If we focus attention on the fastening together of the balloons we find that if two balloons of equal fatness are wound together the four parts adopt a regular tetrahedral distribution. If one balloon is fat and the other thin, the fat parts push the thin parts closer together. Relating the thickness of the inflated balloon to electron density, we can see that the two halves of the fat balloon are equivalent to the unshared pairs of the oxygen atom in a water molecule. These push together the two halves of the thinner balloon which are equivalent to the bonding pairs in the water molecule. This analogy can be used to illustrate several structures and these are discussed in detail elsewhere.

The test of any useful theory is of course its successful application to the prediction of some previously unknown fact. We can best illustrate this by reference to the recently discovered noble gas compounds. When the preparation of xenon difluoride and xenon tetrafluoride was first announced it was tempting to guess

---

[1] H. R. Jones, *Educ. Chem.*, vol. 2 (1965), no. 1, pp. 25–31.

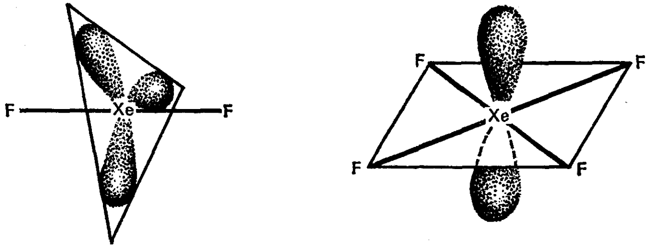


Figure 3 The arrangement of electron pairs in xenon difluoride and xenon tetrafluoride

(or predict) that their structures would be linear and square-planar respectively, as far as the distribution of the xenon–fluorine bonds is concerned (Figure 3), since these molecules contain respectively ten and twelve electrons around the xenon atoms. It was no surprise to many when these structures were confirmed by neutron and X-ray diffraction.

It should be emphasized at this stage, however, that these ideas can only be used qualitatively. In certain circumstances all of the repulsions must be taken into account to obtain the correct structure. Thus, while we can say that the planar structure of XeF_4 is due to the strongest repulsions being between unshared pairs resulting in these being as far apart as possible, on this basis alone it would be tempting to say that the chlorine trifluoride molecule will also be planar since it has two unshared

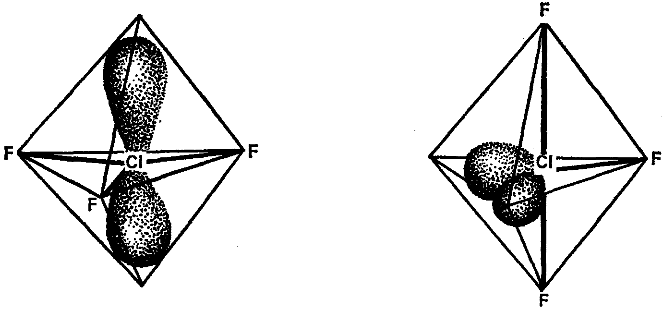


Figure 4 Two of the possible arrangements of electron pairs in a chlorine trifluoride molecule

pairs (Figure 4). A planar molecule would have the unshared pairs at 180° to each other and repulsion between these would be at a minimum. However, the structure of the ClF_3 molecule is T-shaped. It appears therefore that the increased repulsion between the unshared pairs at 120° to each other is compensated for by two of the unshared pair–bond pair angles being 120° instead of 90°.[1] In most other cases the most obvious prediction is the correct one.

Examples of the structures known for the various numbers of pairs of electrons are summarized in Table 1.

While we are discussing single-bonded structures we might mention the trend in bond angles in a series of related compounds, for example water and hydrogen sulphide. These involve the same number of valence-shell electrons and might be expected to have similar structures. The bond angles are 104° 27′ in H_2O and 92° 20′ in H_2S. The lower electronegativity of sulphur compared with that of oxygen means that in the bonding pairs the electrons are pulled more towards the hydrogen in H_2S than in H_2O. The reduced electron density in the region of the bonding pairs near the sulphur atom results in their being pushed even nearer together by the unshared pairs. A similar trend is observed in the series NH_3, PH_3 and AsH_3.

Multiple-bonded structures

In the case of double-bonded systems the structure may be predicted by assuming that the double bond occupies one place among the electrons around an atom. Ethylene is a planar molecule and, with the HCH angle being 120°, the arrangement of bonds around each carbon is the same as if there were only three single bonding electron pairs. We could have arrived at this arrangement by assuming that the carbon atoms form four tetrahedral bonds such that the axis of the double bond and the CH bonds are coplanar (Figure 5). If we use the more customary sigma and pi description of the double bond it appears that the arrangement is determined only by the sigma bonds. A similar situation applies to triple-bonded systems, for example acetylene

[1] This compound is not only of academic importance but is commercially important as a fluorinating agent, for example in the production of uranium via uranium hexafluoride.

Table 1

The Arrangement of Electron Pairs in some Binary Compounds

Number of electron pairs	*Arrangement*	*Number of bond pairs*	*Number of unshared pairs*	*Example*	*Shape of molecule*	*Corresponding hybridization scheme*
2	Linear	2	—	$HgCl_2$	Linear	*sp*
3	Triangular	3	—	BCl_3	Planar (120° bond angles)	sp^2
		2	1	$SnCl_2$ (gas)	V-shape	
4	Tetrahedral	4	—	CH_4	Tetrahedral	sp^3
		3	1	NH_3	Pyramidal (trigonal)	
		2	2	H_2O	V-shape	
5	Trigonal-bipyramid	5	—	PCl_5	Trigonal-bipyramid	sp^3d
		4	1	$TeCl_4$	Irregular tetrahedron	
		3	2	ClF_3	T-shape	
		2	3	XeF_2	Linear	
6	Octahedral	6	—	SF_6	Octahedral	sp^3d^2
		5	1	IF_5	Square-pyramid	
		4	2	XeF_4	Square-planar	
7	Pentagonal-bipyramid	7	—	IF_7	Pentagonal-bipyramid	sp^3d^3

is a linear molecule. The structure is the expected one if we assume that the triple bond occupies one position around the carbon atom. This is the arrangement expected whether we assume that the triple bond is made up of three equivalent bonds or one sigma and two pi bonds.

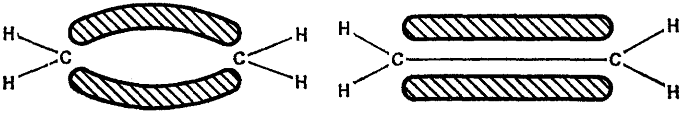
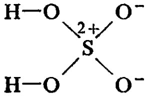
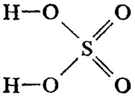


Figure 5 Two alternative representations of the double bond in ethylene

In some compounds we may be a little unsure about the extent of double bonding, for example the sulphuric acid molecule may be represented by either of the extreme structures



or



The true structure is probably somewhere between the two extremes. Whichever structural formula we use, however, we would still arrive at a tetrahedral distribution of bonds around the sulphur atom and an angular arrangement of the bonds around the oxygen atoms of the hydroxyl groups.

Other Descriptions of Shapes

The previous section has been given as a simple working guide to predicting structures. Another approach frequently encountered is that of hybridization.

The reader is probably familiar with electronic configurations of atoms and their relation with the structure of the periodic table. Thus the electronic configuration of carbon is 2. 4 or, in more detail, $1s^2 2s^2 2p^2$. In the outer shell we have two electrons in the $2s$ orbital and an electron in two of the three $2p$ orbitals. Note, however, that this is the arrangement in an isolated carbon atom with the electrons in the lowest energy levels (the ground state). In the proximity of other atoms the energy levels which the electrons occupy are altered by the influence of both the positive nuclei and negative electrons. When a carbon atom combines with four hydrogen atoms, or with four other carbon atoms as in diamond, the four pairs of electrons which the carbon now shares occupy four equivalent orbitals, pointing towards the corners of a regular tetrahedron. We can imagine that on the point of combination the four electrons on the carbon have been rearranged in these new orbitals ready to be shared in the formation of the covalent bonds. The four orbitals in the outer shell of the carbon atoms (one *s* and three *p*) have been replaced by four equivalent orbitals which are hybrids of the pure atomic orbitals. Because it is one *s*- and three *p*-orbitals in this hybridization we call it sp^3 hybridization and the hybrid orbitals will resemble *p*-orbitals more than *s*-orbitals (Figure 6). The hybridization scheme is represented schematically in Figure 7.

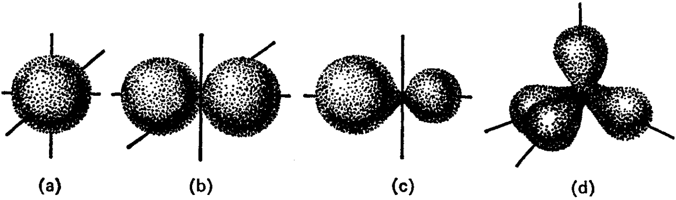


Figure 6 Hybridization of orbitals: (a) an *s*-orbital, (b) a *p*-orbital, (c) an sp^3 hybrid orbital, (d) the distribution of sp^3 hybrid orbitals

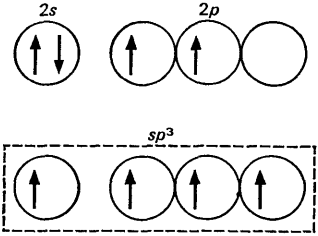


Figure 7 A schematic representation of the sp^3 hybridization of carbon in methane

Other combinations of orbitals may be involved in hybridization schemes. In the case of chlorine trifluoride, which we discussed previously, we saw that the combination of a chlorine atom with three fluorine atoms produced a T-shaped molecule. This resulted from the repulsions between the five pairs of electrons around the chlorine atom which arrange themselves towards the corners of a trigonal bipyramid. The configuration of a chlorine atom is 2.8.7 or, in more detail, $1s^22s^22p^63s^23p^5$. Combination with three fluorine atoms requires three unpaired electrons so an additional orbital must be involved (Figure 8).

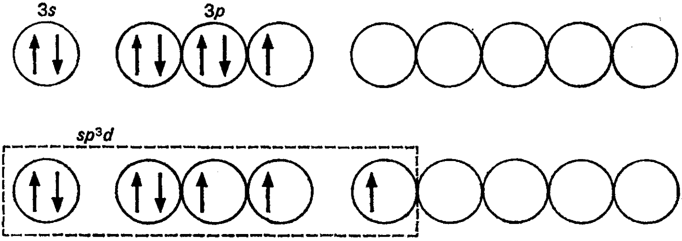


Figure 8 A schematic representation of the sp^3d hybridization of chlorine in chlorine trifluoride

This might be expected to be a 3*d* orbital, but in the presence of the other atoms involved in the combination the pure atomic orbitals are replaced by five equivalent orbitals, and since it is one *s*-, three *p*- and one *d*-orbital which are replaced, the new orbitals are sp^3d hybrid orbitals. These hybrid orbitals are arranged symmetrically around the chlorine atom.

If we wish to work out the structure of any simple compound by this method we must determine what orbitals are involved and use these in a hybridization scheme. Each hybridization scheme will always have the same geometrical distribution of the hybrid orbitals and these will be arranged symmetrically around the atom and as far apart as possible. The various hybridizations are included in Table 1.

The question might be raised at this stage as to whether this treatment is necessary. The short answer is no, if we are only concerned with a qualitative description of the shape. The full hybridization method is, however, a more sophisticated mathematical description of the behaviour of electrons in molecules. Perhaps at this level the main advantage is that it provides a useful shorthand notation in describing the distribution of electrons around an atom. Thus when we say that the electrons around a sulphur atom in sulphur hexafluoride are in sp^3d^2 hybrid orbitals, this implies that there are six pairs of electrons pointing towards the corners of a regular octahedron, that is, the SF_6 molecule is octahedral.

The Structure of Ions

Although the title of this article is 'The shapes of simple molecules' it would be pointless to omit mentioning that the same principles apply in deducing the shapes of simple ions. Thus in ICl_4^- we have six pairs of electrons around the iodine atom, that is, seven from the iodine, four from the chlorine atoms and one giving the excess charge. These six pairs are arranged octahedrally and the two unshared pairs will be arranged above and below the square plane which contains the atoms (Figure 9).

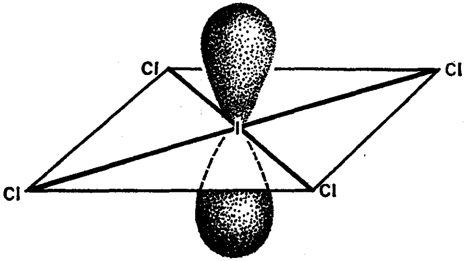


Figure 9 The arrangement of electron pairs in the ICl_4^- ion

Summary

A knowledge of the shapes of molecules is useful in correlating the physical and chemical properties of a compound. The shape may usually be readily deduced from the number of valence-shell electrons. These arrange themselves in pairs which are as far apart as possible because of the repulsions between like charges.

Bibliography

W. E. ADDISON, *Structural Principles in Inorganic Compounds*, Longmans, 1961.

R. J. GILLESPIE and R. S. NYHOLM, 'Inorganic stereochemistry', *Quarterly Reviews*, vol. 11 (1957), p. 339.

G. E. RYSCHEWITSCH, *Chemical Bonding and the Geometry of Molecules*, Chapman & Hall, 1965.

J. E. SPICE, *Chemical Binding and Structure*, Pergamon, 1964.

5 G. Baddeley

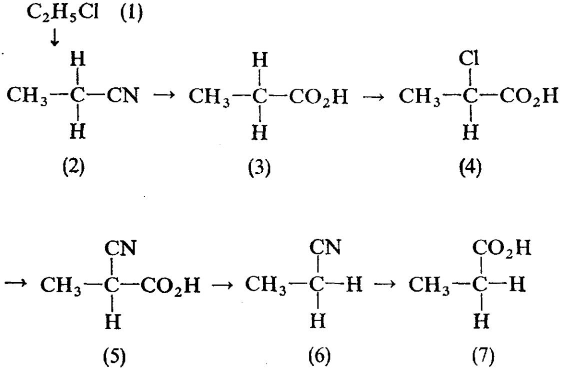
A Fragment of Stereochemistry

G. Baddeley, 'A fragment of stereochemistry', *Education in Chemistry*, vol. 1 (1964), no. 3, pp. 126–35.

This article provides only a part of the nomenclature of stereochemistry and completely ignores the historical development of the subject. Instead, it indicates the type of information on which our ideas concerning stereochemistry are based.

The Tetrahedral Distribution of Bonds

Qualitative analyses followed by quantitative analyses and molecular weight determinations allow us to assign molecular formulae to organic compounds. Every organic compound has a distinct set of physical and chemical properties which differentiates it from every other compound. The physical properties most frequently used in this work include the melting point, the boiling point and interactions with light and other electromagnetic radiations. These interactions are exemplified by the refraction and absorption of light, the absorption of infrared and ultraviolet radiation, and the rotation of the plane of plane-polarized light. Data of the type outlined above enable us to say, without fear of contradiction, that whereas, for example, there are two compounds with the molecular formula $C_2H_4Cl_2$, each of the formulae CH_4, C_2H_6 and C_2H_5Cl represents only one compound. The action of potassium cyanide on ethyl chloride (1) gives propionitrile (2) from which propionic acid (3) can be obtained by hydrolysis. The action of chlorine on this acid can give 2-chloropropionic acid (4) which, by interaction with potassium cyanide, gives 2-cyanopropionic acid (5). This compound, at sufficiently raised temperatures, evolves carbon dioxide and leaves us with propionitrile (6) which is identical in every respect with propionitrile (2) and hydrolyses to propionic acid (7) which is identical with propionic acid (3). This series of reactions does not provide us with a straight (3) and a bent (7) type of propionic acid as it could have done had the four valencies of carbon been directed towards the corners of a square.



$$C_2H_5Cl \ (1) \downarrow \quad CH_3-\overset{H}{\underset{H}{C}}-CN \rightarrow CH_3-\overset{H}{\underset{H}{C}}-CO_2H \rightarrow CH_3-\overset{Cl}{\underset{H}{C}}-CO_2H$$

(2) (3) (4)

$$\rightarrow CH_3-\overset{CN}{\underset{H}{C}}-CO_2H \rightarrow CH_3-\overset{CN}{\underset{H}{C}}-H \rightarrow CH_3-\overset{CO_2H}{\underset{H}{C}}-H$$

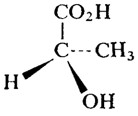
(5) (6) (7)

The hydrolysis of 2-chloropropionic acid (4) gives lactic acid, $CH_3{\cdot}CH(OH){\cdot}CO_2H$, and if we proceed to prepare a number of its salts and study their recrystallization from water, we shall find, sooner or later, that the crystals provided by a particular salt are of two types, both asymmetric (since should Alice retrieve the image of any one crystal from behind the mirror it would not be superimposable on the crystal in front of the mirror) and related to one another as object to mirror image. Given great patience, good eyesight and a steady hand we can separate the crystals of one type from those of the other and thereafter recover separately the lactic acid from each. The two samples of lactic acid thus obtained have identical physical properties except in their effect on plane-polarized light. Both are *optically active*: one, named (+)-lactic acid, rotates the plane of polarization clockwise (relative to the observer) while the other, (—)-lactic acid, rotates it to exactly the same extent but anticlockwise. The ability to rotate the plane of polarized light is not lost when the two lactic acids are in solution: we conclude, therefore, that *the molecules of lactic acid are asymmetric*, a molecule of (—)-lactic acid being the mirror image of a molecule of (+)-lactic acid.

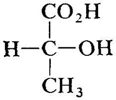
That compounds of the type Ca_4, Ca_3b and Ca_2bc are unique while those, like lactic acid, of the type Cabcd can occur in (+)- and (−)-forms finds its interpretation in the concept that *the bonds attaching a carbon atom to four other atoms are directed from the centre to the four corners of a regular tetrahedron.*

Configuration

An attempt to give a three-dimensional representation of the two forms of lactic acid is shown below (8 and 10). Fischer projections provide a convenient form of three-dimensional representation (9 and 11 are the Fischer projections of 8 and 10). To obtain them, project on to the paper what an eye would see if, as it looks at the central carbon atom, the bonds directed *away from* the observer are in a *vertical plane* while those directed *towards* the observer are in a *horizontal plane.*



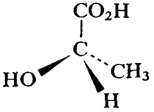
(8)



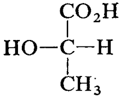
$$\begin{array}{c} CO_2H \\ | \\ H-C-OH \\ | \\ CH_3 \end{array}$$

(9)

(−)-Lactic acid



(10)



$$\begin{array}{c} CO_2H \\ | \\ HO-C-H \\ | \\ CH_3 \end{array}$$

(11)

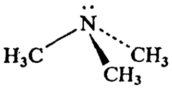
(+)-Lactic acid

When, as in lactic acid, four different atoms or groups of atoms are attached to the same carbon atom, this atom is called an *asymmetric carbon atom*, since there are two arrangements or *configurations* of the atoms and groups attached to it. Pairs of molecules that are non-superimposable and are the mirror image of each other are called *enantiomers*, *enantiomorphs* or *optical isomers*. When enantiomers are present in equal amount, the whole is optically inactive and is called a *racemate*. The lactic acid we prepared from propionic acid by chlorination followed by hydrolysis is optically inactive and is called racemic or (±)-lactic acid.

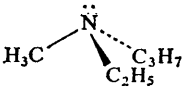
Living organisms usually produce only one enantiomer of a pair; thus the vegetable kingdom provides many optically active nitrogenous organic bases. When one of these bases, for example (−)-quinine from the bark of the cinchona tree, reacts with (±)-lactic acid, two salts are formed which can be represented by (−)-B, (+)-A and (−)-B, (−)-A in which A and B represent the acid and the base respectively. These two salts are not enantiomeric: they have the same configuration in one part and opposite configurations in the other. They are *diastereomeric.* Unlike enantiomers, *diastereomers* have different properties and, often, are conveniently separated from one another by recrystallization. Thus we can separate (−)-B, (+)-A from (−)-B, (−)-A; afterwards we can obtain (+)-lactic acid from the former and (−)-lactic acid from the latter. In this manner (±)-lactic acid has been *resolved* into its optically active enantiomers.

To discover the actual or *absolute* configuration of the lactic acid which rotates the plane of polarized light to the right and of that which rotates it to the left, we must sooner or later, directly or indirectly, have recourse to X-ray diffraction patterns from which we can precisely deduce the spatial arrangement of the parts of a molecule. Formulae 8 and 9 represent (−)-lactic acid and formulae 10 and 11 (+)-lactic acid. The esterification of (−)-lactic acid with methyl alcohol gives the (+)-methyl ester in which the $—CO_2CH_3$ group occupies the same relative position in the molecule as does the $—CO_2H$ group in the molecule of (−)-lactic acid. Thus we are compelled to recognize that two compounds may have similar configurations yet rotate the plane of polarized light in opposite directions.

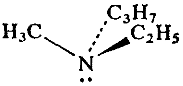
Is the 'tetrahedral distribution' of bonds shown by carbon atoms also a feature of atoms of other elements? Electron diffraction data concerning trimethylamine (12) have shown the molecules to be pyramidal with $C\overset{N}{\diagup\;\diagdown}C$ angles of about 108°, very nearly the same as the $H\overset{C}{\diagup\;\diagdown}H$ angles (109·5°) of methane molecules. We might expect to find optically active forms (13) and (14) of methylethylpropylamine, but optical isomers of this kind have not been found. Spectroscopic data show that the pyramidal molecules (13) and (14) require only about 5 kcal mol^{-1} to enable the pair of unshared electrons to pass through the nitrogen atom and thereby invert the configuration. This amount of energy is readily available at room temperature. When the four pairs of electrons surrounding the nitrogen atom are used to bond the nitrogen atom to *four* different groups, as in methylallylphenylbenzylammonium iodide (15), two optically active forms exist. Clearly, the carbon atom is not the only one which can have its bonds directed from the centre to the corners of a regular tetrahedron, the angle between any two bonds being about 110°.



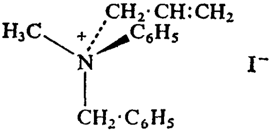
(12)



(13)

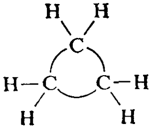


(14)



(15)

When 1,3-dichloropropane is allowed to react with zinc, cyclopropane (16) is formed. What has happened to the interbond angle of 110° ? Perhaps we ought to conclude that bonds can be bent into the sausage shape shown in formula (16). From our knowledge of the molecular shapes of many compounds and of the energy, measured by infrared spectroscopy, to change



(16)

intramolecular vibrations, we are able to deduce that bonds are easier to bend than to stretch.

Conformation

When we construct a model of a molecule of ethylene dichloride we find that, without stretching or bending bonds, we can obtain a variety of molecular shapes by *rotating* one half of the molecule relative to the other half about the bond joining the carbon atoms. Nearly aligning our eyes with the C—C bond, let us begin with the chlorine atoms and hydrogen atoms *eclipsing* one another as shown in (17) and gradually rotate the nearer —CH_2Cl group so that we pass through the shapes represented by (18), (19) and (20). Since no one has found two or more types of ethylene dichloride, we might be led to suppose either that the two halves of the molecule can rotate freely and independently of one another about the C—C bond or that the molecule is restricted to one of the shapes (17) to (20). The following considerations help to resolve this ambiguity. The effects of methyl chloride on electrical fields into which it is placed show that the C—Cl bond provides an electric dipole with a moment (magnitude of the charges × distance apart) of about 2·0 D, where 1 D or a Debye unit =

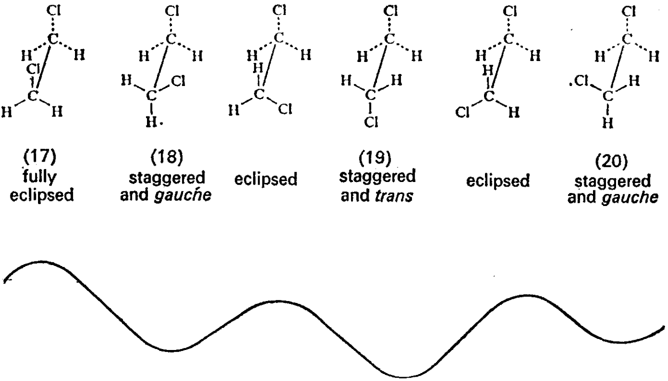
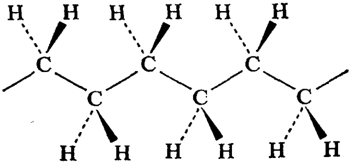


Figure 1 Curve indicating the internal energy change as one —CH_2Cl group rotates

10^{-10} e.s.u. Å. Ethylene dichloride has two such dipoles and the dipole moment of the molecule, which can be measured, is the vector sum of the dipole moments of the two halves. If ethylene dichloride were as represented by (19), its dipole moment would be zero; if as represented by (17), 3·8 D. Free rotation would provide the average value of $\frac{1}{2}(0+3{\cdot}8) = 1{\cdot}9$ D. Experiment reveals that its dipole moment is 1·12 D at 32° and gradually increases with rise in temperature until it is 1·54 D at 271°. The data are compatible with the view, now generally held, that as one half of the molecule rotates with respect to the other (a) the internal energy of the molecule rises and falls as shown by the curve under formulae (17) to (20), (b) the molecule can rest only in the energy hollows, (c) the molecule prefers the hollow provided by (19) to the shallower hollows provided by (18) and (20), and (d) this preference becomes less apparent as the temperature is raised. The various molecular shapes (17) to (20) are called *conformations*; (19) shows the molecule in its *trans* conformation while (18) and (20) show the molecule in its *gauche* or *skew* conformations. In each of these conformations, the bonds radiating from one carbon atom are *staggered* with respect to those radiating from the other.

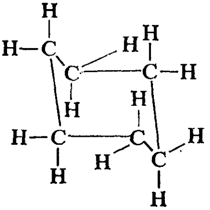
X-ray diffraction patterns produced by paraffin wax show the polymethylene chains to have the *zig-zag* conformation (21) in



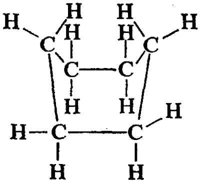
(21)

which every group of four consecutive methylene groups has the staggered and *trans* conformation.

When we construct a model of the cyclohexane molecule, only two conformations are possible if we are to have an interbond angle of 110°: the *chair* conformation (22) and the *boat* conformation (23). The infrared spectrum of cyclohexane shows that at ordinary temperatures the molecules spend a much longer time in the chair than in the boat conformation. Only in the former are the bonds of each carbon atom staggered with respect to those of the neighbouring carbon atoms. The preference for staggered conformations determines molecular shape only when

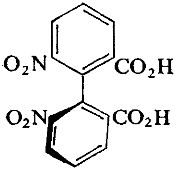


(22)

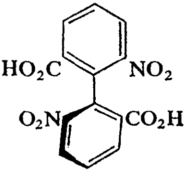


(23)

more important demands on shape are absent. It may originate in repulsive forces between atoms and groups, especially the bulkier ones which are brought close together without being able to bond one to the other. The resolution of 6,6′-dinitrodiphenic acid into its optical isomers (24) and (25) demonstrates that in this compound an amount of energy too large to be readily available at ordinary temperatures is required to enable the substituents in one of the rings to squeeze past those in the other.



(24)

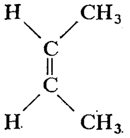


(25)

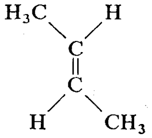
It is noteworthy that the non-planar molecules (24) and (25) are asymmetric although they do not contain an asymmetric atom.

The Coplanarity Requirement of Double Bonds

When 2-bromobutane reacts with bases it can lose the elements of hydrogen bromide and give a hydrocarbon which can be separated into three components boiling at −6°, 1° and 4°C severally. Each component has the molecular formula C_4H_8. The one which boils at −6°C gives formaldehyde and propionaldehyde on ozonolysis and is assigned the formula $CH_3.CH_2.CH{:}CH_2$, while the other two give only acetaldehyde on ozonolysis and are both represented by the formula $CH_3.CH{:}CH.CH_3$. Of these two hydrocarbons, the one with the higher boiling point has a dipole moment and the other has not; consequently, they

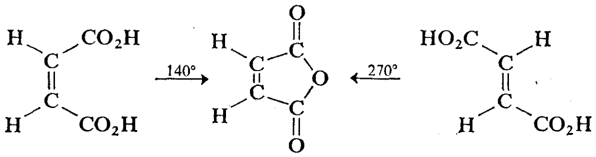


(26)



(27)

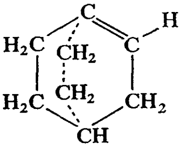
are given respectively the formulae (26) and (27) and named *cis*- and *trans*-but-2-ene. Two compounds are represented by the structure $HO_2C{\cdot}CH{:}CH{\cdot}CO_2H$; one is known as maleic and the other as fumaric acid. When heated to 140°, maleic acid gives maleic anhydride (29) and water while fumaric acid loses water to form maleic anhydride only at temperatures above 270°C; consequently, they are assigned the formulae (28) and (30) respectively. Clearly, the energy barrier to rotation about a carbon–carbon double bond is sufficiently high to hold the four attached atoms or groups in a rigid configuration except at high temperatures. Electron diffraction and spectroscopic studies show ethylene to be a flat molecule with bond angles of about 120°.

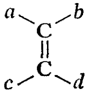


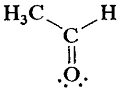
(28) (29) (30)

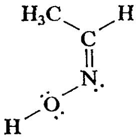
The isomers (26) and (27) and again (28) and (30) are called *geometric* isomers. Isomers that differ from one another only in the way that atoms are oriented in space but are identical with respect to which atoms are bonded to one another are called *stereoisomers*. Geometric isomerism and optical isomerism are both examples of stereoisomerism.

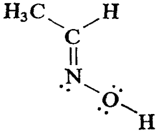
A carbon–carbon double bond requires coplanarity of the two carbon atoms and the four atoms directly attached to them (see 32); understandably, olefins of the type (31) cannot be obtained. Now let us turn our attention briefly to double bonds between atoms other than two carbon atoms. Although there is only one type of acetaldehyde (see 33), this compound condenses with hydroxylamine to give two oximes, (34) and (35). Further,

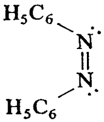

(31)

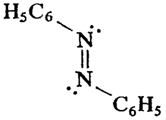

(32)


(33)


(34)


(35)

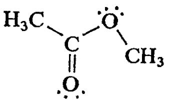

(36)


(37)

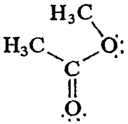
cis- and *trans*-forms of azobenzene, (36) and (37), are well known. In the formulae (33) to (37), unshared pairs of electrons are represented by dots. Clearly, a double bond between atoms imposes planarity even when the two atoms are other than carbon atoms, and is capable of providing geometric isomers when between carbon and nitrogen atoms or between two carbon or two nitrogen atoms.

Conjugation and Conformation

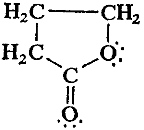
To what extent does the planarity required by a double bond determine the shape of an ester molecule? If the bonding in methyl acetate is correctly represented by (38), only the methyl group attached to an oxygen atom is free to move out of the plane of the rest of the molecule. Perhaps it is free to rotate about the bond joining its oxygen atom to the carbonyl group. The wholly planar conformations (38) and (39) represent the molecule in its least polar and most polar forms respectively. Conformation (39) should provide a dipole moment of about 4·0 D, since this is the moment of γ-butyrolactone (40). Calculation leads us to expect conformation (38) to provide a moment of 1·5 D. Free



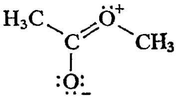
(38)



(39)



(40)

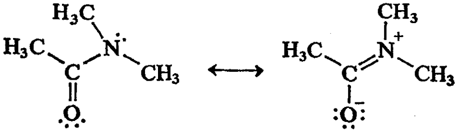


(41)

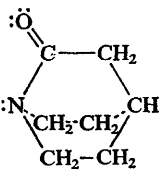
rotation of the methyl group through these two extreme conformations would provide an average moment of 2·7 D. The measured value is 1·7 D and does not change as the temperature of the ester is raised.

We must conclude that methyl acetate is firmly held to conformation (38). The popular interpretation of this conclusion is based on the concept that a pair of electrons, shown in (38) as belonging only to the oxygen atom of the methoxyl group, reaches to some extent into the carbon atom of the carbonyl group. This electronic redistribution, taken to 100 per cent, would result in structure (41); developed to the state of only 10–15 per cent it would provide *partial double bonding* between methoxyl and carbonyl groups sufficient to require the methyl group of the methoxyl group to be near the plane of the rest of the molecule. This interpretation does not discriminate between conformations (38) and (39): perhaps dipole–dipole interaction of methoxyl and carbonyl groups gives preference to (38). The partial double bonding introduced above is referred to as *conjugation* or *electron delocalization*. It lowers the internal energy of any molecule in which it occurs and may require a degree of molecular planarity which is not apparent from the usual formulation. Should we fail to recognize that a molecule like acetdimethylamide (42) must have all but its hydrogen atoms close to one plane if the nitrogen atom and carbonyl group are to be conjugated, we would wonder why, despite the efforts of some of the world's best chemists, the compound to be assigned the formula (43) has not been prepared (cf. 31).

So far, this article has included very little chemistry, since it is concerned only with the shapes of molecules – a topic which these days requires no chemistry. Chemistry is concerned with processes in which bonds are changed, some being broken, others being made, and we must direct our attention to the shape requirements of some of these processes if the word 'stereochemistry' is to appear in the title of this article.



(42)

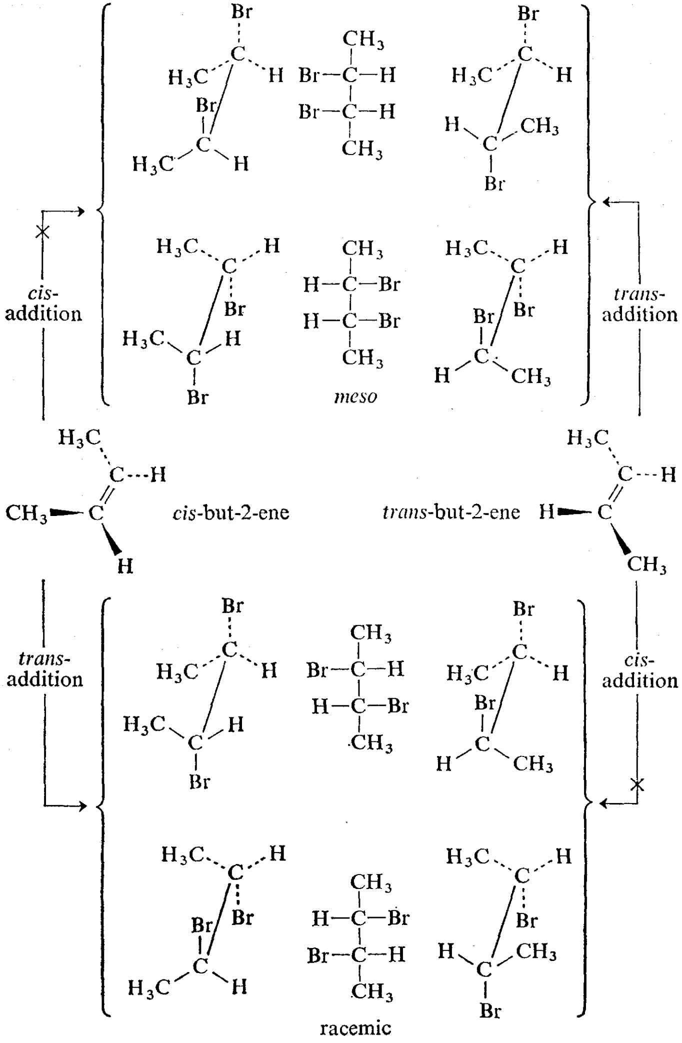


(43)

Additions to Ethylenic Double Bonds

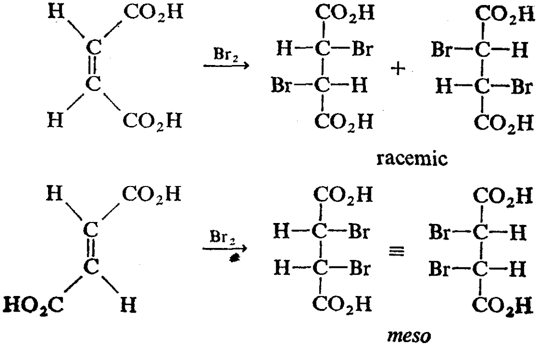
Bromine rapidly adds on to ethylene to give ethylene dibromide, from which ethylene can be regenerated through the agency of zinc:

$$CH_2{:}CH_2+Br_2 \rightarrow CH_2Br{\cdot}CH_2Br \xrightarrow{Zn} CH_2{:}CH_2.$$


cis-addition
trans-addition
meso
cis-but-2-ene
trans-but-2-ene
trans-addition
cis-addition
racemic

Similarly, both *cis*- and *trans*-but-2-ene combine rapidly with bromine to give 2,3-dibromobutane, and if zinc is allowed to react with these two samples of dibromobutane, that from *cis*-but-2-ene regenerates *cis*-but-2-ene while that from *trans*-but-2-ene regenerates *trans*-but-2-ene. Thus we are driven to the surprising though unavoidable conclusion that the 2,3-dibromobutane from *cis*-but-2-ene is different from that from *trans*-but-2-ene. The accompanying chart shows all possible addition products of bromine with *cis*- and *trans*-but-2-ene. Each of the four formulae down the centre of the chart is the Fischer projection of the three-dimensional formulae which flank it. Molecular models save us much mental juggling when we try to assure ourselves that any three formulae of 2,3-dibromobutane which are side by side can be inter-converted merely by rotating one half of the molecule with respect to the other half, and that all three formulae therefore represent the same molecule in different conformations. Do the four rows of formulae represent four different molecules? There are several ways of finding the answer. Let us find it by considering only the Fischer projections.

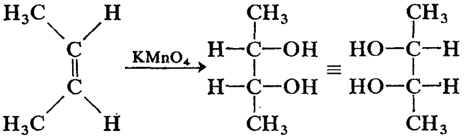
All four projections look different: do they represent four different molecules? If we turn over the lowest one, as we might turn over a hand from the palm-up to the palm-down position, it would then look like the projection immediately above it. But if a Fischer projection is to remain a Fischer projection we must not turn it over in this way, for if we do the bonds in a vertical plane will no longer be directed away from us and the bonds in horizontal planes will no longer be directed towards us. All we can do to a Fischer projection is rotate it in the plane of the paper through 180°. If we do this to the bottom two projections each retains its identity: they represent two different molecules. If we slide between them a mirror which is perpendicular to the plane of the paper we see that each is the mirror image of the other: they are optical isomers. Whether they are formed by *trans*-addition of bromine to *cis*-but-2-ene or *cis*-addition of bromine to *trans*-but-2-ene, they will be formed in equal amount: that is, racemic 2,3-dibromobutane will be formed. When we rotate the topmost projection through 180° it becomes the second one down: the top two projections represent the same molecule. When we place a mirror alongside these two projections we see that each is the mirror image of the other: therefore, they represent a molecule which is not asymmetric and cannot be optically active, even though it contains two asymmetric carbon atoms. They represent *meso*-2,3-dibromobutane. In the *meso* form, both halves of the molecule are asymmetric and, as shown by a vertical mirror so placed as to bisect either of its projection formulae, one half of the molecule is the mirror image of the other half. We see that 2,3-dibromobutane can occur in two optically inactive forms: the racemic form which is a mixture of equal parts of (+)- and (−)-2,3-dibromobutane, and the *meso* form in which all the molecules are identical and optically inactive. *Meso* and racemic forms are diastereomeric, since if we compare a molecule of the former with a molecule of either enantiomer of the latter we shall find the two molecules have the same configuration in one half of their molecules and opposite configurations in the other. The two different 2,3-dibromobutanes obtained by the addition of bromine to *cis*- and *trans*-but-2-ene must be the racemic and *meso* forms, but not necessarily respectively. How can we discover which form comes from *cis*-but-2-ene and which from *trans*-but-2-ene?



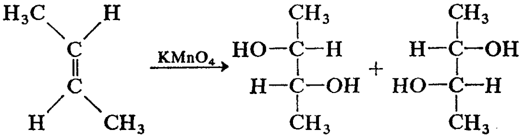
The addition of bromine to maleic acid gives a dibromosuccinic acid which is different from the one obtained by the addition of bromine to fumaric acid. Remembering how we resolved (±)-lactic acid with the aid of an optically active base, let us try to resolve these two dibromosuccinic acids.

Success crowns our efforts only when we work on the one from maleic acid. Thus we learn that maleic acid gives racemic dibromosuccinic acid and that the *meso*-dibromo acid is obtained from fumaric acid. If we are persuaded by analogy, we shall expect the addition of bromine to *cis*- and *trans*-but-2-ene to give racemic and *meso*-2,3-dibromobutane respectively. We see that the addition of bromine across an ethylenic double bond is orderly: *stereospecific* is the adjective which best describes it. The reaction is a *trans*-addition.

Another chemical property of alkenes is their ability to decolorize permanganate. A glycol is formed in this reaction, a hydroxyl group having been added to each of the two carbon



meso-butan-2,3-diol

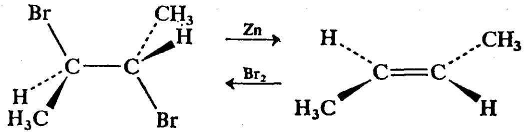


racemic butan-2,3-diol

atoms involved in the double bond. The permanganate oxidation of *cis*- and *trans*-but-2-ene provides two different glycols: only the glycol from *trans*-but-2-ene can be resolved (I wonder how it was resolved). No doubt the reaction of alkenes with permanganate leads to the *cis*-addition of two hydroxyl groups.

Elimination Reactions

Although molecules of *meso*-2,3-dibromobutane can turn from one conformation to another, one half of each molecule being able to rotate with respect to the other half, this compound reacts with zinc to give *trans*- and not *cis*-but-2-ene. Apparently, in this elimination process the molecule is required to have the conformation in which the bromine atoms are most widely separated: that is, the conformation in which the two bromine

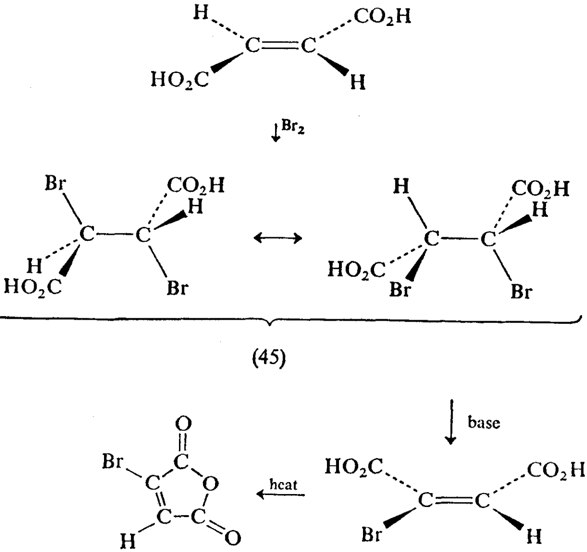


(44)

meso-2,3-dibromobutane

trans-but-2-ene

atoms and the two carbon atoms that link them are all in one plane and the bromine atoms are on opposite sides of the bond joining the two carbon atoms. This requirement is represented by formula (44). The reaction is stereospecific and is suitably described as a *trans*-elimination. This stereospecificity allows the reaction of racemic 2,3-dibromobutane with zinc to give only *cis*-but-2-ene.



(45)

(46)

Molecules which have a hydrogen atom and a halogen atom attached to adjacent carbon atoms can react with base so as to lose the elements of hydrogen halide and acquire an additional bond between the two carbon atoms. What is the stereochemistry of this elimination process?

Having already obtained *meso*-dibromosuccinic acid (45) by the addition of bromine to fumaric acid, let us now, through the agency of a base, remove from it the elements of hydrogen bromide. The product readily forms an anhydride and is therefore bromomaleic acid (46).

Similarly, racemic dibromosuccinic acid, which we obtained by the addition of bromine to maleic acid, gives bromofumaric acid. These reactions are *trans*-eliminations: they require the hydrogen and bromine atoms which are to be eliminated and the two carbon atoms which link them to be in one plane, and the hydrogen and bromine atoms to be on opposite sides of the bond joining the two carbon atoms.

Stereochemistry is concerned with the geometry of chemical reactions, and it helps us tremendously in our attempts to discover their mechanisms. The elucidation of the mechanisms of reactions not only satisfies our curiosity; it helps us to foretell the practical procedure most likely to bring about a required reaction. Although there are numerous reactions, there are but few *types* of reaction, each defined by its stereochemistry and mechanism. The fragment of stereochemistry provided by this article is concerned with the geometry of one or two addition and elimination processes; I hope we are ready and anxious to discover the stereochemistry of other types of reactions, especially that of substitution processes, since they make a vast contribution to chemistry.

Bibliography

D. J. CRAM and G. S. HAMMOND, *Organic Chemistry*, 2nd edn, McGraw-Hill, 1964, ch. 9.

G. HALLAS, *Organic Stereochemistry*, McGraw-Hill, 1965.

J. M. TEDDER and A. NECHVATAL, *Basic Organic Chemistry*, part 2, Wiley, 1967, chs. 23 and 24.

Part Three **Energetics and Kinetics**

The study of the energetic and kinetic aspects of chemical reactions is central to the understanding of chemistry. J. A. Campbell in his article 'Why do chemical reactions occur?' discusses the importance of energy changes in chemical reactions and gives a qualitative answer to one of the most important questions in chemistry in terms of the rate and equilibrium state of the reaction. This leads to a more quantitative treatment in thermodynamic terms involving the concept of entropy as a randomness factor and the use of free energy change as a criterion for chemical change. The author emphasizes that the probability of a reaction cannot be predicted from energetics alone and that the rate of reaction cannot be predicted from a knowledge of the free energy change. An expanded version of this article is now available in book form (see 'References' at the end of the article).

The principles and applications of thermochemical cycles are discussed by J. Dwyer and J. Lee, and A. R. Denaro discusses electrode reactions in terms of the equilibrium and the kinetic approach. A. G. Sharpe develops the concepts of oxidation and reduction quantitatively in terms of standard electrode potentials and discusses the factors which determine their magnitude.

6 J. A. Campbell

Why Do Chemical Reactions Occur?

J. A. Campbell, 'Why do chemical reactions occur?', *Chemistry Today*, O.E.C.D., 1963, ch. 8, pp. 163–81.

The question of why reactions occur is not initially simplified by observing the reactions around us. The iron sides of a blast furnace rust, while inside the furnace rust is changing into iron. Both reactions are certainly occurring. Why sometimes one and sometimes the other? At one spot in the forest a tree rots in the presence of oxygen to form carbon dioxide and water, while in another spot the water and carbon dioxide just released by the rotting are reacting to form new wood and oxygen. The rotting of the one tree and the growth of the other are certainly both occurring. The surrounding conditions seem similar. Yet the two chemical reactions are the reverse of one another. How is it that both can occur in nature?

Conservation Laws

Let us first consider some scientific generalizations which can be shown to be inapplicable to the question of whether chemical reactions will occur or not. These are the conservation laws[1] – the conservation of electric charge, the conservation of elements, the conservation of atoms, the conservation of mass, the conservation of momentum and the conservation of energy in each chemical reaction. Some of these laws may be useful in correlating and describing the chemical changes, but they cannot in general be used to predict chemical reactions.

We shall find that it is very important to consider the energy changes in chemical reactions. The total energy of a system may

---

[1] Exceptions are known to most of the conservation laws as simply expressed above, for example the interconversion of mass and energy. But these exceptions are seldom important in chemical reactions.

always remain constant, yet the system may change if the distribution of the energy varies. Likewise, the number of atoms will be conserved, but they may redistribute themselves into new compounds. The questions are: why do these redistributions occur, what is it that changes and how can reactions occur and still obey the conservation laws?

Mechanism and Extent of Reaction

The question 'Why do reactions occur?' can be clarified if we consider it from two aspects: (1) the rate and mechanism of the reaction and (2) the extent or equilibrium state of the reaction.

For example, a reaction may proceed very rapidly because of a simple mechanism and yet occur to a scarcely detectable extent. The ionization of hydrogen cyanide upon solution in water is an example. The reaction occurs rapidly 'because a simple mechanism is available', but it does not occur to any great extent. Only a slight amount of ionization is observed. Conversely, some reactions occur extremely slowly, yet eventually will consume almost all of the initial reactants. A gaseous mixture of two moles of hydrogen and one mole of oxygen reacts at an imperceptible rate. Millions of years would be required for detectable reaction if the temperature remained low. Yet, over a very long period of time, practically all of the hydrogen and oxygen would react to give water. Water is the more 'stable' substance by far. This reaction occurs almost completely because water is more stable at room temperature than a mixture of hydrogen and oxygen gases, yet the rate of reaction can be minute.

We shall firstly investigate the effects which determine the rates and mechanisms of reactions and then the effects which determine relative stability.

Molecular Collisions

An indisputable prerequisite to chemical reaction is a collision between the reacting molecules. Let us firstly, therefore, examine the types of collisions and the factors that will govern the probability of reaction during the collision.

Distribution of kinetic energy

Let us consider the state of any system of molecules. Under any set of conditions the system will contain a certain amount of energy. Unless the system is at absolute zero, the translational kinetic energy will be distributed in a random fashion among all the molecules of the system. Consider a monatomic gas. The Maxwell–Boltzmann distribution of energy holds, and the distribution of translational kinetic energies ($=\frac{1}{2}mv^2$) may be represented by curves like those in Figure 1. The effect of temperature on the shape of the curve is well known, and is indicated by two separate curves.

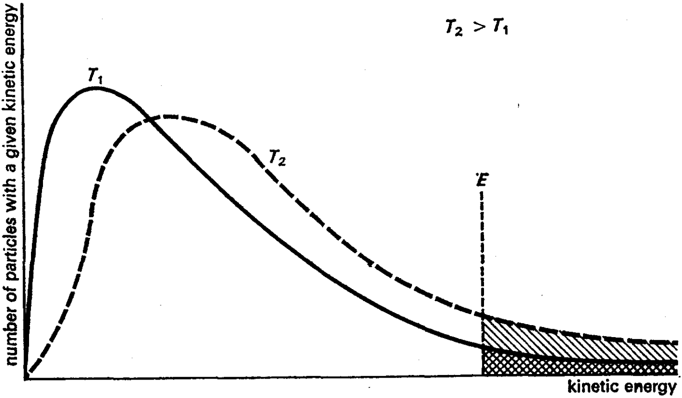


Figure 1 Effect of temperature on atomic (or molecular) kinetic energy distribution. (Note that the number of particles with energy greater than *E* increases rapidly as the temperature rises)

Maxwell showed that such a distribution would result even if at some given instant all of the particles should suddenly achieve the same kinetic energy. Random collisions would quickly lead to redistribution of energies to that represented by the curves. Very few molecules have low energies. The probability of a molecule having a given energy rises to a maximum and then falls off. The decrease in probability in the high energy region is relatively slow, thus there are comparatively large numbers of molecules with rather high energies. The percentage of molecules with these high energies increases rapidly as the temperature of the system is raised.

As a result of this varying distribution of energy among the molecules, molecular collisions of many kinds occur. Some collisions, those between molecules of low kinetic energies, are 'soft' and involve only a gentle push by one molecule on the other. Many collisions, however, are 'hard' and involve considerable impact of one molecule on another. And a few collisions involve very high impact energy due to the high initial kinetic energy of the two colliding particles. This impact energy is particularly high, of course, when two molecules happen to undergo head-on collision. Thus, the nature of a collision is influenced both by the kinetic energies of the molecules involved and by their relative paths at the moment of collision.

It is not possible to apply a Maxwell–Boltzmann distribution curve exactly to condensed phases, that is, liquids and solids, but there is every reason to believe that the general effects are similar to those we have just discussed in the gas phase. Collisions of various total energies will occur, and the effective impact energy will depend both on the energies of the colliding molecules and their paths with respect to one another at the moment of collision.

A minor (but not negligible) effect is that the more energetic molecules, since they are travelling more rapidly, will make more collisions than will the slower moving molecules in the same length of time.

Bond strengths

All atoms tend to attract one another to some extent. The extent, however, varies greatly between different pairs of atoms. The interaction between two helium atoms is so small that they condense and stick together only at the very lowest temperatures, whereas the interaction between two carbon atoms, as in graphite, or between two nitrogen atoms, as in the nitrogen molecule, is extremely strong. These strong bonds dissociate appreciably only at temperatures of thousands of degrees Kelvin. Bond strengths vary from less than one kilocalorie per mole in the case of helium to several hundred kilocalories per mole in the case of the strongest chemical bonds known.

The strength of a bond between any two atoms also depends on the surrounding atoms. The strength of the oxygen–hydrogen bond, for example, will vary depending upon the other atoms to which the oxygen may or may not be bonded. The dissociation energy of the hydroxyl–hydrogen bond (HO—H) is 118 kcal mol^{-1}, and that of the oxygen–hydrogen (O—H) bond 102 kcal mol^{-1}. Clearly the first and second hydrogen atoms are bonded differently. The over-all energy of a bonding oxygen atom and two hydrogen atoms in water is, of course, the sum of the above two values or 220 kcal mol^{-1}. This gives an 'average' oxygen–hydrogen bond strength of 110 kcal mol^{-1}. The individual values, not the average ones, are called 'the bond dissociation energies'.

Bond dissociation energies vary from very small values to quite large values, and are characteristic of the bond being broken as well as the bonding situation to the neighbouring atoms. Only a few exact bond dissociation energies are known, but more information is being acquired all the time. These dissociation values should be used as interpreting reactions rather than the more commonly cited bond energies. The latter are obtained by an averaging process and are ordinarily assumed to be constant regardless of the environment in which the bond may be.

Mechanism of Chemical Reaction

All chemical reactions appear to occur through a collision mechanism or through the absorption of radiant energy (which may be considered as a collision between a molecule and a photon).

Chemical reaction occurs when one particular collision results in the formation of molecules not present prior to the collision. The molecules may be new because bonds have broken, or because bonds have formed, or because both have occurred. Since the collision processes are random, both in the configuration with which the molecules approach one another and the kinetic energies with which they collide, the likelihood of reaction will depend at least on these two effects.

Many reactions have rather stringent orientation requirements. Enzyme reactions are supposed to be a particularly good example of this. The enzyme in general will not react unless the enzyme and substrate form a very well-defined configuration. Even such a simple reaction as that between hydrogen molecules and iodine molecules to give hydrogen iodide becomes quite unlikely, unless the hydrogen and iodine collide in such a manner that the internuclear axes of the two molecules are parallel to one another at the moment of collision. Since not very many collisions are apt to have this particular configuration, it is clear that a high proportion of all molecular collisions will not lead to reaction, but will merely result in a rebound of the colliding particles. Figures 2(a) and 2(b) illustrate a favourable and an unfavourable configuration for the H_2–I_2 reaction.

[Recently the reaction between hydrogen and iodine has been re-investigated (J. H. SULLIVAN, *Journal of Chemical Physics*, vol. 46 (1967), p. 73). He finds that the mechanisms for the thermal and photochemical reaction are the same, thus indicating that the primary step in both reactions is a dissociation of an iodine molecule into two atoms. A fast reaction between an iodine atom and a hydrogen molecule to give H_2I, followed by a slow reaction between this and a second iodine atom to give hydrogen iodide, provides a mechanism consistent with all present data. The discussions in this article of the hydrogen–iodine reaction should be modified in the light of the present knowledge, but it will be found that there is no change in the fundamental principles being discussed.]

In the general case, an activation energy is also required in order for reaction to occur. This may be thought of as the energy required to loosen up the original bonds so that the new bonds can begin to form. The activation energy is normally represented in a schematic fashion by plotting the potential energy of the system against the path of the reaction, or reaction coordinate, as shown in Figure 3. It is conventional to define the potential energy of the reactants as zero. When molecules with less kinetic energy than the activation energy collide, simple rebounding occurs. When two molecules with sufficient kinetic energy to supply the activation energy collide, it is possible for them to surmount the energy barrier and to go on to form products. One must notice, however, that it is not only necessary to have sufficient energy to surmount the activation energy barrier; the geometry of the collision must also be suitable.

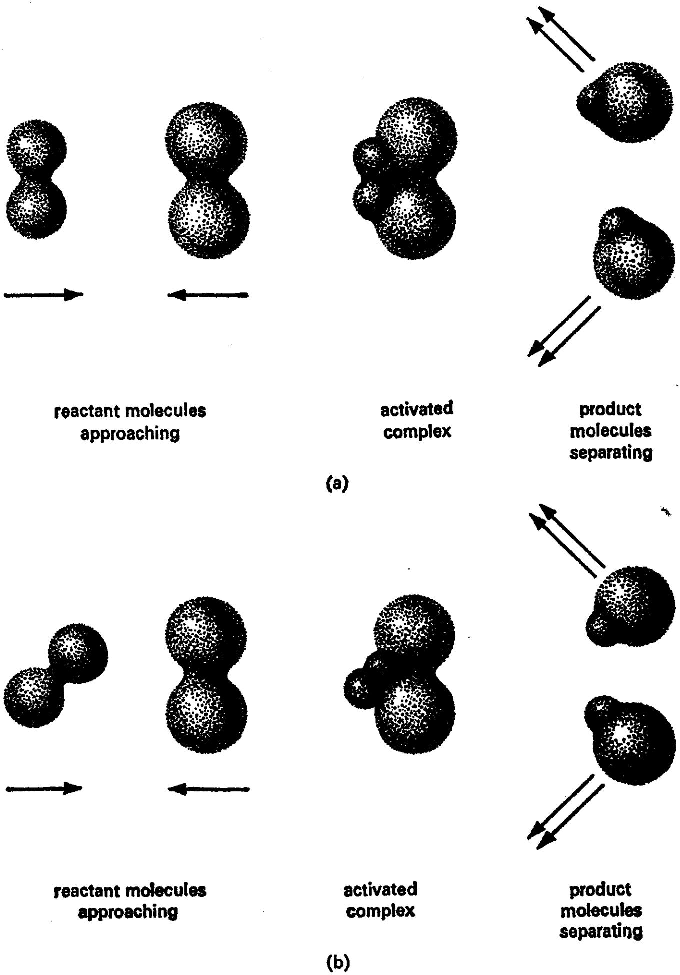


Figure 2 (a) Collision with favourable geometry for reaction: $H_2 + I_2 \rightarrow 2HI$
(b) Collision with poor geometry for reaction: $H_2 + I_2 \rightarrow 2HI$

Thus, each activation energy is characteristic of a particular orientation. Unless very large amounts of energy are available, few reactions occur except those which involve the most favourable geometry of collision.

Normally, the geometrical effect and the activation energy effect can be treated separately. The activation energy effect is, of course, dependent upon the temperature, which determines the kinetic energy distribution among the molecules. The higher the temperature, the more likely it is that colliding molecules will have sufficient energy to traverse the activation energy barrier. On the other hand, a change in temperature ordinarily has no effect on the relative likelihood of various configurations during a collision and, therefore, there is no change in the probability of a suitable geometrical collision when the temperature is changed. This means that the temperature effect on the rate of reaction can be interpreted primarily in terms of the activation energy, and on the likelihood of a collision occurring at all, with no necessity of paying any attention to the configuration requirements during the collision.

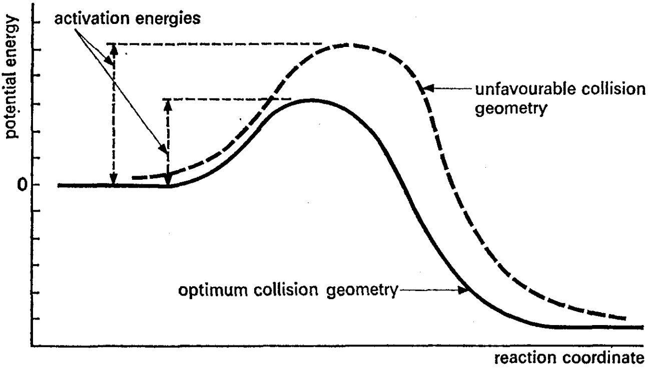


Figure 3 Potential energy diagram for reaction showing effect of collision geometry

Reactions occur then when molecules (a) collide, (b) are in a suitable configuration and (c) have sufficient energy to surmount the activation energy barrier. Thus, if a very high activation energy barrier exists, or if very stringent geometrical requirements must be met, the reaction becomes extremely unlikely in any finite length of time at low temperatures. The slow reaction between hydrogen gas and oxygen gas is a case in point.

Reversible (and Irreversible) Reactions

The curve drawn in Figure 3 shows a typical change in potential energy of a system as it follows the reaction coordinates from 'reactants' to the 'products'. The specific example was the reaction between hydrogen molecules and iodine molecules to give hydrogen iodide molecules. It should now be clear, however, that the hydrogen iodide molecules will have the same type of random motion that characterizes hydrogen and iodine, except that the average molecular velocities will be somewhat different since the masses of the molecules differ. Thus, hydrogen iodide molecules will be colliding with hydrogen molecules or iodine molecules or other hydrogen iodide molecules. If two hydrogen iodide molecules collide with the proper geometry and sufficient energy, then they may surmount the activation energy barrier from the 'reverse' side and re-form a hydrogen molecule and an iodine molecule. If we make the usual assumption that the geometrical requirements for the reverse reaction are the same as those for the forward reaction, then the only difference between the two reactions is in the activation energy term. Whichever of the two (forward or reverse reaction) has the higher activation energy, it will occur less often at a given temperature, since it will be less likely that two molecules will collide with sufficient energy to surmount the activation energy barrier from that side.

The principle of 'microscopic reversibility' is assumed to apply to all chemical reactions. It should be clear, therefore, that the likelihood of a given reaction occurring will depend on the probability of a suitable collision, that is, on the concentrations of the reactants, the configurational requirements for reaction, and on the kinetic energy of the colliding particles (i.e. on the temperature of the system). From this point of view, no reaction is irreversible. It may be very unlikely if the energy requirements are high or if the collision probability is low, but the principle of microscopic reversibility states that the energy requirement is never infinite, nor the collision probability ever zero, and thus there is at least a finite possibility that any chemical reaction will reverse itself. On the other hand the variations in configurational and activation energy requirements will clearly have a marked bearing on determining which chemical reactions occur with the greatest frequency, or which ones are most apparent to an observer.

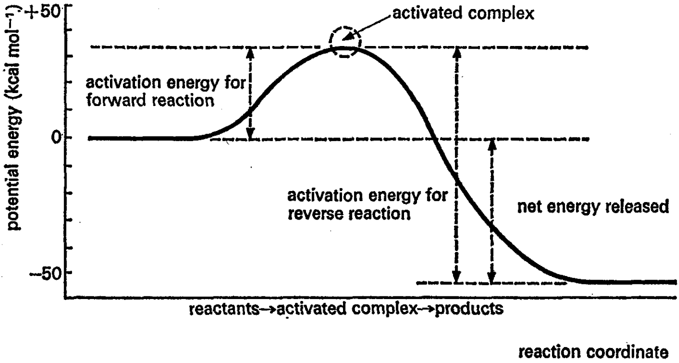


Figure 4 Potential energy diagram for reaction:
$CO + NO_2 \rightarrow CO_2 + NO$
(Equilibrium favours CO_2 and NO)

Catalysis

Many reactions are known whose rates can be changed by introducing further substances. Often these added substances, contrary to usual chemical experience, are not consumed during the net reaction. The amount of the added substance is the same after the reaction as it was initially. Introduction of finely divided platinum into a gaseous mixture of hydrogen and oxygen, for example, leads to the rapid formation of water. Yet the amount of platinum remains unchanged. The platinum has certainly entered into the reaction, but it has not been consumed. Such substances are called *catalysts*.

All catalysts operate by introducing a new path or mechanism for the reaction. Since the old mechanism will also continue, the net rate is always increased by a catalyst. In many cases the new mechanism leads to a much faster rate than the mechanism occurring in the absence of the catalyst. In such a case, the added substance is said to be a 'good catalyst'.

Since rate is determined by orientation effects and by the height of the activation energy barrier, good catalysts must affect at least one of these. In some cases both effects occur. Most commonly, the catalyst provides a new path with a lower activation energy. The lower activation energy allows a considerably more rapid reaction at any given temperature.

The chemical role of the catalyst is to form an intermediate compound, usually with a low activation energy of formation. The intermediate then reacts further to give the stable products of the reaction and regenerates the catalyst. Platinum catalyses the hydrogen–oxygen reaction by adsorbing the oxygen and considerably lowering the activation energy of the reaction with hydrogen molecules. The platinum surface is regenerated when the hydrogen reacts with adsorbed oxygen to give water.

Rate and Equilibrium

We have thus far developed the concept that all molecules are in rapid random motion with kinetic energies related to the absolute temperature, and with bond strengths which may range from very weak to very strong. The random collisions, when they occur with the proper geometry and sufficient energy to overcome the activation barrier, can lead to chemical reaction. Old bonds may be broken and/or new bonds may be formed. The rates of these chemical reactions will depend on the collision frequency, on the geometrical conditions for reaction and on the availability of the activation energy.

The principle of microscopic reversibility indicates that reverse reactions will also occur and that the rates of these reverse reactions will depend on the same effects.

Reactions which have few configurational requirements and a low activation energy will occur at almost every collision. The reactant concentrations will rapidly decrease. Conversely, reactions with restrictive configurational requirements or high activation energies will seldom occur, and the reactant concentrations will tend to remain large. In any given system all possible reactions will occur, but those substances which react least often on collision will tend to have the largest concentrations. Equilibrium is defined as the state reached when the rates of both 'forward' and 'reverse' reactions are identical due to an exact balancing between the activation energies, the configurational requirements and the concentrations.

Tendency to Equilibrium

The net result of the random collisions predicted by the kinetic theory is to bring a system to a state of dynamic equilibrium. At dynamic equilibrium, opposing reactions are occurring but with identical rates. Low energy collisions lead to simple rebounding, whereas high energy collisions may lead to chemical reaction. The relative concentrations of the various species present at equilibrium depend on configurational and activation energy requirements for their formation and reaction.

We might summarize these tendencies in somewhat different terms as follows: The fact that all atoms attract one another to a certain extent would tend to lead to the formation of a large number of chemical bonds around any given atom. On the other hand, once a few bonds have formed to a given atom, it becomes more and more unlikely that it will meet another atom with which it can form a strong bond. The actual number of bonds to a given atom will therefore tend to be such that the *average* bond strength is at a maximum rather than that the *total* bond strength about that atom would be the maximum.

For example, carbon can normally form four very strong bonds. Each succeeding bond becomes weaker and weaker. Thus methane (CH_4) in an atmosphere of hydrogen atoms might indeed form CH_5 molecules for a finite length of time. On the other hand, hydrogen atoms can bond so much more strongly to other hydrogen atoms than they can to methane molecules that the equilibrium situation would find mainly methane molecules and hydrogen molecules present rather than CH_5 molecules. But one must remember that in the equilibrium state there will still be a finite concentration of hydrogen atoms, that these will occasionally collide with methane molecules, and that there might be an occasional opportunity for a CH_5 molecule to exist for a very short period of time.

The existence of any hydrogen atoms at all, of course, is interpreted in terms of the occasional collisions of high energy molecules. During such a collision, a hydrogen molecule may split into two hydrogen atoms. These two hydrogen atoms may then undergo many collisions before they find the opportunity to recombine and form another hydrogen molecule. In the same way, methane molecules will occasionally be decomposed, by collision, into free hydrogen atoms and methyl radicals, or even into methylene (CH_2) and two hydrogen atoms. These latter species get increasingly rare as the number of hydrogens connected to a carbon decreases.

Dynamic equilibrium thus involves competition among all the possible bonding states for the limited number of atoms in the system. The random collisions lead to the formation of all possible molecules, but the equilibrium concentration of each kind of molecule will depend on the chance of its random formation and its bond energy with respect to other possible molecules.

Escaping Tendency

We may interpret the tendency to equilibrium in a slightly different fashion by introducing the phrase 'escaping tendency'. Because of the random kinetic energy, which depends on the temperature, every atom will have a finite opportunity or tendency to escape from wherever it may be into another state. If the atom is bound to a second atom, a high energy collision may free it. If the atom is free, a low energy collision may allow it to combine with another atom. A second low energy collision may let the molecule grow larger, etc. The random nature of the collisions and their very large energy spread will ensure that occasionally all possible combinations are realized, and each atom will spend part of its time in the free state, and part combined with other atoms in various combinations.

The percentage of time an atom spends in any given combination will be strongly dependent on the strength of the chemical bonds holding it in that combination, on the number of possible combinations and on the available kinetic energy (temperature). The stronger the bonds, the longer the time it will be in that particular combination. Therefore, it is most likely that molecules containing strong bonds will be found in any equilibrium system. On the other hand, the higher the kinetic energy (i.e. the higher the temperature) the more likely it is that even strong bonds will be broken. At very high temperatures, it will become extremely likely that a large number of free atoms will be present, and that only simple molecules such as diatomic ones with very strong bonds will be common. In this case collisions are so frequent and so energetic that more complicated molecules, or molecules with weaker bonds, have no opportunity to exist over any length of time.

At very low temperatures, on the other hand, not only will the strong bonds be able to form but also many of the weak ones. As a result methane, for example, will not only exist as CH_4 molecules but the molecules themselves will cluster together to form liquid methane. The weak bonds between the molecules are, at this temperature, sufficiently strong to hold the molecules close together. At very low temperatures, even helium atoms (which exhibit the weakest interatomic forces known) will cluster together to form liquid helium. At these temperatures escaping tendencies are low and high concentrations of complicated aggregations of atoms exist.

The escaping tendency is strongly affected by the temperature. It increases with increasing temperature due to the rapid rise in average kinetic energy of the molecules. Likewise, the escaping tendency increases with concentration. Increasing the concentration of a particular species, in a gas for example, increases the likelihood of its colliding with another reactive substance and of undergoing reaction. Conversely, the formation of strong bonds decreases the escaping tendency of a substance. For example, at low pressures and temperatures, diamond slowly converts to graphite since the bond strength in graphite is, on the average, somewhat stronger than in diamond.

Net reactions will occur until the relative concentrations of each kind of atom, in all the molecular species present at equilibrium, are such that the escaping tendency of that element is the same everywhere. This equilibrium escaping tendency will have a minimum value since substances will always move from regions of higher to regions of lower escaping tendency.

Chemical Systems

We have now answered the question 'Why do chemical reactions occur?' in terms of the kinetic theory, random collision processes, varying bond strengths and various possible states. We have used qualitative description almost entirely. We certainly have not discussed how the decomposition of wood and the synthesis of wood can both occur, or how the rusting of iron and the reclamation of iron from iron oxide also both occur. Clearly we must do more than talk in generalities concerning the tendency to equilibrium, if we are to discuss intelligently such reactions which are clearly going in opposite directions and towards obviously different final states. In order to be more quantitative and to treat such systems with some success, we shall use the language of thermodynamics.

A careful comparison of the conditions on the inside and outside of the blast furnace, or in the rotting and growing trees, will show that the concentrations are comparable. In the case of the rotting and growing trees, the temperatures are also comparable. It is certainly true that the bond strengths are the same, since the same molecules are involved. How then can the escaping tendencies be so different that in one case the wood is reacting with oxygen to produce carbon dioxide and water, and in the other case the carbon dioxide and water are reacting to produce oxygen and wood?

A close study of the two trees will indicate that at least one difference has to do with the energies involved in the reactions. Calorimetric measurements will show that the rotting wood is slowly evolving energy to the outside world, and the growing tree is absorbing energy from the surrounding sunlight. If we apply this observation to the iron system, we notice the same thing. The rusting iron is evolving energy to the rest of the world, and the blast furnace reaction is absorbing energy as it converts iron oxide to iron. The conservation laws, of course, lead us to believe that the total energy of the universe is constant, but it is clear that the energy contents of those portions of the universe in which we are here interested are changing. Before we concentrate on the energy effects, let us firstly develop the concept of a chemical system.

It is convenient to define three types of systems:

1. *Isolated systems.* These are completely uninfluenced by any external effects. In particular, they cannot exchange energy or matter with any other systems. The universe is normally treated as an isolated system.
2. *Closed systems.* These cannot exchange matter with other systems, but can gain and lose energy. A sealed tube alternately placed in hot and cold surroundings would be an example of a closed system. It would gain and lose energy, but the mass of its contents would remain unchanged.
3. *Open systems.* These can exchange both matter and energy with other systems. A beaker of boiling water would be an example. The beaker can gain energy from the flame, lose it by evaporation and lose mass by evaporation. The rusting iron, the blast furnace, and the rotting and growing trees would be treated as open systems.

We normally describe these systems in terms of a set of variables, preferably those which are readily measured and easily manipulated. The most common variables used to describe chemical systems are pressure, volume, temperature, mass and energy.

The Laws of Thermodynamics

The very earliest developments in thermodynamics led to two general conclusions. These are:

1. *The idea of temperature equilibrium.* To be exact, if substance one is in thermal equilibrium with (has the same temperature as) substance two, and substance two is in thermal equilibrium with substance three, then substance one is also in thermal equilibrium with substance three.
2. *The law of the conservation of energy.* The total energy content of any isolated system is constant.

A third major generalization was discovered almost 150 years ago. It is most unfortunate that this generalization is not widely understood and is unused by many chemists. A simple and reasonably exact statement of this principle is that 'all processes in isolated systems of constant volume tend to decrease the order and increase the disorder in the system'. We have already cited one example of this, due to Maxwell, when we pointed out that even should a set of molecules in the gas phase instantaneously have the same individual energies, they would very quickly redistribute these energies to follow the Maxwell–Boltzmann curve shown in Figure 1. The system would move from the ordered arrangement, where every molecule had the same energy, to the more disordered arrangement in which the energies of the molecules differed.

The principle was also invoked implicitly when we interpreted chemical reactions as resulting from the random collision of molecules, and the tendency to a dynamic equilibrium in which all possible configurations were represented, rather than only the most strongly bonded one.

Entropy

This randomness factor can be measured quantitatively and is given the name *entropy*. The entropy change in a given closed or isolated system can be most simply defined as

$\Delta S = k \log_e \left(\frac{\text{disorder in state 2}}{\text{disorder in state 1}}\right)$, where $k \left(= \frac{R}{N_A}\right)$ is the Boltzmann constant[1]. It is not possible for us here to develop the theory of disorder completely, nor can we here apply the entropy concept to all chemical systems. But we can develop a few examples to show its power.

Let us firstly consider the isolated system of gas shown in Figure 5. No energy or mass can leave the system, which consists of two equal volumes separated by a glass partition. The left-hand volume contains one mole of gas. The right-hand volume is completely evacuated. If the barrier is now shattered by the ball,

---

[1] R is the gas constant, N_A is the Avogadro constant.

the gas proceeds to fill the whole container. The system is isolated, hence the change cannot be interpreted in terms of gain or loss in energy; but it can be interpreted in terms of a change of entropy. In state 1, it was certain that every gas molecule would be in the left-hand side of the container. Disorder does exist, however, in the gas. In state 2, the disorder is twice as great since the available volume for gaseous motion is twice as great. It can be shown that the change in entropy is $\Delta S = R \log_e 2$. The system has moved from a more ordered to a more disordered arrangement. The entropy has increased.

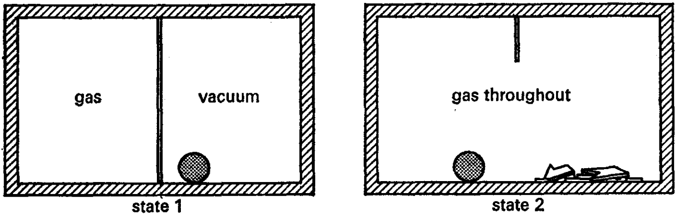


Figure 5 Isolated gas system

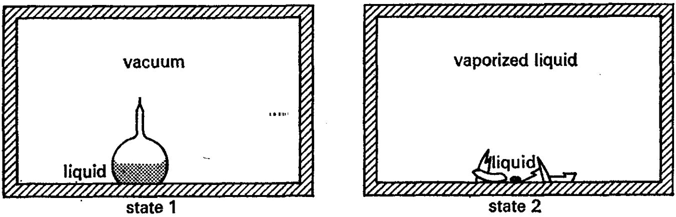
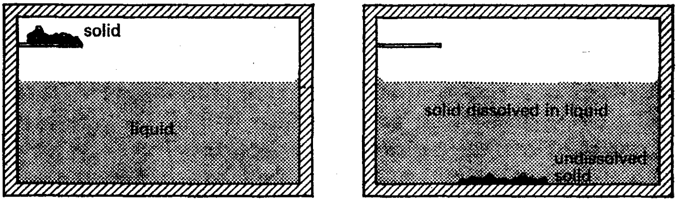


Figure 6 Vaporization of liquid in isolated system

Let us now consider a drop of liquid in a glass bulb mounted in an evacuated isolated chamber as shown in Figure 6. If the glass bulb is broken, the liquid immediately begins to evaporate until the vaporized molecules exert a pressure equal to the vapour pressure of the liquid. Again the total energy of the system is constant, but the drop of water cools off since some of its kinetic energy must go into evaporating liquid. Thus, the drop of water spontaneously loses energy. But again note that the entropy of the system has increased. The disorder in the close molecular arrangements found in a drop of water is considerably less than that found in the gas. In the gas, any given molecule can be found anywhere in the container. Gas disorder, as represented by freedom of motion, is much greater than liquid disorder. The complete calculation of the entropy change is not as simple here as in the example above because not enough is known about the disorder of the various states which the molecules might have in the droplet of liquid.

For a third example, consider Figure 7. Here again we have an isolated system which may be manipulated to drop the crystals of solid into the liquid. The solid spontaneously dissolves. Again the entropy of the system increases. It is easy to see that the entropy will increase, since the distribution of the dissolved particles in the solution is far more disordered than their distribution in the regular crystal packing.



state 1

state 2

Figure 7 Solution of a solid

From these, and many other examples, it has become possible to find a perfectly general criterion for change in any isolated system of constant volume. The criterion is that in any total process occurring in such a system, the entropy of the system will increase. This is to say that such a system is always more disordered after it has changed than it was before the change occurred. We can even say that change occurs because of the natural tendency to increase disorder. But note, the universal increase in entropy is only characteristic of an isolated system of constant volume.

It should be clear that this may allow us to predict which changes will occur if we can decide in advance which is the more ordered and which the more disordered state. In a constant volume isolated system, only changes which take the system toward a more disordered state will occur.

Since the universe is normally considered an isolated system of constant volume, we have a simple criterion for all total changes which we observe in the universe. All such changes result in an increase of disorder in the universe as a whole. We must remember, however, that it is possible to treat small portions of the universe as closed systems or as open systems. The entropy change in a closed system or an open system may be either positive, negative or zero, but the entropy change in the whole universe, including the system we are considering, must always be positive. The universe must always go from a more ordered to a less ordered arrangement in any change.

Criteria for Change

A few criteria for change are listed in Table 1. The first column gives the type of system. The second column lists the conditions held constant during the change under consideration. The third column gives the criteria for the direction of the change.

Chemists operate with systems fitting into all the categories shown, plus other categories which are not described in the table. A very common set of conditions in the laboratory are those of constant pressure and constant temperature. For these conditions, it will be seen that the table lists the criterion that the *free energy* of the system will always tend towards a minimum. For our purposes, the change in free energy in a closed system at constant temperature and pressure may be defined by the equation

$$\Delta G = \Delta H - T\Delta S.$$

ΔH, called the change in heat content, is the quantity of heat which would be measured if the change were carried out in a calorimeter at constant pressure and constant temperature. T is the Kelvin temperature, and ΔS is the change in entropy, or order, defined as above. ΔG, the change in free energy, is the difference between the heat and entropy (times temperature) term. Free energy can also be described as the maximum work which the system can do during the reaction, exclusive of the work which is due to any volume change. We have now developed several thermodynamic ideas. Let us clarify them by discussing some particular examples.

Table 1
Criteria for Change in a Few Types of Systems

System	*Constant*	*Criterion for direction of change*
Isolated	Energy and volume	Disorder (entropy) goes to a maximum. ΔS is positive
Closed	Pressure and temperature	Free energy goes to a minimum. ΔG is negative
Open	No simple general criteria available. In general, system does not reach equilibrium. Change occurs as long as the system is open	

Changes in Isolated Systems

According to Table 1, changes in isolated systems at constant volume always occur in such a direction as to increase the disorder in the system. Since both the energy and the mass content of an isolated system are constant, we can only have energy flow from one place to another, or mass movement from one place to another occurring. When either or both occur the net process must result in a decrease in order within the system.

Thus, in an isolated system, heat can only flow spontaneously from a hotter to a colder region and never the reverse. The energies of the hotter and colder systems become mixed and will not spontaneously 'unmix'. Less order (in terms of the separation of 'hot' and 'cold' molecules, for instance) exists than before.

Order is also associated with molecular structure. Crystals are more ordered than liquids, and liquids more than gases. Polyatomic molecules are more ordered than the simpler molecules from which they may be formed. Thus, if no energy change is involved, spontaneous processes in nature always lead to the decomposition of ordered structures into structures of less order. Buildings tend to fall apart, solids tend to dissolve in liquids, or to evaporate as gases, or to decompose in other ways, for example they rot.

In most actual processes, changes occur as a result of both energy flow and shifts in the order of chemical species. Thus rotting normally is exothermic and also leads to the separation of simple molecules. Both the transfer of energy to other parts of the system and the formation of simpler molecules are spontaneous processes, and the rotting will proceed to completion as long as the energy can leave the rotting log and, together with the other products of the rotting process, enter the rest of the system.

On the other hand, a tree may absorb the high energy of sunlight in a spontaneous fashion and with an apparent large increase in order. There is still a net loss of order since more order is lost through the irreversible process of heat transfer than is gained by photosynthesis in the leaf (the reaction of carbon dioxide and water to give complicated plant material and oxygen). As long as the disorder resulting from the heat transfer is greater than the order resulting from the molecular change in the leaf, the over-all result is a decrease in order and the process is spontaneous.

We see then that chemical reactions in isolated systems will occur if the over-all effect of all energy and molecular change is such as to create net disorder.

1. Any reaction which is both exothermic and creates simpler molecules is probable.
2. Any reaction which is both endothermic and creates more ordered molecules is improbable.

The other combinations: (3) exothermic reactions which create more ordered molecules, and (4) endothermic reactions which create simpler molecules will be probable if there is a net decrease in order as a result of counterbalancing the energy and molecular order effects.

We must talk about the probability of reaction, because the random nature of molecular collisions and the Maxwell–Boltzmann distribution of molecular energies make all reactions possible. A probable reaction is one which occurs to a readily measurable extent.

Here are some examples of each type of reaction above:

1. *Exothermic reactions giving simpler molecules* will always occur: for example combustion of most fuels made of complicated molecules such as wood, coal, petroleum.
2. *Endothermic reactions giving more ordered molecules* only occur when a high energy source is available in which the loss of order is greater than the gain in order in the endothermic part of the system: for example photosynthesis.
3. *Exothermic reactions which create more ordered aggregates:* for example most precipitation reactions, condensation of gases to liquids, and liquids to solids. These only occur when energy loss to other parts of the system creates enough disorder there to balance the gain in order within the exothermic portion of the system.
4. *Endothermic reactions which create disorder* in the same part of the system: for example evaporation of a liquid or solid, most reactions involving formation of a solution, most thermal decompositions. These only occur when the molecular disorder in this part of the system is large enough to counterbalance the energy gain.

Examination of the above will show that exothermic reactions are probable unless there is a large counterbalancing effect of gain in molecular order. Endothermic reactions are, on the other hand, improbable unless there is a large counterbalancing effect due to loss in order. Most observed chemical reactions are therefore exothermic.

Yet it should be equally clear that it is impossible to predict the probability of a reaction from energetics alone. Molecular order must also be considered.

Changes in Closed Systems at Constant Temperature and Pressure

Changes in closed systems are somewhat easier to visualize than those in isolated systems since one need not concern oneself with the energy flowing within the system, but only with its net flow in and out of the system. At constant pressure, the energy change in a system is known as ΔH, positive if energy is gained by the system.

The general relation in a closed system at constant T and P is $\Delta G = \Delta H - T\,\Delta S$, where processes with a negative value of ΔG are probable and those with a positive value of ΔG are improbable.

Considering only the ΔH factor we see that exothermic reactions are probable and endothermic reactions are improbable. Furthermore, as $T \longrightarrow 0$ only the ΔH term is important. Thus at very low temperatures the direction of a reaction is uniquely determined by ΔH.

Now T is always positive so that the sign of ΔS determines the effect of the $T\,\Delta S$ term on ΔG. If ΔS is positive the reaction becomes more and more possible as the temperature rises (assuming ΔH does not change with changing T). Similarly, if ΔS is negative the likelihood of reaction decreases with rise in T. At very large values of T, the $T\,\Delta S$ term uniquely determines the likelihood of reaction.

For most reactions near room temperature $T\,\Delta S$ is small compared to ΔH. Thus at constant temperature and pressure in a closed system, exothermic reactions are probable and endothermic reactions are improbable.

At high temperatures the sign of ΔS is determining. ΔS will be, in general, positive if larger aggregates are decomposing into small aggregates. Thus, for $H_2O \longrightarrow H_2 + \frac{1}{2}O_2$, ΔS will be positive. At high temperatures this reaction will occur to a greater degree than at low temperatures.[1]

ΔG may be calculated, for example, from a knowledge of ΔH, ΔS and their variation with temperature. The sign of ΔH may be predicted from a knowledge of bond strengths. If the bond strengths of the product molecules are stronger than those of the reactant molecules ΔH will be negative. Thus in the reaction $NH_4NO_2 \longrightarrow N_2 + 2H_2O$, both N_2 and H_2O are known to have strong bonds. Thus ΔH is probably negative and the

---

[1] Note that we can extrapolate the last paragraph to indicate that at very high temperatures all complicated molecules will cease to exist and there will be no probable chemical reactions among the atomic fragments to form larger molecules. Thus almost no atomic aggregates will be found at temperatures of 100000K. At this and, yet more so, at higher temperatures, even the atoms disintegrate into electrons and ions due to the high collision energies, so that at 10^6K one has largely nuclei and electrons present in a gas. At still higher temperatures nuclei disintegrate into protons and neutrons.

reaction shown will be probable (though it may be very slow if the activation energy is high) at low temperatures.

Similarly, the sign of ΔS may often be guessed by counting the moles of reactants and products. Since ΔS is positive for a decrease in order, reactions producing the most molecules will probably have a positive sign for ΔS. In the reaction $NH_4NO_2 \longrightarrow N_2+2H_2O$ we would guess ΔS to be positive. Thus $T\Delta S$ is positive and the reaction becomes more and more probable as the temperature rises since $\Delta G = \Delta H - T\Delta S$ and a negative ΔG indicates a probable reaction.

For the reaction $H_2+\frac{1}{2}O_2 \longrightarrow H_2O$, ΔH is negative but ΔS is probably also negative. Therefore this reaction will be probable at lower temperatures, but will become less and less probable as the temperature is raised.

For the reaction $CO+N_2O \longrightarrow CO_2+N_2$ one would guess a negative ΔH since CO_2 and N_2 bonds are so strong. ΔS should be approximately zero. Thus the reaction should have a negative ΔG value at all temperatures and should be a probable reaction. We must of course remember that the reactions will only occur with an appreciable rate if the activation energy is low compared to the average kinetic energy of the molecules.

The sign of ΔG indicates the direction towards equilibrium, while the magnitude of ΔG indicates the distance from equilibrium. Thus if $\Delta G = 0$ the reaction is at equilibrium, whereas if ΔG has a large negative value a good deal of reaction will occur before equilibrium is reached. However, the rate of reaction has no correlation either with the sign or magnitude of ΔG and cannot be predicted from a knowledge of ΔG.

For the reaction $CaCO_3(s) \longrightarrow CaO(s)+CO_2(g)$ we would predict a positive value of ΔH since $CaCO_3$ is tightly bonded, and a positive value for ΔS since a mole of gas and a mole of solid are produced from one mole of solid. Thus the equilibrium will favour $CaCO_3(s)$ at low temperatures and gaseous CO_2 and solid CaO at higher temperatures.

Similar arguments would hold for reactions such as:

$$NH_4Cl(s) \longrightarrow NH_3(g)+HCl(g)$$
$$ZnCl_2(s)+H_2O(g) \longrightarrow 2HCl(g)+ZnO(s).$$

Changes in Open Systems

An open system differs from an isolated or a closed system in that substances can enter or leave. It is impossible in a closed system to have any reaction proceed to completion since the random collisions between the products will occasionally regenerate the reagents. Reactions may be even less complete in a small isolated system, since energy is also conserved and exothermic reactions cannot 'get rid of' the energy they produce and so go further to completion.

In an open system, on the other hand, it is possible in principle to have reactions proceed to completion. Energy can escape to the surroundings thus making a reversal of the reaction through a high energy collision more unlikely. Furthermore, the products themselves may be dispersed or separated so that they cannot collide and re-form the reactants.

Consider again the reaction

$$CaCO_3(s) \longrightarrow CaO(s) + CO_2(g).$$

In a closed system the reaction will proceed to an equilibrium condition where $CaCO_3$ is decomposing at the same rate that CaO and CO_2 are reacting upon collision to form $CaCO_3$. If the CO_2 is swept out of the furnace, say by a current of air, its reaction with CaO becomes impossible and all the $CaCO_3$ will decompose.

It is for this reason that many chemical reactions are carried out under open system conditions, especially if the equilibrium conditions are not favourable to the desired reaction. Thus coal is burned in an open fire box, cement is formed in an open heater, sea water is evaporated in an open pond to obtain salt, blast furnaces operate with a steady throughput of fresh air, petroleum is converted to gasoline in a continuous flow process and ammonia synthesis is accompanied by continual removal of the ammonia in a cyclic process.

Under conditions of an open system chemical reactions may be irreversible, they can 'go to completion', but solely because some of the products are removed and are no longer present to react. In isolated systems, chemical reactions are irreversible solely because energy will not spontaneously flow from a cold to a hot place, not because the molecules cannot react. The reactions do not 'go to completion'. In closed systems, reactions may be reversed simply by reversing the energy flows and they are reversible in any case except at absolute zero. In open systems, reactions may go to completion. But here too they may be reversed. Thus $CaCO_3$ decomposes into CO_2 and CaO completely if heated in such a way that the CO_2 escapes. In principle, the reaction can be completely reversed if the system is cooled and the pressure of carbon dioxide is maintained above the equilibrium value.

The fact that the complete reversal of some reactions is not feasible in closed or open systems depends either on a slow rate or the difficulty of separating one of the reactants or products from the rest of the system, rather than on a fundamental property which can be called 'reversibility' or 'irreversibility'.

We now see that the blast furnace which converts iron oxide to iron, and its outer surface which converts iron to iron oxide, are examples of open systems. On the inside, coke is introduced and oxides of carbon removed so that practically all the iron oxide is converted to iron. On the outside, iron reacts with oxygen to form the oxide. The pressure of oxygen in the atmosphere is far above that which would exist at the equilibrium between iron and iron oxide. Essentially all the metal would eventually be converted to oxide. Only occasionally would an energetic region of the oxide crystal eject some oxygen and revert to iron.

Actually, both growing tree and rotting tree can best be considered as open systems. The growing tree ingests carbon dioxide, water and energy (among other things) and eliminates oxygen while producing tissue. The rotting tree reacts with oxygen liberating energy, carbon dioxide and water (among other things). The principal difference is that energy is gained by the growing tree and lost by the rotting one, but the difference in processes is also partially due to the means each system has of accumulating or eliminating substances such as carbon dioxide, water and oxygen. For example, the rotting proceeds until no wood is left, since the products of reaction (including energy) escape and their concentration never rises to the level where equilibrium could be attained.

Chemical reactions occur, then, because molecules collide in a random fashion, and because no chemical bond is either infinitely strong or of no strength whatever. The substances most likely to form are those with the strongest bonds, but the random nature of collision processes makes the decomposition of even these substances possible.

Actual equilibrium states may generally be described in terms of energy gain and loss, and gain or loss in order.

For the most common type of reactions, those that occur at constant temperature and pressure, the relationship is:

$$\Delta G = \Delta H - T\,\Delta S$$

where ΔH is the heat gained by the system during the change, T is the Kelvin temperature and ΔS is a measure of the increase in disorder during the change. A negative value of ΔG for any change indicates the change is likely; the products will be present in high concentration at equilibrium. A positive value of ΔG indicates the change is unlikely; the reactants will be present at high concentration at equilibrium. Thus exothermic processes leading to a decrease in order are most favoured, and endothermic processes leading to an increase in order are least favoured as possible reactions.

Bibliography

J. A. ALLEN, *Energy Changes in Chemistry*, Blackie, 1965.

P. G. ASHMORE, *Principles of Reaction Kinetics*, 2nd edn, Royal Institute of Chemistry, 1968.

G.C. BOND, *Principles of Catalysis*, 2nd edn, Royal Institute of Chemistry, 1969.

J. A. CAMPBELL, *Why Do Chemical Reactions Occur?*, Prentice-Hall, 1965.

A. A. FROST and R. G. PEARSON, *Kinetics and Mechanism*, 3rd edn, Wiley, 1961.[1]

E. L. KING, *How Chemical Reactions Occur*, Benjamin, 1963.

B. H. MAHAN, *Elementary Chemical Thermodynamics*, Benjamin, 1963.

R. S. NYHOLM, 'The stability of inorganic compounds', *Modern Chemistry and the Sixth Form*, Collins, 1965.

L. E. STRONG and W. J. STRATTON, *Chemical Energy*, Chapman & Hall, 1966.

[1] More advanced.

7 J. Dwyer and J. Lee

Thermochemical Cycles

J. Dwyer and J. Lee, 'Thermochemical cycles', *Chemistry Student*, vol. 2 (1968), no. 2, pp. 46–50, and no. 3, pp. 63–7.

Path-Dependent and Path-Independent Quantities

Some of the key principles of physical science may be described in terms of non-scientific analogy. Consider a multi-storey building, comprising say twenty floors, and suppose that the floors are interconnected by staircases (fire regulations normally necessitate two staircases between adjacent floors) and also by a lift. Each floor has a particular height (above sea level); this is a characteristic feature of the floor. For any pair of floors, say the twentieth and the first, there is a particular *difference* in height. It is quite immaterial whether we walk up one set of staircases or another, or whether we take the lift, our *change of elevation* will be the same – such a quantity is said to be *path-independent*. On the other hand, a quantity which will obviously depend upon the method of ascent will be 'degree of fatigue' (if this could be assigned a quantitative measure). We should feel less tired if we used the lift than if we walked up the steps. The degree of fatigue is *path-dependent*. Notice that although the change of elevation, $\Delta e_{20,1}$ may be expressed as a difference of two heights, one characteristic of the higher (twentieth) floor (e_{20}) and the other characteristic of the lower (first) floor (e_1), the same cannot be said for the degree of fatigue. There is *no* 'fatigue level' for each floor (f_{20}, f_1). One can never write: 'degree of fatigue' on ascending from the lower to the higher floor $= f_{20} - f_1$; if one could, the 'degree of fatigue' would be path-independent. The quantities e_{20} and e_1 (and the other eighteen *e*s) might be termed '*properties of floors*'.

In the science of thermodynamics (a corner-stone of modern chemistry), both path-dependent and path-independent properties are encountered. For example, consider a cylindrical vessel containing exactly one mole of a gas and fitted at one end with a piston – very much like one cylinder of a motor-car engine, except that there are no valves or sparking plug. Let the *pressure*, *volume* and *absolute temperature* of the gas be P_1, V_1 and T_1, respectively. Consider two different ways of changing this *state* to a new state with corresponding quantities P_2, V_2 and T_1, that is, the same temperature but different pressure and volume. The two *paths* are shown in Figure 1 and are designated I and II. *Irrespective of path*, the change of pressure, $\Delta P = P_2 - P_1$, and the change of volume, $\Delta V = V_2 - V_1$; ΔP and ΔV are path-independent. P_1 and V_1 (and T_1) are characteristics of the initial state of the gas, while P_2 and V_2 (and T_1) are characteristics of the final state; pressure and volume are *properties of state*. Conversely, the *work* performed by the moving piston (say in driving round a crankshaft) is quite different for the two paths I and II. It may be shown that the work performed for path I is $RT_1 \log_e(V_2/V_1)$, while for path II it is $RT_2 \log_e(V_2/V_1)$, R being the gas constant; it is assumed here that the gas satisfies the perfect gas equation ($PV = RT$ with V the *molar volume*) and that there are no frictional losses for the piston. For path II, there is zero work associated with the first and third parts, that is, with the changes at constant volume.

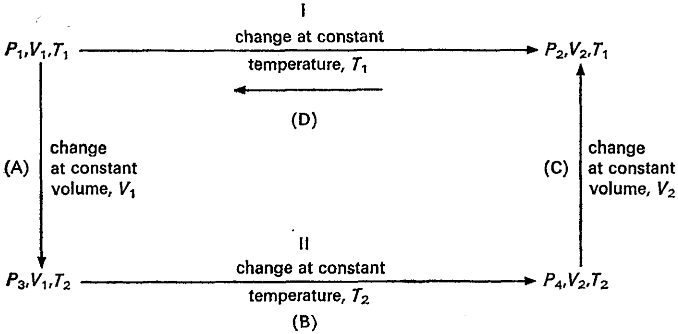


Figure 1 (Note: path II consists of changes A, B and C)

An alternative, but equivalent, approach to the consideration of properties of state is in terms of *cycles*. This may be exemplified by transforming the state characterized by P_1, V_1, T_1 to that associated with P_2, V_2, T_1 by route II and then returning to the original stage by the reverse of route I. The *over-all* change in any property of state for such a cyclic process will be zero. For example, the over-all pressure change is zero. Considering the process in the four stages indicated in Figure 1:

$$0 = \Delta P_{\text{over-all}} = \Delta P_{\text{A}}+\Delta P_{\text{B}}+\Delta P_{\text{C}}+\Delta P_{\text{D}} \qquad \mathbf{1}$$

where $\Delta P_{\text{A}} = P_3-P_1$, $\Delta P_{\text{B}} = P_4-P_3$,
$\Delta P_{\text{C}} = P_2-P_4$ and $\Delta P_{\text{D}} = P_1-P_2$.

If one of the four ΔPs represented here were for some reason difficult to measure, it could be deduced by equation **1**, for example

$$\Delta P_{\text{A}} = -(\Delta P_{\text{B}}+\Delta P_{\text{C}}+\Delta P_{\text{D}}),$$

with the right-hand side of the equation being determined by experiment. In the architectural analogy, one could start at the ground floor (here, in accordance with American rather than British usage, designated the first floor), take the lift to the twentieth floor and then descend to the ground floor via the stairs to the original starting point. Then

$$0 = \Delta e_{\text{over-all}} = \Delta e_{20,1}+\Delta e_{19,20}+\Delta e_{18,19}+\ldots+\Delta e_{1,2} \qquad \mathbf{2}$$

with $\Delta e_{i,j} = e_i-e_j$ generally.
If all the Δes *but one* were experimentally measurable, the other could be calculated from equation **2**.

Thermochemical Cycles

There are several properties of state which are of particular interest in thermodynamics. Four of these have the dimensions of energy: the *internal energy*, U; the *enthalpy*, H; the *Helmholtz function*, A; and the *Gibbs function*, G. The concept of a cycle may be applied to any one of these quantities, but the present article will be concerned only with the first two of these. Unlike the exclusively physical processes involved in the four-stage cycle of Figure 1, internal energy and enthalpy are generally applied to situations involving some (but not necessarily only) chemical change. Two common examples are presented in Figure 2. In each case, there is interest in determining the change

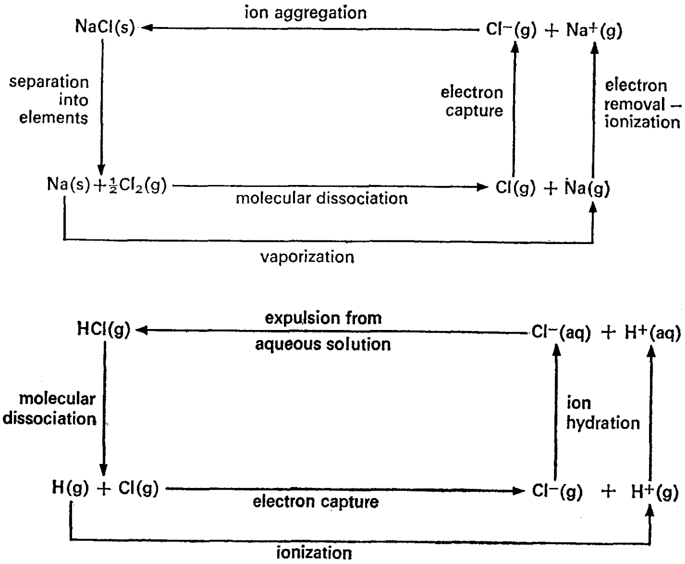


Figure 2

in internal energy or the change in enthalpy for *one* stage of the cycle by use of equations analogous to **1** and **2**.

The internal energy may be conceived in two parts:

(i) the energy of individual molecules, being itself a composite of the energies of moving atomic nuclei and moving electrons, and

(ii) the energy of interaction between individual molecules.

It is well known that the zero of energy may be chosen arbitrarily but it is conventional here to take the energy relative to that of the constituent elements in their normal states at 298·15 K and at a pressure of one atmosphere. The internal energy of a state is *exactly* proportional to the amount of material, for example the energy of 2 g of a substance is exactly twice that of 1 g of it, assuming no change in other properties such as temperature and pressure. Such a property is known as an *extensive* one. In contrast, *intensive* properties are independent of the amount of material; examples of these are temperature, pressure, density (and, in our architectural analogue, the height of a floor, which is independent of the floor area). The internal energy associated with one mole of material will be represented (in accordance with common convention) by the corresponding lower-case letter, u. U and therefore u are dependent upon the temperature and the pressure. The first derivative of u with respect to temperature is the 'molar heat capacity at constant volume' of the material. The dependence on pressure is more complicated involving both the *volume coefficient of expansion* and the *compressibility coefficient*; when the pressure is one atmosphere (and normally for a pure compound), the state is referred to as a *standard* one and the corresponding internal energy symbolized U°. Neither dependence is particularly great but the tendency of some authors to neglect these in consideration of cyclic processes is to be deprecated; in other words, care should be taken to specify the temperature and pressure in the various states. In the case of condensed (that is liquid and solid) phases, however, the dependence of U upon pressure is small and is frequently neglected for common ranges of pressure. One should also specify carefully the physical characteristics of the states, for example (a) solid, liquid, gas, (b) mixture or separated components, (c) mesomorphic forms.

Enthalpy, H, is defined simply as $U+PV$, V being the *total* volume. Again, the property is (a) proportional to the amount, h being employed to symbolize the *molar enthalpy*, and (b) dependent upon temperature and pressure, the first derivative of h with respect to pressure being the 'molar heat capacity *at constant pressure*'. The same caution is strictly necessary when considering enthalpy as with internal energy in cyclic processes. For condensed phases especially, the difference between internal energy and enthalpy is comparatively small; the difference is more important for gases where molar volumes are comparatively large. Whatever the phases are, however, care should be taken not to mix the two properties H and U in a cycle, that is, equations analogous to **1** and **2** should not involve a mixture of ΔH and ΔU for the various steps unless it is properly established that *corresponding* ΔH and ΔU values are equal to a satisfactory degree of approximation.

Experimental Methods

There are several ways in which ΔU and ΔH may be determined for the various steps of a *chemical* cycle. Some of the more important will now be described.

Thermochemical method (*a*)

In fact, this type of method is itself an application of a cyclic process. It involves the use of some form of calorimeter. Although there is considerable variation in the detailed design of such devices, many operate approximately as so-called *isolated* systems in which there is no exchange of material or of heat with the surroundings outside the calorimeter, and for which no work is performed on the surroundings. Under such conditions, operation of the principle of energy conservation[1] ensures the constancy of the *internal* energy of the calorimeter and its contents. Suppose that we commence with n_A moles of a species A at a temperature T_A and a pressure P_A, and that A is then caused to change into n_B moles of B at a temperature T_B and a pressure P_B. Then

$$U[n_A \text{ moles of } A(T_A, P_A) + \text{calorimeter } (T_A)] = U[n_B \text{ moles of } B(T_B, P_B) + \text{calorimeter } (T_B)]$$

that is, $\Delta U = 0$ for the process. (In general, there will be a pressure gradient across the calorimeter walls, but since the variation of internal energy of a metallic solid with pressure is small, this has been neglected here.) Suppose, however, that ΔU is required for the chemical process

$$n_A \text{ moles of } A \longrightarrow n_B \text{ moles of } B,$$

where the state of A is as above but where the state of B is characterized by a temperature T_B and a pressure P'_B. We require ΔU for the process:

$$n_A \text{ moles of } A(T_A, P_A) + \text{calorimeter } (T_A) \longrightarrow n_B \text{ moles of } B(T'_B, P'_B) + \text{calorimeter } (T_A).$$

---

[1] The first law of thermodynamics is a special form of this principle. For a system not exchanging material with the surroundings, the change in the internal energy of the system is equal to the heat absorbed from the surroundings minus the work done by the system on the surroundings.

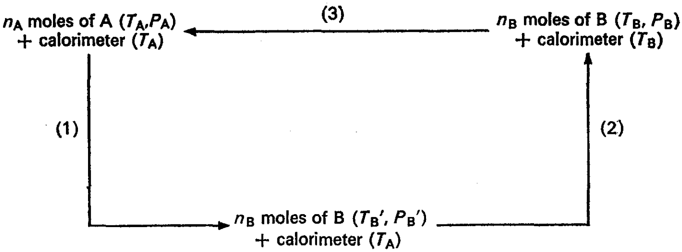


Figure 3

The calorimeter is necessarily in the same state as initially, so that ΔU is entirely associated with the change of calorimeter *contents*. Consider now the three-stage cycle indicated in Figure 3. Then because U is a function of state:

$$0 = \Delta U_{\text{over-all}} = \Delta U_1+\Delta U_2+\Delta U_3 = \Delta U_1+\Delta U_2 \qquad 3$$

(since ΔU_3 is zero) so that $\Delta U_1 = -\Delta U_2$. We can deduce ΔU_2 from a knowledge of (i) specific heat data for the calorimeter, and (ii) specific heat, expansion coefficients and compressibility data for the material B. In a few cases, the contribution of B to ΔU_2 has been tabulated, particularly for $P_B = P_B' = 0$ or 1 atm. The principles involved here are easily extended to the case of two or more reactants (and/or products). For example, a hydrocarbon may be completely oxidized in an excess of oxygen or sodium may be reacted with an excess of chlorine. In these cases, a *known* amount of electrical energy may be necessary to initiate the reaction; this energy appears ultimately as internal energy in the products and a correction must be applied for it. In other cases, mere mixing of reactants may be sufficient to produce reaction. In addition to chemical processes, certain physical processes such as *solution* may be investigated by this means.

If ΔH is required for a transformation typified by stage 1 of Figure 3, an equation analogous to 3 must be employed, thus

$$0 = \Delta H_{\text{overall}} = \Delta H_1+\Delta H_2+\Delta H_3$$

whence $\quad \Delta H_1\,[=\Delta U_1+\Delta(PV)_1] = -(\Delta H_2+\Delta H_3)$.

ΔH_2 may be deduced in a similar fashion to ΔU_2, while ΔH_3 may be readily calculated (by using the definition of H) as follows:

$$\Delta H_3 = \Delta U_3+V(P_A-P_B) = V(P_A-P_B),$$

V being here the volume of the calorimeter, approximated as a constant ($T_B - T_A$ normally being small enough to permit this).

Spectroscopic method (*b*)

The increase in energy required to accomplish certain *molecular* processes may often be provided by electromagnetic (light) radiation. For this purpose, light is to be regarded as a beam of particles or photons each having energy $h\nu$, where ν is the frequency of the light and h is a fundamental constant, Planck's constant. A *single* molecular energy increase is provided by the absorption of a single photon, the energy of which may be calculated from the frequency of radiation required to produce the increase. The occurrence of particular energy increases and the radiation frequency required is indicated by spectroscopy. In some cases, applying particularly to atomic systems, we are concerned with *decrease* of energy and a corresponding *emission* of light, but the principle of energy conservation may still be applied and the molecular transition energy calculated. Such a transition energy is essentially the change in internal energy *per molecule*. From this, the calculation of the change for the whole assembly of molecules, for example for one mole, is readily accomplished, provided that there is no intermolecular contribution. This is the case at a limiting pressure of zero. There are two kinds of processes which are readily investigated by these methods:

1. Ionization of an atom or a molecule. This corresponds to the removal of one or more electrons from the atom or molecule; usually these are the least tightly bound electrons.
2. Dissociation of a molecule. This corresponds to the breaking of a chemical bond to form two fragments. These may be atoms, free radicals or stable molecules; they may be electrically neutral or they may be charged, that is ionic. But spectroscopic observations normally identify the process.

The same type of information can often be obtained by providing energy to molecules with bombarding electrons (rather than bombarding photons) of known energy; this is the field of *electron impact*, which requires a form of mass spectrometer for its implementation. The values for dissociation or ionization energy which are normally derived from spectroscopic methods relate not only to zero pressure but also to zero absolute temperature.

Vapour pressure measurement (*c*)

This method is particularly appropriate to the determination of the enthalpy of vaporization of pure solids or liquids (latent heat); the change of internal energy in such a process is essentially entirely intermolecular. The equilibrium vapour pressure is measured over a range of temperature and the data is processed by means of the well-known Clausius–Clapeyron equation. The simplest form of the equation is

$$\log_{10}\left(\frac{P}{P^\circ}\right) = \frac{-\Delta h^\circ}{2{\cdot}303RT}+\text{constant}, \qquad 4$$

where P is the vapour pressure and Δh° is the enthalpy of vaporization per mole of material at some standard pressure, P°, frequently one atmosphere, that is, the molar enthalpy of a gas – the molar enthalpy of the corresponding liquid or solid (*both* at pressure P°). Thus a plot of $\log_{10}(P/P^\circ)$ against $1/T$ should be linear with gradient $-\Delta h^\circ/(2{\cdot}303R)$, from which Δh° may easily be calculated. Equation **4** assumes temperature independence of Δh° and also perfect gas behaviour. More elaborate forms of the equation, however, include temperature dependence and allow deduction of Δh° for a particular pressure *and* temperature. The method may be readily extended to the solution of a gas in a liquid. If P represents the partial pressure of the gas in equilibrium with the solution, equation **4** (and its extensions) may be applied to the determination of $-\Delta h^\circ$, the enthalpy of solution of the gas.

Calculation (*d*)

There are several situations where calculation of an internal energy or an enthalpy change is satisfactory. Some of these are presented below.

1. *Separation of an ionic lattice into its constituent ions*, for example

$$\mathrm{NaCl(s)} \longrightarrow \mathrm{Na^+(g)+Cl^-(g)}.$$
$$\text{(at zero pressure)}$$

X-ray diffraction analysis of crystalline solids, characterized by three-dimensional regularity, gives complete information on both the arrangement of constituent atoms (charged or otherwise) and the spacing between them. Salts can be regarded as alternating arrays of cations and anions, held together by mainly electrostatic forces; the ions are prevented from coalescing by the operation of an additional repulsive force (arising from interaction of individual electrons and of atomic nuclei) operating at close interionic distances – the Born repulsion. From a knowledge of the crystal structure, application of Coulomb's electrostatic law gives the internal energy of the crystal relative to that of the separated ions. It is, however, necessary to assume a particular dependence of Born repulsion energy upon the separation of ion *centres*; frequently one assumes proportionality to a power n of the separation with $n = -9$ (approximately). In cases where X-ray diffraction has not been applied, the problem may be approached in an approximate way by the use of *ionic radii*, which ascribe to each ion an incompressible spherical shape and a radius characteristic of the chemical identity of the ion and independent of environment. Both the ionic separations and the arrangement may be predicted from a knowledge of the radii of the ions involved; because of the gross approximation implicit in the concept, predictions are subject to some uncertainty.

2. *Application of empirical correlations* – graphical methods. This approach has been particularly exploited in recent years by the Russian scientist Karapet'yants. Several types of correlation have been considered but the principle may be illustrated by an example.

The enthalpy of vaporization of liquid 'inert' gas elements is plotted against the enthalpy of vaporization of 'corresponding' liquid halogens at the boiling point in Figure 4. The relationship is very close to linearity. In other cases, two *different* state function changes for the same compound are plotted. Yet in others, the same function for the same compounds is considered for different temperatures and/or pressures. The relationships are not always linear but generally show sufficient regularity to permit interpolation of 'missing' data.

3. *Conversion of data to different temperatures and/or pressures*

and interconversion of ΔH *and* ΔU *for the same process.* These conversions can be performed quite readily using standard thermodynamic equations and experimental relationships between volume, pressure and temperature of the materials involved. For small changes of temperature and pressure and for condensed 'reactants' and 'products', these conversions may produce only a small modification. For gases, these become significant but the assumption of perfect gas behaviour makes their calculation particularly simple.

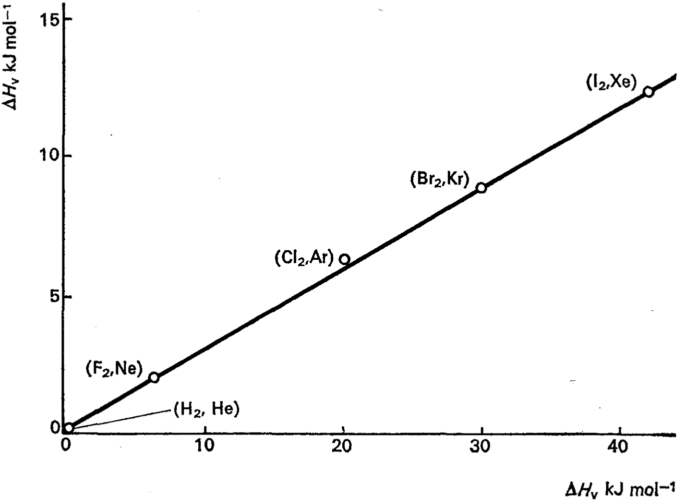


Figure 4[1]

So far we have been concerned with the fundamental principles underlying the application of cyclic processes to the determination of changes in properties of state. Of particular interest here were changes in internal energy and in enthalpy, symbolized by ΔU and ΔH, respectively. Experimental methods for determining the changes occurring in individual stages of chemical cycles were

[1] This figure was first published in *Journal of Chemical Education*, vol. 42 (1965), p. 602 and is reproduced here by permission of the editor of that journal.

outlined. Several applications of the cycle concept will now be described.

The Born–Haber Cycle

Undoubtedly the cycle which has attracted the greatest interest has been that devised by Born and Haber in 1919. It may be applied to a large variety of inorganic salts. Figure 5 shows the situation for sodium chloride. The cycle could equally well be considered in the reverse direction but the consequences are the same. As stated previously, some care must be taken in specifying not only the physical state of the entities involved but also their temperatures and pressures. This matter will be given some consideration here. The cycle is essentially a four-stage process although two of these may be considered in two halves as shown in Figure 5. Let us consider the stages separately.

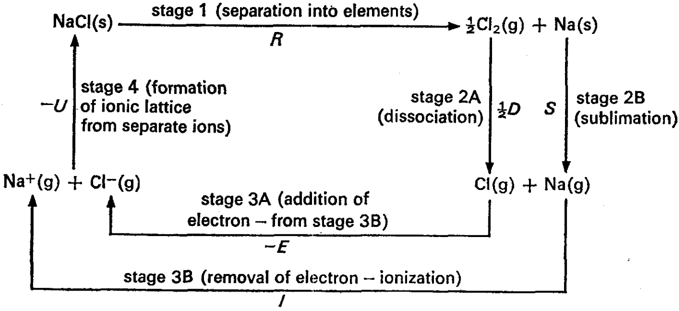


Figure 5

Stage 1: Experimentally the process may be studied in reverse by thermochemical means, that is by the burning of metallic sodium in an excess of gaseous chlorine in a calorimeter (method *a*).

For the purpose of the Born–Haber cycle, stage 1 applies to one mole of sodium chloride (NaCl) and to specified temperatures and pressures of the three species involved; in particular one might choose:

(a) sodium chloride at absolute zero of temperature and at a (defined) pressure near atmospheric – for a solid the latter is not very critical;

(b) chlorine at zero absolute temperature and zero pressure.

Correction of the calorimetric data will be necessary to accommodate these specifications (method *d*). Suppose that the increase in *internal energy* for stage 1 is *R*. This is approximately the same as the *enthalpy* (heat) *of formation* of sodium chloride, except that in the latter case the temperatures and pressures of the various reactants and products are identical. Correction of *R* is straightforward and proportionately small.

Stage 2A: The increase in internal energy for the dissociation of half a mole of Cl_2, both reactants and products being at zero temperature and pressure, may be deduced from spectroscopic observations (method *b*). If the molar dissociation energy is *D*, the required increase in internal energy will be $\frac{1}{2}D$.

B: The process here is the vaporization of a solid (sublimation), and this is most easily studied by vapour pressure measurement – method *c*. The molar enthalpy of vaporization may then be obtained by substitution into the Clausius–Clapeyron equation. By application of appropriate corrections (method *d*) this quantity may be converted into the *internal energy* increase for the sublimation of one mole of sodium, the temperature and pressure of the solid being that appropriate to stage 1, while for the gas the temperature and pressure are both zero. Let the appropriate increase in internal energy be *S*.

Stage 3A: The internal energy *decrease* for this process is known as the *electron affinity* (*E*) of the halogen atom. Direct measurement of *E* is experimentally difficult and the quantity is best determined by making this the single unknown in the cycle – but having once determined *E* for, say, chlorine, the internal energy changes for other stages (particularly 1 or 4) may be determined.

B: The internal energy increase here is the first ionization energy (or potential) of sodium, *I*. It may be determined by spectroscopic means (method *b*). The temperature and pressure for all four species in stage 3 are zero.

Stage 4: The internal energy *decrease* for this stage is the *lattice energy* (U) of the salt. This may be calculated (method *d*) from the known structure of the solid salt. It is assumed here that the attractive forces between sodium and chlorine are entirely ionic or electrovalent, that is that there is no contribution from covalency (electron sharing). However, it is necessary to include in the energy of the solid a repulsive force (Born repulsion) between *all* ions; neglect of this may lead to an error of about 8–16 per cent. It is commonly assumed that the contribution to the repulsion energy from a particular pair of ions is proportional to a power n of the separation of these ion centres in the crystal. The lattice energy is actually proportional to $(1+n^{-1})$ – the proportionality constant being a known function of the crystal structure and ionic charges – so that n is required for a true calculation of the energy. n may be calculated from measurements on the compressibility of the crystal. For sodium chloride, $n = -8{\cdot}2$. For the alkali metal halides in general, n ranges from $-5{\cdot}8$ (for LiF) to $-13{\cdot}2$ (for CsBr).

For the complete cycle:

$$0 = \Delta U_{\text{over-all}} = R+\tfrac{1}{2}D+S-E+I-U. \qquad \mathbf{5}$$

Equation **5** may be applied in a number of ways, the following being some of the more important.

Determination of electron affinities

For the alkali metal halides, where the assumption of complete absence of covalent forces in the crystal is reasonable and where R, D, S and I are all measurable:

$$E = R+\tfrac{1}{2}D+S+I-U. \qquad \mathbf{6}$$

If the electron affinity is required for, say, chlorine, equation **6** may be applied to *all* the alkali metal chlorides. The resulting values of E may then be averaged. For sodium chloride, $R = 409$, $D = 240$, $S = 109$, $I = 496$ and $U = 774$ kJ mol^{-1}, so that $E = 360$ kJ mol^{-1}. (Note: 1 joule (J) = 4·184 thermochemical calories (cal) exactly.) Similar figures may be derived from the remaining alkali metal chlorides. An average value of 362 kJ mol^{-1} is employed for the electron affinity of chlorine. For the other halogens – fluorine, bromine and iodine – average values of 343, 339 and 297 kJ mol^{-1}, respectively, have been obtained. For bromine and iodine it is usual to consider stage 1 in two parts: (i) the production of *solid* metal and molecular halogen as a *liquid* for bromine or as a *solid* for iodine, and (ii) vaporization of the halogen. The electron affinities of the halogens (Cl, Br, I = X) have also been determined by more direct means, in fact by studying the equilibria

$$X + e^- \rightleftharpoons X^-.$$

Values obtained by such methods, which are difficult experimentally, compare favourably with those obtained from the Born–Haber cycle.

Comparison of salt stabilities and prediction of stabilities of unknown compounds

Having determined values of E, equation 5 may be used in the form:

$$R = -\tfrac{1}{2}D - S + E - I + U. \qquad 7$$

R may then be used as an approximate measure of the stability of the salt relative to the corresponding elements in the physical states shown in Figure 1. Use of vaporization energies will then permit the prediction of stabilities relative to the elements in other physical states. In considering the stability of a solid salt *at a particular temperature and pressure*, one ought strictly to determine the change in the Gibbs function, ΔG (or Δg if one mole of salt is involved) for stage 1 with reactants and products at the particular temperature and pressure. Conversion of Δu to Δg involves either

(i) a knowledge of the *entropy* change ($\Delta S = (\Delta H - \Delta G)/T$ or $\Delta s = (\Delta h - \Delta g)/T$ per mole) for stage 1, or

(ii) the employment of a Karapet'yants-type correlation of Δu (or Δh, to which Δu is readily convertible) and Δg (method *d*).

However, R may be used as an approximate measure of comparative stability of salts, large positive values implying high stability and large negative values high instability. The approximation of Δg to Δu *at the same temperature and pressure* has greatest validity when $\Delta g - \Delta u = p\,\Delta v - T\,\Delta s$ is small. For purposes of comparative stability, it is sufficient for this difference to be nearly constant. The interplay of the various terms on the right-hand side of equation **7** may be considered for a series of compounds. For example, suppose we are interested in the relative stabilities of alkali metal fluorides, $E - \frac{1}{2}D$ is constant so that equation **7** becomes

$$R = \text{constant} - S - I + U.$$

In passing from Li to Cs, S decreases from 161 to 78, I decreases from 520 to 375 and U decreases from 1028 to 722 kJ mol^{-1}; the direction of the latter change is compatible with increasing size of ions. Over all, there is a decrease in stability as we pass from LiF to CsF. Actually, the order is reversed for the chlorides, bromides and iodides. The basic reason for this is that for larger anions, the separation of ion centres is larger and the reciprocal of this spacing, to which the lattice energy is proportional, extends over a smaller range.

Similar considerations may be applied to unknown compounds provided that their crystal structures can be assigned to a reasonable degree of approximation.[1] To take an example, one can calculate the change in internal energy for the process:

$$\underset{0\,\text{K},\ \sim 1\ \text{atm}}{XF(s)} \longrightarrow \underset{0\,\text{K, zero pressure}}{X(g) + \tfrac{1}{2}F_2(g)} \qquad (X = \text{He, Ne, Ar, Kr, Xe})$$

if it is *assumed* that the bonding in XF is ionic and that $U_{XF} = U_{X'F}$, where X′ is the 'corresponding' alkali metal, that is the nearest one in the periodic classification. Δu may be calculated as follows: HeF $1{\cdot}07 \times 10^3$, NeF $9{\cdot}0 \times 10^2$, ArF $4{\cdot}4 \times 10^2$, KrF $3{\cdot}1 \times 10^2$ and XeF $1{\cdot}7 \times 10^2$ kJ mol^{-1}. The implication here is that all are highly unstable, although the possibility of stabilization by covalent bonding is not excluded.

The Born–Haber cycle may also be applied to salts involving ions in oxidation states higher than unity. By considering cycles for two different oxidation states, the stability of one state relative to the other may be assessed, for example one can determine Δu or Δh for such processes as

$$MCl(s) \longrightarrow \tfrac{1}{2}MCl_2(s) + \tfrac{1}{2}M(s),$$

---

[1] Rules based on the relative sizes of anion and cation – the radius ratio rules – are of assistance here.

M being a metal such as calcium. Furthermore, the Born–Haber cycle may be applied in a modified form to salts containing polyatomic ions such as K^+CN^- or $NH_4^+Cl^-$. The latter example is shown in Figure 6. From the known internal energy changes

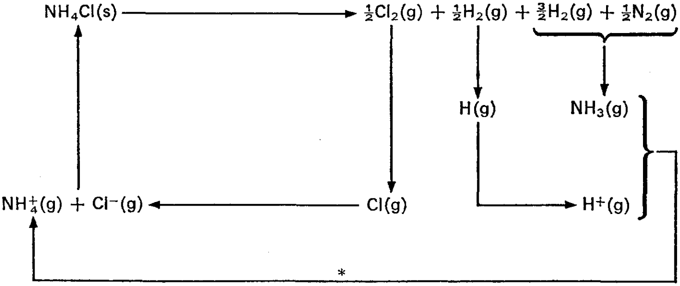


Figure 6

for the remaining stages, the internal energy change for the process marked with an asterisk may be deduced (-870 kJ mol^{-1}).

Determination of lattice energies

In salts comprising ions of large radius and/or in high oxidation states (three or more), there is an appreciable contribution to the lattice energy from covalent forces and calculations based upon electrostatic forces between ions become impracticable. In such circumstances, equation **5** may be employed in the form

$$U = R+\tfrac{1}{2}D+S-E+I. \tag{8}$$

An example of such an application is to tin(IV) oxide, SnO_2. The lattice energy as calculated from equation **8** is $1{\cdot}137\times10^4$ kJ mol^{-1}, which is 530 kJ mol^{-1} greater than that calculated on a *purely ionic* basis.

In the foregoing applications, the Born–Haber cycle has been considered in terms of internal energy changes. However, it could just as easily be applied in terms of enthalpy changes. In fact, this is commonly done for the cycle of Figure 6. Furthermore, the temperatures and pressures of the entities in the various stages may be chosen differently from those values given above. It is, however, essential that the conditions at the end of one stage are the same as those at the beginning of the next. Some lack of care has been shown in the past in this regard.

Other Chemical Cycles

There are several other cycles which are frequently encountered in chemistry. As with the Born–Haber cycle, care should be taken to specify the reactants and products in each stage properly; some error will occur if the temperature and pressure are different for the end of one stage and the beginning of the next. Similarly, one should not mix internal energy changes with enthalpy changes. Either quantity (or any other state property in fact) may be employed, but it must be uniform. Unlike the Born–Haber cycle, the other cycles tend to have a single type of application. The following is a brief description of some of the more important ones.

The dissociation energy of the fluorine molecule

This energy is not directly determinable by spectroscopic means. It can, however, be derived from a consideration of the cycle shown in Figure 7. The internal energy changes involved in the various stages (per mole of MF) are symbolized and determined as shown in the figure. Then

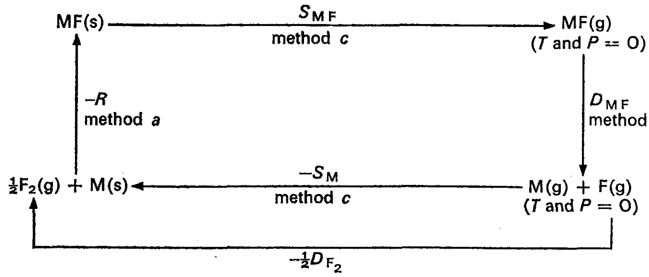


Figure 7

$$0 = \Delta U_{\text{over-all}} = S_{MF} + D_{MF} - S_M - \tfrac{1}{2}D_{F_2} - R$$

and $$\tfrac{1}{2}D_{F_2} = S_{MF} + D_{MF} - S_M - R \quad (\mathrm{M} = \mathrm{K, Rb, Cs}).$$

The currently accepted value of D_{F_2} is 154 kJ mol^{-1}.

Stability of oxosalts

In this category are included hydroxides, carbonates and sulphates. The concept is illustrated in Figure 8 for Group II metal carbonates. Using the established principles, the internal energy changes for the four stages may be interrelated by

$$\begin{tikzcd}
\mathrm{MCO_3(s)} \arrow[r, "R"] & \mathrm{MO(s)} + \mathrm{CO_2(g)} \arrow[d, "U_{\mathrm{MO}}\ \text{method } d"] \\
\mathrm{M^{2+}(g)} + \mathrm{CO_3^{2-}(g)} \arrow[u, "-U_{\mathrm{MCO_3}}\ \text{method } d"'] & \mathrm{M^{2+}(g)} + \mathrm{O^{2-}(g)} + \mathrm{CO_2(g)} \arrow[l, "X"']
\end{tikzcd}$$

Figure 8

$$R = -U_{\mathrm{MO}} - X + U_{\mathrm{MCO_3}}.$$

X is independent of the metal M, being the increase in internal energy when one mole of CO_3^{2-} is formed from CO_2 and O^{2-} (actually at zero temperature and pressure). Thus R is paralleled by the difference in the molar lattice energies for the metal carbonate and the metal oxide. On passing from magnesium to barium (via calcium and strontium), both lattice energies decrease but the decrease is more pronounced for the oxides than for the carbonates; this is due to the larger size of the latter ion. Thus R increases down the group and so, therefore, does the stability; this is reflected in the increase in the decomposition temperature – the temperature at which the pressure of carbon dioxide in the decomposition reaches one atmosphere.

Solvation of ions

Ions in polar solvents, such as water, are not isolated species but are 'bonded' to surrounding solvent molecules. In some cases, these molecules are attached by weak ion–dipole forces (for example Na^+). In other cases *some* of the water molecules are covalently bonded (for example $[Fe(H_2O)_6]^{3+}$), but in either case there is a decrease in internal energy in the solvation process:

$$X^{n\pm}(g) + \text{solvent} \longrightarrow X^{n\pm}\ (\text{solvent}) + \text{excess solvent}.$$

At zero concentration (infinite dilution), the enthalpy decrease per mole of ions is called the *solvation enthalpy*: for water as solvent in particular the appropriate term is *hydration enthalpy.* The solution of a salt may be considered as one stage of a cycle. This is illustrated in Figure 9 for sodium chloride: R is the enthalpy (heat) of solution; H_{Na^+} and H_{Cl^-} are hydration enthalpies. Then

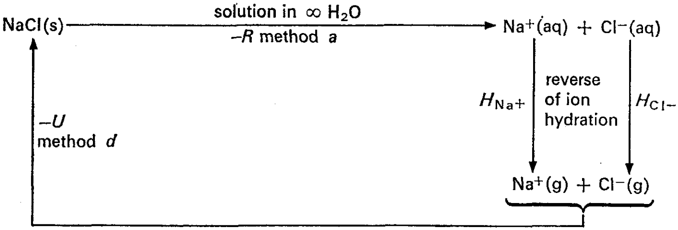


Figure 9

$$H_{\mathrm{Na}^+}+H_{\mathrm{Cl}^-} = R+U \qquad \mathbf{9}$$

from which the *sum* of the cationic and anionic hydration enthalpies may be determined. Parallel study of potassium chloride (or sodium bromide) provides an equation which when combined with equation **9** gives a value for $H_{\mathrm{Na}^+}-H_{\mathrm{K}^+}$ (or $H_{\mathrm{Cl}^-}-H_{\mathrm{Br}^-}$). If ions in different oxidation states are involved, weighted sums and differences may be determined, for example $H_{\mathrm{Ca}}{}^{+}+2H_{\mathrm{Cl}^-}$, $H_{\mathrm{Ca}^{2+}}-2H_{\mathrm{Na}^+}$ (in general $nH_{\mathrm{M}^{p+}}+pH_{\mathrm{X}^{n-}}$, $nH_{\mathrm{M}_n{}^+}-pH_{\mathrm{M}_2^{n+}}$). Differences involving H_{H^+} are frequently quoted, for example $H_{\mathrm{Zn}^{2+}}-2H_{\mathrm{H}^+}$. Semi-empirical methods are available for separating sums and differences into individual H values but this need not concern us here.

Known values of hydration enthalpy sums and differences may be used in cyclic schemes. Indeed they may be used in the scheme of Figure 9 to deduce R, the enthalpy of solution, but this is not very satisfactory as the hydration enthalpy sum and the lattice enthalpy, U, are generally of similar magnitude and both much larger than R. The latter may be subject to appreciable percentage error as a result. Furthermore, it is unwise to use R as a means of comparing solubilities; the Gibbs function change, which is a truer measure, may well show a different order from the enthalpy change, R. Two other cycles involving hydration of ions are

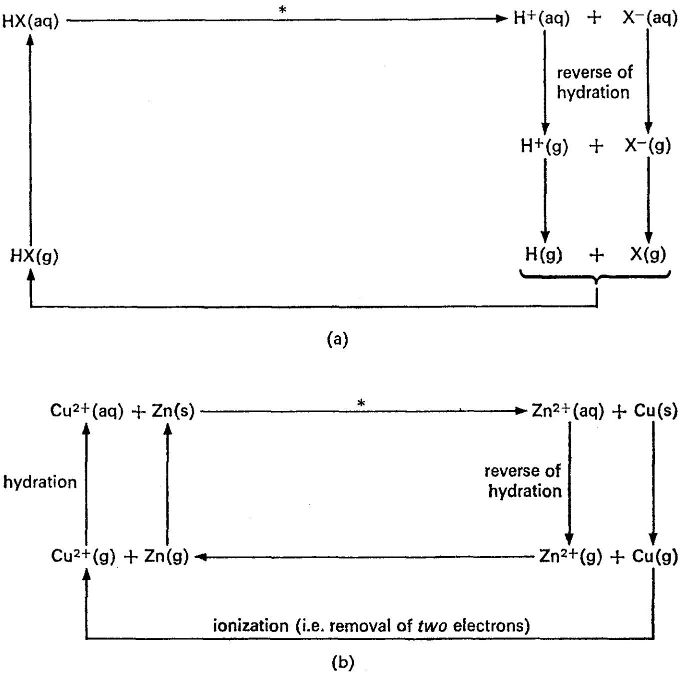


Figure 10

shown in Figure 10. It is common to consider both these cycles in terms of Gibbs function changes, ΔG. The stages of prime interest are indicated by an asterisk in each case. The molar free energy change (Δg) for this stage in Figure 10(a) is a measure of the extent of the process when equilibrium is attained, that is it is a measure of the strength of the acid HX (X = F, Cl, Br, I). A large negative Δg corresponds to extensive dissociation; actually

$$\log_e \text{(acid dissociation constant)} = -\frac{\Delta g^\circ}{RT},$$

where Δg° is the free energy change when the concentrations (strictly, activities) of HX, H^+ and X^- are all 1 M. The order of acid strength (at ambient temperatures) is HI > HBr > HCl > HF, the latter being very weak (at 298 K, the acid dissociation constant is 7×10^{-4} mol dm^{-3}) and the other three almost completely ionized. The order can be rationalized in terms of the Gibbs function changes in the other stages of the cycle. Principally responsible for the weakness of hydrofluoric acid is the large dissociation energy of HF and correspondingly a large Δg for

$$HF(g) \longrightarrow H(g)+F(g).$$

The molar free energy change marked with an asterisk in (b) may be determined from e.m.f. measurements on the Daniell cell for which this is the cell reaction. Similar reactions may be considered in the same way for other electrolytic cells. It should be noted that cycle (b) involves a *difference of hydration Gibbs functions* ($Cu^{2+}-Zn^{2+}$) and that cycle (a) involves a sum of these ($H^{+}+X^{-}$).

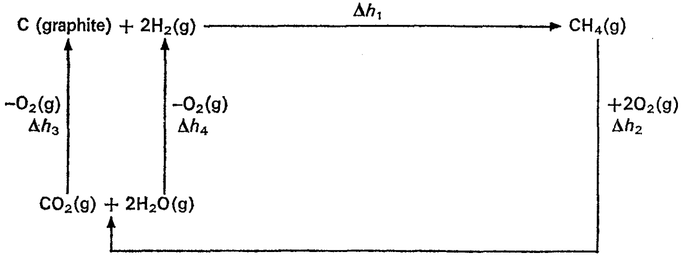


Figure 11

Hess's law of constant heat summation

The familiar Hess's law can be readily understood in terms of cyclic processes. An example is given in Figure 11. On the basis of enthalpy changes, the enthalpy (heat) of formation of methane, Δh_1, is given by

$$-\Delta h_1 = \Delta h_2+\Delta h_3+\Delta h_4.$$

Δh_2, Δh_3 and Δh_4 may be determined experimentally by combustion of methane, graphite and hydrogen, respectively, in a calorimeter.

Bibliography

R. S. Nyholm, 'The stability of inorganic compounds', *Modern Chemistry and the Sixth Form*, Collins, 1965.

L. E. Strong and W. J. Stratton, *Chemical Energy*, Chapman & Hall, 1966.

8 A. R. Denaro

Electrode Processes

A. R. Denaro, 'Electrode processes', *Chemistry Student*, vol. 2 (1968), no. 2, pp. 40–45.

Electrode processes can be considered from two points of view. The equilibrium approach relates to the thermodynamic properties of the systems considered and enables the properties of such systems to be compared. The kinetic approach relates largely to conditions where there is a flow of current and is therefore applicable mainly to practical situations such as electrolysis. There is, of course, some measure of overlap in the two approaches.

The Equilibrium Approach (Reversible Processes)

If a metal is immersed in a solution of its ions, a separation of charge is found to occur and a potential difference is established between the metal and the solution. One of the reactions occurring at the metal–solution interface is

$$M \rightarrow M^+ + e^-. \qquad \mathbf{1}$$

Some of the metal atoms lose electrons and pass into solution as metal ions. The liberated electrons accumulate in the metal which would thus become negatively charged with respect to the solution.

A further reaction which occurs is

$$M^+ + e^- \rightarrow M \qquad \mathbf{2}$$

in which a metal ion in solution travels to the metal–solution interface and abstracts an electron from the metal to deposit as a metal atom. This process would lead to a deficit of electrons in the metal which would thus become positively charged with respect to the solution.

If the velocity of reaction **1** is greater than that of reaction **2**, the metal would acquire a net negative charge making it more difficult for positive metal ions to leave the metal and diffuse into the solution. This state of affairs would have the effect of retarding reaction **1**. Furthermore, the negative charge on the metal would attract the positive ions in the solution thus accelerating reaction **2**. The slower reaction is thus accelerated and the faster reaction is retarded until the speeds of the two reactions are equal. While the speeds of the reactions are changing the negative charge on the metal steadily increases, but when the speeds of the reactions are equal equilibrium is established, electrons are being removed from the metal at the same rate as they accumulate, and a steady potential between the metal and the solution is established.

If the velocity of reaction **2** had been initially greater than that of reaction **1**, the metal would have acquired a net positive charge, thus accelerating reaction **1** and retarding reaction **2** leading once again to the establishment of equilibrium.

The equilibrium between these two reactions is represented

$$M \rightleftharpoons M^+ + e^-$$

the final potential adopted by the metal depending upon the position of this equilibrium which, in turn, depends upon the concentrations (or more strictly the activities) of the species involved. An equilibrium process is often said to be thermodynamically reversible and hence the potential of the metal under the equilibrium conditions described above is known as the reversible electrode potential of the metal.

The concept of electrode potentials can also be applied to gas electrodes: for example, a chlorine electrode where the equilibrium is

$$\tfrac{1}{2}Cl_2 + e^- \rightleftharpoons Cl^-.$$

Chlorine gas is bubbled around an electrode of platinized platinum immersed in a solution of chloride ions. The gas is adsorbed on to the surface of the metal and enters into equilibrium with the chloride ions in solution. In this case the platinum metal acts simply as a source or sink of electrons and takes no chemical part in the reaction.

It should be pointed out that in the examples quoted above, the equilibria involve two opposing reactions – an oxidation reaction and a reduction reaction. This must be so, as the reactions involve a loss or gain of electrons. The equilibria exist therefore between the oxidized and reduced forms of the electrode systems:

$$\mathrm{M(red)} \rightleftharpoons \mathrm{M^+(ox)} + \mathrm{e^-}$$
$$\tfrac{1}{2}\mathrm{Cl_2(ox)} + \mathrm{e^-} \rightleftharpoons \mathrm{Cl^-(red)}.$$

There is another type of electrode system where both the oxidized and reduced forms exist in the solution. Once again an inert metal electrode is present as a source or sink of electrons. An example of such a system is

$$\mathrm{Fe^{2+}} \rightleftharpoons \mathrm{Fe^{3+}} + \mathrm{e^-}.$$

There is no fundamental difference between this example and the first two but such systems where the oxidized and reduced form exist in solution have come to be called *redox* systems and the potentials associated with them are known as redox potentials. Once again, the magnitude of the potential depends upon the position of the equilibrium which will depend upon the concentrations of the species involved.

As there is no fundamental difference between electrode potentials and redox potentials, for the purposes of the subsequent discussion they can be considered together. With redox electrodes and gas electrodes an inert metal is present to act as a source or sink of electrons and, in the case of a metal electrode system, the metal itself can be considered to act in the dual capacity of the reduced form of the system and as an electron source or sink. All electrode equilibria can thus be generalized as

$$\mathrm{Ox} + \mathrm{e^-} \rightleftharpoons \mathrm{Red}. \qquad 3$$

It has already been mentioned that the magnitude of the potential depends to some extent on the concentrations of the species involved. Obviously, the greater the concentration of the reduced form, the more the equilibrium lies to the left, the more electrons accumulate in the electrode and the more negative is the potential. It is important to note at this stage that it is not possible to measure the potential difference between an electrode and a solution. We cannot say, therefore, whether the potential of a particular system is positive or negative in an absolute sense. It is only possible to compare the potential of an electrode system with some arbitrary standard and we are restricted to relative terms in so far as we can only say that the further the equilibrium lies to the left the more negative will be the potential.

While the effect of concentration on the magnitude of the potential is important in practical problems for any individual system, if we wish to compare the properties of different systems we must eliminate concentration effects. A standard value of redox or electrode potentials is chosen which is the value of the potential when the activities of the oxidized and reduced forms are both equal to unity. These potentials are called *standard electrode* or *redox potentials* and are the potentials which the systems exhibit when both the oxidized and reduced forms are present in their *standard states of unit activity*. While this statement has thermodynamic overtones, these can be avoided by merely stating that we choose particular states for the various electrode systems and compare the potentials in these states to arrive at some idea of the comparative chemistry of the systems in these particular states. The states chosen are:

metal electrodes	pure metal, metal ions at unit activity;
gas electrodes	gas at 1 atmospheric pressure, ions at unit activity;
redox electrodes	reduced form at unit activity, oxidized form at unit activity.

We are now in a position to compare the properties of electrode systems by comparing their standard potentials. The potential is that adopted by the electrode due to the equilibrium between the reduced and oxidized forms of the couple as given in reaction 3. If the equilibrium lies to the left, the electrons which are generated collect on the electrode giving rise to more negative potentials. This means that the system readily gives up electrons which is tantamount to saying that it is a good reducing agent. Alternatively, if the equilibrium lies to the right, electrons are abstracted from the electrode which thus adopts more positive potentials. In this case the system readily accepts electrons and is thus an oxidizing agent. The more positive the potential of a system, the better oxidizing agent it is and conversely the more negative the potential, the better reducing agent it is.

This can be illustrated by reference to the standard electrode potentials of zinc and copper:

$$Zn^{2+}/Zn \quad -0{\cdot}763\ V, \quad Cu^{2+}/Cu \quad +0{\cdot}337\ V.$$

As the potential of the copper system is more positive than that of the zinc system, copper(II) ions must be better oxidizing agents than zinc ions. Hence, if the two systems are mixed, copper(II) ions will oxidize metallic zinc:

$$Cu^{2+}+Zn \longrightarrow Cu+Zn^{2+}.$$

A further illustration of more practical importance in oxidation–reduction titrations is obtained from a consideration of the redox potentials of the Fe^{2+}/Fe^{3+} system and the Ce^{3+}/Ce^{4+} system:

$$Fe^{2+}/Fe^{3+} \quad +0{\cdot}771\ V, \quad Ce^{3+}/Ce^{4+} \quad +1{\cdot}61\ V.$$

In this case Ce^{4+} will oxidize Fe^{2+}

$$Ce^{4+}+Fe^{2+} \longrightarrow Ce^{3+}+Fe^{3+}$$

which is the basis of a volumetric method of estimating iron.

An alternative way of regarding the situation is to consider the relative stabilities of the oxidized and reduced forms of a couple. Referring to reaction 3 again, if the equilibrium lies to the left the oxidized form must be more stable. Thus the more negative the potential the more stable is the oxidized form and conversely the more positive the potential the more stable is the reduced form. The whole situation may be summarized as

$$Ox+e^- \rightleftharpoons Red$$

$$\longleftrightarrow$$

More negative potentials	More positive potentials
Better reducing agents	Better oxidizing agents
More stable oxidized form	More stable reduced form

To illustrate this last point compare the potentials of the Fe^{2+}/Fe^{3+} system and the Mn^{2+}/MnO_4^- system in acid solution:

$$Fe^{2+}/Fe^{3+} \ +0{\cdot}771\ V, \ Mn^{2+}/MnO_4^- \quad +1{\cdot}51\ V.$$

As the manganese system has the more positive potential manganese(II) ions are more stable than iron(II) ions and if the two systems are mixed a reaction will occur to give the more stable products manganese(II) ions and iron(III) ions:

$$MnO_4^- + 8H^+ + 5Fe^{2+} \longrightarrow Mn^{2+} + 5Fe^{3+} + 4H_2O.$$

This approach leads on to the effects of complexing in enhancing the stabilities of different oxidation states. The redox potential of the Co^{2+}/Co^{3+} system is +1·82 V which is a very positive potential. Thus the lower oxidation state is very stable while the higher oxidation state is very unstable in aqueous solution. In fact, simple Co^{3+} ions do not exist in aqueous solution as they are rapidly reduced by the water. If, however, the cobalt system is complexed with ammonia to form the hexammine complexes of the ions the resulting redox system has a much more negative potential than that of the simple ion system:

$$Co(NH_3)_6^{2+}/Co(NH_3)_6^{3+} \quad +0{\cdot}10\ V.$$

This more negative potential indicates that $Co(NH_3)_6^{3+}$ ions are more stable than Co^{3+} ions. Complexing has thus stabilized the higher oxidation state. The same is true of the iron system:

$$Fe^{2+}/Fe^{3+} \quad +0{\cdot}771\ V,$$
$$Fe(CN)_6^{4-}/Fe(CN)_6^{3-} \quad +0{\cdot}36\ V.$$

The above considerations cover some of the equilibrium aspects of electrode processes and we can now turn to consider some kinetic aspects.

The Kinetic Approach (Irreversible Processes)

Consider a metal in equilibrium with a solution of its ions having adopted its reversible electrode potential. If such a half cell is coupled with another half cell and the complete cell is connected to a potentiometer circuit which is balanced, the electrode will still be at its reversible potential. This situation is illustrated in Figure 1.

Let us fix our attention on the half cell indicated. If the potentiometer is balanced, zero current flows and hence the point X in the circuit must be at the same potential as the electrode. We can, however, alter the potential of the point X by moving the position of the contact on the potentiometer bridge wire.

If the potential of the point X is made infinitesimally more negative than the reversible potential of the electrode a small potential difference exists between X and the electrode and a small current will flow. Electrons will flow from X to the electrode and on arriving at the electrode will disturb the equilibrium which existed. The system will endeavour to restore equilibrium by consuming these electrons and a net reaction occurs which is

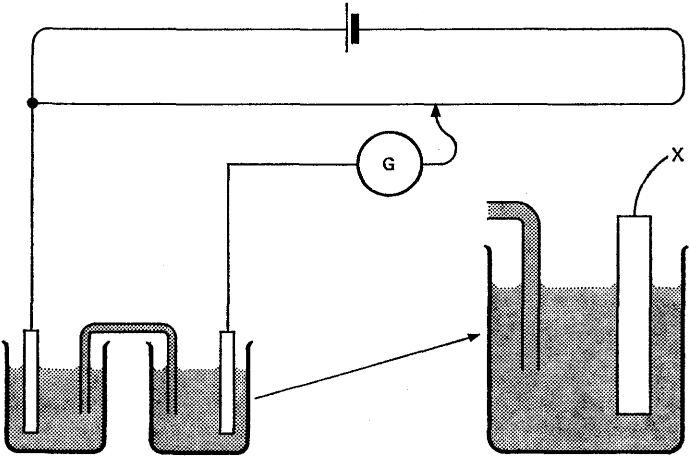


Figure 1 Potentiometer circuit

$$M^+ + e^- \longrightarrow M.$$

In this case, cations (M^+) are travelling to the electrode to combine with electrons and the electrode is behaving as a *cathode*. If the current is small, electrons are consumed as fast as they arrive and the 'stationary concentration' of electrons in the metal is the same as the 'equilibrium concentration' and the electrode maintains its equilibrium potential. Under these conditions the electrode is said to be operating reversibly as a cathode.

If the point X is made appreciably more negative than the potential of the electrode, electrons will arrive more rapidly and if there is a slow stage in the electrode reaction the 'stationary concentration' of electrons in the metal will be greater than the equilibrium value and the potential of the metal becomes more negative than its equilibrium potential. The electrode is said to be *polarized* and to be operating irreversibly as a cathode. The difference between the working potential and the equilibrium potential is called the *overpotential.* We have, therefore, for cathodes where the working potential is more negative than the equilibrium potential

$$E'_C = E_{rev} - \eta_C,$$

where E'_C is the working potential of the cathode, E_{rev} is its equilibrium potential and η_C is the cathode overpotential.

The above situations can be illustrated with the apparatus shown in Figure 2. This consists of a vessel containing water with a wide outlet so that the level in the vessel can be maintained constant. A tube with a constricted outlet near the bottom is mounted inside the outer vessel. This inner tube represents the electrode and the water represents the electrons. At equilibrium molecules of water will pass from the inner tube through the

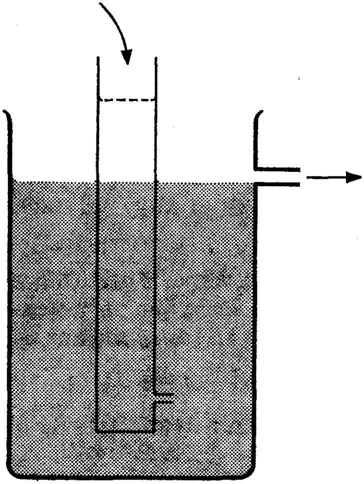


Figure 2 Electrode analogue

constriction into the outer vessel and also in the reverse direction. This is the analogue of the equilibrium

$$M \rightleftharpoons M^+ + e^-.$$

If current flow is now represented by the addition of water to the inner tube, provided the rate of flow is small water escapes from the bottom of the tube as fast as it is added and the volume of water contained in the inner tube remains the same. This is the analogue of reversible operation. If, however, water is added more rapidly and the rate of escape of water is restricted by the constriction, a head of water will begin to build up in the inner tube until the extra pressure of this head of water increases the rate of flow through the constriction. When this rate is equal to the rate of addition of water a steady state will have been achieved with a constant level of water in the inner tube which will be higher than the equilibrium level. This is the analogue of polarization of the electrode and irreversible operation.

To return to the original situation, if the point X is made more positive than the equilibrium potential of the electrode, electrons will flow from the electrode to the point X. This loss of electrons disturbs the electrode equilibrium and the system behaves in such a way as to restore the lost electrons. A net reaction occurs which is

$$M \longrightarrow M^+ + e^-.$$

Metal atoms go into solution as metal ions, which then diffuse away from the electrode. As cations are travelling away from the electrode, it must be behaving as an *anode*. If the current is small the electrode operates reversibly maintaining its equilibrium potential. If the current is large, electrons escape more rapidly than they are formed and the potential of the electrode becomes more positive than its equilibrium potential. Thus, for an anode

$$E'_A = E_{rev} + \eta_A,$$

where E'_A is the working potential of the anode, E_{rev} is its equilibrium potential and η_A is the anodic overpotential.

It is to be noted that overpotential always makes anodes more positive and cathodes more negative than their reversible potentials. Further, for an electrode to behave as a cathode it must be connected to a point in the circuit more negative than its reversible potential, and for an electrode to behave as an anode it must be connected to a point in the circuit more positive than its reversible potential.

It might be useful at this point to clarify the classification of anodes and cathodes as there is frequently some confusion on this point. Consider a zinc–chlorine cell in its standard state. The standard potentials of zinc and chlorine are −0·76 V and +1·36 V respectively. As the chlorine electrode has the more positive potential it will be the positive pole of the cell. If the cell is connected to an external load and allowed to operate spontaneously the situation may be represented as

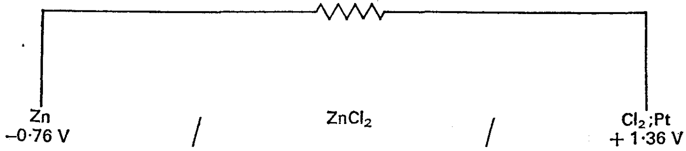


Figure 3

The zinc electrode is connected through the load to a potential (that of the chlorine electrode) more positive than its equilibrium potential and thus behaves as an *anode* but it is still the *negative* pole of the cell. Conversely the chlorine electrode functions as a *cathode* and is the *positive* pole of the cell.

If, however, we wish to electrolyse zinc chloride solution between a zinc electrode and a platinum electrode, connecting the zinc to the negative pole of the external source and the platinum to the positive pole of the external source, chlorine would be liberated at the platinum which would then be the anode and the zinc electrode would be the cathode. In this case the *negative* pole is the *cathode* and the *positive* pole the *anode.* Cathodes should always be defined as the electrode at which negative electricity enters the solution regardless of the polarity of the electrode.

For electrolysis to occur in the above example, the applied potential difference must overcome the opposing e.m.f. of the cell, that is it must be at least infinitesimally greater. This means that the potential applied to the zinc electrode is infinitesimally more negative than its equilibrium potential (cathodic behaviour) and the potential applied to the chlorine electrode must be infinitesimally more positive than its equilibrium potential (anodic behaviour). The minimum value of applied potential difference required for electrolysis is thus equal to the reversible e.m.f. of the cell and may be called the reversible decomposition voltage D_{rev}. Under these conditions any electrolysis current flowing will be negligible. If an appreciable electrolysis current is required the applied potential difference must also overcome any overpotentials associated with the electrode processes and the voltage drop, IR, across the cell. The decomposition voltage D_I required to maintain a current I through the cell is given by

$$D_I = D_{\text{rev}}+\eta_{\text{A}}+\eta_{\text{C}}+IR$$

where $$D_{\text{rev}} = E_{\text{A}}-E_{\text{C}},$$

E_{A} and E_{C} being the reversible potentials of the anode and cathode respectively. Thus

$$D_I = E_{\text{A}}-E_{\text{C}}+\eta_{\text{A}}+\eta_{\text{C}}+IR$$

or $$D_I = (E_{\text{A}}+\eta_{\text{A}})-(E_{\text{C}}-\eta_{\text{C}})+IR.$$

The terms $(E_{\text{A}}+\eta_{\text{A}})$ and $(E_{\text{C}}-\eta_{\text{C}})$ are the working potentials of the anode and cathode respectively and these potentials are called the discharge potentials for the electrode processes. Thus, for a substance to be discharged at a cathode at a given current the potential of the cathode must be less than the appropriate discharge potential. Conversely, for an anodic process to occur at a given current the anode potential must be greater than the appropriate discharge potential. As discharge potentials include overpotential, whose magnitude depends upon current, a discharge potential must always be accompanied by a statement of the current to which it refers.

To have some idea of discharge potentials we need to know the equilibrium potential of the system and the overpotential. The first factor may be calculated from the Nernst equation

$$E = E^\circ+\frac{RT}{zF}\log_e\frac{\text{(oxidized form)}}{\text{(reduced form)}},$$

where the round brackets indicate activities. The second factor is a problem as overpotential varies with current density, electrode material, temperature, etc. It can only be known with certainty from an experimental measurement. It is possible, however, to form some ideas by considering the origin of overpotential.

As has been pointed out polarization and the consequent overpotential result from a slow stage in an electrode process. The discharge of ions at an electrode can be divided into three main stages:

(1) transport of ions to the electrode,
(2) discharge of ions to form atoms,
(3) conversion of atoms to the normal stable form.

Any of these three stages can be the slow stage leading to polarization. If item (1) is the slow stage the resulting overpotential is called concentration overpotential and is more important in dilute solutions, but our interest lies with the overpotential arising from items (2) and (3) which is called activation overpotential.

We can form some opinion of the magnitude of overpotential from a knowledge of the velocity of the electrode reaction. The slower the reaction, the greater the overpotential is likely to be for a given current. In order to consider the kinetics of the electrode processes it may be helpful to revert to the specific case of a metal in equilibrium with its ions. This equilibrium is the result of two opposing reactions:

$$M \rightarrow M^+ + e^-, \qquad 1$$

and

$$M^+ + e^- \rightarrow M. \qquad 2$$

In reaction **1** cations are leaving the electrode which must, therefore, be behaving as an anode. Reaction **1** is, therefore, an anodic reaction which is of course an oxidation reaction as electrons are given up. In reaction **2** cations travel to the electrode to be deposited as metal, and the electrode is thus behaving as a cathode. Reaction **2** is thus a cathodic reaction which is of course a reduction reaction.

At equilibrium reactions **1** and **2** are proceeding at equal speeds, and the electrode exhibits no net cathodic or anodic behaviour. Now the velocity of a reaction can be expressed in terms of current density. As the reaction occurs at an interface, instead of expressing the rate as mol l^{-1} s^{-1} it is expressed as mol cm^{-2} s^{-1}. For reaction **1**, the anodic reaction:

$$\begin{aligned} \text{velocity} &= v_1 && \text{mol cm}^{-2}\ \text{s}^{-1} \\ &= zv_1 && \text{equiv cm}^{-2}\ \text{s}^{-1} \\ &= zFv_1 && \text{C cm}^{-2}\ \text{s}^{-1} \\ &= zFv_1 && \text{A cm}^{-2}. \end{aligned}$$

Thus, representing an anodic current density as i_a,

$$i_a = zFv_1 \text{ A cm}^{-2}.$$

Similarly, for reaction **2**, the cathode reaction, the cathodic current density i_c is given by

$$i_c = zFv_2 \text{ A cm}^{-2},$$

where v_2 is the velocity of the cathodic reaction.

At equilibrium $i_c = i_a$ since $v_1 = v_2$ and we write

$$i_c = i_a = i_0,$$

where i_0 is known as the *exchange current density*. The exchange current density is thus a measure of the velocity of the electrode reactions and will give some idea of the magnitudes of overpotentials. Referring to the analogy depicted in Figure 2, the less the constriction at the bottom of the inner tube, the more readily will water molecules exchange between the inner and outer vessels and the larger the exchange current density. Thus, the greater the value of i_0 for a system, the faster will be the electrode reaction and the less easily polarizable will be the electrode. It follows that the greater i_0 the less will be the overpotential at a given operating current.

Although at equilibrium $i_c = i_a$, when a potential more negative than the equilibrium potential is applied to an electrode, the cathodic reaction is accelerated and the anodic reaction is retarded so that $i_c > i_a$ and the net behaviour of the electrode is cathodic, the net current flowing being equal to $(i_c - i_a)$. Conversely, when a more positive potential is applied to the electrode the anodic reaction is accelerated and the cathodic reaction is retarded so that $i_c < i_a$ and the net behaviour is anodic with a net current of $(i_a - i_c)$.

Overpotential, η, is related quantitatively to the operating current density, i, of a cathode by the expression

$$i = i_0\left[\exp\left(\frac{azF\eta}{RT}\right) - \exp\left(\frac{(1-a)zF\eta}{RT}\right)\right],$$

where a is a constant for a particular system. This expression simplifies at low values of η to

$$\eta = \frac{RT}{zF}\frac{i}{i_0} \qquad \text{(I)}$$

and at high values of η to

$$\eta = b(\log i - \log i_o). \qquad \text{(II)}$$

It can be seen from both equations (I) and (II) that for a given operating current density i, the greater i_0, the less is η.

Values of exchange current densities range from about 10^{-2} A cm^{-2} to about 10^{-14} A cm^{-2} and some values for particular systems at comparable ionic concentrations are given in Table 1.

Table 1[1]
Exchange Current Densities

Reaction	*Electrode*	*$-\log i_0$*
$H^+ + e^- \longrightarrow \frac{1}{2}H_2$	Hg	11·7
	Pt	2·6
$4OH^- \longrightarrow O_2 + 2H_2O + 4e^-$	Pt	12·4
$Cl^- \longrightarrow \frac{1}{2}Cl_2 + e^-$	Pt	2·3
$Cu^{2+} + 2e^- \longrightarrow Cu$	Cu	4·7
$Zn^{2+} + 2e^- \longrightarrow Zn$	Zn	4·7
$Fe^{2+} + 2e^- \longrightarrow Fe$	Fe	8·0
$Ni^{2+} + 2e^- \longrightarrow Ni$	Ni	8·7
$Fe^{3+} + e^- \longrightarrow Fe^{2+}$	Pt	2·3

Overpotential is most important in the cases of oxygen evolution and hydrogen evolution at metals other than platinum. Most cases of metal deposition involve little or no overpotential at ordinary currents except in the cases of iron, cobalt and nickel.

---

[1] *Handbook of Electrochemical Constants* compiled by R. Parsons, Butterworths, 1959.

The very low exchange current density for the discharge of hydrogen on mercury ($\log i_0 = -11{\cdot}7$) should be noted. This accounts partly for the use of mercury cathodes in the electrolysis of brine to produce caustic soda and chlorine.

Bibliography

C. W. DAVIES, *Principles of Electrolysis*, 2nd edn, Royal Institute of Chemistry, 1968.
A. R. DENARO, *Elementary Electrochemistry*, Butterworths, 1965.
E. H. LYONS, JR, *Introduction to Electrochemistry*, Harrap, 1967.

9 A. G. Sharpe

Oxidation and Reduction

A. G. Sharpe, 'Oxidation and reduction', *Chemistry Student*, vol. 2 (1968), no. 1, pp. 16–19.

Introduction

Most readers will already be familiar with definitions of oxidation and reduction in terms of addition and subtraction of oxygen or other non-metallic (electronegative) elements. Thus in the reaction

$$2FeCl_2+Cl_2 \longrightarrow 2FeCl_3$$

there has been an increase in the chlorine content of the chloride and the iron has been oxidized from the iron(II) to the iron(III) state, just as it is in the reaction

$$4FeO+O_2 \longrightarrow 2Fe_2O_3.$$

In science, however, we should always seek to represent situations as faithfully as possible, partly because the search after truth is an intrinsically satisfying operation, and partly because accurate analysis of a situation affords the best basis upon which to make predictions concerning its future behaviour. If we are concerned with the action of dry chlorine on anhydrous iron(II) chloride, the above equation is admirable; but if we are concerned with a reaction in weakly acidic aqueous solution, in which both chlorides are completely ionized and the cations are both present as hydrated ions, a better representation would be

$$2Fe^{2+}(aq)+Cl_2 \longrightarrow 2Fe^{3+}(aq)+2Cl^-(aq).$$

If it was clear that we were writing about aqueous conditions the (aq)s could, of course, be omitted. In this case it is seen that, as in the reaction

$$Na(s)+\tfrac{1}{2}Cl_2(g) \longrightarrow Na^+Cl^-(s),$$

the result of the process is an increase in positive charge on the metal and acquisition of a negative charge by the chlorine atom. Electrons are therefore being transferred from the iron or the sodium to the chlorine, and oxidation and reduction may be redefined as loss and gain of electrons respectively.

So long as we are dealing with species which exist as ions under the conditions being considered, we can always represent oxidation and reduction faithfully in terms of transfer of electrons; for example in the Daniell cell the reactions taking place are

$$Cu^{2+}+2e^- \longrightarrow Cu$$
$$Zn \longrightarrow Zn^{2+}+2e^-$$

giving, over all,

$$Cu^{2+}+Zn \longrightarrow Cu+Zn^{2+}.$$

It should, however, be noted that sometimes transfer of atoms is also inevitably involved; when ethanol is oxidized to acetaldehyde or dichromate reduced to chromium(III) ion, we see from the equations

$$C_2H_5OH \longrightarrow CH_3CHO+2H^+ +2e^-$$
$$Cr_2O_7^{2-}+14H^+ +6e^- \longrightarrow 2Cr^{3+}+7H_2O$$

that hydrogen ions play an essential part in the processes, and that, in both, atoms as well as electrons are being transferred. In balancing equations, it is therefore sometimes simpler to think first in terms of transfer of oxygen, for example

$$K_2Cr_2O_7 \equiv K_2O.Cr_2O_3.3O$$

and $$C_2H_5OH+O \longrightarrow CH_3CHO+H_2O$$

therefore $$K_2Cr_2O_7 \equiv 3C_2H_5OH,$$

than to struggle with a set of so-called *ion–electron* equations and from the two equations given earlier to derive the relationship:

$$Cr_2O_7^{2-}+3C_2H_5OH+8H^+ \longrightarrow 2Cr^{3+}+3CH_3CHO+7H_2O.$$

Sometimes the use of the *oxidation number* method simplifies chemical stoichiometry. In this, everything is represented (often, it need hardly be said, quite unfaithfully) as a purely ionic species; atoms in the elementary state are taken to have zero oxidation number, and all negative charge is placed on the most electronegative element present. Thus potassium permanganate could be represented as $K^+Mn^{7+}4O^{2-}$, and its reduction to a salt containing the Mn^{2+} ion involves a change of five in oxidation number; this would mean, for example, that one MnO_4^- ion would oxidize five Fe^{2+} to Fe^{3+}. In as much as the ion–electron method portrays most accurately the nature of the chemical species present, it is the one most useful for a discussion of general principles.

We may now raise some general points about the study of oxidation and reduction. Is it possible to express oxidizing or reducing powers in a quantitative way? If it is, is the order obtained (or better still, the body of numerical data) independent of the environment being considered, for example whether reactions are carried out between anhydrous substances or in various solvents? Can we predict equilibrium constants for these reactions? Does the sequence tell us anything about the *speed*, as distinct from the *extent*, of the reductions it effects? And is it possible to find out anything about the details of how oxidation–reduction reactions take place? These are some of the questions to which the remainder of this article will attempt to give answers.

Quantitative Aspects of Oxidation and Reduction

If a rod of a metal M is immersed in a solution containing its ion M^{z+}, a potential difference is set up between the metal and the solution, and its magnitude depends upon the metal and ion concerned, the temperature and the activity (or, approximately, the concentration) of the ion in solution. For electrochemical work the standard temperature is taken as 25°C. It is, however, impossible to measure the potential of a single electrode, and hence all potentials are referred to an arbitrarily chosen standard – the standard hydrogen electrode. This consists of a plate of platinum coated electrolytically with platinum black, partly immersed in hydrochloric acid containing hydrogen ions at unit activity; hydrogen gas at a pressure of one atmosphere is bubbled through the solution in such a way that the platinum is in contact with both the gas and the solution; its function is to catalyse the attainment of equilibrium. The potential of the M^{z+}/M electrode is then determined by connecting it to the standard electrode by means of a salt bridge; other electrodes which have previously been calibrated against the standard hydrogen electrode may, of course, replace it, and a calomel electrode (mercury in contact with a saturated solution of mercury(I) chloride in potassium chloride solution of known concentration) is widely used.

The standard potential, E°, associated with the change

$$M^{z+}\ (aq, a = 1) + ze^- \longrightarrow M$$

relative to that for

$$H^+\ (aq, a = 1) + e^- \longrightarrow \tfrac{1}{2}H_2$$

is thus obtained. If both reactions took place on gram-ionic scale the difference in the energy changes involved would then be zFE° joules, where F is the Faraday; substitution of values for the Faraday and the electrical equivalent of heat converts this expression into $23{\cdot}06zE^\circ$ kcal. E° is therefore a measure of the energy change per electron transferred relative to the value for the standard hydrogen electrode, which is taken as zero. (It is worth noting that examination of the energy changes associated with the reactions

$$H^+(aq) \longrightarrow H^+(g) \xrightarrow{+e^-} H(g) \longrightarrow \tfrac{1}{2}H_2(g)$$

shows that the over-all energy change is, in fact, $\Delta H^\circ = -103$ kcal; but this, being difficult to determine, is only an approximate value. Great care is needed when thermochemical data obtained from electrical measurements are combined with those obtained by calorimetry.)

In place of M^{z+} and M, we may consider any system involving an oxidizing agent and its reduced form, and the standard potential of the system is then that for the change

$$\text{Oxidant } (aq, a = 1) + ne^- \longrightarrow \text{reductant } (aq, a = 1)$$

relative to that for the standard hydrogen electrode. If the system is one in which equilibrium is rapidly attained, such as

$$\tfrac{1}{2}Cl_2(g) + e^- \longrightarrow Cl^-(aq, a = 1)$$

or

$$Fe^{3+} + e^- \longrightarrow Fe^{2+},$$

the potential can be measured directly; in other cases, calculation from independently obtained thermochemical data is often possible. Data for several systems are given in the table below; where hydrogen ions also appear in the system, the value of E° is that for unit hydrogen ion activity. The more negative or less positive E°, the harder it is to bring about the change from left to right; combination of two E° values therefore tells us in which direction a reaction such as

$$2Fe^{3+} + Sn^{2+} \longrightarrow 2Fe^{2+} + Sn^{4+}$$

will proceed, and can, indeed, be used to calculate an equilibrium constant, which can be shown to be given by the expression

$$\log K = \frac{\Delta E^\circ \times n}{0{\cdot}059}$$

where n is the number of electrons transferred. Thus a table of standard potentials provides a very useful summary of a large amount of chemistry.

The limitations of such a summary become apparent, however, when we recall how the data are obtained: they are measured, or derived from measurements made, in aqueous solution and therefore relate only to aqueous conditions, and do not hold for anhydrous systems or other solvents.

Ions in aqueous media are not usually present at unit activity, but standard potentials can be converted into those for any other activities of oxidant and reductant by means of the expression (at 25°C)

$$E = E^\circ + \frac{0{\cdot}059}{n} \log \frac{[\text{oxidant}]}{[\text{reductant}]}.$$

However, where complex ion formation or precipitation is involved, it is also necessary to know formation constants or solubility products before activities of ions can be calculated. Although, for example, according to the table, iron(III) should oxidize iodide ion, no reaction occurs in the presence of fluoride; this is because fluoride forms a stable complex with iron(III) but only a weak complex with iron(II). Similarly, cyanide forms a more stable complex with iron(III) than with iron(II) (though both complexes have very large formation constants), and, as

Table 1
Standard Redox Potentials

Reaction	E° (*volt*)
$Li^+ + e^- = Li$	−3·04
$K^+ + e^- = K$	−2·92
$Ca^{2+} + 2e^- = Ca$	−2·87
$Na^+ + e^- = Na$	−2·71
$Mg^{2+} + 2e^- = Mg$	−2·37
$Al^{3+} + 3e^- = Al$	−1·66
$Mn^{2+} + 2e^- = Mn$	−1·18
$Zn^{2+} + 2e^- = Zn$	−0·76
$Cr^{3+} + 3e^- = Cr$	−0·74
$Fe^{2+} + 2e^- = Fe$	−0·44
$Cr^{3+} + e^- = Cr^{2+}$	−0·41
$Co^{2+} + 2e^- = Co$	−0·28
$Sn^{2+} + 2e^- = Sn$	−0·14
$Pb^{2+} + 2e^- = Pb$	−0·13
$H^+ + e^- = \frac{1}{2}H_2$	0·00
$S + 2H^+ + 2e^- = H_2S$	+0·14
$Sn^{4+} + 2e^- = Sn^{2+}$	+0·15
$Cu^{2+} + e^- = Cu^+$	+0·15
$\frac{1}{2}N_2 + 4H^+ + 3e^- = NH_4^+$	+0·27
$Cu^{2+} + 2e^- = Cu$	+0·34
$[Fe(CN)_6]^{3-} + e^- = [Fe(CN)_6]^{4-}$	+0·36
$Cu^+ + e^- = Cu$	+0·52
$\frac{1}{2}I_2 + e^- = I^-$	+0·54
$Fe^{3+} + e^- = Fe^{2+}$	+0·76
$Ag^+ + e^- = Ag$	+0·80
$Hg^{2+} + 2e^- = Hg$	+0·85
$\frac{1}{2}Br_2 + e^- = Br^-$	+1·07
$IO_3^- + 6H^+ + 5e^- = \frac{1}{2}I_2 + 3H_2O$	+1·19
$\frac{1}{2}O_2 + 2H^+ + 2e^- = H_2O$	+1·23
$\frac{1}{2}Cr_2O_7^{2-} + 7H^+ + 3e^- = Cr^{3+} + \frac{7}{2}H_2O$	+1·33
$\frac{1}{2}Cl_2 + e^- = Cl^-$	+1·36
$MnO_4^- + 8H^+ + 5e^- = Mn^{2+} + 4H_2O$	+1·52
$MnO_4^- + 4H^+ + 3e^- = MnO_2 + 2H_2O$	+1·69
$Pb^{4+} + 2e^- = Pb^{2+}$	+1·70
$Co^{3+} + e^- = Co^{2+}$	+1·82
$Ag^{2+} + e^- = Ag^+$	+1·98
$\frac{1}{2}S_2O_8^{2-} + e^- = SO_4^{2-}$	+2·01
$\frac{1}{2}F_2 + e^- = F^-$	+2·80

can be seen, hexacyanoferrate(III) (ferricyanide) is harder to reduce than aquated iron(III) ion.

The effect of hydrogen ion (or, conversely, hydroxide ion) concentration on some potentials is easily understood: if either ion is a reactant or product, its concentration *must* affect the position of equilibrium. Thus we can understand why potassium dichromate is a more powerful oxidant in strongly acidic than in weakly acidic media; and why, since iron(III) hydroxide is much less soluble than iron(II) hydroxide, it is easier to oxidize iron(II) ion in alkaline than in acidic media. Under the latter conditions, the oxidation is represented by the equation

$$Fe^{2+} \longrightarrow Fe^{3+} + e^{-}.$$

In alkaline media, this becomes

$$Fe(OH)_2 + OH^{-} \longrightarrow Fe(OH)_3 + e^{-}.$$

Hydroxylamine, NH_2OH, for instance, reduces iron(III) in acidic solution but oxidizes iron(II) in alkaline solution.

The biggest limitation of all in the representation of inorganic chemistry by a table of standard potentials (and auxiliary data on formation constants and solubility products) lies, however, in the fact that they tell us nothing about the *rates* at which reaction will proceed. This is a limitation of the whole branch of science known as thermodynamics, which is concerned only with the difference in energy between initial and final states, and not with the mechanism by which the change is accomplished. Thus although from a glance at the table we see that under equilibrium conditions acidic permanganate will oxidize water to oxygen, or oxygen will convert metallic silver to silver ions in solution, the rates at which these changes take place are extremely slow. Considerable progress has been made in recent years in elucidating the mechanisms of chemical processes, but we have a long way to go before we can predict the rates of any but a very few of the simplest reactions.

Factors which Determine the Magnitude of Standard Potentials

It is instructive to see what factors determine the values for standard electrode potentials, and in simple cases such as those for an ion M^+ being reduced to a metal M this can readily be done by considering the energy changes associated with the various stages in the sequence

$$\underset{\text{(i)}}{M^+(aq) \longrightarrow M^+(g)} \xrightarrow[\text{(ii)}]{+e^-} \underset{\text{(iii)}}{M(g) \longrightarrow M(s)}$$

These are (i) the hydration energy of the ion, (ii) the ionization energy, and (iii) the energy of atomization. Values for ΔH° for sodium and silver (in kcal) are tabulated below.

	(i)	(ii)	(iii)	*Sum* (kcal)
Na	+97	−118	−26	−47
Ag	+116	−174	−67	−125

The differences in ionization energy and energy of atomization contribute about equally to the over-all difference between the two elements, and the difference in hydration energies is relatively small; if we compare the over-all values with that of $\Delta H^\circ = -103$ kcal for the change

$$H^+(aq, a = 1) + e^- \longrightarrow \tfrac{1}{2}H_2$$

then we see how the Na^+/Na electrode comes to have a very negative standard potential and the Ag^+/Ag electrode a positive one.

A similar comparison may be made for fluorine and chlorine by considering ΔH° for the stages

$$\tfrac{1}{2}X_2(\text{standard state}) \xrightarrow[\text{(i)}]{} X(g) \xrightarrow[\text{(ii)}]{+e^-} X^-(g) \xrightarrow[\text{(iii)}]{} X^-(aq)$$

	(i)	(ii)	(iii)	*Sum* (kcal)
F	+19	−81	−121	−183
Cl	+29	−85	−88	−144

Somewhat surprisingly, perhaps, it turns out that the hydration energy term (iii) is by far the most important factor in making fluorine the most powerful oxidizing agent among the halogens under aqueous conditions; this is undoubtedly the consequence of the small size of the fluoride ion.

Oxidation and Reduction under Non-aqueous Conditions

In a solvent other than water it would, in principle, be possible to measure redox potentials, complex ion formation constants and solubility products, and to develop a quantitative treatment analogous to that for aqueous systems. So far, however, little has been done in this direction. But it is instructive to note that water, being a 10^{-7}M acid, is relatively easily reduced; *E* for the system

$$H^+(aq,\ 10^{-7}\text{M}) + e^- \longrightarrow \tfrac{1}{2}H_2$$

can be calculated as −0·4 V. Furthermore the value of +1·23 V for the system

$$\tfrac{1}{2}O_2 + 2H^+ + 2e^- \longrightarrow H_2O$$

shows that water is fairly easily oxidized. In general, therefore, we can see that reductions in solution are usually best carried out in a solvent less acidic than water (for example ethanol or liquid ammonia), and oxidations in solution are best carried out in solvents such as hydrogen fluoride which are very difficult to oxidize.

But what of reactions which do not involve a solvent at all? Let us consider as a simple example the reaction

$$KCl(s) + \tfrac{1}{2}F_2(g) \longrightarrow KF(s) + \tfrac{1}{2}Cl_2(g).$$

Analysis of this reaction by splitting it into the stages

$$KCl(s) \longrightarrow K^+(g) + Cl^-(g) \qquad \text{(i)}$$
$$Cl^-(g) \longrightarrow \tfrac{1}{2}Cl_2(g) + e^- \qquad \text{(ii)}$$
$$\tfrac{1}{2}F_2(g) + e^- \longrightarrow F^-(g) \qquad \text{(iii)}$$
$$K^+(g) + F^-(g) \longrightarrow KF(s) \qquad \text{(iv)}$$

shows that the total ΔH° of −30 kcal comes mainly from the difference between (i) and (iv); the difference between (ii) and (iii) can be seen from data given already to be only 6 kcal. The energy change when separated gaseous ions are converted into an ionic solid is called the *lattice energy* of the solid, and it is inversely proportional to the interionic distance in the crystal; so the smaller size of the fluoride ion has again been the most important factor. Perhaps this conclusion is not really so surprising: the interaction of a cation and the negative ends of polar water molecules and the interaction of the same cation and anions in a solid are similar processes. There is, therefore, a *general* resemblance between oxidizing powers in the presence and in the absence of water, but because of the complications in aqueous media that have already been mentioned, too close a similarity must not be expected.

The Mechanism of Oxidation–Reduction Reactions in Aqueous Solution

Many inorganic reactions are too fast or too slow for convenient study, but modern physicochemical techniques and the availability of isotopes have made it possible to obtain considerable insight into some oxidation mechanisms in recent years, and to conclude this survey mention will now be made of a few investigations in this field.

Neither hexacyanoferrate(II) (ferrocyanide) nor hexacyanoferrate(III) (ferricyanide) exchanges cyanide at a measurable rate with an aqueous solution containing $^{14}CN^-$, but when a solution containing hexacyanoferrate(II) (ferrocyanide) labelled with radioactive iron is mixed with one containing hexacyanoferrate(III) (ferricyanide) not so labelled, it is found that random distribution of the radioactive iron results immediately. This is thus a case of rapid *electron transfer* between two species, which, it may be noted, are similar in both shape and size, so that no radical geometrical change takes place as a result of the reaction.

The complex ion $[Co(NH_3)_5Cl]^{2+}$ exchanges chlorine with radioactive chloride only very slowly (many complexes of cobalt(III) and chromium(III) are shown by this absence of rapid exchange reactions to be kinetically inert, whereas similar complexes of cobalt(II) and chromium(II) undergo fast exchange reactions). When the $[Co(NH_3)_5Cl]^{2+}$ ion is reduced by Cr^{2+}(aq) the product from the latter is not $[Cr(H_2O)_6]^{3+}$ but the (kinetically inert) chromium(III) complex $[Cr(H_2O)_5Cl]^{2+}$; in this case, therefore, the chlorine atom has been handed over directly in the oxidation reaction, and it is believed that the halogen acts as a bridge in a reaction intermediate of formula $[(H_3N)_5CoClCr(H_2O)_5]^{4+}$. The formation of bridged complexes of this kind is now recognized as being very common, and can often be inferred from kinetic studies. In the Fe^{2+}/Fe^{3+} aqueous system, for example, it is found that the rate of exchange decreases with increase in the hydrogen ion concentration of the solution. Even in this simple instance, therefore, it seems that an intermediate such as

$$[(H_2O)_5Fe(OH)Fe(H_2O)_5]^{4+}$$

formed from $[(H_2O)_5FeOH]^{2+}$ and $[Fe(H_2O)_6]^{2+}$ is involved in one possible reaction path. Halide ions can also act as bridging entities in such reactions, so we see that even simple reactions such as the one with which we started may still repay detailed examination.

Bibliography

W. L. JOLLY, 'The use of oxidation potentials in inorganic chemistry', *J. Chem. Educ.*, vol. 43 (1966), no. 4, pp. 198–201.

W. M. LATIMER, *Oxidation Potentials*, 2nd edn, Prentice-Hall, 1952.

A. G. SHARPE, *Principles of Oxidation and Reduction*, 2nd edn, Royal Institute of Chemistry, 1968.

D. B. SOWERBY and M. F. A. DOVE, 'Oxidation states in inorganic compounds', *Educ. Chem.*, vol. 1 (1964), no. 2, pp. 83–90.

Part Four **Acids and Bases**

Many chemical substances can be classified as 'acids' or 'bases' and various theories have been proposed to account for acid–base behaviour. D. Nicholls traces the evolution of 'Theories of acids and bases' including the Arrhenius, Brønsted–Lowry, solvent system and Lewis theories. The Arrhenius theory is limited to acid–base reactions in aqueous solution, but the other theories may be applied to reactions in non-aqueous solvents. A. K. Holliday gives a general survey of inorganic chemistry in non-aqueous solvents such as liquid ammonia in his article on 'Non-aqueous solvents'. An expanded version of the latter article is available in book form (see 'References' at the end of the article).

10 D. Nicholls

Theories of Acids and Bases

D. Nicholls, 'Theories of acids and bases', *Chemistry Student*, vol. 1 (1967), no. 2, pp. 33–8.

Acids and alkalis have been known for several centuries. At first, acids distinguished themselves from other substances by giving an effervescence with limestone and by turning certain plant dyes from blue to red. Alkalis, too, became grouped together because they had characteristic properties such as affecting the colours of plant dyes and of neutralizing the effects of acids. In 1744, Rouelle introduced the term 'bases' for the more general class of compounds which react with acids to form salts.

Lavoisier was the first to attempt to understand the chemistry of acids. He had found that acids were produced when non-metals such as carbon, phosphorus and sulphur were burned in oxygen and so he proposed that the characteristic properties of acids were due to the presence of oxygen. The realization by Davy that hydrogen chloride contained no oxygen led him to propose that it was hydrogen which was the element characteristically found in all acids. In order to specify more precisely the nature of this hydrogen, Liebig showed that all the then known acids contained a hydrogen atom which could be replaced by a metal. The first quantitatively useful model of acids and bases came during the period 1880–90 when Arrhenius developed his theory of electrolytic dissociation.

The Arrhenius Theory

On this theory, acids and bases are defined as follows:

acid – a hydrogen-containing compound which gives hydrogen ions in aqueous solution;

base – any hydroxyl compound which gives hydroxide ions in aqueous solution.

Thus neutralization is represented by

$$H^+ + OH^- \longrightarrow H_2O.$$

This model was a great step forward. It explained, for example, the constant heat of neutralization of any strong acid with any strong base, as these reactions all involve only the combination of hydrogen and hydroxide ions. It rationalized also the catalytic properties of acids. It had been known for some time that the power to act as a catalyst was another characteristic property of acids; simple experiments soon showed that the catalytic effect of acids increased as the concentration of hydrogen ions increased.

Despite the obvious success here, it soon became apparent that the definition limits all acid–base reactions to water solution. At the beginning of this century Franklin and Kraus were beginning their classic studies of reactions in liquid ammonia as solvent. It was shown that ammonia is very much like water in its solvent properties and that reactions akin to acid–base reactions in water could be carried out in ammonia. 'Acid–base' titrations could be performed with the usual indicators, for example phenolphthalein, so that it became necessary to use the words 'acid' and 'base' in this solvent also. There were other fields of chemistry where acid–base reactions were excluded by the definition. For example, in organic chemistry, many compounds were capable of acting as bases (for example aniline, pyridine and so on), yet these contained no hydroxide ions. Reactions in the gas phase were also excluded by the definition. To overcome these difficulties a more general theory was needed; it was proposed in 1923 independently by Brønsted in Denmark and Lowry in England.

The Brønsted–Lowry Theory

This is sometimes known as the protonic theory – it removes the necessity for a base to provide hydroxide ions and acids and bases are defined as follows:

acid – a species which tends to give up a proton;
base – a species which tends to accept a proton.

This definition now includes all the Arrhenius compounds but considerably extends the scope of acids and bases. It requires no particular type of ion or any particular solvent. Hydrogen chloride becomes an acid by virtue of the fact that it can donate a proton – in the gas phase or in a non-aqueous solvent as well as in water.

An acid–base reaction can be regarded as a competition for a proton by two different bases. In the reaction of hydrogen chloride with water, the resultant chloride ion is a base because it can accept a proton; these acid–base pairs, for example HCl and Cl^-,

$$\underset{\textit{acid}}{HCl} + \underset{\textit{base}}{H_2O} \longrightarrow \underset{\textit{acid}}{H_3O^+} + \underset{\textit{base}}{Cl^-}$$

are called ‘conjugate pairs’. In this reaction water is a base because it accepts a proton; hydrogen chloride is an acid not because it contains H^+ but because it gives up a proton. The reaction proceeds in the direction shown because the chloride ion is a much weaker base than water. The definition provides for three types of acid and three types of base.

Brønsted–Lowry acids

Molecular: $HCl \longrightarrow H^+ + Cl^-$.
Cationic: $[Fe(H_2O)_6]^{3+} \longrightarrow H^+ + [Fe(H_2O)_5(OH)]^{2+}$.
Anionic: $HSO_4^- \longrightarrow H^+ + SO_4^{2-}$.

Brønsted–Lowry bases

Molecular: $NH_3 + H^+ \longrightarrow NH_4^+$.
Cationic: $[Fe(H_2O)_5(OH)]^{2+} + H^+ \longrightarrow [Fe(H_2O)_6]^{3+}$.
Anionic: $HSO_4^- + H^+ \longrightarrow H_2SO_4$.

The equations in these examples are written as simply as possible; the protons are, of course, solvated to H_3O^+, for example, in aqueous solution. Let us look at one of these examples in a little more detail. When iron(III) chloride is added to water, a strongly acidic solution results. Because the first reaction of iron(III) chloride with water is hydration to $[Fe(H_2O)_6]Cl_3$ and since no precipitation of the hydroxide occurs, the equation

$$FeCl_3 + 3H_2O \longrightarrow Fe(OH)_3 + 3HCl$$

(found in many of the older textbooks) does not represent the chemistry of the process. The acidity arises from the proton donor ability of the $[Fe(H_2O)_6]^{3+}$ ion:

$$\underset{\textit{acid}}{[Fe(H_2O)_6]^{3+}} + \underset{\textit{base}}{H_2O} \rightleftharpoons \underset{\textit{base}}{[Fe(H_2O)_5(OH)]^{2+}} + \underset{\textit{acid}}{H_3O^+}.$$

Some species may act either as acids or bases, for example water and the hydrogen carbonate ion:

$$\begin{tikzcd}
 & H_2O \arrow[dl, "-H^+"'] \arrow[dr, "+H^+"] & & & HCO_3^- \arrow[dl, "-H^+"'] \arrow[dr, "+H^+"] & \\
OH^- & & H_3O^+ & CO_3^{2-} & & H_2CO_3
\end{tikzcd}$$

The Brønsted–Lowry approach is particularly useful in aqueous solution as the quantitative aspects of acid–base titrations can be readily understood. The tendency of an acid to give up a proton in water is measured by the equilibrium constant (K) of the reaction

$$HA + H_2O \rightleftharpoons A^- + H_3O^+$$

$$K = \frac{[A^-][H_3O^+]}{[HA][H_2O]}.$$

Hydrogen chloride, for example, has $K = 10^7$ or, as it is more usually quoted, $pK = -7$ (where $pK = -\log_{10} K$). Hydrogen fluoride on the other hand has a pK of about $+3$ showing it to be a weaker acid in water than hydrogen chloride. At first sight we would expect that the more electronegative the group attached to the proton, the stronger the acid will be. This is in fact found to be the case when considering the hydrides in the first row of the periodic table; for NH_3, $pK = 35$; for H_2O, $pK = 16$; and for HF, $pK = 3$. Within the halogen hydracids, however, the unexpected order of acidity (HI > HBr > HCl > HF) is explained by the increased H—X bond strength as we go along this series. Thus the strong H—F bond does not give up the hydrogen ion as readily as the weaker bonded H—Cl.

Despite the quantitative uses of the theory, Brønsted–Lowry acids and bases are still restricted to systems containing the proton. The theory can tell us nothing about acids and bases in non-protonic solvents (for example N_2O_4 or ICl). Because water is such a very reactive chemical, modern inorganic chemistry is frequently concerned with reactions in non-aqueous solvents and there is obviously a need for definitions of acids and bases which are independent of the proton.

The Solvent System Concept

On this theory we have:

acid – a solute giving the cation characteristic of the solvent;
base – a solute giving the anion characteristic of the solvent.

To define what we mean by the cation and anion characteristic of the solvent we must write down the autoionization of the solvent, so that for water and liquid ammonia we write:

$$2H_2O \rightleftharpoons OH_3^+ + OH^-,$$
$$2NH_3 \rightleftharpoons NH_4^+ + NH_2 .$$

Just as substances which produce H_3O^+ in water are acids, so then ammonium salts in liquid ammonia are acidic; soluble hydroxides are bases in water and soluble amides are bases in ammonia. Typical neutralization reactions are then:

$$\begin{array}{ccc} H_3OCl + KOH & \rightarrow & KCl + 2H_2O, \\ NH_4Cl + KNH_2 & \rightarrow & KCl + 2NH_3. \\ \textit{acid} \ + \ \textit{base} & \rightarrow & \textit{salt} \ + \ \textit{solvent} \end{array}$$

Note that this last general equation holds for all solvents, not just for water. Just as in water, acids in liquid ammonia react with metals forming salts and hydrogen:

$$2NH_4Cl + Ca \rightarrow CaCl_2 + H_2 + 2NH_3,$$
cf. $$2H_3OCl + Ca \rightarrow CaCl_2 + H_2 + 2H_2O.$$

Amphoteric behaviour too can be illustrated with liquid ammonia. Zinc amide is soluble in an excess of potassium amide solution in ammonia to form the amido-zincate:

$$Zn(NH_2)_2 + 2KNH_2 \rightarrow K_2Zn(NH_2)_4,$$
cf. $$Zn(OH)_2 + 2KOH \rightarrow K_2Zn(OH)_4.$$

We have chosen liquid ammonia as our example here because, after water, it is by far the most widely studied inorganic solvent. Ammonia is a protonic solvent, however, and we must now see the advantages of the solvent system definition of acids and bases in solvents which do not contain the (solvated) proton. Some of these solvents are listed in the following table.

Table 1
Non-Protonic Solvents

Solvent	*Autoionization*
N_2O_4	$N_2O_4 \rightleftharpoons NO^+ + NO_3^-$
ICl	$2ICl \rightleftharpoons I^+ + ICl_2^-$
BrF_3	$2BrF_3 \rightleftharpoons BrF_2^+ + BrF_4^-$
$SOCl_2$	$SOCl_2 \rightleftharpoons SOCl^+ + Cl^-$

In the solvent, the acids and bases usually arise by reaction of the solute with the solvent. For example, in iodine monochloride the ICl_2^- ion is produced by soluble ionic chlorides:

$$KCl + ICl \longrightarrow K^+ICl_2^-.$$

Thus KCl, or rather $KICl_2$, is a base in iodine monochloride. The iodide cation is produced by soluble chloride ion acceptors (Lewis acids – see later), that is, covalent chlorides which can form complexes with the chloride ion, for example

$$ICl + SbCl_5 \longrightarrow I^+SbCl_6^-.$$

A typical neutralization reaction is then:

$$\underset{\textit{acid}}{ISbCl_6} + \underset{\textit{base}}{KICl_2} \longrightarrow \underset{\textit{salt}}{KSbCl_6} + \underset{\textit{solvent}}{2ICl}.$$

Examples of acid–base behaviour could, of course, be taken from any ionizing solvent. By means of research in these non-aqueous solvents we have been able to extend our chemical knowledge considerably and prepare a large number of new chemical compounds which are unstable in water. The solvent system definition of acids and bases has been of considerable help to chemists working in this field of research. It is, however, still limited to reactions occurring in solution; its usefulness lies in extending acid–base concepts to solvents which do not contain the proton.

All the definitions that we have used so far have had limitations and we have seen how these limitations have been minimized by the introduction of new definitions. In 1923, G. N. Lewis propounded an acid–base theory which embraced all those so far mentioned and yet covered reactions in the gas phase and in inert (non-ionizing) solvents as well.

The Lewis (or Electronic) Theory

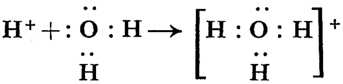
If our definition of a base is to include all such species as OH^-, Cl^-, NH_3 and pyridine, then we must search for the fundamental property that these all have in common. Lewis realized that acidity and basicity could be related to the electronic structure of ions and molecules. Thus all the bases mentioned above have one thing in common – namely a pair of electrons available for chemical bonding. These pairs of electrons are often referred to as 'lone-pair electrons'. Acids on the other hand are substances which have vacant orbitals which can accept the lone pair of electrons from the base. The Lewis or electronic definitions are thus:

acid – an electron pair acceptor;
base – an electron pair donor.

Acids are often just called 'acceptors' and bases 'donors', particularly in coordination chemistry. Neutralization then involves the formation of a coordinate (covalent) bond, for example

$$H^+ + :\underset{\cdot\cdot}{\overset{\cdot\cdot}{O}}:H^- \longrightarrow H:\underset{\cdot\cdot}{\overset{\cdot\cdot}{O}}:H$$

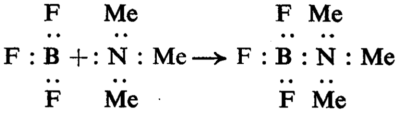
The hydrogen ion has an empty (1*s*) atomic orbital and it 'accepts' a pair of electrons from one of the lone pairs (doubly filled sp^3 orbital) on the oxygen of the hydroxide ion. The species formed in this reaction has two lone pairs of electrons and so will be basic; consequently, it reacts with the hydrogen ion:



$$H^+ + :\underset{\underset{\displaystyle H}{\cdot\cdot}}{\overset{\cdot\cdot}{O}}:H \longrightarrow \left[H:\underset{\underset{\displaystyle H}{\cdot\cdot}}{\overset{\cdot\cdot}{O}}:H\right]^+$$

The H_3O^+ ion so formed is not basic because it is positively charged and consequently will be very reluctant to donate its last lone pair of electrons.

A more familiar example perhaps of coordinate bond formation is the reaction of boron trifluoride (a Lewis acid) with trimethylamine (a Lewis base):



$$\mathrm{F}:\underset{\underset{\textstyle\mathrm{F}}{\cdot\cdot}}{\overset{\overset{\textstyle\mathrm{F}}{\cdot\cdot}}{\mathrm{B}}}+:\underset{\underset{\textstyle\mathrm{Me}}{\cdot\cdot}}{\overset{\overset{\textstyle\mathrm{Me}}{\cdot\cdot}}{\mathrm{N}}}:\mathrm{Me}\longrightarrow\mathrm{F}:\underset{\underset{\textstyle\mathrm{F}}{\cdot\cdot}}{\overset{\overset{\textstyle\mathrm{F}}{\cdot\cdot}}{\mathrm{B}}}:\underset{\underset{\textstyle\mathrm{Me}}{\cdot\cdot}}{\overset{\overset{\textstyle\mathrm{Me}}{\cdot\cdot}}{\mathrm{N}}}:\mathrm{Me}$$

$(\mathrm{Me} = \mathrm{CH_3})$

Again the vacant (sp^3) orbital on boron receives an electron pair from the filled (sp^3) orbital on the nitrogen atom of the amine. The Lewis definition thus includes within its scope the formation of complex or coordination compounds. The presence of any particular solvent is immaterial and whether a neutralization reaction will occur depends solely upon the electronic structures

Table 2
Lewis Acids

Type	*Examples*
Positive ions	H^+, H_3O^+, Al^{3+}, Sn^{4+}, Cu^{2+}
Molecules with unfilled octets	$BeCl_2$, BF_3, BCl_3
Compounds, especially halides, in which the central atom may exceed its octet (i.e. has a vacant orbital of low energy)	$MgCl_2$, $SiCl_4$, PCl_5

Table 3
Lewis Bases

Type	*Examples*
Negative ions	OH^-, Cl^-, CN^-, SO_4^{2-}
Molecules with one or two lone pairs of electrons	NH_3, NMe_3, H_2O, Et_2O
Molecules containing carbon to carbon multiple bonds	$H_2C{=}CH_2$, $HC{\equiv}CH$, C_6H_6

of the reactants. It will be convenient now to tabulate the classes of compounds which are called acids and bases under this definition.

The inclusion of all cations as acids and all anions as bases perhaps makes the concept unnecessarily broad. The reaction

$$Na^+ + Cl^- \longrightarrow NaCl$$

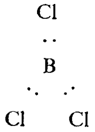
for example, does not involve acceptance or donation of electron pairs (NaCl is bonded ionically) and this then is not regarded as an acid–base reaction. The strengths of cations as Lewis acids varies over a wide range. In general, we can expect the acidity (or coordinating ability) of cations to increase as the charge on the cation increases. Thus Na^+ ions are weaker acids than say Al^{3+} ions; and Sn^{2+} ions are weaker acids than Sn^{4+} ions. In aqueous solutions all of these ions are attached to water molecules (bases) and when crystallized out of solution the ions of higher charge, for example Al^{3+} and Sn^{4+}, retain the water molecules as in, for example, $Al_2(SO_4)_3.18H_2O$ and $SnCl_4.5H_2O$. This is because the bonds between aluminium or tin(IV) ions and water are stronger than those between sodium ions and water; in other words, sodium ions are only weakly acidic, the bonding to water molecules being electrostatic rather than coordinate. Transition metal cations are generally strong Lewis acids and hence form a variety of complex ions with Lewis bases. Again, as water is a base, reactions of transition metal ions in dilute aqueous solution are nearly always reactions of the octahedrally shaped $[M(H_2O)_6]^{n+}$ cations. In the solid state too, many of the chemicals that we often regard as simple salts are really coordination complexes formed by the base water. Blue copper sulphate, for example, is really $[Cu(H_2O)_4]SO_4.H_2O$ in which four of the water molecules are bonded to the copper atom and the other one is hydrogen bonded between the oxygen atoms of two sulphate ions and the hydrogen atoms from two coordinated water molecules. Some examples of acid–base behaviour with cations and anions are the following:

$$Ag^+ + 2CN^- \longrightarrow [Ag(CN)_2]^-$$
$$Cu^{2+} + 4Cl^- \longrightarrow [CuCl_4]^{2-}$$
$$Al^{3+} + 6F^- \longrightarrow [AlF_6]^{3-}.$$

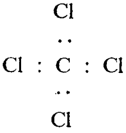
Under the second heading of Lewis acids we have molecules with unfilled octets. These molecules have a tendency to accept pairs of electrons from bases so as to complete their octets. The chlorides of beryllium, boron and carbon have the following structures:

$$\mathrm{Cl : Be : Cl}$$

acidic: needs two pairs to complete octet



acidic: needs one pair to complete octet



$$\begin{array}{c}\mathrm{Cl}\\ \cdot\cdot\\ \mathrm{Cl : C : Cl}\\ \cdot\cdot\\ \mathrm{Cl}\end{array}$$

neutral: has a complete octet

It becomes apparent that beryllium chloride will combine with two molecules of a base to complete its octet; it therefore reacts with two molecules of ether to form a tetrahedral coordination complex:

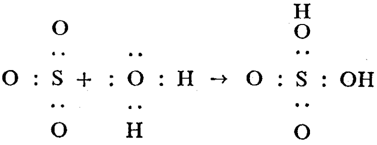
$$BeCl_2 + 2OEt_2 \longrightarrow BeCl_2.2OEt_2.$$

Boron trichloride on the other hand will form compounds with one molecule only of a base. With a chloride ion for example:

$$BCl_3 + Cl^- \longrightarrow BCl_4^-.$$

Carbon tetrachloride, however, will not react with ether, chloride ions or any Lewis base under normal conditions; it is neutral. As a general guide, acidity increases with increasing electronegativity (electron-withdrawing power) of the groups attached to the central atom; BF_3 for example is a stronger acid than $B(OH)_3$.

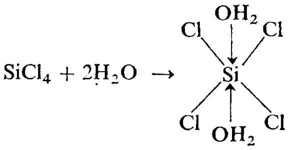
Compounds possessing double bonds between unlike atoms often function as Lewis acids. Typical of this class are CO_2, SO_2 and SO_3. After coordination of water to these molecules a proton shift often occurs (although apparently not in the case of SO_2):



$$\begin{array}{c}\mathrm{O}\\ \cdot\cdot\\ \mathrm{O : S}\\ \cdot\cdot\\ \mathrm{O}\end{array} + \begin{array}{c}\\ \cdot\cdot\\ \mathrm{: O : H}\\ \cdot\cdot\\ \mathrm{H}\end{array} \rightarrow \begin{array}{c}\mathrm{H}\\ \mathrm{O}\\ \cdot\cdot\\ \mathrm{O : S : OH}\\ \cdot\cdot\\ \mathrm{O}\end{array}$$

(Sulphur electrons only shown in SO_3.)

Compounds in which the central metal atom may exceed its octet (or, putting it another way, may use *d*-orbitals for bonding) are found in the second and subsequent rows of the periodic table. Whilst carbon tetrachloride is not acidic, silicon and tin tetrachlorides are acidic; silicon and tin can exceed their octets by using vacant *d*-orbitals to accept electron pairs from bases. Whilst all three chlorides have similar physical properties (colourless liquids) the silicon and tin compounds are violently hydrolysed by water. The first step in this reaction is the coordination of water molecules to the acids:



Hydrolysis occurs in excess of water by elimination of hydrogen chloride:

$$SiCl_4(H_2O)_2+H_2O \longrightarrow SiCl_3(OH)(H_2O)_2+HCl$$

until finally $Si(OH)_4.2H_2O$ (hydrated silica) is precipitated. Carbon tetrachloride does not react with water because the first step, that is the coordination of water molecules, cannot occur.

We have seen many examples of anions acting as bases. Neutral molecules are basic if they have one or two lone pairs of electrons. It is instructive to look at the first two rows of the periodic table to see which compounds fall into this category. In the diagrams below the lines represent bonds to some other element (for example hydrogen) and the statements about the basic nature expected are broad generalizations.

In neutral molecules with three lone pairs of electrons, the acidity depends upon the nature of the element bonded to the halogen. In HF and HCl, because of the great electronegativity differences between the elements, the molecules are highly polar ($H^{\delta+}-Cl^{\delta-}$). The hydrogen atoms, which have only a small share of the electrons in the bond, are able to accept electrons

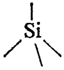
no lone pairs: neutral

one lone pair: basic

two lone pairs: basic

three lone pairs: neutral or acidic

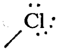
four lone pairs: neutral



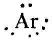
no lone pairs: acidic

one lone pair: basic

two lone pairs: basic

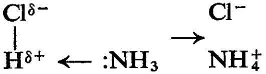


three lone pairs: neutral or acidic



four lone pairs: neutral

from a base if these are readily available, and then rupture of the H—X bond occurs:



$$\begin{matrix} \mathrm{Cl}^{\delta-} & & & \mathrm{Cl}^- \\ | & & \rightarrow & \\ \mathrm{H}^{\delta+} & \leftarrow :\mathrm{NH}_3 & & \mathrm{NH}_4^+ \end{matrix}$$

The base strength of nitrogen compounds depends upon the nature of the three atoms bonded to nitrogen. If these are hydrogen as in NH_3, or carbon as in NMe_3, then the lone pair is readily available for bonding and these are strong bases. With very electron-withdrawing atoms (for example fluorine) bonded to the nitrogen atom, however, the donor ability is much reduced because of the electron withdrawal along the N—F bonds. Thus in decreasing order of base strength we have

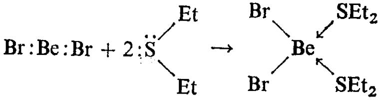
$$NMe_3 > NH_3 > NF_3.$$

Similar arguments apply to oxygen, sulphur and phosphorus compounds.

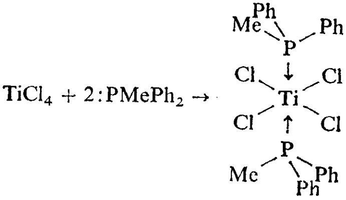
The third type of Lewis base concerns compounds with carbon to carbon multiple bonds. Acetylenes, olefins, allenes and even benzene are capable of forming coordination compounds by donation of electron pairs from the double bonds to transition metal ions (but not to the other Lewis acids). The theory of bonding in these compounds, however, is too detailed to be included in this article.

The great usefulness of the Lewis theory is that it enables us not only to comprehend a large section of chemistry but also to make an intelligent guess as to whether two compounds A and B will react together. If we already know some chemistry of the reactants and some basic stereochemistry, we may also make a prediction about what the formula and structure of the product might be. This does not mean that we are in a position to be able to predict chemistry; indeed the most interesting research is usually that which is concerned with reactions which do not give the expected product. Let us take some examples; we will choose to ask ourselves questions which, as far as the author knows, are unanswered.

Firstly, will beryllium bromide react with diethyl sulphide? Beryllium bromide is an acid; diethyl sulphide, Et_2S, has a sulphur atom possessing two lone pairs of electrons and is therefore a base. A reaction between these two compounds will therefore occur under suitable conditions. The expected product would be tetrahedral $BeBr_2.2SEt_2$ with a completed octet around beryllium:



Secondly, will titanium tetrachloride react with methyldiphenylphosphine, $MePPh_2$ ($Ph = C_6H_5$)? Titanium tetrachloride, like silicon and tin tetrachlorides, is a strong Lewis acid (titanium has vacant *d*-orbitals); the phosphorus atom in the phosphine has one lone pair of electrons and so it is a base. Reaction will therefore occur; the most likely reaction ratio will be 1:2 to give an octahedral complex:



These, then, are examples of the predictions we can make, but they cannot replace actually doing the experiment and they may be wrong. Much more sophisticated theoretical understanding can of course be applied to these problems but for the sixth-form student the Lewis acid–base theory provides an initial guide as to whether or not two chemicals will react together.

The main disadvantages of the Lewis theory are that it is only semi-quantitative and that the strengths of Lewis acids and bases are not constant but vary according to the nature of the compound with which they are reacted. Suffice it to say here that these variations can often be explained using more advanced theories of chemical bonding, for example ligand-field theory.

Conclusions

We have seen how the theories of acids and bases have evolved and how each new definition has allowed the theory to be more generally applicable. It is very important to realize that the theories we have been discussing cannot be classified under headings 'right' or 'wrong'. Each theory serves some particular situation best. A chemist studying quantitative properties of acids and bases in aqueous solution will, by and large, use the Brønsted–Lowry concept, while a chemist using non-aqueous solvents will find the solvent system definition most useful. When reactions are studied in solutions of inert solvents, or in the absence of a solvent, or in the gas phase, then the Lewis theory is the more appropriate. The Lewis theory has been emphasized in this article because this is the one least known by sixth-formers and yet perhaps the most profitable one in terms of rationalizing an ever-increasing number of inorganic facts.

Bibliography

R. P. BELL, 'The use of the terms acid and base', *Quarterly Reviews*, vol. 1 (1947), p. 113.

W. F. LUDER and S. ZUFFANTI, *The Electronic Theory of Acids and Bases*, 2nd revised edn, Dover, 1961.

C. A. VANDERWERF, *Acids, Bases and the Chemistry of the Covalent Bond*, Chapman & Hall, 1965.

11 A. K. Holliday

Non-Aqueous Solvents

A. K. Holliday, 'Non-aqueous solvents', *Chemistry Student*, vol. 2 (1968), no. 2, pp. 50–54.

The mere use of the phrase 'non-aqueous solvents' is an indication of the dominant position occupied by water as a solvent. In inorganic chemistry particularly, water is almost invariably the first choice for a solvent, simply because of its remarkable ability to dissolve electrolytes, that is, substances which are ionic or yield ions in solution. However, much of our organic chemistry would be very difficult or impossible without water, either as a solvent or as a reacting species. It follows that our chemistry as a whole is very much 'water chemistry' – that is, both solvent water and the constituent atoms of water (appearing as species such as H_3O^+, OH^-) occupy a particular and important place in our chemical thinking. It is interesting to speculate about what sort of chemistry would predominate if our environment differed in some fundamental way. Thus it is not too fanciful to visualize another planet, colder than earth, where ammonia (liquid over the range $-78°$ to $-33°C$) predominated – where, in fact, there were seas of ammonia. Some idea of the chemistry which might be important there can be gathered from a book called *The Nitrogen System of Compounds*[1] written about thirty years ago by the American chemist Franklin, a pioneer worker with liquid ammonia as a solvent. Franklin attempted, with some success, to correlate the chemistry of nitrogen-containing compounds with the solvent ammonia in much the same way as we correlate oxygen-containing compounds with water; thus amides (for example sodium amide, $NaNH_2$) were comparable to hydroxides, primary amines (for example ethylamine, $C_2H_5NH_2$) to alcohols (for example ethyl alcohol, C_2H_5OH) and ammonium salts in

---

[1] E. C. Franklin, *The Nitrogen System of Compounds*, Reinhold, 1935.

liquid ammonia to our common acids in water. Further systems based on other solvents can readily be visualized; the real difficulty is that most of our 'water chemistry' is ready-made in the form of naturally occurring substances, whereas ammonia and other 'non-aqueous chemistry' must be built up, often with difficulty, by syntheses starting from simple compounds.

Coming back to earth, why do we need solvents? The simplest answer is that they are needed to dissolve substances and so allow them to react at reasonably low temperatures at a reasonably rapid rate. If we can find no solvent for a solid substance, our only hope is to heat it so that it melts (we hope without decomposition) and then allow it to react with some other solid, itself also molten or soluble in the original melt. In fact, the use of melts as solvents is a rapidly developing field of non-aqueous solvent chemistry, for reasons which appear later. But for the moment, let us look at the general question of solubility – *why* is water so versatile? The answer seems to be that water possesses three important properties:

1. a high dielectric constant,
2. a molecule which is a good electron-pair donor, i.e. a good Lewis base,
3. a molecule with strong hydrogen bonding power towards substances containing N—H, O—H and F—H bonds.

No other solvent seems to possess all three of these properties to the extent that water does; many can be found with two, for example dimethyl sulphoxide with (1) and (2), hydrogen fluoride with (1) and (3) and ammonia with (2) and (3). Other solvents, for example hydrocarbons, have none, and here the solvent power is virtually defined by the phrase 'like dissolves like'; thus the solvent hexane will dissolve other hydrocarbons or hydrocarbon-like solutes, but not much else. Obviously we must generally seek a solvent appropriate to the nature of the solute as far as we can; even so, the actual extent of solubility of a solute in a given solvent is not yet predictable theoretically except in a few favourable cases. Consequently, trial and error methods must still be of importance in finding a solvent, and complicating factors such as *solvolysis* (the more familiar term is hydrolysis, that is solvolysis involving water), where solvent molecules are split by the solute, may restrict our choice still further.

If we do find a suitable solvent, we have still to find out if it is better than, say, water for the type of reaction we wish to carry out. We may conveniently consider reactions under the following headings: metathetic and substitution reactions, acid–base reactions, complex-forming reactions, and oxidation–reduction reactions.

Metathetic and Substitution Reactions

The simplest metathetic reaction can be represented by

$$AB + CD \longrightarrow AD + CB.$$

Usually, we rely upon either AD or CB being insoluble or sparingly soluble; thus in water the familiar reaction

$$AgNO_3 + NaCl \longrightarrow AgCl(s) + NaNO_3$$

fulfils these conditions and silver chloride is readily obtained. In liquid ammonia, for example, rather different solubility relationships prevail, and a feasible reaction is

$$Ca(NO_3)_2 + 2NaCl \longrightarrow CaCl_2(s) + 2NaNO_3$$

with calcium chloride precipitating as an ammoniate (cf. a hydrate). Similarly in liquid sulphur dioxide, the reaction

$$2KBr + SOCl_2 \longrightarrow 2KCl + SOBr_2$$

proceeds quantitatively. These reactions, though interesting, are not particularly useful; there are, however, some useful synthetic reactions of the same type, which cannot be carried out in water because of hydrolysis. One example is

$$(CH_3)_3SnNa + C_2H_5Cl \longrightarrow (CH_3)_3SnC_2H_5 + NaCl.$$

The first compound, sodium trimethylstannate(IV), is hydrolysed by water (it has to be prepared in a non-aqueous solvent, see below); ethyl chloride is virtually insoluble in water. But both reactants dissolve in liquid ammonia, and one product, sodium chloride, precipitates; there is no solvolysis (ammonolysis) and the other product, trimethylethylstannane, is easily recovered from the solvent.

The problem of solvolysis is often a serious difficulty when a substitution reaction involving a substance such as a covalent chloride is required. Suppose, for example, we wish to prepare substitution products $SiCl_3X$, $SiCl_2X_2$, . . ., SiX_4 from silicon tetrachloride, $SiCl_4$. Since this molecule is an acceptor of an electron-pair, that is a Lewis acid, we may expect any donor (Lewis base) solvent to form an addition compound with it. If, however, the solvent molecule is at all reactive, this addition is followed by solvolysis:

$$\begin{aligned}&SiCl_4 + \text{water} \rightarrow [(H_2O)_2SiCl_4] \rightarrow \\ &[SiCl_3(OH)(H_2O)] + H_3O^+ + Cl^- \rightarrow \\ &\ldots \rightarrow Si(OH)_4 \rightarrow SiO_2 + 2H_2O\end{aligned}$$

and we obtain, ultimately, hydrated silica. In liquid ammonia an exactly similar reaction scheme gives us finally $Si(NH_2)_4$ which loses ammonia to give silicon nitride, Si_3N_4. Solvolysis reactions of this kind have an intrinsic interest, but do not help in obtaining the desired substitution product. Sometimes it is possible to use a less reactive donor solvent – we may use ether and the Grignard reagent methyl magnesium iodide to obtain (say) the compound CH_3SiCl_3, or ether and lithium aluminium hydride to obtain the completely substituted hydride SiH_4. A relatively recent method has been to use fused alkali chloride as solvent, for example

$$SiCl_4 + \underset{\substack{\text{potassium}\\\text{cyanate}}}{KCNO} \xrightarrow[\text{melt}]{LiCl/NaCl} SiCl_3CNO + KCl$$

$$(CH_3)_3SiCl + \underset{\substack{\text{sodium}\\\text{azide}}}{NaN_3} \xrightarrow[\text{melt}]{ZnCl_2/KCl} \underset{\substack{\text{trimethylsilyl}\\\text{azide}}}{(CH_3)_3SiN_3} + NaCl.$$

In melts the tendency is for larger groups or ions (for example $SiCl_3$, N_3^-) to associate with similarly large species and the smaller groups or ions (for example Na^+, Cl^-) to associate with similarly small species.

Acid–Base Reactions

In water, the reaction

$$\mathrm{NaOH + HCl \longrightarrow NaCl + H_2O}$$

is more correctly described as the interaction of the basic species OH^- with the acidic species H_3O^+ to give two molecules of water (solvent); we refer to the reaction as an acid–base neutralization. On the assumption that slight dissociation of a solvent into cations and anions occurs, a large number of reactions in non-aqueous solvents have been described as acid–base reactions, and some of these are shown in Table 1.

Table 1

Acid–Base Reactions in Non-Aqueous Solvents

Solvent	*Postulated ionization*	*Example of acid–base reaction*
NH_3 ammonia	$2NH_3 \rightleftharpoons NH_4^+ + NH_2^-$	$NH_4Cl + KNH_2 \rightarrow KCl(s) + 2NH_3$
SO_2 sulphur dioxide	$2SO_2 \rightleftharpoons SO^{2+} + SO_3^{2-}$	$Na_2SO_3 + SOCl_2 \rightarrow 2NaCl(s) + 2SO_2$
N_2O_4 dinitrogen tetroxide	$N_2O_4 \rightleftharpoons NO^+ + NO_3^-$	$NOCl + NaNO_3 \rightarrow NaCl(s) + N_2O_4$
BrF_3 bromine trifluoride	$2BrF_3 \rightleftharpoons BrF_2^+ + BrF_4^-$	$Ag[BrF_4] + BrF_2[AuF_4] \rightarrow Ag[AuF_4](s) + 2BrF_3$

Now the reactions quoted in the last column of the table certainly occur. It is, however, very doubtful if they can all be properly described in terms of acid–base reactions, that is involving solvent cations and anions, because there is very little evidence in some cases for any real self-ionization of the solvent. It will be noted that precipitation of one reaction product occurs in all the reactions, and this, rather than association of solvent ions, may provide the driving force for the reaction. Nevertheless, there is further evidence for this acid–base hypothesis for some solvents (for example, some metals dissolve in dinitrogen tetroxide with evolution of nitrogen monoxide, exactly as some metals in water evolve hydrogen); certainly, the idea has stimulated investigations in a wide variety of solvents and, equally certainly, many of the reactions could not be carried out in water!

A more important concept in acid–base reactions is that of the strengths of protonic acids and bases. We say that hydrogen chloride in water is a strong acid because the proton is completely transferred to the solvent:

$$\text{H—Cl} + H_2O \longrightarrow H_3O^+ + Cl^-.$$

If the transfer is incomplete (for example with acetic acid) we say the acid is weak. If we add hydrogen chloride to a mixture of water and ammonia, we know that the proton will be taken up by the ammonia (to give ammonium chloride) and not by the water – because ammonia is a stronger base than water. If, therefore, we dissolve acetic acid in liquid ammonia, we are not surprised to find that the proton transfer is complete:

$$CH_3COOH + NH_3 \longrightarrow NH_4^+ + CH_3COO^-$$

and acetic acid is therefore a strong acid. We can therefore *enhance the strength of a weak acid by dissolving it in a basic solvent* (and this usually implies a good donor solvent). This fact can be put to practical use – a very weak acid can be dissolved in (say) ammonia or an amine and titrated with a suitable base (see below) using either an indicator or a potentiometric method for the end point. Conversely, if we wish to titrate a very weak base, it can be dissolved in an acidic solvent, and a very suitable one is glacial acetic acid, which dissociates thus:

$$2CH_3COOH \rightleftharpoons CH_3COOH_2^+ + CH_3COO^-.$$

If then we dissolve ammonia – normally a weak base – in acetic acid, the reaction will be

$$NH_3 + CH_3COOH \longrightarrow NH_4^+ + CH_3COO^-$$

that is, the ammonia removes a proton completely from the solvent and is therefore a strong base. Perchloric acid, $HClO_4$, a strong acid in water, remains a strong acid even when dissolved in acetic acid as solvent, and it can therefore be used to titrate very ‘weak’ bases (for example aminoacids, polypeptides, aniline) in acetic acid, the end point being determined either by an indicator (crystal violet) or potentiometrically.

Mention was made above of a 'suitable base' for a liquid ammonia titration. Of course we could use a solution of potassium amide for this purpose – but could we use a stronger base? The answer is 'no' in liquid ammonia, because any stronger base will be solvolysed to give amide ion, just as any base stronger than hydroxide will be solvolysed in water; some examples are:

$$CH_3O^- + H_2O \rightarrow CH_3OH + OH^-$$

$$CH_3O^- + NH_3 \rightarrow CH_3OH + NH_2^-$$

and

$$H^- + H_2O \rightarrow H_2 + OH^-$$

$$H^- + NH_3 \rightarrow H_2 + NH_2^-.$$

Here, then, the strengths of the bases, methoxide ion and hydride ion, are 'levelled' to the strength of the solvent base, OH^- or NH_2^-. Similarly, the strength of protonic acids in water are levelled to the strength of H_3O^+. In water, perchloric and nitric acids are both indistinguishably strong, because both are completely converted to H_3O^+ ions; in acetic acid as solvent, perchloric acid is still strong, that is it is converted completely to $CH_3COOH_2^+$ ions, whereas nitric acid is weak (incompletely converted): in both cases strength is shown in terms of the solvent cation concentration.

We can sometimes avoid this levelling effect by the right choice of solvent – for example, we may use sodium methoxide in alcohol as a very strong base. An even more useful idea is to enhance the basicity of the hydroxide ion by using it as a melt (say of molten potassium hydroxide); we then find that it will neutralize oxides which we usually consider to be basic: for example, oxides of iron are neutralized to form ferrites or ferrates. Even more ingeniously, Professor Jolly of the University of California at Berkeley has recently shown that *solid* potassium hydroxide suspended in an inert solvent is a sufficiently strong base to remove protons from substances which we would not normally consider to be acids at all, for example phosphine:

$$\underset{\substack{\text{solid}\\ \text{in excess}}}{2KOH} + PH_3 \rightarrow KPH_2 + \underset{\text{solid}}{KOH.H_2O}.$$

Notice that the water formed in the reaction is not allowed to become free and so weaken the hydroxide as a base, but is taken up as hydrate water by the excess solid hydroxide. The substance

$(CH_3)_3SnNa$, already mentioned, can be made from trimethylstannane, $(CH_3)_3SnH$, by this method.

Complex-Forming Reactions

If we dissolve copper in aqueous nitric acid, we obtain a solution of copper(II) nitrate which on careful evaporation yields a solid hydrate $Cu(NO_3)_2.xH_2O$. We cannot remove the hydrate water from this salt and obtain the anhydrous nitrate, because the water molecules, being good donors, are too firmly coordinated round the copper(II) ion to be replaceable by a weaker donor such as the nitrate ion. (We can, of course, replace the water molecules by a *stronger* donor like ammonia, to give tetra-amminecopper(II) cations.) To obtain the anhydrous copper(II) nitrate, we can dissolve the copper in the solvent dinitrogen tetroxide (with added ethyl acetate to enhance its dissociation into the ions NO^+ and NO_3^-). The essential reaction is:

$$Cu + 2N_2O_4 \longrightarrow Cu(NO_3)_2 + 2NO,$$

and copper(II) nitrate is finally obtained as a pale blue, volatile and largely covalent solid. If zinc nitrate is dissolved in dinitrogen tetroxide, salts containing the anion $[Zn(NO_3)_4]^{2-}$ are obtained (compare this with the dissolving of zinc hydroxide in excess alkali to give $[Zn(OH)_4]^{2-}$ ions). These examples show how complexes which are not directly accessible in water as solvent are formed by choosing a solvent with the appropriate solvent anion. In a similar way, it is possible to prepare complexes with chloride as the coordinating species (chloride is a weaker donor than the water molecule) by using chloride solvents; thus complexes of the types $[BCl_4]^-$, $[BF_3Cl]^-$ may be obtained, in liquid hydrogen chloride as solvent, from the compounds BCl_3 and BF_3; and the chlorocadmate complexes $[CdCl_3]^-$, $[CdCl_4]^{2-}$ and $[CdCl_6]^{4-}$ by dissolving cadmium chloride, $CdCl_2$, in molten potassium chloride.

Oxidation–Reduction Reactions

In water, any oxidizing agent with a redox potential greater than about 1·23 volts (acid solution) will theoretically oxidize water to oxygen, and any reducing agent with a redox potential below about −0·4 volt (neutral solution) will reduce water to hydrogen. These theoretical limits are, fortunately, not obeyed in practice for many oxidation–reduction systems; thus acidic potassium permanganate should oxidize water to oxygen, but this is so slow in practice that permanganate solutions are effectively stable. Again, many electropositive metals do reduce water to hydrogen but can reduce other solutes simultaneously (for example, reductions by zinc in acid solution). But there are limitations: we cannot safely exploit the powerful reducing action of alkali metals in water, and we cannot obtain powerful oxidants, like fluorine, directly from aqueous solution. There is the further difficulty that under alkaline conditions most metals in aqueous solution are precipitated as hydroxides and this again introduces practical limitations. Perhaps one of the most useful of all the properties of liquid ammonia (and of related solvents) is its power to dissolve alkali metals to give deep blue solutions, which behave essentially as alkali metal cations and electrons associated with the solvent, that is

$$\underset{\text{metal}}{\mathrm{Na}} \xrightarrow{\text{ammonia}} \underbrace{\mathrm{Na^+} + \mathrm{e^-(NH_3)}_x}_{\text{blue solution}}.$$

Although the solvated electrons react very slowly with the solvent to give hydrogen, the solutions are sufficiently stable to be used as reducing agents of great power and versatility. Two examples must suffice to illustrate their use:

$$(1) \qquad \mathrm{C_2H_5Br} + 2\mathrm{e^-} \xrightarrow{\mathrm{NH_3}} \mathrm{C_2H_6} + \mathrm{Br^-} + \mathrm{NH_2^-}$$

$$(2) \qquad [\mathrm{Ni(CN)_4}]^{2-} + 2\mathrm{e^-} \xrightarrow{\mathrm{NH_3}} [\mathrm{Ni(CN)_4}]^{4-}.$$

Reaction (1) is a very simple example of the reduction of an organic compound; many other selective reductions are known which are not easily achieved with other reducing agents. Reaction (2) is also one of many similar reactions; here, the nickel is reduced from oxidation state +2 to zero, and although this zero oxidation state is known in, for example, nickel tetracarbonyl, $\mathrm{Ni(CO)_4}$, it is an unusual one in solution, and cannot be attained directly in aqueous solution.

Another generally useful type of reaction of alkali metals in ammonia is illustrated by the following:

$$(CH_3)_3SnCl+2Na \xrightarrow{NH_3} (CH_3)_3SnNa+NaCl.$$

This reaction gives us a second method of preparing the compound $(CH_3)_3SnNa$, already mentioned; it may be regarded as, firstly, the 2-electron reduction of the unipositive $(CH_3)_3Sn^+$ to the negative $(CH_3)_3Sn^-$, and secondly as a metathetic reaction in which one sodium ion pairs with the newly formed $(CH_3)_3Sn^-$ and the other sodium ion precipitates the chloride ion as sodium chloride.

We have been thinking about sodium metal parting with electrons in ammonia and so becoming a useful reducing agent. The reverse process – reduction of sodium ion to sodium metal – is one of the most familiar processes in a salt melt, for the electrolysis of fused sodium chloride has been the industrial method for making sodium for a very long time. Here, the sodium ion is reduced at the cathode; equally well, oxidation at the anode (giving chlorine here) can be used for oxidations which are not possible in water, so that, for example, electrolysis of a fused alkali fluoride is used to obtain fluorine. It is a sobering thought for the laboratory chemist about to embark upon a study of fused salts as solvents that they have been familiar materials in chemical industry for many decades.

Strongly oxidizing conditions can be achieved in fused salts in other ways. A fused alkali metal nitrate is itself an oxidizing agent, and ions in it are raised to higher oxidation states; thus uranium(IV) (as U^{4+}) is oxidized to uranium(VI) (as UO_2^{2+}) and many other metals are readily converted to their higher oxides. Mention may be made of a reaction sometimes used in qualitative analysis, whereby any manganese compound when fused with a nitrate under alkaline conditions is oxidized to a manganate, that is to manganese(VI). Since dinitrogen tetroxide is in effect a 'liquid nitrate', it is not surprising to find that it also is an oxidizing solvent, and metals dissolved in it usually form nitrates or similar compounds in which they show high oxidation states. Glacial acetic acid is also a useful solvent for oxidations, since it is not itself readily oxidized, and lead(IV) (as lead tetra-acetate), for example, is a stable oxidation state in acetic acid. Finally, it is appropriate to mention a rather bizarre reaction which is carried out in a melt of sodium and aluminium chloride:

$$\underset{\text{metal}}{22\text{Bi}} + \underset{\substack{\text{(as bismuth}\\ \text{trichloride)}}}{2\text{Bi}^{3+}} \longrightarrow 3\text{Bi}_8^{2+}.$$

What is the future for non-aqueous solvents? Much of our knowledge of the chemistry of solvents such as ammonia and sulphur dioxide has been acquired empirically – by a policy of 'try it and see what happens'. On the other hand, our present knowledge of the chemistry of fused salts is much more physico-chemical and fundamental; we know more about the structure of a molten salt than we do about the reactions we might carry out in it. As methods of studying the physical and structural properties of solvents improve, so we are improving our fundamental knowledge of the 'empirical' solvents. There is therefore an increasing correlation between the reactions we observe and the molecular properties of the solvents, and this should be of great help in the future; we should be able to select suitable solvents for specific reactions more readily.

Bibliography

C. C. ADDISON, *Use of Non-Aqueous Solvents in Inorganic Chemistry*, Royal Institute of Chemistry, 1960.

A. K. HOLLIDAY and A. G. MASSEY, *Inorganic Chemistry in Non-Aqueous Solvents*, Pergamon, 1965.

H. H. SISLER, *Chemistry in Non-Aqueous Solvents*, Chapman & Hall, 1965.

Part Five The Chemical Elements

The articles in this section have been chosen to illustrate the way in which the purely descriptive has given way to a more systematic approach to inorganic chemistry. W. E. Addison presents a general theoretical discussion of allotropy in energetic terms and discusses the allotropy of the more common elements. G. G. Schlessinger reviews coordination chemistry, including Werner's theory, Pauling's valence-bond theory and ligand field theory. The nomenclature of coordination compounds and the stereochemistry of complexes are also discussed. G. R. Choppin describes the history and present state of our knowledge about nuclear fission.

12 W. E. Addison

The Physical Basis of Allotropy

W. E. Addison, 'The physical basis of allotropy', *Education in Chemistry*, vol. 1 (1964), no. 3, pp. 144–50.

An element is described as allotropic when it can be obtained in more than one form in the solid state. A compound which exhibits this property is said to be polymorphic, rather than allotropic. In principle, allotropy and polymorphism are the same phenomenon, and some authors believe that no such differentiation in terminology should be adopted. In this article, however, the term allotropy is used in the more common sense – to describe the behaviour of elements only.

Almost one half of the known elements are allotropic, the number of allotropes per element varying from two to six or more. Clearly it is important to study a phenomenon of such widespread occurrence. In the present account a general theoretical discussion is presented, followed by a brief treatment of some of the more common elements.

Energy Considerations

When atoms approach one another from a state of infinite separation and form chemical bonds such that the interatomic separation is about 10^{-8} cm, energy is given out. Since each atom possesses a positively charged nucleus and negatively charged electrons, atoms experience both attractive and repulsive forces as they approach one another. At the equilibrium position at which they come to rest, these forces balance and the system achieves the state of lowest energy that is possible under the prevailing conditions (see Figure 1). Energy must be supplied to cause the atoms to move to alternative positions.

When an element can exist as two allotropic modifications, the arrangements of the constituent atoms are different in the two

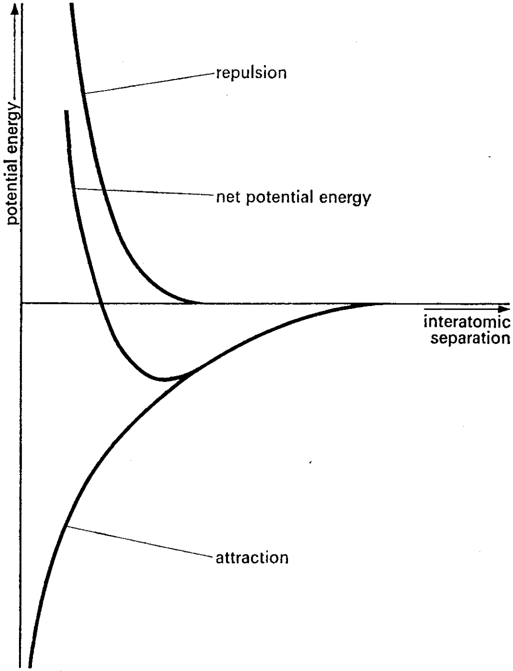


Figure 1

allotropes. Some at least of the interatomic distances are different in the two forms, and the different atomic arrangements correspond to different energies. The difference in energy between the two forms is given by

$$\Delta G = \Delta H - T\Delta S,$$

where ΔG, ΔH and ΔS are the differences in free energy, heat content and entropy of the two forms respectively, and T is the absolute temperature. Two forms may coexist in equilibrium, that is they are equally stable thermodynamically, and have the same energy, that is $\Delta G = 0$ when $\Delta H = T\Delta S$; this condition applies for only one value of T for any pair of allotropic forms. Below that temperature the one form is of lower energy, whereas above it the alternative form is of lower energy.

Suppose that element X can occur as two different forms, X_L and X_H, which are stable at low and high temperatures, respectively, that is below and above T_t, the transition temperature at which both forms are of identical energy and hence can coexist in equilibrium. If the element crystallizes below T_t then X_L is obtained, and, if above, X_H is obtained; either may be transformed to the other on passing through T_t. Such a reversible relationship between two allotropes is said to be enantiotropic, perhaps the best-known example being the transformation between rhombic and monoclinic sulphur. In some transformations, on the other hand, for example white to red phosphorus, the two forms cannot coexist in equilibrium at a given temperature, and the change is not reversed easily. Thus the action of heat on white phosphorus yields the red form, whereas conversion from red to white cannot proceed in the solid state. The relationship between such a pair of allotropes is said to be monotropic – 'one way only'. It can be imagined that the transition temperature for such a pair lies above the melting point of both forms so that the low-temperature form melts before it can change to the high-temperature form. White phosphorus is a high-temperature form of phosphorus; it is thermodynamically unstable under all conditions, and is said to be metastable.

It is obviously important to inquire why such a metastable form can exist for any appreciable time in conditions where it would be expected to alter into the thermodynamically stable form. Such metastable high-temperature forms are almost invariably prepared by chilling rapidly from a high temperature, at which they are thermodynamically stable, to room temperature or below. The atoms in the crystal of the metastable high-temperature form are unable to move into the alternative positions which they would occupy in the stable form. For the atom to move from one position to another it must be imagined that it has to pass through a position which is energetically very unfavourable, and at the low temperature it has insufficient energy to do this. The atom has to overcome an energy barrier, or the process of the transformation has a high activation energy. To understand why a particular process should have a high activation energy it is necessary to consider the structural arrangements of atoms in the two allotropic forms.

Structural Aspects

Considerations of structure facilitate an understanding of how an element can exist in more than one form, and of why some allotropic transformations are rapid and others relatively slow. It is convenient to treat metals and non-metals separately in this discussion of structure.

Three structural arrangements are commonly found in metals: two forms of close-packing (cubic and hexagonal) and the body-centred cubic form. (There are, of course, other metallic structures but they are mostly less regular than these.) At present it is difficult to indicate any correlation between electronic configuration and structure for a given metal. It is not surprising, however, that a number of metals can adopt two of these structures under different conditions.

A considerable number of metals undergo a transformation from an alternative structure to a body-centred cubic structure at an elevated temperature; among these are calcium, strontium, titanium, zirconium, manganese, lanthanum and several of the lanthanide elements. An atom in a close-packed structure has twelve nearest neighbours, whereas an atom in a body-centred cubic structure has only eight. These metals illustrate the statement that, provided the nature of the bonding is the same, a form of lower coordination has a higher energy and is hence a high-temperature form. In a structure of lower coordination there are fewer interatomic close approaches that result in the evolution of energy during lattice formation. The only element which transforms from a body-centred cubic lattice to a close-packed lattice with increase in temperature is iron; it is even more unusual in that it reverts to the same body-centred cubic lattice at a still higher temperature. This behaviour is not easy to explain, but the exception does not detract from the usefulness of the generalization.

The non-metallic elements differ from the metals in that two near-neighbour atoms are held together by an electron pair, whereas there are insufficient electrons in a metal to allow two between every pair of atoms. In a non-metal the number of nearest neighbours – sometimes referred to as the primary coordination – of any atom is determined by the electronic configuration of the element, and variation of this is possible only when there is double bonding in one allotrope. An example of this is carbon. Each carbon atom has four valence electrons and can form four bonds; in diamond each atom has four neighbours and is bonded to each of these by a single bond (sigma bond); in graphite, however, each atom has three neighbours and three of its electrons form three sigma bonds to these. The fourth electron produces partial double-bond character between each pair of atoms (pi bonding).

Since the two allotropes differ in the nature of their bonding, it is not necessary in this case for the form of lower coordination to be the high-temperature form. Indeed graphite is the low-temperature form, and diamond is metastable under normal conditions.

Although multiple bonding in one form is necessary to allow a covalently bonded element to exist in two structural forms which differ in primary coordination, an element can possess two or more structures in which every bond is a single bond. This indicates variation in secondary rather than primary coordination. Such variation requires that around any given atom there should be different distributions of the atoms other than its nearest neighbours. This can be simply illustrated for an element which has a primary coordination of two, such as a Group VI element. From the somewhat hypothetical two-dimensional structures in Figure 2 it can be seen that secondary coordination can be varied in different ways while the primary coordination remains unchanged. Schemes (a) and (b) represent a polymer–monomer relationship; (b) and (c) show different modes of packing of the same monomer, whereas (d) is a different size of monomer molecule.

The second topic on which structural considerations prove informative is the rate of allotropic transformations. During any formation some movements of atoms must occur. In general, it is true to say that the more profound the structural change, the greater is the chance that the transformation in the solid state will be slow. Transformations from one of the common metallic structures to another normally take place rapidly, since these structures are based on layers of atoms stacked on top of each other, and the transformations require only slight movement of

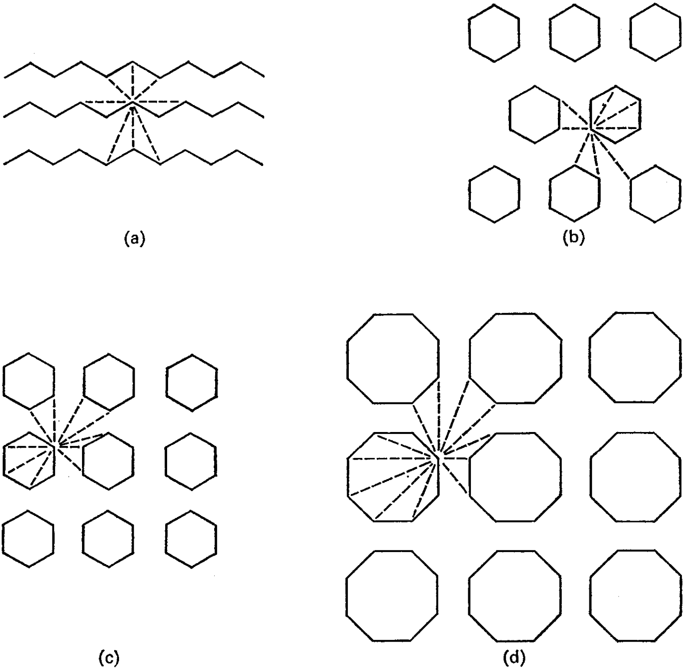


Figure 2

these layers with respect to each other. Transformations of the more complex metal structures are often more sluggish, and they cannot be visualized in such a simple manner.

In the non-metallic elements the presence of localized pairs of electrons between atoms, in contrast to the delocalization in metals, makes movement of bonded atoms with respect to each other a more difficult process. The most difficult transformations, that is those which take place most slowly, are those which involve a change in primary coordination, since this requires the greatest reconstruction. The easiest of the transformations that require a change in secondary coordination is the type represented by (b) and (c) in Figure 2, since such a change does not require the breaking of any electron-pair bonds: bond rupture is necessary in the changes (a)–(b) and (b)–(d).

In addition to structural considerations, it is clear that transformations are likely to occur more rapidly at high temperatures than at low ones, since at the higher temperature the atoms have greater thermal energy.

Allotropic Elements

It is impossible here to deal with the allotropy of the forty or so elements which can exist in more than one structural form, but some of the more common elements may be mentioned to help illustrate the points made above.

Lithium and sodium possess body-centred cubic structures at room temperature, and until recently were thought not to be allotropic. Barrett, however, argued from the premise that low coordination was associated with high temperature and hence that a close-packed form of these elements could be obtained only at low temperatures. He was able to obtain a second form of lithium, with a close-packed lattice and the higher coordination of twelve, by cooling the metal to 77 K. A similar second form of sodium was obtained at a still lower temperature, but attempts to obtain close-packed forms of the other alkali metals, which are also body-centred cubic under normal conditions, were unsuccessful.

Cerium shows unusual features at low temperatures. At room temperature the stable form has a close-packed cubic structure, and this changes to a body-centred cubic structure at 730°C – an example of the not uncommon change from high to lower coordination at high temperatures. Two distinct allotropes can be obtained on cooling the normal form (which is usually designated gamma – by the convention which labels the lowest temperature form alpha). A beta form is obtained at a temperature just below 0°C. This, like the gamma form, has a close-packed structure but with a different stacking sequence known as double hexagonal. The alpha form is obtained by cooling the gamma form to liquid-air temperature and is iso-structural with the gamma form. It is believed to be a distinct allotrope, since the interatomic distances are appreciably shorter than in the gamma form. The electrical behaviour of the alpha form is close to that expected with a tetravalent metal, whereas the gamma form behaves like a trivalent metal in this respect. The change from alpha to gamma cerium can be observed if the temperature is raised sufficiently rapidly to short circuit the formation of the beta form; the reverse change also occurs on rapid lowering of the temperature. It proceeds more rapidly than a transformation involving either the alpha or the gamma form and the beta form, since these involve structural rearrangement, whereas the alpha-to-gamma change involves only a dilatation.

Carbon has long been known to be allotropic. It is notable for the remarkable contrast in physical properties between diamond, which has a continuous tetrahedral network of single bonds, and graphite, which comprises planar trigonal sheets in which each bond has double-bond character. The transformation from one form to the other involves changes in both primary coordination and bond type, and the activation energy is large. The low-temperature form is graphite, but the transformation of diamond to graphite is, perhaps fortunately, infinitely slow; it can be accomplished by heating at 1500–2000°C. The existence of natural diamonds has aroused much speculation as to their mode of formation, and considerable effort has been exerted in attempts to synthesize diamonds from graphite. The difficulty of attaining suitable conditions for the formation of diamond can be gauged from the equation for the transition curve

$$P = 7100+27T$$

which defines the conditions under which both diamond and graphite can co-exist in equilibrium. In this equation, P is the pressure in atmospheres and T the temperature in degrees Kelvin.

It is of interest to speculate whether diamond would yet have been discovered were it not for the fortunate circumstance of its natural occurrence. In view of the extreme conditions necessary for its formation, and also for the formation of the recently discovered low-temperature forms of lithium and sodium, it seems reasonable to assume that still further allotropes remain to be discovered.

A less well-known feature of the allotropy of carbon is that there are two forms of graphite, each of which can be regarded as a distinct allotrope. The more recently discovered beta form differs from the more familiar alpha form only in the sequence in which the sheets of atoms are stacked – the repeat unit is two layers in the alpha form, whereas it is three in the beta form (cf. hexagonal and cubic close-packing in metals). The alpha form is the more stable thermodynamically, but it can be converted to the less stable form by cold working, that is the application of mechanical energy at room temperature. This input of energy produces the higher-energy form at a temperature high enough to prevent reversion to the more stable form. Such a transformation is effected by heating beta graphite to about 1000°C. An interesting way in which the transformation can be accomplished at lower temperatures is for an intercalation compound of beta graphite to be prepared and subsequently decomposed, when alpha graphite results. The intercalated molecule causes the sequence of stacking to be altered, and when the layers must change position again on its removal, they take up the thermodynamically favoured sequence.

Tin exists in two allotropic forms which differ appreciably in both physical properties and structure. The low-temperature form, grey tin, is isostructural with diamond, and non-metallic in properties, whereas the high-temperature white form is metallic in properties and has a unique structure with an irregular coordination in contrast to the regular tetrahedral coordination of the grey form. These two forms exhibit an enantiotropic change at 13·2°C but, on account of the alteration both in coordination and in bond type, it is a sluggish change, so that it is easy to maintain either form in a metastable condition outside its normal stability range. The significance of the transition and its sluggishness are too well known to be elaborated upon here.

Phosphorus is normally prepared by the carbothermic reduction of its oxide at a temperature above the boiling point of the element, and rapid cooling of the product. This process yields white phosphorus, which is the high-temperature form of the element and is metastable at normal temperatures. If phosphorus vapour at a still higher temperature is chilled to liquid-air temperature, a species characteristic of that higher temperature can be isolated in the solid state; this brown phosphorus is also metastable.

White phosphorus comprises tetrahedral molecules in which each atom, as required by the electronic configuration of the element, is bonded to three nearest neighbours. This molecule possesses a bond angle of 60°, and such a small angle is very uncommon. This can be said to cause a strain in the molecule, and this strain corresponds to an excess of energy compared with the thermodynamically stable form, such as is characteristic of all high-temperature allotropic forms. When white phosphorus is cooled it undergoes an enantiotropic transformation to a further allotropic form at −77°C. The two forms of white phosphorus probably differ only in the way in which the P_4 tetrahedra are packed in the crystal. The transformation requires only minor movement of atoms, has a low activation energy and takes place rapidly. It is unusual in being an enantiotropic change between two metastable forms.

Red phosphorus is the product obtained by a monotropic transformation when white phosphorus is heated to a temperature just below its boiling point; the process is sluggish, even in the presence of a catalyst such as iodine. Red phosphorus does not exhibit a completely ordered structure; the detail of the structure is not known, although it is based on the characteristic three-coordination. The physical properties of red phosphorus are those of a material of high molecular weight, so that the tetrahedral P_4 molecules have opened up and interlinked during its formation.

The structure of black phosphorus, which is believed to be the form with greatest thermodynamic stability, is known to consist of buckled sheets in which the atoms are arranged in a trigonal pattern so that each atom preserves three-coordination. The material has features in common with graphite, in which the sheets of atoms are planar. This allotrope was first prepared by the action of heat on white phosphorus under high pressure, and has subsequently been prepared, without pressure, by heating with liquid mercury at 370°. The properties of the two preparations are not identical, perhaps due to the presence of traces of mercury in the latter preparation.

A consideration of the structure of monomeric white and polymeric black phosphorus makes it understandable that the nature of red phosphorus should be an intermediate in which structural order is incomplete. It can be appreciated that the linking of opened-up tetrahedra to form sheets must have a considerable activation energy; the process can therefore proceed more rapidly in the presence of a catalyst, which affords some facile reaction path, or when energy is supplied, to yield the black form. If the right conditions are not provided, the red form with incomplete order in its structure is obtained.

When black phosphorus is heated, some disordering takes place and the less stable red form results. If it is vaporized, P_4 (and above 1000°C, P_2) molecules are formed and the chilling of the vapour yields these molecules in the solid state as white and brown phosphorus, respectively, as metastable species.

Arsenic, like phosphorus, can be obtained as a monomer, As_4, as a polymer, when it behaves as a metalloid, or in an intermediate state.

Sulphur, with its enantiotropic change between rhombic (alpha) and monoclinic (beta), has provided the most extensively studied allotropic transformation. It takes place relatively rapidly and is believed to involve only a change in the relative positions of the S_8 rings with respect to one another. In this it resembles the change between the two forms of white phosphorus, involving change in secondary coordination only.

Four other allotropes of sulphur, which seem to be well characterized, should also be mentioned; they are all metastable in the solid state and yield rhombic sulphur by monotropic changes. Nacreous (gamma) sulphur, obtained for example by the slow oxidation of an alcoholic solution of ammonium polysulphide, is very similar in its properties to rhombic sulphur, to which it is readily transformed. Nacreous sulphur is built up of S_8 rings, but with a different arrangement from that in either rhombic or monoclinic sulphur. It would be very interesting to know what conditions in the preparation influence the formation of this particular metastable allotrope.

Rhombohedral, rho or Engel's sulphur, like the alpha, beta and gamma forms, has the properties typical of a small covalent molecule. Whereas solutions of each of these three deposit alpha sulphur on crystallization, a solution of rho sulphur deposits only itself. This form is obtained by extracting with chloroform a solution of thiosulphate which has been acidified under carefully controlled conditions. Although this allotrope is metastable with respect to alpha sulphur, the transformation is slower than those of beta or gamma sulphur.

X-ray studies have shown that the molecule of rho sulphur is a six- rather than an eight-membered puckered ring, as found in the other monomeric species. It is for this reason that the transformation S_ρ–S_α is sluggish, since it must involve ring opening and subsequent ring closure; the opened-up six-membered rings must form an intermediate chain structure from which eight-membered units break off before cyclization, and evidence of this is afforded by the formation of small amounts of polymeric material as the change proceeds.

It is relevant to inquire why the six-membered ring should be less stable than the eight-membered ring, and why it is formed in this particular reaction. In the smaller ring the non-bonding pairs of electrons are forced more closely together, and this repulsion causes the molecule to have a higher energy. The formation in this particular reaction has been attributed by Bartlett to a mechanism whereby $S_2O_3^{2-}$ ions combine and eliminate SO_3^{2-} ions, thus:

$$2S_2O_3^{2-} \xrightarrow{H^+} HS_3O_3^- + SO_3^{2-}$$
$$HS_3O_3^- + S_2O_3^{2-} \longrightarrow HS_4O_3^- + SO_3^{2-}, \text{ etc.}$$

until finally a sterically favoured cyclic transition state occurs to yield a six-membered ring, that is

$$HS_6SO_3^- \longrightarrow S_6 + SO_3^{2-} + H^+.$$

When liquid sulphur is heated above 160° it becomes very viscous; the viscous material is believed to be a solution of a long-chain polymer in a solvent comprising S_8 monomers. When this viscous liquid is quenched by pouring into cold water, plastic sulphur is obtained; this has the properties of a supercooled liquid, but, when stretched, the chain molecules become aligned and the material becomes a crystalline solid. This fibrous sulphur is a further example, like white and brown phosphorus, of a high-temperature form frozen out at room temperature in a metastable condition. The structure of fibrous sulphur is very subtle, since the spaces between the spirals of the chain polymer are the appropriate size to contain ring monomer molecules. Thus the solute and the solvent molecules of the viscous melt are cunningly interwoven. It can be appreciated that the transformation of fibrous to rhombic sulphur is slow, since a movement of atoms and the breaking and formation of bonds are involved.

A further metastable allotrope is obtained if gaseous sulphur is chilled from 1000°C to the temperature of liquid air. This is a purple solid, which is paramagnetic and is believed to consist of S_2 molecules. When it is allowed to warm up, a mixture of rhombic sulphur and polymer is obtained. The much lower stability of the diatomic molecules of sulphur and phosphorus at room temperature when compared with oxygen and nitrogen is thus emphasized: it is attributed to the weaker p_π–p_π-bonding in elements of the second short period compared with those of the first.

Bibliography

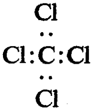
W. E. ADDISON, *The Allotropy of the Elements*, Oldbourne, 1964.

13 G. G. Schlessinger

Coordination Chemistry

G. G. Schlessinger, 'A summer short course in coordination chemistry', *Chemistry*, vol. 39 (1966), no. 6, pp. 8–13, and no. 7, pp. 13–16.

Chemical compounds can be thought of as combinations of different elements involving varying degrees of electron sharing. One extreme, the ionic bond, involves almost complete transfer of charge, as in the formation of Na^+Cl^- from Na and Cl atoms. On the other hand, covalent bonding represents a more or less equal sharing of negative charge, as in carbon tetrachloride, CCl_4:



$$\begin{array}{c} Cl \\ \cdot\cdot \\ Cl:C:Cl \\ \cdot\cdot \\ Cl \end{array}$$

Both carbon tetrachloride and sodium chloride differ radically in physical and chemical properties from their elemental constituents and both are stable chemical entities.

Complex compounds, on the other hand, may be considered as molecular or addition products resulting from the combination of compounds such as those described above to form new materials which likewise differ from the parent substances. The distinction, then, between so-called simple and complex compounds is that simple compounds use elements as building blocks, but complex compounds use preformed compounds.

The transition elements of importance to our discussion fall in Groups VIB, VIIB, VIIIB, IB and IIB of the periodic table; that is, chromium to zinc, molybdenum to cadmium, and tungsten to mercury. They owe their unusual properties to unfilled inner electron shells or, more specifically, the *d*-orbitals.

Complex compounds, of which numerous examples are known, may be divided into two broad classes – normal complexes and penetration complexes. Normal complexes are those which dissociate reversibly in solution. For example,

$$(NH_4)_2CrCl_5.H_2O \text{ or } 2NH_4Cl.CrCl_3.H_2O$$
$$(NH_4)_2CrCl_5 \rightleftharpoons 2NH_4^+ + Cr^{3+} + 5Cl^-$$

and

$$Co(NH_3)_6Cl_2 \text{ or } CoCl_2.6NH_3$$
$$Co(NH_3)_6Cl_2 \rightleftharpoons Co^{2+} + 2Cl^- + 6NH_3.$$

These are double salts because, when dissolved in an appropriate solvent, they show the combined chemical and physical characteristics of each component, ion or molecule.

Penetration complexes are those compounds which, when in solution, exist as ionic aggregates of greater stability than the double salts. For example,

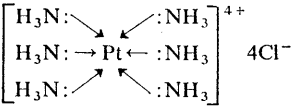
$$K_4Fe(CN)_6 \text{ or } 4KCN.Fe(CN)_2$$
$$K_4Fe(CN)_6 \longrightarrow 4K^+ + [Fe(CN)_6]^{4-}$$

and

$$Rh(NH_3)_6Cl_3 \text{ or } RhCl_3.6NH_3$$
$$Rh(NH_3)_6Cl_3 \longrightarrow [Rh(NH_3)_6]^{3+} + 3Cl^-.$$

The difference between the two classes of complexes is essentially a matter of degree. Both are formed by interaction of a metal (as its salt), called the central atom or ion, with various anions or molecules, called ligands, containing an unshared pair of electrons.

The ligands are attached to the central ion by so-called dative, coordinate covalent, or semipolar bonds using their lone electrons. For example,



The metal plus ligands that surround it comprise the coordination sphere which is shown by the square brackets.

Because of limited space, only penetration complexes or coordination compounds of the elements listed above will be considered.

History of the Theory

The historical development of experimental and theoretical progress in coordination chemistry can be divided into three major periods: the late eighteenth century to 1893; the Wernerian era, 1893 to 1940; and the modern period, 1940 to the present.

Because coordination compounds occur in nature, particularly in dyes and pigments, we can assume that they were used by the ancients, although the ancients had no knowledge of their true nature. For example, alizarin, a red dye prepared from madder root and clay, was used.

Coordination chemistry probably originated in 1704 with the discovery of the pigment Prussian blue by a Berlin color maker, Diesbach. Its preparation, which would make a present-day chemist wince, went like this: equal parts of potassium nitrate, potassium antimonyl(III) tartrate (cream of tartar), and either ox blood or animal flesh were heated together. The product was dissolved in water, treated with iron(II) (ferrous) sulfate, alum (potassium aluminium sulfate), and finally with hydrochloric acid. Today, we simply mix any iron(III) (ferric) salt with potassium hexacyanoferrate(II) to obtain the pigment:

$$Fe^{3+} + K_4[Fe(CN)_6] + xH_2O \longrightarrow KFe^{III}[Fe^{II}(CN)_6].xH_2O + 3K^+.$$

In 1753 Macquer prepared potassium hexacyanoferrate(II) by treating Prussian blue with alkali. About 1760 Lewis reported the use of potassium hexachloroplatinate(IV) (K_2PtCl_6) in refining platinum.

In 1799, Tassaert obtained orange $CoCl_3.6NH_3$ by allowing a mixture of cobalt(II) chloride and aqueous ammonia to stand in air:

$$4CoCl_2 + 24NH_3 + 2H_2O + O_2 \longrightarrow 4[Co(NH_3)_6]^{3+} + 8Cl^- + 4OH^-.$$

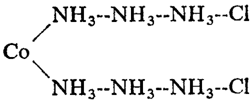
The puzzled French chemist recognized that a new chemical substance was formed, by combination of two stable or fully saturated compounds, which had very different properties from either.

During the next fifty years some activity was centered on other complex cobalt, iron, and platinum compounds, notably the preparation of potassium hexacyanoferrate(III) $[K_3Fe(CN)_6]$ in 1822, tetrammineplatinum(II) tetrachloroplatinate(II) $[Pt(NH_3)_4PtCl_4]$ by Magnus in 1828, and sodium nitrosylpentacyanoferrate(III) (sodium nitroprusside, $Na_2[Fe(CN)_5NO]$) in 1849.

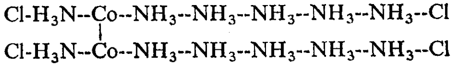
Beginning with Genth's work on the composition of the cobalt(III) ammines in 1847 we enter a fruitful period of empirical work by Fremy, Blomstrand, Gibbs, and particularly Jorgensen in Copenhagen which lasted until 1890. About 1851 Fremy devised color code names for the cobalt ammines. In 1854 Claus recognized the submergence of the normal properties of ammonia (and water) molecules in these compounds but his ideas were not accepted then because sufficient experimental evidence was lacking.

In 1858 and 1865, two events occurred which were to hold back the development of coordination chemistry for almost fifty years. The first was Kekulé's theory of the tetrahedral and chain-like nature of carbon atoms. The second was emergence of the classical hexagonal formula of benzene. The rapid acceptance and verification of these valence principles in organic chemistry caused chemists working with penetration complexes to propose all sorts of unwieldy chain structures for their compounds; this originated with Blomstrand in 1871.

From 1870 to 1890 Jorgensen was the recognized master of the field. He prepared and carefully characterized large numbers of complexes but used and extended Blomstrand's chain formulations. His painstaking work and meticulous observations laid the groundwork for Werner's theory. Typical of some of these earlier formulae were:

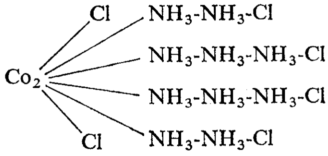


$[Co(NH_3)_6]Cl_2$



$$\begin{array}{l} Cl\text{-}H_3N\text{--}Co\text{--}NH_3\text{--}NH_3\text{--}NH_3\text{--}NH_3\text{--}NH_3\text{--}Cl \\ \qquad\qquad\;\, | \\ Cl\text{-}H_3N\text{--}Co\text{--}NH_3\text{--}NH_3\text{--}NH_3\text{--}NH_3\text{--}NH_3\text{--}Cl \end{array}$$

$[(Co[NH_3]_6)Cl_2]_2$



$[(Co[NH_3]_5Cl)Cl_2]_2$

The work of Raoult and van't Hoff on the properties of non-volatile compounds in solution about 1882 allowed the determination of molecular weights and discarded dimeric (double) formulae such as those above. Also, in 1887, when Arrhenius introduced his concept of ionization, the difference between coordinated molecules (or ions) and fully ionized groups was clarified.

Werner's theory

The stage was thus set in 1891 for Alfred Werner, then only 25 years old, to publish his first epoch-making paper on valence. In it he suggested that the central atom or ion does not necessarily have a small fixed number of valence bonds but that its valence might be exerted over the whole surface of the metal atom and have units of different affinity or strength.

Two years later Werner, then assistant professor at Zurich, had considered the problem further and was led to think about molecular compounds in general. It is said that the answer came to him in a dream. He awoke at two o'clock one morning with the entire problem solved in his mind and went into a room with pen, paper and a box of cigars. He wrote until five the following afternoon to complete his theory. As a result, he was promoted to a full professorship that year and won the Nobel prize in 1913. Over the period 1893 to 1919 a prodigious amount of work came out of his laboratory, all of which has substantiated his postulates and served as the foundation of coordination chemistry ever since.

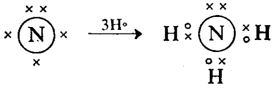
The fundamental principles of Werner's theory can be summarized as follows:

1. Metals possess two types of valence: primary or ionic and secondary or non-ionic (coordinate covalent bond type).
2. Every metal has a fixed number of secondary valences, either four or six (other numbers such as two, five, seven and eight are recognized today).
3. Primary valences are satisfied by anions, while secondary valences may be filled with anions or neutral molecules. In every case the coordination number (C.N.) of the metal must be completed.

4. The secondary valences are directed spatially around the central metal ion. A C.N. of four requires a tetrahedral or planar configuration, while a C.N. of six is directed to the apices of a regular octahedron. These geometries predicted a variety of types of isomerism in penetration complexes; this postulate was the most difficult to prove. It was not until 1914 that Werner finally achieved the resolution of a purely inorganic compound.

In the so-called modern period of coordination chemistry, perhaps the major area of concern has been the gradual elucidation of bonding in penetration complexes.

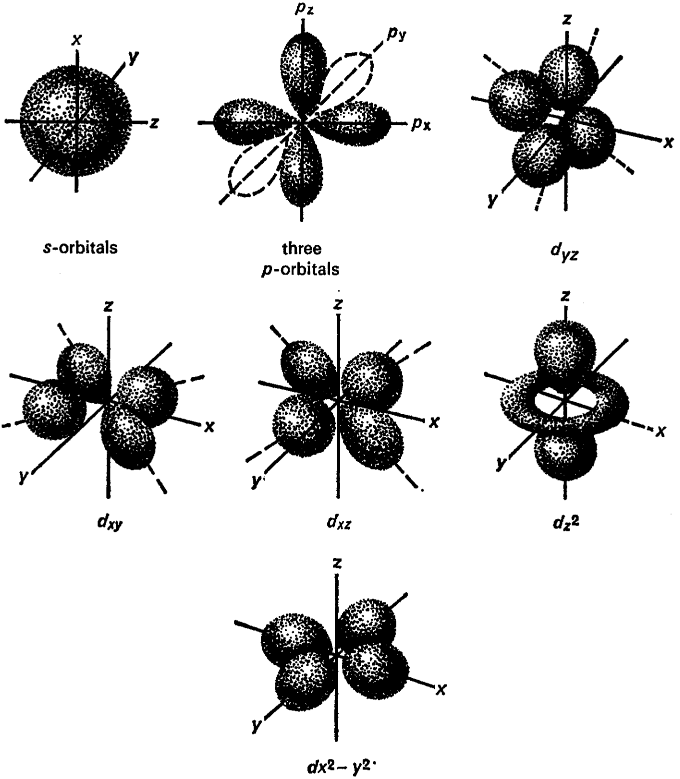
In 1916 G. N. Lewis advanced his ideas on electronic bonding which involved the now elementary octet structure as applied to the valences of all atoms in compounds. For an atom or ion to have all its valence potential satisfied, the outer shell of electrons had to contain eight electrons. For example, for NH_3, the three Hs each contribute one electron:



This principle, which worked well for most simple substances, was extended in the years 1923 to 1927 by Sidgwick and Lowry to coordination complexes. For example, the Co^{3+} ion has 24 electrons and can accept (by semi-polar bonding) 12 more from six coordinating groups, each with a pair, to achieve maximum stability – that is, the noble gas configuration of krypton with 36 electrons. This effective atomic number (E.A.N.) was arrived at by taking the atomic number of the element, subtracting the electrons lost in ion formation, and adding the electrons contributed by coordination. However, many relatively stable penetration complexes gave E.A.N.s which were not those of a noble gas. For example, in $CrCl_3.6NH_3$, Cr has 24 effective electrons; Cr^{3+} has 21; and $6NH_3$ has $6 \times 2 = 12$. This gives a total E.A.N. of 33. Furthermore, this theory was unable to account for the spatial arrangement of ligands around the central metal.

Valence-bond theory

The next major step was due to Linus Pauling who, in 1931, introduced the valence-bond theory which contained three important new ideas. The first was concerned with the geometry of complexes with different C.N.s, explained by the so-called hybridization of electronic orbitals. These orbitals are electron density patterns around the positive nucleus, and each can contain a certain number of electrons. The *s*-orbitals can contain up to two electrons with opposite spins. The *p*-orbitals have a maximum of six electrons, two in each *p*-orbital. The *d*-orbitals can contain up to 10 electrons, two in each *d*-orbital (Figure 1).



s-orbitals

three *p*-orbitals

d_{yz}

d_{xy}

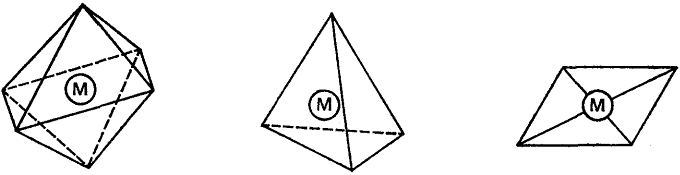
d_{xz}

d_{z^2}

$d_{x^2-y^2}$

Figure 1 *s*-, *p*- and *d*-orbitals

Thus, a combination of one *s*-orbital with three *p*-orbitals resulted in four sp^3 hybridized orbitals directed to the corners of



octahedral tetrahedral square planar

Figure 2 Shape of metal complexes

a tetrahedron. Similarly, dsp^2 hybridization gave a square planar arrangement and d^2sp^3 combination gave a regular octahedral orientation (Figure 2). All these hybrid orbitals were stronger than the simple orbitals alone and all orbitals of a hybridized set have the same energy.

Because in Table 1 the energy differences of the $3d$, $4s$ and $4p$ orbitals are relatively small, the empty s- and p-orbitals may reasonably be involved in hybridization.

The formation of octahedral $[Co(NH_3)_6]^{3+}$ was explained diagrammatically as

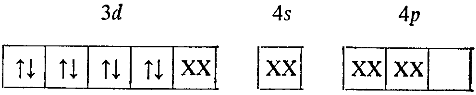
3d					4s	4p		
↑↓	↑↓	↑↓	XX	XX	XX	XX	XX	XX

six d^2sp^3 hybrid orbitals

Table 1
Electronic Distribution of some Transition Metal Ions

Orbitals		3d					4s	4p		
3+ ions	Cr	↑	↑	↑						
	Mn	↑	↑	↑	↑					
	Fe	↑	↑	↑	↑	↑				
	Co	↑↓	↑	↑	↑	↑				
2+ ions	Ni	↑↓	↑↓	↑↓	↑	↑				
	Cu	↑↓	↑↓	↑↓	↑↓	↑				
	Zn	↑↓	↑↓	↑↓	↑↓	↑↓				

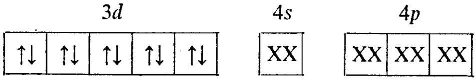
where Xs represent electrons from the nitrogen atoms of the six ammonia molecules. Square planar $[Ni(NH_3)_4]^{2+}$ may be represented diagrammatically as



four dsp^2 hybrid orbitals

Note that the two unpaired electrons of Ni^{2+} have combined in one 3*d* orbital to form the complex ion. This will be explained under magnetic properties below.

Tetrahedral $[Zn(NH_3)_4]^{2+}$ would contain



four sp^3 hybrid orbitals

The second principle of Pauling's valence-bond theory concerned the magnetic properties of many complex salts which might be used to investigate the nature of bonding between the ligand and central ion. If the coordination sphere has one or more unpaired electrons alone in an orbital, the compound will be attracted into a magnetic field and is said to be paramagnetic. With no unshared electrons, the complex will be diamagnetic or repelled by the field.

The movement of an electrical charge produces a magnetic effect, and if we think of an electron as a very small bar magnet, the magnetic moment is given by the strength of the poles multiplied by the distance between them. Polarity is set up in molecules due to variation in electron spin and cloud charge distribution of the bonding orbitals between elements.

The total magnetic moment of a substance may be regarded as the result of all the moments of the individual electrons. It can be measured with a Gouy balance which weighs the substance in a strong uniform magnetic field of known strength and compares this weight to the weight of the substance when it is free of the field.

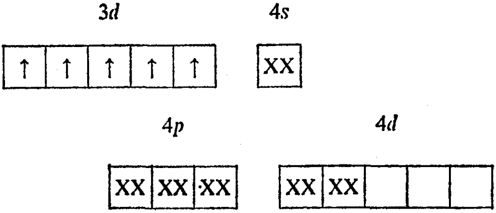
The magnetic moment thus measured is easily related to the number of unpaired electrons by the approximate relationship: number of unshared electrons = $X-(0{\cdot}7$ to $0{\cdot}9)$, where X = magnetic moment in Bohr magnetons (BM). The BM has the value of 5564 gauss $\text{cm}^{-1}\,\text{mol}^{-1}$ and is given by

$$\frac{he}{4\pi mc}$$

where h is Planck's constant, $6{\cdot}63\times10^{-34}$ joule second; e is electron charge, $1{\cdot}60\times10^{-19}$ absolute coulomb; m is mass of electron, $9{\cdot}11\times10^{-28}$ g; and c is the speed of light in a vacuum, $3\times10^{10}\,\text{cm s}^{-1}$.

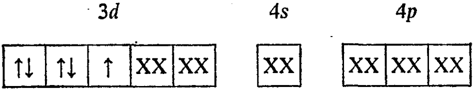
The magnetic criterion is valueless for transition metals with fewer than four or more than seven d-electrons. Thus Cr^{3+} complexes always show three free electrons and Cu^{2+} coordination compounds always show one.

The usefulness of Pauling's approach can be shown by a simple example. For $[FeF_6]^{3-}$:



six d^2sp^3 hybrid orbitals

Here there are five unpaired electrons, and we obtain a so-called outer orbital, ionic or labile complex, in which outer $4d$ orbitals are used. For $[Fe(CN)_6]^{3-}$ the following applies:



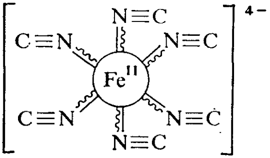
six d^2sp^3 hybrid orbitals

where there is one unpaired electron and an inner-orbital type, covalent or penetration complex, using the inner $3d$ orbitals.

Note that in the last example, as in the case of the above-mentioned ammonia ion, trivalent iron pairs its previously lone electrons in complex formation. Pauling offered no satisfactory quantitative explanation of the differing effect of various ligands on the central ion's bond nature.

The third principle of Pauling's valence-bond theory concerned a type of double bonding between metal and ligands or a reverse electron donation of a partial nature which could spread out the negative charge over the whole coordination sphere. This would result in a more stable structure.

The coordination of various groups with their extra electrons might build up the minus charge on the central metal according to Sidgwick's theory. For example in $[Fe(CN)_6]^{4-}$,



the iron(II) might have up to four excess negative charges. The metal could relieve itself of this electronic burden by partly 'feeding back' some of these charges to the N≡C groups.

The valence-bond theory, for all its utility, was unable to provide quantitative explanations of bond strength and color in complexes. It also led to erroneous conclusions about ionic and covalent structure assignments based on magnetic data.

Electrostatic field theory

From 1929 to 1935, the physicists Bethe and Van Vleck developed the crystal field theory for use in ionic crystal lattice phenomena. Finally in 1951 and 1952, it was applied to coordination compounds by the German chemist Hartmann and the Englishman Orgel as the ligand or electrostatic field theory. It consists of the following simple postulates:

1. The central metal and ligands are hard spherical ions of

opposite charge with the metal as positive. When the ligand is formally neutral, unshared electrons on the constituent atoms are regarded as local charged sites.

2. In the stable structure of the complex the particles are so arranged that oppositely charged species are in contact and like charges are separated as widely as possible.

3. The coordination sphere is energetically stabilized by the difference between all electrostatic attractions and repulsions of ions and molecules involved, and by the electrical interaction between the *d*-orbitals of the transition metal and the electronic field of the ligands.

To begin with, consider an octahedral arrangement with six ligands approaching the central ion along opposite ends of the *x*-, *y*- and *z*-axes. The ligands will be repulsed most strongly by the $d_{x^2-y^2}$ and d_{z^2} orbitals of the metal and least strongly by the d_{xy}, d_{xz} and d_{yz} orbitals because of their geometric orientation (Figure 1).

Free ion	*All orbitals of equal energy (or degenerate)*
Low energy (low repulsion) orbitals	d_{xy}, d_{xz}, d_{yz}
High energy (high repulsion) orbitals	$d_{x^2-y^2}$, d_{z^2}

Similar spatial considerations for tetrahedral and square planar complexes are illustrated in Figure 3.

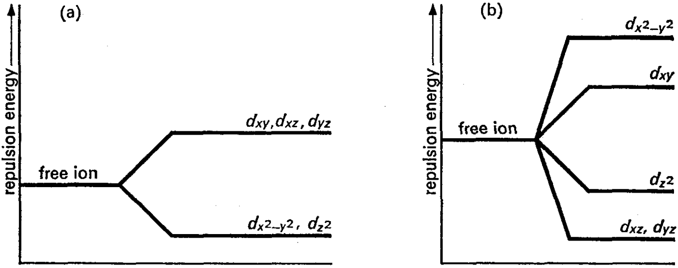


Figure 3 (a) Tetrahedral (b) square planar

Table 2

Electronic Distribution of the *d*-Orbitals for Transition Metal Complexes

Octahedral

		High spin					Low spin				
		d_{xy}	d_{xz}	d_{yz}	d_{z^2}	$d_{x^2-y^2}$	d_{xy}	d_{xz}	d_{yz}	d_{z^2}	$d_{x^2-y^2}$
d^3	Cr^{3+}	↑	↑	↑			↑	↑	↑		
d^4	Mn^{3+}	↑	↑	↑	↑		↑↓	↑	↑		
d^5	Fe^{3+}	↑	↑	↑	↑	↑	↑↓	↑↓	↑		
d^6	Co^{3+}	↑↓	↑	↑	↑	↑	↑↓	↑↓	↑↓		
d^7	Co^{2+}	↑↓	↑↓	↑	↑	↑	↑↓	↑↓	↑↓	↑	
d^8	Ni^{2+}	↑↓	↑↓	↑↓	↑	↑	↑↓	↑↓	↑↓	↑	↑
d^9	Cu^{2+}	↑↓	↑↓	↑↓	↑↓	↑	↑↓	↑↓	↑↓	↑↓	↑
d^{10}	Zn^{2+}	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓

Tetrahedral

d^3	↑			↑	↑				↑↓	↑
d^4	↑	↑		↑	↑				↑↓	↑↓
d^5	↑	↑	↑	↑	↑	↑			↑↓	↑↓
d^6	↑	↑	↑	↑↓	↑	↑	↑		↑↓	↑↓
d^7	↑	↑	↑	↑↓	↑↓	↑	↑	↑	↑↓	↑↓
d^8	↑↓	↑	↑	↑↓	↑↓	↑↓	↑	↑	↑↓	↑↓
d^9	↑↓	↑↓	↑	↑↓	↑↓	↑↓	↑↓	↑	↑↓	↑↓
d^{10}	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓

Square planar

d^7	↑	↑↓	↑↓	↑↓	
d^8	↑↓	↑↓	↑↓	↑↓	
d^9	↑↓	↑↓	↑↓	↑↓	↑

An approaching ligand with a strong tendency to donate its electrons (for example, NH_3 or CN^-) results in a strong field environment and a pairing up of the central ion's electrons (which were previously unpaired in the *d*-orbitals) due to the repulsions set up. Thus, a low spin or spin-paired penetration complex is formed of the inner-orbital or covalent type.

A weakly electron-donating ligand produces a weak field and gives a high spin or spin-free relatively unstable complex with outer-orbital or ionic features. If we use the Hund rule which states that electrons will occupy separate degenerate orbitals, then the electronic distribution of the *d*-orbitals for the transition metals can be summarized in Table 2.

Only the configurations d^4 to d^7 differ in high and low spin complexes and the electronic assignments are similar to Pauling's except with the modification of differing orbital energy levels. However, this last distinction gives the ligand field theory its most useful quantitative aspect in its ability more simply and accurately to evaluate bond strengths, kinetics and color in all types of complexes without recourse to often misleading magnetic data.

Much remains to be done before an all-encompassing coordination theory will evolve; it seems increasingly likely that no simple one will suffice. Nevertheless, the close historical parallel between complex and general chemical bonding theory is interesting. Experimentation and technology must reach a certain stage before a new idea is ripe for general acceptance.

Nomenclature

In 1852, Fremy suggested color code names for cobalt complexes. Later these were frequently used as generic-type prefixes for other metals with similar formulae but with different colors (Table 3). These names are still encountered in the older literature, but they have been replaced by a more systematic method of nomenclature.

Current Nomenclature

The current system of nomenclature, a modification of Werner's rules, involves the following steps. The cation is named first.

Table 3

Early Nomenclature

Type	*Generic Name*	*Color*
$[Co(NH_3)_6]^{3+}$	Luteo	Yellow
$[Ir(NH_3)_6]^{3+}$	Luteo	Colorless
$[Co(NH_3)_5Cl]^{2+}$	Purpureo	Purple red
$[Ir(NH_3)_5Cl]^{2+}$	Purpureo	Pale yellow
$[Co(NH_3)_5I]^{2+}$	Purpureo	Olive green
$[Co(NH_3)_5H_2O]^{3+}$	Roseo	Rose red
$[Ir(NH_3)_5H_2O]^{3+}$	Roseo	Colorless
cis-$[Co(NH_3)_4(NO_2)_2]^+$	Flavo	Brown yellow
trans-$[Co(NH_3)_4(NO_2)_2]^+$	Croceo	Crocus yellow
cis-$[Co(NH_3)_4Cl_2]^+$	Violeo	Violet
trans-$[Co(NH_3)_4Cl_2]^+$	Praeseo	Green

Then the coordination sphere is specified as follows:

Bonded groups are listed with negative groups first, then neutral groups, and positive groups last.

Negative groups end in -o; neutral groups normally have no suffix, except aquo (water); and positive groups end in -ium. Among neutral groups aquo is listed first; ammine (ammonia) last.

The number of each simple ligand is prefixed to the name; for example, di-, tri-, tetra-, penta- and hexa-. For complicated ligand names, bis-, tris-, tetrakis-, pentakis- and hexakis- are used.

The name of the metal is given, with a Roman numeral in parentheses as suffix to indicate the oxidation state. If the coordination sphere is negative, the metal name ends in -ate and the Roman numeral follows.

Some representative examples of current nomenclature are:

$$K[Co(NH_3)_2(C_2O_4)(NO_2)_2],$$

potassium dinitro-oxalatodiamminecobaltate(III);

$$[Co(NH_3)_3(H_2O)Cl_2]HSO_4,$$

dichloroaquotriamminecobalt(III) hydrogen sulfate;

$$[Co(NH_3)_6][Cr(CN)_6],$$

hexamminecobalt(III) hexacyanochromate(III);

$$[Ni(NH_3)_4(NO_2)_2],$$

dinitrotetramminenickel(II); and

$$H[Co(NH_3)_2(C_2O_4)_2],$$

hydrogen dioxalatodiamminecobaltate(III).

Typical Ligands

Table 4 lists some of the more important ligands. A ligand which

Table 4

Some Important Ligands

Position of Key Atom in Periodic Table	*Unidentate*	*Bidentate*	*Polydentate*
Group IV[a]	CO_3^{2-}, CO	CO_3^{2-}, $C_2O_4^{2-}$	
Group V[b]	NH_3, CN^-, NO, NO_2^-, NO_3^-, Amines (RNH_2), R_3P, R_3As, R_3Sb, R_3Bi	Organic 1,2- and 1,3-diamines; for example, $H_2NCH_2CH_2NH_2$	Organic polyamines
Group VI[c]	SO_4^{2-}, SCN^-, $S_2O_3^{2-}$, $SeCN^-$, $TeCN^-$, H_2O, OH^-	SO_4^{2-}, SO_3^{2-}, $S_2O_3^{2-}$	
Group VII	F^-, Cl^-, Br^-, I^-		

R = organic group.

[a] Ligands include all types of alkenes ($>C=C<$), alkynes ($-C\equiv C-$), aromatic hydrocarbons, and cyclopropane, as well as isonitriles (RNC).

[b] Ligands include all types of aromatic nitrogen molecules; for example, pyridine, C_5H_5N.

[c] Ligands include many organic molecules containing oxygen, such as alcohols and phenols (ROH), ethers (ROR), acids (RCO_2H), and oximes ($R_2C=NOH$ using N and O atoms).

occupies more than one position in the coordination sphere is referred to as bi-, tri-, tetra-, penta- or hexadentate. These polydentate ligands are known as chelates (from the Greek *chele*, crab's claw). An example of a bidentate ligand is ethylenediamine, $H_2N—CH_2—CH_2—NH_2$, which can attach itself to a metal ion by the unshared electron pairs on both nitrogen atoms.

Stereoisomerism

Werner suggested three basic spatial geometrical arrangements for complexes: octahedral, tetrahedral and square planar. To simplify the discussion, unidentate ligands are labeled a, b, c and so forth, and bidentate ligands, A, B, C and so forth. The three basic coordination sphere shapes can be projected diagrammatically as shown in Figure 4. The metal atom is in the center of each.

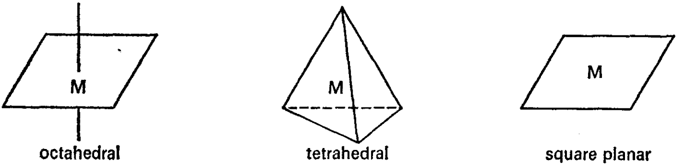
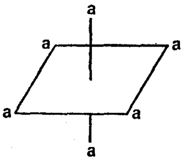


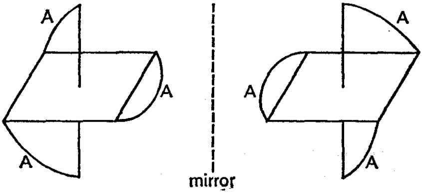
Figure 4

Note that each position is equivalent to any other.

The octahedral model

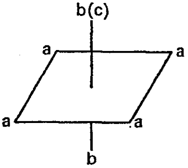
Ma_6 can have only one configuration, but MA_3 can have two arrangements which are mirror images of each other and non-superimposable. Such a pair of isomers is known as a pair of optical isomers or enantiomorphs because one of the few distin-



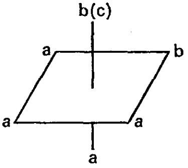


guishing features between members of the pair is their ability to rotate polarized light in equal and opposite directions. The twin which is right-turning is called dextrorotatory or the *d* form, while the left-turning twin is the levorotatory or *l* form. The two together are sometimes referred to as a *dl* pair. An example of Ma_6 is $[Co(CN)_6]^{3-}$ and of MA_3 is $[Co(en)_3]^{3+}$, where en is ethylenediamine, $H_2N—CH_2—CH_2—NH_2$.

Note that the total charge on the complex ion is the sum of all the constituent ions because neutral ligands do not contribute to any change in over-all charge. Ma_5b has no isomers. Ma_4b_2 and Ma_4bc have two isomers.

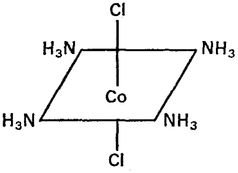


trans I

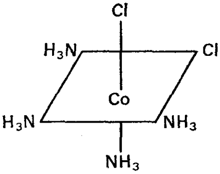


cis II

Structures I and II are so-called geometric isomers which have the same formula but simply differ in the arrangement of ligands around the metal. Geometric isomers differ in both physical and chemical properties in contrast to mirror-image pairs. If the ligands b are grouped next to each other, the name of the complex is prefaced by *cis*. When the two groups are at opposite ends of the octahedron, the prefix *trans* is used. Thus, Ma_4b_2 has a pair of *cis–trans* geometric isomers. An example is the ion $[Co(NH_3)_4Cl_2]^+$. The *trans* form is green and the *cis* form is violet.

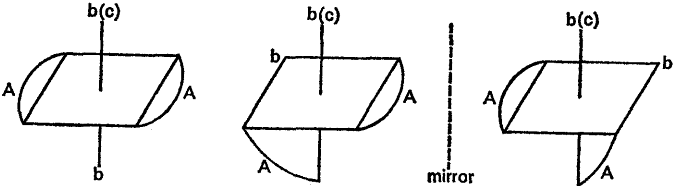


trans (green)



cis (violet)

MA_2b_2 and MA_2bc are the most interesting cases for our purpose. There are three isomers:



trans *d*-and *l*-*cis*

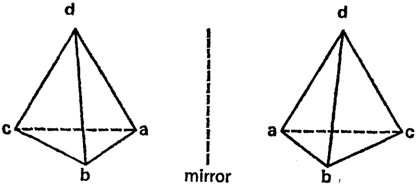
In this instance, the *cis* form exists as a *dl* pair. The ion $[Cr(en)_2Cl_2]^+$ falls into this category and has similar colors to the dichlorotetrammine complex discussed previously. Three isomers are known, of which two are mirror-image structures.

Whenever a *cis* complex is produced in the laboratory, an optically inactive blend of exactly equal parts of *d* and *l* mirror images is formed which is known as a racemic mixture or racemate. These racemates can be separated or resolved by special techniques which make use of the lack of symmetry of the structures.

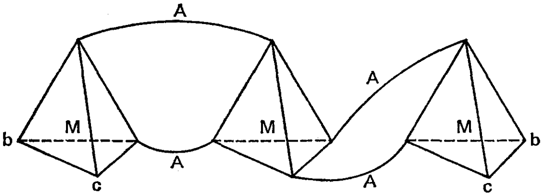
Further treatment of the hexacoordinate octahedral type is left to the ingenuity of the reader. The magnitude of the problem becomes staggering if one considers a compound of type (Mabcdef), for which fifteen pairs of enantiomorphs are predicted theoretically. As a test of your skill and understanding of the problem, work out the number of stereoisomers for the following octahedral complexes and indicate which of them exist as *dl* pairs: Ma_3b_2c, Ma_3bcd, Ma_2bcde, MAb_2c_2, MAb_2cd and MAbcde. The answers are given at the end of this article.

The tetrahedral model

The only case of isomerism in carbon compounds is Mabcd



where a *dl* pair is predicted. This is true, however, only in cases where one central atom is involved in coordination. A trinuclear tetrahedral complex, type $M_3A_4b_2c_2$, with structure

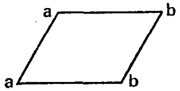


would be expected to exist as two enantiomorphs.

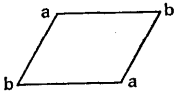
Note that this is also true for $M_3A_4b_2cd$ and M_3A_4bcde, provided the end groups stay attached to a given metal atom. If, however, the various ligands can be interchanged, the number of isomers increases to two *dl* pairs for the first case and three pairs for the second. See if you can work out these examples on paper or, better, with models.

The square planar model

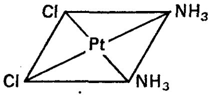
Ma_4, Ma_3b, MAb_2, MA_2 or MAbc exist in only one form. They have no isomers. A compound of form Ma_2b_2 can exist as a *cis–trans* pair:



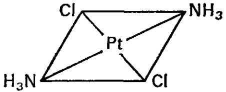
cis



trans



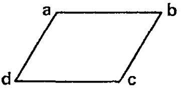
cis

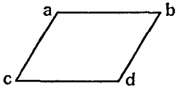


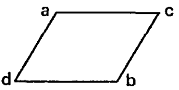
trans

In this *cis* case, however, no enantiomorphic forms are possible. An example of this kind is $Pt(NH_3)_2Cl_2$. Both isomers are yellow, but the *cis* form is somewhat deeper colored than the *trans* isomer. Ma_2bc also has a pair of geometric *cis–trans* forms as above; for example, $[Pt(NH_3)_2(SO_4)(H_2O)]$.

Finally, the type Mabcd should have three geometric isomers:







The synthesis of these three isomers has been realized in the case of $[Pt(NH_3)(NO_2)(C_5H_5N)(H_2NOH)]^+$ and none of the isomers prepared could be resolved into a mirror-image pair.

The three spatial arrangements of inorganic complexes propounded by Werner in analogy to organic compounds have been verified in many cases. However, many predictions of his stereochemical theory have not, as yet, been demonstrated in the laboratory. It is, nevertheless, significant that no evidence contradictory to Werner's coordination postulates has yet come to light. The problem then is one of incompleteness rather than of errors of assumption. The theory is retained, therefore, until it can be systematically disproved.

Other Kinds of Isomerism in Complexes

Coordination isomerism

This phenomenon arises when a set of given ligands is distributed around different central metals. For example, $[Cu(NH_3)_4][PtCl_4]$ and $[Pt(NH_3)_4][CuCl_4]$ are coordination isomers. So are $[Co(NH_3)_6][Cr(SCN)_6]$ and $[Co(NH_3)_4(SCN)_2][Cr(NH_3)_2(SCN)_4]$. These isomer pairs have properties differing completely from one another.

Polymerization isomerism

This term applies to a series of totally different complexes which are multiples of a basic unit compound. Some examples are shown below.

Table 5

No. of units	*Formula*
1	$[Co(NH_3)_3(NO_2)_3]$
2	$[Co(NH_3)_6]^{3+}[Co(NO_2)_6]^{3-}$ and $[Co(NH_3)_4(NO_2)_2]^+[Co(NH_3)_2(NO_2)_4]^-$
3	$[Co(NH_3)_5NO_2]^{2+}2[Co(NH_3)_2(NO_2)_4]^-$
4	$3[Co(NH_3)_4(NO_2)_2]^+[Co(NO_2)_6]^{3-}$
5	$3[Co(NH_3)_5NO_2]^{2+}2[Co(NO_2)_6]^{3-}$

Ionization isomerism

When two different complexes have the same composition but have their ligands arranged differently in the coordination sphere, they are called ionization isomers. For example,

$$[Co(NH_3)_4(NO_2)Cl]^+SCN^-$$

III

$$[Co(NH_3)_4(SCN)Cl]^+NO_2^-$$

IV

$$[Co(NH_3)_4(NO_2)(SCN)]^+Cl^-$$

V

Structure III yields only ionic thiocyanate in solution, while the chloride and nitrite anions are bound as coordinated species and do not give characteristic qualitative tests. Similarly, structure IV yields only nitrite, and structure V chloride.

If this isomerism is brought about by a difference in coordinated *v.* crystal lattice water molecules (water of crystallization), the specific phrase, hydrate isomerism, is used. For example, $[Co(NH_3)_3(H_2O)_2Cl]Br_2$ and $[Co(NH_3)_3(H_2O)ClBr]Br.H_2O$ are hydrate isomers. The first material has all and the second only half its bromine in the ionic form. Ionization and hydrate isomers differ in physical and chemical properties.

Conductance of Penetration Complexes

A readily applicable and widely used method of studying coordination compounds is the measurement of electrical conductivity in water. This approach depends on the fact that coordinated ions, for example Cl^- in $FeCl_6^{3-}$, do not carry current in solutions; thus, the proper assignment of complexed and ionic ligands can be determined. If we choose solutions containing 1 mole of the salt in 1000 liters of water at 25°C, we obtain values in reciprocal ohms, as shown in Table 6.

For example, the conductivity of $CoCl_3.4NH_3.H_2O$ indicates the presence of three ions. This, taken together with the facts that the ammonia molecules are complexed and the coordination number of Co^{3+} is six, leads to the formulation of the substance

Table 6

Electrical Conductivity of 1 Mole of Complex Compounds in 1000 Liters of Water at 25°C

No. ions present	*Conductance range* (Ω^{-1})	*Example*
2	95–120	K^+, $[Co(NH_3)_2(NO_2)_4]^-$ similar to NaCl
3	225–280	$[Co(NH_3)_5Cl]^{2+}$, $2Cl^-$ similar to $BaCl_2$
4	380–435	$[Co(NH_3)_6]^{3+}$, $3Cl^-$ similar to $AlCl_3$
5	520–560	$4K^+$, $[Fe(CN)_6]^{4-}$

Table 7

Some Compounds Exhibiting Unusual Valence States for Metals in the First Transition Series

		Oxidation State		
Metal	*Common Valence*	0	1	3
Cr	2, 3	$Cr(CO)_6$ $Cr(RNC)_6$ $Cr(C_6H_6)_2$	$(C_6H_6)CrI$	
Mn	2	$Mn_2(CO)_{10}$	$K_5Mn(CN)_6$ $CH_3Mn(CO)_5$	$K_3Mn(C_2O_4)_3$
Fe	2, 3	$Fe(CO)_5$ $Fe_2(CO)_9$ $Fe_3(CO)_{12}$ $(RNC)_2Fe(CO)_3$ $(RNC)_2Fe(NO)_2$	$Fe(NO)_2I$	
Co	2	$Co_2(CO)_8$ $(C_2H_2)_2Co_2(CO)_6$	$CH_3Co(CO)_4$ $(RNC)_5CoI$	
Ni	2	$K_4Ni(CN)_4$ $Ni(CO)_4$ $Ni(PCl_3)_4$	$Ni(NO)I$ $K_2Ni(CN)_3(NO)$ $K_4Ni_2(CN)_6$	

as $[Co(NH_3)_4(H_2O)Cl]Cl_2$. This is further substantiated by the observation that only two thirds of the total chlorine present can be precipitated as AgCl with cold silver nitrate.

An important corollary of Werner's ionization theory is that materials such as $[Co(NH_3)_3(NO_2)_3]$ or $[Pt(NH_3)_2(SO_4)(H_2O)]$ should be non-electrolytes or non-conductors in aqueous solution. This is verified by careful conductometric studies.

Other physical methods which are used to determine the number of ions a complex yields in solution, as well as the molecular weight, include freezing point depression, boiling point elevation, and osmotic pressure. These techniques, based on the so-called colligative properties, are treated in physical chemistry textbooks.

Oxidation States of the Transition Metals in Coordination Compounds

One of the many fascinating aspects of penetration complexes is their ability to stabilize lower valence states of transition metals which do not occur in simple salts. Table 7 lists some known compounds exhibiting unusual valence states for metals in the first transition series. Numerous similar complexes are known of other transition elements and may be found in advanced textbooks.

Answer to problem on page 250

Table 8

Compound	*Number of stereoisomers*	*Number of dl pairs*
Ma_3b_2c	3	0
Ma_3bcd	4	1
Ma_2bcde	9	6
MAb_2c_2	3	1
MAb_2cd	4	2
$MAbcde$	6	6

Bibliography

F. Basolo and R. C. Johnson, *Coordination Chemistry*, Benjamin, 1964.

G. B. Kauffman, 'Alfred Werner's coordination theory', *Educ. Chem.*, vol. 4 (1967), no. 1, pp. 11–18.

R. K. MURMANN, *Inorganic Complex Compounds*, Chapman & Hall, 1965.

L. E. ORGEL, *An Introduction to Transition Metal Chemistry – Ligand Field Theory*, 2nd edn, Methuen, 1966.

P. SALTMAN, 'The role of chelation in iron metabolism', *J. Chem. Educ.*, vol. 42 (1965), no. 12, pp. 682–7.

J. SCHULBERT, 'Chelation in medicine', *Scientific American*, vol. 214 (1966), no. 5, pp. 40–50.

14 G. R. Choppin

Nuclear Fission

G. R. Choppin, 'Nuclear fission', *Chemistry*, vol. 40 (1967), no. 7, pp. 25–30.

Nuclear fission is the process whereby a heavy nucleus divides into two fragments of nearly equal mass. This division liberates a large amount of energy, popularly called atomic energy. However, this term is misleading because, unlike energy in a flame, the energy does not arise from changes in the electronic energy levels of the atom but rather from changes in the protonic and neutronic energy levels of the nucleus.

Today, in the post-World War II era, no one can be unaware of nuclear energy and its potential, either for the welfare or destruction of our civilization. It is a complex and fascinating scientific phenomenon in which chemists have played an important role. In this article the history and state of our knowledge about nuclear fission are briefly described.

Background

Discovery of the neutron in 1932 provided scientists with a valuable tool for producing nuclear reactions. The neutron is uncharged and, therefore, when it approaches a nucleus, experiences no repulsion, even though the nucleus has a positive charge. Consequently, neutrons of very low kinetic energies can easily strike nuclei and cause nuclear reactions. Enrico Fermi and his co-workers in Rome realized this and bombarded practically every known element with neutrons (liberated by bombarding beryllium with alpha particles from radium) to produce many new radioactive species.

A characteristic feature of radioactive nuclides produced by capture of a neutron is decay by beta particle (high energy electron) emission which increases the atomic number by one; for example,

$$^{75}_{33}\text{As}+^{1}_{0}\text{n}\longrightarrow ^{76}_{33}\text{As}\longrightarrow ^{76}_{34}\text{Se}+^{\;0}_{-1}\beta.$$

In 1934, Fermi suggested that via beta decay, elements with atomic numbers greater than 92 could be made by bombarding uranium with neutrons. For example, Fermi's group found a new beta activity with a 13-minute half-life which could be separated from all the known elements heavier than lead, including uranium. Therefore, they concluded that this radioactivity must stem from a new element heavier than uranium.

Otto Hahn, Lise Meitner and Fritz Strassman in Germany confirmed that Fermi's radioactivity could not arise from any element between radium and uranium and, using the same logic, these workers found within the next few years a number of other radioactivities which they attributed to new transuranium elements. In fact, it seemed that four such elements had been made which had chemical properties corresponding well with those of homologs of rhenium, osmium, gold and platinum, respectively.

The activity in this area can be gauged by the fact that, in 1938, a review article cited over ninety references supporting discovery of the four new elements. Later, however, these conclusions were proved almost completely wrong. Indeed, the early history of fission is convincing evidence that even the best scientific effort is fallible.

In 1938, Irène Joliot-Curie, the daughter of Pierre and Marie Curie, and P. Savitch investigated a radioactivity of 3·5-hour half-life produced by irradiating uranium with neutrons. This was a puzzling substance. They believed it to be a transuranium element, yet its chemical properties were very different from those reported for what were believed to be transuranium elements. Hahn and Strassman, intrigued by this report, re-investigated the Joliot-Curie and Savitch research. They found radioactivity which was precipitated with barium chloride and, consequently, believed it to arise from an isotope of radium, chemically very similar to barium. To their consternation, however, their 'radium' could not be separated from barium by fractional crystallization. On the other hand, when a compound definitely known to be radium was used, separation was achieved.

Many similar experiments performed with patience and skill led Hahn and Strassman to conclude that their radioactivity, produced as a result of the reaction of neutrons with uranium, was not radium but barium. As chemists they were quite convinced of their results; nevertheless, when, in January 1939, they published their findings, they cautiously spoke of the 'bursting' of uranium (atomic number 92) which had yielded an element such as barium (atomic number 56), much lighter than uranium. It was obvious, nevertheless, that all the so-called 'transuranium' elements were, in fact, isotopes of lighter elements formed by breaking the uranium nucleus.

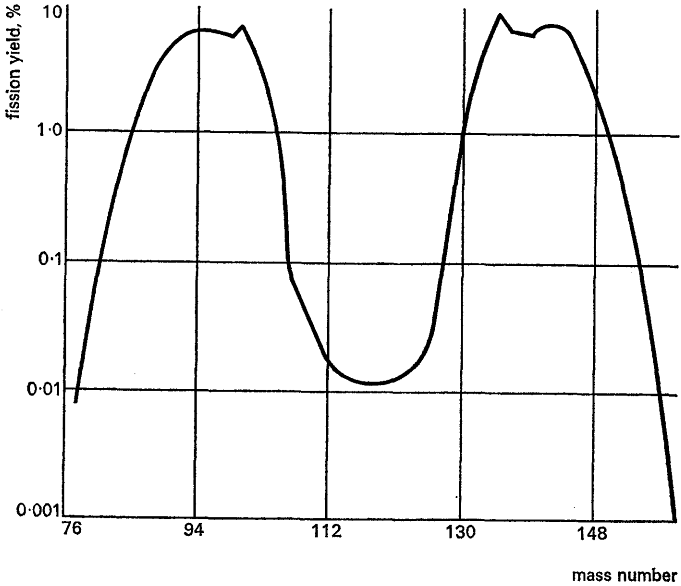


Figure 1 Mass–yield curve for fission of $^{235}_{92}U$ with neutrons of very low energy

Lise Meitner, who had been forced by the Nazi government to leave Germany, was informed of these experiments. Immediately, she and Otto Frisch published a paper giving calculations indicating that if a heavy atom were split or fissioned into two lighter ones, a large amount of energy would be released. Very shortly afterwards, several laboratories confirmed this. A few months later, Niels Bohr and John Wheeler published a paper explaining many features of fission by using a model of nuclear behaviour based on the analogy to a droplet of liquid.

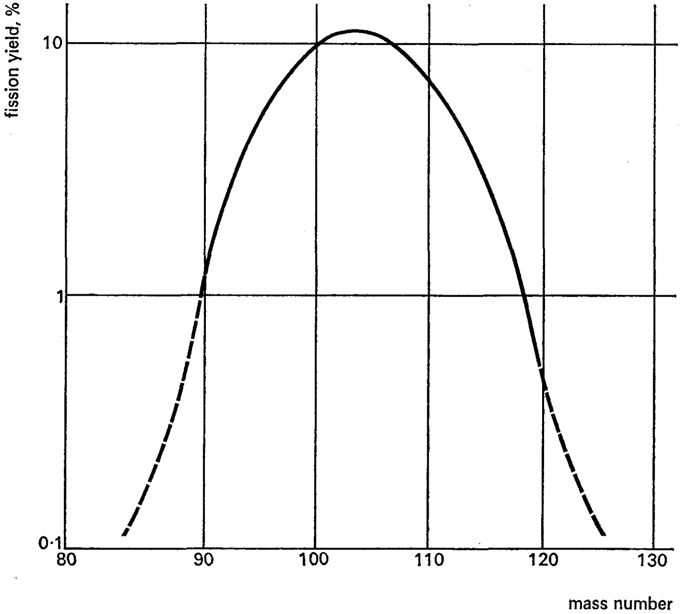


Figure 2 Mass–yield curve for fission of $^{209}_{83}$Bi by deuterons of 22-MeV energy

Within months, the attention of various governments had been drawn to the potential power of a bomb which released fission energy. From then until 1945, although research on fission was cloaked in secrecy, achievements of the Manhattan project in studying nuclear fission were almost beyond comparison. Since then, fission has been studied intensively, and now many thick volumes could be filled with descriptions of the phenomenon.

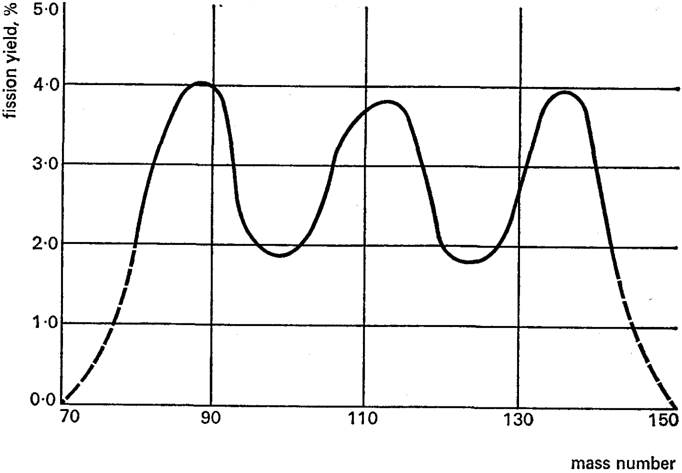


Figure 3 Mass–yield curve for fission of $^{226}_{88}$Ra by protons of 11-MeV energy

Nevertheless, more volumes will have to be filled before sufficient theory is developed to describe satisfactorily the mass and charge division, the energy division and other details of the fission process.

Mass Distribution in Fission

One aspect of fission most studied by chemists is distribution of mass among fragments formed in fission. No matter how nuclei are made to undergo fission, fragments of various masses are formed, which result in production of chemical elements as light as gallium (atomic number 31) and as heavy as gadolinium (atomic number 64).

When uranium fissions from bombardment with low energy neutrons, a plot showing the amount of different masses formed is a curve with two maxima – one near mass number 97 and a second near mass number 137 (Figure 1). These two masses are formed together in the most probable split. Because the mass number of uranium is 235 and that of the neutron is 1, the fissioning system had a total mass of 236 but the sum of the fission products is only 234. The other two mass numbers are accounted for by the emission of two neutrons during fission.

Figure 1 also indicates that only rarely does the fission of uranium-235 by a low energy neutron result in two fragments of equal masses because mass number 117 is the minimum in the curve. Division into fragments of unequal mass numbers is termed asymmetric fission and division into fragments of equal mass number is termed symmetric fission.

Fission of heavy elements other than uranium can be made to occur by bombardment, particularly if we use charged particles such as protons accelerated to high energies – for example, 10 million electron volts (MeV) for thorium and 30 MeV for gold. The mass distribution curve for this type of fission is interesting. At low bombarding energies, it is asymmetric as it is with low energy neutrons. However, as the energy of bombardment is increased, the valley between the peaks of the curve becomes more shallow and, at large enough energies, a single humped symmetric curve is obtained (Figure 2). Thus, the most probable mode of mass split changes from asymmetric at low energies to symmetric at high energies.

This dependence of mass distribution on bombarding energy led to the hypothesis that only two modes of mass division exist in fission – an asymmetric mode favouring a ratio of masses $M_1/M_2 \approx 1{\cdot}5$ and a symmetric mode favouring a ratio $M_1/M_2 = 1{\cdot}0$. The actual mass distribution observed at a particular energy is a result of the relative contributions of these two modes. This two-modes-of-fission hypothesis seemed to receive substantial support from the main distribution curve measured from the fission of radium by protons of 11-MeV energy where three maxima were obtained, two for mass numbers 87 and 136, corresponding to the asymmetric fission mode, $M_1/M_2 = 1{\cdot}6$, and one maximum for 112 corresponding to the symmetric mode, $M_1/M_2 = 1{\cdot}0$ (Figure 3).

Opponents of the two-modes-of-fission hypothesis hold that mass distributions result from many possible modes of fission and that the ratio of M_1/M_2 decreases gradually with increasing energy. The objections which can be made to the two-modes-of-fission hypothesis are sufficiently valid that much more research is needed to explain mass distributions in fission.

Energy of Fission

A simple explanation is available for the large amount of energy released when a nucleus fissions. In nuclear science the heat of formation is termed the binding energy which is expressed in electron volts or million electron volts rather than calories or kilocalories. The heat of formation of the uranium-236 nucleus from its constituent neutrons and protons is 1800 MeV. The heat of formation of a nucleus having a mass number 118 is 1000 MeV. Because two of these nuclei are produced in fission, the total energy release is 2000 MeV. Using these data, an approximate value for the energy released can be obtained by assuming the fission reaction to be:

$$^{235}_{92}\mathrm{U}+\mathrm{n} \longrightarrow [^{236}_{92}\mathrm{U}] \longrightarrow 2\,^{118}_{46}\mathrm{Pd}.$$

The energy released is the difference between the heat of formation of two $^{118}_{46}$Pd nuclei and that of one $^{235}_{92}$U nucleus or $(2\times 1000) - 1800$ MeV. This provides an estimate that 200 MeV of energy are released in fission. This assumes a symmetric division in fission; however, a more complicated calculation gives about the same answer.

The next question concerns how this energy is distributed in fission. With such a large energy release in so disruptive a nuclear process, it is not surprising that fission consists of more than simple division of a nucleus into two fragments. Certain ratios of neutrons to protons (the n/p ratio) lead to stable nuclei. When the ratio is not that necessary for stability, the nuclei are unstable with respect to radioactive decay. The n/p ratio for stability is 1·3 for palladium; consequently, $^{118}_{46}$Pd (n/p = 1·6) has too many neutrons. Because the n/p ratio for $^{235}_{92}$U is 1·6, while the ratio necessary for stability is 1·2 to 1·4 in the elements produced in fission, fission fragments always have too large a value for the n/p ratio. To compensate for this, several neutrons are emitted in the act of fission. However, the number of neutrons emitted is not sufficient to lower the n/p ratios to stable values. To lower the n/p ratios further, the fission fragments, after neutron emission, successively undergo radioactive decay steps in which beta particles are emitted. In beta decay neutrons are converted in the nucleus to protons, and an electron is ejected. As an example, assume two neutrons were emitted in fission so the fission process can be written as

$$^{235}_{92}\mathrm{U} + \mathrm{n} \rightarrow {}^{236}_{92}\mathrm{U} \rightarrow 2\,^{117}_{46}\mathrm{Pd} + 2\mathrm{n}.$$

The $^{117}_{46}\mathrm{Pd}$ then decays by the sequence:

$$^{117}_{46}\mathrm{Pd} \xrightarrow[\beta^-]{\text{very short}} {}^{117}_{47}\mathrm{Ag} \xrightarrow[\beta^-]{1{\cdot}1\text{ min}}$$

$$^{117}_{48}\mathrm{Cd} \xrightarrow[\beta^-]{1\text{ hr}} {}^{117}_{49}\mathrm{In} \xrightarrow[\beta^-]{2\text{ hr}} {}^{117}_{50}\mathrm{Sn}\ \text{(stable)}.$$

The times refer to the radioactive half-life of each decay.

In addition to neutrons, several gamma rays are emitted also in fission. Therefore the fission energy appears as the energy of gamma rays, as the kinetic energy of the emitted neutrons and of the fission fragments, and as the energy of beta decay of the fragments. Table 1 shows how the fission energy is distributed for a total fission energy of 200 MeV.

Table 1
Approximate Fission Energy Balance

	Fission Energy (MeV)
Kinetic energy of fission fragments	170
Energy of gamma rays	16
Energy of beta decay	9
Kinetic energy of emitted neutrons	5
Total	200

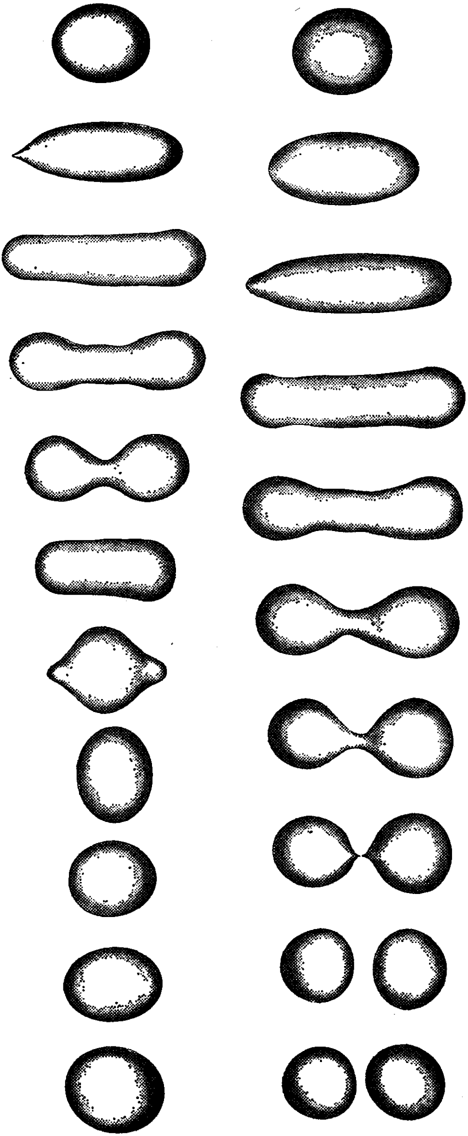
The Liquid Drop Model

Niels Bohr and John Wheeler explained fission just months after its discovery by comparing it to a droplet of incompressible liquid. The surface tension of a liquid causes a droplet to assume a spherical shape, but if energy is supplied in some fashion, this shape is distorted. If the attractive surface tension force is greater than the distorting force, the drop will oscillate between spherical and elongated shapes. If, however, the distorting force becomes larger than the attractive force, the drop will elongate past a threshold point and split (fission). S. G. Thompson and coworkers have photographed an actual drop in which the energy of distortion was supplied electrically rather than by heating (Figure 4).

In nuclei there exist two opposing energies. A disruptive energy results from the mutual repulsion of the protons in the nucleus. This Coulomb energy increases as the number of protons, Z, increases and decreases as the average distance, d_{p-p}, between the protons increases. Because each of the Z protons repels the other $Z-1$ protons in inverse proportion to the interproton distance, the repulsive energy is expressed as being proportional to $Z(Z-1)/d_{p-p}$. For heavy nuclei where Z is about equal to $(Z-1)$, the energy of repulsion is proportional to Z^2/d_{p-p}. For incompressible, spherical nuclei it can be shown that the radius of a nucleus is proportional to the cube root of its mass number, A. In such a sphere, for a constant value of Z, the average distance between protons is proportional to the radius and, hence, d_{p-p} is proportional to $A^{\frac{1}{3}}$. Therefore, the energy of repulsion is proportional to $Z^2/A^{\frac{1}{3}}$.

The attractive energy in nuclei is due to the nuclear forces present between adjacent nuclear particles (nucleons) and is equally strong between two protons, two neutrons, and a neutron and a proton. Nucleons at the surface of a nucleus are not in contact with as many other nucleons as those inside the nucleus and, therefore, do not experience as great a total nuclear force. The total nuclear attractive force is a maximum when the nucleus assumes a spherical shape because in this configuration, a minimum number of nucleons is on the surface. For a sphere the surface area is related to the square of the radius. Because the radius of a nucleus is proportional to $A^{\frac{1}{3}}$ and the attractive energy is related to the surface area, energy of attraction $\propto$ (radius)^2 $\propto$ $(A^{\frac{1}{3}})^2 \propto A^{\frac{2}{3}}$.

The analogy to a liquid drop is simple. The repulsive Coulomb force tends to distort the nucleus in the same way a distorting force does a droplet. The attractive nuclear force tends to make it spherical just as the surface tension does a droplet. The ratio



between the two opposing energies should measure the instability to fission of the nucleus:

$$\text{instability} \propto \frac{\text{energy of repulsion}}{\text{energy of attraction}} \approx \frac{Z^2/A^{\frac{1}{3}}}{A^{\frac{2}{3}}} \approx \frac{Z^2}{A}.$$

This Z^2/A is known as the fissionability parameter because the liquid drop model predicts that the probability of fission should increase with increase in Z^2/A. Figure 5 plots the relation between the half-life for spontaneous fission and Z^2/A showing that half-lives get shorter (fission gets more probable) as Z^2/A increases. In fact, the simple liquid drop model predicts that when Z^2/A has a value of about 45, the half-life to fission should be about 10^{-20} second, or the time of a single vibration of the nucleus. This predicts that elements with atomic numbers greater than 110 to 120 cannot be synthesized. Recently, evidence has indicated that this is likely to be incorrect which is not too surprising because the simple liquid drop model is no doubt too simple to predict correctly all the time.

Summary of Fission Process

Figure 6 summarizes the sequence of events in the over-all fission process. Of great practical importance is the emission of several neutrons. A single neutron causes a fission in which two or three neutrons are emitted; these in turn can cause more fission, etc. Because more neutrons are emitted in each fission than are absorbed to cause fission, a rapidly multiplying process or a chain reaction occurs. If the chain reaction proceeds uncontrolled, an explosion (the atomic bomb) results because of the rapid

Figure 4 Variation in shape of an ordinary drop of water suspended in oil when deformation is induced by voltage applied across the oil. In the left sequence, the drop returned to its initial spherical shape without undergoing splitting (fission). In the right sequence, the initial deformation is sufficiently large that the drop fissions rather than return to a spherical shape. Bohr and Wheeler used the analogy of the splitting of such a drop of liquid to explain nuclear fission [Adapted from R. B. Leachman, 'Nuclear Fission', *Scientific American*, August 1965, pp. 49–59. Copyright © Scientific American, Inc., 1965. All rights reserved.]

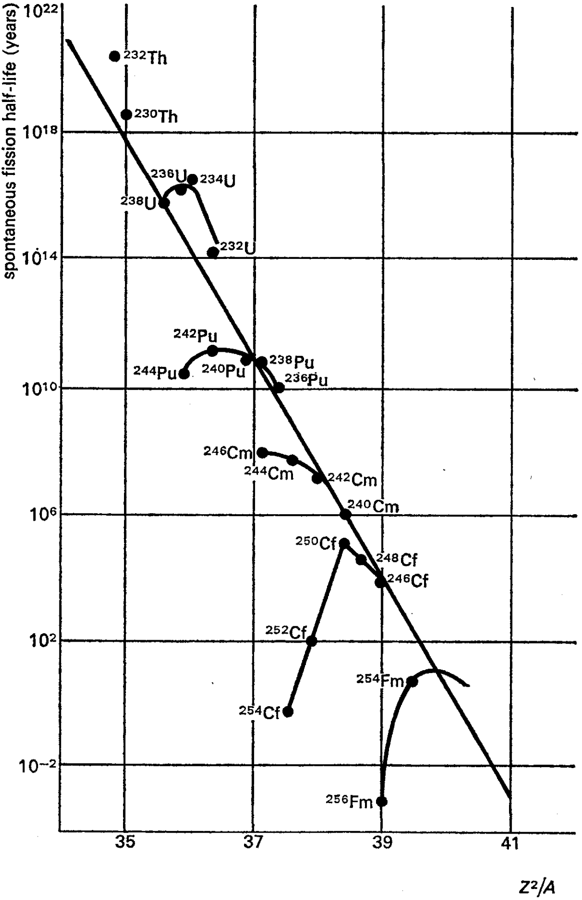


Figure 5 Variation in half-life for spontaneous fission with changes in Z^2/A values

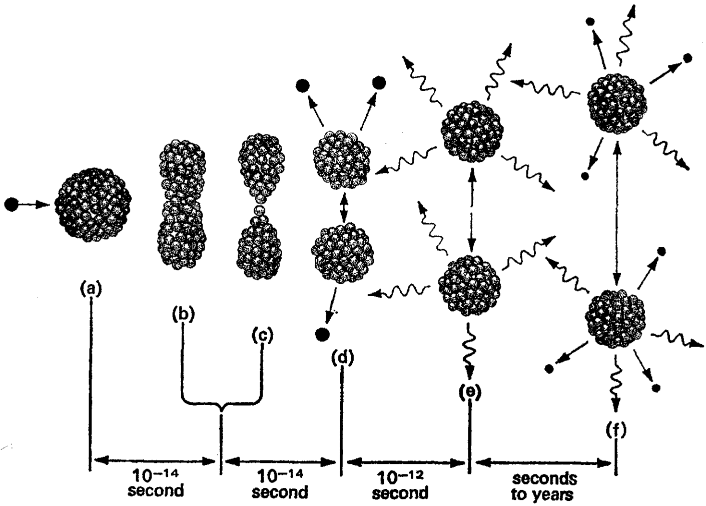


Figure 6 Sequence of events in the fission of a uranium nucleus by a neutron. In (a) the neutron strikes the nucleus and is absorbed, causing the nucleus to undergo deformation, (b). In about 10^{-14} second, one of the deformations, (c), is so drastic that the nucleus cannot recover and fissions, (d), releasing two or three neutrons. In about 10^{-12} second, the fission fragments have lost their kinetic energy and have come to rest, emitting a number of gamma rays. In the final stage, the excess nuclear energy is removed from the fission fragments by the emission of beta particles and gamma rays over a period of time from seconds to years [From R. B. Leachman, 'Nuclear Fission', *Scientific American,* August 1965, pp. 49–59. Copyright © Scientific American, Inc., 1965. All rights reserved.]

release of fission energy. Alternatively, if an element such as cadmium is present to absorb some of the neutrons, thereby controlling the chain reaction as in a nuclear reactor, the reaction rate can be kept equal to the rate of removal of energy so that fission energy is used to do work rather than explode.

In the years just after World War II, it was believed that nuclear energy from reactors would be relatively simple to obtain. However, many complex problems arose and even today only a small fraction of the world's power needs are being met by nuclear power. Progress has been more rapid in the past few years, and it seems likely that rather soon nuclear power will challenge the more conventional power sources as a means of abundant, inexpensive power to satisfy the world's ever-increasing needs.

Bibliography

G. R. CHOPPIN, *Nuclei and Radioactivity*, Benjamin, 1964.
B. G. HARVEY, *Nuclear Chemistry*, Prentice-Hall, 1965.
R. T. OVERMAN, *Basic Concepts of Nuclear Chemistry*, Chapman & Hall, 1965.
G. T. SEABORG, *Man-Made Transuranium Elements*, Prentice-Hall, 1963.

Part Six Organic Reaction Mechanisms

The concept of mechanism is another of the main themes of modern chemistry and G. Illuminati gives an account of the mechanisms of some of the main types of organic reaction in his article 'Organic reactions': acid–base relationships, substitution reactions of alkyl halides, aromatic substitution and elimination, rearrangement and addition reactions are all discussed and concepts such as the reaction intermediate and transition state are introduced. The importance of kinetic studies in the elucidation of reaction mechanisms is emphasized.

In recent years remarkable advances have been made in the study of *inorganic* reactions in aqueous solution. A. G. Sharpe touches on the mechanism of oxidation–reduction reactions in his article on page 178. The reader is also recommended to read an article by A. G. Sykes on 'Kinetics and mechanisms of inorganic reactions in solution' (*Educ. Chem.*, vol. 2 (1965), no. 6, pp. 277–82).

15 G. Illuminati

Organic Reactions

G. Illuminati, 'Organic reactions', *Chemistry Today*, O.E.C.D., 1963, ch. 14, pp. 259–84.

Introduction

This chapter is intended to give a presentation of some of the main types of reactions encountered in organic chemistry. Although, because of the limited space available and, also, of the nature of the present book, a basic knowledge of the systematic classification of organic compounds is assumed here, this chapter is not an outline of an 'advanced', separate course of organic chemistry but rather a presentation of one particular aspect – that is chemical reactivity – of the whole subject matter of organic chemistry.

Reactivity is admittedly one of the most complex problems in the whole of chemistry but it is also one of the most important – in fact, the ultimate one, since a sound understanding of the course of reactions has an essential bearing on many other aspects of chemistry such as choice of synthetic routes, prediction of properties and stabilities under a variety of conditions, rationalization of natural occurrences, search for practical uses, search for lower-cost production, and so on. Since the course of many organic reactions is now well understood through the application of all available experimental techniques, this aspect of chemistry can be mastered quite well, and the new insights can be introduced at any level of learning, no matter how elementary. A few general principles may be helpful in this connexion.

Firstly, any chemical reaction must be made to correspond, as a result of experiment, to a stoichiometric equation to account for the material and chemical balance of an over-all transformation process from a definite set of reactants to a definite set of products.

Secondly, the stoichiometric equation provides a relation between chemical species, answering the question of what products are formed by what reactants: however, it says nothing about how the change takes place. The latter question needs careful and subtle experimental approaches to be worked out and, unless the course of the reaction has been established by experiment, rationalizing of the reaction by arbitrary guesses is to be avoided at all times. Of course a student, after gaining a sufficient background, can make predictions on the basis of analogy to known related reactions, and it is good if he does so.

Thirdly, experimental conditions must be defined in expressing a reaction by its stoichiometric equation. If conditions are changed, the reaction may either not occur at all or may give rise to different products. Conversely, under certain experimental conditions, a single stoichiometric equation may not be sufficient to describe the over-all reaction, since the reactants may react in different parallel ways which are called competitive reactions. Usually, the relative importance of competitive reactions is strongly dependent on experimental conditions. Experimental conditions include temperature, pressure, catalysts, light, solvent and so on. Most organic reactions occur in the condensed phase and may be strongly dependent on the nature of the solvent. The reason why the solvent is so important in organic reactions is that organic compounds, unlike inorganic, are soluble in a great number and wide variety of solvents which are found to affect the course of reaction from one case to another.

Finally, an organic reaction, under a given set of conditions, may appear to be rather complex, as it involves the participation of a number of reacting molecules, with old bonds to be broken and new bonds to be made. Usually, however, a reaction as complex as this is made up of a succession of simpler acts, each involving a limited number of reacting molecules with a very limited number of synchronous bond changes. Depending on the nature of the succession of simple acts the over-all reaction passes through intermediates of varying stabilities. Some of these can be isolated, others cannot.

A good understanding of a chemical reaction requires a knowledge of the succession of all the intermediates involved in the process. This is what is usually meant by 'course of reaction'. The first important stage in elucidating a reaction is the detection or, better, actual isolation of the intermediates of some stability. Mechanism is a more sophisticated term and has come to be reserved for the detailed reaction path by which bond changes occur and for a good estimate of the energetics of such changes. Solving a reaction mechanism completely is a task of considerably greater difficulty than the isolation of the stable intermediates, and usually requires the application of the so-called physical methods of organic chemistry. Even though the mechanism of most organic reactions has been elucidated in its essentials, a great many problems are still open to advanced research on practically every reaction. Keeping this in mind, we outline in the following sections the essential facts and interpretations concerning the nature of some organic reactions in order to give an idea of current trends of thought in this field. To this end, we have limited our considerations to acid–base relationships, substitutions and additions which, although they do not by any means exhaust the reaction types, occur throughout organic chemistry even in the more complex reactions. A basic knowledge of molecular structure has been assumed here.

Needless to say, we do not pretend to describe the reactions as they actually are but, rather, as they are viewed by us today by use of a certain amount of approximation necessary in such a brief account as this.

Acids and Bases[1]

According to Brønsted's definition, an acid is a substance capable of releasing a proton, and a base one capable of accepting a proton. Solvents themselves possess acid–base properties; for this reason the dissociation of acids and bases depends largely on the nature of the solvent. For example, water is slightly dissociated as follows:

$$2H_2O \rightleftharpoons H_3O^+ + OH^-,$$

thus showing a tendency to behave both as an acid and as a base (amphoteric behaviour). When an acid HA or a base B is dissolved in water the following equilibria are established:

$$HA + H_2O \rightleftharpoons H_3O^+ + A^-,$$
$$B + H_2O \rightleftharpoons HB^+ + OH^-.$$

---

[1] See also chapter 10.

The above equilibria show that water, on account of its amphoteric properties, allows both types of dissociation to occur. However, this is not always the case, as we shall see further on.

Now, many organic compounds possess acid or basic properties and, since they dissolve and undergo reactions in a wide variety of solvents, the study of their acid–base relationships with solvent, as well as with other species present in solution, is of fundamental importance and deserves consideration for a deeper understanding of more complex reactions such as those involving acid or base catalysis.

Organic molecules are known to be hydrogen-rich molecules: however, only a few hydrogen atoms located in special positions are 'acidic'. Among the more familiar bonds involving hydrogen, C—H, N—H and O—H, the latter is by far the most commonly present in acidic organic functions.

The main hydroxyl-group-containing organic acids are carboxylic acids, RCO_2H, and phenols, ArOH:

RCO_2H: Formic acid $HCO_2H + H_2O \rightleftharpoons HCO_2^- + H_3O^+$
Acetic acid $CH_3CO_2H + H_2O \rightleftharpoons CH_3CO_2^- + H_3O^+$
Benzoic acid $C_6H_5CO_2H + H_2O \rightleftharpoons C_6H_5CO_2^- + H_3O^+$
ArOH: Phenol $C_6H_5OH + H_2O \rightleftharpoons C_6H_5O^- + H_3O^+$.

The hydroxyl group of alcohols ionizes to a negligible extent in water. This is generally true for hydroxyl groups attached to saturated carbon atoms. Non-saturation can be held responsible for the acid strength of the above-mentioned classes, where the organic anion is especially stabilized through resonance (Figure 1). As a consequence of the organic anion stabilization, the

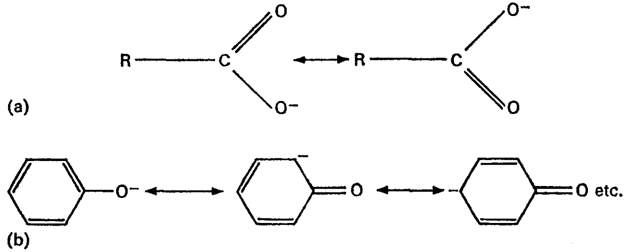


Figure 1 Resonance stabilization for organic anions, (a) carboxylate ion (b) phenate ion

above equilibria will be shifted to the right and the acid strength will be enhanced as compared with that of alcohols.

There are quite a few examples of acidic hydrogens attached to carbon. We may mention cyclopentadiene, β-diketones, and usually all carbonyl compounds containing α-hydrogens. The acid strength of these compounds is, however, very low and its importance lies in the role of acid–base catalysis in organic reactions (see subsequent sections).

The strength of acids is measured by the equilibrium constant related to the dissociation reaction. In organic homologous series, the relative strength of an individual acid with respect to another taken as reference is widely used to describe the influence of structure on dissociation. Suppose HA is any acid in the series and HA_0 the reference acid. The equilibria will be expressed as follows:

$$HA + S \rightleftharpoons A^- + HS^+;\ K = \frac{[A^-][HS^+]}{[HA]}$$

$$HA_0 + S \rightleftharpoons A_0^- + HS^+;\ K_0 = \frac{[A_0^-][HS^+]}{[HA_0]}$$

and:

$$HA + A_0^- \rightleftharpoons HA_0 + A^-;\ K/K_0 = \frac{[A^-][HA_0]}{[A_0^-][HA]}$$

where K/K_0 defines the relative strength in a given solvent S. Tables 1 and 2 list a number of dissociation constants for aliphatic carboxylic acids and for benzoic acids. It will be noted that acetic acid is a markedly weaker acid than formic acid, but an additional methyl group, as in propionic acid, has only a minor effect on the constant. Also, chlorine atoms strongly enhance the acidity when in the α-position but much less so in the β-position.

Table 1
Acid Strengths of Aliphatic Acids in Water at 25°C

RCO_2H	Formic	Acetic	Propionic
$10^5\ K_a$	17·8	1·8	1·3
RCO_2H	Chloroacetic	α-chloropropionic	β-chloropropionic
$10^5\ K_a$	155	147	8·2

Table 2

Acid Strengths of Substituted Benzoic Acids in Water at 25°C

$ArCO_2H$	Benzoic	*m*-methyl	*m*-chloro	*m*-nitro
$10^5 K_a$	6·27	5·35	14·8	32·1

These data show that acid strengths of aliphatic acids may be greatly affected by structural changes, provided that such changes are made in the vicinity of the carboxyl group. More generally, substituent effects on acidity are found to fade rapidly on moving the substituent away from the acidic group:

XCO_2H
CH_2XCO_2H
$CH_2XCH_2CO_2H$
$CH_2XCH_2CH_2CO_2H$
etc.

↓ decreasing order of the inductive effect (see below) of X on acidity

The structure of an organic molecule can affect its chemical properties in several ways, as this and further sections will illustrate. A useful approach is to consider atoms or groups of atoms as 'substituents' for hydrogen in a given reference molecule. Substituent effects on properties are essentially of two different kinds depending on whether they are related to the electronic transmission through bonds and space (electronic effects) or to the steric compression due to the size of neighbouring groups (steric effects). Electronic effects may arise from the polarity of bonds and be transmitted mainly through σ-bonds (inductive effects); or they may be relayed through conjugated systems by *p*- and π-electron displacements (mesomeric or resonance effects; hyperconjugation); finally, they may originate from the peripheral polarity of substituents and the electrostatic field thus produced in the surroundings of the reaction centre (field effects).[1]

The structural correlations noted above in the carboxylic acid series are characteristic of the relay of the inductive effects of substituents through σ-bonds.

As to the direction of the effects, the carboxyl ion stabilization

---

[1] For a detailed description of substituent effects see the 'References' given at the end of this article.

mentioned above is expected to increase, and the acidity to become greater, if electronic charge is pumped out of the anion, and to decrease in the opposite case. This is consistent with the enhanced acidity observed in chloroaliphatic acids and the reduced acidity of methyl-substituted acids. The same trends are found in acids of other series. Thus, Table 2 shows that chloro- and nitrobenzoic acids are stronger than the unsubstituted acid, in agreement with the electron-withdrawing direction of the inductive effects of Cl and NO_2; whereas toluic acids are weaker, in agreement with the electron-releasing direction of the inductive effects of alkyl groups.

Bases, being proton acceptors, must be electron donors. Oxygen in alcohols and ethers and nitrogen in amines are important examples of organic neutral electron donors. In these classes of compounds, these atoms can use unshared electron pairs to make an extra bond with a proton (coordinate bond), thereby increasing their covalency by one unit:

$$\begin{gathered} C_2H_5OH + H_2O \rightleftharpoons C_2H_5OH_2^+ + OH^- \\ (C_2H_5)_2O + H_2O \rightleftharpoons (C_2H_5)_2OH^+ + OH^- \\ C_2H_5NH_2 + H_2O \rightleftharpoons C_2H_5NH_3^+ + OH^- \\ C_6H_5NH_2 + H_2O \rightleftharpoons C_6H_5NH_3^+ + OH^-. \end{gathered}$$

Nitrogen bases are generally stronger than oxygen bases. Also, their strength depends markedly on structure; thus aliphatic primary amines are stronger than aniline. The latter substance is stabilized by resonance in the unionized form (Figure 2), so that the equilibrium for aniline will be shifted toward the left more than the equilibrium for ethylamine. This situation is analogous to that which makes phenol a stronger acid than ethanol. When homologous series of organic bases are examined, relative base strengths, K/K_0, are often used for comparison. Table 3 gives the equilibrium constants for some anilinium ions.

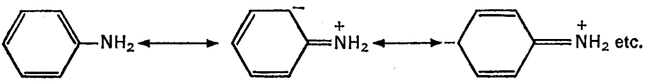


Figure 2 Resonance forms of aniline

Amides have long been recognized to be a separate class of organic nitrogen compounds when compared with amines.

Table 3
Acid Strengths of Substituted Anilinium Ions in Water at 25°C

$ArNH_3^+$	Anilinium ion	*m*-methyl	*m*-chloro	*m*-nitro
$10^5 K_a$	2·40	2·14	47·87	354·9

However, as far as basicity is concerned, the almost complete loss of basic character of the nitrogen is simply due to the electron-withdrawing carbonyl group attached to the nitrogen.

Secondary and tertiary amines also have a pronounced basic character. Change in base strength can be fairly well interpreted in terms of structural effects. A very important class of tertiary amines is that of heterocyclic compounds such as pyridine, pyrimidine, quinoline and others.

Going from uncharged oxygen or nitrogen bases to their corresponding negatively charged ions a large increase in base strength is noted. Here basicity is linked with the tendency of the anion to form an ordinary covalent bond with a proton. Thus, while water is a very weak base, the hydroxide ion OH^- is a strong base. On account of this, alcohols are negligibly dissociated as acids in water solvent, but in the absence of water the equilibrium:

$$C_2H_5OH + OH^- \rightleftharpoons C_2H_5O^- + H_2O$$

is nearly completely shifted to the right. Thus a strong oxygen organic base is generated – the ethoxide ion, which is the conjugate base of C_2H_5OH. Alkoxides are an important class of organic reagents to which we will return later. It is worth recalling that alkali metal alkoxides are best prepared by reaction of the alkali metal with anhydrous alcohol:

$$Na + C_2H_5OH \longrightarrow C_2H_5ONa + \tfrac{1}{2}H_2.$$

The salt can be isolated on evaporation but is rapidly hydrolysed if exposed to moisture because of the strong basicity of the organic anion.

Some nitrogen organic bases closely resemble hydroxide-ion-releasing inorganic compounds. These are the so-called quaternary bases such as tetramethylammonium hydroxide:

$$(CH_3)_4NOH \longrightarrow (CH_3)_4N^+ + OH^-.$$

Their aqueous solutions are strongly basic.

The rates at which acid–base equilibria are attained are usually too high for measurement by conventional methods (conductivity, etc.). However, there are cases where the dissociation is slow. Carbon acids, such as nitroalkanes and α-hydrogen-containing carbonyl compounds, are examples of this kind:

$$CH_3NO_2 \rightleftharpoons H^+ + CH_2NO_2^-.$$

There is a marked degree of electronic charge rearrangement in the anion $[CH_2NO_2]^-$, as shown by the following resonance structures:

$$^-CH_2{-}NO_2 \longleftrightarrow CH_2{=}NOO^-.$$

Slow-dissociating acids involving electronic charge rearrangement in the anions produced have often been referred to as pseudo-acids. However, this distinction has lost much of its validity because the rate criterion is not an absolute one. Thus, for rapidly established equilibria, relaxation methods are now available allowing the determination of the reaction rates. On the other hand, there are many oxy-acids (such as carboxylic acids and phenols) whose reaction rates are very fast, and which involve a good deal of electronic rearrangement in the anions (Figure 1).

We may now give brief consideration to non-aqueous solvents in relation to the state of organic solutes. A number of these show, like water, self-ionization and possess, therefore, amphoteric properties, as in the following examples (ionic products shown on the right):

alcohols	$2CH_3OH \rightleftharpoons CH_3OH^+ + CH_3O^-$	10^{-17}
carboxylic acids	$2CH_3CO_2H \rightleftharpoons CH_3CO_2H_2^+ + CH_3CO_2^-$	10^{-13}
	$2H_2SO_4 \rightleftharpoons H_3SO_4^+ + HSO_4^-$	10^{-4}
	$2NH_3 \rightleftharpoons NH_4^+ + NH_2^-$	10^{-33}

The influence of the above equilibria on the acid–base properties of organic solutes varies greatly from one case to another, depending on the ionic product. Thus, uncharged nitrogen bases behave as weaker bases in alcohols than in water ($K_w = 10^{-14}$), and as stronger bases in acetic acid than in water. For example, aniline is completely converted into the anilinium ion in acetic acid and the much weaker nitroanilines are extensively ionized in this solvent. Sulfuric acid is such a good proton donor that nearly all oxygen- and nitrogen-containing compounds will behave as bases in this reagent. Such weak bases as amides and ethers are completely protonated in this solvent. Sulfuric acid is also able to protonate carboxylic acids and even aromatic hydrocarbons to a measurable extent.

Relative strengths, K/K_0, for compounds of a general series, such as substituted benzoic acids and anilinium ions, depend to some extent on solvent, as the following few examples show.

Benzoic acid (log(K/K_0), relative to benzoic acid)

m-CH_3	−0·06	(water)	−0·09	(methanol)	−0·06	(ethanol)
m-Cl	0·38		0·59		0·63	
m-NO_2	0·72		1·05		1·17	

Anilinium ions (log(K/K_0), relative to anilinium ion)

m-CH_3	−0·09	(water)	−0·19	(methanol)	−0·29	(ethanol)
m-Cl	1·26		1·53		1·55	
m-NO_2	2·13		2·91		3·24	

The effect of the solvent on relative strengths will become still greater if solvents of more varying types and solutes not belonging to the same series are compared.

General Considerations Regarding Organic Reactions

Reversible and irreversible reactions

Organic reactions show very large variations in the rates at which they proceed. Some are reversible, but a much greater number is irreversible. Many (but not all) reversible reactions, such as acid–base reactions, are very fast. They are usually described by their equilibrium constants only, because of the experimental difficulty in following the rates at which equilibria are attained and, therefore, of the lack of this kind of information. On the other hand, irreversible reactions, if slow enough to be measured, can only be described in terms of their kinetic course and have in fact been given a great deal of attention in the last few decades from this point of view. The two ways of description are not equivalent, for reversible reactions are equilibrium-controlled and the composition of the reaction products depends on the relative stability of their components; while irreversible reactions are rate-controlled in the sense that if two or more competitive reactions starting from the same substance are conceivable, the composition of the final reaction mixture depends on the relative rates at which the products are formed, rather than on their relative stabilities.

The reactions which will be illustrated in the following sections are, in most cases, irreversible. In the illustration of reaction mechanisms, which are responsible for reaction rates, frequent mention will be made of the major kinetic observations supporting them.

Intermediates and transition states

We wish to deal briefly with only one particular point from the field of chemical kinetics – this concerns the nature of the intermediates that may be formed in the course of organic reactions. Throughout this chapter, by 'intermediates' we mean true chemical compounds formed in the course of a given process. In fact, in many cases these intermediates cannot be isolated because of their very low stability. Despite this, a reaction intermediate is never to be confused with the concept of 'transition state' discussed in the theory of absolute reaction rates. In principle, the two concepts can be kept distinct from each other in a clean-cut way using the fact that a chemical intermediate corresponds to a minimum, and a transition state to a maximum, in the potential energy diagram of the reacting system (see Figure 9). Also, unlike a chemical intermediate, a transition state involves one or more partly formed and partly broken covalent bonds and is usually accompanied in organic reactions by a marked degree of electronic rearrangement.

Classification of reagents

A final general observation to be made on organic reactions concerns the classification of reagents. In all reactions, reagents fall into three main classes: nucleophilic, electrophilic and homolytic.

Nucleophilic reagents (or nucleophiles) act as donors of unshared electron pairs to a particular atom of the organic molecule ('substrate'), usually carbon, for sharing in covalent bond formation. When a nucleophile acts on a proton, it falls into Brønsted's definition of a base. However, the nucleophilic power of a reagent (or nucleophilicity) does not exactly parallel its basicity. This is due to polarizability effects. Examples of nucleophiles are H_2O, OH^-, ROH, RO^-, CN^-, I^-, NH_3, RS^-, etc. These examples show that a nucleophile can either be a negatively charged ion or a neutral molecule.

Electrophilic reagents (or electrophiles) act as acceptors of unshared electron pairs from a particular atom of the substrate, usually carbon, for sharing in covalent bond formation. They include, as a special case, Brønsted's acids when a proton from the reagent is the actual acceptor. Electrophiles can be positively charged ions, neutral molecules containing electron-deficient atoms (the so-called Lewis acids), or polarizable molecules. Examples of electrophiles are H^+, NO_2^+, carbonium ions, BF_3, Cl_2, Br_2, etc.

Homolytic reagents are radicals using an unpaired electron for covalent bond formation. Examples are $CH_3^\cdot$, $C_6H_5^\cdot$, $(C_6H_5)_2N^\cdot$, Cl, H, etc.

Some Reactions of Alkyl Halides, Alcohols and Related Compounds

One main feature of many classes of organic compounds, such as halides, alcohols, phenols and, more generally, compounds of the type RX, is the high degree of interconvertibility among each other through metathetical reactions. These reactions occur very frequently in organic chemistry and are often of great value in syntheses.

Depending on the nature of the organic compound undergoing reaction ('substrate'), they may be referred to as substitutions at saturated carbon atoms, substitutions at unsaturated carbon atoms, or aromatic substitutions; whereas depending on the nature of the attacking reagent they are called nucleophilic, electrophilic or homolytic substitutions.

In this section we wish primarily to illustrate in some detail those reactions of alkyl halides and related compounds which are classified as nucleophilic substitutions at saturated carbon, because they offer one of the most complete pictures of an organic reaction in relation to the structure of the substrate, to the experimental conditions and to the variety of the methods employed to elucidate their mechanism. Then at the end of the section, we shall give a brief account of alkene-forming eliminations and molecular rearrangements.

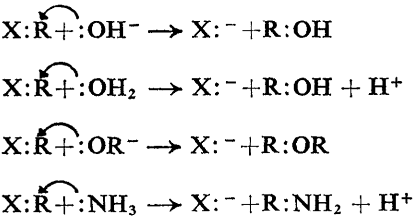
Nucleophilic substitutions at saturated carbon atoms

Some examples of nucleophilic substitutions at saturated carbon atoms for an alkyl halide, RX, are as follows:

1. $RX+KOH \longrightarrow KX+ROH$ (alcohol).
2. $RX+H_2O \longrightarrow HX+ROH$ (alcohol).
3. $RX+KOR \longrightarrow KX+ROR$ (ether).
4. $RX+NH_3 \longrightarrow HX+RNH_2$ (amine).

The scope of the above reactions can be further illustrated by its extension to many substrates other than the halides, such as alcohols, ethers, esters, alkylsulfuric acids, alkyl sulfonates, ammonium and sulfonium salts. Moreover, the list of nucleophiles, other than hydroxide ions, water, alkoxide ions and ammonia, can be extended to carboxylic acids and their salts, primary, secondary and tertiary amines, hydrazine, halide ions, phosphorus trihalides, thiols, sulfides and many others.

The stoichiometry of the reactions, as the four schemes above show, is very simple, but it does not disclose any of the interesting aspects involved in the dynamics of the process. Some additional information can be obtained by consideration of the structure of the bonds involved and of the species present in solution. Thus, the bond involved in the reaction of RX is a covalent bond provided with some degree of polarity; so are the new bonds formed in the substrate products (ROH, ROR, RNH_2). The actual nucleophilic reagents are the hydroxide and the alkoxide ions and the neutral water and ammonia molecules. The product halides are the ionic species X^-. During the reaction the nucleophile uses its unshared electrons for bond formation with an electron-deficient carbon and the leaving group X detaches itself with the electron pair of the broken bond. By use of self-explanatory symbols[1] the above reactions can be rewritten as follows:



$$X{:}R + {:}OH^- \longrightarrow X{:}^- + R{:}OH$$

$$X{:}R + {:}OH_2 \longrightarrow X{:}^- + R{:}OH + H^+$$

$$X{:}R + {:}OR^- \longrightarrow X{:}^- + R{:}OR$$

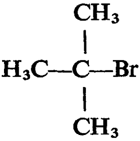
$$X{:}R + {:}NH_3 \longrightarrow X{:}^- + R{:}NH_2 + H^+$$

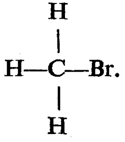
A deeper understanding of the mechanism of these reactions can only be attained by recourse to such experimental techniques as chemical kinetics, structural and medium effects, and to stereochemistry. These studies show that any of the above 'general' reactions can change their character from one compound to another, some of them being very fast while others hardly occur at all.

Let us continue to consider the reaction of alkyl halides. The reaction rates may be found to increase on changing the substrate according to the following order, tertiary > secondary > primary carbon, whereas under other conditions the reverse order is found. To give an idea of how large these effects can be, we may compare the hydrolysis of alkyl bromides in formic acid with the substitution in the same substrates by the iodide ion in acetone.

Reaction:

$R{-}Br + H_2O \longrightarrow R{-}OH + H^+ + Br^-$ (in formic acid)





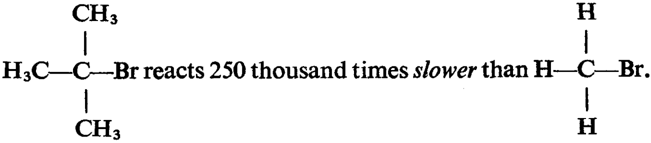
$(CH_3)_3C{-}Br$ reacts 100 million times *faster* than $H_3C{-}Br$.

---

[1] The curved arrows denote the shift of electron pairs during a reaction.

Reaction:

$R—Br+I^- \longrightarrow R—I+Br^-$ (in acetone)



$H_3C—\overset{CH_3}{\underset{CH_3}{C}}—Br$ reacts 250 thousand times *slower* than $H—\overset{H}{\underset{H}{C}}—Br$.

For some halides hydrolysis does not occur readily in water (neutral solution) but on addition of the hydroxide ion, which is a more powerful nucleophilic reagent than water, a reaction occurs at a rate proportional to the concentration of the ion. Thus, ethyl chloride hydrolyses in aqueous solution only in the presence of hydroxides:

$$CH_3CH_2Cl+OH^- \longrightarrow CH_3CH_2OH+Cl^- \qquad 1$$

whereas t-butyl chloride does so readily in aqueous solution, and added hydroxides do not change the rate:

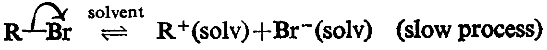
$$(CH_3)_3CCl+H_2O \longrightarrow (CH_3)_3COH+Cl^-+H^+. \qquad 2$$

This kinetic difference allows a distinction between reactions **1** and **2**, although a similar product (an alcohol) is formed. Hydrolysis such as **2** is an example of the special type of substitution whereby the solvent acts as the nucleophile. These are referred to by the general name of solvolyses.

The above-mentioned differences in behaviour are only a few of the outstanding facts concerning nucleophilic substitutions at saturated carbon atoms. They have been interpreted as showing that these reactions can occur by different mechanisms. In fact two essentially distinct mechanisms have been established by kinetic methods, the one involving a two-step process and the other a single-step process. Both the structure of the substrate and the nature of the solvent contribute to an equally large extent to these mechanisms.

The two-step mechanism

The two-step mechanism is favoured by solvents which promote the ionization of the halide:



$$R—Br \overset{\text{solvent}}{\rightleftharpoons} R^+(solv)+Br^-(solv) \quad \text{(slow process)}$$

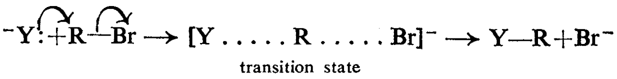
by their ability to solvate the ions thus produced. Solvents differ largely in their ionizing power, water and formic acid being among the strongest, followed by the lower alcohols, acetic acid, acetone and, finally, by the non-ionizing hydrocarbons, to mention but a few. In the above ionization a carbonium ion is produced which is usually a very reactive species, capable of capturing even relatively weak nucleophiles at high speed:

$$R^+ + \text{nucleophile} \longrightarrow \text{product} \quad \text{(fast process).}$$

If the solvent can act as a nucleophile to yield insoluble products, the over-all process is a solvolysis. Typical solvolytic solvents are water and the alcohols yielding, as final products, alcohols and ethers, respectively. If strong nucleophiles are present in the solutions such as OH^- ions, they may effectively compete for the solvent. In both cases the reaction rate remains unchanged, that is equal to the rate of the slow step – the ionization. As a consequence, the reaction rate will be dependent only on the concentration of the organic halide and the reaction will appear to be first order. This mechanism can be briefly indicated by the symbol S_N1 (substitution, nucleophilic, unimolecular).

The one-step mechanism

The one-step mechanism can be formulated as follows:



$$^-Y\!: + R—Br \longrightarrow \underset{\text{transition state}}{[Y\,.....\,R\,.....\,Br]^-} \longrightarrow Y—R + Br^-$$

Y^- being the nucleophile. Unlike the S_N1 mechanism it involves the primary attack of the nucleophile on carbon (provided that the solvent does not ionize the R—Br bond), and a transition state in which Y—R bond making has made substantial progress whilst R—Br bond breaking is still under way. In this mechanism reaction rates are sensitive to the nucleophilic power of the attacking reagents and, in fact, are proportional to their concentrations. Proper experimental conditions can be found whereby the kinetics of the reaction are second order. This order corresponds to the bimolecular nature typical of the mechanism which is indicated by the symbol S_N2 (substitution, nucleophilic, bimolecular). The S_N2 mechanism is more likely to occur in the less ionizing solvents (for example dry acetone). Dependence of the rates on the nature of a nucleophile is shown in Table 4.

Table 4

Influence of Nucleophiles on the Second-Order Constants for the Reactions of β-Propiolactone

Nucleophile	k_2(l mol^{-1} s^{-1})	k_2/k_0[1]
H_2O	$0{\cdot}99 \times 10^{-6}$	1
CH_3COO^-	$3{\cdot}0 \times 10^{-4}$	303
Cl^-	$1{\cdot}8 \times 10^{-4}$	182
Br^-	$6{\cdot}0 \times 10^{-4}$	606
I^-	$3{\cdot}0 \times 10^{-3}$	3030
SCN^-	$3{\cdot}8 \times 10^{-3}$	3830
$S_2O_3^{2-}$	$1{\cdot}9 \times 10^{-1}$	191 000
OH^-	1·2	1 210 000

Some specific structural requirements must be fulfilled if reactions (S_N1 or S_N2) are to proceed at a reasonable speed. Under a number of experimental conditions, such as solvents of intermediate ionizing power, both mechanisms may be possible and variations in the structure of the substrate are then sufficient to allow the reaction to switch from one mechanism to another. The ionization mechanism requires the formation of a carbonium ion of some stability, which can be attained by feeding the electron-deficient carbon with electronic charge. A change in this sense takes place along the alkyl series from methyl to *t*-butyl cations because of the electron-releasing ability of methyl groups (inductive and hyperconjugative effects).

$$\underset{\text{increasing stability of carbonium ions}}{\xrightarrow{H_3C^+ \quad MeH_2C^+ \quad Me_2CH^+ \quad Me_3C^+}}$$

This order of increasing stability of the carbonium ions is the same as that to be expected for a substitution proceeding by the S_N1 mechanism. The steric requirements of such a process are small since the carbonium ion has three coplanar residual covalent bonds (sp^2 hybridization). In contrast to this, the steric requirements are expected to be very important for the S_N2

---

[1] k_2 and k_0 are second-order constants; k_2 for a particular nucleophile and k_0 for water.

mechanism, for it involves a penta-coordinated carbon atom in the transition state. The overcrowding around the carbon atom undergoing substitution tends to hinder the approach of the nucleophile, so the reactivity is expected to slow down from methyl to t-butyl halides.

We are now able to interpret the experimental facts mentioned above. The hydrolysis of alkyl bromides in formic acid is a typical example of an S_N1 reaction, as it takes place in a highly ionizing solvent and speeds up steadily from methyl to t-butyl. The displacement by the iodide ion in acetone is, on the other hand, a typical example of an S_N2 reaction as it takes place in a much less ionizing solvent, and slows down steadily from methyl to t-butyl. Finally, the hydrolysis of ethyl chloride in aqueous solution clearly follows a different course (S_N2) from that of t-butyl chloride (S_N1) in the same solvent, showing that structure alone may sometimes be mechanism-determining.

From the synthetic point of view, it is interesting to note that solvolysing ethyl chloride in t-butanol is not equivalent to solvolysing t-butyl chloride in ethanol to make ethyl t-butyl ether, as it is only in the latter case that the reaction is conveniently fast for preparative work.

The same principles used to interpret the reactivity of haloalkanes have been extended to other types of alkyl halides, such as allyl, benzyl and benzhydryl halides. A great body of experimental knowledge involving kinetic form and solvent and structural effects is available, consistent with the main mechanisms outlined above. We do not need to treat the subject in more detail here except to show that further insight into the substitution reactions at saturated carbon atoms may be obtained by consideration of stereochemical aspects, which in fact offer one of the best pieces of evidence for the S_N2 mechanism.

Substitution and optical activity

When substitutions are carried out on asymmetric saturated carbon atoms under conditions favouring the S_N2 mechanism, the change in optical activity of the reacting substrate can be shown to involve 'inversion' of the carbon atom undergoing displacement. Let us consider, as a typical example, the series of reactions shown in Figure 3.

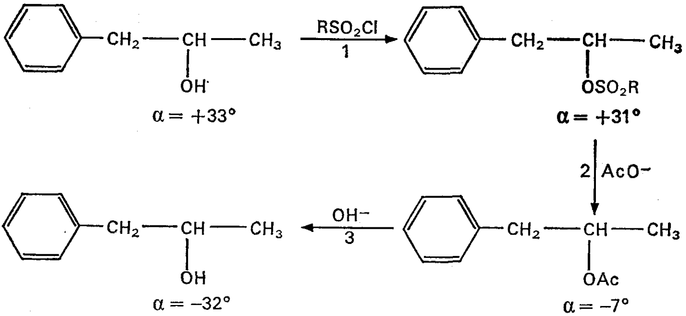


Figure 3 Substitution and optical activity

It is not possible to compare the optical activity of stereoisomers belonging to the same enantiomorphic form unless stereochemical relations have been independently established. In the above scheme what can certainly be said is that starting from a dextrorotatory alcohol we end up with its laevorotatory form. This implies that an inversion act has occurred in one of the three reactions of the series. Reactions 1 and 3 can be excluded as they are known to involve displacement at the oxygen atom rather than at the asymmetric carbon. Inversion must therefore have occurred with reaction 2. Similar conclusions have been reached in all cases where an S_N2 reaction occurs and have been expressed explicitly by Hughes and Ingold in the following statement: 'bimolecular substitutions (S_N2) are invariably accompanied by steric inversion'. This is a generally valid rule and suggests the complete geometry of the process, as shown in Figure 4.

Unlike S_N2 reactions, substitutions proceeding by the S_N1 mechanism are not stereospecific.[1] Nucleophilic attack on a flat carbonium ion can take place on either side of the plane containing it with the result that both enantiomorphs are formed. For example, the hydrolysis of α-phenylethyl chloride yields an extensively racemized alcohol.

Alcohols undergo substitution reactions involving carbonium ions by an acid-catalysed mechanism. A carbon–oxygen bond does not ionize quite as easily as a carbon–halogen bond. In the

[1] A *stereospecific* reaction involves the formation of only *one* stereoisomer rather than a *mixture* of isomers.

presence of acids, the weak oxygen base is protonated and gives rise to the following sequence:

$$R{-}OH + H^+ \rightleftharpoons R{-}OH_2^+ \text{ (rapidly formed equilibrium)}$$
$$R{-}OH_2^+ \longrightarrow R^+ + H_2O \text{ (slow ionization)}$$
$$R^+ + X^- \longrightarrow R{-}X \text{ (rapid),}$$

where the ionization in the second step is favoured by the electron-attracting oxonium group. The reactivity of alcohols in these

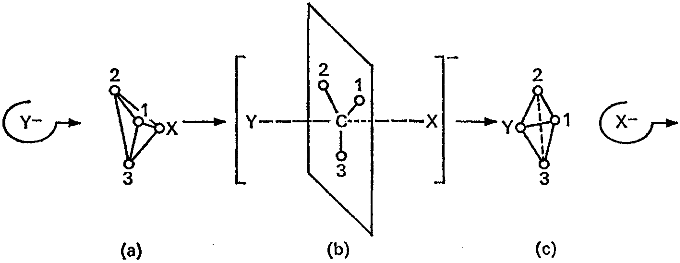


Figure 4 (a) Direction of attack of nucleophile on tetrahedral carbon (sp^3 hybridization); (b) arrangement of the system in the transition state (trigonal carbon, sp^2 hybridization); (c) products with inverted carbon

reactions follows, as expected, the order of stability of the intermediate carbonium ions:

allyl, benzyl > tertiary > secondary > primary.

In the above we have presented a general picture of the main, sharply distinct mechanisms of substitution at saturated carbon. It is, of necessity, an oversimplified picture as it does not include such aspects as those involving retention of configuration, neighbouring group participation and all the still much debated borderline mechanisms where intermediate situations control the reaction. We must give, however, a very brief reference to two types of reactions of alkyl halides and related compounds as they often accompany substitution, that is the elimination reactions and molecular rearrangements.

Elimination reactions

Like the substitutions described above, elimination reactions involve nucleophilic reagents. The substrate is attacked in a different way and is converted to different products. The two reactions can be schematically compared as shown in Figure 5.

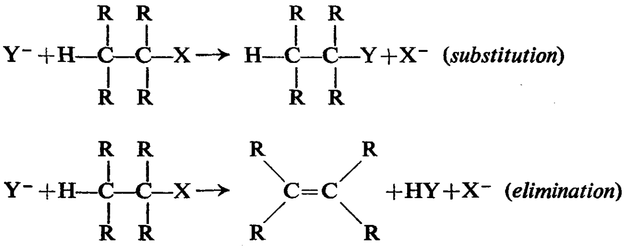
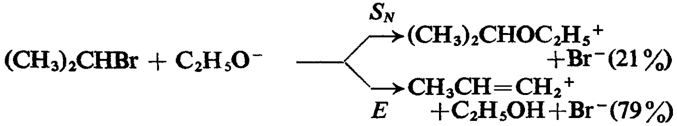


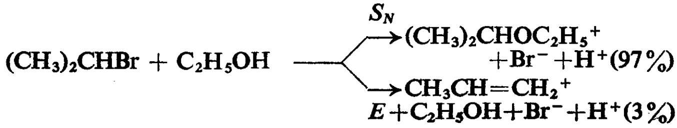
Figure 5 Comparison of substitution and elimination reactions

These two reactions closely resemble each other in many respects. Thus elimination may occur by two main mechanisms, a stereospecific, bimolecular, one-step mechanism (*E*2) and a two-step mechanism (*E*1) involving carbonium ion formation. The structural and stereochemical requirements for the elimination reaction differ appreciably from those of substitution, so there are regions of experimental conditions for the two reactions which overlap each other only in part. The result is that there are cases where either substitution or elimination occurs and cases where one reaction effectively competes with the other. This is shown in the following examples of dehydrohalogenation.

(a) Effect of the reagent



$$(CH_3)_2CHBr + C_2H_5O^- \longrightarrow \begin{cases} \xrightarrow{S_N} (CH_3)_2CHOC_2H_5 + Br^- \ (21\%) \\ \xrightarrow[E]{} CH_3CH{=}CH_2 + C_2H_5OH + Br^- \ (79\%) \end{cases}$$



$$(CH_3)_2CHBr + C_2H_5OH \longrightarrow \begin{cases} \xrightarrow{S_N} (CH_3)_2CHOC_2H_5 + Br^- + H^+ \ (97\%) \\ \xrightarrow[E]{} CH_3CH{=}CH_2 + C_2H_5OH + Br^- + H^+ \ (3\%) \end{cases}$$

(b) Effect of the structure of the substrate

$$CH_3CH_2Br + C_2H_5O^- \begin{cases} \xrightarrow{S_N} CH_3CH_2OC_2H_5 + Br^-(90\%) \\ \xrightarrow[E]{} CH_2{=}CH_2 + C_2H_5OH + Br^-(10\%) \end{cases}$$

$$(CH_3)_3CBr + C_2H_5O^- \begin{cases} \xrightarrow{S_N} \text{no reaction} \\ \xrightarrow[E]{} (CH_3)_2C{=}CH_2 + C_2H_5OH^+ + Br^-(100\%) \end{cases}$$

Example (b) shows that branching is more favourable to elimination than to substitution. This is a typical relation between the reactions of the bimolecular type (S_N2 and $E2$). It also shows that mixing t-butyl bromide with sodium ethoxide is not equivalent to mixing ethyl bromide with t-butoxide although, if substitution occurred in both cases, they would be expected to yield the same ether. In fact, the ether is obtained only by using the latter mixing procedure.

On the basis of the established reaction mechanism, the above examples and many other similar facts have been satisfactorily interpreted. Elimination of water from alcohols (dehydration) is an important reaction of this group. It is an acid-catalysed reaction and is favoured by structures producing stable carbonium ions as in the case of *a*-phenylethanol (Figure 6).

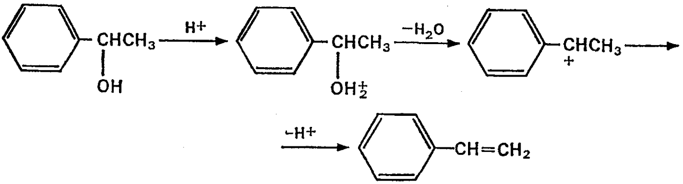


Figure 6 Elimination of water from *a*-phenylethanol

Molecular rearrangements

We finally turn to molecular rearrangements. It has been noted that the ease with which reactions passing through carbonium ion intermediates proceed depends on the stability of the ion; t-butyl, benzyl and allyl carbonium ions are classical examples in this connexion, the stability of the two latter being due largely to resonance (Figure 7).

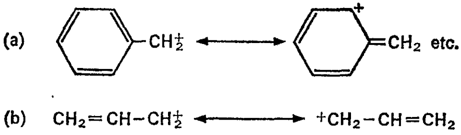
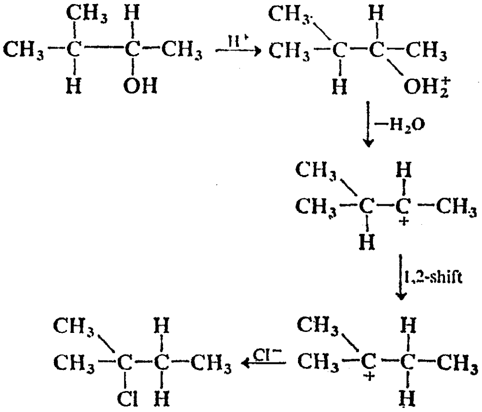


Figure 7 Resonance forms of (a) benzyl and (b) allyl cations

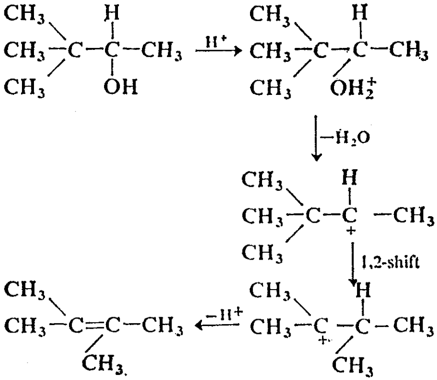
In some cases an unstable carbonium ion rearranges to a more stable one by a 1,2-shift of a migrating group (hydrogen, alkyl or other groups). The reaction will end up with a substitution or an elimination product corresponding to the rearranged ion as the examples below show.

Substitution with rearrangement



2-chloro-2-methylbutane
(and no 3-chloro-2-methylbutane)

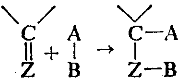
Elimination with rearrangement



This brief description should not lead to the belief that the substituted alkanes are sensitive solely to nucleophilic reagents. Many organic reactions, notably the ones involving the organometallic compounds, can be clearly classified as electrophilic substitutions. Other reactions, such as the important photochemical halogenation of hydrocarbons, are free-radical substitutions.

Some Addition Reactions to Unsaturated Carbon Atoms

The high covalence of carbon, which makes the existence of a large number of double- and triple-bonded compounds possible, is one of the major characteristics of organic chemistry. A double bond, made up of a strong σ-bond and a weaker π-bond, forms a point of attack in the molecule for a variety of reagents, for example:



(Z = R_2C, O,etc.)

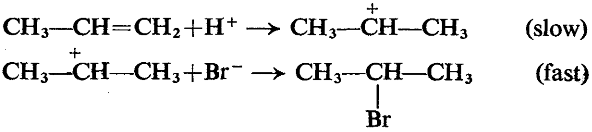
Therefore, addition, like substitution, is a most important organic reaction. Again, such a simple and general scheme as the above is of little value in the understanding of specific reactions.

The course of addition to a great extent depends on the nature of the reagent and the structure of the substrate.

Let us first consider the additions to alkenes and then carbonyl additions. Many reactions of alkenes are electrophilic additions, as in the following examples.

1. Hydrogen halide addition
   $CH_2{=}CH_2 + HBr \longrightarrow CH_3{-}CH_2Br$ (in acetic acid).
2. Halogen addition $CH_2{=}CH_2 + Br_2 \longrightarrow CH_2Br{-}CH_2Br$.
3. Hydration
   $CH_2{=}CH_2 + H_2O \longrightarrow CH_3{-}CH_2OH$
   (in the presence of H_2SO_4).
4. Hypohalous acid addition $CH_2{=}CH_2 + ClOH$
   $\longrightarrow CH_2Cl{-}CH_2OH$.

The π-electrons in the double bond are accessible to electrophilic attack by a proton (reactions 1 and 3), by a cationic halogen (reactions 2 and 4) or other reagents. These reactions usually proceed by a two-stage mechanism through a carbonium ion intermediate. Hydrogen halide addition can be represented as follows:



$$CH_3{-}CH{=}CH_2 + H^+ \longrightarrow CH_3{-}\overset{+}{C}H{-}CH_3 \qquad \text{(slow)}$$
$$CH_3{-}\overset{+}{C}H{-}CH_3 + Br^- \longrightarrow CH_3{-}\underset{\displaystyle Br}{\underset{|}{C}}H{-}CH_3 \qquad \text{(fast)}$$

The electrophilic nature of the initial attack is shown by the rate-increasing effect of electron-donating substituents such as alkyl groups, the highly branched alkenes, $R_2C{=}CR_2$, being very reactive. In hydrogen halide addition, the rate-determining formation of a carbonium ion is shown by the special orientation of the added molecule in the product. For example, isopropyl, rather than n-propyl, bromide is formed exclusively from propylene. This product corresponds to the more stable of the two carbonium ions possible formed by the proton attack, that is:

$$CH_3{-}\overset{+}{C}H{-}CH_3 \text{ more stable than } CH_3{-}CH_2{-}CH_2^+.$$

Similarly t-butyl bromide is formed exclusively from isobutylene:

$$(CH_3)_2C{=}CH_2$$

for $(CH_3)_2\overset{+}{C}{-}CH_3$ is more stable than $(CH_3)_2CH{-}CH_2^+$.

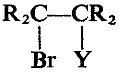
More generally, the electrophilic additions proceed in the direction of the expected order of carbonium ion stabilities, as already observed in the previous section for the S_N1 reactions:

$$\text{tertiary} > \text{secondary} > \text{primary}.$$

The above picture offers an interpretation of the long known Markownikoff's rule, whereby in hydrogen halide addition the hydrogen becomes attached to the carbon atom which holds the larger number of hydrogen atoms. The validity of this rule is linked with the above mechanism of addition. Thus, in the peroxide-initiated addition of hydrogen bromide, the orientation of addition is found to be exactly reversed. In this case, a bromine atom acting as a free radical is produced in the presence of peroxides and a reaction chain is thus propagated:

$$R_2C{=}CR_2 + Br\cdot \longrightarrow R_2\dot{C}{-}CR_2Br$$
$$R_2\dot{C}{-}CR_2Br + HBr \longrightarrow R_2CH{-}CR_2Br + Br\cdot \text{ etc.}$$

In halogen addition, the electrophile is the positive end of a polarized halogen molecule $\overset{\delta+}{X}{-}\overset{\delta-}{X}$. Polarization is induced by the surroundings and by the electron-donor properties of the double bond. In this reaction the electrophilic nature of the attack and the formation of a carbonium ion in the intermediate step has been elegantly illustrated by running the addition by a given halogen, say Br_2, in the presence of a number of nucleophiles, such as Br^-, Cl^-, I^-, H_2O, etc. The product is found to be of the type



$$\underset{Br}{R_2C}{-}\underset{Y}{CR_2}$$

Y being the nucleophile.

Halogen addition is a peculiar reaction in that it is largely stereospecific. Once again stereochemistry is a powerful tool for the elucidation of reaction mechanisms. The basic idea in this connexion has been to examine the optical activity of addition products where asymmetric carbons are produced. Without going into the details of these studies, the results indicate that in bromine addition the intermediate ion is a bromonium, rather than a true carbonium, ion of the approximate structure shown in Figure 8, characterized by a restricted rotation around the σ-bond. Bromide ion would attack from a direction (below the plane) opposite to that (above the plane) of the bromonium ion by a path resembling the inversion encountered in the S_N2 mechanism.

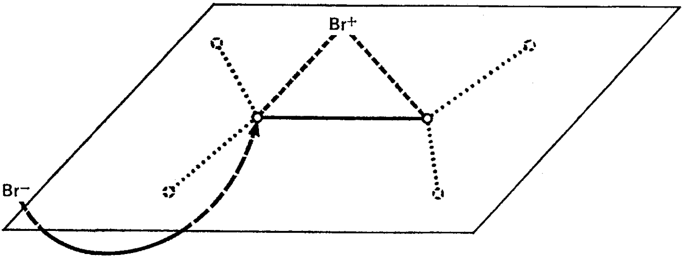


Figure 8 Stereospecificity of bromine addition

Alkenes are much less susceptible to nucleophilic than they are to electrophilic attack except in cases where the C=C double bond is conjugated with electron-withdrawing, unsaturated groups such as NO_2, COR, CN. Nucleophilic addition is, however, the normal type of reaction displayed by C=O double bonds as in aldehydes and ketones. Some examples are reported in the following:

1. Addition of water
   $CH_3CHO+H_2O \rightleftharpoons CH_3CH(OH)_2$.
2. Hemiacetal and acetal formation
   $CH_3CHO+C_2H_5OH \rightleftharpoons CH_3CH(OH)OC_2H_5$ (addition).
   $CH_3CH(OH)OC_2H_5+C_2H_5OH \longrightarrow CH_3CH(OC_2H_5)_2$
   $+H_2O$ (substitution).
3. Addition of bisulfite
   $CH_3CHO+NaHSO_3 \longrightarrow CH_3CH(OH)SO_3Na$.
4. Cyanohydrin formation $CH_3CHO+HCN$
   $\longrightarrow CH_3CH(OH)CN$.

5. Semicarbazone formation
$CH_3CHO+H_2NNHCONH_2$
$\rightarrow CH_3CH(OH)NHNHCONH_2$ (addition).
$CH_3CH(OH)NHNHCONH_2$
$\rightarrow CH_3CH{=}NNHCONH_2+H_2O$ (dehydration).
6. Aldol condensation $CH_3CHO+CH_3CHO$
$\rightleftharpoons CH_3CH(OH)CH_2CHO$.
7. Addition of Grignard reagents
$CH_3CHO+C_6H_5MgBr \rightarrow CH_3CH(C_6H_5)OMgBr$
(addition).
$CH_3CH(C_6H_5)OMgBr+H_2O$
$\rightarrow CH_3CH(OH)C_6H_5+Mg(OH)Br$ (hydrolysis).
8. Cannizzaro reaction
$C_6H_5CHO+C_6H_5CHO + KOH$
$\rightarrow C_6H_5CO_2K+C_6H_5CH_2OH$.

These examples show that many of these reactions are reversible and involve a great variety of nucleophiles leading to carbon–oxygen (reactions 1 and 2), carbon–sulphur (reaction 3), carbon–nitrogen (reaction 5), carbon–carbon (reactions 4, 6 and 7) and carbon–hydrogen (reaction 8) bond formation.

The reactivity of carbonyl carbon toward nucleophilic reagents is explained by the unequal sharing of π-electrons as illustrated by the following resonance structures:

$$RR'C{=}O \longleftrightarrow RR'\overset{+}{C}{-}O^-.$$

This reactivity will be increased by the presence of acids because of the rapidly established equilibrium:

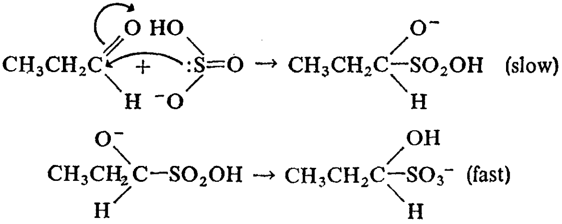
$$RR'CO+H^+ \rightleftharpoons RR'COH^+$$

and of the more important resonance effect in the conjugated acid:

$$RR'CO^+H \longleftrightarrow RR'\overset{+}{C}{-}OH$$

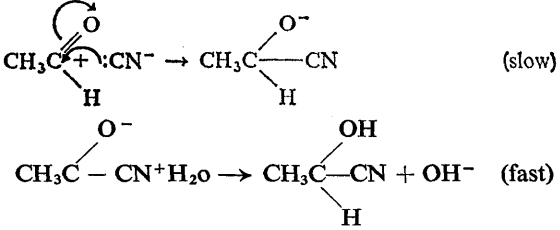
where the contribution of $RR'\overset{+}{C}{-}OH$ is greater than that of $RR'\overset{+}{C}{-}O^-$ in the neutral carbonyl compound. As a result, additions can be acid-catalysed. Base-catalysis is also possible whenever a proton can be abstracted from the reagent to produce a stronger nucleophile.

Usually carbonyl additions are two-step reactions, the first step being the rate-determining nucleophilic attack followed by a rapid neutralization of the anionic charge by a proton or another cation. For example, the addition of bisulfite to propionaldehyde proceeds as follows.



The addition of cyanide by HCN (cyanohydrin formation) requires the presence of traces of strong bases to favour the production of the nucleophile.

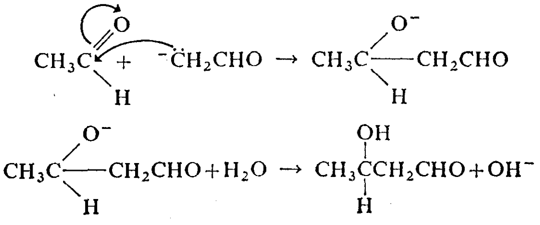
$HCN + OH^- \rightleftharpoons CN^- + H_2O$ (pre-equilibrium)



In the aldol condensation nucleophilic attack of an aldehyde (or ketone) on aldehyde (or ketone) occurs. The reaction is base-catalysed and involves a pre-equilibrium where the nucleophile is produced:

$$CH_3CHO + OH^- \rightleftharpoons {}^-CH_2CHO + H_2O.$$

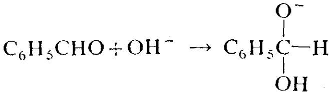
This reaction reveals the acidic character of a C—H bond adjacent to carbonyl (*α*-hydrogen). The next steps involve the addition proper and the neutralization of the anionic charge.



When no *a*-hydrogen is present as in the cases of benzaldehyde and trimethylacetaldehyde, the aldol condensation does not occur. Instead, strong bases catalyse a different reaction involving a simultaneous oxidation of one aldehyde group and reduction of another (Cannizzaro reaction). This reaction is rather more complicated than most additions as is shown by its high-order kinetics:

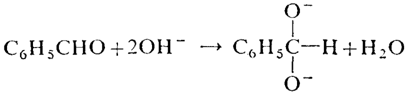
$$\text{rate} \propto [\text{aldehyde}]^2\,[\text{OH}^-] \text{ or rate} \propto [\text{aldehyde}]^2\,[\text{OH}^-]^2.$$

A well-established feature is the hydride ion (H^-) attack on carbonyl carbon. The rate-dependence on hydroxide concentration is related to the formation of the intermediate species I and II whose relative importance depends on the particular reacting system.



$$C_6H_5CHO + OH^- \rightarrow C_6H_5\overset{O^-}{\underset{OH}{C}}-H$$

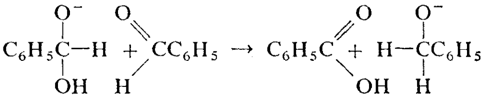
(I)



$$C_6H_5CHO + 2OH^- \rightarrow C_6H_5\overset{O^-}{\underset{O^-}{C}}-H + H_2O$$

(II)

Then the hydride transfer follows, giving in the case of intermediate species (I):



and $C_6H_5CH_2O^- + H_2O \rightleftharpoons C_6H_5CH_2OH + OH^-$

The bimolecular nature of the hydride transfer reaction, with no participation of the medium, has been elegantly proved by running the reaction in heavy water solution and showing that no deuterium becomes attached to carbon in the products.

Mechanism and Orientation in Aromatic Substitutions

Benzene and its derivatives occupy a unique position among organic compounds because of the great stability of the aromatic ring under a wide variety of experimental conditions. For this reason, the majority of aromatic reactions do not destroy 'aromaticity' and are peripheral substitutions. Substitutions of the three main types are possible, but those of the electrophilic type are the most common and also the most impressive ones on account of their spectacular phenomena of orientation which will be illustrated later in this section.

Electrophilic substitutions are capable of converting benzene into a number of derivatives of the general type C_6H_5X. Ignoring experimental details and catalysts, a few examples are given below in the form of stoichiometric equations for a general aromatic compound ArH, where Ar is used to denote any aryl group.

1. Nitration $ArH+HNO_3 \longrightarrow ArNO_2+H_2O$.
2. Chlorination $ArH+Cl_2 \longrightarrow ArCl+HCl$.
3. Bromination $ArH+Br_2 \longrightarrow ArBr+HBr$.
4. Alkylation $ArH+CH_3Cl \longrightarrow ArCH_3+HCl$.
5. Acetylation $ArH+CH_3COCl \longrightarrow ArCOCH_3+HCl$.
6. Benzoylation $ArH+C_6H_5COCl \longrightarrow ArCOC_6H_5+HCl$.
7. Mercuration $ArH+Hg(CH_3CO_2)_2 \longrightarrow ArHgOCOCH_3+CH_3CO_2H$.
8. Sulfonation $ArH+H_2SO_4 \longrightarrow ArSO_3H+H_2O$.
9. Exchange $ArH+D_3O^+ \longrightarrow ArD+HD_2O^+$.

The scope of these reactions is greatly extended by inclusion of substituted and polysubstituted benzenes, some of which react much more smoothly than benzene itself. Benzene reacts in most cases at such a slow rate that the reaction is impractical for preparative purposes or even for measurement. Thus, unless suitable catalysts are available, the reaction is known to occur only in some of the most reactive derivatives. This is, for instance, the case for azo-coupling which only occurs with the most reactive aromatics:

$$C_6H_5N(CH_3)_2+C_6H_5N_2Cl \longrightarrow p\text{-}C_6H_5N{=}NC_6H_4N(CH_3)_2+HCl.$$

Once a given substituent has become attached to the benzene ring, many further derivatives can be obtained by successive, not necessarily substitutive, reactions. Thus, aniline and phenol can be obtained from nitrobenzene, organometallic intermediates from bromobenzene, hydrocarbons from ketones, etc.

One of the main problems concerning reaction mechanisms in aromatic substitution is the identification of the exact nature of the electrophilic entity which is produced in the solution; this may vary depending on the medium and catalyst. The electrophilic nature of the reaction is, however, established beyond doubt by the expulsion of a proton:

$$\mathrm{Ar}\overset{\curvearrowright}{:\mathrm{H}+}\underset{\textit{electrophile}}{\mathrm{X}^+} \longrightarrow \mathrm{H}^+ + \mathrm{Ar}:\mathrm{X}$$

Aromatic substrates do not show any tendency to self-ionization, so aromatic substitutions result from the bimolecular attack of the electrophile on the substrates. Indeed, these reactions depend very much on the nature of the reagent. This point is dramatically illustrated by the difference in reactivity of different chlorinating or brominating agents. Bromination, for example, may involve either of two electrophiles, positive bromine (Br^+, $BrOH_2^+$) or molecular bromine (Br_2). These two species show great difference in reactivity, the former being much more reactive than the latter. Moreover, not only does a given substrate react at different rates with these reagents but it gives a specific response in orientation (see further on). Thus, brominations by Br^+ and Br_2 are to be considered as two distinct reactions.

Lewis acid catalysts act as sources for positive halogen. They have preparative significance for such unreactive substrates as nitrobenzene:

$$C_6H_5NO_2 \xrightarrow{(Br_2,\ AlBr_3)} m\text{-}C_6H_4BrNO_2 \quad \text{(96 per cent).}$$

The complex ($Br_2, AlBr_3$) is an ionic compound of the type $Br^+[AlBr_4]^-$.

Hypobromous acid is another well-established source of positive bromine in either form, $BrOH_2^+$ or Br^+:

$$BrOH + H_3O^+ \rightleftharpoons BrOH_2^+ + H_2O$$
$$BrOH_2^+ \rightleftharpoons Br^+ + H_2O$$
$$BrOH_2^+ (\text{or } Br^+) + ArH \longrightarrow ArBr + H_3O^+.$$

The above scheme illustrates an example of acid-catalysed substitution.

Highly reactive substrates, such as aniline, react so readily with molecular bromine that they yield polysubstituted products:

$$C_6H_5NH_2 \xrightarrow{Br_2} C_6H_2Br_3NH_2 \quad \text{(100 per cent).}$$

Being bimolecular reactions, aromatic electrophilic substitutions show a first-order dependence on each reactant. However, in some cases the reaction is independent of the concentration of the substrate. This happens with nitration in organic solvents (nitromethane, acetic acid) when the substrate is sufficiently reactive (toluene). The electrophilic entity in nitration is the nitronium ion, NO_2^+. The ion is produced in the medium at such a slow rate that its production is rate-determining:

$$\begin{aligned} &2HNO_3 \rightleftharpoons H_2NO_3^+ + NO_3^- && \text{(fast)} \\ &H_2NO_3^+ \rightleftharpoons NO_2^+ + H_2O && \text{(slow)} \\ &ArH + NO_2^+ \longrightarrow ArNO_2 + H^+ && \\ & && \text{(over-all reaction, relatively fast).} \end{aligned}$$

The nitronium ion-forming equilibria are readily established in concentrated mineral acid:

In sulfuric acid:

$$HNO_3 + 2H_2SO_4 \rightleftharpoons NO_2^+ + H_3O^+ + 2HSO_4^- \quad \text{(fast)}$$

In nitric acid:

$$\begin{aligned} &2HNO_3 \rightleftharpoons NO_2^+ + NO_3^- + H_2O && \text{(fast)} \\ &ArH + NO_2^+ \longrightarrow ArNO_2 + H^+ && \\ & && \text{(over-all reaction, relatively slow).} \end{aligned}$$

In these latter cases the rate depends on the concentration of the substrate. Molecular nitric acid, HNO_3, has very little nitrating power.

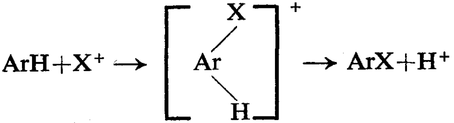
Alkylation, acetylation and benzoylation are reactions of a group known under the general name of Friedel–Crafts substitutions. It must be noted that since in these reactions a carbon–carbon bond is formed they can also be classified as nucleophilic substitutions, provided that the alkyl or acyl compounds are considered as the substrate and the aromatic hydrocarbon as the nucleophile. However, they are actually electrophilic reactions if considered from the point of view of the attack on aromatic carbon. Friedel–Crafts reactions do not occur unless a Lewis acid catalyst is used to effect the production of the electrophilic reagent in the form of a carbonium ion:

alkylation: $CH_3Cl + AlCl_3 \longrightarrow CH_3^+AlCl_4^-$.
acetylation: $CH_3COCl + AlCl_3 \longrightarrow CH_3CO^+AlCl_4^-$.

In the case of particularly stable carbonium ion intermediates the milder Brønsted acids can be used as catalysts:

t-butylation: $(CH_3)_3COH + H^+ \rightleftharpoons (CH_3)_3C^+ + H_2O$.

There is still much debate on the mechanism of attack on the aromatic system. The most likely paths are shown in Figure 9. The intermediate-complex mechanisms:



σ-complex

are likely to occur at least in some cases. Thus the fact that some (but not all) aromatic substitutions show no kinetic isotope effect indicates that the rate-determining step involves no C—H bond breaking as would be expected were the σ-complex formation the rate-determining step.

Finally we may examine the orientation phenomena in aromatic substitutions. For the best appreciation of the well-known orienting power of substituents in the benzene ring, a combination of two main factors should be considered: (1) the specific electronic effects of the substituents and (2) the kinetic factor. The former factor arises from the consideration that, had a given substituent no specific orientation power, the reaction product at all times should be a mixture of isomers corresponding to the statistical distribution:

$$C_6H_5X \xrightarrow{Y^+} o\text{-}C_6H_4XY\ (40\ \text{per cent}) + m\text{-}C_6H_4XY\ (40\ \text{per cent}) + p\text{-}C_6H_4XY\ (20\ \text{per cent}).$$

Products of this composition are never found experimentally. For example, the above reported bromination of nitrobenzene yields the *m*-isomer only, whereas aniline tribrominates with Br_2 with no detectable introduction of bromine into a position *meta* to the amino group. The second factor arises from the observation that benzene derivatives of the type C_6H_5X do not react at the same rate as benzene itself but at a rate that is to a great extent related to the orienting power. Usually, *o*–*p*-orienting substituents have an activating effect whereas *m*-orienting substituents behave in the opposite way. Among the former, we note CH_3, OH, OCH_3, NH_2, $N(CH_3)_2$; among the latter, NO_2, $COCH_3$, SO_3^-, NR_3^+. There are exceptions, however, notably the halogens which are *o*–*p*-orienting but, in nearly all cases, deactivating substituents.

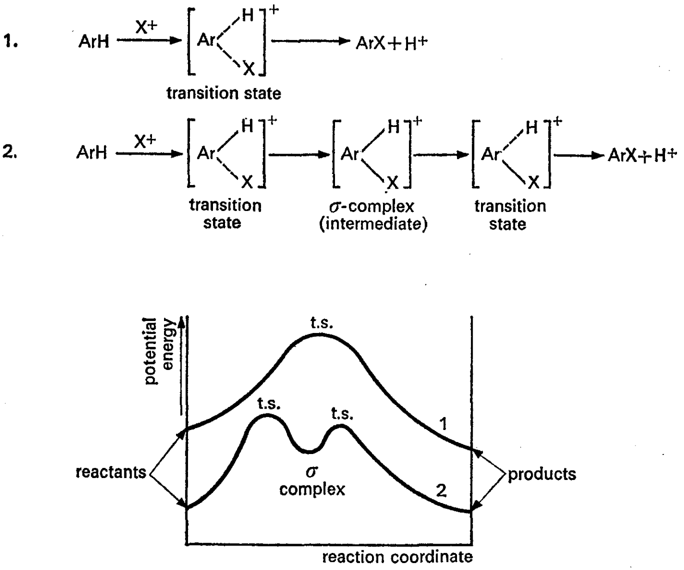


Figure 9 Mechanism of aromatic substitution reactions

Table 5 gives orientation data for the nitration of some substituted benzenes. In order to account for the kinetic factor also, the over-all nitration rates relative to benzene can be combined

Table 5

Isomer Distribution in the Nitration of Monosubstituted Benzenes, C_6H_5X

X	OH	CH_3	Cl	NO_2	$^+N(CH_3)_3$
per cent *o*-	40·0	40·9	29·6	6·4	0
per cent *m*-	0	3·1	0·9	93·2	100
per cent *p*-	60·0	56·0	69·5	0·3	0

with the isomer distribution percentages to yield new, very useful parameters – the partial rate factors, *f*. These tell how many times individual positions of a compound C_6H_5X are activated or deactivated with respect to a single position of benzene. Some of these factors are reported in Table 6.

Table 6

Partial Rate Factors for the Nitration of some Monosubstituted Benzenes, C_6H_5X

X	CH_3	Cl	$CO_2C_2H_5$
f_o	42	0·03	0·003
f_m	2·5	0·001	0·008
f_p	58	0·14	0·001

It can be seen from these few data that partial rate factors fully describe the orientation phenomenon by showing the interplay of isomer distribution with the kinetic factor. It is difficult to obtain similar data for such strongly *m*-orienting substituents as NO_2, $N(CH_3)_3^+$ and for such strongly *o*–*p*-orienting substituents as OH, NH_2 because the reactions are either too slow or too fast for measurement. It will suffice to say that the above substituents often deactivate or activate the benzene ring by factors exceeding 10^6.

The structural origin of orientation lies in the perturbation produced by a given substituent on the electronic arrangement of the benzene ring by effects which are well-established from a large number of properties. These effects will not be discussed in detail here: we want only to make a few points clear in connexion with electrophilic substitutions. Aromatic substrates present a differential distribution of the electronic charge from one non-equivalent position to another; such a distribution is able to explain by itself the gross picture of orientation because the electrophilic reagent will seek the highest electron density positions for preferential attack, even in such apparently exceptional cases as the *o–p*-orientation of halogenobenzenes which occur with deactivation. Indeed, this effect simply results from the fact that the electron-releasing power of halogens through resonance, which is responsible for the *o–p*-orientation, is not quite so strong as to counterbalance the inductive effect which acts in the opposite direction by withdrawing electronic charge from the σ-bonds of the ring and effectively deactivating it. A more precise picture must take into account the intensity of the electronic effects, which in some cases depends largely on the polarization of the system in the transition state. Since an electrophile is an electron-seeking reagent, polarization is particularly important with electron-releasing substituents acting through resonance. This explains, for example, the very high activating power of such substituents as OH, OCH_3, NH_2, etc. in the *para* position (Figure 10) despite their inductive effect acting, again in these cases, in the opposite direction.

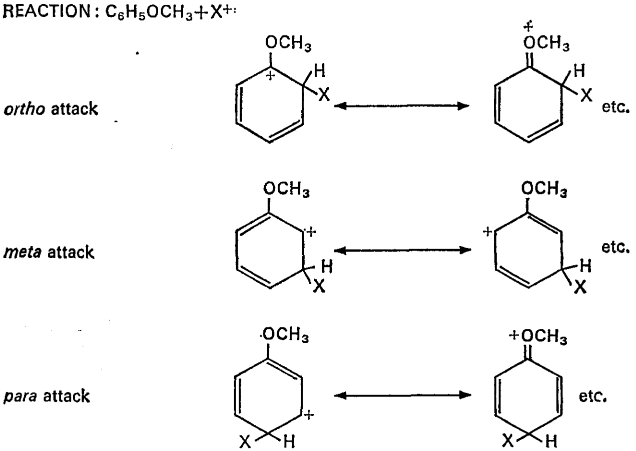


Figure 10 Some relevant resonance structures showing the relationship of orientation to the stability of the σ-complex intermediate

Such polarization effects occur in the aromatic substrate but also depend on the nature of the electrophilic reagent. Accurate experimental data are available for the methyl group which is capable of hyperconjugative electron-release as shown in Table 7.

Table 7

Partial Rate Factors for Attack by some Electrophiles on Toluene

Electrophile	Cl^+	Br^+	Br_2	NO_2^+
f_o	134	76	600	42
f_m	4·0	2·5	5·5	2·5
f_p	82	59	2420	58

The largest variations are observed at the *o*- and *p*-positions which are more sensitive to electron density changes through resonance. *Ortho*-positions deserve special comment because of field and steric effects. The latter have been better established by experiment and can be illustrated, for example, by comparing the f_o values for the chlorination of toluene and t-butylbenzene with Cl_2. The values are 617 and 72, respectively, and show that the attack in the latter case is much more hindered than in the former.

The above are only a few examples of some of the major features at issue in aromatic electrophilic substitutions. We ought to realize, however, that whole groups of aromatic reactions belong to other types which have been neglected altogether.

Finally, we must mention homolytic substitutions, such as phenylation, and nucleophilic substitutions. The reactions in the latter two types are fundamentally different from the electrophilic type because of the different nature of the reagent. Homolytic substitution still involves orientation phenomena as it takes place with the expulsion of the hydrogen atom. The orienting effects in this reaction, however, are much poorer than those observed in classical electrophilic substitutions.

Bibliography

D. J. CRAM and G. S. HAMMOND, *Organic Chemistry*, 2nd edn, McGraw-Hill, 1964.

E. S. GOULD, *Mechanism and Structure in Organic Chemistry*, Holt, Rinehart & Winston, 1959.[1]

J. HINE, *Physical Organic Chemistry*, 2nd edn, McGraw-Hill, 1962.[1]

R. STEWART, *The Investigation of Organic Reactions*, Prentice-Hall, 1966.

P. SYKES, *A Guidebook to Mechanism in Organic Chemistry*, 2nd edn, Longmans, 1965.

J. M. TEDDER and A. NECHVATAL, *Basic Organic Chemistry* (2 vols.), Wiley, 1966, 1967.

1 More advanced.

Acknowledgments

Permission to reproduce the Readings in this volume is acknowledged from the following sources:

Reading 1	Organisation for Economic Co-operation and Development
Reading 2	The American Chemical Society
Reading 3	*Chemistry Student*
Reading 4	*Chemistry Student*
Reading 5	The Royal Institute of Chemistry
Reading 6	Organisation for Economic Co-operation and Development
Reading 7	*Chemistry Student*
Reading 8	*Chemistry Student*
Reading 9	*Chemistry Student*
Reading 10	*Chemistry Student*
Reading 11	*Chemistry Student*
Reading 12	The Royal Institute of Chemistry
Reading 13	The American Chemical Society
Reading 14	The American Chemical Society
Reading 15	Organisation for Economic Co-operation and Development

Index

Advanced science
Readings

These Readings reflect the many profound changes chemistry has seen in the twentieth century, and provide clear and stimulating introductions to the most important new developments. The editor has made his selection, and provided editorial introductions, for the more advanced student at school and the student beginning at university. The major areas in which papers have been selected are atomic and molecular structure, stereochemistry, energetics and kinetics, acids and bases, the chemical elements, and organic reaction mechanisms. A number of suggestions for further readings are provided at the end of each article.

Many of the papers were not originally published in SI units. To bring the book into line with current thinking on the subject, the editor has provided a note on SI units and given the relevant conversion factors.

J. G. Stark is Chemistry Master, Merchant Taylor's School, Liverpool.

United Kingdom 50p 10/–
Australia $1.70
New Zealand $1.70
South Africa R1.20
Canada $2.15

The cover shows a model of a uranium atom in the German Museum, Munich.

Published by **Penguin Education** 14 080161 8

Advanced science
Readings
Modern chemistry
edited by J. G. Stark

14 080161 8