JOHN HENRY

A SHORT HISTORY OF SCIENTIFIC THOUGHT

A Short History of Scientific Thought

A Short History of Scientific Thought

John Henry

palgrave
macmillan

First published 2012 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries

ISBN 978-0-230-01942-3 hardback
ISBN 978-0-230-01943-0 paperback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

10 9 8 7 6 5 4 3 2 1
21 20 19 18 17 16 15 14 13 12

Transferred to Digital Printing in 2012

For my daughter, Eilidh,
with love

Contents

List of Boxes, Figures and Tables

Boxes

Figures

Tables

Acknowledgements

The author and publishers wish to thank the following for permission to reproduce copyright material:

Taylor & Francis, and the late author's wife, Mary Ann Rossi, for Tables 8.1 and 8.2, which have been simplified from J. Bruce Brackenridge, 'Kepler, Elliptical Orbits, and Celestial Circularity: A Study in the Persistence of Metaphysical Commitment', *Annals of Science*, 39 (1982), pp. 117–43 and 265–95.

The Wellcome Library, London, for Figure 5.1, drawing of the skeleton from a fourteenth-century medical manuscript, and Figure 5.3, from Johannes de Ketham, *Fasciculus medicinae* (Venice, 1495). With permission from the Wellcome Trust.

The Trustees of the National Library of Scotland, for Figure 23.2 from James Clerk Maxwell, 'On Physical Lines of Force', in *The London, Edinburgh, and Dublin Philosophical Magazine*, April 1861; J.219–221 PER.

The University of Edinburgh, Special Collections Department, for Figure 4.1, from Peter Apian, *Cosmographia* (Antwerp, 1584); Figure 5.2, from Charles Estienne, *De dissectione partium corporis* (Paris, 1545); Figures 5.4 and 5.5, from Andreas Vesalius, *De humani corporis fabrica* (Basel, 1543); Figure 6.5, from Nicholas Copernicus, *De revolutionibus orbium coelestium* (Frankfurt, 1543); Figure 7.1, from William Gilbert, *De magnete* (London, 1600); Figures 8.1 and 8.2, from Johannes Kepler, *Mysterium cosmographicum* (Tübingen, 1596); Figure 8.3, from Johannes Kepler, *Epitome astronomiae copernicanae* (Frankfurt, 1635); Figure 8.4, from Johannes Kepler, *Harmonices mundi* (Linz, 1619); Figures 11.1 and 17.1, from René Descartes, *Principia philosophiae* (Amsterdam, 1644); Figure 17.2, from James Hutton, *Theory of the Earth* (Edinburgh, 1795); Figure 19.1, from Robert Chambers, *Vestiges of the Natural History of Creation* (London, 1844); and Figure 23.1, from Thomas Young, *A Course of Lectures on Natural Philosophy and the Mechanical Arts*, 2 vols (London, 1807).

Additionally, the author is grateful to Marie-Louise Collard at Wellcome Images, James Mitchell at the National Library of Scotland, and Tricia Boyd at Edinburgh University Library's Special Collections Department, for their help with acquiring these images.

Every effort has been made to contact all the copyright-holders, but if any have been inadvertently overlooked the publishers would be pleased to make the necessary arrangements at the first opportunity.

Introduction

This book provides a brief historical survey of the major developments in scientific thought and the impact of science on Western culture from ancient times to the twentieth century. The aim is to understand the character of the study of the natural world, in each time and place, with its own successes and failures, in terms of the culture of which it was a product. However, because it is a brief survey, it necessarily omits more than it includes, and can only deal superficially with those topics it does cover.

In any attempt to provide an overview there must be a compromise between a number of conflicting factors: the general lessons to be conveyed; the complexity of the historical processes; the availability of sources; the possibility of popularizing technical work; the minimum consideration required to make any given topic meaningful; and above all, the need for brevity. Bearing these factors in mind, it is still possible to omit vast areas of the field without falsifying the enterprise altogether. I believe that what follows does provide the reader with a broad overview of the historical development of attempts to understand the nature of the physical world.

The overall shape and structure of the book derives from three general organizing principles. Firstly, the book concentrates on those topics which can easily be seen to have reflected or affected contemporary perceptions of the very nature of the physical world and of man's role within it. Apart from making the survey feasible, this general principle has two clear advantages: firstly, it ensures that science is seen as part of the broader cultural context of any given society; secondly, since we are concerned with the impact of science outside scientific circles, it avoids the necessity to consider complex technical developments.

Nobody should let ignorance or even fear of science put them off this book. The main aim is not to promote the understanding of science *per se*, but to help in understanding why and how science has become such an important cultural force in the modern world. These questions can only be understood historically. After all, science is now widely regarded as the supreme authority for pronouncing on the way things are; for pronouncing on what is truth. Nowadays, political decisions, courts of law, and even public opinion all defer to scientific authority, but it hasn't always been so. Until the second half of the last century, science always had to defer to religion as the supreme cognitive authority, and one of the things we'll be looking at in this book is how religion came to be supplanted by science. But first, we have to determine where science, or perhaps we should say the scientific approach to understanding the world, came from.

The second general organizing principle is to provide a continuous narrative of the historical development of science. One of the main aims of this book is to show the importance of the cultural context for an understanding of scientific discovery and the establishment of scientific orthodoxy and authority. Scientific knowledge is shown to be an integral part of our culture not something that transcends it; not something that is imposed upon it by supreme geniuses who somehow stand outside the social and political concerns that effect ordinary mortals. One way of demonstrating this is to show how scientific developments in one age stemmed from what had gone before, and subsequently led to developments in the next age. Charles Darwin readily drew attention to the traditional belief that *Natura non facit saltum* – Nature does not make leaps; similarly, I wish to insist that there are no jumps, no discontinuities in history. Copernicus (see Chapter 6) and Einstein (see Chapter 23) are often seen as inaugurating revolutionary breaks with the past – but this does not mean that their ideas sprung fully formed out of their own heads with no reference to anything that had gone before.

The historiography of science – that is to say, the *writing* of the history of science, *the way the history is written* by professional historians – has inevitably become increasingly specialized, and has pursued its quarry (an understanding of the cultural phenomenon known as science) over many different terrains, and into many different corners. One way of offering an introductory book on the history of science would have been to offer overviews of the major themes picked out in recent historiography; the organization of science, for example, or the places in which scientific knowledge is generated, the practice of science and its practitioners, science and imperialism, science and gender, and so forth. In a sense, though, an introductory book on the *historiography* of science is a contradiction in terms. It is important to know the history before learning about the historiography. And besides, an overview of separate historiographical themes could not provide a comprehensive picture of the historical development of science from Ancient times to the twentieth century.

Consequently, this book tries to offer a continuous narrative of the evolution of what we now call science. But, if I have tried to follow E. M. Forster's famous injunction, 'only connect', I have had to impose a narrative structure on the past. I hope my narrative is not only plausible, given the historical evidence on which it is based, but is also as authentic as possible within the restrictions of brevity. There is no denying, however, that major aspects of the history of science are bypassed by the thread of my narrative – I say nothing, for example, about the development of the Periodic Table in chemistry, much less about the concept of valency. There is nothing here about the development of cell theory in biology, and nothing about drifting continents in the history of geomorphology, much less the theory of plate tectonics. I hope, however, that I have said enough to enable readers to pursue the history of these ideas for themselves. If they do so, I hope they will see that these things could have been fitted into my continuous narrative, had I made different choices, in the interests of brevity, as to what to include and what to leave out.

Another organizing principle, which also allows us to keep things briefer than a fuller history would require, has been to focus on the history of ideas, or thought. Our concern is with ways of thinking about the natural world which led successive generations of thinkers to believe that they understood, or to some extent understood, the way the world was constructed and how its parts interacted to result in the natural phenomena we see all around us everyday. One obvious alternative to this approach would have been to focus on the *practice* of science, and to concentrate on the development not just of the experimental method, but of specialist experimental apparatus, and of specialist instruments for observing and measuring the world and its parts. Here again, our focus enables us to compress, and to keep things short. It is not possible to explain the historical development of experimental techniques without also expounding the theories (or perhaps the vague assumptions) that made the experiment seem important. Experiments are designed to test ideas, or to decide between theoretical presuppositions. In the interests of brevity, therefore, we follow the ideas here and for the most part simply report in passing whether ideas were experimentally confirmed or not. The alternative account might well have been richer, but it would also have had to be far longer.

Similarly, we do not look at the various ways in which many of the ideas recounted here led to practical applications. We are interested in why and how James Clerk Maxwell predicted the existence of radio waves, and even in how Heinrich Hertz was able to demonstrate the truth of Maxwell's claims by generating and detecting radio waves in his laboratory (see Chapter 23). We are not concerned, however, with how Guglielmo Marconi exploited these ways of thinking to develop wireless communications. We are interested in why and how Einstein came up with the idea that energy and matter were essentially the same thing, and that, therefore, matter could be converted into energy. We will not, however, pursue the history of how this led to the atomic bomb.

The book begins at what is usually regarded as the beginning with an examination of the contribution of the Ancient Greeks. Ancient Greece produced the earliest thinkers known to have explained the world in a rational and naturalistic (non-supernatural) way, and their approach, one way or another, can be seen to have been repeatedly adopted or adapted for two millennia. After tracing the fortunes of Ancient Greek thought through to the Renaissance, the course considers the phenomenon designated by historians as the Scientific Revolution, when what had been previously an almost entirely contemplative natural philosophy adopted methods and aims which made it more recognizably like modern science. We then follow the story through the period known to historians as the Enlightenment, and then through the nineteenth and into the twentieth century. The story is brought to a close in the first half of the twentieth century. There is no attempt to bring the story up to the present. Along the way we also consider various themes, such as the interaction between Ancient Greek philosophy and Christian theology on early scientific thinking, the development of the experimental method, reasons for the success of Newtonian science, and the changing fortunes of theories of biological evolution.

The major area of focus of the story told here gradually diminishes: firstly, we consider broad attempts to understand the nature of the world system, from the Ancient Greek's concern with the foundations of knowledge to sixteenth- and seventeenth-century concerns with the nature and structure of the cosmos; secondly, we consider eighteenth- and nineteenth-century concerns with the nature of the earth itself, its creation and its development over time; we then narrow our focus to theories of plant and animal development, and ultimately to the development of the human species, before finally descending into the strange world of the atom and the sub-atomic.

One aspect of the attempt to keep the story short which may disappoint readers is the lack of biographical detail about any of the many thinkers discussed. There is no hidden agenda here – I am not trying to suggest that the personal biographies of scientists are irrelevant to their achievements; far from it. It is simply another way of keeping my narrative short. There is no shortage of biographical treatments of scientists, so it should be an easy matter for the reader to satisfy their interests in this regard by supplementary readings. Apart from book-length biographies which exist for almost every thinker covered here, there are a number of biographical dictionaries. For British thinkers an excellent starting point would be the articles in the new *Oxford Dictionary of National Biography*, which is available on-line in major research libraries. For non-British thinkers a good starting point is the *Dictionary of Scientific Biography* (see Further Reading below for full details).

Readers with an interest in the philosophy of science, or sociological theories about science, may also be disappointed by the lack of theoretical underpinning to what follows. There is no all-encompassing theory of scientific or historical change in these pages. I do not subscribe to Thomas Kuhn's theory of the development of science through successive scientific revolutions, or any other single theory of scientific discovery, or scientific consensus formation. As a historian, I am conscious of the dangers inherent in preconceived ideas. As Francis Bacon pointed out, way back in the early seventeenth century (see Chapter 7), theories have minds of their own, so to speak, and force their acolytes to note the evidence that fits and to disregard the rest. There is a still thriving tradition among philosophers and sociologists of science of using historical case studies to illustrate and support (or to refute) the philosophical or sociological theories in question. I am not doing that here – I am merely trying to present a narrative account of how science developed. I leave it to the philosophers and sociologists of science to see what general principles, if any, they can discern in my historical account. I am all too aware that some philosophers and sociologists have been guilty of relying on bad history, or at best, restricted history, which works for their discussion, but cannot be extended to cover different aspects of science. Good history, revealing the importance of contingency in any and every unique state of affairs, tends to undermine any attempts to establish a neat and tidy philosophy of science. Indeed, professional history of science – a very recent profession, beginning effectively in the 1960s – grew out of dissatisfaction with the way philosophers of science were prone to write histories of science to suit their philosophical theories of the way science is.

In this short history, therefore, I have tried to give an account of what happened in history, of how it all began, where it went from there, and how we got here. My ambition was only to provide good history, not to offer anything else besides. As with all historical studies, therefore, it is hoped that it will have made clear the need for a culturally relative perspective when trying to understand the past, while at the same time revealing how difficult it is to escape our own cultural perspective.

FURTHER READING

W. F. Bynum, E. J. Browne, and Roy Porter (eds), *Dictionary of the History of Science* (Basingstoke: Macmillan, 1981).

Geoffrey N. Cantor, J. R. R. Christie, Jonathan Hodge, and R. C. Olby (eds), *A Companion to the History of Modern Science* (London and New York: Routledge, 1990).

Gary B. Ferngren, Edward J. Larson and Darrel W. Amundsen (eds), *The History of Science and Religion in the Western Tradition: An Encyclopaedia* (New York: Garland Publishing Inc., 2000).

C. C. Gillispie (editor in chief), *Dictionary of Scientific Biography*, 16 vols (New York: Scribners, 1970–80).

Arne Hessenbruch (ed.), *Reader's Guide to the History of Science* (London and Chicago: Fitzroy Dearborn Publishers, 2000).

Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962).

H. C. G. Matthew and Brian Harrison (eds), *Oxford Dictionary of National Biography*, 60 vols (Oxford: Oxford University Press, 2004); available on line (with a license) at www.oxforddnb.com.

1

Setting the Scene: Natural Philosophy in Ancient Greece

The earliest attempts to study the world seem to have been carried out for pragmatic ends. There is evidence that early civilizations had some knowledge of plants and other natural things (whether animal or mineral) which could be used for medicinal purposes. More impressively, there is clear evidence that the early civilizations of Mesopotamia and Babylonia made detailed observations of the heavenly bodies, and it seems impossible to deny that prehistoric structures, such as Stonehenge and Newgrange, were built at least in part to reveal precise moments in the cycles of the sun or moon, or both. The building of structures like these, and of impressive monuments and tombs, such as the pyramids erected by various ancient civilizations, show that prehistoric and ancient people were capable of exploiting specially developed technology, and of developing working mathematical techniques, to help them accomplish feats which we would still find difficult today. But our concern in this book is not with pragmatic ingenuity, even if it involves mathematics and technology. Our concern is with attempts to *understand* the phenomena of nature. The aim of this chapter is to uncover how and why humankind first sought not to exploit but to *explain*, the physical phenomena they saw going on around them.

The earliest attempts at explanation of what is going on behind the appearances, however, all seem to rely on a kind of anthropocentrism. The sun travels across the sky because it is carried on a chariot, pulled by horses, but driven by a man. The charioteer is a special kind of man, of course, and so is designated as a god, but otherwise he is pretty much like a man. Similarly, earth and water are two separate kinds of things because the earth is a woman and water is a man, and by mating together they give rise to vegetation and other creatures. Again, this woman is designated a goddess and the man a god, but these procreating gods are clearly modelled on human beings, and their ability to create is based on the familiar human experience of producing a new human being. As more complex and intricate matters needed to be explained, more characters were introduced into the pantheon who interacted with one another in accordance with various power plays, lusts, jealousies, and other all too human foibles. Ancient civilizations explained things in terms of soap operas, not science.

What we are looking for is something that we can recognize as the beginnings of an attempt to explain the natural world in naturalistic terms. This sounds circular, but what we mean is that we need to find a civilization which explained the world, not in human terms, but in its own terms, and in so doing simultaneously defined what are natural terms, and characterized the world explained by such terms as the natural world.

The earliest civilization to do this, as far as we know, was the Mediterranean civilization of the Ancient Greeks. The Ancient Greeks are seen as the first people to develop systematic ways of thought, and they even gave these ways of thinking a name – philosophy. Philosophical thought might primarily be concerned with morals or with the law and with forms of government, but for some thinkers at least it was considered important to develop a philosophy of nature – a way of understanding why the world beyond the realms of human society is the way it is, and how it works.

The first thinker to turn his attention to explaining natural phenomena in natural terms that we know of was Thales (*c.* 624–*c.* 546 BC), who lived in Miletus, a town in the Greek territory of Ionia, which is now on the mainland of Turkey. The period beginning with Thales, then, is seen as a unique turning point in the history of mankind – and one which was to lead to the development of the philosophical investigation of nature with which modern Western science can be seen to be historically continuous.

But if this is so, then it is a truly remarkable turning point, and we must immediately ask ourselves, therefore, *why* there was this unique change, why did it happen in Greece and why during the period from the sixth century BC onwards?

The first thing to say is that **it will not do** to answer this question by saying that the Ancient Greeks were blessed with a remarkably high number of geniuses in their ranks (Box 1.1).

So, why did the Ancient Greeks develop a rationalistic and naturalistic approach to attempts to understand the physical world? You may be surprised to hear that the consensus among classical scholars is that the unique system of politics of Ancient Greece was the major factor.

Box 1.1 General presupposition of intellectual history

Genius is not a good explanation (nor is stupidity)! We need to seek accounts as to why some thinkers saw what others did not see; were able to understand things by ways of thinking that others either could not follow, or deliberately rejected. It is no more helpful to say that Galileo thought of the law of free fall because he was a genius, than it is to say that he thought of it because everybody else at the time was an idiot. It is nearly always possible, after historical research, to offer explanations of why a particular discoverer happened to be able to make the discovery that they did. Usually it is a matter of the discoverer being in the right place at the right time – but we mean here the right mental place, having the right mind-set. Sometimes such explanations must remain tentative and speculative – such is the nature of history – but this is still more enlightening than attributing it to genius, which is simply another way of saying the discoverer discovered what they discovered.

Now, we needn't go into the fine detail of Ancient Greek history but suffice it to say that, unlike other early civilizations, such as Egypt or Babylonia or Persia, where powerful monarchs held sway over all the peoples, Greece was made up of an often unstable mix of independent city states (Athens, Sparta, Corinth, etc.). This unique political organization was largely due to geographical factors. Even on mainland Greece, the development of a city was confined to a narrow valley floor and separated from other cities by the mountainous terrain. Other cities were located on small islands in the Greek archipelago or in small colonies (surrounded by other civilizations) on the coast of what is now Turkey, or the coast of Italy and even North Africa. Travel even between mainland cities had to be conducted by sea. The relative autonomy of these cities, and their small populations, led to political confrontations at a face-to-face level, resulting in a critical independence of thought among the inhabitants which, from the sixth century BC, began to establish itself in democratic forms of governance (replacing earlier oligarchies or tyrannies, which had been resented by significant factions of the citizenry). Although the emphasis on democracy implied a more egalitarian society (at least among those citizens allowed to take part in the democratic processes, which did not, for example, include women or foreigners), this introduced complications into social organization which needed to be managed by complex legislation. One result of this was a population – throughout all the Greek cities – which was politically knowledgeable, critical, and accustomed to participating in government to an unprecedented extent. Arising from this political sophistication was a philosophical sophistication about the nature of law and justice and related ideas.

In Egypt and Mesopotamia, laws were simply regarded by the people as the pronouncements of the supreme authority – the Pharaoh, the Emperor, or whoever. In Greece, since every citizen played some role in government or at least in deciding how they should be governed, the notion of law as an abstract but real entity with a character of its own came to be recognized by the people. So, in other words, laws were not seen merely as arbitrary whims of a transient tyrant, but were seen as 'natural' concomitants of the nature of society – laws were inherent in the nature of society. Without the laws, it was assumed, the society could not function and could not have established itself.

It seems that this critical way of looking at the way society functions, and the role of law in maintaining these functions was carried over into the study of nature. What was unique to Ancient Greek philosophy was the concept of natural law – that the world system itself was not a collection of unconnected things but was a *cosmos* – an ordered system like a well-governed city state, which operated according to the laws of nature, laws which governed the behaviour of all things. These laws of nature were seen as inherent in the nature of the world. And the earliest discussions of the laws of nature betray their origins in analogies with laws of society. So, for example, Anaximander (*c.* 610–547 BC) wrote: 'Things give satisfaction and reparation to one another for their injustice, as is appointed by the ordering of time'. By 'things' here he evidently did not mean things in human affairs, but things in general; he was making a comment about the nature of the world. Similarly, Heraclitus (*c.* 535–*c.* 480 BC) said: 'The sun will not overstep his measures; if he does, the handmaids of Justice will find him out'. We still talk today of inanimate objects 'obeying' the laws of motion and Heraclitus seems to be trying to say that the sun cannot deviate from its heavenly path because it would be breaking the law (policed by 'the handmaids of Justice') for it to do so.

Now, of course, it would be naïve to assume that every ancient Greek saw this clearly and had an instinctive view of nature as a cosmos ruled by natural law. But certainly the leading intellects thought this way and, what's more, schools were established throughout the Greek territories to inculcate these ideas more widely among the population. The tendency towards democracy created a populace who recognized the importance of skilful and persuasive public speaking, and the need for people trained in the ways of public affairs – trained as lawyers, administrators, teachers, and orators – men who could convince others by the force of their arguments, or by the persuasiveness of their words.

The social and political structure of Ancient Greece, therefore, provided an important ecological niche for *philosophers*, in a way that no society had before (or since?). Philosophers established schools to train people in the arts of persuasion by dint of rhetoric or logical argumentation.

It was these same philosophers – at least the most prominent of them – who began to develop the natural philosophy which has been seen as the origin of the scientific approach to knowledge. Thales, for example, was evidently a law-giver, and political leader, as well as a natural philosopher; Anaximander was said to have helped create laws and a constitution for the colony of Apollonia; and Parmenides of Elea (*fl. c.* 480 BC) was said to have written the laws of his city.

The problem of change and the reductionist impulse

The reports we have of the ideas of the very earliest natural philosophers make it clear that they were concerned to find order in the seemingly chaotic world around them. Furthermore, they seemed to have looked for it in an underlying ordering principle; that is to say, they tried to reduce complex realities to a single explanatory principle, or to a small group of explanatory principles.

Obviously, the physical world around us is in a constant state of change (to use my favourite oxymoron), and this presents problems for any thinker seeking to reduce this turmoil to an ordered *cosmos*, ruled by law (*logos*). Indeed, generally speaking the Ancient Greeks seem to have been extremely reluctant to acknowledge changeability. The renowned classical scholar W. K. C. Guthrie explained this in terms of human nature: 'There is simply a deep-seated tendency in the human mind to seek something that persists through change'. But I find this unconvincing. It seems to me that the political background provides a better clue

Box 1.2 Some difficulties of studying Ancient Greek philosophy

First of all, it's all too easy to think of Ancient Greek science as belonging to a single period in history. That is to say, the label 'Ancient Greek' can sometimes be thought of as analogous to talking about 'science in the twentieth century', or 'Victorian science', or some other brief period. In fact Thales of Miletus (624–548 BC) flourished at the beginning of the sixth century BC, while Plato (347 BC) and Aristotle (384–322 BC) flourished in the fourth century BC, and Ptolemy (90–168 AD) and Galen (129–199 AD) strutted their stuff in the second century AD. So, about 700 years separates Thales from Ptolemy and Galen. This is the difference between us and someone living in the thirteenth century (though I grant you the rate of change has been somewhat faster of late).

Another difficulty is that the very earliest natural philosophers, as far as we know, did not write anything down. So we only know what they believe from reports by later thinkers, such as Aristotle. If Aristotle's, or whoever's, account of this earlier thinker is critical, or even hostile, then it is difficult to be sure that we are getting an accurate picture of that thinker. It's as though all of Sigmund Freud's works got lost and the world relied upon me to reconstruct what he thought. Apart from the fact that I don't know everything that Freud wrote or believed, I think that what I do know of him (the Oedipus complex, the vaginal orgasm, women as 'normal hysterics', the ego and the id and the superego and all that kind of stuff) is the most ludicrous farrago of nonsense I've ever heard and should never have been taken seriously for even a moment. I'm hardly likely to give the best account of it, therefore.

Even in the cases of Greek philosophers who did write things down, it's all so long ago that often we no longer have full or reliable versions of what they wrote. All that we have are a few fragments, usually quotations in later Greek writings. The only Ancient thinker whose works we have more or less complete is Plato. One of the richest sources for an understanding of earlier thinkers is Aristotle, but there is even some doubt about the authenticity of his writings. Some texts attributed to Aristotle are in fact not written by Aristotle himself but are Ancient compilations from the lecture notes written by his students. You can imagine how reliable a quotation, supposedly from (say) Thales, is likely to be if it was quoted by Aristotle in a lecture and taken down by one of his students.

Another major difficulty with reconstructing the opinions of the earliest thinkers derives from developments in the Greek language. When Aristotle reports the opinions of some of his predecessors he often uses words like 'element' and 'substance' which linguistic scholars tell us did not exist when the reported philosopher was alive. What Aristotle refers to as the four elements first appeared in the philosophy of Empedocles designated by a word that is closer to our word root – the four roots. Moreover, when we do have reliable quotations from early thinkers the meaning is often difficult to grasp because they lack the kind of technical terms which would only become current later on. Consider the examples we have already seen of Anaximander and Heraclitus trying to invoke the notion of laws of nature.

Nevertheless, there is no shortage of classical scholars who spend their time reconstructing the beliefs of the earliest known natural philosophers.

to Greek ways of thinking. The main difference between man and Gods, according to the Greeks, was the fact that men are ignorant of their own destiny. The emphasis on political awareness and political involvement in Greek society was, as many Greek writers testify, an attempt to gain as much control as possible over one's destiny. One way of doing this, it was believed, was to establish some kind of political stability to counteract the vicissitudes of contingency.

It seems that this kind of thinking was carried over also into their attempts to understand the physical world: there must be some stability, some unchanging truth, behind the changes we see all around us. So, how did the Ancient Greek philosophers find this stability? The simple answer is: by a number of different reductionist strategies.

Let's start with Thales. Now we don't know much about Thales but we do know he suggested that, in spite of appearances, all things in the world are made of *water*. Water is fluid and changeable, and yet always water. It can become something that we think of as a gas (vapour, or steam), or a solid (ice), but still it is water and can change back to its fluid state. It seems that Thales wanted to suggest that all things were made of water in different forms, and so the bewildering complexity of the world was really only one thing.

But there's a big problem with supposing that all things are different manifestations of water: it seems to defy the common sense belief that fire is, as it were, a direct opposite of water. Fire dissipates water, and water destroys fire. So how could fire be explained as some kind of variation on water?

This kind of reasoning led Anaximander, another Milesian, to come up with a notion that all things come from an abstract principle known as 'the Boundless' (or 'the Eternal'). Anaximander describes the world we see as growing out of the boundless in the same way that a tree grows from a seed. Now, notice that a seed looks like undifferentiated substance and yet it gives rise to leaves, fruit, bark, pith, roots, flowers, and so on. So, likewise, the multiplicity of things in the world could all come from the boundless, even though the boundless itself was undifferentiated. Unfortunately, we don't know any more about how Anaximander thought of this.

A third Milesian philosopher, Anaximenes (*fl.* 545 BC), reverted back to a more physical first principle: *Air*. Like Thales he claimed that air could precipitate into water (he was thinking of the formation of clouds and their precipitation into rain), and, in turn, concrete itself into ice. Significantly, however, there seems to be no absurdity in supposing that air could rarefy itself into fire or even light, both of which could be (and often have been) seen as fluids of greater subtlety and rarity than air.

Now these are the earliest examples of natural philosophy and they might seem just like myths to some readers, hardly worth regarding as the beginnings of scientific thought. But the thing to notice here is that they don't invoke the behaviour of the gods, or anything non-natural in their explanations. Even Anaximander's 'boundless' is likened to a seed (though perhaps it should be seen more as a chaotic fluid filling the universe out of which the various things in the world precipitate – but either way, the image is a natural one, not one which depends upon human-like actions).

Furthermore, there are clear indications of rational thinking among these early thinkers. We've already noted Anaximander's rejection of Thales' water principle on the grounds that it couldn't account for fire. But he also proposed that the earth was simply hanging unsupported in space, and that it remained where it was because it was equidistant from all surrounding heavenly bodies. This was a rational response to the theory sometimes attributed to Thales that the earth must be floating on cosmic water – which immediately raises the question as to what contains the water. Similarly, Anaximander argued that mankind must have evolved from fish. His reasoning here was that human babies require a good deal of looking after for a number of years before they can look after themselves, and so humankind could not have appeared suddenly on the earth. The fact that he took it for granted that humans must have first appeared as babies, rather than as fully formed adults (like Adam and Eve) is further testimony to the fact that he was thinking in naturalistic terms.

If these early thinkers tried to reduce the complexity of things to a simple constancy underlying all change, another Ionian thinker is actually famous for claiming that all things were continually in flux. Heraclitus (*c.* 535–*c.* 480 BC) is most famous for insisting that it is impossible to step twice into the same river (because the second time, the water would be completely different – the river is never the same). Even for Heraclitus, however, it was the stability that persisted through change that was the most important thing to grasp. Although our daily experience might suggest to the unreflective mind nothing but change, more careful thought, Heraclitus believed, should lead us to conclude that there must be a coherent plan, a 'determination' which restricts the kinds of change that can take place (the river, after all, will always be constituted by flowing water, not by whisky, or treacle). Heraclitus referred to this determining feature as the *Logos* (which might be translated as 'measure' or 'reckoning', but which later came to mean 'reason' or even 'word'), a principle of order which he believed could be discerned by our experience even as we are bombarded by experience of change: 'For although all things happen according to this Logos men are like people of no experience... [and] fail to notice'.

It is fairly easy to understand the idea that a 'Logos' could be responsible for uniting all things in a coherent complex, but Heraclitus, known even to his contemporaries as an obscure thinker, complicated the issue. He seems to have identified the Logos with fire. This may look like a return to the strategies of Thales and Anaximenes, substituting fire for water or air, but the fact that Heraclitus equated this with a more abstract sounding principle like the Logos makes it more reminiscent of Anaximander's 'boundless'. Furthermore, it is evident that Heraclitus saw fire as an active principle of some kind; fire was not just another form of matter but was a motive force at work in the cosmos. But if fire was also the Logos, the determining principle, it meant that fire did not act randomly, but in a way which ensured that change always remained within the bounds set by the Logos.

Box 1.3 Pythagoras, the Pythagoreans, and the mathematics of the natural world

Pythagoras (*c.* 570–*c.*495 BC) is now one of the most famous of the Ancient Greek philosophers, and yet he was already a shadowy figure, obscured by legend, shortly after his own lifetime. He was not only, for a time at least, a legislator and state administrator, but also came to be recognized by his followers as a religious leader, the founder of a secretive religious brotherhood. As with other religious leaders legendary tales, some credible only to the most gullible believer, were told of him. The result was that even by the time of Plato and Aristotle, about a century after his death, those outside the ranks of his followers were forced to be sceptical of everything claimed about him. We need not consider his religious and moral teachings here, although it seems clear that his work in geometry and mathematics, for which he is still remembered, was seen by him and his followers as an integral part of his religious teachings. One of the major legends about Pythagoras, which may well echo the truth, is that he discovered that the principle musical intervals correspond in some way to simple numerical ratios. The lengths of the string on a monochord with a moveable bridge varied according to the interval. The lengths which produced an octave were in the ratio of 1:2, those which produced the fourth were in the ratio 4:3, and of the fifth, 3:2. From here, the Pythagoreans went on to suppose that all things might be represented somehow by numbers or ratios. As Aristotle reported it: 'they thought its [mathematics'] principles were the principles of all things'. A detailed account of Pythagoreanism would take us too far out of our way, but suffice it to say that numbers, or geometry, and the rules that they both obey, seemed to the Pythagoreans to impose a harmony on what might seem like unruly nature, and a world of unintelligible changeability. If they believed, as Aristotle reported, that 'the elements of number' are 'the elements of all things', then the changes in the natural world should be reducible to harmonious rules. It might seem to us, therefore, that the Pythagoreans discovered what we all now know to be true, namely, that the physical world can be analysed in mathematical terms. But the Pythagoreans did not think the same way as we do (how could they?) and invested number with symbolic and even mystical significance. In part, this

➠

If we move our focus from the Greek territories at the eastern end of the Mediterranean Sea to those on the south of what is now Italy, we come to a very different way of dealing with change. Arguably one of the most prominent of the early Greek philosophers was Parmenides of Elea. Indeed, the impact of Parmenides on subsequent Greek thought, shows just how important it was for learned Greeks to solve the problem of change, and to find unity underlying diversity, and order behind the changeability they noticed all around them. We know little about Parmenides, but we have substantial extracts from his most important work, a poem entitled *The Way of Truth*.

Strange though it may seem, Parmenides developed a philosophy in which he denied the reality of change. In a way, therefore, he confronted head-on the probmay have amounted to what we call 'numerology'; Aristotle said that the Pythagoreans tried to understand concepts such as 'opportunity', 'justice', and 'marriage' by analogy with the alleged characteristics of certain numbers. There were also attempts to allocate numbers to different things by suggesting that a horse, say, required a certain number of points or dots to depict it as unmistakeably a horse, rather than any other four-legged animal (think of how the stars in the constellation of *Ursa major* are supposed to depict a bear). The number of dots became the number of a horse (or a bear, or whatever). This was also associated with a kind of atomistic view of the world, in which unity, the number one, is seen as a physical point in space, and so physical objects can be built up of unit points. As Aristotle said, the Pythagoreans 'supposed real things to be numbers—not separable numbers, however, but numbers of which real things consist'. It is possible that some of these ideas were falsely reported as a result of misunderstanding by non-Pythagoreans, but there can be no denying that the Pythagorean tradition mixed what we would recognize as mathematics with more mystical and magical notions of the supposed significance of numbers. In geometry, their interest in the properties of the so-called 'golden section', led them to adopt the pentagram, the five-pointed star in which the crossing lines divide one another in the golden section, as the badge of their secret brotherhood. In this way, Pythagoreanism led not to mathematical physics but to mathematics becoming a part of the magical tradition (and the pentagram became a sign of sorcery). Even so, the Pythagoreans represent the earliest thinkers to recognize the importance of mathematics for understanding the natural world.

FURTHER READING:

Walter Burkert, *Lore and Science in Ancient Pythagoreanism* (Cambridge, MA: Harvard University Press, 1972).

Dominic J. O'Meara, *Pythagoras Revived: Mathematics and Philosophy in Late Antiquity* (Oxford: Oxford Uniiversity Press, 1991).

H. E. Huntley, *The Divine Proportion: A Study in Mathematical Beauty* (New York: Dover, 1970).

lems associated with the changeability of the world, simply by denying there was any change. In an argument that evidently impressed his contemporaries, he insisted that the only thing that really existed was something called 'the One'. And, he insisted, the One never changes.

I stress that this argument impressed his contemporaries because it is rather difficult for us now to appreciate its force. Indeed, it seems to be merely playing with words. Putting it simply (and rather crudely) what Parmenides said was that there can be no change because it is impossible for something to *become*, that is to come into existence from non-existence. The notion of 'becoming' has no real meaning, Parmenides said, because it implies that something becomes something else, which is to say that something becomes what it is not. But if something becomes what it is not, then it becomes what doesn't exist, which is to become nothing. So, if something cannot become something else, then there can be no change. Change is a logical impossibility.

Parmenides decried those who 'are carried along deaf and blind at once, altogether dazed'; for him they were 'hordes devoid of judgement, who are persuaded that to be and to be-not are the same'. Presumably, what Parmenides meant by this last remark was that most people simply accepted that change occurred – that a state of affairs existed at one time, but later that state of affairs no longer existed. This might seem perfectly reasonable to us, but for Parmenides what this meant was that the 'hordes' were accepting that there was no significant difference between existing and not existing. But of course there is a world of difference, and to glibly accept change was to fail to notice the dichotomy between being and not-being. Accordingly, Parmenides insisted that we must reject change. The only true statement is that a thing *is*:

> One way only is left to be spoken of, that it *is*; and on this way are full many signs that what *is* is uncreated and imperishable, for it is entire, immovable and without end. It *was* not in the past, nor *shall* it be, since it *is* now, all at once, one, continuous.

Parmenides extended these ideas, then, to argue that the appearance of change in the world was merely illusion and that the world was in fact 'One' and unchanging. Classical scholars are still arguing about what he meant by the One, but it seems fairly clear that in his *Way of Truth* Parmenides wasn't interested in details (such as how grass eaten by a cow can become milk), but only in fundamental metaphysical truths. 'Turn your mind away from this path of enquiry', Parmenides said of attempts to understand the details of physical change, 'let not the habit ingrained by manifold experience force you along this path, to make an instrument of the blind eye, the echoing ear, and the tongue, but test by *reason* my contribution to the great debate'.

Now, it might be that you are convinced that it makes no sense whatsoever to claim that change is logically impossible. But consider the paradoxes raised by Parmenides' most famous follower, Zeno of Elea (*fl.* 450). Zeno suggests that it must be impossible to walk across a stadium (or any other distinct distance) because before you can reach the other side, you must pass the half-way point. No problem? But consider the fact that before you reach the half-way point you must reach the point half-way to that point, and before you reach that half-way point you have to get half-way to that one, and so on. If you think about it you ought to be rooted to the spot because you have to pass an infinity of points before you can complete even one step (you have to get half-way to completing one step before you can take the step...). But surely it is impossible to pass an infinity of points in a finite time.

And then there's the famous paradox about the race between Achilles and the tortoise. Achilles gives the tortoise a head start. It doesn't matter how big (or how small) a start the tortoise has, as long as we assume that once the starting gun goes off the tortoise is in continuous unceasing motion, he must have already left his starting point by the time Achilles gets to it. The tortoise is still ahead, therefore, and Achilles has to make up this new distance. This will certainly be a smaller distance than the original head start (because Achilles is very swift) but, even so, by the time Achilles covers it, the tortoise will have already moved on (remember the tortoise is in continuous motion, so must have moved as Achilles was trying to catch up). Achilles is closer now but he still has to make up a short fall – but by the time he makes up this third distance, the tortoise has already moved on ahead. Poor old Achilles never can catch that damned tortoise.

Now, Zeno must have known that he could walk across a stadium in no great time, and if he'd been a betting man he would surely not have bet against Achilles. His arguments don't constitute a proof that motion is impossible, but they do seem to point to a paradox. A paradox works by showing that initial assumptions (such as that motion is unproblematic) lead to a contradictory or irresolvable state of affairs, and so suggest that your initial assumptions must be wrong. Motion is a kind of change and ought to be logically impossible if Parmenides is right. Zeno's paradoxes suggest that, counter-intuitive as Parmenides's claims are, there may be something in them.

Parmenides could be said to have caused something of a crisis in Greek philosophy. The challenge was to accept the metaphysical truth (arrived at by privileging rational thought over experience) that there was only the One, which was unchanging, while also accounting for the myriad different phenomena of the world. Nobody succeeded in this until the atomists, Leucippus (*fl.* 440 BC) and Democritus (*fl.* 420 BC). In Ancient atomism it was assumed that all things were made of just one substance, one kind of matter, but that it appeared in different forms, and can undergo various changes, because the one substance is divided up into atoms which can combine in different ways, break-up and recombine, in such a way as to provide all the appearances of the physical world.

It is important to note, therefore, that these two fifth-century BC thinkers arrived at the essence of what scientists today believe about the atomic structure of the physical world, without the benefit of any experimental apparatus or any of the paraphernalia of modern science, but just by thinking of a way to explain the changeability of the world without conceding that the world was simply an inexplicable chaos. The atomists effectively accepted the reality of the Parmenidean One, but wanted to show how change and motion could nonetheless occur. The One became matter, of which there was only one kind, but diverse changes could be explained by the ways invisibly small particles of that matter combined and re-combined. Furthermore, they could show that Zeno's paradoxes were invalid by insisting that the tortoise can't move continuously but only in indivisible increments, and so Achilles can catch up; and we can easily cross the stadium because space cannot be divided infinitely, but only down to the size of atoms, thus ensuring that we don't have to pass an infinite number of points to get from one place to another.

Ingenious as the atomists were, their ideas seemed as counter-intuitive to many of their contemporaries as those of Parmenides and Zeno. The alternative response was to reaffirm a more down-to-earth common-sense approach to dealing with the problem of change. One of the most influential of these was the system of Empedocles (*fl.* 440 BC).

Empedocles drew upon what seems to have been the popular view of the natural world, namely that all things were made of earth, water, air and fire, and he added to these the notions of love and strife (evidently close to notions of attraction and repulsion, or affinity and disaffinity) to account for the interactions of these four substances. Changes are explained in terms of the combination, dissolution, and re-combination in different proportions of these four 'elements'. In essence it was a return to the kind of reductionist approach developed by Thales and Anaximenes but acknowledging the need for more than one essential ingredient.

The four elements were considered in a flexible way – water was often discussed in terms which we would find closer to talk of a 'fluid principle', and earth was seen as a principle of solidity or heaviness. Analyses of the composition of individual bodies was also carried out in a notional way, rather than a strictly empirical way. Wood clearly consisted largely of earth, but because it floats on water it must also contain air. This can be confirmed by burning the wood, when the air can be seen escaping in the form of smoke. The fact that wood can be ignited indicated that it must also contain fire, which can also be seen escaping, in the form of flames, as the wood burns. A bar of iron, by contrast, cannot be ignited and so contains little or no fire. It will not float and so contains little or no air. It is not pure earth, however, because on heating it melts and turns to fluid, so must contain water. Although much of this kind of analysis was derived from reasoning based on rudimentary, everyday, experience, it was worlds away from the emphasis upon reasoning that we see in Parmenides and his followers.

For Parmenides the physical world was too changeable to be true – truth had to be unchanging and eternal, and what is real, therefore, must be the One, and our knowledge of it must be based on pure reasoning. The physical world, by contrast, by dint of its changeability, can only be deceptive and unreliable, and we can only have opinions about it, but not real knowledge. For Empedocles, however, talk of an unchanging One was neither here nor there; the point was to understand the world, and to do so we had to take seriously the many ways in which it changed and the circumstances of those changes. This could only be done by scrutinizing the world itself, and seeking to interpret what our senses tell us with the aid of reason.

The important thing to notice, then, is that from about 450 BC, there were two contrasting approaches in Greek philosophy – one, exemplified by Parmenides, emphasizing rational thought as the way to arrive at truth, and rejecting experience and the information provided by the senses as misleading; and the other, exemplified by Empedocles, which assumed that the only way to arrive at truth was through a cautious reliance upon the senses and a thorough familiarity with the physical world. These two approaches were each to receive a great champion.

FURTHER READING

Mott T. Greene, *Natural Knowledge in Preclassical Antiquity* (Baltimore: Johns Hopkins University Press, 1992).

W. K. C. Guthrie, *The Greek Philosophers From Thales to Aristotle* (London and New York: Routledge, 1968).

Paul T. Keyser and Georgia L. Irby-Massie (eds), *Encyclopedia of Ancient Natural Scientists: The Greek Tradition and its Many Heirs* (London and New York: Routledge, 2008).

G. S. Kirk, J. E. Raven, and M. Schofield, *The Presocratic Philosophers*, 2nd edn (Cambridge: Cambridge University Press, 1983).

David C. Lindberg, *The Beginnings of Western Science: The European Scientific Tradition in Philosophical, Religious, and Institutional Context, Prehistory to A.D. 1450* (Chicago: University of Chicago Press, 2008).

G. E. R. Lloyd, *Early Greek Science: Thales to Aristotle* (London: Chatto & Windus, 1982).

G. E. R. Lloyd, *Magic, Reason and Experience: Studies in the Origins and Development of Greek Science* (Cambridge: Cambridge University Press, 1979).

2

Plato and Aristotle

The abstract rationalist approach to knowledge, insisted upon by Parmenides, was championed by Plato (*c.* 427–*c.* 347 BC), while the more down-to-earth, common-sense approach of Empedocles was to be championed by Aristotle (384–322 BC). If we exclude the founders of the world's great religions then these two philosophers have proved to be, far and away, the most influential thinkers in the history of the world so far. Even if we dismiss as insignificant their continuing influence (they continue to be studied in university philosophy departments, and ideas from both continue to be taken up by philosophical theologians), they were a major influence from their own time through to the nineteenth century AD, and on the whole of European and North American culture. No other secular thinker comes anywhere near the same level of usefulness for subsequent thought or culture. Alfred North Whitehead, a twentieth-century philosopher whose significance is already being forgotten, once declared that the history of Western philosophy was nothing more than a series of footnotes to Plato. The more you learn about intellectual history, the harder it is to disagree with Whitehead's point.

In order to keep things short, we are going to look at these two thinkers, as with earlier thinkers, from the point of view of the problem of change. How did Plato and Aristotle deal with change? Let us start with Plato.

Plato and the problem of change

Like Parmenides, Plato seems to have regarded change as extremely distasteful. Change is going on all around us, but it ought not to be. The cosmos should be perfect, ordered, and *therefore* unchanging. Plato was clearly a highly religious thinker and underlying this 'therefore' was a belief that the world was created by a divinity as perfectly as it could be, and any change would necessarily result in deterioration – and that is why change ought not to have been allowed by this divine creator. So, Plato is immediately confronted with a problem: Why isn't the world unchanging?

In order to answer this question, what Plato does, putting it crudely and over simply, is to say that in fact the power of reason reveals to us that really the cosmos is (must be) perfect and unchanging, and so the world we see around us, the changeable and changing material world, is *not* the real world.

Implicit in this conclusion are assumptions about the nature of knowledge itself – how we know things. We do not come to know about things through our senses, or our experience, Plato insists, because we are continually deceived by these things. We think we've seen or heard things, when we can't have done, we mishear what people say, we hallucinate, we sometimes think something is happening and wake up to discover that it was only a dream, and so on. What's more, the physical world is constantly changing – you can say: 'I know it's raining outside', but in fact it has stopped raining. So, precisely because the material world is continually changing, you cannot really talk about knowing anything about that world; it is more appropriate to talk about opinion (thinking, believing – 'I think it's raining outside').

We come to *know* through the use of *reason*. Knowledge must be certain and unchanging – if I say that 'I know *x*', then *x* must be true now and always. The major exemplars of knowledge, according to Plato, are mathematics – especially geometry – and logic. It is said, of course, that over the doors of Plato's Academy in Athens (the original Academy from which all others take their name) was written: 'Let no-one enter who is ignorant of geometry'.

But notice this: although geometry deals with abstract entities, logic purports to deal with things in the real world. We make logical propositions and logical arguments about the physical world. So surely this means we can have knowledge of the real world. A formal logical argument would be something of the kind: all men are mortal, Socrates is a man; therefore, Socrates is mortal. This seems to be telling us something true about the material world. And yet, given that we have just agreed that the physical world is changeable and seemingly chaotic, and not really susceptible to claims of the kind 'I know...', how is it possible for us to make a logical argument which seems to establish an undeniable truth (in this case, that Socrates is definitely mortal)?

Plato answers this question by recourse to one of the most remarkable and one of the most influential ideas in Western philosophy. Logic can tell us truths about the real world because the real world, we might say the *really* real world, is not the material world but a world of *Forms*, or *Ideas*.

So, how does this argument work? Plato tries to illustrate it by pointing out how difficult it is to define even so simple a thing as a table.

Suppose I asked you to define a table. You might think this is easy and start to come up with statements like: 'a flat rectangular piece of wood supported horizontally on four legs'. But I could simply say, some tables are round, and you'd have to drop the 'rectangular'. I could also say that some tables are made of glass, or plastic. Banqueting tables have many more than four legs and small round tables sometimes only have one central leg, splayed out at the bottom. If your definition tried to allow for and encompass all the tables you've ever seen, I might still be able to throw doubt upon whether your definition allows for the kinds of traditional tables that might have developed among the Inuit, or the tables that might be produced in the future. So, defining even as simple a thing as a table is by no means easy.

But you'll no doubt be impatient with this and will be thinking to yourself something like: 'yes, yes, but we all know what we mean by a table, we don't have to define it!' This is precisely Plato's point. We do know what we mean by tables and chairs, to say nothing of more complex things like flowers, or 'justice', or human being. Imagine trying to give a definition of a human being, instead of a table. Surely you can see straight away that any fixed definition of a human being is likely to have important consequences. 'An intelligent creature capable of communication in one of the human languages'? So, can we cull all those suffering from brain damage or arrested mental development, on the grounds that they are not human? It would not simply be the parents of such children who would object to a claim like this. Many would reject the original definition. It won't do to say 'any creature born of *human* parents' either, because we are trying to define what we mean by 'human' and so cannot use the term 'human' to define what we mean by 'human'. Although we could argue the point indefinitely, in the end we all know what we mean by 'human', and we are rightly extremely suspicious (or ought to be) of anyone who seeks to impose a fixed definition. It is obvious that those who impose such a definition have some purpose in mind and the definition is serving that purpose.

Plato's point is, then, that we each have a full repertoire of concepts which enables us to function in the world, to discuss it logically, to understand most of what is going on around us, and so on. We know things, even though we cannot give adequate definitions of them. So, this immediately presents us with a problem. *How* do we know these things? How is it that we know what is and what is not a table? And notice that we do not just know particular tables – like the one we always eat our dinner off, or the one we always sit around when we go to our favourite café – we know what it is for something to be a table or not. We can recognize all never-before-seen tables as tables. How is this?

Plato answers this question by saying that what we know is the *form* of a table, or the *ideal* table. All physical tables conform in some way to the ideal table; each of them are poor imitations of the one true table, the ideal table. Since we know the ideal table, we are able to recognize the poor imitations. The ideal table or the form of the table, then, is a kind of blueprint of what a table should be.

The question still remains: *how do we know about the ideal table* (and remember there are also ideal chairs, flowers, an ideal human, and a form of justice, etc.)? How have we come to learn about the ideal table, if all we've ever seen are poor imitations of this one true table?

The answer, according to Plato, is that the real world is the world of the forms, of the ideals, and we, in reality, are inhabitants of that world. Or rather our real, true selves are inhabitants of that world. Here Plato is relying upon another of his fundamental ideas, that our real selves are our *souls*. It is the soul, or *psyche*, which is the inhabitant of the ideal world, and knows all about the forms of things.

Unfortunately, in the material world our souls are trapped in prisons of corruption – the charnel houses of our fleshly bodies. In these circumstances it is possible for the soul to become confused and deceived by the senses, and to continually make errors. But, providing we subdue the temptations of the flesh and allow ourselves only to be guided by the light of reason, we can still arrive at knowledge and truth.

In fact, Plato believed that all education was merely a question of reminding people what they already know. It is not possible to convey new information, only to remind the soul what it already knows (but has forgotten since being trapped in the body). The clearest exposition of this view occurs in Plato's dialogue known as the *Meno*. Here, Plato shows how Socrates (always the principal character in Plato's 'Dialogues') manages to get an uneducated slave boy to construct a square with twice the area of a given square (try this for yourselves!). Socrates does this merely by asking the boy questions, and so prompting him to remember things which he (i.e. his soul) already knows.

To get back to our problem of how we know about tables without being able to define them: our souls remember the ideal table and recognize physical tables as imitations of that ideal. And the same applies for all the other things we encounter in the material world – even abstract phenomena such as justice. In the real world, the higher realm of the Ideas or the Forms, nothing ever changes and so, if we could gain access to our knowledge of this world, we could know the essential relationships between things, and other aspects of the truth.

Box 2.1 The Platonic 'dialogue'

Plato's books are nearly all written in the form of a dialogue between various characters – like a script for a play. Usually, Plato's own views are represented by Socrates (469–399 BC), a leading philosopher who was also Plato's mentor. Various opposing philosophical views, or sometimes unconsidered common-sense views, are represented by other characters in the various dialogues. Usually, the dialogue is known by the name of Socrates' main interlocutor: *The Meno*, or *The Phaedo*, though there are exceptions, such as *The Symposium*, *The Republic*, and *The Laws*. The format allows Plato free reign to present opposing views as fully, or as incompletely, as suits him, and also enables him to show the consequences of particular philosophical views in a way that might not be apparent from a summary account of that view. This became an influential literary form, favoured by a number of subsequent philosophers.

Now, perhaps you can see something familiar in all of this? Plato's notion of an immaterial soul, a resident of a higher, perfect world, trapped temporarily in corrupt flesh, smacks very much of the Christian view of the soul, does it not? Plato discusses the nature of the soul in a dialogue called the *Phaedo*. The discussion also includes an account of what happens to the soul after death. Again the similarities with Christian doctrine are very marked. These similarities are by no means coincidental, and since Plato died over three hundred years before Christ, Plato must be the source of these ideas. The fact is that Plato's philosophy profoundly affected, and was taken up by, the majority of the Early Church Fathers, but especially Clement of Alexandria (150–216), Origen (185–254), and Augustine of Hippo (354–430). So that, although there is no mention of the incorporeal soul in the New Testament, this Platonic notion became a major aspect of Christian teaching thanks to the residual Platonism of the first theologians to forge Christian doctrine.

It is perfectly clear, then, that Plato's doctrines were seen as highly useful for religious thought in Western civilization, but what about their relevance to scientific thinking? It might look as though Plato's philosophy was inimical to an understanding of the *physical* world. But this is not the case.

Although Plato himself did not have much to say upon scientific matters – like his mentor, Socrates, he was much more interested in moral and political issues – his work did implant the suggestion that the best way to understand the physical world was by analysing it in logical and mathematical terms. His most immediate impact here was on the science of astronomy. For Plato the heavenly bodies, stars, planets, etc., being heavenly, were the least corrupt part of the physical world. The heavens ought to be perfect, and therefore unchanging – I say, 'therefore unchanging' because the rational assumption was that if something was perfect, if it changed it could only change for the worse. The perfect heavens ought, therefore to be unchanging. But they noticeably underwent changes every night. In fact these changes are wide-ranging, complex, and to the casual observer bewildering.

You might imagine, therefore, that Plato would have come to the conclusion that his initial starting point – the assumption that the heavens are perfect and unchanging – must be wrong. But, as a rule, philosophers do not give up their cherished ideas quite so easily. What Plato did was to put out a challenge to contemporary geometers and astronomers to 'save the appearances' of the heavens. In other words, Plato wanted astronomers to try to determine how the seemingly baffling motions of the heavenly bodies might be explained in terms of a few geometrical principles. To preserve the perfect and unchanging nature of the heavens as much as possible, Plato assumed that the heavenly bodies were perfect spheres rotating concentrically around the central Earth. The image was not one of planetary bodies moving in orbits around the Earth, but of transparent (and therefore invisible) spheres *staying in one place* (centred upon the Earth) and merely rotating upon their own axes. The blob of light which we know as the planet is regarded as a marker on the face of the sphere, to enable us to perceive the sphere's rotation. Again, in pursuit of a heaven of minimal changes, it was assumed that the spheres would not speed up and slow down, but must always be rotating at a uniform speed.

In fact, it had been clear to astronomers for a long time that the heavenly bodies did slow down and speed up, and in some cases even stopped and went backwards for a while before continuing in their usual direction, and that they sometimes varied in size and brightness. Even so, astronomy seems to have become an important aspect of the work of the Academy. Eudoxus of Cnidus (*c.*

390–c. 337 BC) and Heraclides of Pontus (c. 390–c. 339) worked alongside Plato, responding to his challenge in a highly fruitful way.

Plato did not confine his thinking about the physical world to astronomy and cosmology. Generally speaking, he held out the hope that, in spite of its changeability, the physical world was capable of being understood in geometrical terms. He even subscribed to an atomist version of Empedoclean four-element theory, in which the atoms of each of the elements had characteristic shapes. Specifically, the shapes were four of the five regular solids (that is to say, the only solid geometrical bodies in which all faces were alike – so the cube has eight square faces, the tetrahedron four equilateral triangular faces, and so forth). Atoms of earth were cubes, fire was composed of tetrahedral atoms, air of invisibly small octahedrons and water of icosahedrons.

From this distance it looks as though Plato wanted his atoms to be both material bodies and, at the same time, abstract geometrical entities. The three-dimensional atoms did not simply have triangular shaped faces, rather they were composed of triangles, and the triangles could disassociate and rearrange themselves in transmutation of one substance into another (so, the twenty triangles of an icosahedral atom of water could separate and reform as two atoms of air and one of fire). The atoms, therefore, have to be envisaged not as solid bodies, but as hollow shapes defined by the laying together of two dimensional triangles (or squares in the case of earth).

It is not clear what the actual triangles (or squares) were supposed to be made of – as I say, it looks as though Plato is trying, perhaps somewhat desperately, to ensure that physical entities like atoms can still be amenable, like the motions of the heavenly bodies, to geometrical analysis. In trying to ensure this, however, his atoms begin to look suspiciously non-physical, and perhaps more suited to the realm of *Ideas*, or *Forms*. But after all this is not so surprising from a writer who believed that the real world was an unchanging immaterial world, distinct from the physical world, and revealed not by the senses but by reason. It is worth noting also that Plato describes his cosmology, with its geometrical atoms, in a dialogue known as the *Timaeus*. The cosmology is described in detail to Socrates by the character known as Timaeus, who is introduced as a Pythagorean. It seems clear that Plato was willing to have his own cosmology seen as broadly Pythagorean, and it is hardly surprising, therefore, that his atoms should be supposed to constitute real bodies in the material world, and yet the atoms themselves are not made of matter but of two-dimensional abstract geometrical entities.

The important thing to note about Plato's physics, nonetheless , is that he wants it to exemplify his conviction that truth can only be arrived at by the use of reason (especially by logic and geometry), and as a result, it led subsequent generations of like-minded thinkers to suppose that the world, for all its apparent messiness, can be understood by turning it into abstract and mathematical conceptions.

Aristotle, champion of common sense, and the problem of change

Unlike Parmenides and Plato, who wanted to diminish the importance of the physical world, even to the extent of denying its reality, Aristotle embraced the material world with all its changeability. He was a thoroughly down-to-earth natural philosopher who believed that primarily we learn from the use of our senses, and by interpreting the sensory information in the light of our so-called 'common sense'. Aristotle was generally suspicious of any philosophy which seemed to deviate widely from what was accepted by the common man or woman, and for that reason could not even accept atomism, much less a Platonic higher realm of Ideas. It is surely significant, therefore, that he subscribed to the Empedoclean belief in the four elements – a philosophy which in itself drew upon common folk beliefs about the make-up of the physical world (earth, sea, wind, and fire).

Now Aristotle's philosophy is encyclopaedic in its scope, highly complex, and not easily summed up. The important thing for our purposes is to consider how Aristotle approached the problem of change. Although a reductionist of sorts, Aristotle clearly did not believe that the complexity of the world could be reduced to a single explanatory principle, such as water or air, much less that its existence could be denied in favour of an unchanging One or an unchanging world of Ideas. Aristotle's approach was completely different. Instead of trying to explain change away, Aristotle tried to draw up classificatory schemes to codify all the different kinds of change, or changeability, there are in the world.

Consider, for example, his so-called 'four causes'. Something of a misnomer for us, these are best understood as 'four *becauses*' – that is to say, four possible answers to the question as to why a thing is the way it is. This is typical of Aristotle's approach. If we ask why something is the way it is, or ask what makes a thing what it is, Aristotle says that we might be expecting four, and only four, different kinds of answer – depending upon our particular concern. The classic example is provided by a clay pot. If someone points to a clay pot and asks me why this pot is the way it is, they might mean: why does it look so different from a similar thing made of glass or of wood. In which case, I would say to them, 'because it is made of clay'. Alternatively, they might be wondering why it is the shape it is, and not just a round ball or a flat pancake; in which case I would tell them that it is designed to securely hold a fluid without spilling. Or, perhaps they are wondering how it came to be put into that shape. I could then respond by telling them that potters can easily manipulate clay into shapes like this. The only other thing they could possibly mean by the question, according to Aristotle, is what is the purpose of it being this way: 'I understand it is a pot, but why make pots?' The answer here might be 'to carry water' (as we already noted above), or it might require notice of a deeper purpose – we need to drink to survive and so we need pots to carry water with us if we intend to travel afar.

These four different kinds of answer illustrate Aristotle's four causes: the material cause for something being the way it is, the formal cause (its shape or form), the efficient cause (the thing that shaped it or made it), or the final cause (its end, or purpose). The four causes apply to everything in the world, not just man-made objects. A wolf is the way it is, and not like a tree, because it is made of flesh and bone. It is shaped the way it is for fast running and to be able to kill its prey. The efficient cause of a wolf, as of all living creatures, are its parents, and the in-built principle of growth emergent somehow from the egg and the seed which the parents provide. It is not always clear what the purpose of a living creature might be; perhaps to ensure the propagation of its species? Perhaps to ensure that no ecological niche in nature is left unfilled? Perhaps in the case of wolves to ensure that the populations of other, smaller, creatures do not overrun the vicinity, and strip away too much vegetation as they feed?

Aristotle also lists ten categories of existence – again as a way of providing a systematic classification of the variety of things in the world. The categories are supposed to provide an exhaustive list of the ways by virtue of which a thing could be said to exist: by virtue of its substance, quantity, quality, its relation to other things, its place, time, position, state, its action, or its affection (which means by virtue of its response to the action of other things – a canvas can be painted, and so its existence is affirmed by virtue of the fact that it has an affection to paint, or can be affected by paint).

It sometimes looks as though Aristotle's main concern is with how we talk about things; as though he is not so much concerned with the nature of the world as with how we describe it. This reflects his concern with common sense and with what ordinary people think, or say. He wants his philosophy of nature to be compatible with the way ordinary people talk about the world, and its various familiar features. This is one reason why he will not accept that Plato's realm of immaterial Ideas is the real world. When a woman in the market place talks about the real world, she means the material world, and Aristotle wants to provide a philosophical underpinning for her concept of reality, not Plato's.

As far as Aristotle is concerned, however, his analyses lead him to definite conclusions about how things are, not just how we say they are.

An obscure but important aspect of this can be seen in his distinction between the potential and the actual. This was initially invoked in order to dispose of the Parmenidean objection that something cannot become something it is not, because this implies it becomes something that does not exist, and so in becoming what is not, it ceases to exist. Aristotle simply insists that when one thing habitually changes into another, that thing must have had the potentiality within it to become the second thing. An acorn's existence is actual, but it is already an oak tree in potential. This is not just a way of speaking about acorns and other things which undergo change in order to cheat Parmenides. For Aristotle it reflects a genuine aspect of the nature of things. A stone high up on a ledge has the potential to become a deadly projectile if it falls. Scientists today still talk of the potential energy of things in a static system, and compare this to the kinetic energy of the system when its stasis is no longer maintained. In Aristotle's philosophy the idea of potential was also bound up with his concept of final causes – everything has an end or purpose, although in some cases the purpose remains part of its potential, not part of its actuality. The final cause of an acorn is to produce an oak, and this remains true of the acorn even if it never germinates.

Another seemingly wordy distinction of Aristotle's which makes real claims about the nature of things is the distinction between essential and non-essential properties. You can paint a chair all sorts of different colours but it remains a chair – only the non-essential property of colour changes. You can't, however, set fire to a chair, reduce it to ashes, and claim that it is still a chair – now you've changed an essential property. The colour of a chair is not what makes a chair a chair, but a chair must have a characteristic form or shape making it suitable for sitting. In different cases, however, colour might very well be an essential property of a thing (our contemporary science of spectroscopy, for example, is based on the premise that different chemical substances can be made to emit a spectrum of colours characteristic of that substance – Aristotle couldn't have known that but he would presumably have conceded that in the case of a rainbow, its characteristic colours are an essential, defining, property).

The distinction between essential and non-essential properties led Aristotle to develop another concept – the *form* of something. Things can meaningfully be said to remain unchanged when their form is unchanged. An old ship which has been repaired and patched up so many times that none of its original fabric remains can still be said to be the same ship because it has maintained its form throughout. Any individual physical entity, according to Aristotle, must be made of both matter and form in combination (matter cannot be conceived of unless it has some physical shape, and it makes no sense to talk of the shape of something immaterial). But since matter was considered among the Ancients (again showing their reductionist tendencies) to be one and the same, so that one bit of matter was indistinguishable from any other bit, it followed that the uniqueness of any given object was due to its particular form. In a way Aristotle approaches the views of Thales, or Anaximanes here. Matter, like Thales's water or Anaximenes' air is always the same, but it can manifest itself in all sorts of different ways by taking on the appropriate specific form.

Notice, however, that Aristotle's concept of form is specific to individual objects – he rejects Plato's notion of the universal or Ideal Form. According to Aristotle, Plato makes a mistake when he concludes that because all tables have something in common (which makes them all tables), there must be an actual Ideal Table which singly embodies that common attribute. For Aristotle, every individual table has its own form, which is unique to it, even though it shares many characteristics with the forms of every other table. Providing we bear this in mind, we can talk of the form of a horse, or the form of a beetle, and talk in general terms as to how they differ (the form of a horse has four legs, the form of a beetle has six, etc.), but this is only loose talk – in reality there are only individual horses, or beetles, or tables, with their individual specific forms.

This concern with form, and its role in enabling us to understand the nature of things, led to an important consequence. Aristotle and his followers were unique among the Ancient Greek philosophers in paying serious attention to the variety of creatures in the world. Aristotle himself made many detailed studies of animals, describing their forms in universal terms, but also descending to details of their anatomy and describing how the forms of their various parts contributed to the form of the whole. Aristotle noted that animals do not have both horns and sharp teeth, that carnivores have sharp teeth, but herbivores have flat teeth, and other standard aspects of animal forms. Aristotle's teaching gave rise to the important notion, still an important aspect of physiology and anatomy, that form reflects function. Aristotle's successor as head of the Lyceum, Theophrastus (died *c.* 287 BC), extended his master's work by making similar contributions to the taxonomy of plants, and to the relationship between the forms and functions of their parts.

Aristotle's characteristic approach – seeking to classify and codify the variety of phenomena – had another very important, and useful consequence. He set down in a formal way the different kinds of logical reasoning and the specific rules pertaining to each of them. Where Plato was content to demonstrate philosophical argument in action, so to speak, in his dialogues between Socrates and other thinkers, Aristotle spelled out in precise detail, over a series of books, what constituted a valid argument and what did not. This in itself was to ensure that Aristotle was always regarded by subsequent generations of thinkers as a philosopher of the utmost importance. Subsequently, his logical works were often taken together under the collective title of the *Organon* (*Instrument*).

* * *

Aspects of the thought of both Plato and Aristotle have been repeatedly embraced, and adapted by subsequent generations of thinkers to a far greater extent than any other secular thinker. One way of understanding the usefulness of their thinking for subsequent natural philosophers is to see them as representatives of two strands of thought in science. Plato inaugurated the belief that the best way to understand the world is to analyse it in terms of supposed underlying geometrical and mathematical regularities, or by using a rationalist approach based on abstractions from the messy reality. Aristotle took a much more prosaic and pragmatic view. The world could be understood merely by observing what you see. The way to uncover patterns in things is simply to classify them according to regular principles (ten categories of existence, four elements, four causes, matter and form, classification of animals, etc.).

Underlying the views of both Plato and Aristotle, however, is the unifying belief, derived from earlier Greek thinkers – that the natural world, in spite of its seemingly baffling ever-changing complexity, is a regular, ordered cosmos, governed by natural laws.

The subsequent history of philosophy saw the development of new schools of thought, most notably Epicureanism, Stoicism, and what is now designated as Neoplatonism. Epicureanism was developed by Epicurus (341–270 BC), perhaps best known for his moral philosophy (although this is usually mistakenly equated with hedonism), his natural philosophy revived and extended atomism. The Stoics, who included Chrysippus (c. 280–c. 207 BC), and Posidonius (c. 135–c. 51), are also best known for their prescriptions as to how to live one's life, but they developed an essentially vitalistic natural philosophy in which all things were connected by an all pervading cosmic *pneuma* (breath or spirit). Ancient Neoplatonists, Plotinus (AD 204–270), Iamblichus (c. 245–c. 325), Proclus (412–485), and others, would have seen themselves simply as Platonists, but their version of Platonism, emphasizing the philosophical theology of Plato and extending it into more mystical realms, seems to us so different from Plato's original thinking that we regard it as necessary to distinguish them as Neoplatonists. Again, the emphasis of this philosophy was on how to live a good life, although its distinctive world-view did include assumptions about the nature of the cosmos.

Each of these systems of philosophy will appear later as influences upon important figures in the historical development of science, but the importance of Ancient Greek thought after Plato and Aristotle was not confined to systems of philosophy. The Hellenistic period, as it is known, saw many achievements in the sciences, and especially mathematics, which add to our sense of Ancient Greece as a hotbed of scientific accomplishment.

In astronomy, Aristarchus of Samos (310–c. 230 BC) suggested that the Earth was in motion around the Sun, and calculated the distances between them. Eratosthenes (c. 276–c. 195 BC) developed an ingenious (and remarkably accurate) method of measuring the diameter of the Earth, and proposed the addition to the calendar of an extra day every fourth year (although this wasn't adopted until the establishment of the Julian calendar in 45 BC). Although the importance of their work was not recognized immediately, it was different for Hipparchus (c. 190–c. 120 BC), who discovered the fact that the equinoxes move slowly around the celestial equator (the so-called precession of the equinoxes). This was immediately recognized to be immensely important for the understanding of the heavens. It has even been forcefully argued that Hipparchus' discovery partly inspired the founding of the new religion known as Mithraism, which flourished vigorously before Christianity. Last but by no means least, Claudius Ptolemy (AD 90–168), drawing upon all previous astronomy, created a synthetic account of how to calculate the movements of the heavenly bodies which dominated astronomical science until the sixteenth century.

In mathematics, Euclid (c. 325–c. 265 BC) systematized geometry and arithmetic in what came to be seen as one of the most useful books of all time, his *Elements*. Apollonius (c. 262–c. 190 BC) was nonetheless able to extend geometry in his account of conic sections (the geometry of shapes created by slicing through a cone at different angles, such as the ellipse and the parabola). Archimedes (c. 287–c. 212 BC) was not only a consummate mathematician but also developed statics and hydrostatics, and other aspects of what we would now think of as engineering. Hero of Alexandria (AD 10–70) similarly developed hydrostatic and pneumatic devices which were capable of producing dramatic effects, although these were mostly used only for spectacle and entertainment. Herophilus (335–280 BC) and Erasistratus (304–250 BC) made great advances in the understanding of anatomy, including the anatomy of the brain. Galen (AD 129–c. 200), personal physician to the Roman Emperor and Stoic philosopher, Marcus Aurelius (121–180), drew upon previous theories of medicine to develop a synthesis which dominated medical theory and practice right through to the seventeenth century.

It should be clear from this and the preceding chapter that, for all the incommensurable differences between the culture, and the thought, of Ancient Greece and our own times, something recognizable to us as scientific thinking did emerge there, and made remarkable progress from the sixth century BC to the second century AD. It might be supposed that the Ancient Greeks had stepped a few rungs up a ladder of progress which from now on simply led inexorably upwards. But it was not so.

FURTHER READING

W. K. C. Guthrie, *The Greek Philosophers: From Thales to Aristotle* (London and New York: Routledge, 1968).

Paul T. Keyser and Georgia L. Irby-Massie (eds), *Encyclopedia of Ancient Natural Scientists: The Greek Tradition and its Many Heirs* (London and New York: Routledge, 2008).

David C. Lindberg, *The Beginnings of Western Science: The European Scientific Tradition in Philosophical, Religious, and Institutional Context, Prehistory to A.D. 1450* (Chicago: University of Chicago Press, 2008).

G. E. R. Lloyd, *Early Greek Science: Thales to Aristotle* (London: Chatto & Windus, 1970)

G. E. R. Lloyd, *Greek Science after Aristotle* (London: Chatto & Windus, 1973)

G. E. R. Lloyd, *Magic, Reason and Experience: Studies in the Origins and Development of Greek Science* (Cambridge: Cambridge University Press, 1979).

David Ulansey, *The Origins of the Mithraic Mysteries* (Oxford: Oxford University Press, 1991).

3

From the Roman Empire to the Empire of Islam

The achievements of the Ancient Greeks did not, alas, represent the first steps on a ladder of knowledge extending ever upwards. As far as philosophical and scientific thought was concerned, the Greeks provided a high point of intellectual achievement in the Ancient world, succeeded by a period of much lower intellectual standing. The heights achieved by the Greeks were not to be matched for many centuries. In the Western tradition it was not until the period known as the Renaissance, beginning in the fifteenth century, that knowledge and understanding of the natural world began to rival that of the Ancient Greeks, although, as we shall see, the civilization of Islam had already begun to equal, if not exceed, the Greeks a number of centuries before.

But let us begin with a brief look at what happened after Plato and Aristotle. Plato's thought was considered to be supremely enlightening and useful in his own lifetime and way beyond. His school, the Academy, survived for 900 years, until it was forcibly closed down by the Christian Roman Emperor, Justinian, in AD 529. If it is true that the whole history of Western philosophy is merely a series of footnotes to Plato, we can see it in the fact that subsequent Greek philosophers developed their ideas either as extensions or refinements of Plato's (e.g. Stoics, Sceptics, and later, Neoplatonists), or in opposition to Plato (e.g. Aristotelians and Epicureans).

Aristotle was far less popular amongst subsequent generations of Greek thinkers. Perhaps because he did not offer a simple 'key' to understanding the nature of the universe, the way other philosophical systems did (Plato in terms of the Ideal World, Epicurus in terms of atoms, etc.). Aristotle, remember, offered typologies or classificatory schemes to enable us to consider all the possible facets of any given problem and then would try to deal with each one in their own terms. This evidently was not as popular as providing a philosophical 'hook' on which to hang all phenomena, and all problems. But Aristotle was going to come into his own later, and thanks to his works on logic and the forms of argument, his philosophy was never going to completely disappear from view.

We need not go into the details of later Greek philosophy here, but suffice it to say, that Stoicism, Epicureanism, Scepticism, and Neoplatonism proved to be effective rivals to one another. Of these, the most successful (perhaps testifying once again to the dominating role of Plato and his legacy) was Neoplatonism, the development of which, after it was first systematized by Plotinus (AD 204–70), would have major intellectual consequences.

The moral and political, and often religious, implications of these philosophical movements continued to attract adherents after Greek civilization went into decline and the Mediterranean saw the ascendancy of the Romans. Although the Romans numbered many brilliant engineers, architects, and other pragmatic thinkers among their ranks, they never produced an original philosophical thinker. Arguably the best philosophical minds among the Romans were Lucretius (c. 99–55 BC), who was an Epicurean; Seneca (4 BC–AD 65), and the Roman Emperor Marcus Aurelius (AD 121–80), who were both stoics, and Cicero (106–43) who was a sceptic. Of these, only Lucretius' atomism made a contribution to attempts to understand the natural world.

Over the centuries of Roman domination, therefore, natural philosophy went into decline. What interest there was in natural phenomena was evidently satisfied by compendia or encyclopaedias of facts, such as the *Natural Questions* of Seneca, and the *Natural History* of Pliny the Elder (c. AD 23–79). Aristotle's works, presumably because of the way they were written, seemed to provide much of the material for these compendia.

The Roman Empire began to divide into two halves from AD 284, when Diocletian (244–311) recognized that the Empire could not be ruled by one administration. The Western half maintained its capital in Rome, but the Eastern half established its capital in the newly built Constantinople (named after the first Christian Emperor, Constantine, *c.* 272–337), which is now Istanbul. Although the Eastern half of the Roman Empire, where the *lingua franca* was still Greek, had continuous access to ancient Greek writings, the Western half, centred on Rome, soon lost direct access to Greek sources and had to rely on the derivative compendia and encyclopaedias written in Latin. The speculative and metaphysical aspects of Greek thought disappeared even in educated circles. This was especially true after the Western half of the empire collapsed in the fifth century AD, after repeated invasions by Vandals, Visigoths, Ostrogoths, and other 'barbarians'.

If there was a counter to the lack of interest in natural philosophy among the Romans it was to be found among the early Christians. Augustine, arguably the greatest of the Early Fathers of the Church and certainly one of the major shapers of Western thought, insisted that Christians should learn at least the basics of Greek natural philosophy. He was evidently concerned that Christians should not open themselves up to ridicule by taking the Bible too literally. He was perhaps thinking of Lactantius (*c.* 250–*c.* 325), another Early Father of the Church, who had declared the Earth to be flat on Scriptural grounds. The Ancient Greeks had long since established beyond doubt that the Earth was a sphere, and any educated person reading Lactantius would have simply thought him an ignoramus. It is something of an irony that in our day Lactantius's notion of a flat Earth is all too often taken to be what everybody believed until the early modern period. In fact, Lactantius's flat-earthism was deeply embarrassing to more highly educated Christians, such as Augustine. This aspect of ancient Greek wisdom was never lost. When Christopher Columbus sailed out across the Atlantic ocean he was not concerned that he might sail off the edge of the world, merely that the Earth might be so big that he and his crews would run out of food and water before they reached China (as they might have done if the Americas had not been there to save them).

Augustine was a genuine admirer of Platonic thought and he was not alone among leading churchmen to think this way. Even at a humbler level within the Church, monasteries often included a library which was stocked not just with Christian writings but with the philosophical, historical, and literary works of Ancient Greek and Roman authors. Nevertheless, the fall of the Roman Empire can be seen as marking the beginning in Western Europe of the period which used to be called the Dark Ages, because the 'lamp of learning' burned very low throughout this period. Although this term has fallen out of favour with medieval historians it does serve to show the marked difference, as far as natural philosophy was concerned, between this period and (as we shall see in the next chapter) the High Middle Ages.

The only sources of knowledge about the natural world in these 'dark' ages were the Roman compendia, which by now were all too often compendia based on earlier compendia. If natural philosophy can be said to survive at all, it is maintained at a very debased level by a handful of monks scattered through European monasteries. The fact is that those who wished to pursue higher learning at this time, were hampered not only by socio-political circumstances which did little to allow, much less encourage, contemplative thought, but also by the dearth of philosophical writings available to them. The most useful sources from the point of view of speculative philosophy were a commentary on Plato's *Timaeus* by Chalcidius (*fl. c.* fourth century AD), a Neoplatonic commentary on Cicero's *Dream of Scipio* (actually Book VI of Cicero's *Republic*) by Macrobius (*fl.* AD 400), a survey of the seven liberal arts which were the focus of Roman education by Martianus Capella (*fl.* 410–439), *The Marriage of Philology and Mercury*, and the Latin translations of some of Aristotle's logical works which had been undertaken by Boethius (*c.* 480–524). Subsequent generations of would-be scholars had little to do but to re-work this material, often adding to it their own Christian-inspired glosses. Cassiodorus (*c.* 488–575), for example, produced an *Introduction to Divine and Human Readings*, which included material on the seven liberal arts to represent human readings, and Isidore of Seville (560–636) compiled a vast encyclopaedia, the *Etymologies*, which was intended to sum up all learning (ancient and modern) for the education of churchmen.

Although the roster of significant thinkers is very small during this time, there were two outstanding thinkers from the British Isles. The Venerable Bede (*c.* 673–735), who spent all of his life in the monastery at Wearmouth in Northumberland, is generally agreed to have been the most learned man of his time. Although chiefly known for his work as a historian (*Ecclesiastical History of the English People*, 731), he also contributed to the encyclopaedic tradition with his *De natura rerum* (*On the nature of things*). That he was capable of original thought is evident, however, from his *On the Reckoning of Time* (*De temporum ratione*, written about 723). Bede here resolved a controversy over the annual dating of Easter, which required mastery of astronomical calculation, and while he was at it, he also reported observations which added to contemporary knowledge of the movement of the tides. John Scotus Eriugena (*fl.* c. 845–c. 870) was born and raised in Ireland but is believed to have left to escape increasingly frequent Viking raids, some time after 836. By 850 he was a prominent member of the court of Charles the Bald, King of the West Franks (who reigned from 840–77), most probably as royal physician. He was evidently fully conversant with those aspects of Platonic philosophy which had filtered through to his day and his most original work, the *Periphyseon* or *De divisione naturae* (*On the division of nature*, completed about 866), is essentially an attempt to reconcile the Neoplatonist doctrine of emanation (in which it is supposed that space, light, and ultimately matter are all emanations from the substance of God) with the Christian account of the Creation. The originality of John's work derives from the fact that he was willing to base his arguments solely on reason (showing his Platonic antecedents), which he saw as 'prior by nature' to authority, even if authority was prior in time (he meant Scriptural authority). John's emphasis upon reason was significant for subsequent developments in the history of philosophy. By the tenth century, Western Europe was ready to escape the doldrums of the Dark Ages and there was to be a major revival of learning, and of philosophical thinking. John's *De divisione naturae* was therefore appropriated throughout the High Middle Ages for showing subsequent generations how to bring rational approaches to bear upon the authoritative writings of earlier times.

There can be no denying that from the fifth through to the tenth century natural philosophy barely held on as an aspect of learning in Christendom, hardly being recognized as a worthwhile pursuit. Those who were intellectually gifted either found no opportunity to shine at all in the social and political climate of the time, or invested their intellectual energies in theology, or perhaps the law, or government administration. In spite of the impressive efforts of a handful of churchmen, it is possible that the legacy left by the Ancient Greeks night have been lost forever. The situation was saved, however, by the emergence of a new civilization.

* * *

The civilization of Islam, inspired by the teachings of Mohammed (570–632), began to flourish from the seventh century onwards. The Arabs, invested with a new sense of purpose and self confidence, conquered large parts of the Middle East, North Africa, and eventually Spain and Sicily. More importantly, as far as the history of science is concerned, science began to flourish as it had not done since the time of the Ancient Greeks, and in many respects the science of Islam came to exceed that of the Ancients.

This immediately raises the question as to why – why did the Arabs and the other ethnic groups who contributed to the rise of Islam take such a keen and fruitful interest in the natural sciences?

The evidence suggests that the initial interest stemmed from concerns with pragmatic arts and sciences such as medicine, astronomy, assaying and other chemical procedures. Medical knowledge is always useful for any civilization, but Islam had a particular need for astronomers. Enjoined to face Mecca while praying, Muslims who habitually travelled had to be able to work out the direction of Mecca. Similarly, they had to be able to establish the five times of the day stipulated for praying. Furthermore, as the Arab empire began to establish itself there was a need to mint new imperial coinage to replace the Byzantine coinage in use in the West of their new territories, and the Sasanian, or Persian, coinage that was in use in the East.

Among those who first began to develop the necessary astronomical skills to establish the direction of Mecca from any point, or the necessary skills in assaying gold, it might have been deemed sufficient to do just what was required and no more. There was a similar impetus to develop the mathematical skills to work out the practical consequences all too often arising from the highly elaborate rules of inheritance laid down in Islamic law. It was not long, however, before other educated men began to take things further. It seems likely that the first stimulus to go beyond just what was required for strictly Islamic purposes was nothing more than a competition for jobs in the bureaucracy of the nascent empire.

Just as the new empire wanted a new coinage, so it wanted an Arabized bureaucracy, and inevitably many of the former bureaucrats, under the Byzantine or the Persian administrations, would have been replaced by Arabs. In a bid to win back their former positions – in some cases even high positions of government – those who knew Greek could and did turn to Ancient Greek sources, which they had previously ignored, to learn more about the technical backgrounds to the simple knowledge of astronomy, arithmetic, assaying, and medicine that was in use among the Arab incomers. In turn, of course, they began to pass this more profound knowledge on to their own children.

The Arabization of the administrative systems in former Byzantine and Persian territories and the minting of the new coinage was begun by Abd al-Malik (646–705), the fifth caliph of the Ummayad dynasty. By the time the Ummayads were deposed by the Abbasid dynasty in 750, the promotion of Ancient Greek wisdom was well under way. The age of the Abbasids is often referred to as a 'golden age', and a succession of caliphs, al-Mansur (712–775), who established Baghdad as the new imperial capital, Harun al-Rashid (763–809), and al-Ma'mun (786–833), encouraged and supported the recovery of Ancient Greek writings and their translation into Arabic. Some Islamic thinkers found patronage from enlightened caliphs like these, others from lesser but still high-ranking government officials. It is important to note, however, that the thriving of non-religious intellectual endeavours in Islam was always almost entirely dependent upon a fairly small number of enthusiastic patrons. The significance of this will emerge later.

Perhaps because of the initial concern with pragmatic techniques in the Arab empire, the new converts to Islam did not simply recover Ancient Greek writings for their own sake, but went on to make important advances, in many cases extending their knowledge well beyond anything known to the Greeks. Greek knowledge of assaying and other chemical techniques became what was virtually a new science – still known by its Arabic name, alchemy. Similarly, the use of arithmetical methods to calculate inheritances developed into the new domain of algebra – a kind of problem-solving that had never occurred to the Ancients. They also made great advances in trigonometry, especially spherical or projective trigonometry (again partially stimulated by the need to know how to face Mecca on a spherical Earth). Similarly, they were able not only to master Greek astronomy but also to make significant improvements upon it. For them, as for the Ancient Greeks, this was bound up with the interpretative divinatory art of astrology, and this too was greatly expanded by the Arabs.

The major aspects of these advances would later be taken up by European thinkers and would lead, comparatively rapidly, to what is now called modern *Western* science. If it came to be known as Western science, however, it was undeniably built on *Eastern* foundations (see Box 3.1). The sciences developed in the civilization of Islam were to have a major impact upon the European Scientific Revolution, beginning in the sixteenth century. It is somewhat ironic, therefore, that the power of Islamic thought over European thinkers when they first began to emerge from the 'Dark Ages', in the eleventh and twelfth centuries, manifested itself not in a new enthusiasm for alchemy, or astronomy, much less for algebra or trigonometry, but in an almost overwhelming excitement for Aristotelian natural philosophy. People generally hear what they want to hear, and it seems evident that in the Middle Ages, the European thinkers who became acquainted with Islamic learning saw little to interest them in these pragmatic sciences (the exception, of course, as ever, was medicine, which was immediately taken up by Western thinkers), but they were immediately captivated by the natural philosophy of Aristotle.

We will come back to the reasons for this in the next chapter, but first it is important to understand how *Aristotelian* natural philosophy came to loom so large in the Islamic sources.

The would-be pioneers of Islamic thought were assiduous in recovering Ancient Greek writings and translating them into Arabic. By the ninth century there was a deliberate effort in Baghdad to translate a wide range of Greek works into Arabic. Inevitably, therefore, they included not just medical works, or astronomical works, or other writings on pragmatic subjects, but also the more philosophical works of the great thinkers, such as Plato and Aristotle, and the leading lights of the other philosophical schools. Among the most immediately influential of these were the Neoplatonists. Not only had the Neoplatonists been the most successful school of philosophy in the later Roman Empire, but they were still actively represented in Persia, because the teachers at Plato's Academy had moved to the Persian imperial court in 529, when the Roman Emperor Justinian, forbade the teaching of pagan philosophy.

As the name suggests, the Neoplatonists were closely affiliated to Platonism. Indeed, the designation Neoplatonism is ours, not theirs – they would no doubt

Box 3.1 JUST A FEW LEADING ARAB THINKERS...

Jabir ibn Hayyan (known to the Latin west as Geber) (*c.* 730–*c.* 810). Arab, or Persian, or possibly even merely legendary, he is renowned as a great alchemist. He emphasized the necessity for the experimental method in alchemy and was the discoverer of the mineral acids and other chemical substances.

al-Khwarizmi (known to the Latin West as Algoritmi) (*c.* 780–*c.* 850). Persian mathematician based in Baghdad whose technique for solving what we call quadratic equations, known as al-jabr, gave rise to the new mathematical system of algebra. His book on arithmetic introduced the Indian numerical system to Islam (and then to the West – hence we call them Arabic numerals!), and the decimal system of notation. The word algorithm derives from his name. He also did important work in trigonometry, and in astronomy and geography.

al-Kindi (known as Alkindus) (*c.* 801–873). An Arab working in Baghdad, he was a pioneer in introducing Ancient Greek philosophy into Islamic culture. Involved in the translation of Greek works into Arabic, and in reconciling Greek philosophy with Islamic doctrines. His work in astronomy and astrology led him to develop a theory of rays of influence (*On the rays of the stars*) which was frequently adopted in magical traditions. He also did important work in optics, and in chemistry and pharmacy.

al-Razi (known as Rhazes) (864–932). Persian polymath who made important contributions to medicine. A major figure in the history of alchemy and author of the widely read *Secret of Secrets*, he was also much impressed by Epicurean thought and developed an atomistic theory of matter.

al-Farabi (Alpharabius) (*c.* 870–*c.* 950). Probably Turkish (but possibly Persian) thinker who worked successively in Baghdad, Aleppo, and Damascus. Called the Second Teacher (after the First Teacher, Aristotle), he developed his own philosophy based on elements from Aristotle and Plato and his attempts to reconcile them with Islam. His inventive eclecticism can also be seen in his cosmology which combined Aristotelian theories of causation and Ptolemaic astronomy with Neoplatonist emanation theories to provide a picture of the world system which was physical but infused with deity. He was also significant among Islamic thinkers in seeing the physical world as simultaneously mathematical, and therefore amenable to mathematical analysis and being understood in mathematical terms.

have seen themselves as Platonists, but from our perspective there are sufficient differences to demand a new label. So, the Neoplatonists, seeking to promote their great founding father, Plato, were all too aware of the powerful critique of Plato's views by Aristotle. Furthermore, thanks to the importance of Aristotle's logical works in pedagogy, it was not possible to simply dismiss his views. Accordingly, a major aspect of Neoplatonism was the attempt to reconcile Plato and Aristotle, and to show that Aristotle's critiques were by no means debilitating to Platonism.

Ibn al-Haytham (Alhazen) (965–1039). Persian polymath who lived most of his life in Cairo. Known mostly for his important work in optics and in ophthalmology, but he also made important contributions to astronomy, physics, and mathematics.

Ibn Sina (Avicenna) (*c.* 980–1037). Persian polymath's polymath who wrote many works in a wide range of subjects. His most widely adopted work was in medicine. His *Canon of Medicine* was a mainstay of medical schools in medieval Western universities even into the seventeenth century. It combined traditions of Islamic medicine with Galen's compendious account of Greek medicine and his own experience as a highly successful physician. He was also a major interpreter of the philosophy of Aristotle, especially his metaphysics.

al-Biruni (Alberuni) (973–1048). Persian polymath who worked on mathematics, mechanics, astronomy, astrology, geography, meteorology, and medicine, among other things. He is also credited with pioneering work in the philosophy of science and promoting the experimental method.

Ibn Bajja (Avempace) (d. 1138). Worked in Seville and Granada, died in Morocco. Astronomer and physicist who did important work on theories of motion, which were taken up by Averroes and even Galileo. He was also an important commentator on Aristotle whose interpretations were sometimes embraced by Albertus Magnus and Thomas Aquinas.

Ibn Rushd (Averroes) (1126–1198). Polymathic thinker who lived in Cordoba, and has been seen as the inspiration behind the development of secular thought in Christian Europe – certainly so-called Averroism, in Europe, was equated with extreme naturalism (i.e. explaining all things in entirely naturalistic terms). A major interpreter of Aristotelian philosophy – often simply designated as 'the Commentator' by Western thinkers – he also did important work in astronomy, medicine, and other subjects.

Nasir al-Din al-Tusi (Tusi) (1201–1274). Persian astronomer and mathematician. Said to be the greatest astronomer between Ptolemy and Copernicus, he produced accurate tables of planetary motions based on observations made at his observatory near Maragha (in what is today northern Iran). Inventor of the so-called Tusi-couple which generates linear motion from a combination of two circular motions. Used by al-Tusi to improve on Ptolemaic astronomy and taken up by Copernicus for the same purpose.

As a result, Islamic thinkers had so much trouble distinguishing between what was Platonic and what was Aristotelian in Neoplatonic writings that they even attributed two Neoplatonic works to Aristotle himself. The so-called *Theology of Aristotle* was in fact part of the major work by Plotinus, founder of the Neoplatonists, called *The Enneads*, and the *Book on the Pure Good* (which would later be translated into Latin as the *Liber de causis – Book of causes*) was part of the work by the last great Neoplatonist, Proclus, known as the *Elements of Theology*.

At this point, it is by no means a digression that we find ourselves discussing works of theology, rather than scientific works. Bearing in mind that the great thinkers who are set to make the greatest contribution to philosophy since the Greeks are devout Muslims, it is hardly surprising that they should be excited by the works of the Neoplatonists who essentially extended the religious aspects of Plato's thought, and turned it into a complex theology based on reason, rather than revelation. Given this interest in Neoplatonism, however, there are a number of reasons why Aristotelianism should come to the fore in subsequent Islamic philosophy. The Neoplatonists themselves, especially the later ones, spent as much time discussing Aristotle as they did Plato. Aristotle's works were much more systematic than Plato's dialogues, and they bore titles that indicated what they were about: *Physics*, *Metaphysics*, *On the Soul*, *On Coming to be and Passing Away*, and so forth (compare this with Plato's *Symposium*, *Phaedo*, *Critias*, *Timaeus*, etc.). Furthermore, Plato's works, and the more Platonic examples of Neoplatonic writings, were often overtly theological in their themes and potentially threatening, therefore, to the doctrines of Islam. Aristotle was much less concerned with matters that might be construed as theological and offered highly satisfying intellectual challenges to Islamic thinkers without implications for their religion. It is understandable, therefore, that Aristotle, rather than Plato, should come to be awarded the title of 'first teacher' by Islamic philosophers.

The accolade of 'second teacher' was bestowed posthumously upon the philosopher al-Farabi (*d. c.* 950), about whose life little is known, except that he began his career in Baghdad but moved to Damascus in 942. Al-Farabi's natural philosophy takes its starting point from a cosmology based upon a complex combination of Neoplatonic emanationism (the belief that all things, immaterial and material, are successive emanations – like light, say – from the being of God), Aristotelian theories of causation, and the consummation of technical astronomy which had been achieved by the Ancient Greek astronomer, Claudius Ptolemy (*c.* AD 85–*c.* 165). Taken as a whole, al-Farabi's philosophy was predominantly Aristotelian, particularly in its extolling of the value of Aristotle's logical works, and this set the scene for subsequent Islamic philosophers.

The two most pre-eminent Islamic philosophers were Abu 'Ali al-Husayn ibn 'Abdallah ibn Sina (*c.* 980–1037), a Persian, and Abu al-Walid ibn Rushd (1126–1198), of Cordoba in al-Andalus (now Spain), and although both should be seen as powerful original thinkers, they both took as their starting point, and as their constant guide, the philosophy of Aristotle. It seems fair to say that perhaps their works were taken up to most effect by medieval philosophers in the Latin West. Ibn Sina's name was Latinized to Avicenna, and ibn Rushd's to Averroes, and as such they were frequently invoked, and deferred to, by medieval Western philosophers. Many of Averroes's works were written as commentaries on Aristotle – in which a work by Aristotle would be provided, paragraph by paragraph, with interpolations by Averroes. Throughout the Middle Ages in Europe, Aristotle was often referred to simply as 'The Philosopher', and Averroes was simply referred to as 'The Commentator'.

We need not go into the details of their philosophies here, although it is perhaps worth pointing out that had things been different, instead of being studied today only by specialists in Islamic thought, they might well have been recognized as at least the philosophical equals of Thomas Aquinas (1225–1274), John Duns Scotus (1265–1308), or William of Ockham (*c.* 1287–1347). For our purposes, however, as we try to understand the development of scientific thought, it is important to note that their and their fellow Islamic philosophers' Aristotelianism was to profoundly effect thought in the West as it emerged from the Dark Ages. In broad terms, the Muslim influence can be seen not just in the fact that the Western Christians also focused almost exclusively on Aristotle, rather than any other Ancient Greek philosopher, but also in the fact that Aristotle's philosophy was used to examine, and sometimes to test, theological claims. For the fact is, Islamic thinkers, no less than the Neoplatonists of late Antiquity before them, and their Christian emulators after them, all of whom shared a belief in the power of reason, and of the rational mind's ability to reach truth, were so impressed by the power of Aristotle's thought that they could hardly forbear from the attempt to show how Aristotle's philosophy might be ultimately compatible with the truth as it was told to them by their respective faiths.

Before turning to see how this Arab learning inspired, and shaped, the medieval West, it is worth considering why a pursuit of the historical development of science will soon take us away from the empire of Islam and across to Western Europe. For the fact is, the intellectual heights which the Arabs achieved did not mark the beginnings of an unstoppable progress towards new heights. As with the Ancient Greeks before them, the Arab achievements proved transitory.

The claimed decline of science in Islam is often attributed to a religious backlash against infidel science, particularly after the major religious attack on Avicenna in al-Ghazali's (1058–1111) *Incoherence of the Philosophers*. Alternatively, it is attributed to the destruction of Baghdad by the Mongols in 1258. But both of these events are far too early to be credible causes, since Islamic thinkers continued to make significant contributions to the sciences right into the sixteenth century. Ghazali's *Incoherence of the Philosophers* received a major response, after all, from Averroes in a work entitled *The Incoherence of the Incoherence*. The Mongol invasion of the Islamic empire was undeniably brutal in human terms as well as in terms of the loss of many works of intellectual achievement, but it by no means spelled the end of all Arab science. The fact that one of the leading Arab astronomers, Nasir al-Din al-Tusi (1201–74), not only escaped with his life after the destruction of the fortress of Alamut, but subsequently became the personal astrologer to Hulagu Kahn (*c.* 1217–65), grandson of Genghis Khan (*c.* 1162–1227), and persuaded the Kahn to build him the largest observatory in the world at Maragha, shows that even the Mongols felt the need for expert astronomers. It is perhaps also worth noting that although the major Islamic thinkers often had trouble with contemporary religious authorities, there is no evidence that the decline in Islamic philosophy and science was due to strengthening religious opposition.

The decline seems to have begun in the sixteenth century, and it was at just this time that the unity of the Islamic world came to an end, with the formation of three separate Islamic empires: The Ottomans (*c.* 1453–1920) maintained control of Turkey, the Eastern Mediterranean and parts of North Africa; the Safavids (*c.* 1502–1736) controlled Persia; and further east, the Mughals (*c.* 1520–*c.* 1750) ruled in the Indian subcontinent. It is possible that this led to a weakening of the cultural cohesion of Islam and had a detrimental effect on further attempts to advance learning. If so, it must have been affected by other developments, because there were similar fragmentations in the empire even as early as the tenth century. After the Abbasids overthrew the Umayyads in Baghdad, in 750, a surviving Umayyad prince, Abd ar-Rahman (731–788), was able to flee to al-Andalus and founded Cordoba, which rivalled Baghdad as a major centre of learning and culture in the West. By the tenth century his successor, Abd ar-Rahman III (889–961), was able to declare himself caliph of al-Andalus. Shortly before this, the Fatimid family, who claimed descent from Muhammed's daughter Fatima, had declared their own caliphate, based at Cairo. If anything, these early fragmentations greatly increased the opportunities for patronage of the sciences, so if the later fragmentations of Islam were a cause for the decline of science, something else had obviously changed.

Perhaps the most significant factor lay outside the empire of Islam. The sixteenth century was also the period which saw the discovery by Western Europeans of the New World, and the subsequent voyages of discovery. As well as providing new markets and untapped natural resources in the New World, the discovery of the way around the south of Africa to south Asia and the Far East meant that Europeans no longer had to cross through Muslim territories to trade with further parts of the world, but could circumvent them. Islamic territories lost their commercial power and became consumers of European wealth, rather than suppliers of it. Similarly, after centuries in which Islamic thinkers were in possession of knowledge in advance of that of the Europeans, by the end of the seventeenth century, Islamic thinkers became prominent in their own lands by adopting the latest ideas from Western Europe.

Another undeniably important element in this, which was a contrast between the way things had developed in Western Europe (as we shall see in the next chapter), was the continuing dependence of higher learning in Islam on the patronage of wealthy and powerful individuals. Since patronage was by nature sporadic and not guaranteed, so was the emergence and success of Islamic philosophers; as the leading historian of Islamic science, George Saliba, has recently written: 'the scientific production of the Islamic world was mainly driven by individual genius but only when those geniuses could by accident encounter the right patron who would offer the support'. For whatever reason, patrons of the sciences seem to have become fewer and further between by the sixteenth century. Again, perhaps this was the result of developments outside the territories of Islam. Given that, by the seventeenth century, advanced knowledge of the natural world was available in the West, it must have seemed easier to would-be patrons of scientific learning to import such knowledge at need, rather than to enter into long-term support of an untried thinker, no matter how great his potential. Moreover, the importation of wisdom from the West led increasingly to a shift in attitudes and expectations; for a culture losing its former dynamism and self-confidence, relying on knowledge from Europe must have increasingly seemed the only viable thing to do.

FURTHER READING

Peter Adamson and Richard C. Taylor (eds), *The Cambridge Companion to Arabic Philosophy* (Cambridge: Cambridge University Press, 2005).

Toby Huff, *The Rise of Early Modern Science: Islam, China and the West* (Cambridge: Cambridge University Press, 1993).

J. Jolivet, 'The Arabic Inheritance', in Peter Dronke (ed.), *A History of Twelfth Century Western Philosophy* (Cambridge: Cambridge University Press, 1988), pp. 113–50.

David C. Lindberg, *The Beginnings of Western Science: The European Scientific Tradition in Philosophical, Religious, and Institutional Context, Prehistory to A.D. 1450* (Chicago: University of Chicago Press, 2008).

Richard Rubenstein, *Aristotle's Children: How Christians, Muslims, and Jews Rediscovered Ancient Wisdom and Illuminated the Middle Ages* (New York: Harcourt Brace, 2003).

A. I. Sabra, 'Situating Arab Science: Locality versus Essence', *Isis*, 87 (1996), pp. 654–70.

George Saliba, *Islamic Science and the Making of the European Renaissance* (Cambridge, MA: MIT Press, 2007).

William H. Stahl, *Roman Science: Origins, Development, and Influence to the Later Middle Ages* (Madison: University of Wisconsin Press, 1962).

4

The Western Middle Ages

The tenth century saw the beginnings of a revived and newly wealthy Western Europe. In a period free of barbarian invasions (the Vikings proved to be the last), improvements in agriculture and associated technologies led to a steady increase in the population, which in turn led to a division of labour and increased urbanization as townsfolk could service surrounding populations of landworkers. As confidence grew, so did intellectual aspirations. Western scholars, dimly aware of the glories that had been in Ancient Greece, and aware that the Arabs were in possession of this knowledge, began to try to recover for themselves ancient sources of knowledge. Gerbert of Aurillac (c. 946–1003), later Pope Sylvester II, when he was master of the cathedral school in Reims used contacts in al-Andalus, Muslim Spain, to acquire Latin translations of ancient works which were available there. The trickle of translations turned into a flood after the Christian reconquest of Spain in 1085, and the recapture of Sicily in 1091, as a number of Western scholars devoted their energies to translating newly available ancient writings from Arabic into Latin (see Box 4.1).

We saw in the previous chapter that up to this time there were only a handful of works on natural and secular philosophical knowledge available in Western Europe, and these were often compendia compiled from earlier works, which were themselves incomplete gatherings of Ancient Greek wisdom. These were frustrating times, therefore, for anyone with a burning curiosity about the world, and a craving to learn as much about natural phenomena and their causes as they could. Now, most of you reading this will never have looked at any of Aristotle's philosophical works, but, trust me, if you were as starved of the opportunity to learn as medieval Westerners, and then Aristotle's *Metaphysics*, say, fell into your hands, or his *Physics*, you could hardly fail to be excited, and to crave more. If you were to get your hands on one of Aristotle's works with a built-in commentary every few lines, explaining what it all meant, and why Aristotle thought this was an important issue to discuss, and obviously written by a commentator who was immensely learned in his own right, you would soon come to think that your life was never going to be the same again. Evidently, in Western Europe in the twelfth century, there were sufficient numbers of readers of Aristotle and his Arab commentators who felt this way, and together they forged for themselves new possibilities for an intellectual life in Western Europe.

Box 4.1 Some Latin translators and their major works

Constantine the African (*c.* 1020–87), a monk at the monastery of Monte Cassino, translated the medical encyclopedia of Ali ibn Abbas al-Majusi as *The Complete Book of the Medical Art*, or the *Pantegni*, and Arabic versions of the ancient medical writings of **Hippocrates** and Galen.

Gerard of Cremona (*c.* 1114–87), translated 87 books, including Ptolemy's compendium of astronomy, the *Almagest*; many of the works of Aristotle, including *Posterior Analytics*, *Physics*, *On the Heavens and the World*, *On Generation and Corruption*, and *Meteorology*; Euclid's *Elements of Geometry*; al-Khwarizmi's *On Algebra and Almucabala*; Archimedes' *On the Measurement of the Circle*; al-Farabi's *On the Classification of the Sciences*; the chemical and medical works of al-Razi (Rhazes), Avicenna's *Canon of Medicine*; and Ibn al-Haytham's *Optics*.

Adelard of Bath (*fl.* 1116–42) made the first translation of Euclid's *Elements*, as well as the highly regarded *Introduction to Astrology* of Abu Ma'shar.

Robert of Chester (*fl. c.* 1150) translated the alchemical works of Jabir ibn Hayyan (Geber), arguably the most influential of the alchemists, including the *Kitab al-Kimya* (*The Composition of Alchemy*).

Leonardo of Pisa, known as Fibonacci (*c.* 1170–*c.* 1250) presented the first complete account of the Hindu-Arabic numerical system in his *Liber Abaci* (1202).

Michael Scot (*c.* 1175–1232) translated *On the Motions of the Heavens* by al-Bitruji (Alpetragius), and the important commentaries on the scientific works of Aristotle by ibn Rushd (Averroes).

Arnald of Villanova (1235–1313) translated the Arabic versions of Galen and some of the works of ibn Sina (Avicenna).

William of Moerbeke (*c.* 1215–86) translated philosophical, medical, and scientific texts directly from Greek, rather than Arabic translations. He focused mostly on the works of Aristotle, including the first translation of Aristotle's *Politics* (*c.* 1260). He also translated mathematical works by Hero of Alexandria and Archimedes, and an important Neo-Platonic work, the *Elements of Theology* of Proclus (1268).

It is worth noting that the scholars who went to Spain and subsequently to Byzantium (to acquire Ancient writings in their original Greek rather than in Arabic translation), proved to be highly selective in what they chose to translate. Apart from the usual concern with works that could be seen to offer pragmatic advantages, the translators almost exclusively chose works of natural philosophy. There was an emphasis upon medical works, of course, particularly those of Galen (*c.* AD 129–200), the prolific synthesizer of Ancient Greek medical theory and practice. Mathematical works of various kinds, which were also seen to be useful – including works in astronomy and astrology, and works on optics – were also prominent. But leaving these aside, the major focus was on translations of Aristotelian natural philosophy.

Although the medical and mathematical writings were clearly deemed to be useful in an entirely pragmatic way, it is hard to understand the emphasis upon natural philosophy unless we assume that, right at the outset, there was a strong feeling that works about the natural world, about God's Creation, would complement God's Word in the Scriptures. There was already a vigorous movement among some of the leading thinkers in the Church to support religion with reason. Although the initial impetus was to develop a rational metaphysics, as Aristotle's works became increasingly available the movement drew increasingly on his natural philosophy. But this was, after all, already a feature of the Islamic philosophical writings, particularly those of Avicenna and Averroes. Islamic philosophy was often intended to offer rational accounts of the kinds of ideas which emerged in Islamic speculative theology, in particular how God created the world, how he subsequently related to it, the existence and nature of the soul, and so forth. If philosophy could serve Islam, the Christians thought, it could surely serve Christian theology.

There was undeniably a perpetual current of opposition to the rational approach to religion, often tantamount to anti-intellectualism, as can be seen in Bernard of Clairvaux's (1090–1153) prosecution of Peter Abelard (1079–1142) for a supposedly heretical, rationally inspired, anti-Trinitarianism. Bernard evidently recognized that Peter simply desired knowledge for knowledge's sake, but he regarded this as 'ridiculous vanity'. Bernard triumphed over Peter and was canonized twenty years after his death, but in the long run it was Peter's rational theology which became characteristic of medieval Christianity. In spite of Bernard and those like him, there were always others whose excitement at the discovery of new ways of thinking about ideas, new at least in the Latin West, could not be suppressed. The clearest evidence for the strength of the view that reason and religion should be compatible can be seen in the set-up of the newly emerging universities.

The universities grew out of spontaneous gatherings of students around a teacher of great renown, usually at a cathedral school. Such gatherings consisted mostly of strangers, or foreigners, in the cathedral city, men with no rights or privileges there. Accordingly, they sought collective protection under the law, in the same way that tradesmen and craftsmen did, by declaring themselves to be a 'universitas', or corporation. As the universities grew they were organized into separate faculties, specializing in a particular area of study. The three 'higher' faculties of divinity, law, and medicine, emerged because the burgeoning population of Western Europe had increased need, sadly, of doctors and lawyers, and an increased need of priests.

To begin with, the higher faculties required their students to do preliminary studies in the Arts faculty, so-called because this was originally taken up with the foundational seven liberal arts (preserved from the Ancient Roman educational system): Latin grammar, rhetoric, and dialectic or logical argument, geometry, arithmetic, astronomy, and music. As the works of Aristotle became increasingly available, thanks to the translators' work in Spain and the other centres of translation, the Arts faculties effectively became philosophy faculties, where everyone studied the natural philosophy of Aristotle. The seven liberal arts were now taught to young boys, who would then progress to natural philosophy. Aristotle himself had declared that 'medicine and philosophy are sisters', and the equivalent medical authority, Galen, had declared that the best physician is also a natural philosopher. Schools of Law, similarly, were keen to see their students arrive with a command of logic, which was best acquired through study of Aristotle's *Organon* (the collective name for his logical works). It might seem, however, that the Theology faculties would be less approving of a grounding in pagan philosophy, but such was the excitement caused by the power and sweep of Aristotle's philosophy that, for the most part, they were happy to encourage the new style Arts faculties. Indeed, this indicates the beginning of Peter Abelard's eventual redress against the more anti-intellectual aspects of the Church. Natural philosophy began to be seen as an essential grounding for work in theology, no less than for work in medicine or the law.

It can hardly have been coincidental that universities appeared, uniquely in Western Europe, at just the time that the civilization of Western Europe was experiencing a new flourishing after the social and political insecurity of the 'Dark Ages'. The rise of strong monarchies which were a feature of this time required not just legal reform but a sound basis for its system of law. Accordingly, the importance of the major codification of Roman Law commissioned by the emperor Justinian (527–65), the *Corpus Juris Civilis*, became recognized for the first time, and became the mainstay of the new law faculties. Medical faculties taught from the works of Galen and his Arab commentators, especially Avicenna, and were able to produce elite medical practitioners who could lay claim to mastery of Ancient (and therefore superior) medical wisdom. The theology faculties, of course, studied Scripture and the works of the Church Fathers. If professors in the Arts faculties had ambitions to extend themselves beyond the seven liberal arts, and to claim the same kind of status and expertise as professors in the higher faculties, it seems clear that they too had to engage in the advanced study of ancient texts of their own; the obvious body of Ancient work to choose was that of Aristotle.

Given the almost exclusive focus on Aristotle among the major Islamic philosophers, it is hardly surprising that Western scholars should have followed suit. Even those who went to Byzantium, which still had continuous links with Ancient Greek learning (and where Greek, not Latin, was the *lingua franca*), and who might, therefore, have looked for the works of Plato or other Ancient philosophers, tended to look only for the works of Aristotle. As a result Aristotle dominated Western philosophy. If you look at any medieval philosophy book there are continual references to 'the Philosopher': 'the Philosopher says this...', 'according to the Philosopher...', and so forth. Similarly, there are constant references to 'the Commentator'. There was no need to mention names. The philosopher could only be Aristotle, and the commentator was Averroes. Even without going into the details of twelfth- and thirteenth-century philosophy in Western Europe, it seems impossible to deny that it was shaped by the Arabs, and like Islamic philosophy it focused almost exclusively on Aristotle.

* * *

It was not long, however, before tensions did arise between the faculties. Professors in the Arts faculties began to develop Aristotelian ideas in a way which seemed to undermine the assumptions of the Church. Generally speaking, problems arose when the philosophers insisted, on Aristotelian grounds, that certain physical states of affairs were impossible. Aristotle was adamant, for example, that a vacuum could not exist. Aristotle had no real concept of empty space – for him, extension in three dimensions was a defining characteristic of bodies. Talk of an extended space with nothing in it made no sense to him. A space with defined measurements of length, breadth, and depth, was a space occupied by body. To take away the body from the space was to take away the dimensions from the space, and therefore meant that the space in question no longer existed – not that it was still there but empty. Void space, for Aristotle, was categorically impossible.

The Church took a different view, however. Bearing in mind the supposed omnipotence of God, the Church insisted that if God chose to create empty space, nothing Aristotle or anyone else could say was going to prevent him from doing so. Similarly, Aristotle had insisted that the world as a whole (i.e. the Earth and all the surrounding heavenly spheres, including the sphere of fixed stars) could not be moved. Aristotle's argument hinged upon the fact that the place of something was defined in terms of the bodies surrounding it (so the place of a parked car is defined by the fact that it is sitting by the kerb of the road, say, between two other cars). Since there is nothing surrounding the world, it is impossible to define how far it has moved, or indeed whether it has moved. Accordingly, to talk of the movement of the whole world makes no sense; it is a category mistake based on a failure to realize that the world as a whole is not something whose movement can be defined or understood. Again, theologians simply insisted that if God wished to move the world, he could.

Theologians were willing to concede that there were some things that their omnipotent God could not do. It was accepted that it did not diminish God's omnipotence to say that he could not perform *logical* contradictions. Clearly, he could not create a married bachelor, and he could not defy the law of excluded middle (which says that something either is, or is not, the case – there is no middle alternative). Although Aristotelian doctrines about the void, or about the movement of the world, seemed to be couched as *logical* impossibilities, the instincts of the theologians led them to believe that the real issue was one of *physical* impossibility. There was nothing, however, which they considered to be physically impossible for an omnipotent deity.

Box 4.2 A Selection from the 219 Articles Condemned by Etienne Tempier, Bishop of Paris in 1277

9 That there was no first man, nor will there be a last; on the contrary, there always was and always will be generation of man from man.
34 That the first cause [i.e. God] could not make several worlds.
35 That without a proper agent, as a father and a man, a man could not be made by God [alone].
37 That nothing should be believed unless it is self-evident or could be asserted from things that are self evident.
49 That God could not move the heavens [that is, the world] with rectilinear motion; and the reason is that a vacuum would remain.
90 That a natural philosopher ought to deny absolutely the newness [that is, the creation] of the world because he depends on natural causes and natural reasons. The faithful, however, can deny the eternity of the world because they depend upon supernatural causes.
91 That the argument of the Philosopher demonstrating that the motion of the sky is eternal is not sophistical; and it is amazing that profound men do not see this.
141 That God cannot make an accident exist without a subject [i.e. make a property of something without making the something of which it is a property] nor make more than three dimensions exist simultaneously.
145 That no question is disputable by reason which a philosopher ought not to dispute and determine...
147 That the absolutely impossible cannot be done by God or another agent – an error, if impossible is understood according to nature.
153 That nothing is known better because of knowing theology.
154 That the only wise men of the world are philosophers.
185 That it is not true that something could be made from nothing, and also not true that it was made in the first creation.

Consequently, universities where the Theology faculty was the most powerful, notably Paris but also Oxford, ordered the banning of numerous Aristotelian teachings in 1277 (see Box 4.2). Although the ban was soon lifted, and it had been largely confined to Paris and Oxford (other universities actually tried to tempt students away from Paris by pointing out that they operated no such ban), it nevertheless had a significant effect upon the subsequent development of natural philosophy.

Philosophers who had previously been thoroughly enthralled by Aristotle's ideas and seemed reluctant to deviate from his teachings, now began to think more critically. Although this usually meant seeing the issues in a way which was more in keeping with the Church's views, it nevertheless enabled medieval philosophers to see that Aristotle's beliefs weren't necessarily the last word on every topic, and that there were other possible ways of seeing things. In some cases this even led thinkers to propose other theories about the natural world which were not found in Aristotle's works.

Consider, for example, impetus theory. Aristotle said that something could only move if it was being moved by something else. A book will sit on a shelf until someone picks it up and moves it. But what about projectiles? Once a spear leaves the hand of the thrower, why does it keep on moving? Aristotle's explanation of projectile motion was vague and confusing. Consequently, some medieval philosophers developed the notion of impetus – which supposes that when you throw something you impart an impetus upon it and the impetus keeps it moving until the impetus is all used up, then the projectile falls to earth. The theory of impetus is not found in Aristotle, but it seems to foreshadow our modern theory of inertia. Could it be that the ban on Aristotle led medieval thinkers to break away from Aristotle and develop their own, superior physical theories?

There is no use speculating on what might have been. The fact is that the power of Aristotle's thought and the coherence of his system, were so great that medieval thinkers never did succeed in breaking away from him. What tended to happen was that improvements on Aristotle's theories were simply made to fit into Aristotelianism. The resulting philosophy, developed continually throughout the High Middle Ages, is more properly called *scholastic* Aristotelianism, or simply scholasticism. In the case of impetus theory, for example, instead of leading on to the modern principle of inertia – which simply rejects Aristotle's claim that everything that is moved is moved by something else (the principle of inertia assumes that once something is moving it continues to move until something stops it) – it was simply fitted into Aristotelian theory. As soon as a projectile leaves the thrower's hand, impetus takes over; but it is still true that the moving body is moved by something else, namely the impetus. And whereas inertia is now seen to guarantee perpetual movement until something else prevents further movement, impetus was seen as something which was used up as a body moved, and once exhausted could no longer keep the body moving.

The original Paris ban of Aristotelian propositions included many that were upheld by Thomas Aquinas (*c.* 1225–74), even though Aquinas, unlike some of his contemporaries (most notably Siger of Brabant and Boethius of Dacia, both *fl. c.* 1260, who were thoroughgoing Averroists with a strong dose of atheism thrown in), had tried to reconcile his reading of Aristotle with the doctrines of the Church. Canonized fifty years after his death, and eventually inducted into the ranks of the 'Doctors of the Church', Aquinas largely defined the scholastic Aristotelianism which became not only acceptable to the Church but also, in many respects, a recognized and valued adjunct to Roman Catholic theology. It was thanks largely to Thomas's posthumous influence that the Aristotelian world picture became the Christian world picture (see Figure 4.1). But this would hardly have been possible if it were not already the case that leading Churchmen had a strong tendency to uphold a rational approach towards understanding what they

Figure 4.1 The Aristotelian world-picture as depicted in Peter Apian, *Cosmographia* (Antwerp, 1539). Reproduced with permission from Edinburgh University Library, Special Collections (JA 1050).

Around the central Earth, shown as a single terraqueous globe (see Chapter 6), it depicts the sphere of Air and the sphere of Fire (the belt of flames just below the Moon), showing the Aristotelian assumptions about the natural places of each of the elements. The words around the outside read: 'The ruling heaven, dwelling place of God and all the Elect', showing how this picture of the Universe dovetailed with religious beliefs. The picture shows a cross-section through the heavenly spheres, which should be envisaged as completely surrounding the Earth.

saw as God's Creation. The Roman Catholic Church, after all, even used Aristotle's matter theory to explain the mystery of the Eucharist.

It might be remembered (from Chapter 2) that, according to Aristotle, a particular body consisted of matter and form – neither of which could be found alone (it is impossible to have matter without any form whatsoever, and it is impossible to have a disembodied form). It is the form, however, which endows a body with its specific properties (matter is always the same, until a specific form is imposed upon it). Now, some of these properties are essential to the body in question, and some are merely accidental. Essential properties cannot be changed without changing the body (you cannot re-mould the blade of a knife into a sphere and still call it a knife), but accidental properties can be changed without seriously affecting the body (you can change the handle of a knife from stag's horn to wood).

Remarkably, Roman Catholic doctrine in the Middle Ages appropriated this theory of body to explain what happens to the bread and wine in the Eucharist. The form of the wafer was held to change from the form of bread to the form of Jesus Christ's body. Accordingly, the bread now had the *essential* properties of Jesus's body; however, it still maintained what were deemed to be the *accidental* properties of the bread, and so still looked and tasted like bread (to be sure, bread can vary widely in taste and still be bread, so taste is not an essential property; presumably what is essential to bread is that it is made from raised and baked dough). The important thing to note for our purposes is not merely the ingenuity of this effort to explain the mystery of the Eucharist, but the fact that the Church felt it was important to provide such a rationally based account. By the late thirteenth century, then, there was a vigorous rational tendency among Church leaders. Thomas Aquinas was not just a leading commentator on the philosophy of Aristotle, he was also (and remains) one of the Church's foremost theologians. The theology of the medieval period was no less scholastic than the natural philosophy, and it too deferred to reason.

The alliance between Roman Catholic theology and scholastic Aristotelianism (in which it was routinely held that natural philosophy was the hand-maiden to the Queen of the Sciences, Theology) partly accounts for the fact that scholastic Aristotelianism dominated the Arts faculty curriculum in every university from the thirteenth through to the seventeenth (and in some universities even into the eighteenth) century. In the long run, it was this Holy alliance which marked Western Europe out from both Greek Orthodox Christianity, which held sway in the Eastern part of the former Roman Empire, and from the Empire of Islam. The universities were an invention of the Latin West, and although the precise origins of the universities are obscure, the ubiquitous prevalence of natural philosophy in the curriculum strongly suggests that natural philosophy, as recovered from the Arabs, was immediately recognized as an important support for, and adjunct to, not only medicine, but also theology. As the various universities sprang up in different parts of Europe, natural philosophy was always built into the curriculum, and furthermore, it was always regarded as something that everyone should study before going on to try to master theology, medicine, or law.

In this respect, the Middle Ages in Western Christendom mark a high point in history for the recognition of the importance of the study of the natural world. Furthermore, the flourishing of the universities ensured that natural philosophy always had an important place in Western civilization. We have seen how the flourishing of science in Islam depended entirely upon the interest and the engagement of wealthy and politically powerful patrons. This was inevitably precarious, and even though Islamic achievement in the sciences far exceeded anything produced in Western Europe until the Renaissance, nothing guaranteed its continued existence. In Western Europe, by contrast, the continuity of the universities ensured the continuity of natural philosophy (the earliest universities, Bologna, Paris, and Oxford have recently outlasted the 900 years of Plato's Academy). What's more, since anyone who ever attended university was schooled from the beginning in Aristotelian natural philosophy, it meant that the understanding of the natural world became almost a taken-for-granted aspect of elite culture.

There were some negative aspects to all this, however. Whereas the new civilization of Islam was driven to make significant advances on Ancient Greek achievements, the Latin West was content to recover Ancient Greek wisdom and to endlessly expound upon it. Islam in the early period after its foundation felt the need to prove itself superior to surrounding civilizations, such as that of Persia, or Byzantium, and so was led to improve upon the Ancient knowledge which these other civilizations had been happy merely to preserve. The Latin West was not driven by similar considerations, and scholars emerging from a period in which there was a dearth of opportunities for advanced learning, allowed themselves to become completely, and uncritically, enthralled by Aristotle, Galen, Ptolemy, and a few other Ancient sources of wisdom.

It's a commonplace in the history of science to dismiss the Middle Ages as a vast period of time when virtually nothing of significance happened. There is a tendency to jump from the Ancient Greeks to the Renaissance with perhaps a passing mention of Islam in between. This is undoubtedly unfair, and simply wrong; numerous books have been written which are devoted solely to the history of science in the Latin Middle Ages. Nevertheless, it has to be said that a broad-brush picture of the medieval period can hardly avoid depicting it as a time when scholastic Aristotelianism dominated the scene from the thirteenth through to the sixteenth century (see Box 4.3). It is certainly possible to find examples of innovatory ideas, such as impetus, but these never led to significant changes in the direction of thought, but tended instead to be adapted to fit into the Aristotelian world picture. At best, it seems fair to say, there was an accumulation of doubts and an increasing awareness of inconsistencies which would eventually help to stimulate the sea-change in ideas which occurred during the Renaissance.

Having said that, it is important for the historian of science *not* to skip over the Middle Ages, however. It is here that we find the reason why the study of the natural world did not once again diminish in value and all but disappear, as it had done eventually in the Ancient World, and in Islam. It is here, in other words, that we discover the reason why scientific knowledge, after a number of false starts, became an intrinsic part of civilization and an indispensable aspect of our culture. The perpetual continuity of science was guaranteed when its precarious dependence on individual patrons of the sciences, as in the civilization of Islam, was replaced by the institutionalization of natural philosophy in the Western European universities, and effectively by its intimate association in the university setting with Christian theology.

Box 4.3 Five reasons for the longevity of Aristotelian philosophy in the medieval university system, and for its tendency to remain unchanged

Medieval thinkers assumed that perfect knowledge belonged to the past. Adam, the first man, was supposed to have known all things in the Garden of Eden, but that Adamic wisdom was gradually lost (forgotten) after he and Eve ate the forbidden fruit and were cast out of Paradise. (If you don't know who Adam and Eve were you'd better read the first couple of chapters of the Book of Genesis, the first book of the Bible!) Accordingly, medieval thinkers had no sense of progress, but rather a belief that knowledge must be recovered by looking back to see what people in the past knew; the earlier the thinker, the closer they were to Adam, and so the more they would have still remembered of the Adamic wisdom. So, medieval thinkers did not think of discovering new knowledge for themselves, but of studying earlier thinkers, such as Aristotle.

The two most powerful faculties in the medieval universities were Divinity and Law. In both, the professors studied early texts and demonstrated their expertise by showing the depth of their knowledge of those ancient texts (Scripture or various texts of ancient Roman Law). Accordingly, professors in the other faculties, Arts and Medicine, wanted to emulate professors of Divinity and Law by showing their mastery of Ancient texts. So, professors in the Arts faculties all over Europe showed their mastery of the writings of Aristotle. In Medicine, the equivalent ancient authority was a physician called Galen (see Chapters 5 and 10). Furthermore, there was no incentive in Divinity or Law to introduce innovation. The emphasis in law was on past precedent not reaching new conclusions. The aim in theology was to uphold the earliest traditions of the Church. Again, medicine and natural philosophy within the universities followed suit.

Christianity differed markedly from the other Abrahamic monotheistic religions in recognizing the need for a systematic theology. It was not necessary, and in many respects it was considered inappropriate, to examine and deliberate upon the nature of God in Judaism and in Islam. Questions arising from the faith in these two Abrahamic religions were for the most part matters of law – what would God want us, or expect us, to do in given circumstances? To go beyond this, and to question God's nature, was simply blasphemous. In Christianity, however, thanks to the need to reconcile the notion of a supremely transcendent God on the one hand, with the notion of the man Jesus Christ as God on the other, a completely different set of questions arose which were about the nature of divinity itself. Far from being blasphemous, it was hardly possible to be a Christian without addressing questions like these. The Christian belief in a God that was both one and three (God the Father, God the Son, and the Holy Spirit), which strikes many as simply irrational (or incompatible with monotheism) in fact emerged out of philosophical debates about how God could be both a transcendent being and a man. Accordingly, systematic theology grew out of the origins of Christianity as a matter of course.

The fact that the two most powerful faculties were teaching would-be priests to be good pulpit orators, and would-be lawyers to be good orators in courts of law, together with the focus on Ancient writings had repercussions on the nature of university education. Student performance was judged by their abilities in public debate or 'disputation'. Typically, a class would be given a topic to dispute, such as: 'Whether a vacuum is possible', or 'Whether there can be other worlds than this one'. They were always topics discussed by Aristotle and one student was assigned to defend Aristotle and another to oppose his ideas. If a good student came up with a powerful anti-Aristotelian argument he might win the debate, but nobody thought any more of it than that; everyone knew that next time the same topic was debated the chances were that the outcome would be different. The issue was never whether Aristotle was right or wrong, but who, on this occasion, argued more persuasively in debate. This in itself was a legacy from Ancient Greece; even Aristotle admitted (in his *De caelo*) that 'we are all in the habit of relating an enquiry not to the subject matter, but to our opponent in argument'.

There was simply no alternative to Aristotle. The recovery of Ancient writings had been very piecemeal and only Aristotle's writings were available in anything approaching completeness. Accordingly, university Arts faculty curricula throughout Europe were based entirely on Aristotle's philosophy, and the result was a powerful system of thought which had no rival views. There was simply no alternative perspective, no other standpoint from which to judge any issue in natural philosophy.

Finally, thanks largely to the efforts of Thomas Aquinas (who was subsequently made a saint), Aristotelianism was married to Roman Catholicism and became so closely intertwined with Church doctrine that to attack Aristotle was to attack the Church. This was seen most clearly, perhaps, in the famous case of Galileo's attempts to defend the theory of a moving Earth (see Chapter 9).

Moreover, it soon became clear that it wasn't possible to discuss the nature of God without discussing the nature of his relationship to man and, more importantly from our point of view, without discussing his relationship to the world. This is why natural philosophy became such a close companion, a handmaiden indeed, to Christian theology in the Middle Ages, became institutionalized in the universities, and so came to be seen as an essential preliminary study for anyone aspiring to higher learning. While it was always possible in Judaism and in Islam to dismiss questions about the nature of the physical world as irrelevant to what the culture deemed to be the really important questions, in Christianity it never was. In Christianity the really important questions were not just concerned with God's commandments but also with his nature, and his nature could not easily be discussed without considering his relationship to his Creation – the physical world.

It is important to note that I am *not* saying Judaism and Islam were opposed to science, while Christianity was not. I am simply saying that in Judaism and Islam it was never essential to consider in detail God's relationship to the natural world, and therefore the religious authorities never considered it necessary to support and encourage scientific investigation of the world. In Christianity, however, thanks to the unique circumstance that it worshipped a man as a God, it had to develop a systematic theology, and that theology in turn ensured that the nature of the physical world was always regarded as an important adjunct to understanding the nature of God.

Although theology was never entirely absent from Islam, there was always a great deal of resistance to it. Mohammed's strict monotheism involved the rejection of the incarnation of divinity in Jesus, and the notion of the Holy Trinity, and for many Muslims this meant that there was no need for a theology at all. But even those who recognized the need for an Islamic theology tended to develop an abstract theology which had no internal need to interact with natural philosophy. It was only in Christianity that theology and natural philosophy went hand in hand.

Science and religion are all too often seen today as categorically distinct entities, each with its own irreconcilable worldview. Books which emphasize the supposed unbridgeable gulf between them become best sellers in our entirely secularized world, but the authors of such books are merely representing the prevailing view in our secular society that science and religion are inimical to one another. The historical fact remains, however, that the uninterrupted development of Western science would almost certainly not have been possible if it hadn't been recognized by leading medieval thinkers (all of whom were theologians) as an essential concomitant of Christian theology.

FURTHER READING

Roger French and Andrew Cunningham, *Before Science: The Invention of the Friars' Natural Philosophy* (Aldershot: Scolar Press, 1996).

Edward Grant, *The Foundations of Modern Science in the Middle Ages: Their Religious, Institutional and Intellectual Contexts* (Cambridge: Cambridge University Press, 1996).

Edward Grant, *God and Reason in the Middle Ages* (Cambridge: Cambridge University Press, 2001).

Edward Grant, *Science and Religion, 400 B.C. to A.D. 1550* (Baltimore: Johns Hopkins University Press, 2005).

David C. Lindberg, *The Beginnings of Western Science: The European Scientific Tradition in Philosophical, Religious, and Institutional Context, Prehistory to A.D. 1450* (Chicago: University of Chicago Press, 2008).

Steven P. Marrone, 'Medieval Philosophy in Context', in A. S. McGrade (ed.), *The Cambridge Companion to Medieval Philosophy* (Cambridge: Cambridge University Press, 2003), pp. 10–50.

5
The Renaissance

By the fifteenth century, wide-ranging changes in Western European culture and civilization were taking place – this marked the beginning of the period known as the Renaissance.

It was a period which saw the introduction of three great inventions which were to thoroughly change ways of life. The invention of gunpowder changed the face of battle and the nature of warfare; the magnetic compass made navigation across the oceans possible for the first time and led to voyages of discovery, the extension of trade, and the beginnings of colonialism and imperialism; and the printing press meant that access to religious, philosophical, and literary works was no longer the preserve of the wealthy and powerful, but could be made accessible to almost everyone who took an interest. Developing alongside these innovations and their impacts were other sea-changes in European life. It was a period which saw the decline of feudalism and the rise of mercantilism, and some would argue, even the beginnings of capitalism. Certainly banking was invented at this time. It was also the period which saw the rise of strong city states and the beginnings of the nation state. Similarly, it was a period renowned for the burgeoning of painting, sculpture and architecture. It was also the period of the Protestant Reformation, and, as we shall see, of the scientific revolution.

Many, if not all, of these developments were interconnected – changes in one area affecting, or even leading to, changes elsewhere. Our concern here, however, is only with those changes which directly affected intellectual matters, and in particular efforts to understand the natural world.

The word 'renaissance' means re-birth, but why is this period referred to as a re-birth? Interestingly, unlike the 'Middle Ages', this is not a term of convenience invented by historians. The Renaissance was given this name by learned scholars living at the time, who saw their own age not as an age of innovation, but as an age of re-birth. The re-birth they were referring to was that of ancient wisdom – re-born thanks to the recovery of previously lost works of literature and philosophy by the Ancient Romans and, above all, the Ancient Greeks. After centuries of being content with what had been recovered from the Arabs and from Byzantium in the twelfth century, a number of scholars decided to see if they could recover any of the other works which were often mentioned in the Ancient texts that had survived.

They did not undertake their searches on a whim, of course. The newly wealthy princelings in the Italian City States of the Renaissance wanted to show their superior status by various visible means. These princelings famously commissioned architects to build them elegant palaces, painters to adorn the walls and ceilings with frescoes, and gardeners to create elaborate settings for their palaces. But they also became collectors – of exotic curiosities from far flung parts of the world (displayed in Cabinets of Curiosities – the beginnings of what later became museum collections), and of what were regarded as important books and manuscripts. The leading book collectors appointed their own librarians and charged them with expanding their collections. For the first time, there were now learned men who were paid by their patrons to search for prestigious writings to enhance the patrons' collections.

The starting point for these librarians was in the older libraries which already existed in monasteries, and this proved astonishingly fruitful. Italian scholars who found a few ancient Roman works of poetry and history in local monasteries immediately began a systematic search, first in Italy but later further afield including, eventually, monasteries in the territories of the Greek Orthodox Church, where they found Ancient Greek works.

Monasteries were first established in Europe in the early fifth century AD, and the monks evidently took it upon themselves to establish libraries and preserve the works in them. In the ages before the printing press what this meant was that every now and then a work in a dilapidated condition would have to be recopied by hand. The process would have to be repeated of course when that copy itself began to deteriorate. Euclid's *Elements*, the major summation of Ancient Greek geometry, was written in about 300 BC. The oldest manuscript copy which exists today was written in AD 880. Clearly, whoever wrote that manuscript in the ninth century AD was copying not from a copy produced in 300 BC, but from a manuscript that was perhaps only decades old, and already in danger of rotting away. But that manuscript itself must have been copied from an earlier copy written perhaps in the eighth century. Clearly the line of copies must stretch back to an original copy dating from about 300 BC but which has long since ceased to exist.

It is important to note that there is no evidence the monks were reading these Ancient works which they held in their libraries. They simply venerated books and so ensured that they made fresh copies whenever they were needed, in order to preserve them. In some cases, of course, they did allow works to deteriorate without copying them – usually works of which they disapproved. Although the Renaissance scholars searching the monasteries found many copies of works by Plato and Aristotle, for example, only one copy of Lucretius's *De rerum natura* (*On the nature of things*) was found, and only one copy of the poems of Catullus (*c.* 84–*c.* 54 BC). These are now easily available in modern editions, but all due to the fact that one monastery decided to preserve Lucretius in spite of the fact that his epic poem was clearly atheistic, and one decided to preserve Catullus in spite of the fact that his poems were very sexually explicit. Every other monastery which had copies evidently decided to let them rot away. These two works were discovered by Renaissance scholars just in time; others, alas, had already been consigned to oblivion by the time the scholars began their systematic searches.

The riches that emerged from these attempts at recovery were truly remarkable. Indeed, virtually all the Ancient classical literature we have today was recovered at this time – very little has been added since. The impact of this new literature was profound. In philosophy, for example, the complete works of Plato now became available. Previously only part of his *Timaeus* was available, but scholars could now see the full range of Plato's interests and became thoroughly absorbed by the depth of his arguments. But the writings of other Ancient thinkers also became available, including those of the Epicureans and the Stoics. It was now all too easy to see that Aristotle's way of considering things was by no means the only way.

One very important discovery for the history of Ancient Greek philosophy was a book by Diogenes Laertius (third century AD) called *Lives of the Philosophers*. As the title suggests, this gave biographies of all the Greek philosophers, and an account of their theories and ideas, and even a list of all their writings. Thanks to this we are now able to tell how much of Greek wisdom we are still lacking.

But, there was something surprising about this book for Renaissance scholars. It was quite clear from the comparative length of the chapters, to say nothing of what was written in them, that Aristotle was not much admired by Diogenes Laertius. Aristotle was given less space than Plato, of course, but also less than Pythagoras, less than the atheistic atomist Epicurus, less than the Stoics Chrysippus and Zeno of Citium. It was also clear from Diogenes' discussions that he thought more highly of each of these than he did of Aristotle.

There were other surprises in store. Cicero, the Roman orator whose Latin style was held up as a model to be emulated, wrote a number of philosophical works. Some of these were in the form of Platonic dialogues, allowing Cicero to present the ideas of different philosophers next to one another. One of his most important works, *On the nature of the gods*, was presented as a dialogue between an Epicurean, a Stoic, and a Sceptic. Cicero clearly did not think it was worth presenting the ideas of Aristotle.

Bearing in mind that Renaissance readers of these and other similar works had been totally steeped in Aristotle's philosophy while they were at university, and were used to thinking of him simply as **the** Philosopher, it came as something of a shock to see him being treated as a comparatively lesser thinker. **The** Philosopher was only **a** philosopher. The result was a crisis of thought amongst Renaissance scholars. A crisis brought on by the dawning realization that they had invested all their intellectual efforts on the thought of one thinker, who turned out to be just one among many.

There were different responses to this crisis. Some thinkers abandoned Aristotle in favour of another Ancient source of authority. Plato was the most popular, but some embraced Stoicism, some even defended Epicurus against charges that he was an atheist. Others took a line known as eclecticism, which involved amalgamating the best ideas from different philosophers to come up with a blended philosophy.

Others, however, embraced another Ancient Greek philosophy, and one which could now be seen to have been popular amongst the Ancient Greeks themselves (this could be seen both by the treatment of this philosophy by Diogenes Laertius and by Cicero), namely, Scepticism. Essentially, sceptics rejected all authorities on the grounds that it was impossible to know who was right. After centuries in which it was believed that Aristotle was the supreme authority in philosophy and always had the right answer, a number of Renaissance thinkers now began to embrace a revived scepticism.

For others, however, to whom scepticism was uncongenial, the rejection of authority was accompanied by a new determination to try to discover the truth for themselves.

This was a very dramatic change in intellectual life. Bear in mind that natural philosophers throughout the Middle Ages did not study the natural world, they studied what Aristotle said about the natural world. So, the idea of rejecting authority and studying the natural world itself, so obvious to us, was radically new.

We can see this change in intellectual attitude even in the art of the Renaissance. Medieval painting is emblematic, and symbolic, but Renaissance paintings are much more realistic and representational. The introduction of mathematical perspective into paintings is a clear sign that Renaissance artists wanted to try to depict the world as it really appears in our daily experience. Medieval painters were content to depict symbolic or ritualistic relationships between people and objects in their paintings; in so doing they were representing traditional and authoritative claims about the way things were supposed to be. Renaissance artists, by contrast, wanted the scenes they depicted to look the way they did in the real world.

Another example of the rejection of authority in the Renaissance can be seen in Martin Luther's (1483–1546) break away from the Roman Catholic Church and the authority of the Pope, to set up the first Protestant Church. Famously, Luther said that instead of obeying the authority of the priest, every man should be his own priest. He advocated reading the Bible for oneself – something which was forbidden to Roman Catholics, who were expected to learn what was in the Bible through the intermediary of a priest (just as natural philosophers learned what was in nature through the intermediary of Aristotle). The Bible was for Luther what the natural world was for the new breed of Renaissance natural philosophers – the source which should be studied in its own right, without interference from the preconceived ideas of authority figures.

The dramatic impact which the rejection of authority had in natural philosophy can be seen in changes in the standard of illustrations in botanical and medical works. Consider the example provided by the anatomical textbook of Andreas Vesalius (1514–64), *On the Structure of the Human Body*, published in 1543.

Andreas Vesalius and Renaissance anatomy

Before printing, books were reproduced in hand-written copies. It was important, therefore, to employ the services of a good scribe, a professional writer who could copy accurately in legible handwriting. Scribes were not necessarily also good at drawing, however. When copying a medical manuscript, they would no doubt try their best, but the result might well look like that depicted in Figure 5.1.

Admittedly this is a particularly bad example, but there were many medical manuscripts in which the illustrations were little better. Evidently, the illustrations were not considered to be important - what mattered were the authoritative words, of Galen, Avicenna, or whichever author was being copied.

With the advent of printing, however, this had to change. Printers, realizing it would be embarrassing to publish anatomy textbooks with inadequate pictures, hired artists to supply the pictures. Not every artist was a Leonardo da Vinci (1452–1519), however, studying anatomy in their own right, as he did. The result therefore might be illustrations which although reasonably competent, showed little anatomical detail, as shown in Figure 5.2.

Clearly, it was still the case that what mattered were the verbal accounts of anatomy provided by the authority figure who wrote the book.

Andreas Vesalius showed his credentials as an innovatory Renaissance thinker by turning his scrutiny towards the human body itself, rather than accepting

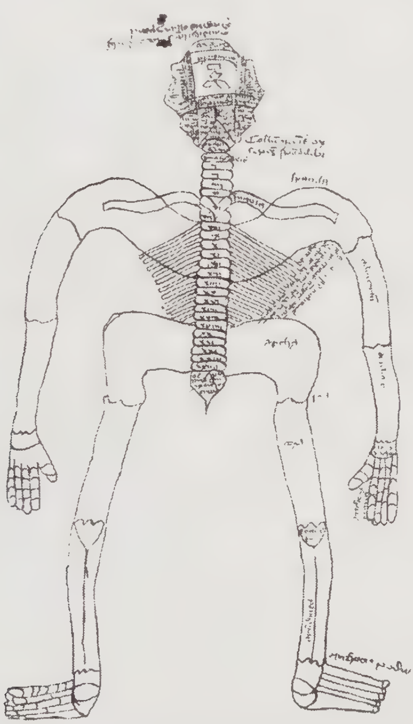

Figure 5.1 Illustration of the skeleton, copied by a scribe, into a fourteenth-century medical manuscript. With permission from the Wellcome Trust.

Figure 5.2 Anatomical drawing of female abdomen from Charles Estienne, *De dissectione partium corporis* (Paris, 1545). Reproduced with permission from Edinburgh University Library, Special Collections (JY 526).

The artist commissioned to illustrate this early sixteenth-century textbook of anatomy was clearly a very good draughtsman (and wanted the readers to know it), but equally clearly he didn't know much about internal female anatomy.

what Galen or any other authority had said. Vesalius was professor of surgery at the University of Padua and he personally dissected human cadavers in his anatomy classes, as his students clustered around him. This in itself was a major innovation. Previously, the professor would stand at his lectern reading from one of Galen's anatomical works, while the body was dissected by a local surgeon who was brought in for the occasion. Frequently, the professor would maintain his distance from the messiness of the cadaver by employing a demonstrator to point out to students the relevant part of the exposed anatomy to which the Galenic text referred (see Figure 5.3).

This was not an arrangement that was likely to uncover errors in Galen's anatomical accounts, or to make new anatomical discoveries. The surgeon who was in the best position to notice discrepancies between what Galen said and what was in front of his eyes may well have been the only person in the anatomy theatre who could not understand Latin. When Vesalius did his own dissections, talking students through the lesson, rather than reading from Galen, he was able to notice over 200 errors in Galen's authoritative account of human anatomy. Galen himself was forbidden by Roman law to dissect human bodies and inevitably drew unwarranted assumptions based on his dissection of apes and pigs. In some cases the errors were highly significant and, if noticed, would have led to the downfall of Galen's physiological theories (as eventually they did); but they were not noticed until Vesalius, Renaissance man that he was, began to look at the human body itself, rather than to accept Galen's authority (we will come back to some of these errors, and their implications, in Chapter 10 below).

Vesalius saw himself as ushering in a re-birth of the correct, and more ancient, way of doing medicine, before it became corrupted by later thinkers:

> Especially after the devastation of the Goths, when all the sciences, formerly so flourishing and fittingly practised, had decayed, the more fashionable physicians, first in Italy in imitation of the old Romans, despising the use of the hands, began to relegate to their slaves those things which had to be done manually for their patients and to stand over them like architects. Then when, by degrees, others who practised true medicine also declined those unpleasant duties – not, however, reducing their fees – they promptly degenerated from the earlier physicians ... And so in the course of time the art of treatment has been so miserably distorted that certain doctors assuming the name of physicians have arrogated to themselves the prescription of drugs and diet for obscure diseases, and have relegated the rest of medicine to those whom they call surgeons, but consider scarcely as slaves. They have shamefully rid themselves of what is the chief and most venerable branch of medicine, that which based itself principally upon the investigation of nature.

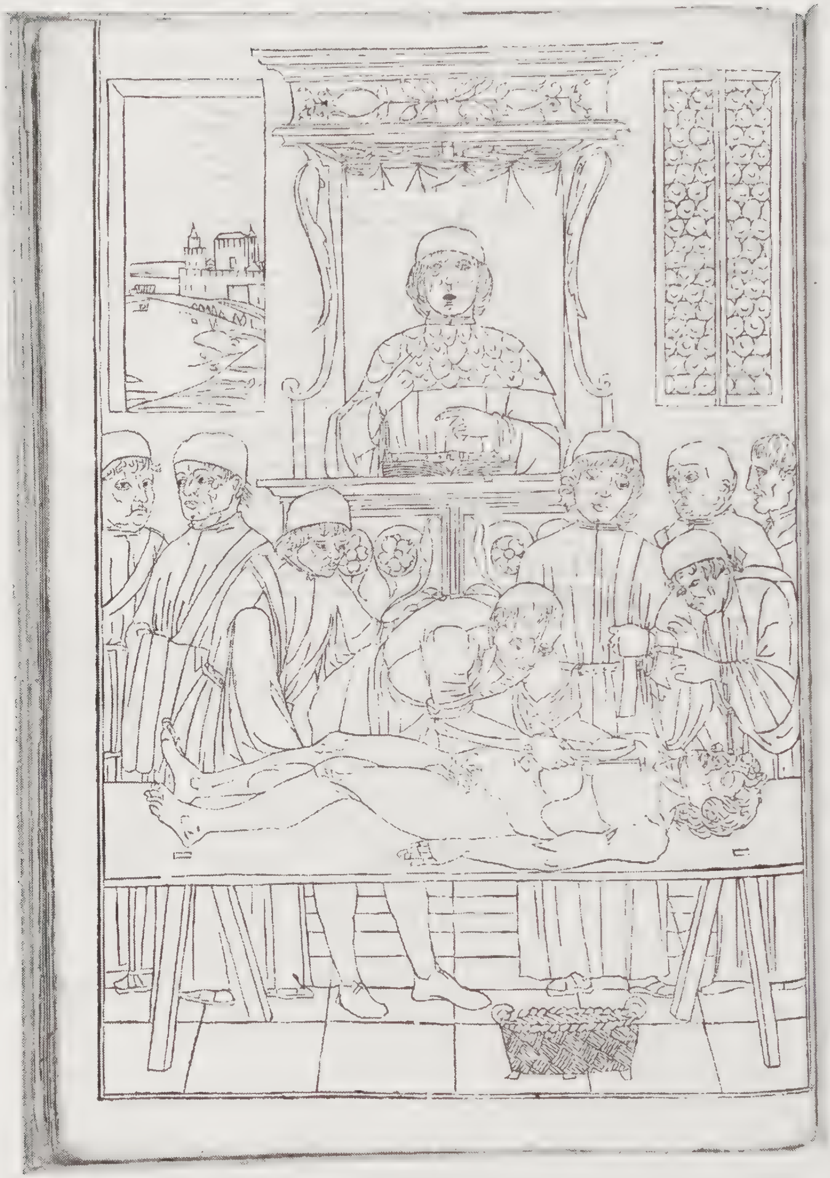

Figure 5.3 Anatomy lecture with dissection as depicted in Johannes de Ketham, *Fasciculus medicinae* (Venice, 1495). With permission from the Wellcome Trust.

As the professor reads from one of Galen's anatomical works, a barber-surgeon cuts open the body, and a demonstrator points to the appropriate anatomical part. The students, then as now, stand around paying little attention.

Figure 5.4 The Fifth Plate of the Muscles, from Andreas Vesalius, *De humani corporis fabrica* (Basel, 1543). Reproduced with permission from Edinburgh University Library, Special Collections Department (Df. 1.52*).

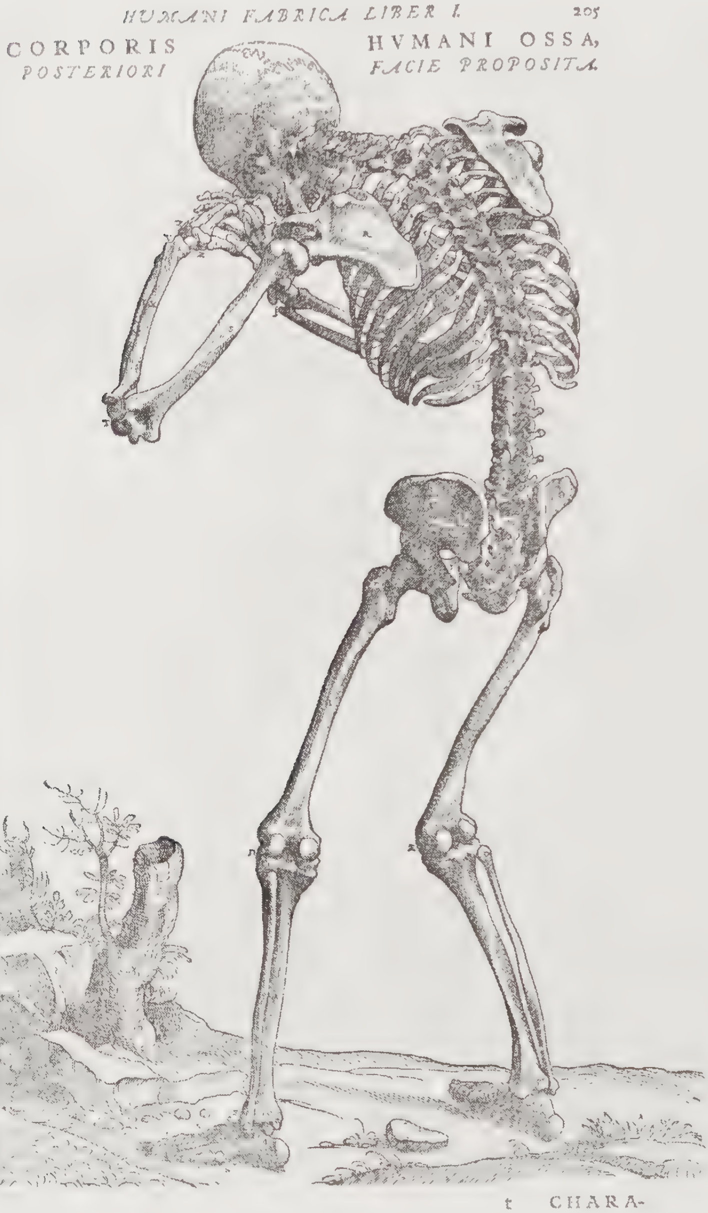

Figure 5.5 Back view of the skeleton, from Andreas Vesalius, *De humani corporis fabrica* (Basel, 1543). Reproduced with permission from Edinburgh University Library, Special Collections Department (Df. 1.52*).

It was these same physicians who, rather than investigating matters for themselves, treated Galen as an oracle:

> coupled with the failure of others to dissect, they have shamefully reduced Galen's writings into brief compendia and never depart from him – if ever they understood his meaning – by the breadth of a nail... So completely have all yielded to him that there is no physician who would declare that even the slightest error had ever been found, much less can now be found, in Galen's anatomical books, although it is now clear to me from the reborn art of dissection, from diligent reading of Galen's books and their restoration in several places that he never dissected a human body; but deceived by his monkeys he frequently and improperly opposed the ancient physicians trained in human dissection.

Furthermore, when Vesalius came to write his own anatomical text book, he did not simply want to substitute his own authority for Galen's. He wanted to be able to provide his readers with something like the experience of performing their own dissections. Accordingly, he hired a superb artist (thought to be Jan Calcar, *c.* 1499–1546, from Titian's studio), and closely collaborated with him to produce a series of illustrations successively penetrating the layers of the body, from the surface muscles to the skeleton. For the first time, an anatomy textbook appeared in which the pictures were not only artistically competent, but also useful for illustrating anatomical detail (see Figures 5.4 and 5.5).

FURTHER READING

Brian P. Copenhaver and Charles B. Schmitt, *Renaissance Philosophy* (Oxford: Oxford University Press, 1992).

Allen G. Debus, *Man and Nature in the Renaissance* (Cambridge: Cambridge University Press, 1978).

Roger French, 'The Anatomical Tradition', in W. F. Bynum and Roy Porter (eds), *Companion Encyclopaedia to the History of Medicine*, 2 vols (London: Routledge, 1993), vol. 1, pp. 81–101.

C. D. O'Malley, *Andreas Vesalius of Brussels, 1514–1564* (Berkeley: University of California Press, 1964).

K. B. Roberts and J. D. W. Tomlinson, *The Fabric of the Body: European Traditions of Anatomical Illustration* (Oxford: Clarendon Press, 1992).

W. P. D. Wightman, *Science in a Renaissance Society* (London: Hutchinson, 1972).

6

Nicholas Copernicus and a New World

Observing the natural world for oneself, rather than taking Ancient authority on trust, was undoubtedly an important aspect of Renaissance intellectual life, and one that was to have far-reaching consequences for our understanding of the natural world. But there were other ways of rejecting Ancient authority, and one of the most important of these was introduced by Nicolaus Copernicus (1473–1543), canon of the cathedral at Frombork in the Kingdom of Poland. It is also a wonderful irony of history that Vesalius published his great work of anatomy, *De humani corporis fabrica*, in 1543, the same year that Copernicus sought to reform astronomy with the publication of his *De revolutionibus orbium coelestium* (*On the revolutions of the heavenly spheres*). For historians of science 1543 was a very good year.

Copernicus tried to convince his readers that the Earth was not stationary in the centre of the world system, as had been accepted on the authorities of Plato, Aristotle, and the astronomer, Claudius Ptolemy, but that it was revolving around the Sun, along with the other planets (only Mercury, Venus, Mars, Jupiter and Saturn were known at that time), and rotating upon its own axis once every 24 hours.

We're all Copernicans now, but somewhat ironically, in view of what we were saying about Vesalius and anatomy in the Renaissance, we accept the motion of the Earth on trust. If we were to judge it by our daily experience we would surely have to admit that the Earth seems to be standing foundationally still and stable. Clearly, Copernicus was not going to succeed in reforming astronomy by appealing in a straightforward way to his contemporaries' personal experience of the natural world. Nevertheless, his epoch-making innovation in astronomy was based on his own detailed analysis of heavenly phenomena. Rather than taking anything on trust (even his own sense of what is moving and what is not), Copernicus allowed himself to be led by the astronomical evidence to reach the only conclusion he felt was compatible with that evidence.

What follows is, of course, a highly simplified account of Renaissance astronomy before and after Copernicus; the aim is simply to indicate how astonishing his achievement was.

Although Copernicus was to open up the medieval world picture and to give rise to ideas about an infinite universe, the starting point for his innovation was a change in the conception of the Earth itself which was made possible by the Renaissance voyages of discovery. Klaus A. Vogel has referred to Christopher

Columbus's voyage across the Atlantic to the 'West Indies' in 1492 as 'the first great experiment in the history of early modern science'. The aim of the 'experiment' was not, as popular consciousness would have it, to determine whether the Earth was round or flat. The aim was to confirm, as Columbus wrote, that 'Earth and water together form one round body'. According to the prevailing scholastic Aristotelian philosophy, the sphere of the Earth was floating in a larger sphere of water, with only the top hemisphere of the Earth protruding. This view was adopted in order to accommodate the Aristotelian view that the four elements should arrange themselves concentrically, with the heaviest, earth, at the centre, and then water, air and fire successively, followed by the fifth element, the element of the stars. On this view, the Earth should be completely surrounded by water, which clearly it is not. Consequently, the scholastics came to the compromise view that the Earth bobbed up through the sphere of water, presenting one hemisphere into the sphere of air.

It was assumed that Europe, Asia and Africa effectively filled this hemisphere, except that they were surrounded by an ocean which merged at the edge of the hemisphere with the greater surrounding sphere of water. The discovery of the West Indies, and subsequently of South America, eventually led to the view that this picture could not be right, and that the spheres of earth and water were in fact combined in a single terraqueous globe. Clearly, this new view of the Earth made it easier for Copernicus to claim that the Earth might be in motion around the Sun – the older view, being based on the assumption that the sphere of earth must be at the lowest point in the world system (because it is the heaviest of the five elements) was intrinsically inimical to the idea that the Earth might not be at the centre of the world system. Before launching forth into astronomical details, therefore, Copernicus presented his readers with a chapter entitled: 'How Earth Forms One Single Sphere with Water', in which he brought to their attention the conclusion arising from the voyages of discovery.

Turning then to astronomy, there were two sets of underlying factors which had to be taken into account. First, Ancient Greek natural philosophy, as represented primarily by Plato and Aristotle, had laid down the fundamental principles that the Earth is stationary in the centre of the world system, the planets move around the Earth in perfect circles, and they do so with perfectly uniform motion (that is to say, they do not speed up and slow down as they move on their circular paths). These assumptions were in turn based on the assumption that the heavens should be perfect and unchanging (remember, the Ancient Greeks found change morally, politically, and aesthetically distasteful). The rotation of a sphere, ending where it begins, is motion without involving change of place, and uniform motion means there is no other kind of change going on (no slowing down or speeding up) apart from the endless rotation of the sphere. Motions could be seen to be taking place in the heavens, but the Greeks reduced them, by these assumptions, as close to unchanging as they could.

This brings us to another important aspect of the way all pre-modern thinkers, including Copernicus, envisaged the heavens. They saw the Earth as being

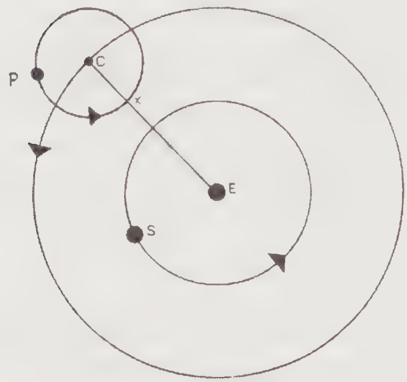

Figure 6.1 Motion of a planet on an epicycle moving around a deferent, according to Ptolemy.

Note that, for the outer planets, the line joining the Earth to the Sun was always parallel to the line joining the planet to the centre of its epicycle. This means that the planet always turns on its epicycle at the same rate as the Sun turns around the Earth. It also means that an equivalent geometrical relationship could be maintained if the planets were envisaged as revolving around the Sun (as Copernicus was to do).

surrounded by a series of nested spheres, like the layers of an onion. References to the sphere of Mars, say, are not references to the spherical globe of the planet, as we might suppose, but to the whole sphere surrounding the Earth, and upon which what we think of as the planet is carried around the Earth. It is as though the planet itself is simply a spot of luminous paint on the surface of the otherwise invisible sphere to enable astronomers to see how the sphere as a whole is moving.

Unfortunately, however, these assumptions hardly reflected what we now know to be the physical reality of planetary motions.

So, the second set of factors to be taken into account when trying to understand Renaissance astronomy is that the planets, including the Earth, move around the Sun. What's more, they move on elliptical, not circular, paths, and they continually speed up and slow down as they make their orbits (they speed up as they approach the part of their elliptical orbit which is closest to the Sun, and slow down as they enter the part of their orbit which is taking them farther away from the Sun).

As well as bearing these things in mind, you should also try to remember that pre-Copernican astronomy was based on the assumption that the Earth was absolutely stationary and so did not rotate on its axis once every 24 hours. This forced the conclusion that the sphere of the fixed stars (fixed in relation to one another, unlike the planets which move independently of one another) rotates

Figure 6.2 Typical observed path of a planet as it passes into retrograde motion (between the first and second stationary points).

around the Earth, once every 24 hours. We now know that this seeming revolution of the stars is an illusion caused by the Earth's rotation.

In view of this, you might suppose that Plato's command to the astronomers of his day to 'save the phenomena' – to explain the complex motions of the heavens in terms of uniform motion on perfectly circular paths would founder and fail. How can the heavens be accurately explained while the astronomers are hampered with false assumptions, incompatible with physical reality? In fact, Ancient Greek astronomers were nothing if not ingenious and the final summation of ancient Greek astronomy, in Ptolemy's *Almagest* (the Arabic title by which it was known to medieval thinkers), came remarkably close to fulfilling the Platonic brief.

Well, somewhat ... there were at least two massive cheats, and a few smaller ones. The first cheat was that Ptolemy made the planets move around the Earth not on a single circle but on a circle revolving around another circle. The planet was envisaged to move on a circular epicycle, while the epicycle moved around a larger circle, called a deferent (Figure 6.1). This was hardly what Plato had in mind, but it enabled Ptolemy to account not only for the fact that planets sometimes seemed farther away from the Earth than at other times (they varied in size and brightness), but also for so-called retrograde motion. The motion of the Earth sometimes made it look as though a planet had changed its usual direction of travel to move backwards for a short spell (just as a slow train might appear to be moving backwards when observed from a faster train overtaking it) – this was called retrograde motion (Figure 6.2). The combination of epicycle and deferent meant that a planet could be made to loop the loop – during which time it would move oppositely to its normal direction of travel (seen in Figure 6.3).

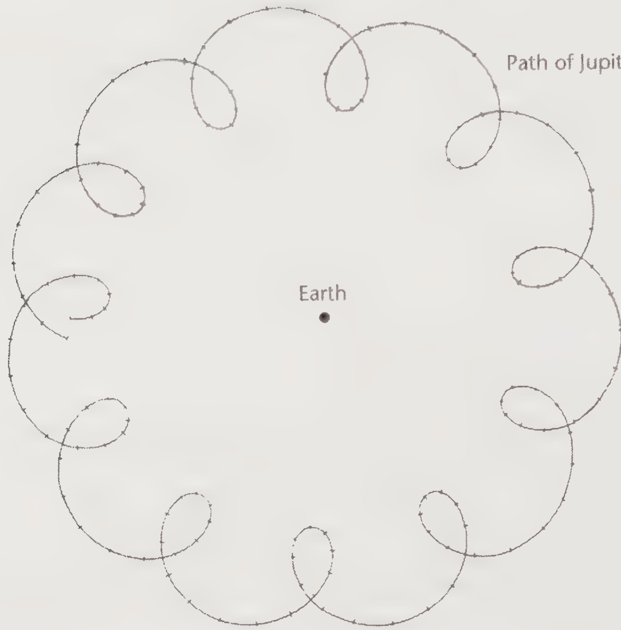

Figure 6.3 The loop-the-loop path supposedly taken by a planet moving on an epicycle.

This shows how the combination of motions upon epicycle and deferent result in a path which accounts for the retrograde motions of the planets, in this case (based upon observations between 1708 and 1720) Jupiter. It can be seen at the left that the planet, moving anti-clockwise, does not quite get back to its starting point after 12 years, according to the Ptolemaic scheme.

The other big cheat was the so-called 'equant point'. This was Ptolemy's way of dealing with the observed fact that the planets (contrary to Platonic assumptions about uniform speed) speed up and slow down as they move across the heavens. Ptolemy had already moved the Earth from the centre of rotation of the surrounding heavenly spheres (one of his smaller cheats), and in so doing had approximated the geometry of an elliptical orbit. Where a circle is defined in terms of a single focus at the centre, an ellipse is defined in terms of two foci (and a circle can be seen as a special case of an ellipse, where the two foci are coincident). So, envisaging the Earth slightly removed from the centre of the heavenly spheres, makes it look almost as though it is sitting at one focus of a surrounding ellipsoid (imagine an ellipse where the two foci are very close together, so that the ellipse is close to being circular – and its corresponding three-dimensional ellipsoid close to being spherical). Ptolemy effectively noticed that motion of a planet appeared uniform with respect to the second focus of such an imagined ellipsoid. Accordingly, he declared that, contrary to Platonic expectations, a planet did not move

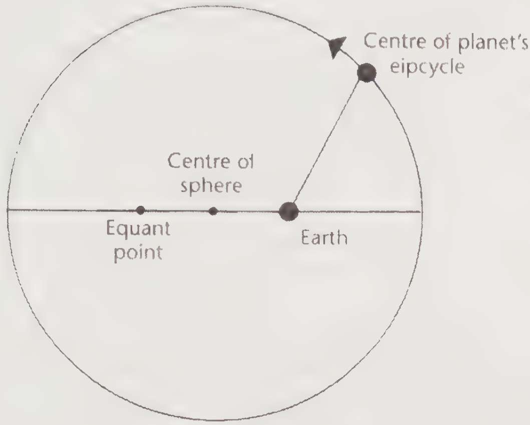

Figure 6.4 The equant point devised by Ptolemy to falsely maintain 'uniform' circular motion.

Ptolemy proposed a point in space, called the equant point, and supposed that the motion of the planet, or the centre of its epicycle, would look uniform in rotational speed when seen from that point. This works very well in explaining the actual observed movements of the planets but it implies that the planet's sphere must be alternately slowing down and speeding up as it rotates around its centre. This was considered to be physically impossible.

with uniform speed around its circular path, but only with respect to the non-central imaginary point in space, which he called an equant point.

It is important to note, however, that Ptolemy did not think in terms of ellipses or ellipsoids, or the two foci of ellipses. I have described it in these terms to indicate that he was in fact approaching close to the correct account later arrived at by Johannes Kepler (1571–1630) (see Chapter 8). Ptolemy himself simply saw it in terms of a surrounding sphere with the Earth's position, and the equant point placed symmetrically on either side of the sphere's centre. The important thing to note is that, although Ptolemy laid claim to 'saving' the supposed uniform motions of the planets, this was always regarded as a betrayal of Platonic principles of uniform planetary motion, and an unacceptable cheat. It meant that the planetary sphere had to rotate sometimes faster and sometimes slower in order to look as though it was moving uniformly with respect to the equant point (Figure 6.4).

Now, it might be supposed that Ptolemy's attempts to 'save the phenomena' would be seen as unacceptably fudging the phenomena, and rejected out of hand. It is clear, for example, that Ptolemy's system of epicycles and deferents, of wheels within wheels, is hardly compatible with the neat picture depicted in Figure 4.1. But in fact Ptolemy's devices formed the basis of the technical art of astronomy from his own day to the late Renaissance. The reason for this is not hard to fathom. Quite simply, Ptolemy's system was accepted *because it worked* – it enabled the calculation of planetary positions both retrospectively and in the future. This was important not only for calendrical purposes, and navigation, but also for drawing up horoscopes and making astrological prognostications.

Even so, the learned were not fooled by it. As the highly respected Andalusian philosopher, Averroes, had pointed out:

> The astronomical sciences of our days offer nothing from which one can derive an existing reality. The model that has been developed in the times in which we live accords with the computations not with existence.

Astronomy could tell us how to calculate planetary positions, but it could tell us nothing reliable about the real, physical, structure of the world system. One important aspect of this inability to describe the real structure of the universe was the fact that it was impossible to allocate a fixed order to the planets in the Ptolemaic scheme. The order of the planets outwards from the Earth was established only by convention, but as long as the relative proportions between a planet's deferent and its epicycle were maintained, the planet could be set at any distance from the centre. This reinforced the view that Ptolemy had merely provided a set of geometrical tricks for calculating astronomical positions, but had offered no information about how things in the heavens really were. As the German astronomer, Georg Rheticus (1514–76), said:

> What dispute, what strife there has been until now over the position of the spheres of Venus and Mercury, and their relation to the Sun. But the case is still before the judge. Is there anyone who does not see that it is very difficult and even impossible ever to settle this question while the common hypotheses are accepted? For what would prevent anyone from locating even Saturn below the Sun, provided that at the same time he preserved the mutual proportions of the spheres and epicycle, since in these hypotheses there has not yet been established the common measure of the spheres of the planets, whereby each sphere may be geometrically confined to its place?

As Copernicus himself put it, Ptolemaic astronomers could not even solve 'the principal consideration',

> that is, the structure of the universe and the true symmetry of its parts. On the contrary, their experience was just like someone taking from various places hands, feet, a head, and other pieces, very well depicted, it may be, but not for the representation of a single person; since these fragments would not belong to one another at all, a monster rather than a man would be put together from them.

One way of seeing the dissatisfaction with Ptolemy of Copernicus and other Renaissance thinkers was in terms of a regrettable separation between astronomy and cosmology. While Ptolemy had sought to establish a set of geometrical exercises which put the art of astronomy on an acceptably accurate and pragmatically useful footing, he had completely lost sight of the Platonic and Aristotelian desire to establish a true understanding of the world system. He was concerned with a practical art, astronomy, not with a true science, cosmology.

Nicolaus Copernicus, cosmologist?

Copernicus, like Vesalius and other Renaissance thinkers, wanted to see a rebirth of the way the Ancients did things. In his case, he wanted to see a rebirth of the Ancient attempts to establish a true cosmology. This was an enterprise that had been all but abandoned by the pragmatist Ptolemy, and by professional astronomers ever since. This was a bold ambition for Copernicus. His reputation was as an astronomer, a mathematician, and as such his contemporaries would have seen him as unqualified to pronounce on natural philosophical matters; cosmology was the preserve of natural philosophers, not astronomers (see Box 6.1).

There was a sharp separation between mathematics and natural philosophy at this time. Aristotle had insisted that the aim of natural philosophy was to explain natural phenomena in terms of physical causes. Mathematics, however, could say nothing about physical causes, and merely offered a special kind of technical description of various phenomena. The geometry of the deferent and epicycle of Venus, and their respective motions, could be used to determine when and where Venus would appear in the sky, but this said nothing about what kept Venus in motion, how it stayed on its epicycle, or why it moved with the speed it did, or indeed whether it really was moving on an epicycle at all. According to Aristotle, therefore, mathematics was incompetent with regard to natural philosophy and, in the scholastic university tradition, natural philosophy was generally pursued without reference to mathematical considerations.

This separation of natural philosophy and mathematics on the one hand seemed to be confirmed by the clear incompatibilities between Aristotelian cosmology and Ptolemaic astronomy, and on the other hand it ensured the perpetuation of the separation between cosmology and astronomy. Copernicus wanted to bring this separation to an end. He wanted to reform astronomy in such a way that it could not only provide accurate calculations of planetary movements and positions, but also reveal the real structure of the world system.

Presumably the initial impetus to reform was the general awareness among Copernicus's learned contemporaries that things were seriously amiss in astronomy. Discrepancies resulting from Ptolemy's inaccuracies had built up so that, for example, the vernal equinox (necessary for computing the date of Easter Sunday) was occurring ten days before it should. What's more, full Moons, as

Box 6.1 THE MEDIEVAL SEPARATION OF ASTRONOMY AND COSMOLOGY

Ptolemy can be seen to have abandoned the tradition of cosmology, in favour of a merely pragmatic astronomy. This separation was reinforced in the Middle Ages when Aristotelian cosmology became an accepted part of natural philosophy, but astronomy was pursued by mathematical practitioners, and regarded merely as an art with some practical uses. Copernicus, wanted to raise the intellectual status of astronomy and to re-unite it with cosmology, as he supposed it was before Ptolemy.

Cosmology

Endeavour to describe and explain the structure of the world system. Characterized as *philosophical.*

Pythagoras (6th century BC)
Plato (427–347 BC)
Aristotle (384–322 BC)

Absorbed into Christian teaching, especially after St Thomas Aquinas (1225–74). See Figure 4.1

Arguments based upon aesthetic and metaphysical considerations, underlying which was a 'religious' conviction that the Cosmos was created by a 'divine' intelligence guided by rational/geometrical principles (i.e. simple, harmonious, non-random, non-chaotic)

Considered to be a **science**

Astronomy

Endeavour to provide detailed underpinning for (and therefore confirmation of) cosmological theories by analysing motions of heavenly bodies in geometrical terms. Characterized as *mathematical.*

Claudius Ptolemy (c. AD 100–65)
All medieval astronomers (until Copernicus).

Arguments based exclusively upon pragmatic considerations, i.e., does it work? By the Christian era the main task for professional astronomers was to calculate the motions of heavenly bodies for the future and the past in order to establish:
(a) times of Church festivals (Easter, etc.)
(b) horoscopes
(c) navigation

Considered to be an **art**

calculated, bore no relation to what could be seen in the night sky (and any idiot can tell when there's a full Moon). So, Easter Sunday was being celebrated each year on the wrong day (according to the rules laid down by the Council of Nicaea in AD 325 – which had declared that Easter Sunday should be the first Sunday after the first full moon to fall on or after the vernal equinox).

Clearly, something had to be done, and in 1513 the Fifth Lateran Council of the Church called for reform of the calendar. Copernicus, who was known to be a formidable astronomer, was invited to take part in this enterprise but evidently declined.

Perhaps he saw this as an entirely pragmatic exercise, in the Ptolemaic mould, and preferred to carry on working alone to re-unite astronomy and cosmology.

There are rival theories as to how Copernicus was led to put the Sun in the centre of the world system and the Earth in motion, but one possibility is that he realized that the system could be simplified by putting all the planets in revolutions about the Sun, while the Sun revolved around the Earth. There was already a long-standing post-Ptolemaic tradition that Mercury and Venus revolved around the Sun, as it moved around the Earth, so extending this to the other planets was not such a hard step to envisage (see also the caption for Figure 6.1).

At some point he also set the Earth rotating on its axis once every 24 hours. This had the advantage that the whole world system, including the sphere of the fixed stars, was no longer required to rotate around the Earth each day every day. Although it might seem absurd to make the Earth rotate on its axis, the alternative supposition, that the sphere of the stars rotated once a day seemed even more absurd, given the almost inconceivably vast size of that sphere. Again, we don't know the sequence of events but Copernicus also realized that the precession of the equinoxes, discovered by Hipparchus in the second century BC and assumed to involve a gradual shift of the whole sky around the Earth, could be explained by a gradual movement in the orientation of the Earth's axis.

Depending upon the order of Copernicus's thinking, having attributed one motion, or possibly two, to the Earth, Copernicus was perhaps emboldened to put it in motion around the Sun, together with the other planets.

The crucial aspect of this last step – putting the Earth in motion around the Sun – was that it immediately suggested that Copernicus had not just a workable astronomy but a seemingly coherent cosmology. Unlike the Ptolemaic system, Copernicus's heliocentric system established a fixed order of the planets. The geometry of the heliocentric world system allowed only one possible arrangement of the planets, and that arrangement precisely matched the order suggested by the increasing solar periods of the planets (*i.e.* the period taken for each planet to complete a revolution of the Sun) – from Mercury's eighty-eight days to Saturn's thirty years.

It was the harmonious nature of heliocentric cosmology that convinced Copernicus that his astronomy must be correct even though it depended upon the seemingly impossible idea that the Earth was in motion. As Copernicus himself wrote:

> Having thus assumed the motions which I ascribe to the Earth ... by long and intense study I finally found that if the motions of the other planets are correlated with the revolving of the Earth, and are computed for the revolution of each planet, not only do their phenomena follow therefrom, but also the order and size of all the planets and spheres, and heaven itself is so linked together that in no portion of it can anything be shifted without disrupting the remaining parts and the universe as a whole.

Figure 6.5 The heliocentric world system as depicted in Nicholas Copernicus, *De revolutionibus orbium coelestium* (Frankfurt, 1543). Reproduced with permission from Edinburgh University Library, Special Collections (JY 730).

In fact, Copernicus's cosmology still depended upon some mathematical messiness. Being inspired by the original Platonic plea to 'save the appearances', Copernicus continued to believe that only perfectly circular and perfectly uniform motions were allowed. Accordingly, he had to use the Ptolemaic devices of eccentrics and epicycles (although Copernicus only required small epicycles – effectively to bring things closer to elliptical orbits, though he did not realize that – whereas in some cases Ptolemy had required huge epicycles). He refused to use the equant point, however, which he saw as a betrayal of the principle of uniform motion. Even so, Copernicus's system wasn't as neat and tidy as he depicted it (Figure 6.5).

Furthermore, in spite of what at first might look like cosmological advantages, it was even less compatible with the Aristotelian world-picture than the Ptolemaic system simply because it required the Earth to be moving.

Ironically, this last point did not matter to the astronomers of Copernicus's day. They just wanted a more accurate way to calculate planetary positions than the old Ptolemaic way, and they certainly found this in Copernicus's *De revolutionibus*. The irony was, however, that hardly any of them believed the Earth

Box 6.2 WAS COPERNICUS A REVOLUTIONARY OR CONSERVATIVE?

Copernicus has been described as a 'timid canon' (he was a canon of the Cathedral at Frauenburg), who was too conservative to initiate the real astronomical revolution, which had to be carried through by others after him.

The arguments for his conservatism are these:

1 He emulated Ptolemy's style of presentation.
2 He relied upon Ptolemy's observations.
3 He used Ptolemy's mathematical techniques (except for the equant)
4 He rejected the use of the equant on the grounds that it betrayed ancient Greek principles.
5 He withheld his work from publication for over forty years.
6 He referred to earlier Greek authorities as the originators of his theory.
7 His book included a Preface in which the work was said to be only hypothetical and not an account of how the world system really is.

In fact, Copernicus can be defended on all of these counts. With regard to the first two, for example, he did this for the benefit of his fellow astronomers. They were all used to using Ptolemy so he followed the same general pattern to help them follow him. He used Ptolemy's examples of how to calculate the movements of a particular planet (based on, say, five successive observations by Ptolemy) to show how his system could come up with just as good results. If he had offered his own observations and completely new examples, astronomers used to Ptolemy's works might not know how to move from Copernicus's unfamiliar demonstrations back to Ptolemy's more familiar examples. There is no evidence that he held his book back. On the contrary, evidence suggests that he was continually working on it and trying to improve it. Contemporaries believed that the Ancients, being closer to Adam, knew more than they did, so if Copernicus hoped to persuade his contemporaries of the truth of his system he had to find evidence that some of the Ancients believed it. The Preface in question was written without Copernicus's knowledge or permission.

What's more, there are two major aspects of his work which suggest he was very definitely a revolutionary thinker:

1 His aim was to reinstate the science of cosmology, not merely to come up with another mathematical system.
2 He privileged mathematics over physics, insisting that the Earth must be in motion because the mathematics demanded it, even though he could not provide a physical explanation of how it could be in motion.

really was in motion. They were used to the idea that astronomy was just a set of calculating techniques which bore no relation to reality. They thought of Copernicus's system in the same way – by pretending the Earth is moving they could arrive at accurate reconstructions of planetary movements, but very few were led to suppose, even for a moment, that the Earth really was, or ever could be, in motion.

This kind of thinking came naturally to astronomers, but if any readers of *De revolutionibus* believed they had to conclude that the Earth is in motion, they would have been put right by the preface added to Copernicus's book by a Lutheran churchman, Andreas Osiander (1498–1552), who was entrusted to supervise the printing of the book in Frankfurt (Copernicus was unable to travel to the German printshop himself). Osiander left his preface unsigned and so readers naturally assumed Copernicus had written it. Osiander reminded people of the traditional separation between astronomy and cosmology, or between mathematics and natural philosophy. Astronomy, he insisted, is just meant to help calculate planetary movements, and cannot tell us anything about the way the universe really is arranged:

> For it is the duty of an astronomer [Osiander wrote] to compose the history of the celestial motions through careful and expert study... *Since he cannot in any way attain to the true causes,* he will adopt whatever suppositions enable the motions to be computed correctly from the principles of geometry for the future as well as for the past. The present author has performed both these duties excellently. For these hypotheses need not be true nor even probable.

The labours of Copernicus's lifetime were reduced by Osiander to hypotheses which were 'not put forward to convince anyone that they are true, but merely to provide a reliable basis for computation'. Copernicus was not a cosmologist, Osiander insisted, he was an astronomer just like Ptolemy, and nothing more.

Copernicus died within months of his book appearing in print, and so he was unable to make it clear to his contemporaries that he really did believe the Earth was in motion. Even so, some readers did see what he was attempting, and were to take up his cause. As far as we can tell, only eleven thinkers before 1600 believed that Copernicus had shown in his *De revolutionibus* that the Earth really was in motion. We'll consider three of these, and their contributions to the Copernican revolution, in later chapters.

FURTHER READING

Michael J. Crowe, *Theories of the World from Antiquity to the Copernican Revolution* (New York: Dover, 1990).

William Donahue, 'Astronomy', in Katherine Park and Lorraine Daston (eds), *The Cambridge History of Science Volume 3 Early Modern Science* (Cambridge: Cambridge University Press, 2006), pp. 562–95.

John Henry, *The Scientific Revolution and the Origins of Modern Science*, 3rd edn (Basingstoke and New York: Palgrave Macmillan, 2008), chapter 2, pp. 12–17.

John Henry, *Moving Heaven and Earth: Copernicus and the Solar System* (Cambridge: Icon Books, 2001).

Rocky Kolb, *Blind Watchers of the Sky: The People and Ideas that Shaped Our View of the Universe* (Oxford: Oxford University Press, 1999).

Thomas S. Kuhn, *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought* (Cambridge, MA.: Harvard University Press, 1957).

John North, *The Fontana History of Astronomy and Cosmology* (London: Fontana Press, 1994).

Klaus A. Vogel, 'Cosmography', in Katherine Park and Lorraine Daston (eds), *The Cambridge History of Science, Volume 3: Early Modern Science* (Cambridge: Cambridge University Press, 2006), pp. 469–96.

7

New Methods of Science

The period from the so-called 'high Renaissance', in the sixteenth century, to the end of the seventeenth century has been seen by historians of science as a period of special significance, when speculative natural philosophy, as it was conducted in the universities by scholastic philosophers, was either drastically reformed, or (depending on your point of view) replaced, by an approach to the understanding of the natural world which was much more recognizably like modern science. Accordingly, historians have called this period the Scientific Revolution. Andreas Vesalius and Nicolaus Copernicus are two of the most prominent figures in this extended period of revolutionary change, and I hope you can already understand why. They should not be seen, however, as unique and isolated geniuses towering over their contemporaries. They were not 'ahead of their time', but very much emergent from their time – even though they undoubtedly emerged further than many of their fellow anatomists and astronomers.

In the last two chapters I used them as examples of Renaissance innovators, rejecting authority, trying to discover the truth by their own efforts of observation or analysis. But I also tried to make it clear that these new approaches were aspects of what can be seen as a characteristically Renaissance endeavour. There were many thinkers in the high Renaissance who played their roles not only in contributing to new discoveries, but also in changing ideas about the best way to gather reliable and accurate knowledge of the natural world.

Vesalius and Copernicus also illustrate for us the two most far-reaching new ways of acquiring natural knowledge; direct observation and mathematical analysis. We will come back to the mathematical approach later, to see how and why the previous separation of natural philosophy from mathematics was replaced by a new amalgamated, physico-mathematical method. For now, we are going to look at the way emphasis upon experience and direct observation transformed understanding of the natural world and led to the rejection of scholastic Aristotelianism.

* * *

It has long been recognized by historians of science that one of the major innovations of the Scientific Revolution was the development of the experimental method. This method is still seen as a defining feature of modern science (nobody accepts the existence of a theoretically proposed sub-atomic particle, for example, until it is actually confirmed experimentally – even if it means building something like the Large Hadron Collider). It is important to note, however, that the suggestion is *not* that the experimental method was newly invented during the Scientific Revolution; to claim this would be as ridiculous as claiming that mathematics was invented in the Scientific Revolution. The point is that natural philosophers appropriated the experimental method from other traditions, and in so doing transformed natural philosophy (formerly effectively an armchair speculative pursuit) into something much closer to what we would recognize as science.

But, what were the traditions where experience, or even the experimental method, were already in use? One of these, to a limited extent, was the medical tradition. We saw in the last chapter that dissections of human cadavers was a feature of medical training throughout the Middle Ages, even though, before Vesalius, it was conducted in a way that made it an ineffectual adjunct to the authoritative words of Galen being read out by the professor. But medical schools also incorporated elements of bedside training for their students, so they could experience the contingencies of medical practice alongside the securities of medical theory. We will come back to developments in medicine later (Chapter 10), so let us look now for other sources of the experimental method.

Judging from the historical record, one of the most important sources for the experimental method was the magical tradition (or perhaps we should say, various traditions that could loosely be lumped together as occult or magical traditions). Alchemy has its own continuous history, for example, from the Ancient Egyptians through to the Ancient Greeks and on through Islamic civilization where it flourished, and into the Western Middle Ages and the Renaissance. We cannot pursue this history here but suffice it to say that it was always pursued in an entirely experimental way (sometimes in a space which the alchemist referred to as his 'elaboratory'), and depended upon a vast array of specialized apparatus.

Alchemy was a highly specialized pursuit, but there was a much more familiar aspect of the occult (of which alchemy was a special case) known as 'natural magic'. This was the predominant form of magic practised in the Latin West and it was simply based on the assumption that there are hidden (or 'occult') connections between things, and that these connected things can affect one another in various ways. The role of the natural magician was to discover these connections and the way connected things affected one another, with a view to exploiting this knowledge by putting it to some practical purpose. Most commonly this would involve knowledge of herbs with a particular medicinal property (many of which were genuinely efficacious), but it could also involve knowledge of how to create optical illusions using mirrors, or even how to accomplish feats of strength using carefully constructed machinery.

These things do not look like magic to us because we have re-designated this kind of knowledge as pharmacy, geometrical optics, or mechanics. The important thing to note, however, is that these things were not included in university natural philosophy, and so accomplishing things with machines could be deemed to be, quite literally, technological wizardry. Scholastic natural philosophy, following Aristotle's classifications, was considered to be concerned with natural phenomena, and machines were clearly artificial. The magicians who were capable of designing and making such machines knew that they worked by exploiting natural powers or forces, such as gravity (just as clockwork is driven by the slow descent of a weight – a weight which has to be wound back up to the top of its travel every so often to keep the clock going), but this still did not earn them a right to be considered in university natural philosophy; machines worked by nature acting under constraint, not nature acting naturally.

Generally speaking, there were two major aspects of the natural magic tradition which set it apart from scholastic natural philosophy. First, it was assumed that knowledge of the hidden connections between things, their powers over one another, could only be discovered – because they were otherwise hidden – by experience. Second, this knowledge had practical uses – either for good or ill – and could be used by its possessor to accomplish desired outcomes. Scholastic natural philosophy by contrast, was generally regarded as 'knowledge for its own sake', if it had any use it was merely to reveal the hand of God in the natural world. Similarly, according to the scholastics, explanations of natural phenomena had to be given in terms of manifest, and indeed undeniable, causes; explanation in terms of occult causes was, by definition, not possible. Scholastics did not deny the existence of occult causes, but they tried to avoid them in natural philosophy.

Scholastic strictures against the occult began to collapse, however, during the Renaissance. There were a number of reasons for this, not least increasing awareness that many natural phenomena simply could not be explained in terms of the so-called manifest qualities, hot, cold, dry and wet (embodied in the four elements, fire, air, water and earth). Another major factor, however, was the discovery, during the Renaissance search for Ancient texts, of a group of writings attributed to Hermes Trismegistus, Thrice-Great Hermes, an Ancient sage evidently identified by the Ancient Greeks with their god, Hermes, and supposed by Renaissance thinkers to be a contemporary of Moses.

These Hermetic writings created great excitement among many Renaissance thinkers because their Ancient pagan author seemed to foreshadow the belief in a Holy Trinity, and hinted at other Christian beliefs. We can now see that this is hardly surprising because these documents were written in the early centuries of the Christian era by anonymous Neoplatonists who were undoubtedly aware of Christian doctrine. Renaissance humanists, however, believing these works long pre-dated the ministry of Jesus, could hardly fail to conclude that the author was a pagan thinker who arrived at Christian truths through philosophical reasoning.

Alongside these Ancient theological writings, there were also a huge number of alchemical and other occult writings that were also attributed to Hermes Trismegistus. Here then, was a supreme Ancient sage, contemporary with Moses, who on the one hand foreshadowed Christian doctrines, and on the other operated as a keen and accomplished natural magician. Given the general assumption that the further back in time a thinker lived, the closer he was to Adam and the Adamic wisdom which gradually began to be lost after the Fall (see Chapter 4), the Hermetic writings, theological and magical alike, came to be seen as a supreme source of Ancient wisdom. Magic, after centuries of condemnation by the Church (because of its association with demons), came to be seen by many leading thinkers as one of the most respected sources of knowledge and wisdom (see Box 7.1).

The scene was set, therefore, for the input of magical ideas and approaches into natural philosophy. But even now, this would not occur spontaneously – scholastic philosophers and university curricula were set in their ways. What was required was a thinker of sufficient stature to be able to persuade contemporary natural philosophers to pay serious attention to the alternative tradition of natural magic and its experimental method. The thinker who emerged to fulfil this important role was an aspiring English statesman during Elizabeth I's reign, who eventually rose to the office of Lord Chancellor in the reign of James I. His name was Francis Bacon (1561–1626).

Francis Bacon and the reform of learning

Bacon has a major place in the history of modern science even though he never made a single scientific discovery and, from a modern perspective, carried out no real scientific work. What is more, seen from our viewpoint, he made a number of bad judgements about the science of his day. He rejected the Copernican theory, for example, dismissed as unreliable the experimental work of his contemporary William Gilbert (whose work is discussed below), and rejected the idea that mathematics could be useful for understanding the workings of the physical world.

His claim to fame in the history of science is that he was an ambitious programmatist and propagandist for a reformed way of doing natural philosophy. Although never carried out, his vision for reform set up resonations in the thought of subsequent thinkers, and thereby played a major role in introducing not only the experimental method but also the idea that knowledge should be practically useful.

One of the keys to understanding Bacon's vision for reform was his up-bringing by his parents. He was raised and educated by his mother until the age of 12. She was the daughter of a leading family of English Protestants, and saw to it that Bacon was thoroughly steeped in Calvinist theology. His father, Nicholas Bacon, was Lord Keeper of the Privy Seal to Elizabeth I and, as such, one of the leading statesmen in her government. Nicholas groomed his son for similar high office in government.

The result of this upbringing was that Bacon combined Calvinistic apocalyptic visions of the end of the world, with ambitions to form a new government

Box 7.1 The theory of magic

The occult powers attributed to things are supposed to be *natural* powers, *not supernatural*. The assumption was that God had bestowed these powers on things when he created them, and therefore, although they were occult, they were nonetheless part of God's creation, and part of nature. Consider the emphasis on nature in the following quotations.

> The whole course of Nature could teach us by the agreement and disagreement of things either so to sunder them, or else to lay them together by the mutual and fit applying of one thing to another, as thereby we do strange works. (Giambattista della Porta, *Natural Magick*, 1589).

> Magicians are careful explorers of nature, only directing what nature has formerly prepared, uniting actives to passives and often succeeding in anticipating results so that these things are popularly held to be miracles when they are really no more than anticipations of natural operations. (Cornelius Agrippa, *Of the Uncertainty and Vanity of the Sciences*, 1526).

> Towards the effecting of works, all that man can do is put together or part asunder natural bodies. The rest is done by Nature working within. (Francis Bacon, *New Organon*, Part I, Aphorism 4, 1620).

> Magic is an art or technique which by using the power in creation rather than a supernatural power produces various things of a marvellous and unusual kind, the reason for which escapes the senses and ordinary comprehension. (Martin del Rio, *Disquisitions on Magic*, 1608).

This *Natural* Magic was held to be the only kind of magic. Popular conceptions of magic today assume that it was based on supernatural powers, but in the premodern period it was believed that only God could perform the supernatural.

'department' concerned with the gathering and exploiting of natural knowledge. The link between these two things was a prophecy in the Book of Daniel, in the Old Testament, that before the 'last times', when the world would be brought to an end, 'many would go to and fro and science will advance'. Bacon believed that the first part of the prediction, 'many would go to and fro', was a reference to the Renaissance voyages of discovery (by Christopher Columbus, *c.* 1451–1506, Vasco da Gama, *c.* 1460–1524, and others) and was therefore already fulfilled. Bacon wanted to initiate the second half of the prediction by ensuring that science would advance. Being a devout believer, Bacon thought the end of the world was a good idea – because after that the faithful would live forever in a blessed state.

His apocalypticism can be seen, for example, in a comment in his most significant work, the *Novum organum*, or *New Organon* of 1620:

Demons, including the Devil himself, did not have supernatural powers – they were God's creatures, and therefore part of nature themselves. All they could do to accomplish things was to exploit their knowledge of natural magic. Because they were immortal beings and had been around since before Adam and Eve, it was assumed that they knew more magic than even the greatest human magus, but apart from that, they had no special powers. Although, being spirits, they could do things that we cannot, such as walk through walls, fly, and so forth.

> Though the Divell indeed, as a Spirit, may do and doth many things above and beyond the course of some particular natures: yet doth he not, nor is he able to rule or command over generall Nature, or infringe or alter her inviolable decrees in the perpetual and never-interrupted order of all generations... For Nature is nothing else but the ordinary power of God in all things created, among which the Divell being a creature, is contained, and therefore subject to that universall power. (John Cotta, *The Triall of Witch-Craft*, 1616).

> ...in nature there be some properties, causes, and effects... most familiar unto him [the Devil], because in themselves they be no wonders, but only misteries and secrets, the vertue and effect whereof he hath sometime observed since his creation... The Devil has exquisite knowledge of all natural things, as of the influences of the starres, the constitutions of men and other creatures, the kinds, vertues, and operation of plantes, rootes, hearbes, stones, etc., which knowledge of his goeth many degrees beyond the skill of all men, yea even those that are most excellent in this kind, as Philosophers and Physicians are. (William Perkins, *Discourse of the Damned Art of Witch-Craft*, 1618).

> Demons operate nothing except by natural application of active forces to the appropriate and proportionate passive objects, which is the work of nature. (Francesco Giuntini, *Speculum astronomiae*, 1573).

> God forbid that we should give out a dream of our own imagination for a pattern of the world: rather may He graciously grant to us to write an apocalypse or true vision of the footsteps of the Creator imprinted on his creatures ... Wherefore if we labour in thy works with the sweat of our brows thou wilt make us partakers of thy vision and thy Sabbath.

The Sabbath that Bacon has in mind here is not at the end of the week, but the ultimate Sabbath at the world's end, when all the faithful can rest forever from their labours.

His desires to set up a civil service department concerned with gathering scientific knowledge can be seen, for example, in the preface he wrote for a piece called *The Interperetation of Nature* (1603):

> Believing that I was born for the service of mankind, and regarding the care of the commonwealth as a kind of common property which like the air and the water belongs to everybody, I set myself to consider in what way mankind might best be served, and what service I was myself best fitted by nature to perform. Now, among all the benefits that could be conferred upon mankind, I found none so great as the discovery of new arts, endowments, and commodities for the bettering of man's life.

As we read on it becomes clear that Bacon's ambition is not to make specific discoveries, but to develop a new method of discovery:

> if a man could succeed, not in striking out some particular invention, however useful, but in kindling a light in nature – a light which should in its very rising touch and illuminate all the border-regions that confine upon the circle of our present knowledge; and so, spreading further and further should presently disclose and bring into sight all that is most hidden and secret in the world – that man (I thought) would be the benefactor indeed of the human race – the propagator of man's empire over the universe, the champion of liberty, the conqueror and subduer of necessities.

Bacon's plan for reform of natural knowledge was called by him *The Great Instauration*, and it really was something that could only be successfully accomplished by the collaboration of many hands and minds. As he wrote in the dedication to King James:

> I have a request to make – a request no way unworthy of your Majesty, and which especially concerns the work in hand; namely, that you who resemble Solomon in so many things ... would further follow his example in taking order for the collecting and perfecting of a Natural and Experimental History true and severe (unencumbered with literature and book-learning), such as philosophy may be built upon – such, in fact, as I shall in its proper place describe: that so at length ... philosophy and the sciences may no longer float in air, but rest on the solid foundation of experience of every kind.

The Great Instauration was to consist of six stages, culminating in a 'New Philosophy', but Bacon only made any headway with two of them: the *New Organon or Directions concerning the Interpretation of Nature*, which was a new method, or logic, for analysing scientific data; and 'The Phenomena of the Universe: or a Natural and Experimental History for the Foundation of Philosophy', which was to be a massive database consisting of everything known about every phenomenon of nature. It was this projected database of all knowledge that required a new ministry in the government, equipped with its own extensive staff of civil servants. The value of Bacon's projected reform of natural knowledge had never been recognized by Queen Elizabeth's administration, and nor was it by James I. Consequently, he had to work alone in gathering his database and, needless to say, he never accomplished much.

His most important publication in the history of science was the *Novum Organum*, published in 1620. The title implies that it is intended to replace Aristotle's *Organon* – the collective title given to Aristotle's various logical works. It was here that Bacon outlined his new method, including his somewhat idiosyncratic form of experimentalism.

The Baconian method

Bacon was convinced that the main source of error in our attempts to understand nature was a tendency to jump to hasty conclusions before all the possibilities have been considered. 'The understanding', he wrote, 'must not therefore be supplied with wings, but rather hung with weights, to keep it from leaping and flying'. In practice, what this meant was a prolonged period of fact gathering under a strict self-denying ordinance, to avoid any premature theorizing to explain the assembled facts. Ideally, it is only when all possible information has been gathered that the next stage of the enterprise takes place. This involves arranging the facts in vast and complex 'tables of discovery' in order to bring out obvious similarities, or dissimilarities, or other relationships between things. By surveying all the digested and organized facts in these tables, Bacon believed that the underlying patterns of things, the underlying connections between things, would be revealed.

Unfortunately, Bacon only managed to provide one example of how his 'tables of discovery' might work. He provided a partial table of as many examples of hot things, and things and processes that generate heat, as he could think of. Looking through them he noticed that the generation of heat always seemed to be associated with vigorous movement. Accordingly, he came to the conclusion that heat is motion.

Remarkably, modern scientists would agree that Bacon was essentially right. The trouble is, however, that the connection between heat and motion can now be clearly explained in terms of the speed of motions of molecules or atoms. When Bacon announced that heat is motion there was no such explanation available to him. He was simply basing his conclusion on inductive grounds. A contemporary might well have been justified in saying that Bacon's pronouncement made no real sense. Heat is heat and motion is motion. The terms cannot be used interchangeably; the motion of a falling body is not the same thing as the heat of a falling body.

But Bacon's insistence on the use of induction was an important innovation. According to Aristotle the deductive syllogism was the only acceptable logic to use in science; this fitted with his demand that natural philosophy should provide explanations based on obvious causes. Bacon pointed out that no new knowledge can ever emerge from this procedure, because the conclusion of a syllogism is always contained implicitly in its premises (see Box 7.2 below on Bacon's logic). In inductive logic, by contrast, the conclusion is a new piece of information that was not known at the outset. This was the form of logic which went hand in hand with Bacon's experimental trial-and-error methods, but it was, of course, subject to a major criticism (which is why Aristotle rejected it). There can never be any assurance that a conclusion arrived at by induction is certainly true. Potentially, any inductive conclusion can be undermined by a subsequent observation. All swans were white, according to inductive experience, until black ones were found living happily in Western Australia.

Bacon was aware of this problem of course, but he hoped to be able to overcome it by devising a form of induction which not only detailed positive cases but also ruled out all possible alternative accounts (in the case of swans, for example, Baconian naturalists would have to systematically search every part of the world to make sure there were no swans of any other colour anywhere; once they had done that, they could be sure that all swans were indeed white). In his exercise to determine the nature of heat and how it is generated, Bacon took care to gather information about cold bodies and things that were supposed to generate coldness to make sure that none of these things operated by means of motion. Sadly, the problem of induction is still being disputed by philosophers today, and Bacon certainly never managed to overcome it. Nonetheless, inductive logic came to be recognized as an important element in scientific thinking, and it will appear again more than once in later pages of this book.

But where does experimentation fit in Bacon's method? Readers with a scientific background will know that most experiments in science are conducted deliberately to test a particular hypothesis – and are therefore in keeping with what is called the hypothetico-deductive method, rather than a Baconian inductive method. Certainly, when we come to look at the experiments conducted by William Gilbert (shortly), Galileo (Chapter 9), and William Harvey (Chapter 10), we will see that they follow the hypothetico-deductive method. But Bacon repudiated this kind of experiment as potentially misleading. It is always possible, he believed, to devise an experiment to prove what you want it to prove. It is possible that what he had in mind here were the experiments which fraudulent alchemists performed to produce gold, or rather something that looked like it.

For Bacon an experiment should simply be another way of gathering detailed knowledge of phenomena, in order to make one's tables of discoveries as comprehensive as possible. As Bacon wrote in *Novum Organum*:

> Experiments ... have one admirable property and condition: they never miss or fail. For since they are applied, not for the purposes of producing any particular effect, but only of discovering the natural cause of some effect, they answer the end equally well whichever way

> they turn out; for they settle the question ... The office of the senses shall be only to judge of the experiment and ... the experiment itself shall judge of the thing.

So, rather than designing an experiment, as Galileo did, to establish that different weights fall at the same speed, an experiment should be set up to measure the speeds at which a huge range of different weights, or perhaps different bodies, fall (in the case of different bodies, of course, specific gravities enter the picture and might confuse matters, but a Baconian should not, could not, decide this in advance and choose to use only lead weights, say).

It seems fair to say that no scientist has ever conducted experiments in the Baconian way: experiments are usually conducted in order to test a particular theory or hypothesis, but Bacon seems to want theory-free experiments, in just the same way that he wants theory-free observations (which many would say are also unrealistic – we bring pre-conceived expectations to all our observations). Nevertheless, Bacon was immensely influential as we will see later, although his ideas were used not so much to do good science as to argue abstractly for how science should be done. Be that as it may, his works were read all over Europe and recognized as powerful pleas in favour of the experimental method (even if the experimentalism that was taken up was not strictly Baconian).

Bacon and natural magic

Because Bacon was the first major philosopher to suggest that knowledge of the natural world should be put to use for the benefit of mankind, and because he was known to extol the advantages of the experimental method, it was all too easy to see him as 'ahead of his time', somehow pre-figuring today's alliance between science and technology. Indeed, one historian of science writing in the 1940s, Benjamin Farringdon, wrote a book on *Francis Bacon, Philosopher of Industrial Science*.

However, as I've already suggested, Bacon's experimentalism and his famous vision that 'knowledge is power' came not from an uncanny prescience about the future, but from the age-old natural magic tradition. There can be little doubt of the impact of magical notions on Bacon. This can be seen, for example, in a list of *Great works of nature for the particular use of mankind*, which Bacon drew up in 1624. Virtually all the desiderata are those things that magicians had always claimed they could do: the prolongation of life, the increasing of strength and activity, the altering of features, [con]versions of bodies into other bodies, force of the imagination (either upon another body or upon oneself), impressions of the air and raising of tempests, deceptions of the senses, and so the list goes on. If Farringdon had been right we might have expected to see a list of desiderata more like those dreamed up earlier by Leonardo: helicopters and other flying machines, carriages capable of moving without horses, tanks, submarines, and

Box 7.2 Bacon's logic

> The logic now in use [Aristotle's deductive logic] serves rather to fix and give stability to the errors which have their foundation in commonly received notions than to help the search after truth. So it does more harm than good. (*New Organon*, Part II, Aphorism 12).
>
> The syllogism consists of propositions, propositions consist of words, words are symbols of notions. Therefore if the notions themselves (which is the root of the matter) are confused and over-hastily abstracted from the facts, there can be no firmness in the superstructure. Our only hope therefore lies in a true induction. (*New Organon*, Part II, Aphorism 14).

Here is an example of the kind of wordy argument Bacon had in mind. The following is an adapted version of an argument in Aristotle's *Physics* but gives the flavour of Aristotle's way of arguing:

> Whether a vacuum is possible:
>
> A cube displaces its own bulk of water if immersed in it, so does it also if the medium be air, only that the displacement is not perceptible to the sense. So, whatever the medium may be it must yield to the intrusive body.
>
> Now, this yielding is impossible in vacuity, which is not a material entity at all, and one must suppose that the dimensionality already there in the place before it was occupied must interpenetrate the equal dimensionality of the intrusive cube when it enters; just as if the water or air should permeate it all through. This means that the cube and the vacuum will be in the same place at the same time. But two bodies cannot be in the same place at the same time.
>
> Therefore a vacuum is impossible.

Notice that Aristotle has lost sight of the fact, even though he says it himself, that a vacuum is 'not a material entity at all'; because he then goes on to object to the

➠

the like. Bacon did not write like a philosopher of industrial science, but as a natural magician, and the useful things he wanted his new philosophy to be able to accomplish were exactly the kind of things that magicians were supposed to be able to do (see Box 7.3).

It is worth noting also that the inductive method was also the logic that was in use in the magical tradition. The magi had always been primarily interested in whether a technique or procedure works, not in trying to explain it. So, for the natural magician a conclusion based on inductive experience fitted the bill.

Bacon was clearly inspired by the magical tradition, by its methods and its pragmatic aims. His ambitions to reform natural philosophy went hand-in-hand impossible situation of having two material entities occupying the same place at the same time, namely, a body and a vacuum (or a body and space). Bacon is right to say that Aristotle's argument 'serves rather to fix and give stability to' Aristotle's erroneous belief that a vacuum is impossible.

> There are and can only be two ways of searching into and discovering truth. The one [deduction] flies from the senses and particulars to the most general axioms, and from these principles, the truth of which it takes for settled and immoveable, proceeds to judgment and to the discovery of middle axioms. And this way is now in fashion. The other [induction] derives axioms from the senses and particulars, rising by a gradual and unbroken ascent, so that it arrives at the most general axioms last of all. This is the true way, but as yet untried. (*New Organon*, Part II, Aphorism 19).

A deductive syllogism looks like this, and the conclusion tells us nothing that isn't already implicit in the initial premises. For Aristotle, however, it provides a clear explanation as to why Socrates is mortal.

> All men are mortal
> Socrates is a man
>
> Therefore: Socrates is mortal

Induction proceeds like this, and provides no explanation, but does result in new (if necessarily tentative) knowledge. This example is taken from a logical textbook in use in Bacon's day, Thomas Wilson's *Rule of Reason* (1551).

> Rehenyshe wine heateth,
> Malvesey heateth,
> Frenchwine heateth,
> Neither is there any wyne that doth the contrary,
>
> Ergo all wine heateth.

with his belief that the age-old ambitions of the magus could also be achieved by bringing their methods to perfection. It was his attempts to provide a philosophical defence of the ways of the magus that had such a major impact on contemporary natural philosophers and helped to establish the Baconian method.

William Gilbert and the origins of the experimental method

In spite of the fact that Francis Bacon's call for the introduction of the experimental method, inductive logic, and the practical usefulness of knowledge about

Box 7.3 MAGICAL THINKING IN BACON

Here are some quotations from Bacon's works, illustrating his concern with magic and what it could do.

> Just as bodies lay themselves open all round to attractive and friendly things and go to meet them, so when they happen on things hateful and hostile they fly from them all round and pull back and withdraw into themselves. (*History of Density and Rarity*, 1624)
>
> We must... investigate the individual and particular friendships and quarrels or sympathies and antipathies of bodies with diligence and care, seeing that they bring with them such a number of useful things. (*New Abecedarium of Nature*, 1622)
>
> For although such things lie buried deep beneath a mass of falsehood and fable, yet they should be looked into... for it may be that in some of them some natural operation lies at the bottom; as in fascination, strengthening of the imagination, sympathy of things at a distance, transmission of impressions from spirit to spirit no less than from body to body, and the like. (*New Organon*, Part II, Aphorism 31, 1620)
>
> The aim of magic is to recall natural philosophy from the vanity of speculations to the importance of experiments. (*The Proficiency and Advancement of Learning*, Book I, 1623)
>
> The End of our Foundation is the knowledge of causes, and secret motions of things; and the enlarging of the bounds of Human Empire, to the effecting of all things possible. (*New Atlantis*, 1624)

the natural world, was inspired by, and heavily indebted to, the natural magic tradition, it might be suggested that this counted for very little because Bacon himself did no real science but was merely a propagandist for how science ought to be done. He could be regarded as a buffer, therefore, between magic and science. In other words, it could be argued that magic did not directly shape the development of science, but merely influenced Bacon, who then developed a philosophy of experimentalism and pragmatism which played a role in the history of science.

The negative force of this kind of opposition to the importance of the magical tradition can be easily dispelled by looking at one of the many natural philosophers who were inspired by the new-found intellectual respectability of magic, after the recovery of the Hermetic writings, and who introduced innovations in science as a result of their magical world-view. We are going to look at one of these now; one who has been chosen because he has long been recognized by historians of science as a 'founding father' of experimental science.

His name is William Gilbert (1544–1603), an Elizabethan physician, who is known as the first thinker to conduct a systematic examination of magnets and their properties, and who discovered, and experimentally established, that the Earth itself is a giant magnet. He was also, like Bacon, heavily indebted to magical practices and occult ways of thinking.

The year 1600 was an auspicious one for Gilbert; he became personal physician to Elizabeth I and President of the Royal College of Physicians in that year, but as if that was not enough, he also published the book which is his main claim to fame, the first detailed study of the phenomenon of magnetism, entitled *Concerning the Magnet* (*De magnete*).

The question immediately arises, why did Gilbert undertake his studies of magnetism and what made him embark on the first book-length study of magnetism?

Earlier generations of historians of science tended to see Gilbert as yet another precursor of modern 'industrial science'. According to this view, Gilbert's interest in magnetism stemmed from his knowledge of the magnetic compass, one of the major inventions of the Renaissance period, which made many of the great voyages of discovery possible. By Gilbert's day the compass was a crucially important feature of naval exploration and, by implication, colonization. England, after all, was beginning to thrive at that time as a dominant maritime power. Given that Gilbert not only discusses navigational matters but also other technical aspects of what might be seen as compass production, such as mining for lodestones (natural magnets) and for iron, it seemed clear to these historians that Gilbert, supposedly like Bacon, foresaw the importance of an alliance between natural knowledge, and engineering and other technical skills, to foreshadow our science–technology complex.

A classic study in this mould is Edgar Zilsel's article on 'The Origins of William Gilbert's Experimental Method'. He takes the line that Gilbert was inspired by patriotic Elizabethan imperialist and economic considerations to provide fellow scholars with the knowledge of magnets, and in particular the magnetic compass, which hitherto was only known by mariners, metal foundry-workers, ore smelters, and others who worked with metals, or lodestones.

For Zilsel, therefore, the origins of the experimental method were to be found not among natural magicians but among artisans and craftsmen who were working with their hands and used to manipulating materials (see Box 7.4).

Zilsel talks about miners, for example, as though they experiment, but he never gives an example where miners can be seen to be performing anything that might look like an experiment. The closest he comes to it is to suggest that miners had a hand in an experiment which was actually carried out by Gilbert. Marking the orientation of a large lodestone in its place in a seam, Gilbert then arranges for the lodestone to be dug out and brought to the surface. He then places the lodestone on a suitably buoyant wooden raft in a large tub of water and notes how the freely moving lodestone aligns itself. Admittedly Gilbert did require the help of miners for part of his experiment (as labourers, to dig out and lift the heavy

Box 7.4 SCHOLARS AND CRAFTSMEN

From the sixteenth century onwards, the Aristotelian natural philosophy, which dominated the curricula in university arts faculties all over Europe, came increasingly under attack. One focus of that attack was the contemplative nature of the Aristotelian philosophy (as it was taught), and the lack of any concern with *practical* knowledge. Some scholars sought to correct this by deliberately seeking out craft knowledge and reporting it to their fellow scholars. One of the major examples of this can be seen in the increasingly economically important area of mining and metallurgy. The first printed account of Renaissance mining techniques, including instructions on the extraction of metals from their ores, how to make cannons and even how to make gunpowder, was the *De la pirotechnia* (1540) of Vanuccio Biringuccio (1480–1539). Written in Italian by a mining engineer who rose to the rank of Director of the Papal arsenal in Rome, it was evidently intended as an instruction manual for artisans working with metals. This can be compared with the *De re metallica* (1555) of Georgius Agricola (1494–1555). Agricola was a humanist scholar who taught Greek at Leipzig University before turning to medicine. Practising in a mining area, and initially interested in the medicinal uses of minerals and metals, he soon developed a compendious knowledge of mining and metallurgy. The fact that the *De re metallica* was published in Latin shows that it was aimed at an audience of university-trained scholars, not at miners or foundry-workers. Furthermore, the book's numerous editions and wide dissemination throughout Europe show that Agricola did not misjudge the audience.

A similar interest in the smelting of ores and the recovery of metals can be seen in William Gilbert's *De magnete* of 1600. Although principally concerned to use the

lodestone for him), but it is clear they were merely acting under his instructions. There is nothing to lead us to believe that the miners had done this kind of thing before, in order to satisfy their own curiosity.

Another stratagem that Zilsel tries, in an attempt to establish that Gilbert copies the technique of experimenting from artisans and craftsmen, is to point out that Gilbert merely repeated experiments performed by two earlier writers on magnets, Pierre de Maricourt (the Latin version of his name is Petrus Peregrinus), who wrote an *Epistle on the Magnet* in 1269, and Robert Norman (*fl.* 1560–84), author of a book called *The Newe Attractive* (1581).

Zilsel presents Pierre as some kind of early artisan or craftsman, but in fact he was famous in the magical tradition and Pierre tells us in his *Epistle on the Magnet* that this is the first part of a bigger work on the magnet and the second part (which, if he ever did get around to writing, no longer exists) is to be concerned with the use of magnets in magical amulets. Undeniably, Pierre did develop all the experimental techniques used two centuries later by both Robert Norman and Gilbert, but, given the use of experiments in the magical tradition, it is hardly surprising to see a writer whose main interest was in magical amulets using experimental methods.

spontaneous movements of magnets to show how the Earth itself might move, Gilbert also took the opportunity to report on all the practical know-how associated with magnets. As well as the metallurgical aspects, therefore, he also wrote at length on the use of the magnet in navigation, with a great deal of extra information on navigation besides. In this he explicitly drew upon the work of Robert Norman (*fl.* 1590), whose accidental discovery of magnetic 'dip' provided a means of determining latitude even when the heavenly bodies were obscured by clouds or fog.

Although there undoubtedly was an increased awareness of craft know-how and a willingness among scholars to accept and exploit its practical usefulness, it is important to avoid over-stating the case. During the 1930s and 40s a number of social historians seemed to forget the role of the scholars in this and to suggest that modern science owed its origins to the working man. The historian Edgar Zilsel (1891–1944) even went so far as to argue that the experimental method had been developed by artisans. The knowledge of craftsmen and artisans was taken up by scholars during the Scientific Revolution but it was chiefly the scholars' idea to do so; it was not something that was imposed upon them by the craftsmen. The main concern of scholars was to discover new ways of establishing certain knowledge, to replace the newly realized inadequacies of Ancient authority.

FURTHER READING:

Pamela H. Smith, *The Body of the Artisan: Art and Experience in the Scientific Revolution* (Chicago: University of Chicago Press, 2004).

Edgar Zilsel, *The Social Origins of Modern Science* (Dordrecht: Kluwer Academic, 2000).

Robert Norman fits Zilsel's claims much better. He was a former mariner, who had turned to making compasses for a living. Norman's book *The Newe Attractive* certainly provides descriptions of experiments to demonstrate the points he wants to make, but what Zilsel fails to mention is that Norman actually says in his *Newe Attractive* that the reason he did experiments was because 'learned men' advised him to do so. So, Norman cannot be used as an example of an artisan routinely doing experiments. Presumably his learned advisers knew about experimentation from the magical tradition, and in the case of experimentation with regard to magnets, they probably knew it from Pierre's *Epistle*, which had been printed for the first time in 1558 – before that, of course, it was only known in manuscript).

In spite of its status as a classic paper in the history of science, there is no solid evidence in Zilsel's article on Gilbert's experimental method that that method was suggested to Gilbert by artisans or craftsmen. Certainly, such men worked with their hands, but they were not manipulating the materials they worked with in order to discover truths about the nature of the world, or in order to try to understand how the world system worked. The men who were pursuing this kind of enterprise were not artisans, but alchemists, and others in the natural magic tradition.

Gilbert himself was certainly driven to explain the workings of the world. Indeed, it is obvious from his book on the magnet that his main aim was to explain one particular pressing problem in natural philosophy. This problem was a pressing one for Gilbert because he happened to be one of the very few thinkers who believed in the physical truth of the Copernican theory. That is to say, he did not simply regard the Copernican theory as useful for calculating heavenly movements; he was one of the earliest natural philosophers to believe that the Earth really was in motion, spinning on its axis, and revolving around the Sun.

As such, he inherited a problem that Copernicus never solved (indeed, a problem which Copernicus never even discussed). The problem was this: *How* does the Earth move? What is it that keeps the Earth in perpetual motion?

Effectively, Copernicus left this problem to succeeding generations of natural philosophers. Indeed, he drew upon the age-old division between natural philosophers (concerned with cosmology) and mathematicians (concerned with astronomy), to give himself an excuse for not tackling this problem. Effectively, he told readers of his book that he was a mathematician and the mathematics of the world system told him that the Earth must be in motion, but he left it to natural philosophers to discover and explain how it was kept in motion.

According to Aristotelian physics the motion of the Earth was, of course, impossible. The element earth was defined as that element whose natural place was at the centre of the world system. This definition could then serve in Aristotelian syllogisms to explain why heavy (earthy) bodies fell towards the Earth. These little earthy bodies had a natural tendency to return to their natural place and so fell directly towards the centre of the Earth. Consequently, the very idea of displacing millions of tons of earth – the whole Earth itself – from the centre and expecting it to remain out there without immediately crashing to the centre was absurd.

It was also a golden rule of Aristotelian physics that anything in motion must be being moved by something else (there was no concept of inertia – the idea that once something is set in motion it will carry on moving until something stops it). So, if the Earth was moving, there would have to be something perpetually pushing it, or pulling it. The only other alternative was to suppose that the Earth was capable of moving itself, like an animal. In Aristotelian natural philosophy, self-movers, such as animals and humans, were still moved by something else, namely their souls. Indeed, animal motion was one proof of the existence of animal souls: animals can move without being pushed or pulled by something outside them, so they must have internal souls which move them. Copernicus actually said that the planets (now including the Earth) 'fly in the heavens like birds', but he did not take this any further.

It is not known how Gilbert hit upon the solution to this problem, but it seems clear that he realized that magnetism could provide the answer and undertook to set this down. The result is his remarkable book: *A New Philosophy concerning the magnet, and magnetic bodies, and the great earthly magnet* (*De magnete, magneticisque corporibus, et de magno magnete tellure, physiologia nova*).

The inspiration behind Gilbert's argument is pretty clear: lodestones and magnetized iron unfailingly orientate themselves in a particular direction by performing a spontaneous rotatory movement. The seeming spontaneity of this (nobody has to push a magnetic needle to make it spin towards the north) indicates that magnets are self-movers. Accordingly, if the Earth itself was a giant lodestone, presumably it would also make the required Copernican rotatory movements in the same spontaneous way.

Gilbert's *De magnete* is dedicated to testing this hypothesis, or rather, it is dedicated to showing that the Earth *is* a giant magnet, and therefore can, and does, move as Copernican theory demands. (This, by the way, is why Bacon spurned Gilbert's efforts. He saw Gilbert as jumping to hasty conclusions, and instead of devising experiments simply to help record facts about the behaviour of magnets – although there were some experiments of that kind – he carefully designed experiments to prove the truth of his hypothesis.)

So, how did Gilbert prove his Copernican hypothesis? He could not argue from authority, especially not Aristotelian authority, because his hypothesis about a self-moving magnetic Earth was an entirely new idea. And, obviously he could not demonstrate the motion of the Earth by any direct means. What he could do, however, was argue by analogy. In this case, by analogy with the behaviour of small lodestones that can be manipulated in experimental demonstrations. The problem here was that very few people at this time were familiar with the behaviour of magnets. Always regarded as occult objects, they were rarely seen until they began to be used in mariners' compasses, and even then, only a few on board ship would get to see them.

Gilbert's *De magnete*, therefore, is not just concerned with the task in hand (proving the motion of the Earth by analogy with magnets), but takes it upon itself to familiarize readers generally with the nature and behaviour of magnets. What's more, Gilbert describes their behaviour in such a way that he invites his readers to test what he says for themselves. He tells them, in other words, how to repeat his experiments. Gilbert even puts asterisks in the margin of his book at the points where experiments are described, to help readers quickly locate them. Furthermore, he varies the size of these asterisks depending upon what he takes to be the importance of the experiment – large asterisks for important experiments, that should certainly be performed by committed readers, smaller asterisks for those experiments which are less significant.

It is perhaps worth noting, however, that Gilbert nowhere lays claim to having invented the experimental method. He boasts in his preface about the innovatory nature of what he calls his 'magnetic philosophy', but he does not boast of a new method. In describing his system of using asterisks to indicate the importance of experiments he nowhere offers an apology for introducing strange procedures. He clearly regards his experiments as an obvious way to proceed – presumably because he was so familiar with the literature of natural magic that he failed to note that his readers may not have been equally aware of this way of doing things.

Because Gilbert's ultimate aim is to show that the Earth itself is a giant magnet, he performs his experiments not with bar magnets, which might seem to be the most convenient shape to use, but with spherical lodestones, which Gilbert has had made by cutting them on a lathe. In a cunning move, he refers to these spherical magnets as 'terrellae', which is the Latin for 'little Earths'. So, each time he instructs his readers to perform a certain action with a *terrella*, he is telling them to perform that action with a tiny version of the Earth.

The *De magnete* makes it impossible for any reader to deny that a magnet has a spontaneous occult power of what Gilbert calls 'verticity', a turning and orientating power. Gilbert is then able to draw on every educated reader's knowledge of Aristotelian philosophy to point out that, because the magnet can move itself, it must have a soul. He even points out that the magnet has a superior soul to the human soul, because it cannot be deceived or misled:

> The human soul uses reason, [he wrote,] sees many things, investigates many more; but, however well equipped, it gets light and the beginnings of knowledge from the outer senses, as from beyond a barrier – hence the very many ignorances and foolishnesses whereby our judgements and our life-actions are confused, so that few or none do rightly order their acts. But the Earth's magnetic force and the formate soul or animate form of the globes, that are without senses, but without error and without the injuries of ills and diseases, exert an unending action, quick, definite, constant, directive, motive imperant, harmonious, through the whole mass of matter ... Yet these movement's in Nature's founts are not produced by thoughts or reasonings or conjectures, like human acts, which are contingent, imperfect, and indeterminate, but connate in them are reason, knowledge, science, judgement, whence proceed acts positive and definite from the very foundations and beginnings of the World: these, because of the weakness of our soul, we cannot comprehend. Wherefore, not without reason, Thales, as Aristotle reports in his book *De Anima* [*On the Soul*], declares the lodestone to be animate, a part of the animate mother Earth and her beloved offspring.

Gilbert is then able to clinch his argument by showing that the Earth itself is a huge version of the *terrellae*, the spherical lodestones, that he has been describing in his book.

He does this thanks to a newly discovered magnetic phenomenon called magnetic dip. This was discovered by the compass maker Robert Norman, and described in his *Newe Attractive*. But Norman could offer no explanation for it. Norman noticed that magnetic needles not only orientate themselves in a north–south direction, but they also point down as though pointing into the Earth. The magnetic needle is, of course, pointing directly to the magnetic pole, through the curvature of the Earth, and therefore points down below the horizon. Norman concluded from this that the orientating power of the magnet must reside in the magnet itself. It was previously assumed that the magnet pointed to the Pole Star, and that the power might reside in the Star, drawing the point of the magnet towards it. Norman, seeing that the needle pointed down into the Earth was able to discount the power of Polaris, and assumed only that the magnet had its own power.

Gilbert, by contrast, was able to reproduce the phenomenon of magnetic dip in miniature, using tiny pieces of wire suspended above the surface of a *terrella*. At the horizon of the terrella, the wire would remain horizontal, but as the wire was moved closer to one of the poles it would dip down, successively more steeply, pointing across the curvature of the *terrella* directly to the pole. At the pole, the wire would point vertically downwards, straight into the *terrella*. The reproduction of magnetic dip over the surface of a spherical lodestone, revealed to Gilbert that Norman's discovery proved the Earth itself was a spherical lodestone (see Figure 7.1).

But, if the Earth is a spherical lodestone, and lodestones have souls which enable them to move themselves, then the Earth could rotate on its axis once every 24 hours, and could revolve around the Sun once every year. It could even vary the orientation of its axis once every 26,000 years, to cause a precession of the equinoxes. An Earth that is proved to be animate can move itself, and it seems

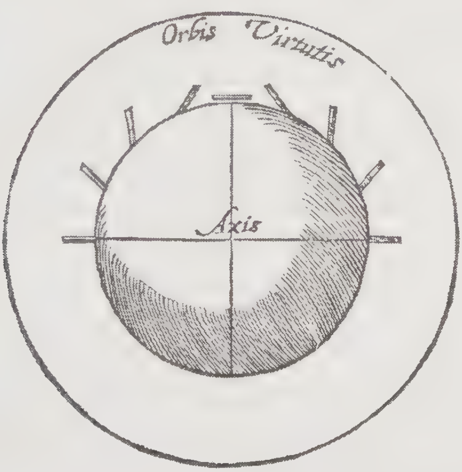

Figure 7.1 Explanation of the phenomenon of magnetic 'dip' according to William Gilbert, *De magnete* (London, 1600). Reproduced with permission from Edinburgh University Library, Special Collections (JY 192).

The north–south axis is shown horizontally here and the equatorial line vertically. A magnetized needle balanced on a point remains horizontal at the equator, but as it is moved successively north or south it points directly at the pole, located beneath the surface of the Earth, and so dips downward. At the poles, the needle points straight down.

perfectly reasonable to suppose that it must move itself in the ways that Copernicus has suggested.

Although Gilbert's theory fits in with the Aristotelian principle that an animate creature can move itself, it is by no means an Aristotelian argument. Gilbert's animistic approach to the world system is entirely magical, and Gilbert can be seen as using the experimental method of the natural magicians to establish a new claim in natural philosophy, namely that the Earth (contrary to Aristotle's teachings) can be, and is, in motion.

FURTHER READING

Stephen Gaukroger, *Francis Bacon and the Transformation of Early Modern Philosophy* (Cambridge: Cambridge University Press, 2001).

John Henry, 'Magic and Science in the Sixteenth and Seventeenth Centuries', in *A Companion to the History of Modern Science*, edited by G.N. Cantor, J.R.R. Christie, J. Hodge, and R.C. Olby (London and New York: Routledge, 1990), pp. 583–96.

John Henry, 'Animism and Empiricism: Copernican Physics and the Origins of William Gilbert's Experimental Method', *Journal of the History of Ideas*, 62 (2001), pp. 99–119.

John Henry, *Knowledge is Power: Francis Bacon and the Method of Science* (Icon Books, 2002).

John Henry, 'The Fragmentation of the Occult and the Decline of Magic', *History of Science*, 46 (2008), pp. 1–48.

Stephen Pumfrey, *Latitude and the Magnetic Earth* (Cambridge: Icon Books, 2002).

Paolo Rossi, *Francis Bacon: From Magic to Science* (London: Routledge & Kegan Paul, 1968).

Edgar Zilsel, 'The Origins of William Gilbert's Scientific Method', *Journal of the History of Ideas*, 2 (1941), pp. 1–32.

8

Bringing Mathematics and Natural Philosophy Together: Johannes Kepler

Gilbert's *De magnete* stimulated subsequent debate over the nature of magnets but it also served as a clarion call to those few thinkers who, like Gilbert, believed in the Copernican system as a new cosmology (and not just as a new hypothetical astronomy, useful simply for calculating heavenly movements). We saw earlier that Copernicus's *De revolutionibus* did not immediately create a major upheaval in astronomy and natural philosophy. The traditional separation between the pragmatic art of astronomy and the contemplative philosophical science of cosmology continued undisturbed, with the result that Copernicus was seen as making a contribution only to astronomy. By presenting a natural philosophical explanation of how the Earth could move, Gilbert showed what needed to be done if Copernicanism was ever to be accepted as an account of the world system which was physically true. This was an important lesson even for those who could not, or would not, accept Gilbert's animistic picture of the world.

Gilbert's ideas were most prominently taken up by the two thinkers who did more than anyone else in persuading contemporaries that the Copernican theory was true, the German mathematician and astronomer, Johannes Kepler (1571–1630), and the Italian mathematician and would-be natural philosopher, Galileo Galilei (1564–1642). Both Kepler and Galileo were to explicitly draw upon Gilbert's work on magnetism in their own published work, and in so doing, undoubtedly introduced Gilbert to many new readers.

Magic and mathematics: Johannes Kepler

Kepler's great achievement was to discover what are now known as the three laws of planetary motion: that planets move in ellipses, not perfect circles, that the line joining the Sun and a planet sweeps out equal areas in equal periods of time (a law which defined how a planet moved around its orbit – speeding up and slowing down as it went), and that the sidereal period of a planet (known with great accuracy since antiquity), squared, is directly proportional to its average distance from the Sun, cubed (basing all measurements on the so-called 'astronomical unit', that is the distance from the Earth to the Sun, this can be reduced to $T^2 = r^3$, where T is the sidereal period and r is the mean distance from the Sun). The most interesting aspect of the story, however, is how Kepler arrived at these laws, which enabled planetary positions to be calculated straightforwardly and correctly.

Kepler was one of the very first converts to Copernicanism, and really believed the Earth was in motion. He was impressed, as Copernicus was, by the fact that the heliocentric system allowed the order of the planets to be fixed, and their respective distances from the Sun to be clearly established by geometry. Remember that, in the Ptolemaic system this was impossible – the planets could be in any order, and were only given an order by convention. Copernicus offered a properly geometrically defined cosmos, and that was good enough for Kepler.

But Kepler then went on to raise other questions arising out of this feature of Copernicanism. From our perspective, Kepler was trying to answer questions which we would not regard as scientific questions at all. For us, they seem the wrong questions to ask, but for Kepler, these were the crucially important questions. The questions were these:

- Why are there only six planets? (Nobody at this time knew of Uranus and Neptune.)
- Why have they been placed at the particular distances they are from the Sun?

He tells us that he was led to ask these questions in the first place because he had already seen the wonderful resemblance, as he saw it, between the motionless things in the cosmos (namely, the Sun, the sphere of the fixed stars, and the space in between), and the Holy Trinity: Father, Son and Holy Spirit. What was the cosmic significance of six planets? Why not seven, or ten, or a hundred? And how did God reach a decision as to where to put them? It wouldn't have been necessary to ask this question if God had placed them at equal intervals outwards from the Sun, but their spacing differed markedly from this obvious aesthetic ideal.

Kepler first saw the answer to these questions in what he took to be a revelation from God, while he was teaching his students about astrological conjunctions (when two planets appear to coincide against the backdrop of the stars).

Noticing from his drawing that it looked as though the orbits of Jupiter and Saturn might be set at the distances they are by interposing a triangle between them (see Figure 8.1), he tried this for the other planets, using other geometrical figures.

Eventually, he realized he should be thinking in three dimensions rather than two (he was, after all, dealing with heavenly spheres, not circles), and tried separating the planetary spheres with the five regular polygons. Remarkably, this seemed to work. By placing a cube inside the sphere of Saturn, so that its corners touched the sphere, Kepler could then inscribe a sphere inside the cube, just touching each face of the cube, which was remarkably close, proportionally, to the sphere of Jupiter. A tetrahedron inside the sphere of Jupiter defined the size of the sphere of Mars; a dodecahedron inside the sphere of Mars defined the sphere of the Earth, and so on.

The beauty of this – and for Kepler it was to a large extent its geometrical beauty that convinced him of its truth – was that this nesting of solids in this

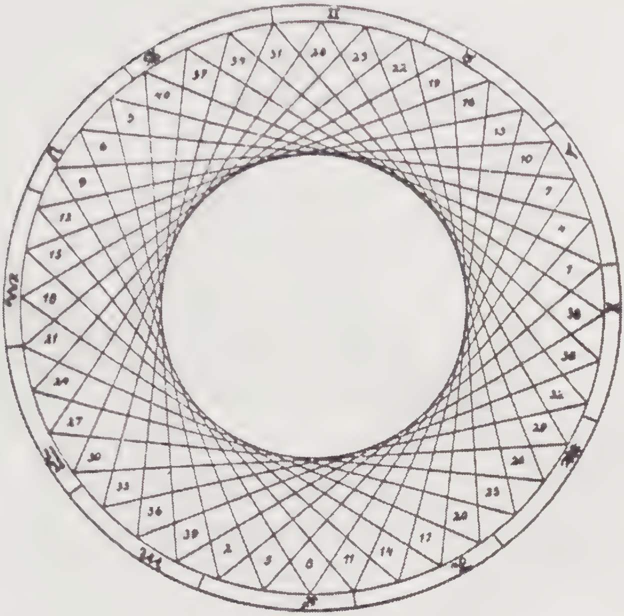

Figure 8.1 Kepler's depiction of how successive conjunctions of Jupiter and Saturn move around the sky; from his *Mysterium cosmographicum* (Tübingen, 1596). Reproduced with permission from Edinburgh University Library, Special Collections (JA 429).

Drawing a picture like this on the blackboard for his students, showing how successive conjunctions of Jupiter and Saturn move around the sky (position 1 is at 3 o'clock, position 2 at about 7 o'clock, 3 at about 10 o'clock, and 4 just above position 1, and so on around the circle), Kepler noticed that the proportions between the outer circle and the inner circle were close to the proportions between the spheres of Jupiter and Saturn. This was to lead him to God's 'geometrical archetype'.

way, not only explained why the planets were placed at the distances from one another that they were, but it also explained why God had to stop after he'd created six planets. The point is that there are only five regular polygons – that is to say only five polygons in which all the faces are the same (a cube has eight square faces; a tetrahedron, four equilateral triangular faces; a dodecahedron, twelve regular pentagonal faces; and so on). Not even God could create another regular polygon – the rules of geometry simply did not allow it. So, after God had placed the sphere of Mercury inside an octahedron, he had used all of the regular solids; he could not, therefore, define where to put another planet unless he repeated himself. So, he stopped.

It was no doubt an added bonus for Kepler that the five regular solids were also known as the Platonic solids, because Plato had also used them in his account of the cosmos in his *Timaeus*. Plato had suggested that the four

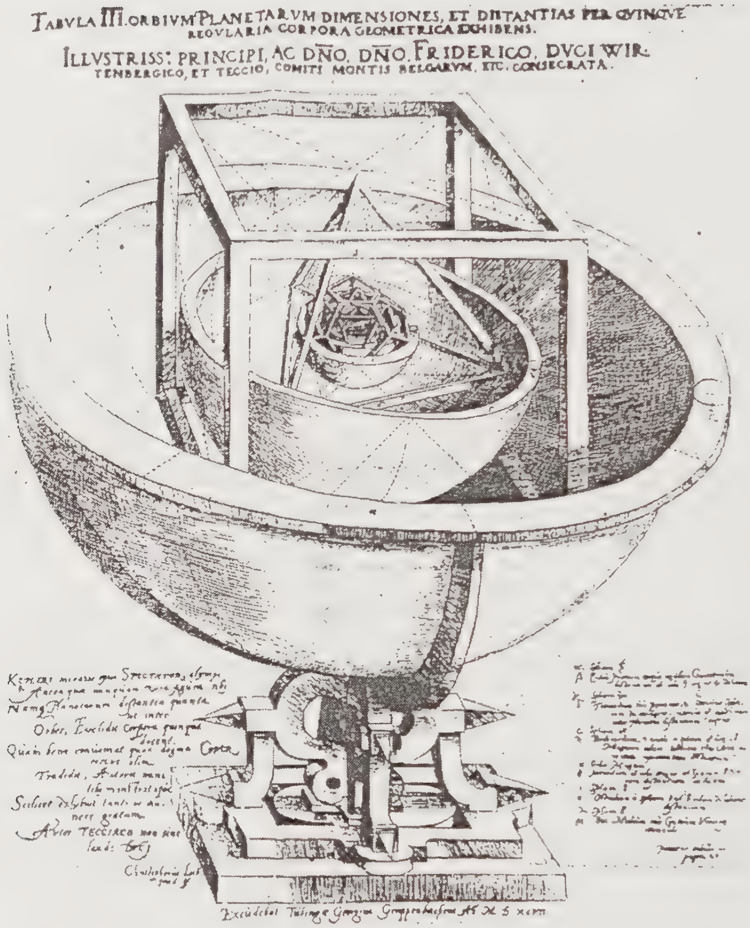

Figure 8.2 Depiction of a model of Kepler's geometrical archetype, from his *Mysterium cosmographicum* (Tübingen, 1596). Reproduced with permission from Edinburgh University Library, Special Collections (JA 429).

Illustration from Kepler's *Cosmographic Mystery*, published in 1596, in which he described what he believed to be God's geometrical archetype, or blueprint for making the universe. The picture shows how the planetary spheres are nested alternately with the five regular, or Platonic, solids, so providing the sizes of the spheres and limiting their number to six.

elements existed in a particulate form, each with characteristically shaped particles: cubes for earth, icosahedrons for water, octahedrons for air, and tetrahedrons for fire (dodecahedrons were said to represent the cosmos as whole) (see Chapter 2). Kepler certainly subscribed to a Platonic, or Neoplatonic, world-view and saw his geometrical archetype as an extra confirmation of it (Figure 8.2).

Kepler knew the distances between the planets according to Copernicus's calculations, and he could calculate what distances they ought to be if they are separated, as he suggested, by the nested polygons. Believe it or not, the two calculations were remarkably close. For Kepler, however, who believed he was 'thinking God's thoughts after him', close wasn't good enough. What Kepler needed were much more accurate observations – surely then his calculations would fit as precisely as God had made them.

Kepler knew exactly where he could lay his hands on accurate observational records. A Danish nobleman by the name of Tycho Brahe (1546–1601) had a reputation for unprecedented accuracy (using instruments of his own design), and he had recently taken up residence in Prague (in 1597) as astronomer to the court of the Holy Roman Emperor, Rudolf II (1552–1612).

Tycho's earliest claims to fame were his observations of what we now call a supernova, which appeared in 1572, and which for a time was so bright that it could be seen in daylight. Bearing in mind that, according to Aristotle, the heavens are perfect and unchanging, a new star suddenly appearing in the heavens ought not to happen. The general assumption, therefore, was that, in spite of superficial appearances, this new 'star' was in fact some kind of atmospheric phenomenon. Meteors and comets had long been dealt with in this way – both were meteorological, or atmospheric, phenomena, like rain and hail.

Tycho Brahe, however, calling on a network of professional astronomers arranged for the observation of this new light in the sky from different parts of Europe, and established that it had no parallax. This meant that it must be located, not in our atmosphere, but in the heavens.

A simple way to understand parallax is to consider what you see when you walk down a street when the moon is visible in the sky. If you look at a street lamp as you walk down the street, you will have to keep changing the direction of your gaze. At first, say, it is ahead of you, a little later on it appears alongside you, and later, if you wish to still see it you have to look backwards. Meanwhile, however, if the Moon was alongside when you started walking down the road, it will still be alongside you when you get to the bottom of the street. The Moon seems to follow you. The lamp-post has parallax, the Moon does not. In short, parallax is the apparent change of position of an object when it is observed from different places. If something is a long way away, a small change in the location of an observer will seem to the observer to make no difference to how he sees the distant object.

Tycho established that the new light in the sky had no parallax and must, therefore, be in the region of the heavens, and not below the Moon, as it would be if it was an atmospheric phenomenon. Inspired by this, he also decided to check the parallax of a comet which appeared in 1577. Sure enough, Brahe was able to establish that comets had no parallax and so must in fact be heavenly phenomena. What's more, he established that they followed a path which meant they must be cutting through the planetary spheres. This seemed to suggest that the spheres may not exist – certainly they did not exist as solid crystalline spheres, as some scholastic philosophers claimed.

Tycho might have used this as circumstantial evidence in favour of Copernicus, who had denied the solidity of the spheres on the grounds that the Earth is moving through space without being embedded in some kind of sphere. But Tycho was no Copernican. In fact, he was driven to ever more accurate observations of the heavenly bodies precisely because he felt that these would reveal the correct way to re-unite astronomy and cosmology, without introducing a patent absurdity such as the motion of the Earth.

He did eventually come up with his own system, known as the Tychonic system, in which all the planets went around the Sun, while the Sun went around the stationary Earth (this is thought to have been a half-way stage through which Copernicus went, on the way to his system). Tycho was not much of a mathematician, however, and he invited Kepler, known to be the most accomplished mathematical astronomer around, to work as his assistant, hoping that he would have the wherewithal to really make the Tychonic system work.

Although Tycho needed Kepler's mathematical expertise he was wary of him. Kepler, was not only a Copernican but one with whacky ideas like the geometrical archetype. Accordingly, Tycho allowed Kepler to see and use only his data on Mars, thereby setting Kepler the task of working out the orbit of the most difficult of the planets to solve (its proximity to the Earth made it particularly tricky). Even after Tycho's death, the year following the start of their collaboration, his heirs maintained this restricted access.

Kepler said that 'he could have died ten times' while trying to solve the orbit of Mars, but eventually he realized that the only orbit that fitted Tycho's data was an ellipse (he'd seen this earlier and then rejected it, but eventually had no choice but to come back to it). He published the results in his *New Astronomy* (*Astronomia nova*), of 1609, which included the first two of Kepler's three laws of planetary motion.

Now you might suppose that Kepler's discovery of elliptical orbits for the planets would force him to abandon his geometrical archetype. The assumption there, after all, was that the planets were moving on perfect spheres. But the geometrical archetype was too beautiful, and seemed so right, that Kepler would not give it up easily.

What Kepler did do was to ask himself why God would have chosen ellipsoids instead of spheres (spheres, after all, as the Ancient Greeks had seen, made the most sense in aesthetic terms), and why would he have chosen to make planets speed up and slow down, when he could have arranged for them to move uniformly? Kepler was still trying to 'think God's thoughts after him'.

Kepler's first attempt to deal with these questions appeared in the *New Astronomy* itself and it was here that he drew heavily on the work of William Gilbert. Kepler needed to avoid the same fate that had befallen Copernicus – he needed to ensure that his *New Astronomy* was not dismissed as nothing more than a set of calculating procedures which had no real relevance to physical reality. He took pains, therefore, to offer a *causal* explanation of how planets move – a physical explanation which simultaneously accounted for the elliptical orbits and

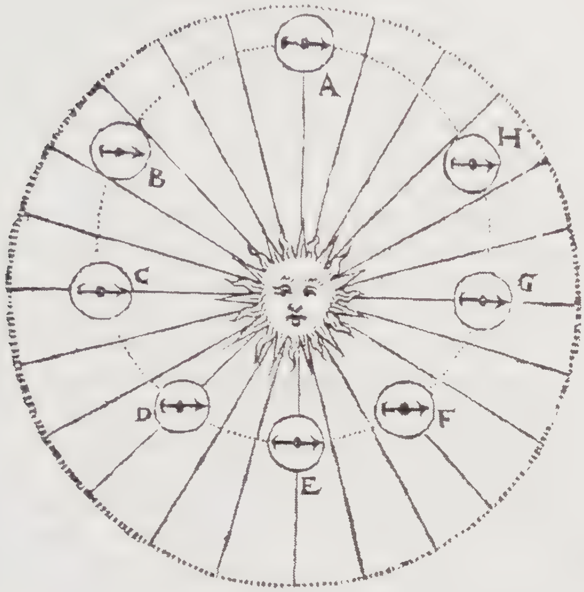

Figure 8.3 Kepler's explanation of elliptical orbits based on magnetic attraction and repulsion acting between the sun and the planets; from his *Astronomia nova* (Prague 1609). Reproduced with permission from Edinburgh University Library, Special Collections (JA 845).

The planet is shown in successive positions around the Sun with a superimposed compass needle showing the direction of the planet's magnetic field. Kepler assumed that the Sun acted like a magnet with one pole, which would sometimes attract the planet and sometimes repel it. This also accounted for the continuous acceleration and deceleration of the planet.

the variation in speed. Aristotelian natural philosophy dealt in causal explanations, not in mathematical technicalities, and so Kepler the would-be cosmologist had to follow suit.

Kepler even announced his intentions in the full title of the book: *A New Astronomy based upon Causes, or, Celestial Physics* (*Astronomia nova aitiologetos, seu physica coelestis*), and famously insisted that his intention was

> Such that I say that the celestial machine is not the equivalent of a divine animate being, but the equivalent of a clock ... so that nearly every variety of motion stems from a single corporeal magnetic force, just as all of the motions in a clock stem from a single weight.

Kepler's celestial physics would have been impossible without the work of William Gilbert, but he managed to turn Gilbert's animated world into an entirely physical (and for him, even mechanical) scenario.

Kepler assumed that all of the planets were global magnets, just like the Earth, and that the Sun was a unique kind of magnetic body with only one pole (say, just a north pole, without a corresponding south pole) as shown in Figure 8.3. These assumptions now allowed him to suggest that occult magnetic forces linked the Sun to each of the planets, and as the Sun rotated on its axis it swept the planets around with it. The further the planet was from the Sun, the slower it moved, because the magnetic force diminished over distance.

But the magnetic force between Sun and planet was also affected by the orientation of the planet. If the planet's south pole was closest to the single north pole of the Sun, the attraction would be strong, and a planet would accelerate towards the Sun, but if the planet moved into a position where its north pole was closer to the Sun magnetic repulsion would occur (like poles of magnets repel), slowing the planet's acceleration, and even causing the planet to move away from the Sun. Kepler's magnetic world system accounted for elliptical orbits and the continuous speeding up and slowing down of the planets.

But Kepler did not stop here. He still wanted to confirm that he had discovered God's geometrical archetype, the blueprint God had used when he created the world. To do so, however, he had to find a way back from ellipses to circles, or from ellipsoids to spheres. Already steeped in Neoplatonic ways of thought, Kepler turned to another famous aspect of the magical tradition to help him answer his new questions: why ellipses, and why a continual variation in speed of each planet? He turned to the Pythagorean idea of the 'music of the spheres'. Pythagoras believed that the celestial bodies, as they turned, made a heavenly music which we couldn't hear on Earth, but which would be audible to the Creator, God.

Kepler lived in the age of polyphonic music and he immediately realized that a planet moving constantly in a perfect circle could only generate a monotone, but a planet whose speed varied could generate a range of notes. Using the latest astronomical data he calculated the notes that each of the planets would generate (Figure 8.4).

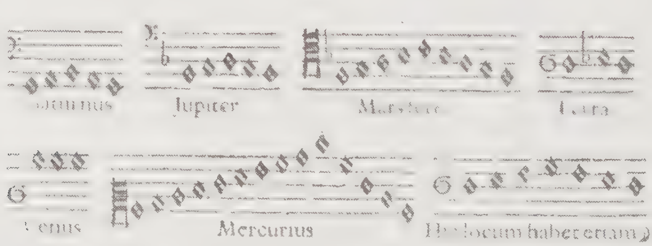

Figure 8.4 Kepler's depiction of the notes played by each planet as they move around the Sun, from his *Harmonices mundi* (Linz, 1619). Reproduced with permission from Edinburgh University Library, Special Collections (JY 112).

The notes made by the different planets as they speed up and slow down around their elliptical orbits. Venus generates a monotone because its orbit is circular.

Kepler went on to develop what he called the musical archetype. He considered many different ways in which adjacent planets would generate musical sounds, for example, comparing the sounds generated when one was moving at its maximum speed (at the perigee of its orbit – when it was closest to the Sun) and the other at its slowest (at its apogee – furthest from the Sun). Unbelievably, Kepler found astonishingly close correlations between musical ratios and the ratios between the speeds of the planets (see Table 8.1), and once again he believed he really had reconstructed God's musical archetype.

During the course of this work, Kepler found many hitherto unnoticed relationships between the planets, but one was particularly significant (and was dubbed Kepler's third law of planetary motion by later astronomers). The sidereal period of a planet (the period taken by the planet to move from a particular fixed star, taken as a marker, back to that same star), which had been known with great accuracy since antiquity, could be used to establish the planet's mean distance from the Sun, that is to say the average between its maximum (apogee) and minimum (perigee) distances. This allowed Kepler to return to perfect (and accurately calculated) circles, or spheres, and so he could finally confirm the accuracy of the geometrical archetype – and believe it or not, *he could confirm it* (see Table 8.2). The planets known to the Ancient world really do seem to be separated by just the distances required to fit invisible Platonic solids between them.

Kepler's attempts to confirm that he had thought God's thoughts after him seemed to him to have been undeniably vindicated. In the judgement of history, however, Kepler's real achievement was the discovery (entirely incidental for Kepler) of the three laws of planetary motion.

By the time Kepler came to write his final great work, the *Harmony of the World* (*Harmonices mundi*, 1619) he believed that few people would understand, or believe, his ideas. He hoped, however, that perhaps within a hundred years somebody would come along who would understand:

> Behold the die is cast, I am writing a book for my contemporaries or – it does not matter – for posterity. It may be that my book will await its readers for a hundred years. Has not God Himself waited six thousand years for someone to contemplate His work with understanding?

Kepler's ghost did not have to wait so long, in less than seventy years another mathematician would come along and turn Kepler's magnetic celestial physics into a physics based on the universal principle of gravitation. Interestingly, this mathematician, as we shall see, was also committed to Neoplatonic and magical ways of thinking, and was fascinated by the music of the spheres. His name was Isaac Newton.

Table 8.1 Kepler's musical archetype

Kepler used Tycho Brahe's data to establish the distance covered by each of the planets when at their aphelion (furthest from Sun, and moving at its slowest) and perihelion (nearest to the Sun, and moving at its fastest). He then compared the ratio between these two measured distances with the ratio between various musical consonances to see if there was any agreement. The result was a remarkably close correlation.

Planet	Ratio between seconds of arc covered per day at aphelion and perihelion	Closest musical consonance	Agreement of data and musical theory
Saturn	106 : 135 4 : 5 (=108 : 135)	Major third	Very good
Jupiter	200 : 330 5 : 6 (=275 : 330)	Minor third	Very good
Mars	1574 : 2281 2 : 3 (=1521 : 2281)	The fifth	Good
Earth	3423 : 3673 15 : 16 (=3448 : 3678)	Semitone	Excellent
Venus	5690 : 5857 24 : 25 (=5690 : 5927)	Diesis	Very good
Mercury	9840 : 23040 5 : 12 (=9840 : 23616)	Octave and a minor third	Good

He then compared the ratios between the distances covered by adjacent planets when one was at apehelion and the other at perihelion (and vice versa), with the ratios of the musical consonances. Again it all looked too close to be mere coincidence and so Kepler was convinced he had discovered that God must have used musical harmonies during the creation of the world system.

Adjacent planets	Ratio between seconds of arc covered per day at aphelion and perihelion	Closest musical consonance	Agreement of data and theory
Saturn (aphelion)/ Jupiter (perihelion)	106 : 330	Octave and a fifth 1 : 3 (=110 : 330)	Good
Saturn (perihelion)/ Jupiter (aphelion)	133 : 270	Octave 1: 2 (=135 : 270)	Perfect agreement
Jupiter (a)/Mars (p)	270 : 2281 1 : 8 (285 : 2281)	Triple octave	OK
Jupiter (p)/Mars (a)	330 : 1574 5 : 24 (=345: 1574)	Double octave and minor third	OK
Mars (a)/Earth (p)	1574 : 3678 5 : 12 (=1593 : 3678)	Octave and minor third	Good
Mars (p)/Earth (a)	3678 : 3423 2 : 3 (=3679: 3423)	Fifth	Perfect agreement
Earth (a)/Venus (p)	3423 : 5857 3 : 5 (3513 : 5857)	Major sixth	Good
Earth (p)/Venus (a)	3637 : 5690 5 : 8 (=3780: 5690)	Minor sixth	Good
Venus (a)/Mercury (p)	5690 : 23080 1 : 4 (5640: 23080	Double octave	Very good
Venus (p)/Mercury (a)	5857 : 9841 3 : 5 (=5844: 9841)	Major sixth	Very good

Note: These tables have been simplified from the tables presented in J. Bruce Brackenridge, 'Kepler, Elliptical Orbits and Celestial Circularity', *Annals of Science*, 39 (1982).

Table 8.2 The agreement of Kepler's geometrical archetype to reality (after applying Tycho Brahe's data, and Kepler's third law of planetary motion)

Adjacent planets Platonic solid	Separated by between size of orbits	Theoretical ratio between size of orbits	Observed ratio
Mercury Venus	Octahedron	86 : 122	88 : 122
Venus Earth	Icosahedron	122 : 153	121 : 153
Earth Mars	Dodecahedron	153 : 192	145 : 192
Mars Jupiter	Tetrahedron	192 : 577	192 : 577
Jupiter Saturn	Cube	577 : 1000	635 : 1000

Note Kepler would have wanted you to notice that the least accurate here (Earth/Mars and Jupiter/Saturn) were the most accurate in the musical archetype. Indeed, he even noted that neither the geometrical archetype nor the musical archetype were exactly right, because God wanted to use both and so had to compromise slightly in each case. At least, that's what he seems to be implying in *Harmonices mundi*, when he writes: 'From this it appears that the exact proportions of the distances of the planets from the Sun were not taken from the five regular solid figures alone: for the Creator does not depart from his Archetype, the Creator being the true source of Geometry and, as Plato wrote, always engaged in the practice of Geometry'.

FURTHER READING

Peter Barker and Bernard R. Goldstein, 'Theological Foundations of Kepler's Astronomy', *Osiris*, 16 (2001), pp. 88–113.

J. Bruce Brackenridge, 'Kepler, Elliptical Orbits, and Celestial Circularity: A Study in the Persistence of Metaphysical Commitment', *Annals of Science*, 39 (1982), pp. 117–43 and 265–95.

Alexandre Koyré, *The Astronomical Revolution: Copernicus–Kepler–Borelli* (Ithaca: Cornell University Press, 1973).

John North, *The Fontana History of Astronomy and Cosmology* (London: Fontana Press, 1994).

Victor E. Thoren, *Tycho Brahe: The Lord of Uraniborg* (Cambridge: Cambridge University Press, 1990).

James R. Voelkel, *Johannes Kepler and the New Astronomy* (Oxford: Oxford University Press, 1999).

9

Mathematics and Mechanics: Galileo Galilei

Like Kepler, Galileo can be seen (has been seen) as a Platonic or Pythagorean thinker, emphasizing the mathematical approach to understanding the world. As he wrote in his *Dialogue on the Two Chief World Systems* (1632); 'I know perfectly well that Pythagoreans had the highest esteem for the science of number and that Plato himself admired the human intellect and believed that it participates in divinity solely because it is able to understand the nature of numbers. And I myself am well inclined to make the same judgement'.

But it is important to note the very considerable differences between Galileo and Kepler. Kepler regarded geometry itself and the geometrical patterns of the cosmos as clues to the workings of the divine mind – a means of interpreting the signs which God gave to man to help him in his sacred duty of veneration towards the Creator. As such, he believed that this kind of geometrical analysis was best exemplified by, and so best employed in trying to understand, the heavens (astronomy) and light (optics), both of which phenomena seemed to Kepler to be closest to God. Kepler's attitude to mathematics was in keeping with Neoplatonic and associated views in which mathematics was affiliated to religious and magical visions of the harmony of creation.

For Galileo, by contrast, mathematical or geometrical analysis was applicable to all aspects of the natural world, including phenomena on Earth. Furthermore it was useful not so much for understanding God's mind, as for understanding the world itself. Galileo regarded the world as a machine whose operations could be analysed in mathematical terms.

> Philosophy is written in this grand book the universe, [he wrote in his *Assayer*] which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and to read the alphabet in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures, without which it is humanly impossible to understand a single word of it; without these one wanders about as in a dark labyrinth.

To us, Galileo could be said to be much more recognizably 'modern' in his thinking than Kepler, but Galileo saw himself as reviving the approach of the Ancient mathematician Archimedes (287–212 BC).

The works of Archimedes had only recently been recovered by Renaissance scholars and their importance was just being recognized by leading mathematicians. Previous assumptions, encouraged by Aristotle and scholastic philosophers, held that mathematics was only relevant to our understanding of very specific aspects of the natural world, such as astronomy, and the behaviour of light rays (optics), both of which could be reduced to exercises in geometry. Otherwise, mathematics was just too abstract to have any relevance to the physical world. It came as a rude awakening, therefore, to see Archimedes deliberately applying a process of extreme abstraction to the physical world (almost seeming to diminish its reality) in order to make it amenable to mathematical analysis. Archimedes' stratagem was only successful, however, in dealing with static systems (levers, pulleys, balances, etc.), and hydrostatic systems (bodies floating in water). Galileo's ambition was to extend the Archimedean approach to deal with moving systems, and to develop a new science of motion, what we would call a new science of 'kinematics'. This enterprise went hand in hand with Galileo's other, over-riding, ambition to reject and depose Aristotelian physics.

The difference between Kepler's and Galileo's attitudes to mathematics went hand-in-hand with other differences in attitude. Kepler had no qualms about occult traditions; he was, for example, a committed astrologer who tried to reform astrology to make it more accurate. Although he did not accept Gilbert's animism, he was happy to develop a theory of the cosmos involving the occult force of magnetism operating across vast distances of empty space.

Galileo, by contrast, rejected all notions of 'action at a distance' and, as we shall see, avoided assumptions which relied upon the supposed existence of invisible and intangible forces. Although he admired Gilbert, and discussed his *De magnete* in his own most famous work, the *Dialogue on the Two Chief World Systems* (1632), it was Gilbert's experimental method that inspired him, not his speculations about the nature of magnets, or the Earth.

There was, however, one major and important similarity between Kepler and Galileo. They were very much alike in their desire to bring mathematics and natural philosophy together. We have already seen how Kepler blurred the distinction between astronomy and natural philosophy, or physics, by providing a physical account of how the planets moved in his *New Astronomy*. He also declared that he was embarking on this kind of amalgamation of two distinct disciplinary traditions when he declared in the title that the *New Astronomy* was also a *Celestial Physics*. Galileo, in spite of being a mathematician, worked on the theory of motion, a topic normally covered in natural philosophy. Furthermore, Galileo never really worked as a professional astronomer, and when he took up the cause of Copernicanism he was much more concerned with its cosmological and physical implications than with its astronomical details. Indeed, Galileo was so concerned to be known as a natural philosopher that when his fame allowed him to take up a position at Cosimo de Medici's Court in Florence, he explicitly asked to be designated as mathematician *and* natural philosopher.

In taking their respective stances, Kepler and Galileo were by no means unique. They were merely leading figures in a much wider movement towards what has been called the 'mathematization of the world-picture'. It used to be supposed, in older literature in the history of science, that natural philosophers of great genius recognized how important mathematics was and used it to help them understand the workings of the world. It is now realized, however, that it was largely mathematicians, men who were relegated to a lower rung of the intellectual ladder, who began to impose their way of seeing the physical world on natural philosophers. In so doing, of course, they were making a play for greater recognition and enhanced status, in intellectual life, at court, and in the wider society. This was indeed a major aspect of the Scientific Revolution and it is important to note that Kepler and Galileo are used here to represent this much wider movement, encompassing mathematical practitioners of many different kinds, and at many different levels of late Renaissance and early modern society.

Galileo's earliest work, in accordance with his aims to extend the Archimedean approach, was concerned with motion and associated phenomena. One of the most successful aspects of this work was the confirmation of what had long been suspected (even by some Aristotelians), namely, that all bodies fall at the same rate, thus rejecting the Aristotelian assumption that the heavier a body is, the faster it falls. Using carefully prepared inclined planes with smooth grooves, and highly polished metal balls which fitted the grooves, Galileo was eventually able to establish that the distances covered during descent increased in accordance with the odd numbers. If 1 represents the distance fallen in the first interval of time, in the second interval the body will fall a distance of 3 units, and in the third interval it will fall 5. What this means is that after one time interval of fall the distance covered is 1 unit, but after two intervals the total distance covered will be 4 (1+3), after the third interval a total distance of 9 (1+3+5), and so on. Accordingly, the increasing speed of descent is directly proportional to the square of the time elapsed, and to the square root of the distance fallen (Galileo hereby corrected his original assumption that the increase of speed would be proportional to the distance).

He also realized that the Aristotelian separation between natural and unnatural or violent motion was completely misleading. According to Aristotle, natural motions were only of three kinds: downwards to the centre of the Earth for earthy bodies; upwards for light bodies composed predominantly of air or fire; and uniformly in perfect circles for heavenly bodies. All other motions were unnatural. The standard view was that if a projectile was sent flying through the air the violence of the motion imposed on the projectile would prevent the projectile undergoing its natural motion. Only after the imposed motion came to an end would natural motion take over – and the projectile would fall to Earth. Galileo realized, however, that a projectile was subject to gravity as soon as it left the hand, the bow, the cannon's muzzle, or whatever. The resulting trajectory of a projectile, therefore, always derived from the nature and direction of the motion imposed by the projector, *and* the simultaneously acting natural motion affected by gravity. He even managed to establish that the resulting trajectory was always a parabola, or part of a parabola.

Although he taught mathematics at the Universities of Pisa, and subsequently Padua, Galileo was never an astronomer. But he seems to have welcomed the Copernican theory as another stick with which to beat Aristotle. Besides, his interest in developing a new theory of motion led him to consider, as William Gilbert and Kepler had, the problem of how the Earth and the other planets move. If he could explain what keeps the Earth in motion he would simultaneously demonstrate the inadequacy of Aristotelian theory.

But the first development in Galileo's Copernican campaign came not from work on the theory of motion but from discoveries he made with the newly invented telescope. Galileo didn't invent the telescope but having learned of it (and possibly having seen one) he worked out how to make his own. He experimented with lenses and he was able to increase the magnification from 8 times in August 1609 to 20 times in November, and finally to 30 times in January 1610.

Turning the telescope to the heavens, Galileo made genuinely surprising discoveries, and published them quickly in a short book, which he called *Siderius nuncius* (*Starry Messenger*, or *Message from the Stars*, 1610). This rapidly made a huge impact, and brought Galileo wide acclaim, or for some, notoriety.

Galileo's observations did not (could not) prove the truth of the Copernican theory, but they added hugely to the circumstantial evidence in its favour (see Box 9.1). They also drove more nails into the coffin of Aristotelianism. As a result, Galileo was inspired to press his advantages further, and conceived a plan to write a *System of the World*. He evidently had two main aims in this: first, to try to remove all the philosophical or physical objections to a moving Earth, and second, to present what he thought was his own proof that the Earth must be in motion.

Galileo's own proof for the Earth's motion seems to have come to him as a eureka moment, worthy of Archimedes himself, and it was to have the most profound effect on his way of thinking about all natural phenomena. Galileo believed he could explain the tides by recourse only to the movement of the Earth. A less confident thinker might have put this tidal theory forward as an alternative to other possible explanations of the tides. For Galileo, however, his new theory not only signified that he could explain the tides in terms of the Earth's motion, but also that motion alone was sufficient to explain all physical phenomena. Galileo's physics was almost entirely couched in kinematic terms – presupposed motions were used to explain all effects. Galileo excluded as much as he could any discussion of forces, and certainly tried to avoid explanation in terms of forces. His objection to force was not to notions like force of impact, or force of percussion, this kind of force derived its power from motion and was therefore in keeping with Galileo's philosophy. What he objected to were the prevalent occult notions of force, often said to emanate in some mysterious way from bodies (magnets were surrounded by an *orbis virtutis*, a sphere of virtue or force, the Moon similarly, had its sphere of activity, and so forth). Galileo clearly

Box 9.1 Galileo's telescopic observations

Siderius Nuncius (*Starry Messenger*), 1610:

Mountains, valleys, seas on the moon:

Therefore the Earth is not the only earthy body flying through space. Aristotle's dichotomy between terrestrial (everything made of the four elements) and celestial (everything made of the fifth element, or quintessence) cannot stand.

Moons of Jupiter:

Therefore the Earth is not the only planet to be the centre of its own separate system of rotation.

Large number of stars invisible to the naked eye:

Therefore many stars are further away than others. This throws doubt on the existence of the sphere of fixed stars, and so makes it less plausible to suggest that the stars all revolve around the Earth once every 24 hours. It also adds credence to the Copernican assumption that the fixed stars are so far away that they still do not show parallax, even though the Earth changes its position by many millions of miles as it revolves around the Sun. It may also suggest that space is infinite in extent and that the stars are infinite in number.

Letters on Sunspots, 1613:

Venus has phases like the Moon:

Therefore Venus revolves around the Sun (but this is also compatible with the Tychonic system, in which the planets revolve around the Sun as the Sun revolves around a stationary Earth).

The Sun has blemishes which 'come to be' and 'pass away':

Therefore the Sun is not perfect and unchanging. The movement of the sunspots show the Sun is rotating on its axis and so the Earth is not the only body in the solar system to rotate.

regarded such forces, including gravity, as occult entities and wanted to develop a physics that had no need of such inexplicable notions.

Galileo's tidal theory was based on an analogy with a barge bringing drinking water across the lagoon to the city of Venice. Imagine what happens when the barge hits a sand bank; the barge will stop, or dramatically slow down, but the fresh water inside it will carry on moving and will be seen to sink in level at the back of the barge, as it sloshes up at the front of the barge. If the Copernican theory is true, Galileo goes on, this sudden speeding up and slowing down of the Earth must be happening every twelve hours, and will therefore cause the seas to slosh around in the same way – this is the phenomenon we call the tides. But Galileo immediately goes beyond this to argue that, since there is no other possible explanation of the tides (he dismisses the occult influence of the Moon out of hand), the fact that the tides exist actually prove that the Earth really is in motion.

But, what did Galileo mean by saying that, according to Copernican theory, the Earth must be speeding up and slowing down every 12 hours? What he had in mind was the variation in speed caused by the alternative addition and subtraction of the Earth's revolutionary and rotatory motions. As the Earth moves around the Sun from east to west, the hemisphere of the Earth away from the Sun is also moving from east to west as the Earth rotates on its axis, but the hemisphere nearest to the Sun, is necessarily moving in the opposite direction. The speed of the hemisphere away from the Sun, therefore, is determined by adding the two westward motions together. The speed of the hemisphere nearest the Sun, however, is determined by subtracting the eastward rotation from the westward revolution (see Figure 9.1).

Galileo eventually published this attempted proof of the Earth's motion in a book known as the *Dialogue on the Two Chief World Systems* (the two in question being Ptolemy's and Copernicus's) in 1632. He had written a preliminary account of the tidal theory in 1616 but was delayed from publishing it as a result of hindrances from the Roman Catholic Church (about which more in the next section). But the delay enabled Galileo to develop ways of removing all explanations in terms of mysterious forces from his natural philosophy.

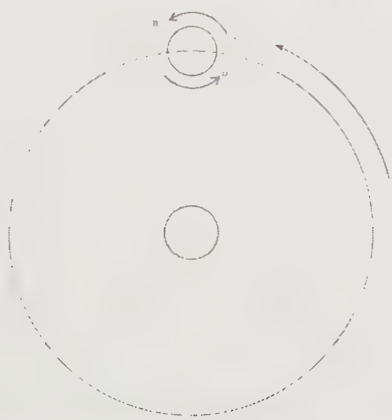

Figure 9.1 Galileo's attempt to explain tidal movements in terms of the movement of the Earth.

The explanation relied on changes in the relative motions of the surface of the Earth as it rotated on its axis and revolved around the Sun. The rotation of A to B is added to the general westward revolution of the Earth, but the rotation from C to D is subtracted, so that the two surfaces A–B and C–D are moving at different speeds – the change from one speed to the other causes a sloshing about of the Earth's seas.

Consider, for example, his ingenious explanation as to what keeps the Earth in perpetual motion. Remember, this is a major problem for Copernican theory. The argument is developed in the *Dialogue* between a character called Salviati, who plays the role of Galileo's mouthpiece, and Simplicio, the defender of Aristotelianism.

Salviati/Galileo asks Simplicio to imagine a situation where a bronze ball resting on a highly polished sloping plane is released. What will happen? Clearly, the ball will accelerate down the slope. What if it is pushed up the slope? It will gradually decelerate as it moves up the slope. But, what if the plane is perfectly horizontal? If the ball is set moving on the plane (and we imagine there is no friction to interfere), what will happen? The ball will not accelerate, nor will it decelerate, because there is no reason for it to do either one of these things. It will simply keep moving at the same speed.

So, now Galileo asks what is a perfectly horizontal plane on the Earth? The answer is a plane that remains at a constant distance from the Earth's centre of gravity. In fact, such a plane would actually be curved – following the curvature of the Earth. So, set a ball rolling on a circular path centred on the Earth's centre and the ball will continue to revolve around the Earth indefinitely.

I hope you can see that Galileo has presented a situation where there is no gravitational attraction to take into account. On a slope, gravity is a factor, causing deceleration or acceleration, but when there is no slope, when the surface upon which a body moves is always equidistant from the centre of rotation, motion is unaffected one way or the other by gravity, and will carry on, unchanging in speed, indefinitely (providing we disregard friction and other disturbances – remember, Galileo is thinking like an Archimedean abstractionist). Of course, we might object that Galileo has not really removed gravity entirely, because his explanation requires the horizontal surface to counteract gravity. But it is evident that Galileo believed no such physical surface was really necessary. In order to illustrate the point for Simplicio, he had had to talk in terms of a ball rolling on an actual surface, but the real conclusion to be drawn from this illustration, Galileo believed, was that a body moving uniformly in a perfect circle around the Earth would continue to move like this indefinitely.

The same argument is extended, by implication, to the Earth itself, rolling forever in a perfect circle around the Sun. This is certainly ingenious, but it raises major problems of interpretation for the historian. On the face of it, it seems impossible to believe Galileo could have been serious about it. The whole argument depends upon the claim that the planets must all be moving in perfect circles, with uniform circular motion, around the Sun. This is the Platonic dream, but any astronomer would have known that such a scheme simply doesn't fit the observations. This is why Ptolemy had to develop the whole paraphernalia of eccentrics, epicycles and equants. What's more, in 1632, Kepler's *Astronomia nova* had been available to astronomers for over twenty years. And he had shown there that planets move in ellipses, not perfect circles, and their speeds vary throughout their orbits.

It seems clear that Galileo must have believed that the deviations of the actual planets from the uniform circular motion, which he claims to have demonstrated in the *Dialogue*, derived merely from incidental factors caused by messy realities. Galileo's tidal theory demanded tidal shifts on a twelve-hour cycle, but the reality was much more complex and indeed scarcely fitted Galileo's theory at all. Accordingly, Galileo acknowledged that many extraneous factors, such as the depths of the different seas, or bodies of water, the natures of the sea floors, the orientation of the body of water in relation to the motion of the Earth, and perhaps even the surrounding topography, complicated the issue to such an extent that the underlying motive force of the tides (the motion of the Earth) was in danger of being completely obscured. It seems clear that Galileo must have thought in a similar way about the motion of the Earth and the other planets. A rocky terraqueous globe such as the Earth should not be expected to move as smoothly as a perfect uniform sphere might. Planets might, therefore, slow down and speed up, and might deviate from a perfectly circular path, but fundamentally they are kept in their perpetual cycles because of the principle that uniform circular motion around a centre of gravity will continue indefinitely.

Leaving that aside, we might also wish to challenge Galileo to tell us more about these centres of gravity and the mysterious power they have over things moving towards or away from them, but Galileo will have no such chat. Close reading of the *Dialogue* shows that anywhere that gravity or magnetism or some other mysterious force or power is mentioned, Galileo either dismisses it outright, or if he can't do that (as in the case of gravity) he simply curtails discussion by saying that this isn't the time or place to discuss these matters.

Galileo's determination to write forces, like gravity, out of the picture can even be seen in his discussion of free fall, and the acceleration of falling bodies. In his final work, *Discourses on Two New Sciences* (1638), he presented the matured versions of his earlier work on motion. With regard to acceleration in free fall he had this to say:

> When ... I observe a stone initially at rest falling from an elevated position and continually acquiring new increments of speed, why should I not believe that such increases take place in a manner which is exceedingly simple and rather obvious to everybody? If now we examine the matter carefully we find no addition or increment more simple than that which repeats itself always in the same manner... [T]hus we may picture to our mind a motion as uniformly and continuously accelerated when, during any equal intervals of time whatever, equal increments of speed are given to it... [H]ence the definition of motion which we are about to discuss may be stated as follows: A motion is said to be uniformly accelerated, when starting from rest, it acquires, during equal time-intervals, equal increments of speed.

In other words, assuming that Nature conforms to herself, and always operates the same way from moment to moment, we can be sure that what happens in free fall is the same from moment to moment. Notice that Galileo does not talk about acceleration due to gravity, he simply talks about increments of speed, which a body acquires. Instead of explaining *why* a body speeds up, he just describes in what way it speeds up: if it speeds up this much in one second, it will speed up by the same amount in the next second. Galileo looks as though he is providing us with a law of nature, regarding free fall, but he is explaining nothing. Furthermore, throughout the book, when talking about acceleration due to gravity he always refers to it simply as 'natural acceleration' – it just happens naturally; but if you want to know why, don't ask.

Galileo's natural philosophy is deeply problematic but that did not prevent it from becoming widely recognized, at least by some, as more useful than any philosophy that had gone before it. His experimental methods were taken up by his immediate followers working in Italy, and by others throughout Europe. But, perhaps more to the point, his rejection of anything that smacked of the occult, and the associated attempt to explain things entirely in terms of motion, and bodies in motion, was to have a major impact on the next generation of natural philosophers.

Galileo and the Church

Before moving on from Galileo, it is important to say something about his clash with the institution of the Roman Catholic Church. The so-called Galileo Affair, a *cause célèbre* if ever there was one, has come to stand as representative of what is generally taken to be an undeniable fact, namely, that science and religion cannot, and do not, get along. In fact, as we shall see, the Galileo Affair was the result of a unique set of circumstances and was very much a 'one-off' – it cannot, therefore, be made to stand as the exemplification of an inevitable clash between supposedly incompatible world-views.

Let us begin by reminding ourselves that the Copernican theory appeared in print in 1543. What was the response of the Catholic Church? They ignored it. Copernicus had dedicated his *De revolutionibus* to the Pope, and before that he had been invited to take part in the Church's proposed calendar reform. We also know that a Catholic theologian, Giovanni Maria Tolosani (1470–1549), wrote an objection to the Copernican theory in 1544, but concentrated not on the supposed Biblical arguments against the theory but the Aristotelian arguments. It seems clear from all this that what Copernicus had done was known at a high level within the Church, but still they chose to ignore it.

Martin Luther, initiator of the Protestant Reformation, is known to have objected to Copernicus as 'a fool who would turn the world upside down'. Perhaps the difference in attitude stemmed from the fact that Luther knew Copernicus was really serious about claiming the Earth was moving. Luther would have learned this from colleagues of Georg Rheticus, who published the first outline of Copernicus's ideas, the *Narratio prima*, in 1540, and who taught at the University of Wittenberg, where Luther was also based. Catholic leaders, by contrast, were more likely to have been fooled by Osiander's preface (see Chapter 6) into thinking that Copernicus wasn't seriously claiming the Earth moved (even though Copernicus's dedication to the Pope made it clear that he was).

So, the Copernican theory continued to be used by professional astronomers from 1543, but nobody else paid much attention to it. This state of affairs continued for about seventy years, until the Roman Church decided it needed to make an official pronouncement against the Copernican theory in 1616.

So, what had happened immediately before the ruling of 1616 which had suddenly forced the Church to pay attention to this not so new astronomical theory in a way that they hadn't before? The most succinct, but nonetheless correct, answer to this question is that Galileo happened.

Most historians are agreed that things might well have gone differently if Galileo had been more cautious, more tactful, and more willing to take things slowly. Indeed, there can be no doubt that a major factor in Galileo's downfall (and it was *Galileo's downfall* – but unfortunately he took the Copernican theory down with him) was his unfailing ability to make enemies. In particular, Galileo made enemies of an important group within the Dominican Order, a significant faction within the Jesuit Order, and finally, he made an enemy of the Pope himself. To make enemies of these people was unwise, to say the least.

When Pope Paul V called for an official ruling of the Church on Copernicanism in 1616, he did so because public disputes between a vociferous group of Dominicans on the one hand, and Galileo and his followers on the other, were becoming embarrassing for the Church. The ruling, decided by theologians, not natural philosophers, much less astronomers, almost inevitably went against Copernican theory, declaring it to be heretical with regard to religion, and foolish with regard to philosophy.

Galileo's friend Francesco Stelluti warned Galileo against making enemies of the Jesuits in a letter he wrote in 1620: 'It would be a business of which you would never hear the end if you picked a quarrel with those Fathers, for they are so many that they could take on the whole world, and, even if they are wrong, they would never concede it'. But Galileo continued to pour scorn on leading Jesuit astronomers at the Collegio Romano (the Jesuit university), especially in his *Assayer* of 1623.

Perhaps even more unfortunate, however, was Galileo's falling out with Pope Urban VIII, who had succeeded Paul V, and who, as Cardinal Maffeo Barberini, had been a good friend to, and genuine admirer of, Galileo. It was Urban VIII who, after much importuning by Galileo, finally gave him permission to write a book discussing the Copernican theory. Galileo wanted to write his *Theory of the Tides* (see above) and present it as a proof of the Earth's motion, but because of the 1616 ruling the Pope could not allow that. As an alternative, Urban generously suggested that Galileo should discuss his theory in the context of a discussion of the pros and cons of the Ptolemaic and the Copernican theories. What the Pope had in mind was a fairly neutral non-committal discussion, and he also asked Galileo to finish the book with a formulaic ending which was always recognized as conceding that the Church should have the last word. Essentially the formula was that man can never be sure how things are in nature because God, thanks to his omnipotence, can do things any way he chooses.

Galileo wrote his book, as we've already seen, as a dialogue (modelled on Plato's dialogues) but with the best will in the world towards Galileo, it is impossible to see it as a balanced and neutral account of the *Two* Chief World Systems (Ptolemaic and Copernican) – it is a glorious plea on behalf of Copernicus. Galileo did finish with the formulaic ending requested by the Pope. The trouble was, he used the character in the dialogue known as Simplicio to say these words. Simplicio's name was well chosen, and throughout the book he had been made to look like a simpleton. When the Pope saw Galileo's *Dialogue on the Two Chief World Systems* he is reputed to have said: 'He has put my words in the mouth of a fool'.

Whether this is true or not, certainly the Pope was extremely upset by Galileo and even after Galileo's death he refused repeated requests by various individuals and groups to be allowed to erect a statue in Galileo's honour. It was only after Urban VIII himself was dead that Galileo began to be honoured in this way in Catholic countries.

If anyone should still think that, even so, a clash between Galileo, as the representative of science, and the Church was inevitable, no matter how carefully Galileo had played his cards, they should consider the major trial documents and the main issue of the trial.

Galileo was not on trial simply for being a Copernican. The Copernican theory had been put on the Index of Prohibited Books in 1616 'until it should be corrected'. The list of corrections had been issued shortly after (in 1620), and Catholics could then own a copy of the book, provided they wrote the corrections by hand into their copies of the book. There are many such hand-corrected copies of *De revolutionibus* in libraries all over Europe.

A major issue at Galileo's trial was whether he had been instructed, after the ruling in 1616, never to defend the Copernican theory *in any way whatsoever*. If he had, then he had deliberately deceived the Pope in asking for permission to write the *Dialogue* – especially in the light of the pro-Copernican way he had written the *Dialogue*. If he had not been instructed this way, as Galileo claimed, then he hadn't done anything wrong. Now, it so happens that a crucial document in the Vatican file seems to suggest that he *was* instructed after the 1616 ruling never to discuss the theory in any way, but there is something dubious about this document.

It is clear from other documents in the file that Pope Paul V in 1616 gave detailed instructions to the inquisitors as to how to deal with Galileo. These instructions read like this:

> His Holiness has directed the Lord Cardinal Bellarmino to summon before him the said Galileo and admonish him to abandon the said opinion; and, *in case of his refusal to obey*, that the Commissary is to

> enjoin on him, before a notary and witnesses, a command to abstain altogether from teaching or defending this opinion and doctrine and even from discussing it; and, *if he do not acquiesce therein*, that he is to be imprisoned.

There is strong evidence to suggest that Galileo did immediately agree to abandon his opinion. The evidence is in the form of an affidavit given to Galileo at the time by Bellarmine. It reads like this:

> We, Roberto Cardinal Bellarmino, having heard that it is calumniously reported that Signor Galileo Galilei has in our hand abjured and has also been punished with salutary penance, and being requested to state the truth as to this, declare that the said Signor Galileo has not abjured, either in our hand, or the hand of any other person here in Rome, or anywhere else, so far as we know, any opinion or doctrine held by him; but that only the declaration made by the Holy Father and published by the Sacred Congregation of the Index has been notified to him, wherein it is set forth that the doctrine attributed to Copernicus, that the Earth moves around the Sun and that the Sun is stationary in the centre of the world and does not move east to west, is contrary to the Holy Scriptures and therefore cannot be defended or held.

The trouble is, the actual injunction against Galileo in 1616, as it is recorded in the file, matches neither the Pope's instructions, nor Bellarmine's letter. The relevant part reads like this:

> At the palace, the usual residence of the Lord Cardinal Bellarmino, the said Galileo, having been summoned and being present before the said Lord Cardinal, was, in presence of the Most Reverend Michelangelo Segizi of Lodi, of the Order of Preachers, Commissary-General of the Holy Office, by the said Cardinal warned of the error of the aforesaid opinion and admonished to abandon it; *and immediately thereafter*, before me and before witnesses, the Lord Cardinal being still present, the said Galileo was by the said Commissary commanded and enjoined, in the name of His Holiness the Pope and the whole Congregation of the Holy Office, to relinquish altogether the said opinion that the Sun is the centre of the world and immovable and that the Earth moves; *nor further to hold, teach, or defend it in any way whatsoever, verbally or in writing*; otherwise proceedings would be taken against him by the Holy Office; which injunction the said Galileo acquiesced in and promised to obey.

It looks as though Michelangelo Segizi, perhaps frustrated that Galileo was being let off too easily, ignored the Pope's instructions and took it upon himself to move immediately to the second stage of the Pope's instructions, even though Galileo had acquiesced to the first part. Presumably Bellarmine himself was also outraged by Segizi's behaviour and readily agreed to provide Galileo with the certificate which corrected the false impression given by the Injunction. Neither Bellarmine's nor Segizi's signature is on the Injunction, which was only witnessed by two servants in Bellarmine's household. It is possible that Bellarmine had no knowledge that Segizi was adding to his disobedience by putting this injunction in the file, but if he did know, he clearly refused to sign it.

At Galileo's trial the inquisitors tried to get to the bottom of this, but unfortunately, by 1633 Bellarmine, Pope Paul V, and even Segizi were all dead. Not all of the inquisitors at Galileo's trial in 1633 signed Galileo's condemnation – presumably they felt it was possible that Galileo never had been told not to '*defend it [Copernicanism] in any way whatsoever*'. In the end, however, the majority of the inquisitors decided that the Injunction of 1616 gave the correct account, and condemned Galileo for breaking his undertaking never to defend the Copernican theory, and for deceiving Pope Urban VIII when he asked for his permission, effectively, to break the terms of the Injunction.

In view of all this, it is wrong to see the Galileo Affair as proving that science and religion must inevitably clash, or are fundamentally inimical to one another. We can see from a letter written by Cardinal Bellarmine in 1615 that the Church might well have conceded that Biblical statements which seem to suggest the Earth is stationary are not meant to be taken literally. Bellarmine insisted, however, that it would be necessary to have an irrefutable proof that the Copernican theory was true before the Church would make such a concession:

> I say that if there were a true demonstration that the Sun is at the centre of the world and the Earth in the third heaven, and that the Sun does not circle the Earth but the Earth circles the Sun, then one would have to proceed with great care in explaining the Scriptures that appear contrary, and say rather that we do not understand them than what is demonstrated is false. But I will not believe that there is such a demonstration, until it is shown to me ... and in case of doubt one must not abandon the Holy Scripture as interpreted by the Holy Fathers.

Unfortunately, the Galileo Affair went out of hand before any reasonable decisions about the truth or otherwise of the Copernican theory could be reached. Perhaps Galileo believed he had a proof to convince Bellarmine, but if that proof depended upon his theory of the tides, he was simply wrong. If Galileo had negotiated more cautiously, and less arrogantly, he might have eventually prevailed, but his behaviour forced Paul V to call for a quick official ruling, and once that went against Copernicus, Galileo found himself fighting a losing battle.

It would be wrong, however, to give the impression that the blame was all Galileo's. One of the Commissioner's appointed to report on Galileo's *Dialogue* before his trial wrote that Galileo had 'wrongly attributed the existing ebb and flow of the sea to the non-existent immobility of the Sun and motion of the Earth'. There is no sign here, or in any of the commissioners or inquisitors, that Galileo's arguments were carefully considered. Instead of regarding the immobility of the Sun and motion of the Earth as the very points which needed to be decided upon, they simply declared them, right from the outset, non-existent. Galileo's arguments were never given a fair trial.

FURTHER READING

Mario Biagioli, 'The Social Status of Italian Mathematicians, 1450–1600', *History of Science*, 27 (1989), pp. 41–95.

Annibale Fantoli, *Galileo. For Copernicanism and for the Church*, translated by G. V. Coyne (Notre Dame: University of Notre Dame Press, 1994).

Maurice A. Finocchiaro, *The Galileo Affair: A Documentary History* (Berkeley: University of California Press, 1989).

John Henry, 'Galileo and the Scientific Revolution: The Importance of His Kinematics', *Galilaeana*, 8 (2011), pp. 3–36.

Ernan McMullin (ed.), *The Church and Galileo* (Notre Dame: University of Notre Dame Press, 2005).

Michael Sharratt, *Galileo: Decisive Innovator* (Oxford: Blackwell, 1994).

William R. Shea, *Galileo's Intellectual Revolution* (London: Macmillan, 1972).

David Wootton, *Galileo: Watcher of the Skies* (New Haven: Yale University Press, 2010).

10

Practice and Theory in Renaissance Medicine: William Harvey and the Circulation of the Blood

We saw earlier how Andreas Vesalius rejected the usual way of presenting anatomical studies to university medical students, where the focus was on following Galen's text (Galen being the Ancient authority in medicine, equivalent to Aristotle in natural philosophy), and chose to do dissections himself and to teach anatomy directly from the human cadaver. Vesalius was a pioneer but he was not a lone genius single-handedly reforming anatomy and physiology. Remember, we looked at him in Chapter 5 as an example of those Renaissance thinkers who rejected Ancient authority and began to look at the natural world for themselves. There were many other figures we might have singled out – it is well known, for example, that Leonardo da Vinci made many anatomical studies. Leonardo is not an important historical figure, however, in the sense that his studies and drawings were unknown to all but a few people, and so he contributed nothing to the accumulation of anatomical knowledge.

We are going to look now at one example of the successive accumulation of anatomical knowledge, in this case leading to the important discovery of the circulation of the blood. Before we start, however, it is important to note that this discovery, on the face of it, ought immediately to have signified the utter ruin of the Galenic system of medicine which had formed the basis of all medical theory and practice throughout the Middle Ages and the Renaissance. In order to understand this, it is important to see just how coherent the Galenic system was, and how interdependent were all its parts (see Box 10.1).

In the first edition of his *On the Structure of the Human Body* (*De humani corporis fabrica*, 1543), Vesalius noted that he had discovered over 200 errors in Galen's anatomy. He wasn't counting here the fact that he had been unable to detect the pores, in the interventricular septum of the heart (the thick wall of flesh dividing the right side of the heart from the left), which was an essential feature of Galen's physiological system. So strong was Galen's authority that even Vesalius preferred to follow Galen on this matter, rather than the evidence of his own eyes. He did, however, express his astonishment at how fine these pores must be:

> We are greatly forced to wonder at the skill of the Artificer of all things [he means God] by which the blood sweats through passages that are invisible to sight from the right ventricle to the left.

By the time of the second edition of his book, in 1555, however, Vesalius had become a bit more bold:

> I have never come upon even the most obscure passages by which the septum of the heart is traversed, albeit that these passages are recounted in detail by professors of anatomy seeing that they are utterly convinced that the blood is received into the left ventricle from the right. And so it is, that I am not a little in two minds about the office of the heart in this respect.

Even here, Vesalius, did not go very far. He was not inspired to pursue independent research on the function of the heart, nor did he draw any conclusions about the wider implications of his doubts. He was content merely to signal that there might be a problem.

One of Vesalius's successors as professor of surgery at Padua, Realdus Columbus (1516–59), did take up the challenge, and proposed the lesser, or pulmonary, circulation. He realized that the blood travelled from the right ventricle of the heart to the left, not through pores in the interventricular septum, but by crossing through the lungs.

There were three strands to Columbus's argument. First, the pulmonary artery (or arterial vein) which leaves the right ventricle, heading for the lungs, was too large if all it was required to do was to deliver nutrients to the lungs. Second, the pulmonary vein (venous artery) which comes into the left ventricle of the heart from the lungs is always full of blood, not air. Whereas, according to Galen this vein brought air from the lungs into the left ventricle to be concocted together with venous blood to form arterial blood. Third, the valves in the heart are fully competent and will not allow two-way flow in the pulmonary vein. Whereas, according to Galen, after concoction of arterial blood in the left ventricle, waste products were taken back to the lungs via the pulmonary vein (as fresh air was coming in) in order to be breathed out.

Columbus concluded that the blood entering the right ventricle of the heart must be conveyed to the lungs to pick up a vital spirit which is generated in the lungs. The blood, carrying the vital spirit, then returns to the left ventricle for distribution around the body. But he held little hope he would be believed: 'I fear that this new use of the lungs, which no anatomist dreamed of till now, must seem a paradox to the unbelievers and the Aristotelians'.

If this was half way to the discovery of the full circulation, another crucially important contribution to the story was made by a later professor of surgery at Padua, Hieronymus Fabricius (1537–1619), when he discovered valves in the veins. Thinking in essentially Galenist terms, however, Fabricius did not see that the valves ensured flow in only one direction. He simply thought that they slowed down the venous flow, to prevent blood rushing to the feet, say, before the knees could extract the natural spirit they needed from the passing blood.

Box 10.1 THE GALENIC SYSTEM OF ANATOMY, PHYSIOLOGY AND MEDICINE

The human body is constituted of **four humours** in different combinations for different parts – in exactly the same way that everything else in the physical world is constituted of *four elements*. The elements and the humours correspond to one another. The cold/dry principles are the element earth, and the humour melancholy; the cold/wet principles are water and phlegm; the hot/wet principles are air and blood; and the hot/dry principles are fire and bile.

All disease is seen in terms of an imbalance in the **four humours**, and the doctor's role is to restore the balance. Each of the humours is associated with a particular **organ**, and three of the major organs of the body are associated also with a particular subtle fluid principle, or **spirit**, which is then distributed through the body via one of the three *separate* vascular systems (veins, arteries, nerves).

The SPLEEN is the seat of MELANCHOLY or black bile (the earthy or cold/dry principle). An excess of melancholy can be countered by administering cathartics or laxatives.

The LIVER is the seat of CHOLER, or yellow bile (the fiery or hot/dry principle). An excess of bile can be countered by phlebotomy (cutting open a vein, since the liver is the source of venous blood and the venous system) or by administering emetics. The liver is responsible for converting chyle (produced in the stomach by the *concoction* of food) into venous blood, containing NATURAL SPIRIT. This spirit provides nutrition for the parts of the body and is distributed through the VEINS.

The HEART is the main seat of BLOOD (the airy or hot/wet principle). An excess of blood can be countered by phlebotomy (in this case cutting open an artery), or by applying leeches. The left ventricle of the Heart is responsible for *concocting* venous blood from the liver, containing natural spirit, into arterial blood, containing VITAL SPIRIT. This spirit was believed to provide the vital principle and the living heat to the parts of the body and was distributed via the ARTERIES.

More importantly, Fabricius was engaged in restoring an Aristotelian anatomical project at Padua, one which in classic Ancient Greek style was highly reductionist (see Chapter 2). While Galen and all other medical writers were concerned only with *human* anatomy, Aristotle, and now Fabricius, were interested in form and function in 'the animal'. Aristotle was not interested in the human liver, or the dog's liver, or a cow's, he was interested in 'the liver'. What is the liver for in the animal constitution; what do livers do in general? Fabricius, therefore, was led to a study of *comparative* anatomy, looking at particular organs or body parts across a wide range of animals. This kind of research was unique to Padua.

In 1602, a keen medical student at Cambridge University, by the name of William Harvey, decided to continue his studies at the very best medical school. Accordingly, he went to Padua, and found himself being taught by Fabricius.

The BRAIN is the main seat of PHLEGM (the watery or cold/wet principle in the body). An excess of phlegm can be expelled in various ways: by diuretics, sudorifics, expectorants (to expel urine, sweat, or phlegm, all of which are watery); or in the case of an accumulation of phlegm in a boil or other abscess by lancing and cauterizing. The Brain *concocts* arterial blood, containing vital spirit, into nervous fluid (the fluid supposedly contained in the nervous system), containing ANIMAL SPIRIT. This spirit was believed to provide the parts of the body with sensitivity and the ability to move, and was distributed via the NERVES (which were believed to constitute a continuous vascular system like the veins and the arteries).

The series of Galenic 'concoctions':

FOOD enters the **stomach** and is concocted into CHYLE.

CHYLE goes via the portal vein to the **liver** and is concocted into venous BLOOD.

Venous BLOOD, carrying *natural spirit*, is transported to all parts of the body to nourish the parts. During this distribution, some of the blood that arrived at the right ventricle of the heart seeps through pores in the interventricular septum (the wall between the right and left ventricles) into the left ventricle...

Venous BLOOD in the left ventricle of the **heart** is concocted into ARTERIAL BLOOD, containing *vital spirit*.

Arterial BLOOD and *vital spirit* are distributed to all parts via the arteries to vivify and heat the parts of the body. During this distribution some of the arterial blood arrives at the *rete mirabile* (a fine network of blood vessels) in the base of the brain (in fact there is no *rete mirabile* in humans – one of Galen's mistakes deriving from the fact he wasn't allowed to dissect humans).

BLOOD is concocted in the **brain** (specifically in the *rete mirabile*) into NERVOUS FLUID, containing *animal spirit*. *Animal spirit* is distributed to all parts of the body via the nerves to sensitize and to move the parts of the body.

Harvey has long been recognized as unusual among contributors to the Scientific Revolution in so far as he saw himself as a loyal follower of Aristotle and his methods. While men like Bacon, Galileo, and others were busy denouncing and rejecting Aristotle, Harvey declared him to be his supreme mentor.

Harvey's commitment to Aristotelianism undoubtedly stemmed from his work with Fabricius on his Aristotelian anatomical project. Padua University in general remained a major centre for Aristotelian natural philosophy (this is one reason why Galileo left) at a time when elsewhere Aristotelianism was falling out of favour. Even the Medical Faculty favoured Aristotelian anatomical and physiological theories over those of Galen. Harvey was readily swept up in the same enterprise, and when he returned to England in 1600 he started doing his own research on animals in the same comparative way. His main interest was on the generation – we would say, reproduction – of animals but this led him to consider the role of the heart and blood in animals.

It was the comparative aspects of Harvey's research which enabled him to perform not just dissections, as medical researchers interested only in humans must do, but also *vivisections*. He performed vivisections on fish, amphibia, snakes, and dogs. This meant that he was observing simpler, single-chambered, hearts, and had a better chance of seeing what was going on. He could also slow down the heart rate of cold-blooded animals by putting them on ice and watching their hearts beating. He could also watch the heart of a dog as it died; seeing it slow down and so, again, he was able to make better sense of what was going on as the heart beats.

He was also able to make clever use of ligatures on both arteries and veins, in different positions relative to the heart to observe the different effects. Noticing that ligatured veins became engorged with blood on the opposite side of the ligature from the heart and appeared pale and empty on the side nearest to the heart, while ligatured arteries became engorged on the side nearest to the heart and empty on the other side of the ligature, Harvey thereby established that blood in the veins was travelling towards the heart, while in the arteries it was flowing out from the heart. He was also able to see that the valves in the heart and in the veins only allow a flow of blood in one direction. In the case of the veins this meant that blood could only travel towards the heart, an impossibility in Galen's system, where venous blood had to travel out from the liver to the feet, say, or along the arms out to the fingers.

No physician pursuing research in the Galenic manner could have seen any of these things – he would have confined himself to dissecting dead humans, and it's very hard to understand how the heart operates if you are looking at a dead body!

Sometimes Harvey's comparative approach proved misleading, however. He also studied lungless animals, for example, and mammal embryos (in which the non-functioning lungs are by-passed) and concluded that in general lungs cannot play an essential role in the heart–blood system. He assumed that the lungs were merely for ventilating and cooling the blood and heart, to prevent over-heating. He saw the heart–blood system as essentially a closed system, a self-contained system keeping the body alive.

Admirers of Harvey's experimental techniques and careful observations are baffled by the fact that he never remarks upon the very clear differences in appearance between venous blood returning to the right ventricle of the heart, and the vivid oxygenated blood leaving the lungs. The fact is, Harvey only noticed things that fitted in with his own conception of animal anatomy and physiology (Francis Bacon had a point after all – we should try to gather facts without any preconceptions). Lungs couldn't be part of the heart-blood system, he was convinced, and so he failed to notice anything significant about the changing appearance of the blood as it crossed the lungs.

As if to compensate for Harvey's inadequacy in this regard, earlier commentators have emphasized what they have regarded as a *quantitative* element in Harvey's argument. Harvey estimated how much blood would be ejected from the left ventricle of the heart at each heartbeat, and how many beats there were in a minute, say, and concluded that the total amount expelled demanded that there must be a continual circulation of the same blood. In the Galenic system, blood ultimately derived from food being converted into chyle, and venous blood and then arterial blood, but we simply didn't eat enough to account for the amount of blood expelled by the left ventricle. It doesn't do to exaggerate, however, and it is somewhat misleading to present Harvey as some kind of mathematical biologist. A glance at his argument will show that he proceeds not by measuring how much is expelled from the heart as it contracts, but by suggesting to his readers 'as a reasonable guess that a fourth, fifth or sixth part, and at least an eighth, is sent into the artery', and that therefore every heartbeat expels 'an ounce or three drams or one dram'. The argument is all the more forceful because it is based not on precise measurements but on obviously under-estimating the amounts involved.

All in all, Harvey's *On the Motion of the Heart and Blood* provides what looks like an unassailable case in favour of the circulation of the blood. Harvey showed that venous blood and arterial blood were not, as Galen supposed, two separate systems, one emanating from the liver, the other from the left side of the heart. Furthermore, the two different kinds of blood did not 'ebb and flow', like the tides, in their separate systems, as they were gradually consumed in the different parts of the body (to be replenished by freshly made supplies in liver and left ventricle respectively). According to Harvey, blood, newly revivified in the left ventricle, was pumped out into the aorta and on through the rest of the arterial system to the extremities of the body. Eventually the exhausted blood passed from the arteries into the veins and returned to the right side of the heart. In mammals the blood was then passed across the lungs to enter into the left ventricle of the heart where it was revivified and sent out again.

In spite of Harvey's impeccable experimental techniques and his persuasive reasoning, however, he did not immediately convince his contemporaries.

The problem was that Harvey's new theory completely shattered the whole of the Galenic system but offered nothing to replace it. If Harvey was right then Galen was not just wrong about the heart but wrong about the liver, and the brain, and all the other interlocking parts of his physiological theory. If the liver did not send natural spirits around the body via the nerves, were there any natural spirits? And if not, were there vital and animal spirits? What was respiration for? But Harvey said nothing about the liver, or the brain, the spirits, the lungs, or anything other than the heart–blood system.

Furthermore, Galen's system was also closely tied to the theory of disease and the correct way to treat it – maintaining the balance of the four humours. Doctors had been making a very good living for centuries by following Galen's theories of disease and therapy – how could this be possible if Galen was wrong? Surely, it must be Harvey who is wrong.

The contemporary response was recorded by John Aubrey (1626–97), writing in the late 1660s:

> I have heard him say that after his Booke of the Circulation of the Blood came out, that he fell mightily in his Practize and that 'twas believed by the vulgar that he was crack-brained; and all the Physitians were against his Opinion... All his Profession would allowe him to be an excellent Anatomist, but I never heard of any that admired his therapeutique way.

Similarly, as Viscount Conway wrote to his daughter in 1651:

> I hear that you have a great opinion of Doctor Harvey. I thinke that you do well to love and respect a person of his merite for I thinke he hath deserved extreamly well of all learned men, for what he hath found out... but in the Practicke of Physicke [i.e. the practice of medicine] I conceive him to be too mutch, many times, governed by his Phantasy... to have a Physitian abound in phantasie is a very perilous thing; occasions in disease are very often suddaine, therefore one ought to have a Physitian that should be governed only by his judgement.

The excellence of Harvey's discovery was acknowledged and yet he was still dismissed as an unreliable physician. These two attitudes didn't match up and there was obvious tension. Underlying this tension was the realization that Galen offered a complete and coherent system which made sense of many aspects of the world we experience around us, but Harvey just offered a nifty account of the heart and blood.

Eventually, a group of followers began to conduct work that was supplementary to Harvey's, concerned with respiration, the stomach, the liver, and the brain. Galen's neat and interlocking system of anatomy and physiology was gradually supplanted, and yet, Galen's medical therapy, based on restoring and maintaining the balance of the four humours lingered on even into the nineteenth century. Replacing a system lock, stock, and barrel, is no easy feat.

FURTHER READING

Andrew Cunningham, 'Fabricius and the "Aristotle Project" in Anatomical Research at Padua', in A. Wear, R. K. French and I. M. Lonie (eds), *The Medical Renaissance of the Sixteenth Century* (Cambridge: Cambridge University Press, 1985), pp. 195–222.

Andrew Cunningham, 'William Harvey: The Discovery of the Circulation of the Blood', in Roy Porter (ed.), *Man Masters Nature* (London: BBC Books, 1987), pp. 65–76.

Andrew Cunningham, *The Anatomical Renaissance: The Resurrection of the Anatomical Projects of the Ancients* (Aldershot: Scolar Press, 1997).

Roger French, *William Harvey's Natural Philosophy* (Cambridge: Cambridge University Press, 1995).

Andrew Gregory, *Harvey's Heart: The Discovery of Blood Circulation* (Cambridge: Icon Books, 2001).

Lois N. Magner, *A History of the Life Sciences* (New York: M. Dekker, 1979), chapter 5, pp. 115–36.

Andrew Wear, 'The Heart and Blood from Vesalius to Harvey', in R.C. Olby *et al.* (eds), *Companion to the History of Modern Science* (London: Routledge, 1990), pp. 568–82.

11

The Spirit of System: René Descartes and the Mechanical Philosophy

The same kind of tension between accepting Harvey's restricted innovation in anatomy, and maintaining a complete and coherent system of medicine, even if it was incompatible with newly acquired knowledge, could also be seen in the case of early modern natural philosophy. Copernicus, Tycho Brahe, Gilbert, Kepler, Galileo, and many other less well known figures, had shown that there were very serious problems with the scholastic Aristotelian system, but nobody had offered a replacement system. It was all too piecemeal. Francis Bacon had recognized that what was required was an ambitious reform of the whole of natural philosophy, but he had never managed to bring anything to fruition. There were a series of Renaissance thinkers who took on the challenge of trying to offer complete replacement systems; men like Bernardino Telesio (1509–88), Francesco Patrizi (1529–97), Petrus Severinus (1540–1602), and a few others. But none of them succeeded in persuading contemporaries that they had all the answers.

This is where René Descartes (1596–1650) comes in. He developed a complete system of natural philosophy which was as comprehensive in its coverage as Aristotle's, and he was by far the most successful in persuading contemporaries that he did indeed have all the answers. Descartes is better known today as a philosopher rather than as a scientist, but he started out, like Kepler and Galileo, as a mathematician but, also like them, he moved across into natural philosophy. His new natural philosophy did not stop at a new theory of how the planets move, however, or even at a new theory of motion in general, but went on to cover all conceivable topics. As Descartes said at the end of his *Principles of Philosophy* (*Principia philosophiae*, 1644 – he meant *natural* philosophy), 'no phenomena of nature have been omitted from this treatise'. He did not mean this literally, of course. He meant that he had shown the principles by which all phenomena can, *in principle*, be explained. It was just a matter of applying his precepts to any phenomena you choose, and understanding will follow.

Needless to say, Descartes made many errors and ultimately failed, but he left a lasting mark on modern intellectual life, and it can be argued that our scientific world-view today owes more to him than to anyone. Our aim here, however, is to try to understand how he made such a big impact on his contemporaries.

We've already heard about what has been seen as a 'crisis of European thought' after the rediscovery of Ancient alternatives to Aristotle made it clear that he was not the oracle of truth, and various new discoveries in geography (that there are people living in the antipodes) and astronomy (changes in the heavens, mountains on the Moon, etc.), confirmed it. Among many of the learned all over Europe, confusion and depression set in. Who knows what to believe any more? Francis Bacon reminded his readers that Pontius Pilate had asked 'What is truth?', but, as Bacon went on, 'he would not stay for an answer'. The same was true for many in the early seventeenth century – they believed there was no point waiting for an answer because nobody could say what was true. Scepticism began to be a prominent way of thinking.

This tendency towards scepticism was exacerbated by the fact that Renaissance scholars had discovered that scepticism was a very prominent philosophy among the Ancients. Indeed, some Ancient sceptics went so far as to insist that certain knowledge of anything was impossible to achieve. Others took the less nihilistic line that with our current state of knowledge there is insufficient information or evidence upon which to base any firm conclusions and so we have to suspend belief until more evidence is forthcoming. The result is that, all over Europe, scholars, either through despair or in some cases anti-establishment mischievousness, turned to scepticism. The cry went up from Oxford to Padua: 'I only know one thing and that is that I know nothing'; and from Uppsala to Naples that 'Only one thing is certain, that nothing is certain'.

Now, it is easy to see that this kind of scepticism, arrived at as the result of a crisis among leading secular thinkers, is likely to cause a new crisis of its own. If you start applying sceptical strictures to all doctrines and ideas then Christianity itself is in danger. Indeed, as a matter of fact, the new religious pluralism resulting from the Lutheran Reformation, but rapidly followed by the establishment of rival Protestant Churches, such as Calvinism, had stimulated yet more scepticism – once there was only the Holy Mother Church of Catholicism, now there were a number of alternatives. Again, where did the truth lie?

It is no coincidence that this was just the time when atheism begins to emerge in European intellectual culture. There was no hint of atheism throughout the Middle Ages (heresy, yes, but not atheism), but in the late Renaissance it was attacked by orthodox thinkers for the first time (and indeed the word 'atheism' was coined for the first time). We do not see the atheists themselves in the historical record – they would have been highly covert and secretive – but we see orthodox attacks on atheism as though it was a genuine presence in contemporary life.

Among the faithful, therefore, scepticism became public enemy number one. But how can you fight it? How can you persuade someone to accept your arguments as valid and your conclusions as true when the whole point of their philosophy is that nothing is certain or reliable? Such a sceptical thinker would simply mock everything you say. The trick, therefore, is to find an argument so obviously true that to deny it would be foolish. If you could come up with such an argument, and get your sceptic to accept that he cannot deny it, then maybe you can go on to persuade him that other claims are equally undeniable.

When Descartes appeared on the scene, with his first publication, *A Discourse on the Method of Rightly Directing One's Reason and of Seeking Truth in the Sciences* (1637), his aim was to do just that. Descartes's most famous argument, perhaps the most famous philosophical argument of them all, 'I think, therefore I exist', was not offered as a proof that he, Descartes, existed. The point was that this was an argument that no sceptic could deny. Not even the most recalcitrant and nihilistic sceptic could declaim 'I think, but that's no guarantee that I exist'. As soon as you admit you are thinking, Descartes has got you.

The sceptical crisis, then, formed the background to this famous philosophical argument. 'I think, therefore I exist' was offered as an argument which no sceptic could reject. The sceptic now has to admit that something is certain and undeniable (namely, his or her own existence). But where do we go from here?

Descartes's next step was to point out that in our ideas we have an idea of perfection. But where did this idea come from? It can't be from ourselves because none of us is perfect. It can't be something we've abstracted from our experience, because we've never actually come across anything perfect in our experience – we've never seen a perfect thing, or tasted something perfect. The only explanation for our idea of perfection is that it must have been planted in our minds by a perfect being.

From here, Descartes makes a very controversial move. He insists that this perfect being must actually exist. He doesn't just base this conviction on the fact that we have an idea of perfection (which can only have come from this perfect being), but also on the nature of perfection itself. *If this being is perfect then it must exist*. Let's face it, a fast car that really exists is better than a supposedly even faster car that only exists in my imagination. If I offered you a real Porsche as a gift and said: 'Or you could have this imaginary car here which is superior to the Porsche in every way', and I was pointing to thin air as I said this; you would opt for the real Porsche, wouldn't you?

So, it makes no sense, Descartes insisted, to say this being is supremely perfect in every way, except for the trifling matter that he doesn't exist. If he doesn't exist he isn't perfect – he isn't superior to anything else you can think of – and therefore you weren't thinking properly about a perfect being. When you do think properly about a perfect being, Descartes wanted to claim, you'll see that that being must really exist.

This is known as the ontological argument for the existence of God and was first formulated in a slightly different way by St Anselm of Canterbury (d. 1109) in the eleventh century. It is usually held to be an invalid argument, although it is still being debated by some philosophical theologians. For our purposes, however, it is sufficient to note that Descartes was convinced he had here another argument which the sceptic could not refute (although he must have been disillusioned by the barrage of objections raised against this argument by his contemporaries – even very religious contemporaries).

Let us assume, nevertheless , that Descartes has proved the existence of God. It is now an easy matter for him to proceed. He argued that God would not allow us to be systematically deceived, and so, providing we use a suitably careful method for checking our facts, and establishing whether something is true or not, we can be sure that our conclusions are reliable.

Descartes set out the required method in his *Discourse on the Method*. It was based on the method in use in geometry with a few modifications, and boiled down to four general procedures:

> The first was never to accept anything as true if I had not evident knowledge of its being so; that is, carefully to avoid precipitancy and prejudice, and to embrace in my judgement only what presented itself to my mind so clearly and distinctly that I had no occasion to doubt it.
>
> The second, to divide each problem I examined into as many parts as was feasible, and as was requisite for its better solution.
>
> The third, to direct my thoughts in an orderly way; beginning with the simplest objects, those most apt to be known, and ascending little by little, in steps as it were, to the knowledge of the most complex; and establishing an order in thought even when the objects had no natural priority to one another.
>
> And the last, to make throughout such complete enumerations and such general surveys that I might be sure of leaving nothing out.

It should be noted that, like Francis Bacon, he decided to present a method of doing science before he presented any actual science. Unlike Bacon, however, Descartes already had his system of philosophy fully worked out. He had been about to publish it in 1633 when he heard about the condemnation of Galileo. Since Descartes's system was Copernican, and Descartes himself was a devout Roman Catholic, he did not want to go against the ruling of his Church and so he suppressed publication of his system and wrote the *Discourse on the Method* instead.

In a way, then, the *Discourse* was an announcement to the world that this author was not only able to defeat scepticism and describe the perfect method for arriving at the true philosophy, but that he already had developed this philosophy. He also included, therefore, a few examples of what his method could produce, showing how he could explain various carefully chosen natural phenomena. The exercise in self-publicity worked, and when he did publish his system, in 1644, it was rapidly taken up as the best hope of a fully comprehensive system to completely replace Aristotelianism, and Galenism, in their entirety.

Cartesianism and the mechanical philosophy

Descartes claimed that his system of philosophy was arrived at by using his method. Accordingly, it should all have conformed to the claim 'never to accept anything as true if I had not evident knowledge of its being so; that is, carefully to avoid precipitancy and prejudice, and to embrace in my judgement only what presented itself to my mind so clearly and distinctly that I had no occasion to doubt it'. The reality seems much more like an elaborate flight of fancy, although it has to be admitted that it is breathtakingly ingenious. Let us consider how Descartes built up his system.

Taking a reductionist line to make sense of the bewildering complexity of the world around us, Descartes asked himself what was the real essence of things in the physical world. All things have various qualities: size, shape, texture, weight, colour, temperature, and so on. But not all of these are part of the essence of a thing. The temperature of an object may change without changing the object (though there are exceptions), or the colour can change, but it remains essentially the same object. Some qualities can be reduced to other qualities and so cannot be fundamental. Texture, for example, can be reduced to fine detail in the shape of the surface of an object (a flat surface makes the object feel smooth, a finely corrugated surface makes it feel rough).

But, Descartes noted that you can never escape the fact that all objects have a shape. That is, they all exist in three dimensions. Even a lump of wax, or clay, which can be moulded into an indefinite number of different shapes can never be moulded in such a way that it has no shape at all – if that did happen the wax would have disappeared, or ceased to exist. So, *extension* must be part of the essence of physical reality.

Descartes identified material objects with extension in space, consequently, void space was held to be a contradiction in terms. Extended empty space was impossible for Descartes – an extended space must be a material body.

Descartes also noted that reality is not unchanging and so motion must also be part of the essence of physical reality. Consequently, everything can be explained in terms of matter and motion, and nothing else. Even qualitative changes can be explained in terms of matter in motion – Descartes's system was similar to atomism in this regard (atomism explained everything in terms of the arrangement and re-arrangement – which entailed motion – of atoms) but he was not an atomist because he believed that the invisibly small particles of matter which he took to constitute all things, were infinitely divisible.

Descartes believed he could therefore reject all occult influences and phenomena, because he could explain them entirely in terms of matter in motion. In this he is very reminiscent of Galileo who, as we have seen, also hoped to explain everything in terms of motion, with no talk of mysterious powers or forces.

Take the example of magnetism. Descartes insisted that a stream of invisibly small particles flows out from each pole of a magnet, and circulates around, to re-enter the magnet at the other pole. The magnet itself has invisibly small pores throughout its substance, through which these particles can flow. Iron has similar pores. It is the stream of particles that is responsible for magnetic phenomena – so a magnet does not operate by an occult 'action-at-a-distance', but by the contact action of the particles.

But how did he account for both attraction and repulsion? He assumed (remember his method, and his claim 'never to accept anything as true if I had not evident knowledge of its being so') that the particles exist in two forms – they are all shaped like little screws but some have a right-hand thread and others have a left-hand thread. Similarly, the pores through which they stream are threaded, with either a right or left twist. If a right-hand threaded particle meets right-hand threaded pores, they stream through, pulling by their screw-like action the iron or magnet towards the magnet from which they had emanated. If right-hand threaded particles meet left-handed pores, however, the result is that the stream of particles pushes the magnet away, which of course looks like repulsion.

Denial of void or vacuum enabled Descartes to develop what is called his 'vortex' theory – as a body moves forward it has to displace the medium through which it moves, let's say air, in front of it. But since there is no empty space for this displaced air to move into, a kind of chain reaction takes place, with air displacing air adjacent to it, and so on. But such displacements don't extend throughout the whole universe, because as the body moves forward it is in danger of leaving an empty space behind it – so this space is immediately filled by air moving into it. But the space vacated by this air also has to be filled, and so what happens is we get a circulation of the air around a moving body, moving from the front of the body around to the back. This circulation is a vortex.

The vortices themselves play an important role in Cartesian explanations. Take the solar system, for example. The whole system is a giant vortex, or whirlpool, of jostling invisibly small particles. As it rotates centrifugal forces occur (think of a sling with a stone whirled around – the stone wants to fly away and so the sling becomes taut), and streams of particles continually move from the centre out to the periphery. At the periphery of our world system, there are other worlds systems (see Figure 11.1), and the outward streams of particles from adjacent world systems meet and are rebounded back to where they came from. So there is also a continual downward stream of particles, rebounding back to the centre.

Descartes uses these two streams, outward and inward, to explain the stable orbits of the planets, and the downward stream to explain gravitational attraction – when a body falls it is because it is pushed down by these continually descending streams of particles.

So, in spite of Descartes insistence on the importance of 'systematic doubt' as a starting point for his method, he develops a system of ingenious, but unworkable, fantasy.

It is easy to see flaws in his system. If gravity is caused by a downward stream of particles pushing things down, why do things fall if they are released underneath a table, say? Shouldn't the table shield the object from the downward stream of particles? Descartes's reply is to insist that the particles are so small and fine that they can pass through the table. But this raises the question: if the particles pass through the table, why don't they pass through the object they are meant to be pushing down to the ground? Can they actually push on something if they are so fine that they can pass through a table?

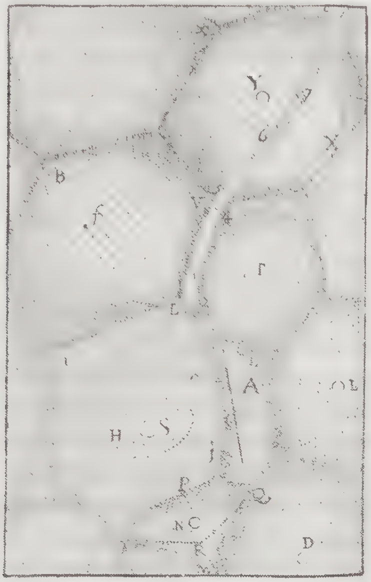

Figure 11.1 Depiction of part of the universe, showing various solar systems as vortexes of swirling matter; from René Descartes, *Principia philosophiae* (Amsterdam, 1644). Reproduced with permission from Edinburgh University Library, Special Collections (JA 398/1).

The Cartesian universe showing several vortexes whirling around different stars (D, S, F, etc.). Note how adjacent vortexes impinge upon one another and ultimately affect one another.

Consider also the fact that Descartes claims there can be no new motion in the world. God set the universe moving at the Creation, Descartes supposed, but after that motion has simply been transferred from one particle to another in collisions. A stationary particle set in motion by being hit by a moving particle, takes its motion from the moving particle, which therefore necessarily slows down, or perhaps even stops. Descartes drew up precise rules for the transfer of motion, but we need not go into such details to realize that there are severe problems with this part of Descartes's system. Let us give Descartes the benefit of the doubt and assume that the vortex theory of planetary motions, and of gravity, works because the vortex was set in motion at the Creation and has continued ever since. But what about the mini-vortexes around magnets? We have seen how a magnet is supposed to work as a result of streams of particles emerging out of one pole and swirling around to re-enter the magnet at the other pole. It seems reasonable to suppose that some of these magnets were set in motion at the beginning of the world by God. But what about the case of a piece of iron, rubbed by a magnet so that it becomes magnetized. Are we to suppose the streams of particles now swirling around the newly magnetized iron were set in motion by the rubbing motion of the original magnet? Can this really work? It is certainly impossible to see how it can follow from Descartes's rules of collision. The same is also true of the dramatic effect of gunpowder. Are we to suppose that the motion gunpowder can impart to a heavy cannonball was in turn transferred to the gunpowder by the flickering motion of the flame which ignited it? If not, then are we to suppose that the chemical process by which the gunpowder was made somehow locked motion into the gunpowder, only to be released on ignition? Can this really work? Again, it does not seem compatible with Descartes's rules of collision.

Even so, in spite of these and numerous other obvious implausibilities, Descartes's system was undeniably successful, being embraced by natural philosophers all over Europe. Indeed, his 'mechanical philosophy', (Box 11.1) as it came to be known, really did take over from Aristotelianism as the system of philosophy which seemed to offer the best way of understanding the natural world. The reason for this seems pretty clear. By 1644, when Descartes finally published his system, everyone knew that the Aristotelian system was untenable, but nobody could give it up until a new, alternative, and equally comprehensive system was available. Descartes's system, for all its faults, was the first one that seemed to have any real chance of replacing Aristotelianism, lock, stock and barrel.

Another crucially important aspect of Descartes's philosophy was its codification of precise laws of nature. Talk of laws of nature had been commonplace since the Ancient Greeks but the notion was entirely vague and indiscriminate. It was a law of nature that bees make honey, and that sparks fly upwards. But because Descartes hoped to explain everything in terms of matter and motion, he could concentrate on codifying the laws by which motion, in collisions, was transferred from one moving body to another. It would have been inconceivable in earlier natural philosophical systems to stipulate a restricted number of laws, from which all else followed. Descartes's system did not just allow the formulation of a restricted number of laws, it actually demanded them as a means of making the system cogent. Descartes's three laws of nature did not fully capture the workings of the world (Descartes hampered himself by trying to explain everything in terms of motion) and were soon superseded by Newton's laws (see Chapter 13), but it is fair to say that in propounding them Descartes set the scene for all future physical science. Precise laws of nature, which can be used to make predictions about the future behaviour of physical systems, are now a *sine qua non* for the physical sciences, but before Descartes they did not exist. (It is important to note in this regard that Galileo's law of free fall, and Kepler's laws of planetary motion, were designated as laws not by Galileo or Kepler, but by later thinkers who saw them as conforming to Cartesian demands as to what laws of nature should look like.)

As we saw at the beginning of this chapter, Descartes himself was confident that his system was capable of explaining all natural phenomena. He even extended his claims to cover the living world and so saw it not only as a replacement for Aristotelianism but also as a replacement for Galenism in medicine. Consider, for example, Descartes's theory of the heart/blood system (see Box 11.2).

Box 11.1 Descartes's mechanical philosophy

1 A new theory of matter

- All properties and qualities of a body are due to its shape, size and motion, but this sometimes has to be understood in terms of the shapes, sizes, arrangement and motions of its constituent particles.
- All objects considered to be composed of invisibly small particles of matter ('atoms', or 'corpuscles').
- Action-at-a-distance is impossible.
- There are no such things as 'occult qualities', 'powers', 'sympathies' and 'antipathies', 'vital' principles or 'animate' principles, in matter or anywhere else.
- Distinction drawn between primary qualities (such as shape, size, arrangement and motion), and secondary qualities, such as heat, colour, texture, and others which were all regarded as reducible to primary qualities (e.g. atoms or corpuscles do not have a colour but their different speeds of motion give rise to the sensation of different colours in the eye/brain).

2 A new concept of causation

- Aristotelian explanation in terms of four causes (material, formal, efficient, and final) was rejected.
- Only efficient causes are regarded as intelligible (i.e. what immediately brings about a new state of affairs), and they must be understood in terms of contact actions, collisions, or inter-meshing of material bodies (causation analogous to clockwork, or snooker).
- Given the assumption about the atomistic or corpuscularian structure of bodies, and rejection of occult powers, efficient causation often summed up in terms of 'matter in motion'.
- The only valid concept of 'force' is 'force of impact' or 'force of motion'.
- Efficient causes were codified in the form of three highly specific laws of nature.

3 A new notion of the nature of 'nature'

- All phenomena, including biological phenomena, considered to be explicable in 'mechanistic' terms, by appeal to mechanical models or analogues. The world as a giant clockwork.
- Previous distinctions between the 'natural' and the 'artificial' rejected. Machinery is no longer the province of natural magic, rather than natural philosophy. Nature is God's handiwork, superior to man-made machinery, but not different in kind.

For Descartes, animals had no souls, no living principle which could make them move. They were simply complex machines. He rejected the Aristotelian position that the self-movement of animals proved they had souls by pointing to the fact that clocks and other automata were capable of moving themselves, but had no souls. Nevertheless, Descartes did still believe that humans have souls. In the case of the human soul, its main function is to think. The soul is

Box 11.2 Descartes on the heart

Descartes is aware of Harvey's discovery but he has to adapt it to suit his purposes. Remember that Harvey believes the heart/blood system is a self-contained system which is the living principle in the body. As such, he believes that the heart and blood are the seat of the soul, and can be moved by the soul - so no further explanation of why the heart continues to beat is required.

Descartes needs to find a mechanical explanation for the continued beating of the heart, invoking only matter and motion rather than something like a soul. To do this he makes a partial return to Galenic theory. According to Galen the active stroke of the heart is expansion (called *diastole*), Harvey showed by impeccable experiments that the contraction of the heart (*systole*) was the active stroke (expansion occurring simply as the heart relaxed). Descartes now disregards these experiments of Harvey's and assumes that *diastole* is the active stroke.

According to Descartes, a few drops of blood drip into the left ventricle of the heart from the pulmonary vein – this blood is coming from the lungs where it has been cooled down. The interior of the left ventricle is very hot, Descartes says, and so the cool blood, when it hits the heat, rapidly (almost explosively) expands, causing the heart to dilate. The expanding blood bursts forth into the aorta (the main artery), and goes off into the arterial system. The heart now collapses (which looks like contraction) and a few more drops of blood enter, and the explosive expansion takes place again, sending more blood off into the aorta.

The model is a bit like the internal combustion engine. Instead of a spark igniting the incoming fuel (blood), we have a fire burning in the left ventricle. This fire was declared by Descartes to be burning without light, but to be otherwise the same as the fire with which we are familiar. It came to be known as 'Cartesian fire'.

Descartes insisted that his account of the heart followed as necessarily from the arrangement of its parts as the movements of a clock followed from the arrangement of its cogs and wheels. And yet, it was incompatible with Harvey's account, and with many of Harvey's experiments with ligatures in vivisected animals, which showed that the contraction of the heart was the active stroke, not its expansion (which occurred as the heart went flaccid).

But, Descartes could not accept Harvey's account, because he would then have to admit he couldn't explain what kept the heart beating.

the seat of reason. But because we have free will (unlike animals who just react to the world on instinct), our souls can cause our bodies to move in particular ways.

This is Descartes one major concession that the mechanical philosophy cannot explain everything. The soul, rational and immortal, belongs to another kind of existence. The physical world is the world of *res extensa*, extended things, while the soul is a *res cogitans*, a thinking thing. Remember that for Descartes an extended thing is necessarily a material thing, and the soul is not material, so it cannot be extended, but it must be thinking.

On the face of it, there were advantages to this aspect of Cartesianism. The dualist separation between body and soul fitted in with the dualism of the Platonized Christianity of the Roman Catholic Church. But Descartes's dichotomy between *res cogitans* and *res extensa* seemed to many thinkers to open up too big a gulf. How can a completely immaterial thinking entity move a material entity? An immaterial soul can't push against a fleshy carcase and make it move – it would simply pass through whatever it pushed against, like a ghost.

Unfortunately for Descartes, a number of his followers chose to solve this problem by simply declaring that there was no such thing as the human soul. We, like the animals, are just machines. Freedom of the will is merely an illusion – we all just react on instinct. These same followers of Descartes did not stop there. They also declared that there was no God. According to Descartes, God created the world and set the whole kit and caboodle moving. After that, however, God did not have to do anything. He imposed laws of nature on the world system and everything carried on in accordance with the laws of nature with no further help from God. It was an easy step from here to suppose that perhaps the world and its laws had always existed, and that God did not exist.

So, although Descartes's system was set up, using the '*cogito*' argument and the ontological proof of the existence of God, to combat scepticism and atheism, it was appropriated by atheists as their system of choice – an entirely mechanistic system devoid of both God and the soul. This is an aspect of Descartes's legacy which still flourishes – in our secular world most of us believe in a mechanical world.

FURTHER READING

Desmond Clarke, 'Descartes' Philosophy of Science and the Scientific Revolution', in John Cottingham (ed.), *Cambridge Companion to Descartes* (Cambridge: Cambridge University Press, 1992), pp. 258–85.

Peter Dear, *Revolutionizing the Sciences* (Basingstoke: Palgrave, 2001), chapter 5, pp. 80–100.

Stephen Gaukroger, *Descartes' System of Natural Philosophy* (Cambridge: Cambridge University Press, 2002).

John Henry, *The Scientific Revolution and the Origins of Modern Science*, 3nd edn (Basingstoke: Palgrave, 2008), chapter 5, pp. 69–84.

Genevieve Rodis-Lewis, 'Descartes' Life and the Development of his Philosophy', in John Cottingham (ed.), *Cambridge Companion to Descartes* (Cambridge: Cambridge University Press, 1992), pp. 21–57.

Genevieve Rodis-Lewis, *Descartes, His Life and Thought* (Ithaca: Cornell University Press,1999).

Martin Tamny, 'Atomism and the Mechanical Philosophy', in R.C. Olby *et al.* (eds), *Companion to the History of Modern Science* (London: Routledge, 1990), pp.597–609.

R. S. Westfall, *The Construction of Modern Science: Mechanisms and Mechanics* (Cambridge: Cambridge University Press, 1977), chapters 2 and 3.

12

The Royal Society and Experimental Philosophy

The mechanical philosophy effectively superseded and replaced the Aristotelian philosophy which had thrived in the university curriculum throughout Europe since the foundation of the universities in the thirteenth century. Descartes by no means had everything his own way, however. The very apparent problems with the Cartesian system meant that a number of other thinkers tried to improve on it and so there were a number of rival versions of the mechanical philosophy. Alternatives were developed, for example, by Pierre Gassendi (1592–1655), a French priest; Thomas Hobbes (1588–1679), more famous for his work of political theory, *Leviathan* (1651); and Sir Kenelm Digby (1603–65), an English Roman Catholic philosopher.

Arguably, the most important alternative version, however, was developed by a group of English natural philosophers who came together during the so-called interregnum in England (the period in between the execution of Charles I by a group of English parliamentarians in 1649 and the restoration of Charles II in 1660). At the Restoration, these and other like-minded natural philosophers formed the Royal Society, one of the first ever scientific research institutions, and made their so-called 'Experimental Philosophy' characteristic of the Society.

The full name of this new institution was The Royal Society of London for the Promotion of Useful Knowledge, but since it was the first ever 'Royal Society' of anything, it was usually referred to simply as the Royal Society. It is still going strong, of course, and now its official title is simply The Royal Society.

Perhaps the most famous founder members were Robert Boyle (1627–91), Robert Hooke (1635–1703), and Christopher Wren (1632–1723), but there were many other less well-known members. Although the leading members of the Society subscribed to a natural philosophy that was loosely mechanistic in its orientation, they wanted to dissociate themselves from the predominantly rationalistic approach of Descartes, and so dubbed their own natural philosophy, the experimental philosophy. But this experimental philosophy was not modelled on the experimentalism of Gilbert, or Galileo, or even Harvey (which was what we would now call hypothetico-deductive in its approach). The Royal Society explicitly claimed to be promoting a much more Baconian way of doing experiments. In other words, they claimed that they had no preconceived ideas, hunches, hypotheses, or the like, but that they simply performed experiments to observe what would happen in certain circumstances, and to add to the sum total of natural knowledge. They were, in short, gathering facts, not promoting particular preconceived theories. Accordingly, the Society saw itself as trying to bring about Francis Bacon's Great Instauration, by setting up a collaborative enterprise to gather as much knowledge of the natural world as it could.

So, the question immediately arises, why did this group of mechanical philosophers prefer to keep quiet about their mechanistic predilections, and to present themselves instead as though they were Baconians, working without any theoretical preconceptions (mechanistic or otherwise)? To answer this question we need to understand the historical background in Restoration Britain.

Ever since the thirteenth century, when Thomas Aquinas (1225–74) had successfully reconciled the pagan aspects of Aristotelianism with Roman Catholic theology, natural philosophy had been regarded as a 'handmaiden' to the 'Queen of the Sciences', theology. In other words, natural philosophy was used to bolster and support theology. Now, this was acceptable to all while there was only one theology – represented by Roman Catholicism. But after the Reformation, when alternative versions of Christianity began to flourish alongside Catholicism (Lutheranism, Calvinism and other Protestant movements), there were corresponding attempts to appropriate natural philosophy to support each particular theology.

In England, things came to a head during the Civil War, and the following interregnum period, after the execution of Charles I (1649) and before the Restoration of Charles II (1660). During these years, the Church of England was proscribed and so the Church courts and Church censorship were suspended. As a result, large numbers of radical Protestant sects began to flourish in England. These included Ranters, Levellers, True Levellers, The Family of Love (or Familists), Baptists, Anabaptists, and Quakers, to mention a few.

These radical sects were often highly subversive in their social and political ambitions, and were regarded with great suspicion by orthodox, and conservative believers. Characteristic attitudes among the sects included illuminationism – the belief that religious truths came to everyone directly by divine illumination (so that the most ignorant could claim to be in possession of as much religious truth as any Oxbridge-educated bishop) – and antinomianism – the belief that the true believer was above the law, and that, therefore, they could sin without it being a sin (because they were in a state of divine grace). Clearly, these were subversive and what would later be called anarchistic views.

Some of these sects, in time-honoured fashion, used the latest ideas in natural philosophy to support their claims about the true religion. For example, the sect of Mortalists, who denied the immortality of the soul, and insisted instead upon the bodily resurrection of all the dead on the Day of Judgment, used Harvey's ideas about the soul being in the blood (which they interpreted in a materialistic way), to support their ideas. Other sectarians embraced the alchemical worldview of the radical Swiss alchemist, Paracelsus (1493–1541), to support illuminationist versions of theology. The result was that orthodox thinkers saw innovatory ideas in natural philosophy as nothing more than ideas which were deliberately developed to support some outlandish and subversive, religio-political position.

Box 12.1 New organizations for science and the contested role of universities

The new appearance of formal societies or academies devoted to the study of nature has been recognized by historians as another important characteristic of the Scientific Revolution. In what Bernard de Fontenelle (1657–1757), secretary of the Académie Royale des Sciences from 1697, called the 'new Age of Academies', groups of thinkers began to come together to collaborate in the new understanding of the natural world. In some cases the group was called together by a wealthy patron with an interest in natural knowledge and its exploitation. One of the earliest of these was the group of alchemists, astrologers and other occult scientists brought together at the court of Rudolf II (1552–1612) in Prague (which included Tycho Brahe, and Kepler), another was the Accademia dei Lincei (Academy of the Lynxes), founded by the marchese di Monticello, Federico Cesi (1585–1630). The evident attractiveness of such collaborative enterprises can also be seen in the astonishing interest shown by scholars all over Europe in the Brotherhood of the Rosy Cross, whose intended reforms of learning, based on alchemy, Paracelsianism and other occult ideas were announced in two manifestos which appeared in 1614 and 1615. In fact, to the disappointment of those like Descartes, who tried to make contact with them, the Brotherhood seems to have been as fictitious as Salomon's House, Francis Bacon's ideal research institute described in his *New Atlantis* (1627). If Rosicrucianism came to nothing, however, Bacon's vision, as we have already seen, was to have profound effect. The two most important Scientific Societies, the Royal Society of London and the Parisian Académie Royale des Sciences, were both based on the collaborative, empirical investigation of nature and the intention to pragmatically exploit natural knowledge which were characteristic of Bacon's vision.

Both of these major societies can be seen to have grown out of less formal groupings of experimentally minded natural philosophers, of which there were a number scattered throughout Europe. These two major societies, based in the two leading cities of Europe, London and Paris, were to have enormous influence upon the subsequent development of science. Their new style of pursuing natural knowledge, and the impressive content of their new discoveries, were broadcast in the journals which they immediately began to produce, as well as in the publications and correspondence of their members.

➡

Thomas Hall (1610–65), a Presbyterian minister (famous for decrying dancing around the maypole and other springtime revels), for example, spoke of the threats to society of a contemporary 'Familiasticall-Levelling-Magicall temper'. He thereby lumped together the subversive ideas of the Familists (who advocated free love), the Levellers (who were proto-communists, suggesting all men were equal, or on a level), and alchemical or magical world views. These were all linked in his mind, and were all bad.

Indeed, radical sectarian ideas became so prominent that some contemporaries began to think in terms of what would now be called 'conspiracy theory'. Conser-

The deliberately reformist attitudes of the early scientific societies, and their public pronouncements of their methods and intentions in journals and other publications, mark them out as completely different from the universities. It used to be said that the universities during this period were moribund institutions, completely enthralled by traditional Aristotelianism, and blind to all innovation. This has now been shown to be completely unjustified, and the important contributions of some members of university Arts and Medical Faculties to scientific change has been reasserted. Nevertheless, it seems fair to say that it was usually individual professors who seemed innovatory, not the institutions to which they belonged. If there were exceptions to this it was in the smaller German universities, where the local prince might hold greater control over the university by his patronage. A number of such universities introduced significant changes in their curricula. In particular, the introduction of what was known as *chymiatria* or chemical medicine (embracing Paracelsianism and rival alchemically inspired forms of medicine) radically transformed a number of German universities. Even so, for the most part it remains true to say that the European universities in general seemed slow to change and institutionally committed to traditional curricula, even if individual professors might seem innovatory. In the case of the new academies or societies, however, the institutions themselves seemed innovatory, and they had a much greater effect on changing attitudes to natural knowledge.

FURTHER READING

Luciano Boschiero, *Experimental natural Philosophy in Seventeenth Century Tuscany: The History of the Accademia Del Cimento* (Dordrecht: Springer, 2007).

John Gascoigne, 'A Reappraisal of the Role of the Universities in the Scientific Revolution', in D. C. Lindberg and R. S. Westman (eds), *Reappraisals of the Scientific Revolution* (Cambridge: Cambridge University Press, 1990), pp. 207–60.

Roger Hahn, *The Anatomy of a Scientific Institution: The Paris Academy of Sciences , 1666–1803* (Berkeley: University of California Press,1971).

Michael Hunter, *Establishing the New Science: The Experience of the Early Royal Society* (Woodbridge: Boydell Press, 1989).

James E. McClellan III, *Science Reorganized: Scientific Societies in the Eighteenth Century* (New York: Columbia University Press, 1985).

vative thinkers began to believe that the kind of subversive ideas propagated by the sectarians were deliberately spread among the ill-educated in England in order to promote subversion and to fragment society. The bogeyman responsible for this conspiracy against sound English Protestantism was, of course, the Roman Catholic Church. This conspiracy theory was summed up by Thomas Barlow, later Bishop of Lincoln:

> It is certain this New Philosophy (as they call it) was set on foot, and has been carried on by the Arts of Rome … for the great Writers and

> Promoters of it are of the Roman Religion: such as Descartes, Gassendus, Du Hamel, Mersennus, etc. ... What divisions this New Philosophy has caused amongst Protestants in Holland and England, cannot be unknown to any considering person.

So, the new natural philosophies, including the mechanical philosophy of Descartes, were associated on the one hand with radical sectarianism, which was seen as highly subversive and anti-social; and on the other with Roman Catholic conspiracies to overthrow the English Government and to re-establish Catholicism in England.

As if that was not enough, the new philosophy was also associated with atheism. We saw at the end of the last chapter that Descartes's system, for example, was appropriated by atheists. Furthermore, the rival mechanical philosophy developed in England by Thomas Hobbes was also regarded as atheistic. Hobbes was a committed materialist and his version of the mechanical philosophy lacked the dualistic acceptance of the immortal soul that Descartes insisted upon (see previous Chapter). Accordingly, Hobbes was widely regarded as an atheist, and the mechanical philosophy as a materialist and atheistic philosophy.

The leading natural philosophers in interregnum England, men like Robert Boyle, were acutely embarrassed to see the new natural philosophies besmirched in this way. Most of the thinkers in England, who were committed to developing the new philosophy, were devout religious believers of a generally conservative kind. A significant number of them had even taken up natural philosophy during the time the Anglican Church was proscribed, and were hoping that one day the Church of England would be re-established.

This eventually came about in 1660 with the restoration of the monarchy, and the immediate re-instatement of Anglican power in Church courts and censorship. Most of the radical sects faded away. Boyle, Wren, and others who had been sitting out the interregnum period pursuing experimental science, now decided to form a scientific society, for fruitful collaboration, and they asked the newly restored monarch, Charles II, to be their patron. The Royal Society was formed.

The trouble is, the image of natural philosophy had become so tarnished during the interregnum that the Royal Society was, right from its inception, regarded with some suspicion. It was accused of being a crypto-Catholic enclave – intended to distract the best minds from affairs of state while Jesuits inveigled their way into the country. As one member of the Society reported, this collaborative enterprise of some of the greatest naturalists in the country was regarded by many as nothing more than 'a Company of Atheists, Papists, Dunces, and utter enemies to all learning'.

Consequently, the founders of the Royal Society decided they had to defend themselves from such charges and restore the good name of natural philosophy. But because of what had gone before, this was no longer an easy task. What they had to do was to show that all earlier natural philosophies were only able to support Catholicism, or Mortalism, or whatever religious views were in question, by being twisted or perverted to suit that religion. We would say that they had to show that all earlier versions of natural philosophy were *ideological* (but they didn't have this word in the seventeenth century).

Second, they had to show that *their* version of natural philosophy is not in the least bit ideological; that it has no axe to grind, no hidden agenda, to promote a particular religion. To pull this off, of course, they cannot use sophisticated and subtle arguments, claiming in a complex way that they have no hidden agenda, that would look downright suspicious. So, they have to try to ensure that their natural philosophy is *obviously* unbiased and ideologically neutral. It should be plain for all to see, not something that people have to be persuaded of by protracted argument.

Now this is by no means an easy trick to pull off, and yet to a large extent the early members of the Royal Society managed it. So, how did they do so?

First, they went back to the methodology recommended by Francis Bacon, who had died in 1626. They presented the Royal Society as a belated attempt to bring about 'Salomon's House', the ideal scientific research institute which had been described by Bacon in his utopian work, *New Atlantis* (1627). They claimed they were simply trying to bring about Bacon's Great Instauration.

This suited their purposes very well because of Bacon's well-known emphasis on purely factual knowledge. As we have seen (Chapter 7), Bacon believed the best way to proceed was to gather facts on a wholesale scale, without any preconceived theoretical assumptions. Accordingly, Thomas Sprat, author of *The History of the Royal Society of London* (1667), an apology for the Society commissioned by its members, said that the members 'confin'd themselves to no order of subjects', but endeavoured only to 'heap up a mixed mass of experiments'.

Furthermore, the members of the society claimed that their experiments were not intended to test a preconceived hypothesis, as in Gilbert, Galileo, Harvey, or Descartes (such as his account of the heart, for example), but merely to establish the 'matter of fact'. Experiments were conducted without any theories, just to see what actually happens.

This Baconian method was presented not only in Sprat's *History* (which was really a manifesto of the Society, rather than a history), but in all the writings of the fellows of the Royal Society. Robert Hooke (1635–1703), for example, included in his famous book of microscopical observations, *Micrographia* (1665), a Preface 'To the Royal Society' in which he wrote:

> The Rules you have prescrib'd your selves in your Philosophical Progress do seem the best that have ever yet been practis'd. And particularly that of avoiding Dogmatizing, and the espousal of any Hypothesis not sufficiently grounded and confirm'd by Experiments. This way seems the most excellent, and may preserve both Philosophy and Natural History from its former Corruptions.

There were two important outcomes of these efforts to clean up the image of natural philosophy. First, it helped to establish the claim of science to present so-called objective and unbiased knowledge. Second, it meant that English natural philosophers could reintroduce occult ideas into their version of the mechanical philosophy. Descartes had been so committed to excluding occult qualities and powers that he was forced to introduce entirely speculative ways of explaining such phenomena in terms of matter in motion (remember, for example, his claims about magnets sending out perpetual streams of screw-threaded particles, some with a right-hand thread and some with a left-hand).

Members of the Royal Society, by contrast, were simply able to say that experiments with magnets revealed, as a matter of fact, that a magnet can attract a piece of iron across a distance. No experiment could reveal the screw-threaded particles described by Descartes, or the screw-threaded channels in magnets or iron through which the magnetic particles were supposed to flow. Descartes did not deal in undisputed facts, therefore, but in hypotheses, and these hypotheses might have been formulated to promote some hidden agenda – he did, after all, try to provide an account of how his philosophy could be used to support Catholic claims about the Eucharist. The Royal Society, by contrast, could simply point to the fact of magnetic attraction, and felt no compunction to make claims about how magnets work.

The culmination of this trend, and its importance for the history of science, can be seen in the famous response by Isaac Newton (1642–1727) to criticism of his natural philosophy from the Cartesian philosopher, G. W. Leibniz (1646–1716). Leibniz pointed out that Newton, in his *Principia mathematica* (1687) gave no explanation of how gravity works (clearly, Leibniz wanted some kind of explanation in terms of streams of invisible particles pushing things downwards, as Descartes had suggested). In the second edition of his great book (1713) Newton simply insisted, 'I do not feign hypotheses'. Leibniz had accused Newton of allowing occult qualities back into the mechanical philosophy, because his notion of gravity was simply an occult power. But Newton responded that

> hypotheses, whether metaphysical or physical, or based on occult qualities, or mechanical, have no place in experimental philosophy.

Descartes's mechanical explanation of gravity, in terms of invisibly small descending particles, was every bit as hypothetical (and as implausible) as any occult explanation. By contrast, Newton was effectively saying to Leibniz, 'I just deal in facts'. Newton had shown that gravity works in accordance with a precise mathematical law (the inverse-square law), and so 'really exists', but any attempt to explain how it works would simply be guesswork, and not in keeping with the Royal Society's Baconian ideals.

FURTHER READING

Marie Boas Hall, *Promoting Experimental Learning: Experiment and the Royal Society, 1660–1727* (Cambridge: Cambridge University Press, 1991).

Peter Dear, '*Totius in verba*: Rhetoric and Authority in the Early Royal Society', *Isis*, 76 (1985), pp. 145–61 (also available in P. Dear (ed.), *The Scientific Enterprise in Early Modern Europe*, pp. 255–72).

John Henry, 'England', in R. Porter and M. Teich (eds), *The Scientific Revolution in National Context* (Cambridge: Cambridge University Press, 1992), pp. 178–210.

Michael Hunter, *Science and Society in Restoration England* (Cambridge: Cambridge University Press, 1981), Ch.1, pp. 8–31.

Michael Hunter, *Establishing the New Science: The Experience of the Early Royal Society* (Woodbridge: Boydell Press, 1989).

James McClennan III, 'Scientific Institutions and the Organization of Science', in Roy Porter (ed.), *The Cambridge History of Science, Volume 4: Eighteenth-Century Science* (Cambridge: Cambridge University Press, 2003), pp. 87–106.

Margery Purver, *The Royal Society: Concept and Creation* (London: Routledge & Kegan Paul, 1967).

Lewis Pyenson and Susan Sheets-Pyenson, *Servants of Nature: A History of Scientific Institutions, Enterprises and Sensibilities* (London: Fontana, 1999).

Paul B. Wood, 'Methodology and Apologetics: Thomas Sprat's *History of the Royal Society*', *British Journal for the History of Science*, 13 (1980), pp. 1–26.

13

Experiment, Mathematics and Magic: Isaac Newton

Isaac Newton is widely regarded as one of the supreme scientific geniuses of all time. Scientists usually include him in their top three or four, alongside Albert Einstein (1879–1955) and Niels Bohr (1885–1962), or Charles Darwin (1809–82). This is all the more remarkable bearing in mind that he lived in the seventeenth century – Einstein and Bohr had the advantage (with regard to scientific achievement) of living in the twentieth century.

His reputation is based almost entirely on just two books. The *Philosophiae naturalis principia mathematica* (*Mathematical Principles of Natural Philosophy*, 1687; 2nd edn, 1713; 3rd edn, 1726), which is regarded as an exemplary text in what would now be seen as mathematical physics. And *Opticks, or A Treatise of the Reflections, Refractions, Inflections and Colours of Light* (1704; Latin edn, 1706; 3rd edn, 1717), which is regarded as an exemplary text in experimental science.

In later life Newton claimed that he did his best work when he was a junior fellow at Trinity College, Cambridge, at the age of 23. His so-called *annus mirabilis* (marvellous year) took place from 1665/6, when he went back home to Grantham, near Lincoln, while Cambridge University was closed down due to plague. Newton himself described his achievements like this (I've added some explicatory comments in square brackets):

> In the beginning of the year 1665 I found the method of approximating series & the Rule for reducing any dignity of any Binomial into such a series. The same year in May I found the method of Tangents of Gregory & Slusius, & in November had the direct method of fluxions [*i.e.* he invented differential calculus!] & the next year in January had the Theory of Colours [*i.e.* much of the material he published in 1704 in *Opticks*] & in May following I had entrance into the inverse method of fluxions [*i.e.* he invented integral calculus!]. And the same year I began to think of gravity extending to the orb of the Moon, & having found out how to estimate the force with which a globe revolving within a sphere presses the surface of the sphere: from Kepler's rule of the periodical times of the Planets being in a sesquialterate proportion of their distances from the centres of their Orbs [a reference to Kepler's third law of planetary motion], I deduced that the forces which keep the Planets in their Orbs must be reciprocally

> as the squares of their distances from the centres about which they revolve: & thereby compared the force requisite to keep the Moon in her Orb with the force of gravity at the surface of the Earth, & found them answer pretty nearly. All this was in the plague years of 1665 and 1666. For in those days I was in the prime of my age for invention & minded Mathematicks & Philosophy more than at any time since.

This was recalled by Newton when he was an old man but since Newton was one of those characters who horded his own notes and papers, scholars have been able to reconstruct when he did what, and for the most part these claims are substantially true. He really did have an *annus mirabilis*.

However, what is not quite true is his suggestion here that he came up with the universal principle of gravitation at this time ('I deduced that the forces which keep the Planets in their Orbs must be reciprocally as the squares of their distances from the centres about which they revolve').

Certainly, at this time he did realize that the force which makes an object on Earth (say, an apple) fall to the ground, and the force which keeps the Moon in orbit around the Earth were the same. *But this was not an original idea of Newton's.* Any Cartesian was already used to the idea that the fall of an object due to gravity and the orbit of a planet was explained in the same terms. It was the first generation of Newton hagiographers, men like William Stukeley (1687–1765), and Henry Pemberton (1694–1771), who (conveniently) forgot Descartes's achievement and presented Newton's discovery of 'universal gravitation' as a eureka moment.

It seems clear, however, that what Newton did while he was at home in Lincolnshire during his *annus mirabilis* was simply to use his mathematical skills to test whether Descartes's assumption was right. This is what he meant when he said he 'compared the force requisite to keep the Moon in her Orb with the force of gravity at the surface of the Earth, & found them answer pretty nearly'.

It is clear from Newton's papers that he was thinking at this time entirely in Cartesian terms. That is to say, he accepted the Cartesian assumption that the planets were kept in their stable orbits as a result of the balance of two forces – one outwards from the centre (centrifugal) and one inwards to the centre (centripetal). The Moon's revolutions around the Earth resulted in a centrifugal force (just as a stone in a sling whirled around one's head has a tendency to move outwards from its centre of rotation), but this was countered by the ever-descending streams of particles (generated in accordance with Descartes's vortex physics) pushing the Moon towards the Earth. An apple (or any object on Earth) was not sufficiently affected by centrifugal force to counter the force of the descending particles, however, and so would fall in accordance with the centripetal force. Descartes himself took it on trust that this scenario must be correct; he never tried to provide the mathematics.

Newton was undoubtedly a brilliant mathematician, a careful experimenter, and an obsessive character who could solve problems, as he once said, 'by continually thinking unto them'. Even so, his story is not a tale of a lone genius who achieved what he did because he was a genius. As always, it is possible in Newton's case, no less than in the case of any other 'genius', to explain how he came to make the discoveries he did. We've already seen at the end of the last chapter how the characteristic approach to experimental philosophy developed by the Royal Society helped to make Newton the kind of thinker he was. And we've just seen that, during his *annus mirabilis*, he used his consummate mathematical skills to test a Cartesian hypothesis, and seemed to confirm it.

We already have, therefore, the basic ingredients required to explain Newton's success. Remember, the mechanical philosophy of Descartes confined itself to mechanical, contact action, as the only means of causation. Everything was explained in terms of matter in motion, and matter can only be moved by impact or collision with other bits of matter in motion. So, when it came to explaining occult forces or powers like gravity and magnetism, Cartesians had to concoct elaborate fantasies about screw-threaded particles and so on.

But, remember also, that the Royal Society rejected such hypothetical explanations, based only on imaginative guesswork, because they wanted to convince everyone that they dealt only in undeniable facts. Accordingly, they were able to simply point to gravitational and magnetic attraction and to declare them to be matters of fact, which could easily be codified in detail by experimental investigation – or, as in Newton's case, by mathematical analysis.

The Royal Society approach can be seen as a return to the tradition of natural magic. The magus was often forced to admit that he did not know how something operated (and so simply called it occult, or hidden), but he could show by experimental demonstration that it did operate this way. This return to natural magical ways is not so surprising if we bear in mind that the Royal Society took their inspiration from Francis Bacon, and that Bacon himself took inspiration from the natural magic tradition (see Chapter 7).

What we see in Newton's work, therefore, is the perfect blend of mechanistic thinking with magical thinking.

The driving force behind Newton's scientific work was to understand the operations of the secret powers of matter (this was manifested mostly in his alchemical work, which he spent far more time doing than mathematics or physics). As a result of his work in this area he concluded that there must be forces of attraction and repulsion operating between particles of matter.

When Newton based his *Principia mathematica* on three laws of motion, as Descartes had done in his *Principia philosophiae*, he introduced the concept of force into his laws. Descartes, by contrast, had been careful to exclude the occult notion of force from his laws (see Box 13.1). Descartes's laws are now held to be completely wrong, Newton's still form the basis for much of physical science (excluding those phenomena which can only be understood in terms of relativity and quantum theories).

Newton freely admitted he could *not* explain how gravity worked, but he could show in great detail how it worked in different situations – even being able to make mathematically precise predictions of planetary movements.

Box 13.1 COMPARISON OF DESCARTES'S AND NEWTON'S LAWS

Descartes's Laws of Nature (1644) [Notice that there is no mention of force]

1 That each thing, as far as is in its power, always remains in the same state; and that consequently, when it is once moved, it always continues to move.

2 That all movement is, of itself, along straight lines; and consequently, bodies which are moving in a circle always tend to move away from the centre of the circle which they are describing.

3 That a body, upon coming into contact with a stronger one, loses none of its motion; but that, upon coming in contact with a weaker one, it loses as much as it transfers to that weaker body.

Newton's 'Axioms, or The Laws of Motion' (1687)

1 Every body perseveres in its state of being at rest or of moving 'uniformly straight forward' except in so far as it is compelled to change its state by forces impressed.

2 A change in motion is proportional to the motive force impressed and takes place along the straight line in which that force is impressed.

3 To any action there is always an opposite and equal reaction; in other words, the actions of two bodies upon each other are always equal and always opposite in direction.

Therefore, Newton's achievement lay in combining the mechanical philosophy with a belief in the reality of occult forces and powers. His belief in the reality of these occult forces stemmed from his own work in alchemy, of course, and from a well-established tradition in English natural philosophy, beginning with William Gilbert, made philosophically respectable by Francis Bacon, and enshrined by the Royal Society. This was a tradition in which bodies were assumed to have secret or hidden powers which could be demonstrated and studied by experiment, even if they could never be explained in mechanistic terms.

Indeed, there is a clear indication that Newton's achievement was not based on inexplicable genius, but on the fact that, as he himself said, he stood 'on ye sholders of Giants'. Robert Hooke, the leading experimental philosopher in the Royal Society, had already concluded from his reading of Kepler's *New Astronomy* that planetary orbits could be explained, not in terms of a balance of a centrifugal and a centripetal force, as Descartes supposed, but simply in terms of inertial tangential motion bent into a closed orbit by a single attractive force operating inversely as the square of the distance between the Sun and the planet. He wrote to Newton in 1679, asking his opinion of this theory. Newton failed to notice the importance of Hooke's idea until 1685, when Edmond Halley (1656–1742), also wanting an opinion on Hooke's hypothesis, visited Newton and recounted it. It was only at this point that Newton produced a mathematical proof of Hooke's hypothesis. The act of proving it mathematically turned it, in the judgement of history (but speaking personally, not in my judgement), from Hooke's idea into Newton's idea.

Prior to this point, as the leading Newton commentator R. S. Westfall, pointed out, Newton had thought of attractive and repulsive forces only in alchemical contexts, as micro-forces operating between the invisibly small particles of bodies, but from this point on (thanks to Hooke and Halley) he saw these forces as relevant to the whole of physics, including cosmology.

Having provided Halley with mathematical proof that Kepler's laws of planetary motion could be explained on the supposition of a single attractive force operating between the Sun and the planets, and varying inversely as the square of the distance between them, Newton was persuaded to write the *Principia*, including this and other exercises in the mathematics behind physical phenomena.

Newton and the secret powers of matter

I said above that Newton was driven throughout his scientific career to understand the operations of the secret powers of matter. Let me now try to demonstrate this. Along the way, we will also see that Newton was affected not only by alchemical and other magical traditions, but that his thought was also profoundly shaped by his religious beliefs, and that these things were inextricably bound together in his thought.

Beginning with his ideas on the secret powers of matter, consider, for example, his own account of the main point of the *Principia* as presented in the Preface:

> The whole burden of philosophy seems to consist in this: from the phenomena of motions to investigate the forces of nature, and then from these forces to demonstrate the other phenomena...
>
> I deduce here the motions of the planets, the comets, the Moon, and the sea. I wish we could derive the rest of the phenomena of Nature by the same kind of reasoning from mechanical principles, for I am induced by many reasons to suspect that they may all depend upon certain forces by which the particles of bodies, by some causes hitherto unknown, are either mutually impelled toward one another... or are repelled and recede from one another. These forces being unknown, philosophers have hitherto attempted the search of Nature in vain; but I hope the principles here laid down will afford some light either to this or some truer method of philosophy.

Notice that Newton can imply that he is 'reasoning from mechanical principles', and yet he can go on to talk of explanations in terms of 'unknown forces' – no continental Cartesian could have written this way; it would have seemed contradictory.

But what did Newton mean by these 'many reasons' which made him think of 'certain forces'? What did he have in mind?

One important answer to this question was provided by the fact that Newton's mathematics made it clear that the force of gravitational attraction depended upon the entire masses of the bodies involved. If Cartesian accounts involving streams of descending particles pushing a body to Earth were true, then we might expect the surface area of the body to be a relevant factor; but gravity clearly worked on the innermost depths of a body. As Newton wrote of gravity in the General Scholium added to the second edition of the *Principia*:

> This is certain, that it must proceed from a cause that penetrates to the very centres of the Sun and Planets, without suffering the least diminution of its force; that operates, not according to the quantity of surfaces of the particles upon which it acts, (as mechanical causes use to do,) but according to the quantity of the solid matter which they contain, and propagates its virtue on all sides, to immense distances...

But among the other 'many reasons' that Newton had for believing in occult powers of matter we can consider his work on optics. At the end of the *Opticks*, Newton included a number of speculations, which he called 'Queries'. A number of the Queries clearly show that Newton appropriated alchemical ideas in which light is an active principle in the universe, a kind of driving force behind all motions.

> Do not bodies and light mutually change into one another? And may not bodies receive their most active powers from the particles of light which enter their composition? ... Now since light is the most active of all bodies known to us, and enters into the composition of all natural bodies, why may it not be the chief principle of activity in them?

Newton spent a great deal of his time in studying alchemy. But he was not trying to convert lead into gold – he was trying to discover the secret sources of motion and activity in matter. Consider one of his earliest alchemical works, usually referred to by Newton scholars as 'The Vegetation of Metals', written in 1669 or 1670:

> And thus perhaps a great part if not the whole mass of sensible matter is nothing but Aether congealed & interwoven into various textures... Note that tis more probable ye aether is but a vehicle to some more active spirit & ye bodys may bee concreted of both together, they may imbibe aether as well as air in generation & in yt aether ye spt is intangled. This spt perhaps is ye body of light 1 because both have a prodi-

> gious active principle both are perpetual workers 2 because all things may bee made to emit light by heat.

Newton still drew upon these ideas in 1675 when he sent his 'Hypothesis explaining the Properties of Light' to the Royal Society in 1675 (note that this wasn't published – Newton wasn't aware at this time of the Society's emphasis upon facts, and rejection of 'Hypotheses'):

> Perhaps the whole frame of nature may be nothing but various contextures of some certain aetherial spirits, or vapours, condensed as it were by precipitation, much after the manner that vapours are condensed in water, or exhalations into grosser substances, though not so easily condensable; and after condensation wrought into various forms; at first by the immediate hand of the Creator; and ever since by the power of nature; which, by virtue of the command, increase and multiply, became a complete imitator of the copies set her by the protoplast. Thus perhaps may all things be originated from aether... light and aether mutually act upon one another.

But light was not just linked to alchemy in Newton's mind but to other aspects of the magical tradition as well. Consider, for example, his account of the colours of the spectrum.

From Aristotle to Descartes the general assumption was that coloured lights were corrupted versions of pure light, which was of course the *white* light of the Sun. But Newton's experiments with prisms led him to believe that white light was *not* the pure form of light. On the contrary, the pure forms were the different colours of light revealed by passing white light through a prism. It was the result of blending the coloured lights of the spectrum together which created white light. This claim proved highly controversial. It seemed to make no sense. Everyone assumed that God had made white, sunlight, the original pure form of light.

Consequently, Newton made a move that was very reminiscent of Kepler. In just the same way that Kepler asked himself why God might have chosen elliptical orbits in preference to celestial circles, Newton asked himself why God might have chosen to make white light the result of a blending of coloured lights. To get the answer he turned (as Kepler might have done) to the ancient Pythagorean tradition of the Music of the Spheres.

Newton claimed in Book I, Part 2, of the *Opticks* that a projected spectrum shared the same ratios between the colours, as the octave did between its seven notes.

> I held the Paper so that the Spectrum might fall upon [it]... whilst an Assistant, whose Eyes for distinguishing Colours were more critical than mine, did by Right Lines... drawn cross the Spectrum, note the

> Confines of the Colours, that is of the red..., of the orange..., of the yellow..., of the green..., of the blue..., of the indico..., and of the violet... And this operation being divers times repeated both in the same, and in several Papers, I found that the Observations agreed well enough with one another, and that the [spectrum was]... by the said cross Lines divided after the manner of a Musical Chord... and so to represent the Chords of the Key, and of a Tone, a third Minor, a fourth, a fifth, a sixth Major, a seventh and an eighth above that Key: And the Intervals... will be the Spaces which the several Colours (red, orange, yellow, green, blue, indigo, violet) take up.

Marking the seven different colours along the spectrum, he insisted, matched the positions where you would have to bridge a monochord of corresponding length to sound the seven notes comprising the octave. Newton even provided a diagram. Threfore, white light was a glorious harmony of all the other colours 'sounding' together. In fact, in Newton's Optical Lectures, which he delivered to students at Cambridge from 1670 to 1672, Newton says that he could only distinguish five colours in the spectrum, but he deliberately added indigo and orange, 'in order to divide the image into parts more elegantly proportioned to one another'. Francis Bacon would not have approved.

Interestingly, we all believe there are seven colours in the rainbow because Newton was an obsessive believer in the magical Pythagorean tradition of the Music of the Spheres. Have a proper look next time you see a rainbow – you will never be able to see seven colours. Not even Newton could, but he needed there to be seven colours to fit his Pythagorean obsession, and the seven notes in the octave. We have subsequently gone along with it simply on the authority of Newton – he's a great scientist and so there really must be seven colours, even though we can only see five, or at best maybe six. The fact is, we believe in the seven colours of the rainbow because Newton was a latter day Pythagorean magus. As John Maynard Keynes famously wrote: 'Newton was not the first of the age of reason. He was the last of the magicians... the last wonder child to whom the Magi could do sincere and appropriate homage'.

This Pythagorean tradition also occurs in other parts of Newton's science. In unpublished manuscripts (although they were intended for publication in a second edition of the *Principia*) he claimed that the Pythagoreans used the theory of the harmony of the spheres to describe the Sun-centred universe, with six planets. By discussing it in these terms they could let adepts know the truth, while hiding the truth from the ill-educated. Newton also claimed that the Pythagoreans knew the inverse square law and the universal principle of gravitation.

> By what proportion gravity decreases by receding from the Planets the ancients have not sufficiently explained. Yet they appear to have adumbrated it by the harmony of the celestial spheres, designating the

> Sun and the remaining six planets... by means of Apollo with the Lyre of seven strings... But by this symbol they indicated that the Sun by its own force acts upon the planets in that harmonic ratio of distances by which the force of tension acts upon strings of different lengths, that is reciprocally in the duplicate ratio of the differences... In general terms, if two strings equal in thickness are stretched by weights appended, these strings will be in unison when the weights are reciprocally as the squares of the lengths of the strings. Now this argument is subtle, yet became known to the ancients. [Pythagoras] understood by means of the harmony of the heavens that the weights of the Planets towards the Sun were reciprocally as the squares of their distances from the Sun.

The religious dimension to this should already be apparent. Newton's musical cosmos revealed, as Kepler's did, the design of an intelligent Creator. But Newton did not leave this as nothing more than an implication in his work. He tied his studies of ancient Pythagorean philosophy in with other studies he was heavily engaged in – namely studies on the original religion, as taught to man directly by God.

Newton began his Scriptural studies in 1672 and, as with everything else he set his mind to, he became obsessive. He was soon spending even more time on this than he did on alchemy.

He came to the conclusion that the original religion, as taught to Adam, and which involved worship of the one true God, became corrupted into idolatry (which Newton saw as the worst sin) after the death of Noah. The true religion was reinstated by Moses, but again became corrupted into idolatry. God then sent Jesus to restore the original religion, but corruption set in again. Instead of worshipping the one true God, as they had been instructed by Jesus, Christians too became idolaters and began to worship Jesus as God. Compared to earlier religions, Newton said, Christianity 'was not more true, and did not become less corrupt'. As Newton concluded, 'the world loves to be deceived'.

It seems clear that Newton believed it was his duty to try to recover the original religion once again, and he seems to have believed that he was on the right lines precisely because his science also fitted in with ancient wisdom (in this case Pythagorean), and so his Scriptural studies were also leading him back to ancient wisdom.

Of course Newton kept all this secret – if he'd said in public that he rejected Trinitarianism and believed the Christianity of his day to be 'not more true, and... not... less corrupt' than pre-Christian religions he would have been in serious trouble.

Even so, Newton wanted desperately to know whether there was anybody else out there who had reached the same conclusions as he had. Perhaps there were many like him? And so, Newton alluded to his religious ideas, including his belief that the true religion first went wrong in the generation after Noah, in the final paragraph of the *Opticks*. The meaning of this paragraph has only just become clear in the last few years, as a result of research into Newton's religious papers. Before that it was utterly baffling, and yet readers simply assumed Newton was making a casual pious remark. In fact, he was showing, in the final paragraph of the *Opticks*, that as far as he was concerned his religion, his natural philosophy, and his alchemy and other magical beliefs were all interconnected:

> And if natural Philosophy in all its Parts, by pursuing this Method, shall at length be perfected, the Bounds of Moral Philosophy will be also enlarged. For so far as we can know by natural Philosophy what is the first Cause, what Power he has over us, and what Benefits we receive from him, so far our Duty towards him, as well as that towards one another, will appear to us by the Light of Nature. And no doubt, if the Worship of false Gods had not blinded the Heathen, their moral Philosophy would have gone farther than to the four Cardinal Virtues; and instead of teaching the Transmigration of Souls, and to worship the Sun and Moon, and dead Heroes, they would have taught us to worship our true Author and Benefactor, as their Ancestors did under the Government of Noah and his Sons before they corrupted themselves.

FURTHER READING

I. Bernard Cohen and G. E. Smith, *The Cambridge Companion to Newton*, Cambridge: Cambridge University Press (2002).

I. Bernard Cohen and R. S. Westfall, *Newton: Texts, Backgrounds, and Commentaries* (New York: Norton, 1995).

B. J. T. Dobbs, 'Newton's Alchemy and his Theory of Matter', in P. Dear (ed.), *The Scientific Enterprise in Early Modern Europe* (Chicago: University of Chicago Press, 1997), pp. 237–54.

Jan Golinski, 'The Secret Life of An Alchemist', in John Fauvel *et al.* (eds), *Let Newton Be!* (Oxford: Oxford University Press, 1988), chapter 7, pp. 147–68.

A. R. Hall, *Isaac Newton, Adventurer in Thought* (Cambridge: Cambridge University Press, 1996).

John Henry, 'Newton, Matter and Magic', in John Fauvel *et al.* (eds), *Let Newton Be!* (Oxford: Oxford University Press, 1988), chapter 6, pp. 127–46.

John Henry, 'The Fragmentation of the Occult and the Decline of Magic', *History of Science*, 46 (2008), pp. 1–48.

Rob Iliffe, *Newton: A Very Short Introduction* (Oxford: Oxford University Press, 2007).

Lord Keynes, 'Newton the Man', in Royal Society, *Newton Tercentenary Celebrations* (Cambridge: Cambridge University Press, 1947), pp. 27–34.

Frank Manuel, *A Portrait of Isaac Newton* (Cambridge, MA: Harvard University Press, 1968), chapter 6.

J. E. McGuire and P.M. Rattansi, 'Newton and the "Pipes of Pan"', *Notes and Records of the Royal Society of London*, 21 (1966), pp. 108–43.

R. S. Westfall, *Never at Rest: A Biography of Isaac Newton* (Cambridge: Cambridge University Press, 1980).

14

The Newtonian Enlightenment

The so-called Scientific Revolution could be said to have reached its zenith with Isaac Newton. The impressive authenticity of his achievement came to be so widely recognized that many believed the true natural philosophy had been achieved in all its essentials, and now all that was required was to fill in the actual details. We saw earlier (Chapter 11) how Descartes claimed in his *Principia philosophiae* that 'no phenomena of nature have been omitted from this treatise', but it was Newton's *Principia*, and his *Opticks*, which seemed to early eighteenth-century thinkers to have really explained all phenomena.

Newton's achievement was so extraordinary, so unprecedented, that it became exemplary for the following generation of natural philosophers. So much so, that to a large extent it shaped and characterized the succeeding age. The eighteenth century has been dubbed by historians the Age of Reason, and the period of The Enlightenment. For the majority of historians, who pay scant regard to scientific developments, the characterizing features of the Enlightenment are the rejection of traditional religious and political systems, and of what was deemed to be superstition, in favour of secularism, republicanism and liberalism, and new concepts such as the natural rights of man (and of woman), individualism and personal liberty.

There is a case to be made, however, for saying that all these aspects of the Enlightenment grew out of, and ultimately can be traced back to, the new widespread belief that scientific knowledge and the application of scientific methods will enhance, even perfect, our understanding in all areas of life. Intimations of Newton's role in the Enlightenment can be seen in the famous epigrammatic epitaph written for him by the English Enlightenment poet Alexander Pope (1688–1744):

> Nature and Nature's Laws lay hid in Night;
> God said 'Let Newton be!' – and all was light.

For the French intellectual, Voltaire (1694–1778), writing in his *Philosophical Letters* of 1734, Newton was not just the greatest scientist, but the greatest man who had ever lived.

> Not long since the trite and frivolous question following was debated in a very polite and learned company, viz., Who was the greatest man,

> Caesar, Alexander, Tamerlane, Cromwell, etc.? Somebody answered that Sir Isaac Newton excelled them all. The gentleman's assertion was very just; for if true greatness consists in having received from heaven a mighty genius, and in having employed it to enlighten our own mind and that of others, a man like Sir Isaac Newton, whose equal is hardly found in a thousand years, is the truly great man. And those politicians and conquerors (and all ages produce some) were generally so many illustrious wicked men. That man claims our respect who commands over the minds of the rest of the world by the force of truth, not those who enslave their fellow-creatures: he who is acquainted with the universe, not they who deface it.

Newton's impact was so great that it affected intellectual life way beyond the confines of science. His impact was not only far greater than that of any earlier natural philosopher, but also different in kind. Copernicus's achievement had been regarded with apprehension when it finally began to be known outside the confines of professional astronomers – it was seen as shattering the order of the old system and to be threatening religion. According to the metaphysical poet John Donne (1572–1631), thanks to Copernicus 'The sun is lost, and the earth, and no man's wit, Can well direct him where to look for it', and this was not just an astronomical matter, it signified that everything upon which we based our values was 'all in pieces, all coherence gone'. These attitudes fed into the sceptical crisis that followed the Renaissance (Chapter 11).

Newton's work, by contrast, inspired optimism. There was a general belief that Newton's method could be applied not only to science but also to moral philosophy, to politics, and to economics, so that even society itself could be run on scientific lines. They took Newton at his word, when he wrote at the end of the *Opticks*, 'if natural Philosophy in all its Parts, by pursuing this Method, shall at length be perfected, the Bounds of Moral Philosophy will be also enlarged'.

It is important to note, however, that Newton was very largely the cipher for a whole movement; he was the embodiment of the success of the new natural philosophies, and the ultimate exemplar of where the pursuit of natural knowledge could lead. We can easily see that it was *not* Newton, the man, who inspired the Enlightenment, but Newton, the supposed embodiment of the Age of Reason.

If Enlightenment thinkers had been inspired by the real Newton, the Newton we saw in the previous chapter, the Enlightenment would now be remembered as an age of alchemy, of natural magic, and of esoteric religion. The Newton who proved to be such a hero to eighteenth-century intellectuals was very much a Newton of their own invention. The triumphs of the *Principia* and of the *Opticks* were seen as the culmination, and proof of the soaring potential, of the new approaches to the natural world that had been developed by Francis Bacon, René Descartes, the Royal Society, and its French equivalent, the Académie des Sciences, and others in the new scientific movement. Accordingly, they saw him simply as a supreme mathematical physicist, and the thinker who had finally discovered the correct version of the mechanical philosophy.

Their Newton was very different from the real Newton we looked at earlier. This Enlightenment image of Newton became so characteristic of the new philosophy that it became reified and entrenched. The Newton we saw in the previous chapter has been uncovered, with great difficulty, by historians of science who have managed to see Newton in his own terms, and have refused to be led astray by the Enlightenment view of him. Even now, however, there is enormous resistance among other historians and philosophers of science to any suggestions that Newton accepted occult notions. A recent study of Newton as a philosopher, for example, takes it for granted that Newton could not possibly have accepted the supposedly untenable idea of occult action at a distance even though Newton himself was perfectly explicit in his use of this notion.

Two of the leading 'manifestos' of the Enlightenment (that is to say, books which can now be seen as setting the scene for, and summing up, the Enlightenment ethos), Voltaire's *Philosophical Letters* and The *Preliminary Discourse to the Encyclopaedia* (1751), written by one of the editors of the *Encyclopaedia*, the mathematician Jean Le Ronde d'Alembert (1717–83), presented a triumphalist intellectual history of their own time, which began with Francis Bacon, paid due tribute to Descartes as a magnificent failure, and culminated in their vision of Newton. According to D'Alembert, when Newton 'appeared at last', he 'gave philosophy a form which apparently it is to keep'. Newton had found the true philosophy, and its method, and it was then just a matter of time before all problems would be solved (see Box 14.1 on page 162).

Since we are dealing with an invented Newton, it is hardly surprising that he appears in different versions in different parts of Europe. The English and the French, for example, formed images of Newton to suit their own views of what Enlightenment should mean. In Britain, Newton's ideas were embraced by the orthodox and used to bolster religion; while in Catholic France, where anti-clericalism was rife, Newton's ideas were taken up by mechanical philosophers with secularizing, or even atheistic, tendencies.

The different attitudes of the British and French can be understood in the light of the recent Civil War in England, and the period of turmoil known as the interregnum. For the elite in Britain, mostly conservative and orthodox thinkers, the proliferation of numerous radical sectarian groups in the interregnum made it all too clear that this was a period of godlessness, a period characterized by lack of religion amongst the masses. For the intellectual elite in France, looking at Britain from the outside, the situation in interregnum Britain was clearly the result of religious fanaticism (represented in different ways by the various sects), and served as a lesson as to how dangerous independent religious thinking can be.

Given the fact that natural philosophy had always been used to support religion, and given the success of Newton's natural philosophy, it was inevitable that orthodox theologians in Britain should begin to use Newtonian natural philosophy to try to prove the existence of God (as Descartes had done in response to the sceptical crisis). It was in the eighteenth century that we see the flourishing in Britain of what is called natural theology – the use of the intricate wonders of the natural world to prove the existence of God as its obviously intelligent Creator. The more detailed this kind of study of nature was, the better – in those days it wasn't the Devil who was in the details, but God.

In France, by contrast, free-thinkers and other opponents of the Church used the new authority of Newtonian science in support of rationally based systems of moral philosophy – that is to say, ethical systems in which good behaviour was promoted as the rational thing to do, and not just as a means of avoiding punishment in hell-fire (which was emphasized in the moral teachings of the Church). These French thinkers took encouragement from Newton's claim, in the closing words of the *Opticks*, that the bounds of moral philosophy might also be enlarged by pursuing Newton's method in natural philosophy (see full quotation at the end of the previous chapter). According to them, religion was developed on purpose as the mainstay of a corrupt political and social order, and was based on terror, by threatening transgressors with hell-fire, and was opposed to science.

The enlightened man, they believed, rejected morality based on terror and threats of hell-fire, and extolled the virtues of a civil religion, a religion of man in society, where all the doctrines are laid down according to reason, where the basis for morality is to do unto others as you would have them do unto you (which, ironically, is actually Christianity's 'golden rule') and which is intended to preserve the freedom, and the rights of the individual as much as possible. This is what Newton meant, they believed, when he said moral philosophy could be established by following the methods of natural philosophy. As Condorcet wrote in 1795:

> The sole foundation for the belief in the natural sciences is this idea, that the general laws directing the phenomena of the universe, known or unknown, are necessary and constant. Why should this principle be any less true for the development of the intellectual and moral faculties of man than for the other operations of nature?

Eighteenth-century Newtonianism: forces and active fluids

We will come back to Newton's importance even beyond science later on (Chapter 16), but for now let us return to the historical development of science, and look at the way eighteenth-century natural philosophers took up the challenge implicit in the Preface of Newton's *Principia*, to explain all phenomena of nature in terms of 'certain forces by which the particles of bodies, by some causes hitherto unknown, are either mutually impelled toward one another... or are repelled and recede from one another'.

Box 14.1 TESTS OF NEWTON'S NATURAL PHILOSOPHY

Initially derided by Continental natural philosophers (mostly under the thrall of Descartes) for re-introducing occult qualities (chiefly gravitational attraction) into natural philosophy, French and German physicists eventually came to concede the superiority of Newton's *Principia mathematica* (1687) over Descartes's *Principia philosophiae* (1644). The main factors leading to the change of attitude were the following:

If Descartes was right, the Earth should be shaped like an oval spinning on one of its narrow ends (think of an egg spinning on its point), but if Newton was right it should be like an oval spinning on one of its long sides (like an egg on its side, or like a pumpkin).According to Descartes the greatest inward pressure on a sphere rotating at the centre of a vortex should be around its equator, deforming it to look like an egg on its point; according to Newton the equator of a rotating plastic sphere should experience the greatest outward (centrifugal) tendency, deforming it to the shape of a pumpkin.

It was decided by a group of French collaborators to test this by measuring the length of a degree of latitude as close to the North Pole as possible, and as close to the Equator as possible. This required two expeditions. One group set off north, to Lapland, the others went to what is now Ecuador in South America.

The group that went north were away for about a year (1736–7) and came back with some Finnish beauties who became, for a while, the toast of Parisian society. Meanwhile in the Amazonian jungle the other group had a very tough time and only made it back after nearly ten years (1735–44). When the results of their respective measurements were eventually compared, it was obvious that Newton was right.

Further confirmation came in 1749 with the solution of the so-called three-body problem. Even Newton couldn't fully explain the motions of the Moon, because it was affected by the simultaneous gravitational pull of the Earth and of the Sun. The leading Swiss mathematician, Leonard Euler (1707–83), had solved this problem in 1747 in such a way that he concluded Newton's theory was incorrect. Euler soon realized, however, that he (Euler) had made a mistake, which when corrected proved Newton right. Meanwhile, two leading French mathematicians, Alexis Clairault (1713–65), and Jean Le Rond D'Alembert (1717–83), independently solved the three-body problem, by using Newtonian assumptions.

Early in 1759, Halley's Comet returned. This had been predicted by Edmond Halley (1656–1742), based on noting historical records of a comet being sighted every 75 years. But Clairault used Newtonian mathematics to establish in advance the precise date (or impressively close to it).

FURTHER READING

Neil Safier, *Measuring the New World: Enlightenment Science and South America* (Chicago: University of Chicago Press, 2008).

Mary Terrall, *The Man who Flattened the Earth: Maupertuis and the Sciences in the Enlightenment* (Chicago: Chicago University Press, 2002).

There was also a variation on this basic idea, again developed by Newton, in which a subtle fluid, for which Newton borrowed the name 'aether' (which had long been used to refer to the material of the heavens), was used to explain all physical phenomena. The idea was that occult forces like gravity and magnetism, and pervasive entities like light, could be communicated across the universe as vibrations, or waves, in the aether.

Now some historians have regarded this as proof that Newton was a proper mechanical philosopher after all – and not really interested in magical, occult nonsense, like actions at a distance. But, this won't wash, because the aether, as described by Newton, is very clearly an aether which depends upon repulsive forces between its particles (in short, it allows Newton to reduce attractive and repulsive forces to just repulsive forces). Waves or pulses are not transmitted through the aether by collision, as in a Cartesian plenum, but by repulsive forces acting between the particles at a distance. As a particle moves closer to another particle, Newton proposes, that particle is repelled and has to move away, and so a shock wave is transmitted.

Newton needed an aether that was very thin and subtle, so that the planets could move through it without being slowed down by a drag effect, but it also had to be capable of transmitting light at a fantastic speed, and therefore needed to be very rigid, with no slack in the system. These two requirements might seem incompatible, but for Newton an aether made of excessively tiny particles (small even in comparison with other atoms) which were spaced far apart from one another (with vacuum in between) had the right level of tenuity to offer no resistance to the planets moving through it. But, by supposing that the particles are spaced so far apart because of a very powerful repulsive force operating between them, the 'rigidity' requirement was also met. This offered the right kind of rigidity, in the sense that a slight movement of one particle immediately effected surrounding particles.

So, in the speculative 'Queries' which Newton added to the end of his *Opticks*, he gave two suggestions for further research:

1 All physical phenomena can be explained in terms of attractive and repulsive forces operating between their constituent particles.
2 All phenomena can be explained in terms of active subtle fluids, or aethers, permeating the whole universe.

Now, some historians have discerned two separate traditions following from these suggestions, but the historical reality seems to be somewhat more messy with eighteenth-century thinkers participating in a spectrum of thought from forces at one end to aethers at the other, and taking up positions which draw selectively on both ideas.

One of the major sources from which other philosophers drew in the development of aether theories was Hermann Boerhaave (1668–1738), professor of chemistry and *materia medica*, at Leyden University in the Netherlands. According to

Boerhaave, fire was an active substance which pervaded the whole universe – penetrating even into the inmost recesses of solid bodies – and endowing matter with many of its activities. Accordingly, fire was seen as the driving force in the universe. Fire is expansive, self-diffusive, and it was capable of expanding everything else. It was an easy candidate for Newton's aether, and could also be seen as the source of repulsion and expansion in the universe. But Boerhaave also accepted that gravity was at work as the source of attraction and contraction.

Many identified Boerhaave's fire with light (remember, Newton himself talked of light being an active principle within matter) but a more promising identification seemed to be provided by the newly discovered phenomenon of electricity. As Edmond Halley wrote in 1731:

> Electricity is the confirmation of Sir Isaac Newton's notion concerning the existence of an universal medium which he sometimes calls the aether, at other times an electric spirit and which he apprehended was the cause of the phenomena of gravity, of light, of heat, and of electricity.

Here, then, was a clear case of a mysterious kind of fire which existed in all things, and could in some cases, be made to reveal itself with spectacular and powerful effects.

The generation of static electricity was an accidental discovery made during experiments with the air-pump which had been developed by Robert Boyle and Robert Hooke in an attempt to decide between rival mechanistic theories (Descartes believed a vacuum was impossible, others insisted that the atomic or corpuscular constituents of the world moved through a void). A partial vacuum in the chamber of the air-pump occasionally gave rise to glowing light (the phenomenon now exploited in strip lighting) and investigation of this led in turn to the discovery of effects caused by static electricity. This soon led to experiments on electricity itself, and to the development of new specialist apparatus for electrical experiments.

These electrical experiments were so spectacular in their effects that it spawned a new profession – the itinerant experimental philosopher, travelling from town to town, drawing ticket-buying crowds to see demonstrations of these electrical effects.

John Wesley (1703–91), the founder of the Methodist movement – a new religious movement which proved extremely popular among many in eighteenth-century England – after seeing one of these shows published a book called *Desideratum: Or, Electricity Made Plain and Useful* (1760). But his idea of the usefulness of electricity does not coincide with ours. He was not imagining artificial lighting, or machines driven by electricity. He was thinking of its usefulness for demonstrating the power of God in nature. He was struck, for instance, by the way electrical sparks could be made to emerge from water (usually by lowering a sword-point towards the surface of water which had been highly charged with static electricity). Everyone assumes that fire and water are opposites, but here was a form of fire (Wesley thought) which comes out of water. Surely, this shows that God can do anything? Similarly, Joseph Priestley (1733–1804), a leading chemist and a Minister in the Unitarian Church, saw electricity (and the air-pump) as a way of demonstrating the power of God in nature.

The phenomenon of electricity also led to a number of speculative philosophical systems, but all of them can be seen to manifest one or another of Newton's two suggestions for explaining all physical phenomena (see Box 14.2). The same two Newtonian suggestions can also be seen to be called upon in the developing theory of electricity itself. The historical development of the science of electricity is highly complicated, and has a huge cast of characters, so what follows is merely to illustrate this general point.

One of the earliest theories of electricity (in keeping with different effects resulting from friction on resinous material and friction on glass, but also conveniently reflecting Newtonian ideas) was that there must be two different kinds of electrical fluid – one responsible for attraction and another for repulsion. But this was simplified by Benjamin Franklin (1706–90), later to make a name for himself as an American statesman. He assumed that all bodies contained electric fluid (of which there was just one kind), and in some cases the fluid could be pumped out of a body, making the pressure of the electrical fluid 'negative', and pumped into another body, giving it 'positive' pressure.

The electrical fluid was supposed to differ from ordinary matter because its particles mutually *repelled* one another (like the particles of Newton's aether), but its particles strongly *attracted* ordinary matter. This did pretty well at explaining many electrical effects, but the big problem was the fact that two *negative* bodies repelled one another. If both had a deficiency of the self-repelling electric fluid, this ought not to happen.

A German natural philosopher by the name of Franz Ulrich Aepinus (1724–1802), who lived and worked in St Petersburg in Russia, overcame this problem by suggesting that in fact particles of ordinary matter were also mutually self-repellent. Ordinary matter only seemed to gravitate (rather than dissipate itself by repulsion) because of the electrical fluid contained in bodies. When bodies were ordinarily saturated with electric fluid the balance between attractive and repulsive forces was slightly towards the attractive, and the bodies 'gravitated' toward one another. When bodies became negatively or positively charged, the electrical effects over-rode what had been previously attributed to gravity, but even gravity itself could now be seen, according to Aepinus, to be electrical in nature.

The result of this kind of speculation – which seemed to suggest that Newton had been misled in regarding gravity as a separate force – were attempts to discover the law describing the variation between force and distance of two electric charges. Aepinus couldn't manage it, but Priestley and others guessed it would follow the inverse square law like gravity. This was eventually confirmed experimentally by John Robison (1739–1805), professor of natural philosophy at Edin-

Box 14.2 Eighteenth-century speculative philosophical systems

Herman Boerhaave (1668–1738), *Institutiones et experimenta chymicae* (1724)

Argued that fire is responsible for expansive activity in nature, and is contained in all matter. So, matter gravitates in accordance with Newton's universal principle, but also contains a repulsive power, endowed by its internal fire.

Stephen Hales (1677–1761), *Vegetable Staticks* (1727)

Acepted Newton's view that the particles of gases must repel one another, and sought to explain chemical processes in terms of interaction between repelling particles of gases and attracting particles of other bodies.

John Rowning (*c.* 1701–71), *Compendious System of Natural Philosophy* (1735)

Argued that particles attract and repel each other alternately at different distances. Used this to explain hardness and softness, cohesion and liquefaction, elasticity, etc.

Gowin Knight (1713–72), *An Attempt to Demonstrate that All the Phenomena in Nature may be Explained by Two Simple Active Principles, Attraction and Repulsion* (1748)

Did exactly what it claimed on the title page.

Roger Joseph Boscovich (1711–87), *A Theory of Natural Philosophy* (1758)

Argued that bodies were not made of matter, but merely of geometrical points which are endowed with inertia, and with the power of attraction up to a certain boundary (close to the geometrical point), when the power of attraction switched to a repulsive power (and results in the sensation of solidity).

Joseph Priestley (1733–1804), *Disquisitions on Matter and Spirit* (1777)

Introduced Boscovich's theory into Britain, and used it to argue that the world is not divided into passive matter and active spirit, but that matter is active with attractive and repulsive forces.

Bryan Higgins (1737–1820), *Experiments and observations relating to... the matter of fire and light... and other subjects of chemical philosophy* (1786)

Argues that particles of matter are hard and globular and surrounded by atmospheres of fire, with the result that the particles can manifest both attractive and repulsive forces.

James Hutton (1726–97), *Dissertation on Different Subjects in Natural Philosophy* (1792)

Argues for two kinds of matter: gravitating matter, which attracts, and solar matter, which repels (the latter being manifested most clearly in light, fire, and electricity). Even suggests that inertia derives from the balance of attractive and repulsive forces in matter.

burgh, and proved mathematically by Henry Cavendish (1731–1810), an English aristocrat and eccentric's eccentric, in 1771.

Nevertheless, all of these thinkers (and others not mentioned) can be seen as trying to explain electricity in terms of attractive and repulsive forces, either operating between particles of all bodies, or between particles of subtle active fluids. The culmination of this tradition can be seen in the work of Michael Faraday (1791–1867), who conceived of the so-called electro-magnetic field. In 1844 he published 'Speculation touching Electric Conduction and the Nature of Matter'. Drawing upon ideas first put forward by Roger Joseph Boscovich (1711–87), a Croatian Jesuit inspired by Newton, and introduced into this country by Priestley, Faraday came close to accepting an idea hinted at by Newton, and developed by Boscovich and Priestley (see Box 14.2). The idea, hinted at in Newton's 'Queries', was that the concept of matter could be replaced entirely by a concept of force. When we touch a piece of wood with a finger we have a sensation of solidity, but that could simply be the result of a force pushing against our finger. The resulting picture of body was one in which attractive forces (due to gravity), at close distances, suddenly flipped over into repulsive forces:

> As in Algebra, [Newton wrote,] where affirmative Quantities vanish and cease, there negative ones begin; so in Mechanicks, where Attraction ceases, there a repulsive Virtue ought to succeed.

Boscovich suggested that particles of matter were simply geometrical points endowed with inertia. These points were held to exhibit a strongly repulsive force at close distance, which at a given distance flipped over into an attractive force (varying in accordance with Newton's universal principle of gravitation).

Faraday is famous for introducing the concept of fields of force into physics, but it was a notion that clearly arose from this Newtonian tradition. 'Particles of matter', Faraday wrote, should be seen as nothing more than 'centres of force', and the forces should be seen as extending indefinitely (albeit diminishing exponentially) from each 'centre'. The result was this:

> The view now stated of the constitution of matter would seem to involve necessarily the conclusion that matter fills all space, or, at least, all space to which gravitation extends (including the Sun and its system); for gravitation is a property of matter dependent upon a certain force, and it is this force which constitutes the matter. In that view matter is not merely mutually penetrable, but each atom extends, so to say, throughout the whole of the solar system, yet always retaining its own centre of force... Hence matter will be continuous throughout

Faraday envisaged his magnetic and electric fields in much the same terms – a force field extending throughout the universe. Some historians of science have assumed that Faraday was rejecting 'Newtonian atomism' here and therefore must have been drawing upon an eighteenth-century anti-Newtonian tradition. This view can only be based on ignorance of, or a failure to understand, Newton's position as put forward in the 'Queries' in the *Opticks*. Properly understood, it seems perfectly clear that the science of electricity as it developed throughout the eighteenth century and into the nineteenth, was a thoroughly Newtonian science.

FURTHER READING

I. B. Cohen, *Franklin and Newton: An Inquiry into Speculative Newtonian Experimental Science and Franklin's Work in Electricity* (Cambridge, MA: Harvard University Press, 1966).

C. C. Gillispie, *The Edge of Objectivity* (Princeton: Princeton University Press, 1960), chapter 5.

T. L. Hankins, *Science and the Enlightenment* (Cambridge: Cambridge University Press, 1985).

John L. Heilbron, *Electricity in the 17th and 18th Centuries* (Berkeley: University of California Press, 1979), chapters 8, 12, 13, and 14, pp. 229–49 and 290–343.

P. M. Heimann, 'Ether and Imponderables', in G. Cantor and M.J.S. Hodge (eds), *The Conception of Ether: Studies in the History of Ether Theories* (Cambridge: Cambridge University Press, 1981), chapter 1, pp. 61–83.

Simon Schaffer, 'Natural Philosophy and Public Spectacle in the 18th Century', *History of Science*, 21 (1983), pp. 1–43.

R. E. Schofield, *Mechanism and Materialism* (Princeton: Princeton University Press, 1970).

Arnold Thackray, *Atoms and Powers: An Essay on Newtonian Matter-Theory and the Development of Chemistry* (Cambridge, MA: Harvard University Press, 1970)

Robinson M. Yost, 'Pondering the Imponderable: John Robison and Magnetic Theory in Britain (c. 1775–1805)', *Annals of Science*, 56 (1999), pp. 143–74.

15

The Chemical Revolution: Priestley and Lavoisier; John Dalton and Beyond

What we would think of as chemical theories and practices, but which were then regarded as alchemy, played a crucially important role in the Scientific Revolution. Although Descartes showed no interest in, and little knowledge of, chemical phenomena, other leading natural philosophers in the development of the new philosophies did. Indeed, knowledge of alchemy was one of the main stimuli to alternative, non-Cartesian, versions of the mechanical philosophy. Those with some knowledge of alchemy found it hard to accept the Cartesian theory of matter as completely passive and inert, moved only by collision with other particles of matter, which were in turn set in motion by collisions with other parts of matter (and so on, *ad infinitum*). For alchemists, matter seemed to interact in ways which seemed beyond the scope of such mechanistic principles.

The most forceful alchemically inspired alternative view of matter, in which particles of matter were assumed to have some intrinsic activity of their own, was of course Isaac Newton's. In the widely admired 'Queries' which he appended to the *Opticks*, Newton used numerous chemical examples to illustrate what he meant by explaining all phenomena in terms of attractive and repulsive forces between particles, and to strongly imply the truth of such explanations.

Newton drew on a medieval tradition in alchemy, which already incorporated particulate theories of matter, and so found it easy to amalgamate this with mechanistic theories. This was not the only alchemical tradition, however, and the historical development of what we now call chemistry was complicated by a recent re-working of ideas deriving from the sixteenth-century alchemical and medical innovator known as Paracelsus (1493–1541), and which was widely adopted by other chemists.

This theory was developed by two German physicians, Johann Joachim Becher (1635–82) and Georg Ernst Stahl (1659–1734), who emphasized the various different characteristics of matter which alchemists had discerned, such as earthiness, acidity, causticity, combustibility, metallicity, and so forth, and declared them to be chemical 'principles'.

The most important aspect of the theory for our purposes is the Stahlian concept of *phlogiston*, the principle of inflammability and combustibility. This concept derives from Aristotelian theories about the nature of fire and from alchemical theories about sulphur, as the principle of combustion. According to both theories, when something ignites or bursts into flame, it shows that there must have been fire, or sulphur, as part of its composition. Heat a piece of wood, and it eventually bursts into flame. Therefore, wood must contain fire, or sulphur. Heat a bar of iron and it does not burst into flame, instead it melts. Therefore, iron does not contain fire, but it does contain the fluid principle of water (if you are an Aristotelian) or of mercury (if you are a Paracelsian).

So, as a substance burns, it gives off fire, or sulphur, or in Stahl's system, *phlogiston*, to the atmosphere. Now, it is important to note, of course, that today we believe that as something burns it is not expelling something which it contains, but rather it is taking something out of the atmosphere - namely what we call *oxygen*.

If we burn something inside an enclosed space, say a bell-jar in a chemistry laboratory, we know that eventually the substance will cease to burn when all the oxygen inside the bell-jar is used up. But, Stahl believed that in this case the substance ceased to burn because it had expelled so much phlogiston out into the bell-jar that the air in the bell-jar had become saturated with phlogiston and could not absorb any more. So, on the one hand we have a theory which assumes exhaustion of oxygen from the atmosphere, and on the other hand we have a theory that assumes saturation of the surrounding atmosphere by phlogiston.

The key to understanding phlogiston theory is that it is often a reversed mirror-image of our current oxygen theory. It might seem odd to us, that Stahl opted for the complete opposite of what we believe – but it was because this was the idea that flowed most naturally from the Aristotelian theory (which in turn had also been the source of the alchemical sulphur theory) – in which fire was an *ingredient* of combustible substances which was seen to escape as the substance burned. Furthermore, in Stahl's day our alternative theory was not available – it was first formulated by Antoine Lavoisier (1743–94), as we shall see in this chapter.

Before going any further we should note the mirror-image accounts for what happens when what was then called a metal calx (often what we would call a metal oxide) is heated. We would now say that what happens is that the metal calx gives off oxygen into the surrounding atmosphere, and leaves behind the metal. So:

$$\text{Metal oxide} \rightarrow \text{Metal + oxygen}$$

According to Stahl's mirror-image theory, what happens is that as we heat the metal calx, phlogiston (supplied either by some adjacent fuel, like charcoal, or even from the air itself, which always contains some phlogiston) is combined with the metal calx to give rise to the metal.

$$\text{Phlogiston + metal oxide} \rightarrow \text{metal}$$

For most metals this is a reversible reaction, and if the metal that resulted from heating the calx is then heated itself, it combines with oxygen in the atmosphere and oxidises back into the metal calx. At least that's what we would say. According to Stahl, the heated metal gives off phlogiston into the atmosphere, and returns back to the elemental form, which is the metal calx.

Now, notice that for Stahl the calx (or oxide) of the metal, is regarded as the elemental form. For us, the metal is the element, and when it is in the form of a calx it is combined with another element, oxygen. It seems completely wrong to us to think of the metal oxide as elemental, but one way of understanding this is to substitute 'natural' for 'elemental' in our description. So, for Stahl the calx of the metal is regarded as the natural form. With a few exceptions, such as gold and lead, metals are more often than not dug up from the earth in the form of a mineral salt, what we call a metal ore. The ore has to be processed to get the pure metal out of it. Given that Stahl is still thinking in terms of *earth* as one of the four Aristotelian elements, for him, whatever we dig up out of the earth is the natural (or elemental) form. The metal is something that is made from the natural earthy form by performing a chemical transmutation on it. It should be noted, therefore, that Stahl's phlogiston theory differs from Aristotelian theory as recounted above, in assuming that metals *do* contain the combustion principle, phlogiston (whereas Aristotle assumed metals *do not* contain fire, or not much of it). Although Stahl was obviously aware that metals do not burst into flames when heated, he needed to invoke phlogiston to explain the difference between metal ores and their metals, and suggested that it provided metals with their lustrous appearance compared to the stony appearance of ores.

Tension between Newtonian ideas and Stahlian ideas fed into what has been called the 'Chemical Revolution', which took place in the second half of the eighteenth century. This tension can be seen, for example, in the famous disagreement between the English Newtonian natural philosopher, Joseph Priestley (1733–1804), and the French chemist, Antoine Lavoisier.

Although Priestley always upheld phlogiston theory, and Lavoisier developed an alternative theory which would depose the Stahlian view, it would be wrong to see it as a dispute simply over phlogiston. Both thinkers were simultaneously indebted to Newtonian ideas, and it was their different resolutions of the tension between Newtonianism and Stahlianism which accounted for their final positions.

We are going to concentrate on Priestley and Lavoisier here, but it is important to note that the Chemical Revolution was not just fought out between these two men. Lavoisier himself was always resentful when he heard talk of the 'French chemists', because he tended to see himself as the single-handed hero of the Revolution. But the fact is, there was a significant group of chemists developing the new French-style chemistry. Lavoisier may have been the most prominent in this group, but he was not alone.

Similarly, it is equally unfair to regard Priestley as the only thinker who refused to accept the new French chemistry, holding out for phlogiston. To begin with there were many like him, although few held out as long as he did. Although in Germany, where Stahl's writ ran high, there were also few converts to Lavoisier's *chimie*. The fact that Lavoisier's chemistry won the day, and forms the basis for modern chemical theory, should not blind us to the fact that there was a lot of implausibility in the theory to begin with, and many aspects of it which seemed irreconcilable with well-known experimental results.

In what follows, nevertheless, we are going to look at the work of Priestley and Lavoisier to see what they were doing, and why both, in spite of the huge differences between them, can be seen as pursuing an essentially Newtonian enterprise.

Joseph Priestley and pneumatic chemistry

Priestley is often referred to today as a chemist, but he saw himself simply as a Baconian natural philosopher, though he was also clearly indebted to Newton. One of his major interests was in the physiology of respiration, and as such he was interested in the nature of air. In particular, he wanted to know how it was that good air could be spoiled or vitiated, so that it would no longer support life, or combustion. Following on from that, he also wondered whether it was possible to restore goodness to vitiated air, so that it would once again support life and fire. This enterprise was in itself inspired by Newton who had speculated in the 'Queries' in the final book of his *Opticks* that the ability to support life may be another active principle inserted into matter, or some matter, by God.

Being a devout believer, Priestley thought that the goodness of air was part of God's benevolence, and that, because we all manage to keep breathing, there must be some natural process by which its goodness is perpetually restored. Pursuing this idea, he eventually discovered that plants can restore the goodness of air (we now say that plants absorb carbon dioxide and expire oxygen). But before that, he included the new phenomenon of electricity in his investigations. One reason for Priestley's interest in electricity, which we saw in Chapter 14, was because he wondered if electricity could restore goodness to vitiated air. Accordingly, he performed experiments involving the creation of electrical sparks in bad air, followed by testing to see whether the air had been changed.

There were other reasons for an interest in air at this time. Newton had also speculated in the 'Queries' about changes of state, from solid to liquid, and in particular, to what he called vapours or fumes (the word 'gas' was to be brought into common use later, by the French chemists). Newton's interest stemmed from the fact that the transition from solid and liquid into a vaporous state suggested a switch from the dominance of attractive forces between particles to repulsive forces between the particles. Newton tentatively suggested that air might be combined with solids or liquids, thereby providing an internal source of mutually self-repelling particles, which might account for solids giving off vapours (the 'vapour' of the solid actually being carried off by escaping air).

These speculations were pursued experimentally by Stephen Hales (1677–1761), who heated many different solids and liquids and was often able to collect large amounts of what he called 'fixed air' – by which he simply meant air that had been 'fixed' in these substances. Hales no doubt produced different gases during these experiments but he was thinking only in terms of atmospheric air and did not subject them to chemical scrutiny. He was content merely to measure the quantities involved.

The Scottish natural philosopher Joseph Black (1728–99), however, noted in 1756 that the 'fixed air' he obtained from heating magnesia alba (magnesium carbonate) was different from common air (it was what we call carbon dioxide). It soon became apparent that there were yet more kinds of air; in 1766, for example, Henry Cavendish discovered 'inflammable air' by collecting the gas given off when sulphuric acid attacks a metal (which is now called hydrogen). In early 1772, Priestley joined in the experimental enterprise of producing and classifying these new kinds of air, and over the next two years discovered 'nitrous air' (nitric oxide), 'marine acid air' (hydrochloric acid gas), 'alkaline air' (ammonia), 'vitriolic acid air' (sulphur dioxide) and 'dephlogisticated nitrous air' (nitrous oxide).

Most significantly for later developments in France, Priestley discovered what we now call (thanks to those developments in France) oxygen. But Priestley isn't yet thinking in terms of different gases – he is still thinking of different varieties of 'air'. But the result of routine tests reveals to him that this particular version of air actually supports life, and combustion, better than common air does. Flames burn longer, brighter and more vigorously in it, and small creatures, such as mice or birds, can live in a confined atmosphere of it for longer than they can in ordinary air.

Now, Priestley was an extremely devout religious believer and believed that his science went hand-in-hand with his religion. So his discovery of this superior kind of air presents him with a problem. Surely God would have given us the best air to breathe? God is good, so why would he fob us off with an inferior product? God's good air must be the best air. It followed, therefore, that what Priestley had discovered was not a natural substance but only something that was artificial – an artificial substance that seemed better than God's good air, but which wasn't natural. Priestley's judgement of the artificial nature of this air was reflected in the name he gave it. He called it dephlogisticated air – air with the phlogiston (artificially) removed from it.

This made sense according to the phlogiston theory. Remember that when something burns it is supposed to give off phlogiston into the surrounding atmosphere – and burning ceases when the surrounding atmosphere is saturated with phlogiston. So, if something burns in dephlogisticated air, it will take longer to saturate that air with phlogiston than it would if it was ordinary air with some phlogiston already in it. So, things burn longer in dephlogisticated air, and likewise little creatures can breathe in it longer (expiring phlogiston). It all makes perfect sense.

Importantly, however, Priestley has established that 'atmospheric air is not an unalterable thing', and 'not an elementary substance but a composition' (specifically, something which normally includes phlogiston in its composition).

Nonetheless, he regards common air as the natural form of air. Returning to his question about why God did not provide dephlogisticated air for us to breathe, Priestley speculated that we would live too fast and die sooner, just as candles burn harder and faster in this kind of air. He concluded that 'the air which nature has provided for us is as good as we deserve'.

Lavoisier and the chemistry of air

Meanwhile, over in France, Antoine Lavoisier was led by lectures on Hales's work by Guillaume François Rouelle (1703–70) to consider whether the air supposedly 'fixed' in a substance was chemically combined with it, or perhaps simply trapped in pores or pockets in the sample. Another possibility, however, was that the fixed air was 'factitious', that is to say, newly produced during the addition of heat to decompose the body. This derived from an earlier theory that fire was the only one of the elements which consisted of mutually self-repellent particles (in keeping, for example, with Boerhaave's views), as opposed to the view that saw both fire and air as self-diffusive.

Following on from this, in about 1766, Lavoisier seemed to have the beginnings of a conception of heating in which fire was driven off, leaving a non-expansive version of air (that is to say, air whose particles were not mutually self-repellent, and which could therefore, easily enter into solid or liquid compositions). Clearly, Lavoisier did not see the full significance of this, but he had a foreshadowing that heating, or combustion, could be seen in terms of the absorption of air, rather than the expulsion of the matter of fire, or phlogiston.

In 1772 Lavoisier was given access to a large burning lens belonging to the Académie des Sciences and used it to repeat Hales's investigations. He also used the lens, however, to test puzzling claims that the products of combustion were, at least in some cases, heavier than the original samples before burning. According to the phlogiston theory, combustion results in the expulsion of phlogiston, and so lead calx, for example, ought to be lighter than metallic lead (which is supposed to be a combination of lead calx and phlogiston). By careful weighing of samples, before and after burning with the lens, Lavoisier was able to establish that combustion *always* resulted in gain of weight. Lavoisier was now convinced of what he had only dimly glimpsed before – that during combustion, phlogiston is expelled from a body but air, or something in the air, replaces the phlogiston, and results in increase in weight.

In 1773 Lavoisier announced at a meeting of the Académie that this would lead to an 'almost complete revolution' in chemistry. It was then that his critics began to pay attention.

Priestley, for example, denied that the diminution of the volume of surrounding air during combustion in a confined space signified absorption of part of the air into the combusted sample. Phlogiston, he insisted, simply had the effect of reducing the elasticity of the air, and so its volume decreased.

Priestley was thinking like a good Newtonian here. Air particles have a tendency to flee from one another, thanks to repulsive forces operating between them. As air becomes vitiated, however, its particles lose some of their repulsive power and they fall closer together, thus reducing the volume they occupy.

Priestley told Lavoisier about his discovery of dephlogisticated air in 1774, and the following year Lavoisier tried it for himself. The experiment exploited the unique fact that simply heating red calx of mercury on its own resulted in the production of metallic mercury, and the expulsion of this newly discovered 'air'. Heating the red calx of mercury with charcoal – the normal experimental procedure for converting a calx to the metal – resulted, as with all other metal calxes, in the production of fixed air (that is, what we now call carbon dioxide). Lavoisier surmised that metal calxes, when heated, gave off dephlogisticated air, but that this always combined with the charcoal to form fixed air. He was then convinced that the new 'air' discovered by Priestley was not merely common air with phlogiston taken out, but was an active chemical agent in itself, capable of combining with metals, and with charcoal, to constitute different substances.

Lavoisier also showed that after strongly heating metallic mercury in a confined volume of atmospheric air, one fifth of the air was used up (as the red calx formed), leaving a gas which would not burn or support life, which he called azote (later called nitrogen). Lavoisier was now convinced that atmospheric air consisted of a mixture of this new gas and azote.

This led Lavoisier to develop not only a new theory of combustion but a new theory of chemistry. When phosphorus combines with this new kind of air it creates phosphoric acid, and subsequent investigations led Lavoisier to believe this new air was the principle of acidity, and in 1779 he began to call it *oxygen* (acid-former). Two or three years later (1781–82) he drew up a brief presentation of his new system in which oxygen was just one of many 'elements', including the known metals, sulphur, phosphorus, water, and various other substances. Lavoisier hereby entirely rejected earlier conceptions of the elements and defined his new elements operationally; that is, for him any substance which could not be broken down further by chemical analysis was considered to be an element. We still think in the same terms today.

Lavoisier's system did not sweep all before it, however. There were a number of problems confronting his interpretations of what was going on in chemical reactions, and it was all too easy to interpret other reactions in a way that was consistent with phlogiston theory.

Consider the fact that the 'inflammable air' discovered by Cavendish in 1766 came to be seen as phlogiston, or almost entirely phlogiston. It was noticed by Priestley that heating lead calx (lead oxide) in an atmosphere of inflammable air resulted in the complete absorption of the inflammable air, and the complete conversion of the calx into metallic lead. This is exactly in accordance with the Stahlian theory:

$$\text{Phlogiston (inflammable air)} + \text{lead calx} \rightarrow \text{lead}$$

Little wonder that Priestley thought he had discovered an experimental proof of phlogiston theory. This was especially persuasive, given that Lavoisier could not at first explain this.

Similarly, Lavoisier could not explain how inflammable air was given off when an acid acts on a metal. According to Lavoisier, sulphuric acid acting on iron ought to look something like this:

$$SO_3 + Fe \rightarrow FeSO_3$$
Acid + iron → iron calx (but no inflammable air)

It was only after the discovery by Henry Cavendish that inflammable air (hydrogen) burned in dephlogisticated air (oxygen) to create water, which Lavoisier heard about in 1783, that Lavoisier's system was vindicated. The combustion of inflammable air could now be seen, as Lavoisier's system demanded, as a combination with oxygen. Furthermore, the production of water in the process now enabled Lavoisier to see that water was not an element but a combination of hydrogen (water-former), as he now called inflammable air, and oxygen. He could now understand also the reduction of lead calx in inflammable air, and the reaction of an acid on a metal.

We now write these reactions like this:

$$4H_2 + Pb_3O_4 \rightarrow 3Pb + 4H_2O$$
inflammable air (hydrogen) + lead calx → lead + water

$$SO_3 + Fe + H_2O \rightarrow FeSO_4 + H_2$$
Acid + iron + water → iron calx + inflammable air (hydrogen)

When Priestley first performed his experiment of heating lead calx in an atmosphere of inflammable air (phlogiston), he performed it over water. As the inflammable air combined with the calx to form metallic lead, the water rose up in the encompassing jar, filling the jar completely. It was the complete absorption of the inflammable air (signified by the water filling the whole space), in accordance with Stahl's theory, that led Priestley to suppose Stahl was right. Clearly, Priestley could not have noticed the formation of water (created by the inflammable air/hydrogen combining with the oxygen in the lead calx). Subsequently, however, Priestley performed the experiment over a bath of mercury, and watched the mercury rise up to fill the space vacated as the inflammable air combined with the calx. On this occasion he did notice the appearance of some water droplets in his apparatus. But he merely regarded this as water which had previously been trapped in the lead calx. Furthermore, he continued to take this line with regard to other relevant experiments. While Lavoisier insisted that Cavendish had shown that water was made of hydrogen and oxygen, for example, Priestley insisted that Cavendish had merely shown that water was a constituent of either inflammable air, or dephlogisticated air, and was left over as the inflammable air burned in the dephlogisticated air. We can now see that Priestley was wrong, but he was not stupidly wrong – he might well have been the one who was interpreting the experimental results correctly.

Similarly, Priestley denied Lavoisier's fundamental claim that oxygen was a defining ingredient of all acids. In this case he was right. So called marine acid could never be persuaded to give up oxygen, and we now know it is a compound of hydrogen (the real principle of acidity) and chlorine.

By 1785, nevertheless , Lavoisier felt able to publicly and uncompromisingly reject the phlogiston theory, and a real revolution in chemistry began as French chemists took one side or another. Younger chemists, notably Claude Louis Berthollet (1748–1822) and Antoine François de Fourcroy (1755–1809), and a number of mathematicians and physicists took sides with Lavoisier.

Why mathematicians and physicists? This brings us to Lavoisier's Newtonian credentials. Lavoisier had been involved in ongoing attempts to establish the so-called 'elective affinities' of different substances.

It was long known, of course, that many chemical reactions take the form of displacement, or swapping of ingredients. So if a compound is mixed with another substance a displacement might take place:

$$AB + C \rightarrow AC + B$$

Similarly, if two compounds are mixed, a swap might take place:

$$AB + CD \rightarrow AC + BD$$

Chemists simply said that A had a greater affinity for C than for B, and so forth. Newton, of course, suggested that the affinities could reveal, and be explained by, attractive forces operating between the particles. In these cases, the attractive force between A and C was greater than that between A and B.

The Newtonian nature of the enterprise can be seen in the comment of the French mathematician, Pierre Louis Moreau de Maupertuis (1698–1759) in 1752:

> Though Astronomers were the first to employ attractions, Chemistry has now recognized the necessity of them, and the most famous chemists today admit attractions and use them no less than astronomers do.

Lavoisier collaborated with the Newtonian physicist, Pierre-Simon Laplace (1749–1827), in the hope of extending what had previously been merely a classifying enterprise (drawing up tables of affinity, listing which substances reacted with others, etc.) into something a bit more theoretical and predictive. As Lavoisier wrote in his 'Memoir on Affinity' of 1785:

> Perhaps one day the precision of the data might be brought to such a perfection that the mathematician in his study would be able to calcu-

> late any phenomenon of chemical combinations in the same way, so to speak, as he calculates the movement of the heavenly bodies.

Lavoisier never succeeded in this, but his new system of chemistry encouraged such attempts, and perhaps for that reason attracted mathematicians to his side.

Furthermore, it was this work on affinities which enabled Lavoisier to recognize dephlogisticated air, not merely as a 'factitious' laboratory creation, but as a natural chemical substance in its own right. A major aspect of trying to establish chemical affinities involved drawing up so-called 'affinity tables'. Each chemical reactant was designated at the top of a column, and below it, in supposed order of affinity, the column was filled with the names of all the reactants known to combine with it. At first, of course, the reacting substances were solids or liquids. After the flurry of new 'airs' were discovered, however, Joseph Black pointed out that the airs would need to be fitted into the tables.

Given his extensive work on affinities, Lavoisier must have been used to thinking of each of the substances at the top of the columns in his tables (and for that matter the substances listed in the columns – each of which, after all, topped their own columns elsewhere in the table) as natural substances in their own right. Accordingly, when he tried to accommodate the different 'airs' in the tables, he had no trouble thinking of them as separate independent substances. Compare this with Priestley's position. Remember, Priestley's major concern was with understanding the nature of God's good air, and as such he had a strong tendency to think of air as elemental in some way, and variations to be just that – merely variations of the one true air, not separate kinds of completely different airs.

In their own ways, therefore, both Priestley and Lavoisier were pursuing Newtonian research programmes. Priestley was trying to understand the nature of atmospheric air in terms of repulsive forces between its particles, and seeing these forces as somehow linked with the principle of life and combustion. Lavoisier, by contrast, was trying to understand all chemical reagents in terms of reactive affinities, due to attractive and repulsive forces between the particles. The trajectories of their careers, and the work they produced, were very different and led them to very different conclusions, and yet both their trajectories can be seen to have been moving through the Newtonian space of Enlightenment science.

There is one other important aspect of Lavoisier's success which must be mentioned. Arguably Lavoisier's eventual triumph depended upon his collaboration in 1787 with Louis Bernard Guyton de Morveau (1737–1816), with a view to reforming chemical nomenclature. Guyton had been calling for reform for a long time and recognized the opportunity which Lavoisier's new system provided for a new nomenclature contrived to simultaneously indicate and endorse the nature of that system. In its essentials, this is the nomenclature still in use today. Operationally defined elements were given simple names and compounds were given binomial designations indicating their constituents: lead calx became lead oxide. Other terms were used to indicate the amount of oxygen in an acid, or a salt: sulphurous acid had less oxygen than sulphuric acid and their salts were sulphites or sulphates respectively. Bearing in mind that this was a vast improvement upon the completely unsystematic way that chemical substances were named previously (often based on appearances, such as butter of antimony, star regulus of antimony, or flowers of sulphur) it is hardly surprising that younger chemists began to adopt the new nomenclature. As they did so, the triumph of Lavoisier's chemistry was assured. As Joseph Black pointed out, using Lavoisier's nomenclature was tantamount to using Lavoisier's system.

Return to atomism: John Dalton

Lavoisier's *Elementary Treatise of Chemistry* (*Traité élémentaire de chimie*, 1789) set the seal on his, and his collaborators' achievements, but when Lavoisier was beheaded (as a so-called tax farmer – the French Crown would save itself the trouble of collecting taxes in a particular region by selling the right to collect the taxes to a local entrepreneur; Lavoisier was one such opportunist) during the 'Terror' which followed the French Revolution – the Chemical Revolution was still not completed.

In particular, the Newtonian dream, of establishing just how particles of matter combine with one another, displace, and are displaced by other substances, and so on, was still absorbing a lot of intellectual energy. Chemists began to focus their attention on the only quantifiable thing they had to work with – weight.

This had been part of the chemical tradition since Joseph Black started using the chemical balance to weigh the ingredients of chemical reactions, before and after the reactions. The technique was made famous by Lavoisier, who used it to defend his interpretation of the experiment in which oxygen or dephlogisticated air (depending on your point of view) had been discovered. But the trouble is, weights of reacting substances don't necessarily provide clear answers to the way they interact. Consider, for example, the controversy which flared up between Joseph Louis Proust (1754–1826) and Claude Louis Berthollet (1748–1822).

Proust insisted that chemical substances always interact with one another in fixed proportions (e.g. the ingredients in common salt are always in a 1:1 ratio – one part sodium to one part chlorine – $NaCl$). Berthollet, on the other hand, believed things could combine in an indefinite range of ratios – even to the extent of sometimes combining in such a way that the chemical properties of the compound vary.

Part of the problem here was that Proust based his conclusions on experimental analysis of dry compounds, whereas Berthollet relied on measurements of substances in solution. Berthollet was perfectly right, of course, in insisting that you can form a solution of salt in water, say, in a wide range of proportions between the salt and the water.

So, did Berthollet have conclusive proof that what Proust called his 'Law of Constant Proportions' was invalid? Or, was Proust correct in concluding that solutions are not proper chemical compounds, precisely because they do not obey the law of constant proportion? Proust's law was more in keeping with the Newtonian view of atomic particles having particular affinities, or powers of attraction, and so eventually Proust won the field, and Berthollet became one of history's losers.

The focus on weights of interacting substances also played an important part in the next stage in the development of chemistry, which entailed the re-establishment of Newton's atoms at the centre of the stage.

The operationally defined elements of the French chemists had led them away from atomism. Regarding as elements anything that could not be further broken down in chemical reactions, Lavoisier drew up a long list of elements including hydrogen, oxygen, nitrogen, the metals, and many more.

Now, this did not sit square with the traditional view of atomism. The basic assumption of atomism had always been that all atoms were made of the same matter. Coal, salt, strawberries, iron, and sulphur, differed only in so far as their constituent atoms had different shapes and sizes, and were arranged in a specific characteristic way. The matter of which each individual atom was composed, however, was always considered to be exactly the same (the Platonic tradition was an exception – Plato believed there were different atoms of earth, water, air and fire; but this never really caught on – see Chapter 2).

If Lavoisier's different elements were all composed of the same kind of atoms, surely the atoms could be separated and recombined to turn what was oxygen, say, into carbon. But this could not be achieved by any chemical manipulations (which is precisely why Lavoisier insisted his elements were elements). Lavoisier was committed to the view, therefore, that oxygen was categorically different from carbon, and from hydrogen, etc. They could not simply be variations based on different arrangements of atoms, because if they were it would be possible to convert one into another. The Lavoisierian conclusion was that elements did not share atoms in common, they were completely different from one another.

Lavoisier did not explicitly deny the existence of atoms – he was too much of a Newtonian ever to do that – but he simply assumed that chemistry was in a state of development which made talk of atoms premature and misleading. At the present state of knowledge, Lavoisier seemed to imply, we must put aside discussion of atoms (which must remain hypothetical entities) and concentrate on what we can know by experiment. Chemical experimenting leads us to believe that there are a significant number of different substances which cannot be analysed or broken down further and so must be assumed (for the time being at least), to be elemental.

This is where the Manchester-based Quaker, John Dalton (1766–1844) comes into the picture. Putting it simply, Dalton combined Lavoisier's multi-element theory with atomism by suggesting that there must be as many different kinds of atom as there are elements in Lavoisier's scheme. So, instead of there being only one kind of atom, which combines in different ways to make oxygen, nitrogen, etc., there are atoms of oxygen, hydrogen, etc., which are all different from one another.

Although this might seem blindingly obvious to us (since it is effectively the way we are taught chemistry in school), it was in fact a radical move – previously, ever since the Ancient Greeks, atoms were all made of the same substance; now some were made of oxygen, and others were made of phosphorus, and so on. So, how was it that Dalton was able to make this radical move when nobody else could see it? In fact, it was an idea that arose naturally out of Dalton's Newtonian inspired attempts to understand the nature of atmospheric air.

As Lavoisier's theory of oxygen began to replace phlogiston theory, it became apparent that only part of the atmosphere was oxygen – air became unbreathable, and incapable of supporting combustion, long before it was entirely used up. Lavoisier had been led by his own experiments to conclude that air was composed of oxygen and nitrogen. Subsequent investigations revealed that air also contained carbon dioxide and water vapour.

The question now arose as to whether these separate ingredients were simply mixed together, jumbled up with one another, or whether they were combined together in a chemical compound. The ingredients were duly weighed to see whether the atmosphere conformed to Proust's law. Given that the constituent gases and water vapour all had different weights, and yet they had not separated out into layers, there was a tendency to believe that they were all combined in a chemical compound.

But Dalton, who had made a study of the formation of dew, was convinced that water vapour was not chemically locked up in the atmosphere, but was merely mixed in with the other ingredients. Dalton began to doubt the chemical compound theory, therefore, but this meant that he had to confront the problem as to why the different gases in the atmosphere did not separate into layers in accordance with their different weights.

Dalton came up with a solution to this problem which should be recognized immediately as thoroughly Newtonian in its assumptions. Dalton said that the particles or atoms of one gas repel one another, as Newton believed, and so the gas spreads itself out to fill any available space. However, particles of one gas do not interact at all with particles of another gas (unless, of course, there is a chemical reaction between them). Dalton wrote:

> When two elastic fluids, denoted by A and B are mixed together, there is no mutual repulsion between their particles; the particles of A do not repel those of B as they do one another. Consequently, the pressure or whole weight upon any one particle arises solely from those of its own kind.

Now, it seems to follow from this assumption that particles of one gas are qualitatively different from those of another gas. If A and B were made of atoms of the same substance they ought to repel one another as much as they repel themselves. So, Dalton already found himself implicitly invoking different kinds of atoms. Atoms of A must be qualitatively different from atoms of B. It was an easy step from here to assuming that all atoms of different elements are different from one another.

Having arrived at this new atomistic interpretation of Lavoisier's elements, Dalton speculated upon how they might combine with one another in chemical combination. Once again, he clung to Newtonian assumptions about attractive and repulsive forces operating between particles. When a single atom of one element interacts with a single atom of another element, the two atoms attract. But what happens in the case of a single atom of one element combining with two atoms of another element? The problem here is that the two atoms which seek to combine with the original atom will repel one another.

Dalton is now led to picture his combining atoms in physical space. Clearly, a chemical compound which we might write as A_2B would be arranged physically in such a way that the two atoms of A were as far as possible from one another, and so would be on directly opposite sides of B. Similarly, a compound which we would write as DE_3, would be arranged to maximize the distance between the mutually repellent particles of E – if D is a sphere, the particles of E would be arranged around it, separated by 120° between them.

Because of the repulsive forces assumed to exist between atoms of the same kind, Dalton concluded that chemical compounds which involve three or more particles of the same element combining with one particle of another were much less likely than simpler compounds (involving only two or three particles in combination). The more mutually repulsive particles were involved, the harder it would be for the attractive force to keep them in combination.

This led Dalton to a set of rules about combinations. These unfortunately led in some cases to false conclusions. He knew only one combination of hydrogen and oxygen, and concluded therefore that the form of the combination must be the simplest: one hydrogen atom and one oxygen atom (HO as opposed to H_2O). Similarly, by the same reasoning he assumed ammonia must be one atom of nitrogen and one of hydrogen (NH instead of NH_3). Using the experimentally determined relative weights of each ingredient in these compounds to enable him to arrive at a so-called atomic weight (the comparative weight of a single atom of each element), led to numerous inconsistencies, and confusions. It took a long time for subsequent generations of chemists to sort this out.

Meanwhile, one of the leading chemists in Britain, Humphry Davy (1778–1829), objected to the multiplication of elements in Lavoisier's, and therefore Dalton's, theorizing. Thanks to new analytic techniques (chiefly using the newly discovered current electricity, as opposed to the static electricity which had only been known before) Davy had discovered that a number of substances which Lavoisier had assumed to be elements were not elements at all (it was surely inevitable that Lavoisier's definition of element as something that had never been analysed into anything else was bound to run up against new analytic techniques one day).

This in turn inspired Davy to object to the idea of God using over thirty (the number of Lavoisier's elements) different kinds of building block to build the world, when traditional atomism showed how it could be done with just one building block. Davy said that he wanted to reinstate

> that sublime idea of the ancient philosophers, which has been sanctioned by the approbation of Newton... namely that there is only one species of matter, the different chemical as well as mechanical forms of which are owing to the different arrangement of its particles.

It is ironic, of course, that Dalton, who undoubtedly saw himself as a thoroughgoing Newtonian, should be criticized for deviating from Newtonian atomism. In one way, of course, Davy was right – Newton never dreamed of Daltonian qualitatively different atoms; but seen from another perspective, emphasizing the Newtonian notion of attractive and repulsive forces between particles, Dalton was as much a Newtonian as Davy. The same was also true, of course, of Priestley and Lavoisier, each very different, but both in their own way Newtonians. Newtonianism, like Christianity, could and did mean different things to different thinkers. It was all part of the process of creating out of Newton the alchemist, the Newton who was the embodiment of the Age of Reason and the beacon of the Enlightenment.

Davy's Newtonianism was also apparent, incidentally, in his theory of how the Voltaic pile worked. The Voltaic pile was the name given to the first means of generating current electricity – based on the pile of alternating discs of zinc and silver (separated by card soaked in seawater) which Alessandro Volta (1745–1827) used in his first electric battery. According to Davy:

> Is not what has been called chemical affinity merely the union or coalescence of particles in naturally opposite states? And are not chemical attractions of particles and electrical attractions of masses owing to one property and governed by one simple law?

As it turned out, the apparently very different views of Dalton and Davy, on the nature and number of the elements, were easily reconciled with one another in what became known as Prout's hypothesis.

In his efforts to determine relative atomic weights, Dalton had designated the weight of hydrogen as one, and all other relative weights had been found to be close to other integers. William Prout (1785–1850), an Englishman who trained as a physician at Edinburgh University, suggested in 1815 that Dalton's seemingly different atoms could all be made up of different combinations of hydrogen atoms. So, there was only one kind of building block after all – the hydrogen atom; but this combined in different numbers to make each of the Daltonian (or Lavoisierian) atoms. This hypothesis proved highly fruitful in stimulating further research. It led to improved techniques of determining atomic weights (in order to test the hypothesis), and it gave impetus to the search for a system of classification of the elements which eventually led to the Periodic Law and the Periodic

Table of the elements. It also pointed the way to attempts to determine the structure of the atom.

The subsequent development of chemistry is, needless to say, highly complex, and would require a proper discussion of profoundly technical details to do it justice. We will leave it, therefore, at this point. But, it should be clear that when chemistry emerged as effectively a new science in the eighteenth century – significantly different from the alchemy that had gone before – that new science was a Newtonian science.

FURTHER READING

Bernadette Bensaude-Vincent, 'Lavoisier: A Scientific Revolution', in M. Serres (ed.), *A History of Scientific Thought: Elements of a History of Science* (Oxford: Blackwell, 1995), pp. 455–82.

William H. Brock, *The Fontana History of Chemistry* (London: Fontana, 1992), chapter 3, pp. 87–127 and chapter 4, pp. 128–72.

J. B. Conant, 'The Overthrow of the Phlogiston Theory: The Chemical Revolution 1775–1789', in Conant and Nash, *Harvard Case Histories in Experimental Science* (Cambridge, MA: Harvard University Press, 1957), Vol. I, Case 2.

Jan Golinski, 'Chemistry', in Roy Porter (ed.), *The Cambridge History of Science, Volume 4: Eighteenth-Century Science* (Cambridge: Cambridge University Press, 2003), pp. 375–96.

Henry Guerlac, 'The Background to Dalton's Atomic Theory', in H. Guerlac, *Essays and Papers in the History of Modern Science* (Baltimore: Johns Hopkins University Press, 1977), chapter 14, pp. 217–42.

D. P. Mellor, *The Evolution of the Atomic Theory* (Amsterdam: Elsevier, 1971), chapter 4, pp. 44–64.

C. E. Perrin, 'The Chemical Revolution', in R.C. Olby *et al.* (eds), *Companion to the History of Modern Science* (London: Routledge, 1990), chapter 17, pp. 264–77.

Isabelle Stengers, 'Ambiguous Affinity: The Newtonian Dream of Chemistry in the Enlightenment', in M. Serres (ed.), *A History of Scientific Thought: Elements of a History of Science* (Oxford: Blackwell, 1995), pp. 372–400.

Arnold Thackray, 'The Origin of Dalton's Chemical Atomic Theory: Daltonian Doubts Resolved', *Isis*, 57 (1966), pp. 35–55.

A. Thackray, *Atoms and Powers: An Essay on Newtonian Matter Theory and the Development of Chemistry* (Cambridge, MA: Harvard University Press, 1970), chapter 8, sections 4–6.

Stephen Toulmin, 'Crucial Experiments: Priestley and Lavoisier', *Journal of the History of Ideas*, 18 (1975), pp. 205–20.

16

Newtonian Optimism: Natural Theology and Natural Order

Over the next few chapters we shall be looking at the background from which Darwin's theory of evolution emerged. You might suppose that this will finally take us away from the Newtonian tradition, which we've been looking at in the past few chapters. If so, you would be wrong. Eighteenth-century Newtonianism was even a factor in the background to Darwinism – indeed, it could be said to be one of the most important aspects of that background.

I mentioned in Chapter 14 that, in Britain in the eighteenth century, Newtonian natural philosophy was used by leading thinkers to prove the existence of God and to support sound religion. The idea was that the intricate design of the universe, revealed by Newton's work, could be used to prove the existence of a supremely intelligent designer. This was the heyday of the tradition of what was called natural theology.

This enterprise chimed so well with attitudes that were in the ascendancy in the Anglican Church that there seemed to be a 'Holy Alliance' between Anglican rational theology and Newtonianism. Newton's authority, for some leading Churchmen at least, seemed almost to replace the authority of the Bible. So, at the end of the seventeenth century the Archbishop of Canterbury, John Tillotson (1630–94), could say in one of his sermons:

> Nothing ought to be received as a revelation from God which plainly contradicts the principles of natural religion... Nothing ought to be received as a divine doctrine or revelation without good proof that it is so.

Like-minded Anglicans, immediately saw that Newton's natural philosophy could be appropriated to support such claims on behalf of a rational, natural religion. Samuel Clarke (1675–1729), for example, close friend and follower of Newton, but also personal chaplain to Queen Anne, argued that morality could be established with the same certainty as Euclid's geometry:

> There is no congruity or proportion in the uniform disposition and correspondent order of any bodies or magnitudes, no fitness and agreement in the application of similar and equal geometrical figures one to another, so plain as the fitness of God's receiving honour from

> his creatures... To deny either in word or action that I should do for another what he in the like case should do for me is as if a man should contend that, though two and three are equal to five, yet five are not equal to two and three.

If religious thinkers emphasized this, it was even more prominent among lay thinkers. Philosophers like John Locke (1632–1704), David Hartley (1705–57), and David Hume (1711–76), among others, all expressed a wish to establish the principles of morality as certainly as Newton had established the principles of mechanics. This very pronounced movement in eighteenth-century British thought has been referred to by historians as 'Moral Newtonianism'.

As the empiricist philosopher, John Locke (1632–1704) wrote in his opinion-forming *Essay Concerning Human Understanding* (1690):

> Upon this ground it is that I am bold to think that morality is capable of demonstration, as well as mathematics since the precise real essence of the things moral words stand for may be perfectly known, and so the congruity and incongruity of the things themselves be certainly discovered in which consists perfect knowledge.

Similarly, the moral philosopher David Hartley paraphrased Newton in his *Observations on Man* (1749):

> The proper method of philosophising seems to be, to discover and establish general laws of action, affecting the subject under consideration, from certain select, well defined and well attested phenomena, and then to explain and predict the other phenomena by these laws. This is the method of analysis and synthesis recommended and followed by Sir Isaac Newton... It is of the utmost consequence to morality and religion that the Affections and Passions (feeling and emotion) should be analysed into their simple compounding parts, by reversing the steps of the Associations which concur to form them. For thus we may learn how to cherish and improve good ones, check and root out such as are mischievous and immoral.

David Hume described his *Treatise of Human Nature* (1739) as 'an attempt to introduce the Experimental method of Reasoning into Moral Subjects'. He was one of the first to develop the psychological concept of the 'association of ideas' (the way in which one idea can, or indeed must, lead to other ideas) which he saw as 'a kind of attraction, which in the mental world will be found to have as extraordinary effects as in the natural, and to show itself in as many and as various forms'.

Other moral philosophers developed, by analogy with Newtonian attraction and repulsion, the notion that all human behaviour could be analysed in terms of 'seeking pleasure' (attractive) and 'avoiding pain' (repulsive). In these views, any tendency towards excessive selfishness would be tempered, it was thought, by the 'common sense' or 'rational', or sometimes even 'experimental', realization that cooperation with one's fellows was required to optimize one's own chances of gaining pleasure and avoiding pain.

This can be seen, for example, in Alexander Pope's *Essay on Man* (1732):

> Two Principles in human nature reign;
> Self-love to urge, and Reason to restrain;
> Self-love, the spring of motion, acts the soul;
> Reason's comparing balance rules the whole.
> The same Self-love, in all, becomes the cause
> Of what restrains him, Government and Laws.
> For, what one likes if others like as well,
> What serves one will, when many wills rebel?
> How shall we keep, what, sleeping or awake,
> A weaker may surprise, a stronger take?
> His safety must his liberty restrain:
> All join to guard what each desires to gain.
> Forc'd into virtue thus by Self-defence,
> Ev'n Kings learn'd justice and benevolence:
> Self-love forsook the path it first pursu'd,
> And found the private in the public good.
> So two consistent motions act the Soul;
> And one regards itself, and one the Whole.
> Thus God and Nature link'd the gen'ral frame,
> And bade Self-love and Social be the same.

In other words, self-love and reason work together for the smooth running of society. The atheist David Hume said that egoism and social responsibility cannot be contradictory for we are all members of society; but the Bishop of Bristol, Joseph Butler, shared the same sentiment: 'Self-love in its due degree is as just and morally good as any affection whatever'.

Bernard Mandeville (1670–1733) in his *Fable of the Bees* (1714) had satirized this moral philosophy by arguing that all virtues are merely the apparent observable effects of what are really extremely selfish causes. But the point of the *Fable* is that, no matter what the underlying cause might be, these private vices do give rise to public benefits. So, once again, the smooth running of society, and the cooperation of everyone on which that smooth running is based, actually depends upon people trying to maximize their own enjoyment, while simultaneously avoiding pain.

Newton's suggestion at the end of the *Opticks* that 'if natural Philosophy in all its Parts, by pursuing this method shall at length be perfected, the Bounds of Moral Philosophy will be also enlarged' even led to what has been called moral calculus. The assumption was not just that morality could be established on rational and experimental grounds but that it could be analysed mathematically. For David Hume, for example,

> The moral evil or vice of a given action is as the degree of misery and number of sufferers; so that, that Action is best which accomplishes the Greatest Happiness for the Greatest Number.

This in turn was used to support political theory in the eighteenth century. Consider Joseph Priestley writing in 1768:

> The 'grand criterium' for settling all political questions should be: 'the good and the happiness of the members, that is to say of the majority of the members, of a State'. For, this one general idea, suitably followed up, throws the greatest possible light on the whole system of politics and of morals, and of theology.

The same thinking was advocated by William Paley (1743–1805), whose philosophical works were taught at Oxbridge well into the nineteenth century.

Now, this kind of new moral philosophy might have been used for reform (certainly Priestley wanted to use it that way), and arguably this kind of rationally based morality was a major force for reform in pre-Revolutionary France. In France, ideas of natural morality placed emphasis upon the freedom of the individual, the natural rights of man, and democracy, and the notion that man, by use of his reason, could progress in moral goodness as much as he could progress in scientific knowledge.

In England, however, a deeply conservative nation, these ideas of a natural morality were used to defend the *status quo* – the way things were, are now, and ever shall be...

This was particularly apparent in eighteenth-century British attempts to deal with the so-called 'problem of evil'. The problem was, why is there suffering in the world? If God is good, and God created the world, why did he allow there to be so much suffering? Why didn't God create a perfect world, with no suffering?

The usual gambit to solve this problem, from the days of the Early Fathers, was simply to say that we cannot understand the working of God's mind – 'God moves in mysterious ways, his wonders to perform'.

But, this gambit cannot be allowed if you are assuming, as eighteenth-century thinkers did, that God made the world according to a rational plan, and that God's plan has been discovered by Newton using his own reason. It would be incompatible with these assumptions to say that we don't understand the working of God's mind. The whole point of Newton's discoveries, and the natural theology being erected upon them, is that Newton *did* understood God's mind, and demonstrated it to the rest of us using mathematical proofs.

Therefore, eighteenth-century thinkers developed a new way of answering the problem of evil. Their reasoning went something like this:

- God is supremely and perfectly good – by definition he is incapable of doing evil.
- God has created the universe, as Newton has shown, by following the laws of reason, logic, mathematics, etc.
- These two assumptions can't be incompatible with one another. If there was any incompatibility God would know it in advance (he is omniscient), and because of His goodness would have chosen not to create the world according to reason.
- Furthermore, if the laws of reason made it possible for God to make more than one kind of world, he must, because of his goodness, have chosen to make the best of all possible worlds.
- Therefore, this is the best of all possible worlds.

The answer to the problem of evil is clear – we might think there's a lot of unnecessary suffering in the world, but it must be necessary to have this amount of suffering. God must have created the best possible world, so any other world presumably would have had even more suffering.

This idea was mercilessly and brilliantly satirized by Voltaire in his short novel, *Candide* (1759), although Voltaire associated it with the German philosopher and great rival to Newton, G. W. Leibniz. Indeed, it is important to note that there is no evidence to suggest that Newton himself would have subscribed to this kind of thinking. But we are dealing here, remember, not with Newton as he really was, but as he was reconstructed by Enlightenment thinkers.

The crucial political point for most conservative British thinkers was that, if this is the best of all possible worlds, *there is absolutely no need to change it*. All the evil and suffering of the world is part of God's plan, and must serve some benefit for the whole system – remember the moral calculus again (see Box 16.1).

Paul Hazard, a distinguished intellectual historian once referred to Alexander Pope's *Essay on Man* (1733) as the most famous 'declaration of faith' of 'this new version of Christianity'. Pope hits the nail right on the head when he writes that what seems like evil is in fact for universal good, and that 'Whatever is, is right':

> For me, health gushes from a thousand springs;
> Seas roll to waft me, suns to light me rise;
> My foot-stool earth, my canopy the skies.
> But errs not Nature from this gracious end,
> From burning suns when livid deaths descend,
> When earthquakes swallow, or when tempests sweep
> Towns to one grave, whole nations to the deep?
> No ('tis reply'd), the first Almighty Cause
> Acts not by partial, but by gen'ral laws;
> Th' exceptions few; some change since all began:
> ...
> The gen'ral ORDER, since the whole began,
> Is kept in Nature, and is kept in Man.

Cease then, nor ORDER Imperfections name:
Our proper bliss depends on what we blame.
...
All nature is but Art, unknown to thee;
All Chance, Direction, which thou canst not see;
All Discord, Harmony not understood;
All partial Evil, universal Good:
And, spite of Pride, in erring Reason's spite,
One truth is clear, WHATEVER IS, IS RIGHT.

Box 16.1 THE OPTIMISTIC PHILOSOPHY OF DR PANGLOSS IN VOLTAIRE'S *CANDIDE*

Pangloss gave instruction in metaphysico-theologico-cosmo-looney-gology. He proved admirably that there cannot possibly be an effect without a cause and that in this best of all possible worlds the Baron's castle was the best of all castles and his wife the best of all possible Baronesses.

It is clear, said he, that things cannot be otherwise than they are, for since everything is made to serve an end, everything necessarily serves the best end. Observe: noses were made to support spectacles, hence we have spectacles. Legs, as anyone can plainly see, were made to be breached, and so we have breeches. Stones were made to be shaped and to build castles with; thus My Lord has a fine castle, for the greatest Baron in the province should have the finest house; and since pigs were made to be eaten, we eat pork all the year round. Consequently, those who say everything is well are uttering mere stupidities; they should say everything is for the best. (Voltaire, *Candide,* 1759)

Some similar opinions to those of Pangloss in 'serious' writers:

Horse dung smells sweet because God knew that men would often be in its vicinity. (George Cheyne, 1705)

Horseflies were created so that men should exercise their wits and industry to guard themselves against them. (William Byrd, 1728)

The louse was invented to promote cleanliness. (Reverend William Kirkby, 1732)

Apes and parrots were created for men's mirth, and singing birds on purpose to entertain and delight mankind. (William Byrd, 1728)

That the chicken was created for the benefit of mankind is evident from the fact that it shows perfect contentment in a state of confinement. (William Swainson, 1754)

Cattle and sheep were given life in order to keep their meat fresh until we shall have need to eat them. (William Swainson, 1754)

So, just as God has established the physical world with the laws of motion so that we see the self-regulating universe in which we live; likewise, he has established human societies with laws of social interaction (analogous to the natural laws of attraction and repulsion) which give rise to the most harmonious social order.

The eventual result of all this was a social and political philosophy based on the fundamental principle known as *laissez-faire*. This is a principle of non interference; an injunction to leave well (and not so well) alone, on the assumption that the system is self-regulating, and any attempt to improve it will only result in making it worse – perhaps disastrously so.

The seminal text of *laissez-faire* 'political economy' as it was then called, was *The Wealth of Nations* (1776) by Adam Smith (1723–90), a leading light of the so-called Scottish Enlightenment. This argued that the most prosperous society would result from the free enterprise of individuals (pursuing self-love, tempered by reason), and that state intervention is unnatural, and because of that, likely to introduce discord into the system. It is important to leave things as they are, Smith believed, because God has made human nature the way it is to ensure that human society works as smoothly as the Newtonian cosmos (which God also created, of course).

One of the most important thinkers in this tradition of political economy, as far as the history of science is concerned, was an Anglican clergyman by the name of Thomas Malthus (1766–1834). Malthus published an *Essay on the Principle of Population* in 1798. It was written as a deliberate warning to the Prime Minister, William Pitt (1759–1806), not to go ahead with his plan to reform the Elizabethan Poor Law. Famines in Ireland meant that Irish were coming over to Britain to seek work but they were literally dying in the streets. Since poor relief in England was still based on statues introduced when Elizabeth was on the throne (1558–1603) and confined to members of Church parishes, any incomer who was not an accepted member of a parish was ineligible for aid. Pitt was intending to change this by reforming the Poor Law.

Malthus claimed, however, that human populations increase in a geometrical ratio (2, 4, 8, 16, 32...), while the amount of food available to feed them could at best only be increased in an arithmetical ratio (2, 4, 6, 8, 10...). The difficulty of acquiring food, Malthus concluded, must be part of God's plan for keeping the population in check. New poor laws would only disrupt the balance of nature, allow more to survive now, and to propagate themselves. Consequently, the population of the poor would increase and, when the resources of even the new poor law couldn't cope, there would be even more suffering and death than at present. The shortage of food, Malthus wrote, is an 'all pervading law of nature... and man cannot by any efforts of reason escape from it'.

Now, it may seem to us that all this, from Pope to Paley, via Adam Smith and Thomas Malthus was just a cynical exercise in justifying the exploitative capitalist, free-enterprise political economy which served their particular social class so well. Certainly, this is the way Karl Marx saw it while he was living in London. For him, this way of thinking amounted to a conspiracy of the bourgeoisie against the working classes, and he urged the lower orders to rise up against them – 'The proletarians have nothing to lose but their chains', he and Friedrich Engels wrote in the *Manifesto of the Communist Party* (1848).

But, of course, Pope, Smith, Malthus, and the rest did not see themselves as conspirators against the proleteriat. They seem to have genuinely believed that their system of *laissez-faire* political economy was the equivalent in the socio-political sphere of the system of Newton in physics. In just the same way that Newton had discovered objective truths about the physical world, they believed they had discovered the natural laws which ensured the smooth running of society. They did not see themselves as imposing a subjective moral and political system, but as revealing the political system established by God, and maintained by objective moral and political laws.

Let us turn, finally, to Charles Darwin. On his return from the voyage on HMS *Beagle* in 1836, Darwin set to work trying to solve the problem of the origin of all the diverse species of plants and animals in the world. Evidence from the fossil record suggested that species were not all created at once but appeared successively on the Earth. The problem was, where had newly arrived species come from? So, when Darwin began to work on this problem he started by looking at the facts:

> I worked on true Baconian principles, and without any theory collected facts on a wholesale scale, more especially with respect to domesticated productions, by printed enquiries, by conversation with skilful breeders and gardeners, and by extensive reading.

He continued working in this Baconian, fact-gathering way for over a year. Then:

> In October 1838, that is, 15 months after I had begun my systematic enquiry, I happened to read for amusement, Malthus on *Population*, and being well prepared to appreciate the struggle for existence which everywhere goes on from long-continued observation of the habit of animals and plants, it at once struck me that under these circumstances favourable variations would tend to be preserved, and unfavourable ones to be destroyed. The result of this would be the formation of new species. Here, then, I had at last got a theory by which to work.

So, Darwin's principle of natural selection and the idea of survival of the fittest were directly inspired by this tradition of natural theology in which all was for the best in the best of all possible worlds, and 'nature red in tooth and claw' was part of God's plan. Although this idea cannot be said to have come directly from Newton himself, it was certainly developed by men who saw themselves as Newtonians – developing Newtonian ideas in the sphere of political economy. It should also be noted that this way of thinking about the natural economy was not the preserve of atheists – far from it – it developed among the devout, and even among churchmen themselves. Malthus and Paley, both Anglican clergymen, saw these ideas as proving the supreme intelligence of the Creator as persuasively as the universal principle of gravitation.

FURTHER READING

P. J. Bowler, 'Malthus, Darwin and the Concept of Struggle', *Journal of History of Ideas*, 37 (1976), pp. 631–50.

Peter Bowler, *Evolution: The History of an Idea*, 3rd edn (Berkeley and Los Angeles, University of California Press, 2003).

B. G. Gale, 'Darwin and the Concept of a Struggle for Existence: a Study in the Extrascientific Origins of Scientific Ideas', *Isis*, 63 (1972), pp. 321–44.

C. J. Glacken, *Traces on the Rhodian Shore* (Berkeley: University of California Press,1973), chapters 8 and 11.

C. Raven, *John Ray* (Cambridge: Cambridge University Press, 1950), chapter 17.

D. R. Oldroyd, *Darwinian Impacts* (Milton Keynes: Open University Press, 1983), chapter 5, pp. 61–72.

D. Ospovat, 'Darwin after Malthus', *Journal of the History of Biology*, 12 (1979), pp. 211–30.

Emma Spary, 'Political, Natural and Bodily Economies', in N. Jardine *et al.* (eds), *Cultures of Natural History* (Cambridge: Cambridge University Press, 1996), pp. 178–96.

Basil Willey, *The Eighteenth-Century Background* (Harmondsworth: Penguin, 1962), chapter 3, pp. 47–59.

R. M. Young, 'Malthus and the Evolutionists: the Common Context of Biological and Social Theory', *Past and Present*, 43 (1969), pp. 109–45; also in R. M. Young, *Darwin's Metaphor* (Cambridge: Cambridge University Press, 1985).

17

The Making of Geology: From James Hutton to Charles Lyell

This chapter looks at a new science, unknown before the seventeenth century, which owes its origins to the mechanical philosophy and especially to the Newtonian version of the mechanical philosophy. This is the study of the Earth itself as a natural object – the science of geology.

Before the seventeenth century the Earth itself was simply taken for granted and barely considered as a topic for discussion within natural philosophy. For Aristotle the Earth had always existed, pretty much in the form it is in now (or as it was in Aristotle's day). In the Christian tradition, the Earth was held to have been created by God but again it was generally assumed it had been created as it is now and had not changed significantly.

Insofar as there was any theorizing or speculating at all about the Earth it was usually regarded as a massive organism, in which there were subterranean processes analogous to life processes. Mother Earth was seen as, in some way, involved in the fertility and fecundity of the life on the Earth. One of the most commonly discussed aspects of this idea was the belief that minerals in the Earth would re-grow and replenish themselves. Mines that were becoming hard to work were closed down for a few years, in the same way that farmers would allow heavily worked fields to 'lie fallow' for a year or two. The assumption was that the minerals, often found in 'veins' in the earth, would once again be found in abundance in the mine after this period of recovery.

The advent of the mechanical philosophy saw a change – particularly with regard to what is now called geomorphology, how the Earth came into existence, and how it came to have the topography we now see. Remember that Descartes explained the solar system in terms of his 'vortex' theory. At the Creation, God set matter in circular motion about a centre, to create a huge whirlpool of particles of matter. Descartes then suggested that his three laws of nature were sufficient to account for the fact that the particles of matter crushed in the centre of the vortex became so heated by friction that it began to give off heat and light, forming a sun; and that lesser conglomerations of matter, further from the centre, where the matter was less crowded together, formed planets.

Descartes, who proudly boasted that his mechanical philosophy could explain everything, gave an account of the formation of the Earth. The details are not important (and anyway not entirely coherent) but the story involves the formation of a spherical body with a central core, surrounded by spheres of differing

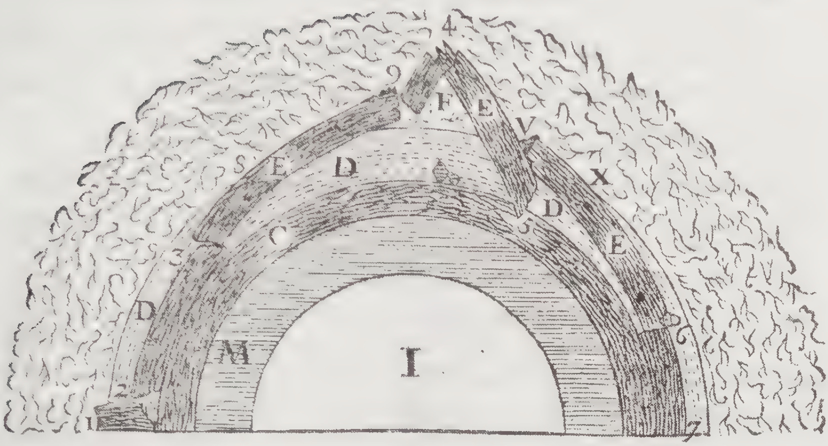

Figure 17.1 Depiction of the formation of the topography of the Earth; from Descartes, *Principia philosophiae* (Amsterdam, 1644). Reproduced with permission from Edinburgh University Library, Special Collections (JA 398/1).

An outer crust, E, originally surrounding a lower sphere of water, D, has broken up and collapsed down into D. In some places (labelled 7, 6, 2, 3) the crust has fallen below the water, while in others it forms land masses above the water; in other places, broken pieces of the crust form mountains.

densities, as particles of matter of different sizes are variously affected by centrifugal and centripetal forces.

But the Earth isn't a perfect sphere – the land and sea in spite of different densities lie intermingled, and the land surface isn't perfectly spherical. So, Descartes supposed that a sphere of dry land formed over a sphere of water, and that this outer shell of spherical dry land eventually broke up and collapsed into the watery sphere below, leaving a jumble of land masses, some quite jagged and mountainous, sticking up here and there through the water (see Figure 17.1).

Descartes's approach to the formation of the Earth was taken up by Thomas Burnet (*c.*1635–1715), who had taught Descartes's philosophy at Cambridge before becoming Master of Charterhouse School. Burnet's aim in his *Sacred Theory of the Earth* (1681), was to use Cartesian philosophy in support of the Bible. Religious sceptics took delight in arguing that forty days and nights of rainfall, heavy as that rainfall might be, would not suffice to flood the whole world, including the mountain tops. Burnet appropriated the Cartesian scheme to suggest that an initially smooth Earth, without any mountains, might well be temporarily flooded when the smooth surface first collapsed into the waters beneath it. Rather than deviating from Scripture, Burnet's Cartesian scenario actually solved a puzzling feature of the Biblical account:

> In the six hundredth year of Noah's life, in the second month, the seventeenth day of the month, the same day were all the fountains of the great deep broken up, and the windows of heaven were opened. (Genesis 7, 11)

As well as the heavens opening (with rain), the flood was caused by the fountains of water coming up from below as the Earth's surface was broken up.

Burnet was also able to claim that the Earth today was not as God originally fashioned it. It is now a 'Rude and Ragged Ruin', which God would never have deliberately created:

> If the sea had been drawn around the earth in regular figures and borders, [he wrote] it might have been a great Beauty to our Globe, and we should have reasonably concluded it a work of the first Creation, or of Nature's first production; but finding on the contrary all the marks of disorder and disproportion in it, we may as reasonably conclude, that it did not belong to the first order of things... God has destroy'd the first World.

The perfect, smooth sphere that Descartes described was more in keeping with what God would have designed.

Clearly Burnet was reflecting the aesthetic sensibilities of his day. His idea of Paradise, where Adam and Eve dwelt before the Fall, was like the formal, geometrical gardens favoured by the gardeners in the palaces and grand houses of Europe. Nature before the Fall was not jumbled, unkempt, and wild, but was smooth, geometrically regular, manicured, like a formal garden.

This kind of consideration of the Earth as something which came into existence in one form, and has subsequently changed in ways that can be explained in terms of natural philosophical theories, also meant that fossils were taken seriously for the first time. Fossils had simply been regarded as 'tricks' or 'sports' of nature – shapes that looked like plants or animals in the rocks, but which were just accidental features of the rocks. Hardly anybody gave them a second thought. In the seventeenth century, however, Robert Hooke (1635–1703) argued that they were signs of the remains of once living creatures, and John Woodward (1665–1728), used common features of fossil-bearing rocks to provide evidence that Noah's Flood (Genesis, 6–9) was a real event.

Woodward noticed that rock formations were frequently stratified, or layered (and particularly so if they contained fossils). He was also conscious of the familiar fact that fossils of what seemed to be sea creatures were often found miles from any sea. So, he assumed that the waters of the great Flood dissolved, or perhaps carried in suspension, much of the previous land, and that as the waters subsided, the different minerals and rocks precipitated out in order of decreasing specific gravities. So, the densest and heaviest rocks precipitated first, and the lightest last. Similarly, the carcasses of animals settled in each layer of rock, with the heaviest in the lowest formations, and the smaller, lighter, animals in the highest deposits. Given the scanty knowledge of fossil remains at this time, especially the few known remains of giant creatures later to be known as dinosaurs, Woodward's theory seemed plausible. But notice that what Woodward's theory means is that all fossils and all rock layers are of roughly the same age.

These early beginnings gave rise to a tradition of trying to show how natural philosophical assumptions about the formation of the Earth could be made to look compatible with the Biblical account of early history. This flourished chiefly in Britain, where, as we've seen, natural theology attempted to use the natural world to prove the existence and power of its supposed Creator.

The situation in France, where many leading intellectuals were anti-clerical if not irreligious, was rather different. One of the great French Enlightenment thinkers who took an interest in the formation of the Earth was George Louis Leclerc, Comte de Buffon (1707–88). Buffon assumed that the solar system was formed as a result of a comet colliding with the Sun. The supposed comet knocked gobbets of molten matter out of the Sun which went into orbit around the Sun, and then cooled to form the planets. Obviously, this means that the Earth, in its early history, would have been much hotter.

Seeking a way to estimate the age of the Earth, Buffon embarked on a series of experiments on the rates at which spheres of different substances, and different sizes cool down from raised temperatures. He used these as a basis for calculating the time needed for the Earth to cool down from white heat to a temperature suitable for life to develop. He published these results in his *Introduction to the History of Minerals* (1774). According to Buffon, a sphere of iron the size of the Earth, in a molten state, would take 4,026 years to solidify to the centre, would be safe to touch after 46,991 years, and would reach its present temperature after 100,696 years. He believed that the Earth would cool quicker than iron, however, and so concluded that the Earth was at least 74,832 years old. Privately, Buffon thought this figure was too low.

Now, of course, this figure is laughably small compared to what we now believe, but given that the more usual approach, prior to this, was to use Biblical scholarship and interpretation to estimate the age of the Earth (and resulted in an assumed age of about 6,000 years), it shows that Buffon was concerned to arrive at the age of the Earth using only naturalistic evidence and arguments.

In 1778 Buffon extended this work and gave an account of what he called the 'Epochs of Nature'. Here, Buffon elaborated on the already established (though not uniformly accepted) tradition that the six days of creation actually referred to six periods, or stages, of time (Newton, for example, had pointed out that since the Sun and Moon were created on the third day, it wasn't possible to speak of real 'days' before that, and if we assume that the Earth's rotation on its axis began to build up from an initially non-rotating Earth, then days could again be assumed to last as long as we like).

For Buffon the first epoch was the age of the molten Earth, the second saw the appearance of rocks cooling and hardening on the surface. The third epoch saw the appearance of seas and oceans, as the Earth cooled sufficiently to allow the condensation of water vapour in the atmosphere. The fourth epoch saw the appearance of more land masses built up by volcanic activity. Tropical animals (adapted to the hot conditions) spread all over the Earth in the fifth epoch. The continents separated in the sixth epoch, and humankind appeared on the earth in the seventh epoch.

It should be noted that even a French thinker like Buffon was concerned at this time to make his theories compatible to some extent with the Biblical account. The Church in France was, after all, still very powerful. Whatever the success of this attempt to forestall religious criticism, Buffon's work certainly drew criticism from his fellow philosophers. It was seen as simply too speculative. His experiments intended to estimate the rate of cooling of the Earth were all well and good, but Buffon couldn't possibly establish that the Earth had been thrown out of the Sun.

The next move in this tradition was made by Pierre Simon Laplace (1749–1827), whom we have already come across as a one-time collaborator with Lavoisier. Laplace was a consummate Newtonian mathematician and he used his mastery of Newtonian cosmology to develop the so-called 'nebular hypothesis'. The idea was that the solar system emerged out of a vast cloud of luminously hot gas, which would spontaneously rotate, and as it cooled gravity would ensure that the sun and the planets were formed out of it. This would also account for the fact that all the planets rotate around the Sun in the same direction, are all in roughly the same plane of rotation around the Sun, and (as far as was then known) all rotate on their axes in the same sense.

Laplace's mathematics was formidable and difficult to dismiss, and what's more it even attracted indirect observational evidence. The German émigré astronomer (who had settled in England), William Herschel (1738–1822), actually observed fuzzy celestial objects which he saw as nebulae (now known as galaxies). He also observed triple, and double star clusters, which he took to be new solar systems in the making, seen at different stages of their development. It now seemed likely, therefore, that the Earth had gradually cooled out of a swirling cloud of hot gas, in accordance with the principles of gravity and other features of Newtonian fluid dynamics.

No beginning and no end: James Hutton and the endless history of the Earth

Meanwhile, back in Britain what has been regarded as a kind of 'Bible Studies School of Geology' was continuing to thrive – that is to say, geology which was seen as fitting in with, and therefore confirming, the book of Genesis and, in particular, the story of Noah's flood. But a more secular attempt to explain the nature of the Earth was proposed by James Hutton (1726–97), a successful gentleman farmer who had been able to retire and set up home in Edinburgh.

Recognizing the importance of Baconian strictures against premature theorizing, Hutton did not speculate about the formation of the solar system or even about how the Earth might have formed, he concentrated instead on trying to interpret the evidence provided by the detailed topography of the Earth.

Hutton was a dedicated Newtonian and as such can be seen as participating in the Newtonian tradition of natural theology. He was not, however, a fully fledged Christian. He was what is known as a Deist. That is to say, he believed in the existence of God on the basis of the signs of intelligent design supposedly manifested in the natural world, but rejected the teachings of revelation as simply the imaginative and superstitious invention of ordinary men. We can also see in Hutton the belief, discussed in the last chapter, that this world must be the best of all possible worlds, and that what seems to be discordant, inharmonious, destructive, degenerative, and so forth, must in fact be all part of God's benevolent plan.

Hutton had studied various new and experimental farming techniques in East Anglia, near Great Yarmouth. He then put the same techniques to work on a small farm in the Borders, near Duns. This farm was sufficiently successful that he could lease it out and retire to Edinburgh, and spend his time as a gentleman of leisure.

Hutton's background in farming gave him an insight into the importance of erosion, or denudation, for replenishing the fertility of the soil. Rocks which offered no support for life, once broken down by erosion to form new soil, enabled plant life to flourish. Erosion, therefore, was a beneficial natural process. However, it was also realized that, potentially, erosion could spell the end of life as we know it. As Joseph Black pointed out in 1787: 'without being counteracted, the inequalities of the Earth's surface would be levelled in the course of time to a perfect plain, and mostly covered up by the sea or by collections of stagnating water'.

But Hutton's deism, and his belief that this is the best of all possible worlds, meant that he refused to believe that the world contained the seeds of its own destruction. It followed, therefore, that the seemingly destructive force of erosion must be part of a process of ensuring the continued fertility of the Earth, and the continued abundance of plant and animal life. As Hutton wrote:

> It has already been our business to show that the land is actually wasted universally, and carried away into the sea. Now, what is the final cause of this event?

Hutton's reference to final cause shows the still lingering relevance of Aristotelian ideas. One of Aristotle's four causes, the final cause refers to the purpose of something. What is the ultimate purpose of this perpetual wasting of the land? It was the attempt to answer this, and closely related, questions which prompted Hutton to write his great work, his *Theory of the Earth* (1795):

> Is it in order to destroy the system of this living world, that the operations of nature are thus disposed upon the surface of this earth? Or, is it to perpetuate the progress of that system, which, in other respects,

> appears to be contrived with so much wisdom? Here are questions which a Theory of the Earth must solve.

Hutton went on:

> The object of my Theory is to show that this decaying nature of the solid earth is the very *perfection* of its constitution as a living world... we shall thus be led to admire the wisdom of nature, providing for the continuation of this living world, and employing those very means by which, in a more political view of things, this beautiful structure of an inhabited earth seems to be necessarily going into destruction.

The Earth, then, is a self-repairing, self-correcting mechanism. Indeed, Hutton insists that:

> If no such reproductive power, or reforming operation, after due inquiry, is to be found in the constitution of this world, we should have reason to conclude that the system of this earth has either been intentionally made imperfect, or has not been the work of infinite power and wisdom.

Needless to say, Hutton couldn't accept either of these two alternatives. For him, the earth is the work of infinite power and wisdom, and would not have been made imperfect by such a being, but must be perfect.

According to Hutton, the world works by the balance of attractive and repulsive forces. Newton's two great principles are here once again. Erosion is the manifestation of the inexorable workings of gravity – causing rain and river water to fall onto the rocks and so erode small particles which gradually descend, eventually to settle at the bottom of the sea.

But there is also an expansive force, the force we can see in the working of heat. Inside the Earth is a Newtonian subtle-fluid of heat which performs two functions – first, it fuses the eroded continental debris, which has been washed to the bottoms of the oceans, to form new rocks; and second, it exerts an expansive force, tending to push the ocean floors up, to make new land.

Now, unfortunately Hutton could not provide much evidence for the expansive part of his cycle. Obviously, he could point to volcanoes as indicators of the internal heat of the earth:

> A volcano should be considered as a spiracle to the subterranean furnace, in order to prevent the unnecessary elevation of land, and fatal effects of earthquakes.

But he had great difficulty persuading his contemporaries that heat could consolidate loose material into rocks. After his death, his follower Sir James Hall (1761–

1832) managed to provide some experimental evidence that this was possible by heating calcium carbonate powder under pressure, to produce artificial limestone.

Even the erosive part of Hutton's cycle demanded a lot of his audience, however, because of the time factor involved. Hutton's cycle of erosion to the sea and the raising of new land from the sea floor required inconceivably vast amounts of time. Hutton's response was simply to insist that it must be so:

> Our fertile plains are formed from the ruins of mountains; and those travelling materials are still pursued by the moving water, and propelled along the inclined surfaces of the earth... The immense time necessarily required for this total destruction of the land must not be opposed to that view of future events, which is indicated by the surest facts, and the most approved principles. That Time, which measures everything in our idea, and is often deficient to our schemes, is to nature endless and as nothing.

This concept of indefinite, almost endless, time was bound to draw the anger of devout Christians, who were used to thinking of the history of the earth in Biblical terms, but Hutton went on to make things worse. The cyclical nature of his theory meant that it was impossible to gather any evidence about the origin of the world. There are no unaltered primitive rocks in Hutton's theory. The deeper down a rock is, only testifies to the number of times it has been broken down by erosion, and re-consolidated:

> The inferior mass [of rocks] must have undergone a double course of mineral changes and displacements; consequently, the effect of subterranean heat or fusion must be more apparent in this mass, and the marks of its original formation more and more obliterated.

This, and other considerations, led Hutton to draw his famous conclusion:

> For having, in the natural history of this earth, seen a succession of worlds, we may from this conclude that there is a system in nature; in a manner as, from seeing revolutions of the planets, it is concluded, that there is a system by which they are intended to continue those revolutions. But if the succession of worlds is established in the system of nature, it is in vain to look for anything higher in the origin of the earth. The result, therefore, of our present enquiry is, that we find no vestige of a beginning – no prospect of an end.

Hutton was making a methodological point, about what could be inferred from the evidence of the rocks, but devout contemporaries saw it as an atheistic statement: 'it is in vain to look for anything higher in the origin of the earth'. To them, Hutton seemed to be saying that the world was not created, and would continue for ever – both of which went against Christian teachings (Box 17.1).

Box 17.1 HUTTONIAN UNCONFORMITIES AND THE CONCEPT OF DEEP TIME

Figure 17.2 A Huttonian unconformity, depicted in James Hutton, Theory of the Earth (Edinburgh, 1795). Reproduced with permission from Edinburgh University Library, Special Collections Department (SC 6306).

This famous picture depicts what is called a Huttonian unconformity. The vertical strata at the bottom of the picture are overlaid by a later sequence of horizontal

Initially, therefore, few thinkers took up Hutton's ideas. He had a devoted band of followers in Edinburgh, but his ideas were marginalized elsewhere in Britain, where geology was still conducted in partnership with the Bible. The value of his work would only be recognized when a later geological theorist felt a pressing need, for reasons of his own, for a geological theory which depended upon very slow processes and, therefore, vast and virtually unending periods of time. We will come to this shortly.

Catastrophism in geology

Although Hutton's ideas about the vast timescale required to explain Earth's topography were rejected, all geologists were now acknowledging that the Earth must be a great deal older than the 6,000 years of Biblical chronology. Even religiously devout geologists were regarding the six 'days' of creation as an allegorical strata, in a way that makes it perfectly clear they are unconformable – that is to say, they are evidently not parts of the same formative process, but evidence of two separate processes. Indeed, correctly interpreted, there is evidence here of four successive landscapes.

Landscape one is gradually eroded by wind, rain, and running water, and washed down to the bottom of an ocean where *horizontal* strata are laid down.

Landscape two is formed when these strata are gradually forced upwards to emerge out of the sea. Imagine that in this process the horizontal strata are bowed severely upwards, so that they form a very narrow arch – curved on top, but vertical, or nearly so, at the sides. Landscape two now becomes subject to erosion and eventually the top part of the 'arch' is washed away leaving two sets of *vertical* strata. During the erosion of landscape two these vertical strata have become submerged once more under a sea.

A third landscape is then eroded and washed down to the bottom of this new sea. The eroded material from this former landscape covers the exposed ends of the vertical strata of what was once landscape two.

Landscape four, which we see in the picture, is formed when the horizontal strata resulting from the erosion of landscape three, and the vertical strata lying beneath them, are pushed up by subterranean heat.

Bearing in mind that the processes of erosion and up-lift are so slow as to be barely noticeable in a human lifetime, it is clear that this picture implicitly depicts inconceivably vast periods of time. Furthermore, we cannot conclude that there was no landscape before landscape one. That too, might have been pushed up from earlier strata, eroded from an earlier landscape, laid at the bottom of an earlier sea. As Hutton wrote, 'we find no vestige of a beginning...'.

way of talking about something more like the epochs of nature described by Buffon.

Evidence for the great age of the Earth was manifest everywhere. One of the most impressive examples was Jean-Etienne Guettard's (1715–86) realization in 1751 that the Puy de Dôme in the Massif Central region of France was part of a chain of extinct volcanoes. It was clear from the lava flows which were discovered in the area that the volcanic activity had once been highly intense, and yet there was no hint of this in any of the ancient legends or myths of the area, much less the historical chronicles. So, all of this volcanic activity must have pre-dated the appearance of the ancient peoples known to have lived here. Some lava flows could be seen to have been diverted by pre-existing river valleys, and to have been subsequently eroded through as the river re-established itself – the picture was of landscapes of immense age and slow formation.

Although geologists were agreed that the traditional timescales had to be extended, there was otherwise little agreement between them as to how the surface of the Earth came to be the way it is. Towards the end of the century, however, a renowned teacher emerged who was able to inspire a unity of effort among geologists. This was Abraham Gottlob Werner (1749–1817).

Werner was a professor at the School of Mines in Freiberg, and what he did, effectively, was to remind geologists of the importance of what was seen as the 'Baconian' enterprise of fact-gathering. The study of the rocks is at too early a stage for a geological Newton to emerge, Werner was suggesting, and what was needed was a period of collaborative effort to build up a body of reliable facts.

Werner showed the way by publishing his own *Short Classification and Description of the Rocks* (1787). This can be seen as the geological aspect of the eighteenth-century vogue for classifying (we've already seen how the work on affinity tables in chemistry served to provide data for classifying chemicals and their reactions, and we shall look at developments in botany and zoology in the next chapter).

Earlier attempts to classify rocks had focused on their mineralogical or chemical features (Lavoisier had been involved in this enterprise early in his career). Werner, however, suggested that it was more important to classify them according to age, and to order of deposition. This simple idea was so useful that it seems clear that geologists immediately recognized its importance for their newly emerging science. The period from the 1790s to the 1830s was a period of intense activity in surveying, mapping, and comparing different regions, in order to provide a solid database for geological theorizing.

Now the age of formation of rocks ought to correlate with order of deposition – the deepest rocks must have been laid down first, and so ought to be the oldest. But things are rarely that simple, and in some places it was apparent that the rocks had been folded, twisted, shunted, crumpled, and so on, so that in some places the order was the reverse of the order in other places, or layers were doubled up – and such effects were particularly dramatic in mountainous areas. These facts demanded explanations (it is hard to stick to Baconian fact-gathering when confronted with such odd results), and so various hypotheses were put forward to explain mountain building, and other dramatic aspects of the scenery. These hypotheses often drew upon the dramatic forces of volcanoes and earthquakes. It seemed intuitively obvious, however, that many of the observed features would have required far more powerful forces than those seen in present day earthquakes and volcanic eruptions.

This problem did not detain geologists for long. After all, it was generally agreed, following the Newtonian tradition promoted by Buffon, Laplace, and others, that the Earth was formed by cooling down from solar temperatures. It followed, therefore, that in the past it was much hotter, and that its surface was likely to be more plastic and malleable, and that its interior would be subject to more violent and powerful forces than the later, cooler, Earth.

These kind of assumptions led to a dominant tradition in geology which is called catastrophist geology. The general guiding principle of this tradition was that topological features could all be explained in terms of cataclysmic events, and that such cataclysmic events were likely to be comparatively common phenomena at earlier stages in the Earth's formation. Cataclysmic floods were also included in these explanatory scenarios; the assumption being that sudden upward shifts in sea beds, or sudden deposition of land adjacent to a sea, could cause rapid influx of water on to dry land. There seemed to be strong evidence in favour of such powerful floods. Massive boulders (now known as 'erratics') which are found on a valley floor many miles from the rock formations from which they must have originated could not have been moved by even a fast flowing river, but they might have been swept along in a catastrophic flood. In fact, we now know such boulders can be carried by glaciers, but the concept of ice ages, in which glacial activity was far more extensive than now, was not developed until 1837 (by the Swiss naturalist Louis Agassiz, 1807–73). Given the array of evidence in its favour, catastrophism was virtually unanimously accepted by geologists from the 1790s until the 1830s.

Fossils and the origin of species

Another major support for catastrophist geology came from the fossil record. Since some fossils are characteristic of the rocks in which they are found, the fossils could be used in the Wernerian classificatory schemes to provide relative ages for the rocks. The first really successful accomplishment in this new palaeontological tradition was the collaborative investigation of fossils in the Paris basin, published in 1811 by Alexandre Brongniart (1770–1847), a geologist, and Georges Cuvier (1769–1832), who was an expert zoologist. Similarly, William Smith (1769–1839), who published the first geological map of England (which included Wales and even parts of Scotland) 1815, had also used fossils to provide the relative ages of rocks.

It became apparent in this and other work on fossils that the older the fossils were, the less similar they were to any forms still alive today. But, more significantly for geological theorizing, there seemed to be clear disjunctions between the flora and fauna found in different levels of rock. Sometimes, for example, it almost looked as though land animals alternated in successive rocks with sea creatures.

The obvious way to read these disjunctions was to assume some catastrophe had overtaken the animals whose remains were found in one layer of rock, and they were replaced by a completely new set of creatures, whose remains were then found in the next layer. So, entire populations were wiped out, only to be replaced by a population of completely different animals and plants. This seemed to confirm catastrophism – how else could entire populations be wiped out? As a result catastrophist geology and palaeontology went hand-in-hand as the most powerful research tradition in geology.

One of the most important outcomes of this work was the almost unanimous consensus among geologists that there was a *progression* of life forms revealed in the fossil record. So, in the oldest rocks there are no fossils, then we find only very simple, primitive, creatures, then the fishes begin to appear, subsequently reptiles and what were eventually designated as dinosaurs, and then mammals, and finally mankind. This looked to late eighteenth and early nineteenth-century geologists like a progression from primitive life forms to ever more complex forms, culminating with the comparatively recent emergence of mankind.

This immediately gave rise to a new problem. If different kinds of life form appeared on Earth at different times, as the fossil record seemed to suggest, where did these new forms come from? What, in short, was the origin of species?

Now, it might be supposed that there was an obvious solution to this problem – namely, that the later life forms *evolved* from the earlier forms. It is important to note, however, that the majority of thinkers at this time did *not* believe the subsequent more advanced forms had grown out of the earlier forms, by some evolutionary process.

On the contrary, catastrophist geology seemed to imply that one population was cataclysmically wiped out and then replaced by a new set of creatures. The evidence of the fossils, therefore, seems to go against any idea that one population could transform into a different population of new creatures. Such an evolution of one set of creatures into another set would require a long time scale without any catastrophes occurring which are likely to destroy the gradually shifting population. So, belief in *progression* of life forms, did *not* entail belief in evolution. The progression of forms was seen as taking place in a sudden step-wise fashion: first, there are only cephalopods on the Earth, then fishes suddenly appear, and so on. Evolution by contrast, requires a gradual transformation to take place (so that some cephalopods actually turn into fish). The fossil record made it impossible to deny step-wise progressionism; but hardly anybody (and no geologists) believed in evolution.

The scenario envisaged by most geologists went something like the following. When the Earth is first formed it is too hot for life to survive. Only gradually, as the Earth cooled, did the lowest forms of life appear: lichens, mosses, etc. clinging to the earliest rocks. As conditions on Earth gradually change, animals which are capable of surviving in the prevailing conditions emerge. Each successive population of creatures are assumed to make changes of their own to the environment. For example, the world-wide tropical rain forests of the so-called Carboniferous period help to reduce the carbon dioxide in the atmosphere by absorbing it, and then, when the dead plants form first peat, and then coal, the carbon dioxide is 'fixed' in the coal. Only then could higher animals appear, because the atmosphere is now richer in oxygen.

It was also assumed that in some cases at least, the emergence of new animals might be dependent upon the prior extermination of earlier animals. The dinosaurs must have served some ecological purpose, creating new conditions much as the carboniferous rain forests had, but then they became extinct and presumably left the way clear for mammals to take over the world. With the appearance of mammals as the dominant kind of life on land, mankind could finally emerge; exploiting other mammals for food, first by hunting, and subsequently by domestication.

The point is, of course, that, in spite of seeming cataclysms wiping out whole species, and even entire genera, everything is for a purpose, and 'all is for the best, in the best of all possible worlds'. It is not clear why dinosaurs were created only to be subsequently wiped out, but there must have been a reason, and it must in some way have been important for the ultimate emergence of humankind.

Clearly, this general view is strongly affiliated to religious assumptions. It is important to note, however, that those religious assumptions are based on the universal belief, among orthodox religious believers and among deists who rejected the doctrines of revealed religion alike, that God has created the world but has contrived for its continuation and its development entirely in accordance with laws of nature which he has established. The scenario just sketched out, from an Earth initially too hot for life to the eventual emergence of humans, was envisaged as an entirely natural process (or a vast complex of interconnected natural processes) – explicable and understandable *in principle* by the workings of natural laws (even if, in practice, we don't yet know all the relevant laws).

Earlier historians of science have shown an unfortunate tendency to assume that catastrophist geologists and other contemporary thinkers believed that the new species of creatures appeared on Earth at the appropriate times in its history because God directly created them. In all my reading of nineteenth-century writers, I have never seen such a claim unequivocally expressed. Natural philosophers, ever since the Middle Ages and the foundation of natural philosophy as a university discipline separate from theology, have offered *naturalistic* explanations of the phenomena they consider. It would have been a remarkable turnaround to find natural philosophers in the nineteenth century abandoning this centuries-old tradition and simply declaring that new species were directly created by God. It was one thing to point to the intricacies of the way the natural world worked and use them to prove the existence of a creator God; it would have been quite another to simply say the natural world worked in a particular way because God made it that way – this would have been a betrayal of the whole ethos of natural philosophy. It was acceptable to say that there was a divine purpose behind all natural change; but it was never considered acceptable to say that God brought about any given change by direct intervention.

The demand for a naturalistic but non-evolutionary explanation of the successive appearance of new life forms was extremely hard to fulfil. The Swiss natural philosopher Charles Bonnet (1720–93), and the French naturalist Jean-Baptiste Robinet (1735–1820), for example, independently drew upon contemporary theories of animal generation (or reproduction) in which it was supposed that females carried 'germs' which gave rise to succeeding generations. Both thinkers tried, but failed, to develop notions in which the germs of all potential species were capable of independent existence, and somehow gave rise to new populations of suitable species after geological catastrophes. The idea was that the germs of all species always existed, but only developed when times were ripe for them. In keeping with naturalistic explanations, appropriate environmental conditions were assumed to provide the trigger for germs to develop.

No such speculations proved persuasive and the majority of nineteenth-century naturalists retreated into Baconianism, declaring that theorizing about the origin of species was premature. Accordingly, geologists preferred to dismiss the mystery of the origin of species as beyond the realms of current science. As William Whewell (1794–1866), a leading scientist in his own right, but also a thinker who pondered the niceties of correct scientific method, said of the question of the origin of species: 'To this question, men of real science do not venture to return an answer'.

There were thinkers, working mostly in the botanical and zoological traditions, who were not content with this and began to develop ideas of transformism – the evolution of one creature into another by some natural process of change. For thinkers like Whewell, however, which seems to have included the majority of contemporary naturalists, theories of transformism were pseudo-scientific, based on jumping to false conclusions as a result of a failure to understand the significance of the evidence against it.

Charles Lyell and uniformitarian geology

In geology, therefore, catastrophism dominated the early decades of the nineteenth-century as what Kuhnian philosophers of science would call the ruling paradigm. But in 1830 the Scottish geologist Charles Lyell (1797–1875) published the first volume (of three) of his eagerly anticipated *Principles of Geology*. Here, Lyell – who had previously been as much a catastrophist as everyone else – explicitly rejected catastrophism and began to promote a rival view of Earth history, usually known as uniformitarianism (or, sometimes, 'actualism').

Lyell cannily defended this view on methodological – indeed Baconian – grounds. Catastrophism, he said, was based on conjecture, on premature theorizing – the assumption (which could not be established empirically) that in the past the Earth was racked by catastrophic forces and events. His proposed alternative approach, by contrast, argued only on the basis of forces and processes seen to be actually at work at present. Uniformitarianism was based on three assumptions:

1 The laws of nature have not changed over time.
2 The kinds of causes operating on the Earth now have not changed, and have always acted the same way.
3 The intensity of the causes acting on the Earth have not changed over time.

These assumptions are often used, even today, to claim the methodological superiority of uniformitarianism over an excessively speculative catastrophism. But, if we remember that everyone at this time assumed that the Earth started out as a body cooling down from solar temperatures, the assumption that the intensity of causes operating may have been stronger in the past makes perfect sense.

Indeed, any catastrophist would have been entirely justified in demanding of Lyell why he didn't think this was the case.

In fact, Lyell could only reject the idea that subterranean forces were more vehement in the past by rejecting the Newtonian scenario of an Earth gradually precipitating out of a nebula of hot gas. Accordingly, Lyell went back to James Hutton's theories and insisted that the Earth's surface features are all explicable in terms of the same levels of erosion we see occurring today, and very slow uplift caused by subterranean heat. Lyell was also committed, therefore, to a vast, almost endless, timescale. Effectively, Lyell insisted, the Earth is in a steady state. Far from being affected by cataclysmic events, all its processes are extremely slow and balanced out – as mountains are washed to the sea in one part of the Earth, a gradual uplift creates mountains, or at least dry land, in another part of the world (in fact, mountain building was one of the major unsolved problems for Lyell's geology).

The big problem facing Lyell's new view was the fossil record. If there were no catastrophes, why did these periodic mass extinctions occur? Lyell somewhat desperately pointed to a couple of fossil finds of what were agreed to be mammals, or mammal-like creatures, living in the age of dinosaurs. He used these to undermine the progressionist claims of the catastrophists. He insisted that the impression of a progression of life-forms was illusory, based on jumping to conclusions from a hopelessly inadequate database of fossil remains. Inevitably, Lyell suggested, the dominant and therefore most numerous species at any given time leave more fossils. But that doesn't mean there are no other creatures alive at that time, just that they failed to leave fossil remains, or so few that we have not yet found them. So, just as there were mammals living in the age of dinosaurs, so there might be dinosaurs living today, Lyell said, albeit in small numbers. Lyell's ideas gave rise to science-fiction stories such as Jules Verne's *Journey to the Centre of the Earth* (1864) and Arthur Conan Doyle's *The Lost World* (1912), in which dinosaurs were found still thriving in remote places. But for Lyell, it was important to insist on this to maintain his view that the Earth was essentially in a steady state. The crucial thing for him was to reject the standard belief that there was a historical *progression* of life forms, from simple and primitive to humans via successively more complex forms in between.

Now, we know from his notebooks that Lyell himself was a catastrophist until 1827 when something happened which made him change his mind, and which led him to resuscitate Hutton's work and to develop the idea that the Earth, its environment, and its creatures, were in a steady state, rather than in a continual process of progression. So, what happened? Why did Lyell suddenly feel the need to reject catastrophist geology, and the progressionism which seemed so obvious from fossil remains, and to develop instead his uniformitarian geology?

The answer lies in the problem of the origin of species. Lyell believed in God but, like all good natural philosophers, he could not bring himself to believe that God would repeatedly intervene directly in nature to create new species as and when needed. So, he started to look around to see if he could discover a theory which explained the origin of new species in naturalistic terms. In 1827 Lyell came across a book written by the French biologist, Jean-Baptiste Lamarck (1744–1829), called *Philosophie zoologique* (1809). It was reading this which caused Lyell to radically change his geological outlook from catastrophism to uniformitarianism.

It is clear that, on reading Lamarck, Lyell decided that he had to stop this evolutionary theory from gaining any foothold. Lyell saw the clear links between Lamarck's version of evolution and geological ideas about the progression of the life forms in the history of the world. So, Lyell developed a geology which undermined the theory of progressionism. It seems clear that Lyell hoped that his uniformitarian geology would dispel the illusion that there was a historical progression of life forms, and so remove the principle ground upon which Lamarck's evolutionary theory was based. If all creatures were always the same throughout the ages, merely changing in terms of population density, rather than in biological morphology, then talk of evolution was completely superfluous.

So, why did Lyell recoil so strongly from Lamarck's work? Because Lyell's religious principles got the better of him. He couldn't bear the idea of humans being descended from apes. Consider, for example, this extract from Lyell's journal, written in 1858:

> If the geologist dwelling exclusively on one class of facts, which might be paralleled by the existing creation [arrives] at conclusions derogating from the elevated position previously assigned by him to Man, if he blends him inseparably with the inferior animals and considers him as belonging to the earth solely, and as doomed to pass away like them and have no farther any relation to the living world, he may feel dissatisfied with his labours and doubt whether he would not have been happier had he never entered upon them and whether he ought to impart the result to others.

The irony in all this is that catastrophism is often dismissed by historians of geology as bad science which was tainted by the attempt to see geology in Biblical terms – the most recent catastrophe to befall the world had for a time been equated with Noah's Flood by early catastrophist geologists. Lyell's uniformitarianism, by contrast, is upheld as good science, based on sound methodological principles, and untainted by religious considerations. This is a travesty of the real history.

In fact, catastrophists had good scientific reasons for their conclusions. The latest catastrophe for which they saw evidence was in fact the last ice age, but not knowing about ice ages (a concept introduced by Louis Agassiz and others in the late 1830s but not accepted until the 1870s), they interpreted many of the effects which are now known to have been formed by glaciers as the result of a cataclysmic flood. Bearing in mind that the last ice age ended about 10,000 years ago, it would not have been so absurd for nineteenth-century thinkers to equate it with a Flood reported in the Bible to have affected the world about 6,000 years ago. Having said that, it is worth noting that even religiously orthodox geologists tended to reject the identification of Noah's Flood with a real geological catastrophe by the early decades of the nineteenth century. William Buckland (1784–1856), devout author of *Reliquiae Diluvianae* (*Relics of the Flood*) in 1823, had abandoned any concern to confirm the Scriptural Flood by 1830. Similarly, Adam Sedgwick (1785–1873), devout Christian clergyman and leading geologist, explicitly denounced Scriptural geology in 1833. So, it is unfair to suggest the catastrophists were deviating from scientific evidence as it was then known. In rejecting a progression of life forms in the fossil evidence, however, Lyell almost certainly was. Furthermore, as we've seen, his motives for doing so were not entirely scientific; Lyell developed his steady-state theory in response to the stimulus of what he saw as the unsavoury religious implications (leading to Lamarckian evolution) of progressionism in geology.

It is important to note, however, that it would be wrong to conclude that Lyell was engaged on a deliberate conspiracy here – deliberately concocting a scientific theory that he knew could not be correct, but which he saw as the only way to foil the irreligious Lamarckian position. As a rule conspiracy theories are to be avoided, and certainly in this case, there is no evidence to suggest Lyell was insincere in his new scientific convictions. The point is, however, that scientific evidence does not speak for itself: it always has to be interpreted and it is simply not possible for any human interpreter to exclude all their preoccupations, prejudices and preconceived opinions when making those interpretations. Scientists are humans too, and interpret things in accordance with their own preconceived ideas, as I hope has been apparent throughout this book. It seems more in keeping with the evidence, therefore, to suppose that as Lyell read Lamarck he became genuinely convinced that Lamarck was jumping the gun in basing his arguments on a supposed historical progression of life forms. Subsequently, the more he looked into the geological evidence, the more he became genuinely convinced that there was no foundation for a belief in evolutionary progression.

FURTHER READING

Michael Bartholomew, 'Lyell and Evolution: An Account of Lyell's Response to the Prospect of an Evolutionary Ancestry for Man', *British Journal for the History of Science*, 6 (1972), pp. 261–303.

Michael Bartholomew, 'The Singularity of Lyell', *History of Science*, 17 (1979), pp. 276–93.

Geof Bowker, 'In Defence of Geology: The Origins of Lyell's Uniformitarianism', in M. Serres (ed.), *A History of Scientific Thought: Elements of a History of Science* (Oxford: Blackwell, 1995), pp. 483–505.

Peter J. Bowler, *The Fontana History of the Environmental Sciences* (London: Fontana, 1992), chapters 4 and 6.

Dennis R. Dean, *James Hutton and the History of Geology* (Ithaca: Cornell University Press, 1992).

C. C. Gillispie, *Genesis and Geology* (New York: Harper & Row, 1959), chapters 2–5.

R. Grant, 'Hutton's Theory of the Earth', in L J. Jordanova and R.S. Porter (eds), *Images of the Earth* (Chalfont St. Giles: British Society for the History of Science, 1979), chapter 2, pp. 23–38.

J. C. Greene, *The Death of Adam* (Ames: Iowa State University Press, 1959), chapter 3.

Martin Guntau, 'The Natural History of the Earth', in N. Jardine *et al.* (eds), *Cultures of Natural History* (Cambridge: Cambridge University Press, 1996), pp. 211–29.

David Oldroyd, *Thinking about the Earth: A History of Ideas in Geology* (London: Athlone, 1996), chapter 6, pp. 131–44.

Rhoda Rappaport, 'The Earth Sciences', in Roy Porter (ed.), *The Cambridge History of Science, Volume 4: Eighteenth-Century Science* (Cambridge: Cambridge University Press, 2003), pp. 417–35.

18

The History of Plants and Animals: Successive Emergence or Evolution?

We have seen how geologists, while insisting upon a historical 'progression' in life forms, vigorously denied any suggestion that animals could evolve from one form into another. And we saw why catastrophist geologists took this line. So, now we are going to look at what those studying plants and animals made of the seeming progression of plant and animal forms as revealed by the fossil record, and the sudden appearance of new species in the fossil record.

The starting point for understanding their attitudes is a familiar one to us: just like all other natural philosophers, those who studied flora and fauna believed that there was some rational order in the living world. God had created all things according to a regular, harmonious plan, and there should be, therefore, a discernable pattern in what seems to be a baffling diversity of plant and animal creations.

The basic pattern was assumed to be linear and hierarchical, and was usually referred to as the Great Chain of Being. As the Neoplatonist philosopher Macrobius (*fl. c.* 430) put it, in the fifth century AD:

> Since the single radiance of God illumines all and is reflected in many mirrors, placed in series, and since all things follow in continuous succession, degenerating in sequence to the very bottom of the series, the attentive observer will discover a connection of parts, from the supreme God down to the last dregs of things, mutually linked together and without a break. And this is like Homer's golden chain which God, he says, bade hang down from heaven to earth.

This notion was able to find support (though often in what we would regard as a naïve way) from seemingly transitional life forms: certain fungi were seen as transitional between rocks (especially those with a fibrous structure) and other vegetables; mudskippers or certain amphibians were transitional between fishes and land-based reptiles; flying fish between fish and birds; and so forth. The point was that God had left no gaps in the system of creation – in the best of all possible worlds, all possibilities of existence were fulfilled.

It followed that the attempt to understand the arrangement of all the different species was an attempt to understand not simply the Creation itself, but the God of the Creation. Taxonomy, the attempt to classify God's creation, was not simply an exercise in convenience, but was intended to reveal *real* relationships between the creatures and, as Kepler had said with regard to relationships in the heavens, to 'think God's thoughts after him'.

It evidently seemed intuitively obvious to early taxonomists (perhaps influenced by the political systems of which they were a part) that the system of nature was a single hierarchical sequence. The difficulty with setting out the creatures in a linear sequence, however, was that it is very difficult to arrive at a consensus as to what the exact order should be. The ongoing controversy among naturalists was exacerbated by the discoveries of new creatures found in newly discovered countries, such as the Americas. It is estimated that John Ray (1628–1705), one of the leading taxonomists of his day, knew only about 1,500 species in total, while the Swedish naturalist, Carl Linnaeus (1707–78), knew about 5,600 species of quadrupeds, to say nothing of other creatures.

Carl von Linné, or Linnaeus, whose ambition to know God led him to devote his life to discovering the natural order of the Creation, abandoned the attempt to provide a single linear sequence and assumed that God had opted instead to arrange creatures in groups. Starting with the age-old tradition – distinct from the Great Chain of Being tradition – that there were three separate 'kingdoms': animal, vegetable, and mineral; Linnaeus developed the idea of separate classes (mammals, birds, fishes, etc.), orders (carnivora, insectivora, rodents, primates), genera (canis, ursus, felis, panthera), species, and varieties. Linnaeus's binomial system of classification, designating genus and species, is still in use, of course. The genus *felis*, for example, includes *felis catus* (common domestic cat) and *felis lynx* (the lynx), while the genus of the large cats, *panthera*, includes *panthera leo* (lion), and *panthera tigris* (tiger). The brown bear is *ursus arctos* and the polar bear is *ursus maritimus*. Varieties are different manifestations of the same species – the Great Dane and the chihuahua are varieties of the domestic dog (*canis lupus familiaris*).

Linnaeus believed that his new system of classification was close to revealing the original patterns used by God in creating the diversity of creatures in the world, even though on a number of occasions he had had to resort to what he recognized were only artificially subjective distinctions between creatures. Such artificial, as opposed to natural, distinctions were intended to be temporary, making it easier to see the real God-given distinctions in due course. As Francis Bacon had suggested, truth can emerge from error more easily than from confusion, and Linnaeus evidently hoped this might be true.

Linnaeus first published his *Systema natura* in 1735, and repeatedly revised it up to 1774, after which time he was incapacitated by a stroke. His taxonomic system proved to be immensely useful and his ideas on the origin of new species, therefore, could hardly fail to attract attention. The fact is that after a lifetime of studying different plant and animal forms, Linnaeus gradually came to believe that 'species were the daughters of time'. He assumed that the process of species formation was in part due to changes in the environment. Species were assumed to be created by God so that they were perfectly adapted for their environments. But if environments change over time (as the new science of geology was suggesting they did) then surely species might adapt themselves to the new conditions. Linnaeus came to believe that God originally created only a single species to represent each genus, and the multiplicity of species gradually emerged in much the same way that varieties emerged within species. It seems, therefore, that what Linnaeus had in mind was a gradual transformation of one form into another, or into several other forms.

Insofar as Linnaeus and his followers considered the means by which these changes might take place, they focused on hybridization. Cross-breeding between genera, they supposed, might lead to new species. Unfortunately, hybrids are almost always sterile and cannot propagate themselves, but Linnaeus and his followers simply had to assume that it may be possible, given time, for repeated hybridizations to result in a hybrid form that was capable of procreating itself.

Linnaeus's main rival as a naturalist was the Comte de Buffon (1707–88), whom we have met before (see Chapter 17). As the author of a monumental, 44-volume work on natural history, the *Histoire naturelle*, Buffon discussed at length almost every aspect of evolutionary thought. Unfortunately his discussion is sporadic and piecemeal, being spread over several volumes, rather than systematically considered in one place. Furthermore, in spite of conceding much to the evolutionary approach, Buffon eventually rejected it.

The main factor leading him to deny the possibility of evolution was the sterility of hybrids. Like Linnaeus, he assumed that the obvious way to form new species was by analogy with the way new varieties are produced – by selective breeding. But, unlike Linnaeus, he could not bring himself to overlook the fact that varieties always remained within the same species (that is, they never led to new species – chihuahuas, beagles, and Dobermanns are all still dogs), and that hybrids across species were always infertile (and therefore could never lead to viable new species).

Judging from the *Histoire naturelle* as a whole, it is clear that Buffon subscribed to the belief that species were real, distinct forms, and yet, in the very first volume, he explicitly denied this view, and simultaneously rejected the Linnaean enterprise. Consider, for example, this famous passage:

> Thus surveying successively and in order all the various objects that make up the universe, and placing himself at the head of all created beings, Man will be surprised to see that one can descend by almost insensible degrees from the most perfect creature down to the most unformed matter, from the best organised animal to the purely brute mineral matter, he will recognise that these imperceptible gradations are nature's handiwork; and he will find these gradations not only in size and form, but also in manners of movement and generation, and in the successions of all species. If the meaning of this idea be fully apprehended, it will be clearly seen that it is impossible to draw up a general system, a perfect method, for natural history... Nature proceeds by unknown gradations, and consequently cannot wholly lend herself to these divisions – passing as she does from one species to another

> species, and often from one genus to another genus, by imperceptible shadings... Objects of this sort, to which it is impossible to assign a place, necessarily render vain the attempt at a universal system... in reality individuals alone exist in nature, while genera, orders, classes, exist only in our imagination.

Buffon seems to call here for a return to the single linear Chain of Being, and to dismiss the reality of species. It was perhaps the attempt to reconcile these forcefully stated views with the fact that elsewhere Buffon treats species as real, which led to the first systematic attempt to develop a theory of evolution. This evolutionary theory was worked out by one of Buffon's most devoted followers, Jean Baptiste Pierre Antoine de Monet, known as Lamarck (1744–1829), and it turned the linear ladder of nature into what has been seen as a perpetually moving escalator of nature.

Now, Lamarck was one of the first thinkers to coin the new word 'biology' to describe what was essentially a new science – 'the study of the origin and development of living organisms'. Before this, there were botany and zoology, two branches of 'natural history', which were essentially factual and descriptive enterprises, devoid of any theoretical content. Biology, as Lamarck now conceived it, was intended to be concerned with a more theoretical approach to the understanding of living things, and was seen by Lamarck as a new branch of natural philosophy or physics:

> A sound Physics of the Earth should include all the primary considerations of the Earth's atmosphere, of the characteristics and continual changes of the Earth's external crust, and finally of the origin and development of living organisms. These considerations naturally divide the Physics of the Earth into three essential parts... Meteorology... Geology... and Biology.

Lamarck's belief that the study of life on Earth must also involve the study of the environment – via meteorology and geology – confirms him to have been a follower of Buffon. It was also through Buffon that Lamarck became introduced to Newtonian ideas:

> The naturalist should turn to laws of nature, and in particular to universal attraction which constantly works to bring particles of matter together, to form bodies, and to prevent their molecules from dispersing; and [he should turn also] to the repulsing action of subtle fluids in a state of expansion; an action which... changes in a number of ways the state of closeness of bodies.

It is clear, therefore, that what we see in Lamarck is a thinker with a firm belief that everything can be explained in naturalistic, or physicalistic terms – even life itself and all its varieties.

Now, remember, geologists have been trying to explain in naturalistic terms the gradual changes in the Earth's surface and the successive habitats which it provides, but they have deferred discussion of the successive origin and diversity of plant and animal species revealed in the fossil record. The problem of the 'origin of species' is recognized as an inexplicable 'mystery of mysteries', and geologists and the majority of natural historians have to a large extent abandoned any attempts to explain it: as William Whewell had written, 'To this question, men of real science do not venture to return an answer'.

For Lamarck this was evidently unacceptable. Lamarck wanted to fill in the plant and animal side of the geologists' naturalistic picture – he believed that the answer to the mystery of mysteries is evolution or what was often referred to as 'transformism' – the transformation of one species into another. He was not in fact the first to develop evolutionary ideas – others who had seen evolution as the obvious solution to the problem of the origin of species included Benoit de Maillet (1656–1738), Denis Diderot (1713–84), and Erasmus Darwin (1731–1802) (Charles Darwin's uncle); and, as we have seen, Linnaeus and Buffon both discussed it. But Lamarck was the only one of these thinkers to develop a fully worked-out system of evolutionary biology, which he presented in his *Philosophie zoologique* of 1809.

Lamarck's theory of evolution

Lamarck took as his starting point the spontaneous generation of life. Now, the notion of spontaneous generation was always controversial among natural philosophers and theologians, and it was rejected by many. But the majority seem to have accepted it because the empirical evidence in its favour was overwhelming. It is not such a familiar sight to us nowadays, but in the days before refrigerators, or indeed before Louis Pasteur (1822–95) alerted us to the existence of microscopic germs, it would have been hard to avoid witnessing the phenomenon of maggots spontaneously emerging out of rotting foodstuffs. We now know that these maggots hatch from eggs laid on the foodstuff by flies, but this idea had not been established at this time, so the undeniable empirical result strongly suggested the truth of spontaneous generation of life.

Lamarck believed that the fine gradations between different kinds of living things which his teacher Buffon spoke about, could be extended back so that there was a very fine line between inanimate matter and the most rudimentary form of life. Observations of the lowest forms of living creatures (like amoebae and other single-celled organisms) led him to believe that it could generate spontaneously from gelatinous matter (water and heat was required, according to Lamarck, to enable gelatinous matter to become living matter). Note that Lamarck did not believe that maggots could spontaneously generate – they were too highly organized – only very primitive forms of life. Older traditions of spontaneous generation held that mice could spontaneously generate in household waste, and that crocodiles spontaneously generated out of fertile Nile mud, but Lamarck clearly rejected these folk beliefs.

Having begun life by spontaneous generation from warm wet gelatinous matter, Lamarck assumed that the same 'power of life' that had transformed the gelatinous matter, would continue to work, causing newly living matter to develop into successively more complex, and more advanced, forms of life. Accordingly, what Lamarck wanted to do was to establish the actual *order in time* in which all known organisms were produced. In other words, Lamarck wanted to do for living forms what contemporary geologists were doing in the case of the rocks – to establish the precise chronological order of their formation. Clearly, if Lamarck could succeed in this he would be able to bring an end to disputes about the precise linear sequence in the Chain of Being by establishing once and for all the correct order of creatures.

According to Lamarck, however, and his theory of a continuously operating 'power of life', the Great Chain of Being was not static but was a continually moving staircase bearing all creatures into higher levels of complexity. Lamarck believed that the driving force keeping this escalator moving, the power of life, was stimulated by the 'subtle fluids', such as heat and electricity, in the surrounding environment, or within the plant and animal bodies themselves.

Lamarck saw these subtle fluids as effectively shaping the internal features of organisms by a process akin to erosion in geology. The subtle fluids moving incessantly in the soft parts of the organisms, he believed,

> characteristically clear for themselves passages, settling places and outlets; they create canals and hence various organs; they vary the canals and organs either by different movements or by different fluids... they enlarge, lengthen, divide and build up these canals and organs by the materials which form and which are constantly separated out from the moving fluids.

Thinking along these lines, Lamarck eventually developed his four laws of development:

> *First law*: by virtue of life's own powers there is a constant tendency for the volume of all organic bodies to increase and for the dimensions of their parts to extend up to a limit determined by life itself.

> *Second law*: the production of new organs in animals results from newly experienced needs which persist, and from new movements which the needs give rise to and maintain.

> *Third law*: the development of organs and their faculties bears a constant relationship to the use of the organs in question.

> *Fourth law*: everything which has been acquired, or changed, in the organization of an individual during its lifetime is preserved in the

> reproductive process and is transmitted to the next generation by those who experienced the alterations.

Given the intractability of the problem of the origin of the species – its status not just as another run-of-the-mill scientific puzzle but as the mystery of mysteries – it might be supposed that Lamarck's theory would have been accepted with open arms by his contemporaries; or, at the least, that it would have been taken seriously as pointing the way to a solution to the mystery. In fact, this was not the case. Lamarck's theories were hardly given a fair hearing and tended to be dismissed out of hand.

Almost inevitably, there was no shortage of objections to Lamarck's reliance on spontaneous generation – although it wasn't until the 1870s, decades later, that spontaneous generation was disproved by Louis Pasteur (1822–95). His analogy between geological erosion and the circulation of Newtonian fluids sculpting ever more complex internal configurations in plant and animal bodies seemed much too fanciful to many. But the loudest protests were simply directed against the very idea that species could change into one another – that is to say, against the idea of evolution or transformism itself. It seems that opposition to transformism was too deeply entrenched for Lamarck's arguments to make much of an impression.

Lamarck's major critic on this score was the leading zoologist and palaeontologist, George Cuvier (1769–1832). A comparative anatomist who had gained a reputation for unequalled knowledge of animal forms, Cuvier had long since come to the view that different parts of animal bodies depended in their functions on other parts, so that an entire organism was a supremely delicate balance of perfectly adapted interacting parts. According to Cuvier, therefore, any change in one of the parts of an organism was liable to throw the whole creature out of balance, and result in its inevitable destruction. So, transformism was simply impossible.

This view of Cuvier's fitted nicely with ideas about nature as a whole which had been forcefully promoted by Linnaeus and his followers. Linnaeus had extended traditional notions from natural theology about the intricate design of individual creatures and talked about the intricate balance of all the interacting parts of the natural world. For Linnaeus, the natural world was a cyclic course of life and death in which every species fulfils its destined task in the service of the whole. Insects prevent plants from taking over the whole world, while birds prevent insects from becoming too dominant. God had ensured the smooth continuation of the system of nature by building in such checks and balances. The delicate 'balance of nature' is always maintained.

But Cuvier also undermined Lamarck's ideas by misrepresenting them. Lamarck believed that the environment affected organisms directly, so that they reacted to changes in the environment automatically (in the same way as they reacted to hunger or thirst). But it is not entirely clear that Lamarck was thinking just in terms of individuals here. His second law seems to imply a gradual process affecting many individuals over time: 'The production of new organs in animals results from newly experienced needs which persist, and from new movements which the needs give rise to and maintain'. Given the fact that even geologists who pondered the problem of the origin of species assumed that the adaptation of species to their environments was brought about in some way by the environment itself, Lamarck's second law looks, in the context of the time, unobjectionable.

According to Cuvier, however, Lamarck claimed that animals changed as a result of their 'wants' or their 'desires' – a view that was regarded as absurdly naïve (which it would have been if Lamarck had actually held it). Lamarck talks of animals' 'needs' but this does not necessarily carry the same implications. A modification of form and function brought about by a need can be interpreted in a naturalistic way (an animal finding itself in a habitat where water is scarce, for example, developing over many generations the ability to survive long periods without having to drink); but to suggest that a modification of form can take place because the animal wants it to seems fundamentally unscientific. It is important to note here that I am *not* arguing that Lamarck's ideas were correct, I am simply trying to show that Lamarck's doctrines might easily have been found scientifically acceptable by contemporary readers if they had been fairly represented by contemporary commentators. The development of a camel-like creature, say, as a result of the gradual adaptation, and transformation, of a species to a harsh environment is by no means an absurd idea.

Unfair as this criticism of Lamarck was, it proved too useful to those who wished to oppose evolution to be seen to be inadequate. It was repeated by Charles Lyell in his critique of Lamarck. Similarly, in the 1840s when Darwin developed a theory of how instinctive behaviour could be passed from parents to offspring he insisted that this was a new idea because he saw it as an *unconscious* process, whereas Lamarck (Darwin believed) had suggested this could only happen by a deliberate act of the will. Once again, Lamarck was abused by his critics. In fact, Lamarck himself had tried (evidently in vain) to emphasize the unconscious nature of the needs, dictated by the environment, which he supposed to have an effect on animal morphology.

Cuvier also brought the fossil record into play against Lamarck. In keeping with catastrophist precepts, Cuvier insisted that the record showed entire populations had been wiped out in the past, and replaced by a population of completely different plants and animals. There was no evidence of one organism changing gradually into another. Besides, Cuvier pointed out, if Lamarck was right there would be no such thing as extinct species. According to Lamarck the 'escalator' of nature was still in operation, and the 'power of life' continued to exert its influence on living creatures. Accordingly, if dinosaurs had emerged in the past as a result of the power of life transforming something lower down the escalator of nature into dinosaurs, then why do we not see dinosaurs today? Surely, dinosaurs should always be replaced on the escalator by this on-going process, and could never become extinct.

On the first point, Lamarck simply pointed to the inadequacy of the fossil record. The fossil beds that have been discovered seem to give an impression of mass extinction events in which one set of flora and fauna is replaced by a completely different set, but this could simply be a misleading impression deriving from incomplete fossil records. Ironically, this same response to fossil evidence was used by Charles Lyell in order to deprive Lamarck of what he (Lyell) saw as the palaeontological support for Lamarck's theory.

Lamarck's answer to Cuvier's second point was also perfectly reasonable. Dinosaurs may no longer be appearing on the escalator of nature, so to speak, because of wider environmental changes. A change in the environment since the age of the dinosaurs might result in the fact that lower creatures which once transformed into dinosaurs now transform into something else. Presumably, the 'something else' is better adapted for the new environment than dinosaurs and so they effectively replace dinosaurs in the continuous scale of nature. Putting it crudely, we could say that primitive animals which once evolved into dinosaurs, after a significant change in the environment, now transformed into birds (there is actually a current theory that birds evolved from certain kinds of dinosaur). So, dinosaurs would no longer be replaced on the escalator of nature, and would die out, but birds, which had never appeared on the escalator before, now seem to have an unassailable position.

Finally, of course, there were the religious objections to Lamarck's theory – determined opposition to the idea that a human being was just a transformed version of an ape. This was an idea, as we have already seen (Chapter 17), which did not just affect devout Christians, but repelled the sensibilities even of a deist like Charles Lyell.

If evolution was ever to be accepted, it needed a new approach – alas for Lamarck, he was not the man to succeed in establishing it.

FURTHER READING

Peter Bowler, *Charles Darwin: The Man and His Influence* (Oxford: Blackwell, 1990), chapter 2, pp. 17–32.

Peter Bowler, *The Fontana History of the Environmental Sciences* (London: Fontana, 1992), chapter 5.

Peter Bowler, *Evolution: The History of an Idea*, 3rd edn (Berkeley and Los Angeles, University of California Press, 2003).

L. Eiseley, *Darwin's Century* (London: Gollancz, 1959), chapters 1 and 2.

C. C. Gillispie, *Genesis and Geology* (New York: Harper & Row, 1959), chapter 6.

F. C. Haber, 'Fossils and the Idea of a Process of Time in Natural History', in B. Glass, O.Temkin and W. L. Straus, Jr. (eds), *Forerunners of Darwin, 1745–1859* (Baltimore: Johns Hopkins University Press, 1968), pp. 222–61.

D. Ospovat, 'Perfect Adaptation and Teleological Explanation: Approaches to the Problem of the History of Life in the mid-nineteenth Century', *Studies in History of Biology*, 2 (1978), pp.33–56.

M. Rudwick, *The Meaning of Fossils* (Chicago: Universaity of Chicago Press, 1976), chapter 4.

There's a good website on the history of evolutionary ideas: http://www.ucmp.berkeley.edu/history/evolution.html.

19

Religion and Progress in Victorian Britain: Hugh Miller versus Robert Chambers

Lamarck's theory of evolution failed to win many converts among the ranks of natural philosophers. His ideas were taken up to a limited extent in Britain, but not by botanists and zoologists, but by various radical political thinkers. Lamarckism suited their purposes because it seemed to provide scientific underpinning for their reformist political views.

The basic starting-point of these radical political reformers was anti-clericalism, a determined opposition to the Church, which they regard as nothing more than a propaganda machine for the old authoritarian order. Accordingly, these political radicals developed a fiercely atheistic stance. Consider, for example, the ideas expressed in the *Oracle of Reason*, a weekly broadsheet produced in Bristol in the 1840s. The epigraph on the front page of every issue read: 'Faith's empire is the World, its monarch God, its ministers the priests, its slaves the people'. It declared itself in its first issue to be 'at War not with the Church but with the Altar, not with forms of worship but with worship itself, not with the attributes but the existence of deity'. Its proud boast was that it was an 'exclusively atheistical print'.

There were a number of similar broadsheets, such as the *Free Thinkers' Information for the People*, *The Investigator*, and *The Movement*. One of the main aims of this kind of literature, and the movements behind them, was to deprive its working-class readership of what Karl Marx called 'the opium of the masses', namely religion. There is a consensus among historians today that social conditions in England in the first decades of the nineteenth century were such that a revolt of the working classes might well have occurred. The reason it didn't, they are agreed, is because of the socio-political power of the Church, especially John Wesley's Methodist movement, which was highly popular with the working classes, and emphasized that they would get their reward for enduring their tribulations in this life in the afterlife – all will be well in the sweet by and by. Marx saw this, and referred to religion as an opiate, sedating the working classes. The radical political movements also saw it, and did all they could to undermine religion and all talk of an afterlife.

Therefore, the radical reformers denied the immortality of the soul and insisted on a materialistic view of life: 'Life, or vital function, is nothing more than a certain series of phenomena matter exhibits' (*Oracle of Reason*, 1840). Similarly, the claim that everything in the world is all part of God's plan (and shouldn't be changed), that 'Whatever is', as Pope said, 'is right' (see Chapter 16), was vigorously rejected. In the *Oracle of Reason*, in 1841, the editor declaimed that:

> To say eyes were made to see is as ludicrous as to say that stones were made to break heads, legs were made to wear stockings, or sheep to have their throats cut.

Remember, Voltaire's satire against the notion that 'all is for the best, in the best of all possible worlds', in his novel *Candide*? Here, we have radical reformers decrying these same views more urgently (in the following case from the *Free Thinkers' Information for the People* in 1841):

> To the atheist, a moth in a candle's flame, or a poor fly in the fangs of a spider, is a *proof* that the world could not have been designed by one being, infinitely wise, infinitely good, and infinitely powerful.

As the editor of *The Investigator* wrote in 1842:

> All nature cries aloud against the idea of a benevolent deity. This worse than ridiculous – this vilely pernicious teaching, the atheist rejects with contempt and disgust.

Similarly, for the editor of *The Movement* (1844):

> It is the most impudent of assumptions that all is for the best in a world which savage men have deluged with crime. Why did not the gentry's God think it proper to design less suffering and more enjoyment, less hypocrisy and more sincerity...?

In order to get away from this kind of defence of the political *status quo* the political radicals dropped any ideas which preserved the role of God in nature, and turned instead to schemes such as that of Lamarck, in which matter itself has built-in powers to produce the world we see around us. William Chilton (1815–55), editor of the *Oracle of Reason*, developed an atheistic theory of matter in an editorial of 1843:

> There must be in the essence of matter a capacity, when combined in certain forms, to produce specific results. The principle of life must be essentially inherent in the whole system and every particle thereof.

The following year he referred to Lamarck's theory of progressive transformism as the theory of 'regular gradation' and argued that:

> The theory of regular gradation, or the change of one mode of natural phenomenon into another, without supernatural interference, is in

> direct opposition to the almost universally received opinions of all countries and ages... yet, stripped of religious prejudice, philosophy must admit that the inherent properties of 'dull matter'... are good and sufficient to produce all the varied, complicated, and beautiful phenomena of the universe.

The radical press went on to use Lamarckism to support their own ideas about the establishment of democracy by universal franchise – the working classes were seen as being compelled by a power or energy in nature to rise up the social scale, just as animals, according to Lamarck, were continually rising up the scale of nature.

Needless to say, these ideas were very disturbing to the defenders of the political *status quo*, which was just about everyone from the middle classes upwards. The significance of this for the history of science was that would-be biologists working in Britain now had other reasons to steer clear of Lamarckism. Because Lamarckism had been appropriated by politically subversive groups to provide naturalistic support for their socio-political views, every Victorian biologist, as well as every geologist, effectively distanced themselves from Lamarck's ideas. The only exceptions were those who sympathized with radical politics. The most notable of these was Robert Grant (1793–1874), who became professor of zoology at University College London in 1827, shortly after informally teaching Lamarckism to Charles Darwin while he was a student in Edinburgh.

But all this was only part of the story of evolutionary ideas in Victorian Britain. The idea of progress, as a kind of all-pervasive law of nature, was extremely popular with the Victorian *bourgeoisie*. The Victorian middle classes, after all, including many of the most religiously devout among them, were the very ones who were perpetuating the belief that 'a moth in a candle flame, or a poor fly in the fangs of a spider' were all part of God's plan. Survival of the fittest, as extolled in the political economy of Thomas Malthus (see Chapter 16), was seen by the Victorian middle classes as a natural law which guaranteed progress – the weak and inadequate go to the wall, while the fit and strong progressed through hard-work.

These progressionist ideas were implicit in the geologists' discussions of the role of extinct creatures in making the world suitable for subsequent life-forms. God had created the world in such a way that the habitat only gradually became suitable to support higher and higher life forms, until eventually it became suitable for humankind – progress was all part of the plan, even though that plan also entailed much suffering (a suffering that was seen as unavoidable even in a perfect world – though, of course, it was far more unavoidable to some than it was to others). Given the general ethos of progress in Victorian Britain, it is easy to see why geology books, with their confirmation of step-wide progression as a real feature of the natural world, regularly topped the best-seller lists.

The Victorian obsession with progress became explicit in a book published anonymously in 1844 under the title: *Vestiges of the Natural History of Creation*.

This was written by Robert Chambers (1802–71), one of the two brothers who founded the Chambers publishing firm in Edinburgh, but his authorship was not known until after his death. The *Vestiges* brought together the ideas of natural progression in a number of different fields and forged them together in support of a theory of evolution.

Chambers believed there must be an all-pervasive law of development, a universal principle analogous to Newton's universal principle of gravitation, and the *Vestiges* was intended to demonstrate this, not just in biology, but in cosmogony, geomorphology, and so forth. Aimed at the Victorian middle classes, it became a best-seller and effectively took evolutionary ideas away from the radicals and showed how useful it could be to the bourgeoisie. Repeatedly revised to strengthen its arguments and to respond to negative criticisms, the *Vestiges* went through four editions in seven months and by 1860 had sold 24,000 copies.

The argument of the *Vestiges of the Natural History of Creation*

The book opens with a discussion of the nebular hypothesis of Laplace (see Chapter 17). The Earth as a planet is presented as evolving out of a cloud of hot gas – the process shows increasing organization, from a chaotic gas to a planet made up of numerous different substances, manifested in a multiplicity of ways, different rocks, water, and so on. This in itself shows the inherent progressionism built into the system, but here of course the progress is the result of continually evolving changes.

Chambers then goes on to talk about the evolution of the Earth itself as it gradually cools down. He drew here upon the latest ideas in geology and palaeontology, and of course emphasized the progressionism evident from the successive life forms found in the fossil record. He was able, for example, to quote Adam Sedgwick (1785–1873), Professor of Geology at Cambridge: 'There was a time when cephalopoda were the highest type of animal life, the primates of the world... Fishes next took the lead, then Reptiles... Mammals were added next, until Nature became what she is now, by the addition of man'. For Chambers this was 'a wonderful revelation to have come upon the men of our time':

> The great fact established by it is that the organic creation, as we now see it, was not placed upon the earth at once:–it observed a PROGRESS.

Chambers differed from Sedgwick and other geologists, however, because he believed the progression of life forms was not simply a step-wise phenomenon, showing increased complexity over time but implying nothing about how the successively more complex forms actually came into existence. Chambers perhaps had the advantage of being an amateur, and an autodidact who had gleaned all he knew from reading books. But whatever the reason, he had no qualms whatsoever about putting forward an evolutionary scheme to his readers. While practising geologists and naturalists were avoiding seeing the 'elephant in the room', and were talking about the mystery of mysteries, Chambers simply accepted the obvious – the origin of species can be explained by evolution.

Accordingly, he went on to remove objections to theories of evolution. In a nice irony, he was able to undermine geological claims about the successive catastrophes revealed in the fossil record, which supposedly wiped out all life forms and therefore precluded gradual evolution, by using the arguments and the evidence put forward by Charles Lyell. This was ironic, of course, because Lyell had developed these arguments not to support evolution but in an attempt to undermine the progressionism which he now saw (having read Lamarck) as leading to transformism. Lyell had argued for the steady state of the world, insisting that the supposed progression of life forms in the fossil record was merely illusory. But now, Lyell's ideas were used to reject claims that repeated catastrophes made gradual evolution impossible, and to suggest that the endless time described by Lyell was just what was needed for evolutionary change to take place.

Chambers also drew on the work on animal forms done by Richard Owen (1804–92), Britain's leading comparative anatomist – the English Cuvier, as he was sometimes called. Owen had noticed what he called homologies – the same basic 'blueprint' used in different ways in different creatures. The 'paddles' of a whale and a seal, and the wing of a bird match the bones in our arms, the wings of a bat the bones in the human hand. The 'giraffe, with its long neck', Chambers wrote, 'has, in that part, no more bones than are to be found in the neck of an elephant or pig'. Owen saw this merely as a sign that God restricted himself in the Creation to a limited number of basic patterns (presumably a self-denying ordinance so that God could show his ingenuity). For Chambers, however, the homologies were clear evidence that creatures had evolved in different ways from a common ancestor.

Chambers wanted to insist that once God had created the cloud of hot gas and the laws of nature at the Creation, he had no need to intervene again. So, Chambers, like Lamarck before him, had to rely on the spontaneous generation of life. Here, he was able to draw on recent experiments in which living spider mites seemed to be created during electro-chemical experiments (passing electric currents through various solutions). Although controversial, these experiments, and their results, were endorsed by no less a figure than T. H. Huxley (1825–95), one of the country's leading biologists.

Finally, Chambers drew upon work in comparative embryology carried out first by Friedrich Tiedemann (1781–1861), and then confirmed in 1824 by Etienne Serres (1786–1868), which showed that the brain of mammals in the developing embryo pass successively through stages where it resembles the brain of a fish, then a bird, then a reptile, before developing the features of a mammal brain. This work had appeared in earlier evolutionary literature and, accordingly, had been discussed by Lyell in his *Principles of Geology*. Lyell acknowledged that this work showed the 'unity of plan' of all vertebrates but insisted that it offered

226 HYPOTHESIS OF THE DEVELOPMENT OF THE VEGETABLE AND ANIMAL KINGDOMS. 227

SCALE OF ANIMAL KINGDOM (The numbers indicate orders:)		ORDER OF ANIMALS IN ASCENDING SERIES OF ROCKS		FŒTAL HUMAN BRAIN RESEMBLES, IN
			1 Gneiss and Mica Slate system	
RADIATA (1, 2, 3, 4, 5)		Zoophyta; Polypiaria	2 Clay Slate and Grawacke system	
MOLLUSCA (6, 7, 8, 9, 10, 11)		Conchifera; Double-shelled Mollusks	3 Silurian system	1st month, that of an invertebrated animal;
ARTICULATA	*Annelida* (12, 13, 14)	Crustacea		
	Crustacea (15, 16, 17, 18, 19, 20)	Annelida		
	Arachnida & Insecta (21—31)	Crustaceous Fishes	4 Old Red Sandstone	
VERTEBRATA	*Pisces* (32, 33, 34, 35, 36)	True Fishes	5 Carboniferous formation	2nd month, that of a fish;
	Reptilia (37, 38, 39, 40)	Piscine Saurians (ichthyosaurus &; Pterodactyles; Crocodiles; Tortoises; Batrachians	6 New Red Sandstone	3rd month, that of a turtle;
	Aves (41, 42, 43, 44, 45, 46)	Birds		4th month, that of a bird;
	Mammalia: 47 Cetacea; 48 Ruminantia	(Bone of a marsupial animal)	7 Oolite	
	49 Pachydermata	Pachydermata (tapirs, horses, &c	8 Cretaceous formation	
	50 Edentata			
	51 Rodentia	Rodentia (dormouse, squirrel, &c	9 Lower Eocene	5th month, that of a rodent;
	52 Marsupialia	Marsupialia (racoon, opossum, &c		6th month, that of a ruminant;
	53 Amphibia			
	54 Digitigrada	Digitigrada (genette, fox, wolf &c		7th month, that of a digitigrade animal;
	55 Plantigrada	Plantigrada (bear)	10 Miocene	
		Cetacea (lamantins, seals, wha		
	56 Insectivora	Edentata (sloths, &c.)		
		Ruminantia (oxen, deer, &c.)	11 Pliocene	
	57 Cheiroptera			
	58 Quadrumana	Quadrumana (monkeys)		8th month, that of the quadrumana;
	59 Bimana	Bimana (man)	12 Superficial deposits	9th month, attains full human character.

Figure 19.1 Table showing the history of progress in different aspects of the natural world; from Robert Chambers, *Vestiges of the Natural History of Creation* (London, 1844).

The 'Scale of Animal Kingdom' is Cuvier's, while the other columns show 'the wonderful parity' between that scale and the succession of fossils, and the 'foetal progress' of the brain.

'no support whatever to the notion of a gradual transmutation of one species into another'. Chambers begged to differ (Figure 19.1).

This covers the technical aspects of Chambers book, but there was another important ingredient. It was entirely representative of the tradition of what has been called 'Cosmic Toryism' (see Chapter 16) – the belief that all is for the best in the best of all possible worlds, which as we've seen was being decried by the political radicals. Chambers didn't actually chance upon the notion of survival of the fittest, but he came pretty close:

> Knowing that much evil does unavoidably befall us from no fault of ours, we are apt to feel that this is a dreary view of the Divine economy… But it may be that, while we are committed to take our chance in a natural system of undeviating operation, and are left with apparent ruthlessness to endure the consequences of every collision into which we knowingly or unknowingly come with each law of the system, there is a system of Mercy and Grace behind the screen of

> Nature... It is necessary to suppose that the present system is but a part of the whole, a stage in a Great Progress.

And elsewhere:

> Everywhere we see the arrangements for the species perfect; the individual is left, as it were to take his chance against the mêlée of the various laws affecting him. If he be inferiorly endowed, or ill befalls him, there was at least no partiality against him. The system has the fairness of a lottery, in which every one has the like chance of drawing the prize.

The *Vestiges of the Natural History of Creation* (which title, by the way, alludes to, and rejects, the famous remark of James Hutton, that there is no evidence left from which we can discover the very origins of things: 'we find no vestige of a beginning') became a runaway best-seller in Victorian Britain. The book-buying middle classes clearly lapped it all up. In 1845 Prince Albert (1819–61) read it aloud to Queen Victoria (1819–1901). It can be seen, therefore, as a book which set the seal on the tradition of deistic natural theology which we've traced from Alexander Pope's *Essay on Man*, through Adam Smith's *Wealth of Nations* and *laissez-faire* political economy to Thomas Malthus's *Essay on the Principle of Population*.

By showing how the law of development could be seen at work in the formation of the solar system, the formation of the topography of the Earth, the origins of life, and the subsequent development of plants and animals, up to humankind, Chambers seemed to establish beyond any doubt that progress really was built into the system. Accordingly, Victorian *laissez-faire* political economy, and its associated moral values, in which poverty and suffering was unavoidable for the lower orders of society, was seen once again to be all part of God's plan, and was confirmed as the *natural* social and political schema, from which it was futile, or dangerous, to deviate.

Nowadays , the picture that is usually painted for us in popular history, is that Darwin's *Origin of Species* (1859) came as a great shock to the Victorian public and was condemned on all sides. This is simply not true. Darwin's book sold out within hours of its appearance in book shops. It was already rumoured to be even better than the earlier best-seller, *Vestiges*. Again, the public wanted to lap it up.

Furthermore, the Anglican Church was not opposed to it either. Thomas Malthus, who gave Darwin the idea for the principle of natural selection and survival of the fittest, was an Anglican clergyman, and the tradition that this was the best of all possible worlds was first developed by theologians seeking to solve the so-called 'problem of evil' – why there is suffering in God's creation. The question arises, therefore, as to why we have the impression that the Church was opposed to Darwin's theory?

It is because, in our secular world, historians and other commentators have chosen to lump all religions together, and dismiss all their followers as credulous believers in mumbo-jumbo and incapable of accepting scientific truths. The fact is that the kind of evolutionary ideas put forward by Chambers (and later Darwin) did indeed meet with religious opposition. This opposition came only from certain religious groups, however, while other groups were happy to accept evolutionary theories. Secular commentators, uninterested in distinctions between different religious groups, have tended to let the religious opponents of evolution stand representative of *the* religious reaction to evolution. Apart from anything else, by suggesting that 'religion' itself was (and supposedly still is) opposed to evolution, they can go on to use this to suggest that 'religion' is opposed to 'science', and use this to confirm their view that religious believers are irrational, ill-educated, downright foolish, and so on.

It is perhaps worth pointing out that historians are now generally agreed that the Victorian age was the first age of 'mass' atheism. If atheists were not quite in the majority in Victorian Britain there were certainly vast numbers of them. So, many of the readers who bought *Vestiges*, or Darwin's *Origin of Species* would have been atheists who did not care about the religious implications of evolution. Furthermore, given that both the *Vestiges* and the *Origin* emerged, at least in part, out of the Anglican tradition of natural theology (of which Malthus's *Principle of Population* was a part), the majority of religious believers had little or no reason to oppose it. In fact, those religious believers who *did* oppose evolution were in the minority, not just when compared with Victorian society as a whole, but also when compared with other religious believers.

Let us consider, by way of example, just one of Chambers's most vigorous critics, who did object to evolutionary ideas on religious grounds. This was the former stonemason, turned popular science writer, Hugh Miller (1802–56). Miller was a leading member of the Free Presbyterian Church, which broke away from the Church of Scotland in 1843, when the Church of Scotland became for the first time the *established* Church, and so, like the Anglican Church in England, became effectively an arm of government (or so it certainly seemed to opponents of establishment).

Members of the breakaway Free Church of Scotland were concerned not to be seen as affiliated to the government in Britain, but wanted to remain free to criticize and oppose the government. In Miller's case there can be no doubt that one of the reasons he opposed government was because he was opposed to *laissez-faire* economics and all that went with it. As editor of *The Witness*, the leading newspaper in Scotland at that time, Miller was a tireless campaigner against social injustice, and incessantly called for government aid to the poor.

Furthermore, Miller simply did not accept the theological tradition which subscribed to the view that there was no point changing the world because God had created the best one possible – even though it was a world where suffering for many was unavoidable. His response to the problem of evil was simply to say that we cannot understand God's mind, and therefore cannot know why there is suffering in the world. Miller liked to repeat in his own publications one of his favourite lines from the poet William Cowper (1731–1800): 'God moves in mysterious ways...' But if we cannot understand why there is suffering, what we can do, he believed, is try to alleviate the suffering of our fellow humans.

Miller's rejection of the rationalist solution to the problem of evil is clear from his attack on *Vestiges* in his *Footprints of the Creator* (1849) – the title of which alludes to a line from Francis Bacon's *Great Instauration*. Miller not only attacks Chambers's palaeontological claims (Miller's own area of expertise) but he also explicitly states that we cannot know 'why it is that evil exists in the universe of the All-wise and All-powerful'. What makes Chambers's 'developmental hypothesis', as Miller calls it, unacceptable is the fact that it condemns many to be 'lost and degraded creatures', just because that is the way things are. Miller rejects Chambers's excuse for political inactivity which he quotes from Chambers (but which could have come straight out of Thomas Malthus): 'It is hard for the sufferer, but what can we say against the course of nature?' Miller insisted that suffering should be alleviated by 'moral endeavour' and 'unshaken faith', and advocated numerous political reforms.

But the fact is, Miller was *not* a representative Victorian thinker – his kind of religious belief was already becoming a minority view. It was a minority view not just because atheism was becoming increasingly embraced by the mass of Victorian society, but because most religious believers did not subscribe to Miller's views about the power of moral endeavour, but preferred to believe in a God who had supposedly made the best of all possible worlds, including the unavoidable suffering of the poor. Miller, a would-be reformer and champion of the poor, and opponent of evolution, out of step with the dominant vicious Victorian values, committed suicide on 23 December, 1856.

FURTHER READING

Peter Bowler, *The Invention of Progress: The Victorians and the Past* (Oxford: Backwell, 1989).

Peter Bowler, *The Fontana History of the Environmental Sciences* (London: Fontana Press, 1992), chapter 8, pp. 306–37.

Peter Bowler, *Evolution: The History of an Idea*, 3rd edn (Berkeley and Los Angeles, University of California Press, 2003).

Adrian Desmond, *The Politics of Evolution: Morphology, Medicine, and Reform in Radical London* (Chicago: University of Chicago Press, 1992).

Loren Eisley, *Darwin's Century* (London: Gollancz, 1959), chapter 4, pp. 91–108, and chapter 5, pp. 117–40.

John Henry, 'Palaeontology and Theodicy: Religion, Politics and the Asterolepis of Stromness', in M. Shortland (ed.), Hugh Miller and the Controversies of Victorian Science (Oxford: Clarendon Press, 1996), pp. 151–70. (The full text of this is available on the Web at: http://www.ssu.sps.ed.ac.uk/research/henry/henry_palae.html.)

Arthur Lovejoy, 'The Argument for Organic Evolution before the Origin of Species, 1830–1858', in Glass, *et al.*, *Forerunners of Darwin, 1745–1859* (Baltimore: Johns Hopkins University Press, 1968), chapter 13, pp. 356–414.

James A. Secord, *Victorian Sensation: The Extraordinary Publication, Reception, and Secret Authorship of* Vestiges of the Natural History of Creation (Chicago: University of Chicago Press, 2000).

20

Bringing it All Together? Charles Darwin's Evolution

Charles Darwin (1809–82) began his higher education in the Edinburgh Medical School in 1825, and in 1826 his interest in natural history was noticed, and cultivated, by Robert Grant (1793–1874), one of the extramural teachers in Edinburgh who made their living by offering additional tuition to students in the Medical School. Grant was a convinced Lamarckian and, like Lamarck, was a specialist on marine invertebrates. Grant introduced Darwin to the study of these humble creatures, showing him techniques for dissecting them. Grant helped Darwin think in terms of generation, both of individuals and species, and introduced him to continental thinking on the relevance of embryological anatomy to the laws of life. Darwin always remembered Grant's admiration of Lamarck and returned to an examination of Grant's ideas in his own transmutation notebooks of 1836–7. Furthermore, Darwin later made a reputation for himself as the leading expert on barnacles, a marine invertebrate he had first studied with Grant.

Darwin was training to be a doctor at Edinburgh but pulled out because he couldn't stand the sight of blood. He then went to Cambridge to train for the Anglican ministry. It was in Cambridge that he also became thoroughly steeped in the *laissez-faire*, natural theology tradition. One of Darwin's major sources here was William Paley (1743–1805), whose textbooks formed the mainstay of the Cambridge curriculum. As Darwin wrote in his *Autobiography* (posthumously published in 1892):

> In order to pass the BA examination, it was also necessary to get up Paley's *Evidences of Christianity*, and his *Moral Philosophy*. This was done in a thorough manner, and I am convinced that I could have written out the whole of the *Evidences* with perfect correctness, but not of course in the clear language of Paley. The logic of this book, and, as I may add, of his *Natural Theology*, gave me as much delight as did Euclid. The careful study of these works, without attempting to learn any part by rote, was the only part of the academical course which, as I then felt, and as I still believe, was of the least use to me in the education of my mind. I did not at the time trouble myself about Paley's premises; and taking these on trust, I was charmed and convinced by the long line of argumentation. By answering well the examination

> questions in Paley, by doing Euclid well, and by not failing miserably in Classics, I gained a good place among the *Hoi polloi* or crowd of men who do not go in for honours.

Paley explicitly drew attention to the work of Thomas Malthus and the notion of the struggle for survival in chapter 26 of his *Natural Theology* (1802):

> Mankind will in every country breed up to a certain point of distress. That point may be different in different countries or ages... but there must always be such a point, and the species will always breed up to it. The order of generation proceeds by something like a geometrical progression. The increase of provision, under circumstances even the most advantageous, can only assume the form of an arithmetic series. Whence it follows, that the population will always overtake the provision, will pass beyond the line of plenty, and will continue to increase, till checked by the difficulty of procuring subsistence... Such difficulty therefore, along with its attendant circumstances must be found in every country.

Paley refers the reader to Malthus in a footnote, and goes on to link these ideas to the traditional response to the problem of evil, namely that all is for the best, in the best of all possible worlds.

Darwin did sufficiently well at Cambridge that J. S. Henslow (1796–1861), the professor of botany at Cambridge, recommended him as a gentleman companion for the Captain of a Royal Navy ship embarking on a protracted government expedition. The appointment was also to offer Darwin scope, as Henslow wrote, for 'collecting, observing, & noting any thing worthy to be noted in Natural History'. Darwin was approved by the Captain, and found himself engaged upon a five-year voyage around the world on *HMS Beagle* (1831–36).

Captain Fitzroy gave Darwin a copy of Volume I of Lyell's *Principles of Geology*, which had just appeared. Darwin received the later volumes by post during the voyage. Lyell, as we have seen, was opposed to evolution, and during his time on the Beagle Darwin did not dissent, although he unwittingly gathered evidence that was later to change his mind. Meanwhile, Lyell's *Principles* offered a number of congenial ideas for Darwin's developing thought. As a result of reading the *Principles of Geology* Darwin came to accept the following precepts:

1. All phenomena must be explained in terms of vigorous scientific naturalism.
2. The age of the Earth is effectively unlimited (the Huttonian view of time, revived by Lyell).
3. The environment is continually, slowly changing (Lyell's uniformitarian view, opposed to catastrophic interruptions in the historical development of the Earth).

4. The 'species problem' (i.e. what is the origin of new species in the fossil record?) must be solved in terms of:
   a) Providential adaptation (that organisms are suited for, or adapted to, the environment in which they live);
   b) environmental determinism (that the environment somehow drives the adaptations).

Early in 1837, the year after his return to England, Darwin's four specimens of mockingbirds, which he'd collected on separate islands in the Galapagos archipelago, were declared by John Gould (1804–81), a specialist on birds, to be separate species. Gould also pointed out that a collection of twelve birds which Darwin thought were of different kinds, were all finches. Darwin was then able to establish (with some difficulty) that they too consisted of distinct species specific to different islands in the Galapagos. So, here were a number of species of finch which were close to one another, and close to finches on the South American mainland (600 miles away), and yet were undeniably different. The islands were all alike in terms of habitat, weather, etc. and so environmental determinism could not explain all the differences. But direct creation by God was also ruled out – why would God create different species on adjacent islands, when the same species could have thrived on any of the islands? Darwin now became convinced that the only explanation for this kind of speciation was evolution, or what he called 'descent with modification'. As he wrote in the *Origin of Species* (1859):

> Seeing the gradation and structure in one small, intimately related group of birds, one might really fancy that, from an original paucity of birds in this archipelago, one species had been taken and modified for different ends.

Primed perhaps by the work of his uncle, Erasmus Darwin, who had written in his *Zoonomia* of 1794 that every creature has a 'faculty of continuing to improve by its own inherent activity, and of delivering down those improvements by generation to its posterity, world without end!', and by the long hours he'd spent working with Grant, Darwin was now able to see past Lyell's negative criticisms of evolution and to embrace it as the most likely solution to the problem of the 'origin of species'.

From early 1837, therefore, Darwin became a convinced evolutionist and set out to establish its truth. He began working, as he said in his *Autobiography* (see Chapter 16), in a Baconian way – gathering facts without any preconceived theory. But he soon stumbles across the mechanism he needs, which he calls natural selection, analogous to the artificial selection used by breeders, to make new varieties of pigeons, dogs, roses, etc. This idea came to him as a result of reading Malthus's *Principle of Population*, and indicates the way Darwin had thoroughly absorbed the natural theology tradition in which even the suffering in the world is supposedly for the good of the whole system:

> I worked on true Baconian principles and without any theory collected facts on a wholesale scale, more especially with respect to domesticated productions, by printed enquiries, by conversation with skilful breeders and gardeners, and by extensive reading... In October 1838, that is, 15 months after I had begun my systematic enquiry, I happened to read for amusement, Malthus on *Population*, and being well prepared to appreciate the struggle for existence which everywhere goes on from long-continued observation of the habit of animals and plants, it at once struck me that under these circumstances favourable variations would tend to be preserved, and unfavourable ones to be destroyed. The result of this would be the formation of new species. Here, then, I had at last got a theory by which to work.

But Darwin didn't publish right away. Knowing from his earlier work in geology, and from Lyell, and no doubt from the negative response to Lamarck, that evolution was regarded as a hopeless position to defend, he decided to build up as strong a case as he possibly could. He felt he was getting close to being able to publish in 1844, when the *Vestiges* hit the book shops.

Now, although the anonymous *Vestiges* was a 'sensation', proving extremely popular with a significant proportion of the general public, it was howled down by nearly all scientists. The problem was that the work was evidently written by an author who was not a practising scientist. There was no hint of original research in the book, rather it was a clever compilation of doctrines and ideas taken from various different scientists and combined in a way that was deemed by scientific critics to be invalid, or at least inappropriate. The implication was that the relevant specialists knew that their work could not be used to support evolution, and only a rank amateur with a shaky grasp on the significance of the work he appropriated could have abused their work in this way. The fact is, however, that the scientific work which Chambers drew upon did point to evolution, but trained scientists had been so indoctrinated with the idea that evolution was a thoroughly unscientific doctrine, that they could never see it.

For Darwin, watching from the wings, the response to the *Vestiges* was alarming. Darwin believed that he himself was doing the same thing that the author of the *Vestiges* was accused of: he was simply compiling results from other people and putting them together in support of evolutionism. As Darwin's botanist friend, Joseph Dalton Hooker (1817–1911), had written to him, only one who had studied many species was qualified to discuss their origin. Accordingly, Darwin decided that, before going any further with his book on natural selection, he must make himself an undisputed expert on some aspect of natural history. It was at this point in his career, therefore, that he took up the study of barnacles which was to engage him for the next eight years, dissecting and describing all known species.

Darwin finally came to write his *Origin of Species*, as the result of what he described as a 'bolt from the blue' in 1858. Darwin received a letter from a naturalist working in the Malay archipelago by the name of Alfred Russel Wallace

(1823–1913). In this letter, Wallace asked Darwin's opinion of his solution to the 'species problem', which he then outlined for Darwin. Wallace's solution was the principle of natural selection – Darwin's own idea, which Wallace had arrived at independently. Interestingly, Wallace had arrived at the idea while he was convalescing after a bout of malaria. As he lay in bed he thought about the implications of a book he'd read some time before. The book was Thomas Malthus's *Essay on the Principle of Population*.

This might seem like an extraordinary coincidence, but in fact it was not so remarkable. Darwin and Wallace were not the only ones to come up with the principle of natural selection. It had been suggested in an article published in the *Transactions of the Royal Society* in 1813 by a physician called Charles Wells (1757–1817), and it had also been stated by a fruit grower called Patrick Matthew (1790–1874) in 1831, in a book *On Naval Timber and Arboriculture* (about growing trees to keep the navy supplied with ships).

Matthew's comment on Darwin's *Origin of Species* is revealing – it shows that the basic premise was indeed a typically Victorian way of thinking, and that the Malthusian ethos really was deeply embedded in Victorian values. As Matthew wrote of the principle of natural selection in a letter to the *Gardener's Chronicle* in 1860:

> To me the conception of this Law of Nature came intuitively as a self-evident fact, almost without an effort of concentrated thought. Mr Darwin seems to have more merit in the discovery than I have had – to me it did not appear a discovery. He seems to have worked it out by inductive reason... While with me it was by a general glance at the scheme of Nature that I estimated this select production of species as an *a priori* recognizable fact – an axiom, requiring only to be pointed out to be admitted by unprejudiced minds of sufficient grasp.

No wonder that T. H. Huxley, who had previously been dead-set against evolutionary theories (he dismissed the *Vestiges* in a very savage review), said when he read the *Origin of Species*: 'How stupid not to have thought of it before'.

Anyway, the 'bolt from the blue' forced Darwin's hand. After negotiations between Darwin's friends and Wallace, a 'joint' paper by Darwin and Wallace was cobbled together and read before the Linnaean Society in London, and then Wallace agreed to wait until after Darwin published before he said anything else on the matter. Darwin then dropped the massive book called *Natural Selection* that he was writing, and hurriedly prepared for publication instead, the *Origin of Species*.

The scientific response to Darwin's *Origin of Species*

What was the response to Darwin's theory from his fellow scientists? We've seen that evolutionary theories were regarded with little or no respect by geologists and most naturalists. Lamarck's theories were misrepresented to make it easy to dismiss his theories, and Chambers's *Vestiges* was vilified by scientists. Did Darwin get the same treatment?

No. Darwin's reputation as a careful research worker and a committed scientist (not some easily ignored amateur outsider like Chambers) – Darwin was by now recognized as the world's leading authority on barnacles – meant that evolution was now taken seriously for the first time by those working in the biological sciences. Others began to take it up and to research different aspects of the natural world to test, or extend the theory. As a result, the theory of evolution went from strength to strength. But Darwin did not have things all his own way. Although the theory of evolution began to thrive as never before, Darwin's own unique

Box 20.1 SCIENTIFIC OBJECTIONS TO DARWIN'S ORIGIN OF SPECIES FOLLOWING SHORTLY AFTER ITS PUBLICATION IN 1859

1. Variations in Darwin's scheme are simply said to happen – this is the chance element in his theory. Sir John Herschel objected to this 'law of higgledy-piggledy'. Darwin couldn't explain why variations (favourable and unfavourable) take place – this is one reason why he needs a very long time-scale. Even so, this wasn't too devastating a criticism because he was able to say, as Newton had of gravity, we know variation is a fact (like gravity), even if we can't explain it. Nobody is a precise replica of either of their parents, there is always some variation.
2. The fossil record was used first of all to re-iterate the catastrophist worldview, and to insist that long periods of uninterrupted evolutionary development had never occurred. Later, with the discovery of the fossil remains of archaeopteryx – a link between reptiles and birds – and the reconstruction of the evolution of the horse from small dog-like *eohippus* to modern *equus*, the fossil record was used to argue that there was a directionality in evolution, suggesting there is some positive guiding force at work, not just natural selection. T. H. Huxley regarded the series of fossils in the genealogy of the horse as 'demonstrative evidence of evolution' but he meant progressionist evolution, not just evolution based on natural selection. Darwin in the first instance pleaded that the fossil record was incomplete and so misleading. In the second instance he could only insist that natural selection was in itself a progressionist factor.
3. William Thomson, Lord Kelvin (1824–1907), leading physicist and pioneer thermodynamicist, calculated the age of the Earth on the assumption that it was cooling down from a hot ball thrown out of the Sun. He concluded that the Earth could only have had a solid crust for the past 100 million years. This was devastating for Darwin's theory which required much longer time-scales. Darwin claimed that the most recent rocks in England, the chalk downs of the South-East, were 300 million years old. The laws of physics at this time were regarded as virtually infallible and so most believed Darwin's theory must be wrong, and insisted that this proved there was a more positive, purposeful force at work in evolution to drive it along quicker than Darwin's natural selection allowed.

contribution to evolutionary theory – the principle of natural selection – did not fare so well.

There is great irony in the fact that evolution, which had been long rejected, was immediately accepted as a result of Darwin's *Origin of Species*, but Darwin's mechanism for explaining how evolution took place was regarded as essentially inadequate to account for evolution. The problem was that natural selection was seen as only a *negative* factor. It might account for extinctions – for the extermination of animals and plants which fail to adapt to changing environments, but it was felt that it could not explain the appearance of new adaptations, which enable a creature to thrive better than before, and to transform that creature into a new, never-before-seen kind (see Box 20.1).

Darwin was vindicated in 1906 when John William Strutt, Lord Rayleigh (1842–1919), discovered radioactivity and it was realized that the Earth had a radioactive core and so generated its own heat. This threw out all calculations based on the assumption that the Earth was simply passively cooling down. Clearly, the Earth could be as old as Darwin wanted it to be. Some people use this as an argument to show what a great genius Darwin was, but given that Darwin couldn't have known he was going to be vindicated in this way, most contemporaries were justified in seeing his intransigence not as a sign of genius but as the sign of a sadly closed and foolish mind.

4 Darwin himself acknowledged the greatest problem to be how favourable variations can be passed on to offspring. Heredity at this time was based on the assumption that characteristics of both parents were blended in the offspring. Fleeming Jenkin (1833–85), soon to be professor of engineering at Edinburgh University, in a review of the *Origin* published in 1867 said that a favourable variation would soon be lost in a breeding population because it would be like dropping a single drop of black paint into a bucket of white paint. Darwin himself retreated towards Lamarckism in response to this – arguing that perhaps variations occurred in many individuals, who could breed amongst themselves and pass on the variation. The clear implication of this was that there is some guiding force directing evolution.

5 Another objection was that the slow evolution of certain organs didn't seem workable in a nature where survival was always a struggle. The classic example is the bat's wing, consisting of a membrane stretched out over grossly elongated fingers. Before this became a serviceable wing, wouldn't the proto-bat be fatally hampered by its elongating fingers? Can an eye be useful without eyelids, tear glands, etc? Surely an eye and a bat's wing both show that evolution operates in a positive way, developing something for almost immediate use, and advantage? Darwin's slow process can't be correct. Again, Darwin acknowledged here that there might be some Lamarckian power at work, speeding things up.

In spite of all these objections, critics didn't reject evolution, which was now strongly entrenched. But they tended to reject the Darwinian view of evolution based solely on the exceedingly slow process of natural selection.

This latter, highly important, aspect of evolution seemed to require a more *positive* force at work than natural selection, to ensure progression. Without ever acknowledging it, post-Darwinian evolutionists seemed to be embracing a Lamarckian , or a *Vestiges*-like, picture of evolution, in which there was a positive force at work, a kind of 'power of life', or 'law of development', pushing creatures up the escalator of nature. Once again, this can be seen as a triumph for the progressionist views of Victorian Britain, sweeping Darwin's natural selection aside, and demanding something to explain the perceived evidence for progressionism, and thereby to justify its moral and political – its Malthusian – implications.

Furthermore, it is important to note that even Darwin could not help thinking the same way – that is, that he too subscribed to a view of progress built into the natural world. In modern Darwinist theory the supposed *lack* of progressionism in Darwin's theory marks it out as superior to rival theories such as those of Lamarck and Chambers. Consider, for example, this quotation from one of the most prominent of recent Darwinian theorists, the late Stephen Jay Gould, writing in 1995:

> The most serious and pervasive of all misconceptions about evolution equates the concept with some notion of progress, usually inherent and predictable, and leading to a human pinnacle. Yet neither evolutionary theory nor life's actual fossil record supports such an idea. Darwinian natural selection only produces adaptation to changing local environments, not any global scheme of progress.

Accordingly, a number of commentators have even tried to argue that Darwin himself was not a progressionist. That he alone stood apart, a great genius transcending the ethos of his own times, while the foolish Victorian thinkers who surrounded him remained obsessed with the notion of progress. But Darwin was as much a Victorian as those with whom he lived and worked. What separated him from most of his contemporaries, was not a different attitude to progress – he was as much obsessed with the idea of progress as any of his contemporaries – but his faith in the *positive* power of natural selection to enact progress. This can be seen quite clearly in the *Origin of Species*:

> The more recent forms must, on my theory, be higher than the more ancient; for each new species is formed by having had some advantage in the struggle for life over other and preceding forms.

When he says 'higher', Darwin clearly means more advanced, more perfected. Clearly, life forms are getting better as natural selection proceeds. Consider also the wonderful closing words of the *Origin*: in case you miss them, I have put the crucial phrases in italics. Note also that, at the end of the first paragraph he says natural selection works 'for the good of each being'. Plainly, it cannot be supposed that natural selection worked for the good of species that were rendered extinct. Darwin was only considering those beings which are the beneficiaries of the process so far – most notably mankind. There are hints here that this is the best of all possible worlds and if natural selection has *not* acted for the good of some specific beings, then those beings simply do not count (they were only created for the subsequent benefit of thriving species). For Darwin, therefore, natural selection is a force guaranteeing progress:

> As all the living forms of life are the lineal descendants of those which lived long before the Silurian epoch, we may feel certain that the ordinary succession by generation has never once been broken, and that no cataclysm has desolated the whole world. Hence we may look with some confidence to a secure future of equally unappreciable length. And as natural selection works solely by and for the good of each being, all corporeal and mental endowments will tend *to progress towards perfection.*
>
> It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent on each other in so complex a manner, have all been produced by laws acting around us. These laws taken in the largest sense, being Growth with Reproduction; Inheritance which is almost implied by reproduction; Variability from the direct and indirect action of the external conditions of life, and from use and disuse; a Ratio of Increase so high as to lead to a Struggle for Life, and as a consequence to Natural Selection, entailing Divergence of Character and the Extinction of *less-improved forms*. Thus from the war of nature, from famine and death, *the most exalted object* which we are capable of conceiving, namely, the *production of the higher animals*, directly follows. There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless *forms most beautiful and most wonderful have been, and are being, evolved.*

It seems perfectly clear, therefore, that Darwin, like any other Victorian gentleman, believed in progress and that the natural world had a built-in capacity for self-improvement. He seems to have been in a minority of one, however, in believing that natural selection was sufficient on its own to bring about that improvement. For virtually all of his contemporaries, natural selection was too negative a principle to have positive effects, and so must be supplemented by a positive force of progress, or improvement.

FURTHER READING

Peter J. Bowler, *The Fontana History of the Environmental Sciences* (London: Fontana, 1992), chapter 8, pp. 306–78.

Peter J. Bowler, *Charles Darwin, The Man and His Influence* (Oxford: Blackwell, 1990 and Cambridge: Cambridge University Press, 1996).

P. Bowler, *Evolution: The History of an Idea*, 3rd edn (Berkeley: University of California Press, 2003), chapters 6 and 7.

W. Coleman, *Biology in the Nineteenth Century* (Cambridge: Cambridge University Press, 1977), Ch.4.

Stephen J. Gould, 'Ladders and Cones: Constraining Evolution by Canonical Icons', in Robert B. Silvers (ed.), *Hidden Histories of Science* (New York: New York Review of Books, 1995), pp. 37–69.

J. C. Greene, *The Death of Adam* (Ames: Iowa State University Press, 1959), chapters 5 and 9.

E. Manier, *The Young Darwin and his Cultural Circle* (Dordrecht: Reidel, 1978), pp. 29–31 and 64–6.

D. R. Oldroyd, *Darwinian Impacts* (Milton Keynes: Open University Press, 1980), chapters 7 and 10.

Robert M. Young, *Darwin's Metaphor: Nature's Place in Victorian Culture* (Cambridge: Cambridge University Press, 1985) (the complete text of this is available on-line at: http://www.shef.ac.uk/~psysc/darwin/dar.html).

21

Darwinian Aftermaths: Religion; Social Science; Biology

According to Thomas Henry Huxley, writing in 1887:

> The most potent instrument for the extension of the realm of natural knowledge that has come into men's hands since the publication of Newton's *Principia* is Darwin's *Origin of Species*.

The subsequent history of the biological sciences makes it easy to agree with him, even in the light of more recent developments in physics. Darwin's *Origin* is one of the most significant texts in the history of science and it had a huge impact, becoming a major resource for subsequent thinkers in a number of different fields. We will consider here how it was taken up in three major areas of contemporary and subsequent culture; religion, the social sciences, and biology.

The religious reaction

In spite of a prevailing image that Darwinism was inimical to religious belief *tout court*, there was no unified religious reaction to Darwin's theory, any more than there had been to earlier attempts to promote evolutionary theories. The Roman Catholic Church made no official pronouncement, and the various factions of Protestantism took different lines, varying from no reaction to a vigorous denunciation. We saw at the end of Chapter 19 that although those factions which were opposed to evolution were in the minority, they have tended to be seen as representative of the general response by modern secular historians who have little or no interest in theological distinctions among Christians, and in some cases have a vested interest in presenting religious belief in general as inimical to science.

If we consider the response of the official, established, Church of England, for example, we'll see that Darwin's *Origin of Species* was by no means seen to be the devastating bombshell that subsequent generations of secularist scientists and historians have gleefully assumed it to be. Ecclesiastical historians, by contrast, have suggested that Darwinism was very much a side issue as far as the leaders of the Anglican Church were concerned. Anglicanism was confronted by much more serious threats to its traditions.

The first thing to note is that the Victorian Age is generally acknowledged by historians to be the first age of mass atheism. The Church was already struggling to be taken seriously in a society in which, as the poet Matthew Arnold perceptively wrote, 'The Sea of Faith', once at the full, was now retreating in a 'melancholy, long, withdrawing roar'. For many other Victorians, as for Arnold, the world 'Hath really neither joy, nor love, nor light,/Nor certitude, nor peace, nor help for pain'.

But the Church was also racked by internal troubles. The Bible itself had been subjected to the analysis of philologers, studying the linguistics of the Bible, and historians, studying the authenticity of its accounts, since the early eighteenth century. A French philologist, Jean Astruc (1694–1766), had demonstrated in 1753, for example, that the book of Genesis, supposedly dictated by Moses, had actually been compiled from two separate, philologically distinct, sources (which are now designated by scholars as source J and source E).

Following on from this kind of scholarship, the Bible's standing as a historical source began to be scrutinized. In 1784 the German philosopher G. E. Lessing (1729–81) wrote *A New Hypothesis concerning the Evangelists considered simply as Human Historians*. Perhaps the culmination of this trend was *The Life of Jesus Critically Examined*, written by the German scholar D. F. Strauss (1808–74) in 1836, but translated into English in 1846, by Mary Anne Evans (1819–80), better known as George Eliot. Strauss established that the gospels were all written long after the death of Jesus, the earliest at the beginning of the second century AD.

In view of this serious scholarship on the Bible, the more learned and intellectual members of the Church of England were already inclining to the view that the Bible was irrelevant to any debates about the physical world. It was clearly a text written by ordinary men, who were ignorant of science, no matter how inspired they were by divine grace.

There were still many Anglican clergymen, however, who were unaware of these developments in Biblical scholarship. The cat was let out of the bag, however, in 1860 with the publication of a book under the snappy title, *Essays and Reviews*. This consisted of seven articles by seven different authors (dubbed by some appalled conservative thinkers as *septem contra Christum* – seven against Christ), each addressing a different aspect of new religious scholarship. The authors included Frederick Temple (1821–1902), then headmaster of Rugby School, and later to be Archbishop of Canterbury, Benjamin Jowett (1817–1893), Master of Balliol College Oxford, and Baden Powell (1796–1860), professor of geometry at Oxford and father of Robert Baden-Powell who later founded the Boy Scout movement. Baden Powell took as his theme the relationship between current science and religion, including Darwin's theory, which he thoroughly endorsed.

The book caused a sensation among devout believers but not because of the endorsement of Darwin, but because of its acceptance of the latest Biblical scholarship, with its implication that the Bible was not the direct word of God. If there was a crisis of faith in Victorian Britain, it was caused by *Essays and Reviews*, not by the *Origin of Species*.

Indeed, devout religious thinkers had two ways of avoiding any confrontation between Darwin and religious belief. The first way we can call separatism – the claim that science and religion are separate, and completely different approaches to understanding the world. This attitude was nicely summed up by Baden Powell in his article in *Essays and Reviews*:

> Scientific and revealed truth are of essentially different natures, and if we attempt to combine and unite them, we are attempting to unite things of a kind which cannot be consolidated, and shall infallibly injure both... In physical science we must keep strictly to physical induction and demonstration; in religious enquiry, to moral proof; but never confound the two together.

Similarly, Sir Leslie Stephen (1832–1904), philosopher and father of Virginia Woolf (1882–1941), declared that he was 'willing to grant Darwin everything, because it is unimportant – the one main fact is that somehow or other, I am here'. In other words, Stephen was unconcerned about how mankind came into existence, because it had no bearing (he believed) on the undeniable fact that man was a moral being.

The second way of dealing with Darwinism was simply to incorporate it into the tradition of natural theology. This wasn't difficult to do, of course, since Darwinism itself grew out of the tradition of natural theology which had been developing in Britain since the eighteenth century. Remember, Malthus himself was an Anglican clergyman and a natural theologian. Churchmen were able to regard the Darwinian theory as just a more refined version of the traditional natural theologies, which they had been taking for granted ever since the days of Newton. God created the world in such a way that the laws of nature will ensure the development through time of all phenomena, to result in the world we see today.

So, for example, Charles Kingsley (1819–75), professor of history at Oxford, and novelist (he was the author of the once popular children's book, *The Water Babies*, 1863), wrote to Darwin, as soon as he had read the *Origin*:

> We know of old that God was so wise that he can make all things; but behold he is so much wiser even than that, that he can make all things make themselves... I have gradually learnt to see that it is just as noble a conception of deity, to believe that he created primal forms capable of self-development into all forms needful *pro tempore* [for the time] and *pro loco* [for the place], as to believe that he required a fresh act of intervention to supply the lacunas which He himself had made.

Note that Kingsley's mention of 'forms needful' for particular times and places shows that he is aware of the geological tradition, in which it was supposed that earlier life forms contributed to the development of the Earth's environment, eventually making it suitable for mankind.

Baden Powell spoke in similar terms in his article in *Essays and Reviews*:

> A work has now appeared by a naturalist of the most acknowledged authority, Mr Darwin's masterly volume on the *Origin of Species* by the law of natural selection, which now substantiates on undeniable grounds the very principle so long denounced by the first naturalists – the origination of new species by natural causes: a work which must soon bring about an entire revolution of opinion in favour of the grand principle of the self-evolving powers of nature.

All this, however, simply pointed to what Baden Powell called a 'grander view of the Creator'.

Similarly, Henry Drummond (1851–97), a leading Scottish theologian, declared Darwin's theory to be 'a real and beautiful acquisition to natural theology'. Likewise, A. H. Strong (1839–1921), the leading Baptist theologian of the age, in his *Systematic Theology* of 1907, wrote that 'We grant the principle of evolution, but we regard it as only the method of divine intelligence'.

None of this should seem surprising. It was, after all, essentially the view that Darwin himself held before he lost his faith. Consider, for example, this quotation from his notebooks, written in 1842:

> It accords with what we know of the law impressed on matter by the Creator: that the creation and extinction of forms, like the birth and death of individuals, should be the effect of secondary laws. It is derogatory that the Creator of countless systems of worlds should have created each of the myriads of creeping parasites and slimy worms which have swarmed each day of life on land and water on this globe.

God did not directly create every slimy and vile creature, Darwin wrote, but he did anticipate their usefulness in the system as a whole, and so arranged for the laws of life to ensure that they emerged from earlier species by evolution.

There was no element of hypocrisy, therefore, when Frederick Farrar (1831–1903), Archdeacon of Westminster Abbey, in his funeral oration for Darwin in 1882 said:

> This man on whom for years bigotry and ignorance poured their scorn, has been called a materialist. I do not see in all his writings one trace of materialism. I read in every line the healthy, noble, well balanced wonder of a spirit profoundly reverent, kindled into deepest admiration for the works of God.

Farrar did not know that Darwin had indeed abandoned Christianity not long before he published the *Origin*. But even if he had known, Farrer would still have been perfectly correct in believing that it was possible to be a Darwinist, and still be a believer in God. It is also worth remarking that Darwin was given a state funeral and buried in Westminster Abbey (Box 21.1).

Frederick Temple, who had been one of the contributors to *Essays and Reviews* in 1860 became Archbishop of Canterbury in 1896. Before that, in 1884, he had delivered the annual Bampton Lectures on Religion and Science, and had declared that 'Darwin's theory left the argument from design even stronger than before'. The argument from design, of course, was the argument that the complex phenomena of the world showed every sign of having been designed by an intelligent creator, and therefore proved the existence of God.

Darwin's book sold out within hours of going on sale in 1859. This could only have happened as a result of the general public buying it – it would not have sold out if only naturalists were buying it. The question immediately arises, why was it so popular?

The answer clearly lies in the fact that the educated reading public in Victorian Britain recognized it as setting the scientific seal on the tradition of *laissez-faire* political economy and the tradition of natural theology which went hand-in-hand with it, in which everything was part of God's plan and worked itself out by natural law, and progression was built into the system. For believers and non-believers alike, a version of this philosophy was dominant (for believers, God played an important part, but atheists simply assumed laws of nature were

Box 21.1 Darwin's atheism?

Echoing the feelings of political radicals that 'a poor fly in the fangs of a spider, is a *proof* that the world could not have been designed by one being, infinitely wise, infinitely good, and infinitely powerful', Darwin wrote in 1860 to a fellow naturalist: 'I own that I cannot see as plainly as others do, and as I should wish to do, evidence of design and beneficence on all sides of us. There seems to me too much misery in the world. I cannot persuade myself that a beneficent and omnipotent God would have designedly created the *Ichneumonidae* with the express intention of their feeding within the living bodies of caterpillars, or that a cat should play with mice.'

The ichneumonida is a wasp whose grubs must feed on living flesh, and so the wasp paralyses a caterpillar with its sting and lays its eggs inside the body of the caterpillar. When the eggs hatch, the wasp grubs eat their way out of the still living body of the caterpillar.

It is important to note, however, that Darwin finally abandoned his Christian faith only in 1851, after the death of his favourite daughter, Annie, aged ten (but many years after he became a committed evolutionist). It is a mistake to assume that to be a believer in evolution is to be an atheist, Darwin managed for many years to be both a believing Christian and an evolutionist, and might have continued to remain within the Church if his young daughter had not been afflicted by fatal illness. Furthermore, in answer to a letter Darwin explicitly stated that one could be 'an ardent Theist & an evolutionist', and that he himself had 'never been an atheist'.

inherent in the nature of things, and the tendency for things to improve and progress was just a natural tendency, not a God-given one).

There was, however, an alternative view; or rather, we should speak of alternative views. One theologian writing in 1865 said that if Darwin's theory is true then the Bible is 'an unbearable fiction', and all Christians have been duped by 'a monstrous lie'. It seems clear that this author knew nothing about the Biblical scholarship which had shown that in some sense the Bible was fiction. Here then, we have a more fundamentalist approach. Similarly, an editorial in the *Family Herald* newspaper in 1871 declared that 'Society must fall to pieces if Darwinism be true'.

Professor Charles Hodge (1797–1878), President of the prestigious Princeton Theological Seminary, wrote that:

> Darwin's theory is that hundreds or thousands of millions of years ago God called a living germ, or living germs, into existence, and that since that time God has had no more to do with the universe than if he did not exist. This is atheism to all intents and purposes.

Hodge was a Calvinist, and so very much outside the tradition of natural theology of those who believed that God was obliged to create the best of all possible worlds. Calvinists insisted upon the complete freedom of God's will, and believed that he could make the world any way he liked, without having to make sure it is the best possible world. Calvinists saw the argument that God must, because of his Goodness, create the best possible world as a constraint on God's freedom, and therefore an unacceptable argument. Calvinists took the alternative approach to the problem of evil, namely that we cannot understand God's mind and cannot fathom the depths of his intents and purposes. Hodge's theology was closer to Hugh Miller's, whom we looked at as a critic of *Vestiges of the Natural History of Creation* (Chapter 19).

The Victorian belief in progress – and that the world had built-in principles of advance and self-improvement began to fade with the decline and fall of the Victorian Empire, one of the most extensive empires the world has ever seen. The First World War (1914–18) sounded its final death knell. During this time Darwinian scientists began to realize that Darwin had (almost certainly unwittingly) come up with a mechanism which could undermine the claim that the world was designed by a supreme intelligence. Natural selection could result in something that looked as though it must have been designed, but in fact was the result of the fortuitous survival of random events and changes over vast periods of time.

Consequently, scientists increasingly dismissed the design argument for the existence of God. The mainstream churches, accordingly, such as Anglicanism and Roman Catholicism, whose leaders were highly educated men, moved away from the design argument, and took an increasingly separatist line – insisting that religion operated in a different sphere from science, and was concerned with moral values, not natural truths (Box 21.2).

The more minority churches, notably various evangelical Protestant sects, nowadays often referred to as fundamentalist, by contrast, took up the design argument and tried to use it against Darwinism. This approach flourished particularly in the United States of America, but it did so for a very specific reason. The constitution of the United States allowed for freedom of religion, but forbade religious teaching in schools – all schools had to be completely secular.

In order to get around this, a number of evangelical groups in the US hit on the idea of presenting Creationism as a scientific hypothesis. Scientifically inclined members of these churches compiled evidence and arguments in support of the claim that the Biblical account of the Creation was literally true. They were careful to present their views, however, as only hypothetical. Creationism was presented as a *scientific hypothesis*.

This went hand-in-hand with a concomitant attempt to argue that Darwinism was also only a hypothesis, and not proven scientific fact. They tried to establish this by direct criticism of Darwinian theory, and by exploiting the lack of consensus amongst philosophers of science as to what constitutes good science, or good scientific method. The final stage in the argument of these religious groups was to claim that Creationism and Darwinism should therefore be given equal consideration in school science classes.

So far, the evangelicals have failed to persuade the legislature in the US to accept this claim, and to allow teaching of Creationism in schools. The background to these developments, in any case, is not so much an attempt to disprove Darwinism as an attempt to win the right, in spite of the US Constitution, to teach religious beliefs in state schools.

Darwinism and social theory

We saw in Chapter 16 how Newton's achievement made such an impression on eighteenth century thinkers that it inspired many to believe not only that Newton's method could solve all the problems of natural philosophy or science but also that the same method could be used to develop a new science of man – embracing psychology and moral theory, and theories about the correct running of society.

Perhaps the single most effective aspect of this extension of the Newtonian method was Adam Smith's development of the principles of political economy in his book the *Wealth of Nations* (1776). Smith argued that the principle of individual self-interest together with economic laws of supply and demand took the place of universal gravitation and the laws of motion in the study of how societies functioned. The result of this was *laissez-faire* theories in politics and economics. Assuming that the laws of political economy – exactly like the laws of nature – were established by God, and so all that followed from them was necessarily good (remember Alexander Pope's comment—'whatever is, is right'), Smith and subsequent political economists advocated non-interference with these laws, and therefore minimal political intervention. This is what *laissez-faire* meant.

Box 21.2 If Darwinism was not seen as inimical to all religious belief, why does modern popular consciousness take it for granted that it was?

Part of the reason is simply that in our secular world-view it is easy to dismiss religion as superstitious nonsense and to intuitively assume that it must be opposed to the scientific world-view. Even many secular historians are guilty of this; having no interest in studying theology or what religious believers actually believe, they assume science and religion must be incompatible, opposed views.

But another major factor in current perceptions is the deliberate campaign against religion launched by a secularist group of scientists in the late Victorian period. To a large extent this campaign was a strategy to increase the professional status of scientists and to establish an intellectual 'territory' which only scientists were deemed fit to control.

At the opening of the nineteenth century British science was characterized by amateurism, aristocratic patronage, negligible government support, limited employment opportunities, and only peripheral inclusion within the Church-dominated universities. What official posts there were in science were often filled by clergymen who had made a reputation for themselves as experts in natural knowledge through their concern with natural theology. By the middle of the century the numbers of secular would-be scientists had grown apace but such men found it difficult, in the face of clerical competition, to land jobs and make careers in science. Accordingly, a group of secular scientists deliberately moved away from the prevailing assumption that science and religion were complementary, and sought to emphasize instead the importance of scientific expertise irrespective of religious considerations. To counter the entrenched assumptions about the fruitful alliance between science and religion some of them evidently felt they had to fight dirty. Francis Galton wrote a book on

This led directly to Thomas Malthus's *Essay on the Principle of Population* (1798), which in turn led to Darwin recognizing the mechanism he needed to explain evolutionary change, namely, what he called natural selection, but which Herbert Spencer (1820–1903), another *laissez-faire* social theorist, called survival of the fittest.

Let us spell out what has happened here. Social theorists, such as Smith and Malthus, were trying to reinforce their theories by claiming that they were based on sound scientific principles (analogous to Newton's principles), and derived from a proven scientific tradition (Newtonianism). But, by the middle of the nineteenth century, such claims about links between social theory and hard science were looking pretty tenuous – they'd moved a long way from anything found in Newtonian physics. But then along came Darwin, who borrowed an idea or two from Thomas Malthus and used it to found what T. H. Huxley, at least, saw as the greatest single contribution to scientific theory since Newton. The result was that once again social theory could claim to derive from the best scientific tradi-

English Men of Science: Their Nature and Nurture (1674) in which he declared, in the teeth of the evidence against him, that the 'pursuit of science is uncongenial to the priestly character'. John Draper (1811–82), professor of chemistry and botany in New York University, and the founder of Cornell University, Andrew Dickson White (1832–1918), presented selective and downright distorted historical accounts to reveal the hitherto barely noticed *History of the Conflict Between Religion and Science* (Draper, 1874), and *The History of the Warfare of Science with Theology in Christendom* (White, 1896). Others, such as T. H. Huxley, the Royal Institution's professor of physics John Tyndall (1820–93), the mathematician W. K. Clifford (1845–79), and Darwin's friend Joseph Hooker (1817–1911), made concerted efforts to promote professionalization in science and to exclude clerical interests. The famous science journal, *Nature* (still going strong) was founded in 1869 by the astronomer and spectroscopist Norman Lockyer (1836–1920) as part of this same campaign. Darwinism also figured prominently in this movement as natural selection came to be seen as a way of explaining how the natural world might look as though it has been designed, but is in fact the result of random variations subject to selective pressure from the habitat. Darwinism thereby became explicitly atheistic, denying the need for an intelligent creator.

Given that all this was taking place at a time when atheism was becoming increasingly prominent, this new view of science as opposed to religion, with Darwinism as the chief illustrative example, found a ready audience and has since become firmly entrenched in popular consciousness.

FURTHER READING:

Frank M. Turner, 'The Victorian conflict Between Science and Religion: A Professional Dimension', in his *Contesting Cultural Authority: Essays in Victorian Intellectual Life* (Cambridge: Cambridge University Press, 1993).

tions. As Herbert Spencer said, 'The truth of survival of the fittest, which biology borrowed from sociology was returned with vast interest'. Social theory, in other words, gained immeasurably from the fact that Darwin had originally borrowed his idea from social theory.

Following the publication of the *Origin of Species*, therefore, a new designation emerged for this kind of political economy and its associated socio-political theory: Social Darwinism.

The basic premise of Social Darwinism is that mankind, as a social animal, has developed in accordance with the principle of survival of the fittest. However, there are a number of different ways in which this basic idea has been pursued by different thinkers, depending upon which particular axe they wanted to grind.

In one aspect, for example, it simply extended the same kind of ideas found in Malthus's *Essay on the Principle of Population* – arguing against poor relief, on the grounds that it would, in the long run, do more harm than good. Consider, for example, Herbert Spencer's comment in his *Social Statics* of 1851:

> Pervading all nature we may see at work a stern discipline, which is a little cruel that it may be very kind. That state of universal warfare maintained throughout the lower creation, to the great perplexity of many worthy people, is at bottom the most merciful provision which the circumstances permit of... The poverty of the incapable, the distress that comes upon the imprudent, the starvation of the idle, and those shoulderings aside of the weak by the strong, which leave so many 'in shallows and in miseries', are the decrees of a large, far-seeing benevolence. It seems hard that an unskillfulness which with all his efforts he cannot overcome, should entail hunger upon the artisan. It seems hard that a labourer incapacitated by sickness from competing with his stronger fellows, should have to bear the resulting privations. It seems hard that widows and orphans should be left to struggle for life or death. Nevertheless, when regarded not separately, but in connection with the interests of universal humanity, these harsh fatalities are seen to be full of the highest beneficence – the same beneficence which brings to early graves the children of diseased parents, and singles out the low-spirited, the intemperate, and the debilitated as the victims of an epidemic.

Notice that the work from which this quotation comes is called *Social Statics* and, as the title suggests, it is concerned with how society continues to function in a stable, or static or perfectly balanced, state. Spencer sees the principle of survival of the fittest as crucial to this steady state of society. Notice also that it was published in 1851, eight years before the appearance of the *Origin of Species*. This would not have been called Social Darwinism when it first appeared, therefore, but with hindsight it can be seen to be fully representative of that later movement. This clearly shows that Social Darwinism, like Darwinism itself (previously stated by writers like Charles Wells, Patrick Matthew, and Alfred Russel Wallace), was 'in the air' in some sense, in Victorian Britain.

Consider, for a post-Darwinian example, Sir Leslie Stephen, father of Virginia Woolf and leading Victorian man of letters, writing in his 'Agnostic's Apology' of 1893:

> Charity, you say, is a virtue; charity increases beggary, and so far tends to produce a feebler population; therefore this moral quality clearly tends to diminish the vigour of a nation. The answer is, of course, obvious, and I am confident that Professor Huxley would so far agree with me. It is that all charity which fosters a degraded class is therefore immoral.

As a result, Social Darwinism is used to support the competitive ethos of Victorian capitalism, and to extol famous 'Victorian values' such as being cruel to be kind, and the belief that charity begins at home.

Similarly, the American millionaire J. D. Rockefeller (1839–1937) defended the cut-throat world of big business in Darwinian terms:

> The growth of a large business is merely survival of the fittest... The 'American Beauty' rose can be produced in the splendour and fragrance which bring cheer to its beholder only by sacrificing the early buds which grow up around it. This is not an evil tendency in business. It is merely the working out of a law of nature and a law of God.

This is a perfect example of the major assumption of Social Darwinism – that human-made institutions, such as business and industry, are *not*, as you might imagine, ruled by laws and rules made by those humans who set them up, but they are ruled by inescapable *natural laws*. This was built into this way of thinking ever since Alexander Pope, Adam Smith, and the original Newtonian thinkers who set all this in train.

Consider also this example from William G. Sumner, a professor at Harvard University who used to boast that he would hand his professorship over to anyone who could give his lecture courses better than he could. In his 'Reply to a Socialist', he wrote:

> When we talk of 'changing the system', we ought to understand that that means abolishing luck and all the ills of life. We might as well talk of abolishing storms, excessive heat and cold, tornadoes, pestilences, diseases and other ills. Poverty belongs to the struggle for existence, and we are all born into that struggle.

These are individualistic, and capitalistic, applications of Social Darwinism. But there was also a nationalistic dimension to these same theories – a concern for survival of the fittest nation (or ethnic group).

Consider Walter Bagehot (1826–77), who published a book entitled *Physics and Politics, or, Thoughts on the Application of the Principles of 'Natural Selection' and 'Inheritance' to Political Society*, in 1872. Bagehot pointed out that strong nations have always dominated their weaker neighbours, and in so doing have contributed to the *progress* of civilization. Inferior nations have either been eliminated, or take up the superior culture of their conquerors.

Social Darwinism came to be used, therefore, as a 'scientific' justification for imperial colonialism. This was the time, of course, when Britain's Empire was the most extensive the world had ever seen. It was also the time when the native American culture was being systematically suppressed (usually by dint of genocide) by the superior culture of Anglo-Americanism. The Reverend Josiah Strong (1847–1916), founder of the conservative social gospel movement and an enthusiastic American imperialist, similarly defended the colonization of the Philippines on the grounds that it was 'America's duty to the world in general, and to the Filipinos in particular'. Charles Darwin himself, writing to a correspondent in 1881 said,

> I could show fight on natural selection having done and doing more for the progress of civilisation than you seem inclined to admit... The more civilised so-called Caucasian races have beaten the Turkish hollow in the struggle for existence.

In Germany, early in the twentieth century, Social Darwinism was even used to defend the idea of aggressive warmongering. War was regarded as a necessary aspect of human progress – the fittest nation survives and eliminates inferior stock. For example, General Friedrick von Bernhardi (1849–1930) in his book, somewhat ominously entitled *Germany and the Next War* (and published in 1911, three years before the outbreak of the First World War), wrote:

> War is not merely a necessary element in the life of nations but an indispensable factor of culture, in which a truly civilized nation finds the highest strength and vitality... War gives a biologically just decision, since decisions rest on the very nature of things... It is not only a biological law, but a moral obligation, and, as such, an indispensable factor in civilization.

Notice that the General brings *biology* into his claims – war does not depend upon who has the biggest stockpile of weapons, and the biggest standing army; no, it depends upon the biology of the victors, and results in 'a biologically just decision'.

You might think this only shows that the German military were capable of cynically exploiting contemporary scientific ideas for their own ends, but that 'science' itself, as an institution, was above all this. But the leading light of this kind of theorizing was Ernst Haeckel (1834–1919), who was the leading biologist and Darwinist in Germany, and who insisted that 'politics is applied biology'. Haeckel was bitterly disappointed by Germany's defeat in the First World War, but his ideas lived on and were instrumental in the formation of the Nazi ideology, with its belief that the Germans were a superior race of human beings.

Let's move on to another aspect of Social Darwinism – the claim that the different races of mankind represent different levels of progress, or advancement, up the evolutionary ladder. Again, this is an idea that pre-dated Darwinism proper, and actually fed into Darwin's own ideas. Following on from Darwin's publication of the *Origin of Species*, however, these ideas became even more scientifically respectable. Charles Lyell, for example, wrote that 'each race of man has its place, like the inferior animals'. It was a commonplace belief even among progressionist geologists and others before the advent of evolutionary theory, that there was a hierarchy of primates from the Caucasian white at the top down to the African negro, the Australian aborigine, and on to the gorilla, and so on.

Consider this quotation from Haeckel's *History of Creation* (1868):

> In order to be convinced of this important result, it is above all things necessary to study and compare the mental life of wild savages and of children. At the lowest stage of human mental development are the Australians, some tribes of the Polynesians, and the Bushmen, Hottentots, and some of the Negro tribes. In many of these languages there are numerals only for one, two, and three: no Australian language counts beyond four. Very many wild tribes can count no further than ten or twenty, whereas some very clever dogs have been made to count up to forty and even beyond sixty.

The history of these notions is clearly disgraceful, being used to justify not only colonization but also slavery. Unfortunately, these ideas are not dead yet, being still upheld by contemporary racists.

Another manifestation of this way of thinking originated in psychiatric theory. In 1866 Dr John Langdon Haydon Down (1828–96) wrote a paper entitled 'Observations on an Ethnic Classification of Idiots'. Down suggested that many so-called congenital idiots – that is to say, those born with what are today called 'learning difficulties', but which then were seen as 'mental retardation' – displayed anatomical features which were not found in their own parents, but were in fact characteristic of 'lower' races. Some congenital idiots he said looked Ethiopian, others Malayan, and most famously, he said: 'A very large number of congenital idiots are typical Mongols'.

Describing a case of 'mongolism' he wrote, 'it is difficult to realize that he is the child of Europeans'. Nowadays, such children are referred to as having 'Down's syndrome', but this is very new; even when I was a young boy, in the 1960s, we had never heard of Dr Down and always referred to such people, without thinking, as Mongols.

Incidentally, even the old designation of 'mental retardation' fits in with Social Darwinist views,. The word 'retard' means to slow down the progress of something. So a mentally retarded person is one whose progress up the evolutionary ladder has been slowed down.

There were even experimental efforts to establish the places of different races, or different types of person, on the ladder of evolution. A group in France calling themselves craniometers (brain-measurers), developed ways of weighing brains, or of measuring the capacity of skulls so that the size of the brain could be calculated. This group of experimenters compared the sizes of the brains of various humans, and also the brains of the great apes.

Women did not do too well in these studies. Gustave le Bon (1841–1931), for example, wrote in 1879 in his *L'homme et les sociétés: leurs origines et leur histoire* that:

> In the most intelligent of races, as among the Parisians, there are a large number of women whose brains are closer in size to those of gorillas than to the most developed male brains... All scientists are

> agreed that women represent the most inferior forms of human evolution.

Le Bon was not a total idiot, so presumably he did not mean to imply that Parisians should be seen as a separate race, merely that the most intelligent races can be assessed by comparing them with Parisians.

Le Bon and other craniometers claimed that they were taking into account the smaller body size (on average) of women compared to men, but they clearly did not do it properly. This is apparent from this comment by the leader of the craniometers, Paul Broca (1824–80):

> We might ask if the small size of the female brain depends exclusively upon the small size of her body... But we must not forget that women are, on the average, a little less intelligent than men... We are therefore permitted to suppose that the relatively small size of the female brain depends in part upon her physical inferiority and in part upon her intellectual inferiority.

This is what is called 'begging' the question – assuming the truth of the answer you want to arrive at, in the process of giving an argument in favour of that answer. Here, Broca wants to use the smaller brain size of women to prove that women are mentally inferior to men, but he has the problem of taking account of smaller body size. Instead of trying to work out how to compensate for average smaller body size, he simply says, we already know women are mentally inferior to men, so we can safely assume that their smaller brain size is not just due to their smaller body size, but must also be due to the fact that they are less intelligent. 'QED', as geometers used to say when they succeeded in providing a proof.

These kinds of assumptions about the inferiority of certain peoples, or persons, led to another spin-off 'science', eugenics.

One of the chief developers of this new 'science' was Darwin's cousin, Francis Galton (1822–1911). The premise of eugenics was that the human race, or some part of it, could improve itself by selective breeding, in the same way that farmers bred better strains of cattle by selective breeding. Even Charles Darwin himself discussed this in his *Descent of Man* (1871):

> I have sometimes speculated on this subject: primogeniture is dreadfully opposed to selection; suppose the first-born bull was necessarily made by each farmer the begetter of the stock!

'Positive eugenics' advocated the careful selection of marriage partners by males of high ability to assure high-achieving children (which the parents should have in large numbers). 'Negative eugenics' – arguably the most prominent aspect of the movement – tried to limit the birth rate among what were regarded as inferior types. Some American States in the early decades of the twentieth century passed legislation on the compulsory sterilization of those below a certain level of intelligence (and intelligence tests were devised to measure intelligence). And, of course, the Nazis in Germany introduced a massive sterilization programme in the 1930s.

There have been many attempts by historians and philosophers of science, and by Darwinian scientists, to argue that Social Darwinism was not a proper science, but was merely a political ideology dressed up as science. Even if this is true, however, it was not apparent to many of the people who advocated these ideas in the late nineteenth and early twentieth centuries. For them, these ideas were genuinely scientific – they grew directly out of Darwinian biological theory and simply applied the same thinking to humankind in society.

The implication of the attempt to reject the scientific credentials of Social Darwinism is that its advocates were involved in a conspiracy to twist Darwinian precepts and to use them to deceive people into thinking that there were scientific grounds for supposing some races, or some kinds of people, were inferior to others. Unfortunately, no such twisting was required. Darwin, and the other evolutionists before and after him, routinely believed that some races were more advanced than others, and so had progressed higher up the evolutionary escalator, or occupied a higher branch on the branching tree of evolution. Like Broca and Le Bon, who already believed they were superior to women, Caucasian evolutionists believed they were superior to negroes. As we've seen, evolutionary ideas did not make progressionist assumptions possible; it was progressionist preconceptions which made the eventual acceptance of evolutionary theories possible.

The fact is, Social Darwinists genuinely believed that their ideas on racial superiority were correct because they were supported by a wealth of scientific evidence as to how the world is. Whether we like it or not, many scientific ideas are dangerous, and ideas can be, have been, and no doubt continue to be, regarded as scientific without being true.

Post-Darwinian biology

There were two major problems for the theory of evolution in the decades leading up to the end of the nineteenth century: the age of the Earth, and the means of hereditary transmission of variations. The first was resolved, unexpectedly, as a result of the discovery of radioactivity.

Although the theory of evolution met with serious opposition from Lord Kelvin's calculations about the maximum age of the Earth, so did contemporary geology. Indeed, it was at a meeting of the Glasgow Geological Society in 1866 that Kelvin had declaimed that:

> A great reform in geological speculation seems now to have become necessary. British popular geology at the present time is in direct opposition to the principles of natural philosophy.

Darwin was by no means alone, therefore, in suggesting that perhaps the geological evidence should override the abstract calculations of the thermo-dynamicists. Darwin never relented, although most evolutionists simply took this as another clear indication that natural selection was insufficient on its own to account for the current state of nature, and that a faster-acting law of progressive development must also be at work.

By the early years of the twentieth century, however, Kelvin's calculations, based on assumptions about the rate of cooling of the Earth, were rejected after Ernest Rutherford (1871–1937), one of the pioneers of the study of radioactivity, suggested that the Earth's core could be radioactive, thus providing an internal source of heat which invalidated calculations based on assumptions of continuous cooling from a high temperature in the past.

Meanwhile, developments in the theory of heredity ensured that it too became a powerful support for evolutionary theory. Here, however, the story is more complicated.

It is often assumed that the rediscovery, around 1900, of the work of the Austrian monk, Gregor Mendel (1822–84), which he had originally conducted in the 1860s, led, via the Mendelian revolution, to the neo-Darwinian synthesis of Darwin's natural selection with genetics, the most powerful research tradition in modern biology. Indeed, there is even a tendency to regret the lack of communication of Mendel's ideas; suggesting that if Darwin had only known of Mendel's work, he could have seized upon the solution to his difficulties and the progress of science would not have had to wait for Hugo de Vries (1848–1935), Carl Correns (1864–1933), and William Bateson (1861-1926) to independently rediscover, and promote Mendel's ideas.

Unfortunately, this is another of those myths which is not only not true, but actually obscures the real truth. The fact is, if Darwin had seen Mendel's paper, 'Experiments on Plant Hybridization', when it appeared in 1866 he would never have thought to use it in support of his own theory. Mendel's paper was an empirical *refutation* of evolution.

You might recall from Chapter 18 that Linnaeus came to believe that perhaps new species could be formed by hybridization. The problem with this is that hybrids across species are sterile and won't propagate. But plant and animal breeders make new varieties by selectively cross-breeding other varieties. New varieties formed in this way (within the same species) can then be propagated, provided Great Danes are only allowed to mate with Great Danes, and Chihuahuas only with Chihuahuas, and so on. Linnaeus evidently thought it possible that, after many generations, a new variety might consolidate itself as a new species.

Mendel set out to test this hypothesis by hybridizing different varieties of garden peas over a ten-year period. It seems clear that Mendel wanted to prove that Linnaeus's hunch had been correct. Disappointed by the outcome, however, he wrote at the end of his paper that his results were not so clear that they had to be 'unconditionally accepted'.

The end result of Mendel's labours was to demonstrate that hybrid forms always revert back to the parental forms (or they die out completely due to reduced fertility). Hybridizing a yellow smooth pea with a wrinkly green pea, Mendel noted that the off-spring were all smooth yellow peas. When he cross-bred two of these second generation plants, however, wrinkly green peas re-appeared among the off-spring. Admittedly, there were fewer of the wrinkly green peas (they were outnumbered three to one), but they were there. The result, therefore, was that there was no blending of properties to produce, say, a smooth green pea or a wrinkly yellow one; but there was always a reversion back to the original parental types. After ten years, Mendel seemed to be proving that Linnaeus's hunch about the evolution of new species from hybrids was misconceived.

If Darwin had seen this, he would not have seized upon it as the answer to his problems with regard to the heritable transmission of new variations. Darwin's theory did not in any way depend upon hybridization, and besides, Mendel's conclusions were against the possibility of evolution.

But if this is so, how is it that later evolutionary biologists, such as De Vries and William Bateson, were able to use Mendel's work?

Let us start with Bateson, who had Mendel's paper translated into English and published a defence of Mendel's principles in 1902. Bateson was interested in the seemingly random variations which were the starting point for natural selection. While thinking about the processes by which such variations might be preserved, he came to the negative conclusion that the habitat could not exert the kind of pressure on living organisms that Darwinian theory seemed to demand.

It was perfectly possible, after all, to transfer species of plants and animals to completely different habitats, and they could survive perfectly well, for many generations, without the need for any fortuitous variations to help them adapt to the new conditions. As Bateson wrote in his revealingly titled, *Materials for the Study of Variation: Treated with Especial Regard to Discontinuity in the Origin of Species* (1894):

> We knew all along that species are approximately adapted to their circumstances; but the difficulty is that where the differences in adaptation seem to us to be approximate, the differences between the structures of species are frequently precise. In the early days of the Theory of Natural Selection it was hoped that with searching the direct utility of such small differences would be found, but time has been running now and the hope is unfulfilled.

Consequently, Bateson concluded that natural selection was comparatively unimportant, and that there must be some other causative factor which brings about variations, and thereby directs evolutionary change. As the title of his book makes plain, he also believed that this causative factor caused variations that were not slight, gradual, changes, but could be comparatively large discontinuities.

But Bateson now confronted a new problem. It had to be possible for such major, discontinuous, changes in organisms to be passed on to their off-spring.

According to the prevailing view of heredity, however, the features of both parents were blended in the progeny. So, a deer-like creature born with a much longer neck than its fellows would have to mate with a normal run-of-the-mill deer-like creature, and so its off-spring's neck would be a blend of long and short – presumably intermediate in length. Any advantage in having a long neck would soon be lost. This was the problem pointed out by Fleeming Jenkin (see Box 20.1).

Accordingly, Bateson started to do his own research on the hybridization of plants, until, in 1900, he learned of the work of Mendel. The crucial thing about Mendel's work for Bateson, of course, was that it dismissed the notion of blending inheritance. Progeny preserved parental features in an unmixed way. It should be noted, therefore, that Mendel's work was taken up not by a Darwinian, seeking to prove Darwin right, but by an evolutionary theorist who rejected natural selection, and the idea of gradual slight variations.

The same was also true of Hugo de Vries, another resuscitator of Mendel's work. Seeking to speed up evolution in response to Kelvin's restricted timescale, de Vries proposed that species undergo occasional rapid bouts of what he called 'mutations', rather than slight variations (de Vries was effectively a catastrophist in biology). These sudden bouts of mutations could also account for the discontinuities of the fossil record, and even, he suggested, provide a large number of individuals with the same mutation, who would breed only among themselves. But, as a failsafe, he exploited, as Bateson did, the Mendelian principle that parental features (mutations included) were passed on undiluted and undiminished to the progeny.

Mendelism and Darwinism (as opposed to these other forms of evolutionism) were only reconciled to one another as the result of the in-put from a completely separate research programme. This was the research programme known as biometrics, which emerged out of the work of Francis Galton.

Darwin's cousin, and one of the founders of eugenics, Francis Galton, had also been concerned to understand how variations were passed on to subsequent generations. Using rabbits as his experimental animals (good choice), Galton pioneered the use of statistical analyses and came to the conclusion that natural selection could *not* produce permanent results by small variations in rare individuals.

Consider the normal statistical distribution of any given phenomena along what is known as the bell-curve. The top-most point of the curve, representing the value in question of the majority of individuals, is also what Galton called the 'focus of regression'. So, parents who have above-average height, or above-average intelligence, will tend to have children of average height, or average intelligence.

Evolutionary change could only be said to have occurred if the focus of regression was shifted (so that the average height of the population shifted from 5 feet 6 inches to 5 feet 8 inches, for example). In his *Hereditary Genius* of 1869 Galton provided an analogy with a multi-faceted stone:

> The mechanical conception would be that of a rough stone, having, in consequence of its roughness, a vast number of natural facets, on any one of which it might rest in 'stable' equilibrium. That is to say, when pushed it would somewhat yield, when pushed much harder it would again yield, but in a less degree; in either case, on the pressure being withdrawn it would fall back into its first position. But, if by a powerful effort the stone is compelled to overpass the limits of the facet on which it has hitherto found rest, it will tumble over into a new position of stability, whence just the same proceedings must be gone through as before, before it can be dislodged and rolled another step onwards.

Variations represented the rocking to and fro of the stone on one of its faces. But the stone always returns to the equilibrium position. For evolutionary change to occur, something would have to flip the stone over on to another of its faces.

As a result, Galton rejected Darwinian gradualism in favour of a more abrupt form of evolutionary change brought about, not by natural selection, but by some other unknown process. Leading followers of Galton, however, were able to use his statistical approach in support of Darwin. The mathematician Karl Pearson (1857–1936), and the biologist W. F. R. Weldon (1860–1906), working together, showed that small variations could accumulate in a population and shift the focus of regression. Given that Darwin himself had suggested that species were not defined by hard and fast natural distinctions, but only in statistical terms, Pearson and Weldon could argue that such statistical changes could be said to lead to speciation.

In Britain, rivalry (even animosity) between Bateson on the one hand, and Pearson and Weldon on the other, ensured that their two approaches were regarded as incompatible alternatives. The subsequent generation of researchers in experimental genetics, however, began to realize that the statistical surveys produced by the biometricians (as Pearson and Weldon's school was called) could be used to clarify the bewildering complexity of genetic factors (evidently Mendel had been extremely lucky in selecting peas as his experimental material – few other plants display such simple and clear-cut hybrid forms).

From the 1920s onwards, therefore, we see the beginnings of the grand synthesis between Darwinist biometrics and Mendelian genetics which has resulted in the triumph of modern neodarwinism.

During this process history was re-written so that Mendelian genetics, originally offered as empirical evidence *against* the possibility of evolutionary change, came to be seen as the completion and consummation of Darwin's speculations which had unfortunately been hitherto overlooked. The fact of the matter is that empirical results do not speak for themselves, telling us unequivocally how nature is. To be understood, Mendel's results had to be subject to interpretation, but interpretations, even rational scientific interpretations, can be, and are, affected by initial starting points, by preconceptions, and by prejudices.

FURTHER READING

G. E. Allen, *Life Science in the Twentieth Century* (Cambridge: Cambridge University Press, 1979), chapter 5.

Robert C. Bannister, *Social Darwinism: Science and Myth in Anglo-American Thought* (Philadelphia: Temple University Press, 1979).

Peter Bowler, *The Eclipse of Darwinism* (Baltimore: Johns Hopkins University Press, 1983), chapter 8.

Peter Bowler, *Theories of Human Evolution: A Century of Debate, 1844–1944* (Oxford: Blackwell, 1986)

Peter Bowler, *The Invention of Progress: The Victorians and the Past* (Oxford: Backwell, 1989)

Peter Bowler, *Evolution: The History of an Idea*, 3rd edn (Berkeley: University of California Press, 2003).

Peter Bowler, *The Fontana History of the Environmental Sciences* (London: Fontana, 1992), chapter 10.

L. A. Callender, 'Gregor Mendel – An Opponent of Descent with Modification', *History of Science*, 26 (1988), pp. 41–75.

Jean-Marc Drouin, 'Mendel in the Garden', in M. Serres (ed.), *A History of Scientific Thought: Elements of a History of Science* (Oxford: Blackwell, 1995), pp. 506–25.

J. C. Greene, *Darwin and the Modern World View* (Baton Rouge: Louisiana State University Press, 1961).

Greta Jones, *Social Darwinism and English Thought* (Brighton: Harvester Press, 1980), chapter 2, pp. 10–34.

Robert C. Olby, *Origins of Mendelism* (London: Constable, 1966).

D. R. Oldroyd, *Darwinian Impacts* (Milton Keynes: Open University Press, 1983), chapters 16 and 18.

F. M. Turner, 'Victorian Scientific Naturalism', in C. Chant and J. Fauvel (eds), *Darwin to Einstein: Historical Studies on Science and Belief* (Harlow: Longman, 1980), pp. 47–68.

22

Beyond Newton: Energy and Thermodynamics

Newtonianism might be said to have reached its zenith in Napoleonic France in a school of physics led by the mathematician Pierre-Simon Laplace (1749–1827) and the chemist Claude Louis Berthollet (1748–1822). They and their followers sought to explain all phenomena in terms of attractive and repulsive forces acting at a distance between particles, or in terms of subtle fluids whose particles were mutually repulsive, but attractive towards ordinary matter. The fluids, often referred to as the imponderable fluids, because their presence could not be detected by weighing, derived from Newton's speculations in the queries at the end of the *Opticks*, and included heat, light, electricity, and magnetism.

At the height of its influence from 1805 to 1815, Laplace's Newtonianism was beginning to be eclipsed by the 1820s. In France, the overthrow of Napoleon Bonaparte (1769–1821) and the restoration of the Bourbons seems to have stimulated intellectual change in science as much as in other spheres of intellectual activity. Certainly, Laplace and his circle were closely identified with Napoleon's regime and Laplace himself was an unpopular figure in the early years of the Bourbon Restoration.

In French science, his eclipse can be seen in the new treatments of two of the supposed imponderable fluids, heat and light, by Joseph Fourier (1768–1830) and Augustin-Jean Fresnel (1788–1827) respectively. Fourier provided a mathematical account of the conduction of heat which owed nothing to what he saw as unsubstantiated assumptions about the nature of caloric, the name given to the imponderable fluid of heat; while Fresnel's work seemed to repudiate the dominant Newtonian concept of light as particulate, establishing it to be not only a wave but a transverse wave; that is to say, a wave of the kind you can make in a rope by moving one end up and down – or from side to side. This theory could not even be made compatible with Newton's speculations about an aether which was assumed to transmit light, electricity, etc. by longitudinal waves (such as sound waves, which are formed by pulses of denser and rarer air travelling directly outward from the source).

Another factor in the decline of Laplacian Newtonianism, however, and one which was international in scope, was the increasing awareness by physicists and other scientists of the new importance of engines and engineering.

The Industrial Revolution had been underway since the second half of the eighteenth century but in its beginnings this had little or nothing to do with scientific knowledge or practice. Input from science was hardly required, for example, for the industrializing developments in textile production – cotton gins, spinning jennies and mules, flying shuttles, etc. It is often assumed that the development of the steam engine depended upon discoveries made with the air-pump, and the discovery of latent heat by Joseph Black. But even this is highly dubious. As far as we can tell Thomas Newcomen (1663–1729) developed the steam engine for pumping water out of mines by trial and error. If there was any input from science it was at the level of someone informing him that a piston in a cylinder could lift a huge weight if a vacuum was created in the cylinder, and that a vacuum could be created in the cylinder by filling it with steam and then cooling the cylinder to make the steam condense. A theoretical description of just such a 'steam engine' had been described by Denis Papin (1647–*c*.1712) in about 1702 and its import could have been passed on to Newcomen.

The Industrial Revolution, therefore, proceeded apace without paying any attention to the latest science; and scientists, for their part, carried on without paying much attention to developments in industrialization. Steam pumps were in common use in mines all over the country by the second half of the eighteenth century but they hardly impinged upon the consciousness of physicists. This all began to change, however, after 1829 when the Liverpool to Manchester railway opened. Steam locomotives had been in use sporadically since 1804 when Richard Trevithick (1777–1833) developed one for hauling heavy goods at an ironworks in Wales, but again, few people, including natural philosophers, knew of them until the public Liverpool to Manchester railway caused a sensation.

This seems to have marked a turning point. Certainly, the 1830s saw a marked change in attitudes among scientists. Seeing the importance of engineering projects for society as a whole, physical scientists began to think more like engineers and to consider the implications of engineering know-how for their science, and the relevance of their science to engineering in ways that had been confined only to rare individuals before. Furthermore, many of the leading contributors to developments in the physical sciences were engineers by profession. It is hardly surprising, therefore, that the principles of mechanics, and mechanical effects, took centre stage in natural philosophy (Box 22.1).

An important aspect of this was a self-conscious move away from Newtonian actions at a distance. The only acceptable forces were forces of motion, or forces of impact, and so electricity, magnetism, heat, and other forces came to be seen in terms of the motions and contact actions of particles. Where Newtonians had previously simply spoken of forces of attraction between particles but said nothing about what, if anything, was going on in the space between the particles, Faraday filled the intervening space with lines of force which quickly came to be hypostatized as real physical entities, without which attraction could not take place. As in the days of Descartes, the world system came to be seen to work like a giant machine, and all the moving parts had to be considered to physically intermesh with one another.

Box 22.1 Scientists and engineers in Victorian Britain

Further evidence for the claim that natural philosophy changed its character and became closer to the ethos of mechanics and engineering in the period after 1830 can be seen from the fact that scientists spent time trying to distinguish themselves from engineers in a way that had never been seen as necessary before. It was one thing for scientific thinking to be based on engineering precepts, it was quite another, however, for scientists to be seen as no different from engineers. Engineers, for their part, tried to lay claim to being superior to pure scientists. Accordingly, explicit efforts were made to define the differences between scientists and engineers.

William Sewell (1804–74), Oxford don and educationalist, regretted the fact that 'deep thinking' was 'out of place in a world of railroads and steamboats, printing presses and spinning jennies', but Samuel Smiles (1812–1904), celebrator of Victorian achievement, especially by the common man, thought it was a cause for celebration that the 'great mechanics' were *not* natural philosophers or mathematicians, but learned 'their practical knowledge in the workshop, or acquired it in manual labour'. John Tyndall, however, who had managed to inveigle himself as 'scientific adviser' to the Board of Trade, took the opportunity to denigrate the engineers who were consulted by the Board as having closed minds and being diffident toward the encouragement of new scientific ideas. Tyndal resigned his position when the Board of Trade favoured the practical engineers over speculative men of science. Tyndall subsequently became something of a spokesman for the superiority of the 'man of science', motivated only by 'the noble excitement of research, and the joy attendant on the discovery of natural truth'. Admitting that the scientist 'cares little for practical ends', Tyndall denigrates the opposite aims of the engineer 'whose object is mainly industrial' and is motivated by the desire 'to make money' and establish a monopoly. Engineers, furthermore, only know 'the conditions, not the reasons, of success', and do not understand the causal principles underlying their techniques. 'Our science', however, 'would not be worthy of its name and fame if it halted at facts, however practically useful, and neglected the laws which accompany and rule the phenomena'.

At a time when science and engineering came to share the same ethos, not to mention many of the same techniques, whether experimental or mathematical, those who wanted to ensure that they maintained separate identities had to deploy insistent and none too subtle rhetoric.

FURTHER READING

Thomas F. Gieryn, 'Boundary Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists', *American Sociological Review*, 48 (1983), pp. 781–95.

Edwin T. Layton, 'Mirror-Image Twins: The Communities of Science and Technology', *Technology and Culture*, 12 (1971), pp. 562–80.

This new version of the mechanical (if not the mechanics') philosophy was especially noticeable in Britain, of course, where the Industrial Revolution began, and where it had the most dramatic effects on society. The characteristic science which developed in Britain was noticeable even at the beginning of the twentieth century. Commenting on a new theory of electricity by the English physicist Oliver Lodge (1851–1940), the French philosopher and scientist Pierre Duhem (1861–1916) noted that:

> In it there are nothing but strings which move around pulleys, which roll around drums, which go through pearl beads, which carry weights; and tubes which pump water while others swell and contract; toothed wheels which are geared to one another and engage hooks. We thought we were entering the tranquil and neatly ordered abode of reason, but we find ourselves in a factory.

Even in Bourbon France, however, the steam engine played a far larger role in the development of theoretical science than science had ever played in the development of the steam engine. In his *Reflections on the Motive Force of Fire* (1824), Sadi Carnot (1796–1832), a military engineer, drew an analogy between a water-wheel and a steam engine, and concluded that the work produced by the engine depended only on the fall of temperature during its cycle, from hot to cold, in the same way that the water-wheel's power depended on the distance fallen by the water driving it. Carnot saw the fall of the imponderable fluid of caloric from the hot cylinder of the steam engine to the cold condenser, as the equivalent of the fall of water from the top to the bottom of the water-wheel.

For others, however, there was a crucial difference between the steam engine and the water-wheel. While the *movement* of water drove the *movement* of the wheel, it was a change of temperature that drove the movement of the steam engine. The latter looked more like the conversion of heat into motion, rather than the translation of one kind of movement into another. This different way of looking at the steam engine was encouraged by other contemporary discoveries in which one kind of force seemed to be converted into another.

From conversion to conservation

The accidental discovery by Hans Christian Oersted (1777–1851) in 1820 that an electric current in a wire deflected a magnetic needle was taken up by André Marie Ampère (1775–1836) in France and Faraday in Britain and apart from suggesting the conversion of electricity into magnetism (and vice versa), it also led to the development of the electric motor, in which electricity and magnetism were turned into motion. The discovery of photosensitive chemicals and the invention of photography suggested that light could be converted to chemical forces, as Humphry Davy had earlier suggested that chemical forces in a voltaic pile were converted into electricity.

For the engineering mentality, however, the important thing was to produce motive force, or work, and one of the most important aspects of this enterprise was undertaken by James Prescott Joule (1818–89).

The son of a brewer in Salford, in the heartland of the Industrial Revolution, and a former student of John Dalton, Joule tried to embark on a career in physics (in the late 1830s) by studying and trying to improve the efficiency of electric motors. During the course of this work he noted, and studied, the production of heat as electricity flowed along a wire. This in turn led Joule to reject the standard (Newtonian) view that apparent heating effects in a system were to be explained simply in terms of the movement of the imponderable fluid, caloric. Joule came to believe that electricity was actually converted into heat.

Behind Joule's thinking on these matters was a conviction that 'the grand agents of nature are, by the Creator's fiat, *indestructible*, and that wherever mechanical force is expended, an exact equivalent of heat is always obtained'. This committed Joule to rejecting the prevailing interpretation of Sadi Carnot's analogy between the steam engine and a water-wheel. The drop in temperature from boiler to condenser, which was seen as equivalent to the fall of water across the water-wheel, seemed to imply that heat was simply lost or destroyed. 'Believing that the power to destroy belongs to the Creator alone', Joule concluded that 'any theory which... demands the annihilation of force is necessarily erroneous'.

Joule now turned away from electrical machines to try to establish that heat was never lost but was always converted into motive force. His most famous experiment, conducted in 1845, sought to establish the conversion of mechanical movement into heat. A descending weight, whose motive force could be easily established, was made to drive the rotation of a paddle wheel submerged in a known volume of water in a container. The rotation of the paddle wheel generated friction in the water and caused an increase in its temperature, which could be carefully measured. Joule concluded that the motive force exerted by the fall of 890 pounds over a distance of one foot was equivalent to the amount of heat required to raise the temperature of a pound of water by one degree Fahrenheit.

Hitherto Joule had encountered marked indifference to his work, and rejection of his papers by leading journals, but in 1847, after a brief account of his paddle-wheel experiments at a meeting of the British Association for the Advancement of Science, he drew the favourable attention of William Thomson (1824–1907), professor of natural philosophy at Glasgow. Thomson, later to become Lord Kelvin, was already an important figure in Victorian science and by 1850 Joule was elected a Fellow of the Royal Society. Then, as now, it's not what you know that counts, but who you know.

Joule's work on the mechanical equivalent of heat played a major part in the efforts of Thomson and others to establish the doctrine of the conservation of what they now called *energy*. Previously, the word 'energy' had simply been used loosely as a synonym for force or power, but Thomson and like-minded physicists and engineers now used it to signify the entity which was quantitatively conserved as one force was converted into another. Energy, therefore, was a mathematical entity. When a particular amount of electricity, say, was converted into a particular amount of heat, it could be shown mathematically that something was conserved throughout; that before and after the conversion process the books balanced, so to speak. The abstract something that was conserved was declared to be energy.

This attracted criticism from a number of contemporaries. Energy was dismissed by some as a mathematical fiction, unlike force whose physical effects could be felt. Thomson, however, had already (1841–2) worked out a successful mathematical analysis of electrostatic induction, along the same lines that Fourier had used with regard to the conduction of heat, and so he had already moved some way towards a physics that was more mathematical than tangible. We can see here just one manifestation of an increasing trend towards mathematical abstractions as the stuff of physics (see Box 22.2), which was to reach its zenith in quantum mechanics (see Chapter 24).

Thomson shared Joule's conviction that nothing that God created could be destroyed by merely human means. And yet although it was clear from Joule's work that heat might be turned into mechanical work, there were numerous occasions when heat did no such useful work (for example, when a pan of boiling water is left to cool down after boiling an egg). Furthermore, Thomson was sceptical of Joule's assumption that, just as his paddle-wheel experiment showed motive force could be converted into heat, heat could be converted entirely into motive force.

In 1851 Thomson announced his 'Dynamical Theory of Heat' which resolved these problems by assuming that energy can be 'lost to man irrecoverably though not lost in the material world'. Only God could create or destroy energy, and the best that humankind could do is exploit transformations of one form of energy into another. During these transformations, however, some energy was dissipated beyond man's control:

> no destruction of energy can take place in the material world without an act of power possessed only by the supreme ruler, yet transformations take place which remove irrecoverably from the control of man sources of power which... might have been rendered available.

It seems clear that Thomson saw this view as confirming the Calvinist theology of his Presbyterian upbringing in Glasgow. Rejecting the progressionism of Anglican natural theology which had recently been popularly expounded in the *Vestiges of the Natural History of Creation* (see Chapter 19), Thomson emphasized the gradual dissipation of energy into forms that made it irrecoverable to humans and so suggesting an eventual decline (culminating in the so-called 'heat death' of the universe when everything in the universe evens out at the same temperature) rather than a triumphant progress (see Box 22.3).

Box 22.2 PHYSICS – CONTINENTAL STYLE

We have already noted that British science was seen by Continental scientists as excessively dependent on physical analogies and 'models' – scientific explanation had to be couched in terms of the easily visualized behaviour of everyday objects. P. G. Tait even devised an experimental demonstration of the stability of smoke rings to support Thomson's claims about vortexes in the aether.

Another way of looking at this is to note that, on the Continent, physicists were more abstract in their approach, concentrating more on providing a mathematical analysis of the phenomena.

This can be seen, for example, in the very different ways that the second law of thermodynamics was stated by Thomson, in Britain, and Rudolf Clausius (1822–88), the leading creator of thermodynamics in Germany (he developed the important notion of entropy). Thomson's form of the law ran like this: 'It is impossible, by means of inanimate material agency, to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest of the surrounding objects'. We don't have to understand this; we just have to note that its concern is with physical bodies in physical surroundings. Now compare this with Clausius's statement of the law: 'The algebraic sum of all the transformations occurring in a cyclical process can only be positive, or, as an extreme case, equal to nothing'. The closest this gets to a physical set-up is in the phrase 'cyclical process', but even this is highly abstract.

Pierre Duhem, who saw British physics as tainted by the factory floor, described in his *Aim and Structure of Physical Theory* (1914), how a German or a French physicist would attempt to understand the forces operating between two electrically charged bodies:

> [They] will by an act of thought postulate in the space outside these bodies that abstraction called a mathematical point and, associated with it, that other abstraction called an electric charge. He then tries to calculate a third abstraction: the force to which the material point is subjected. He gives formulas... From these formulas he deduces... Finally, he integrates all these elementary forces according to the rules of statics; he then knows the laws of the mutual actions of the two charged bodies.

Thomson's two laws of thermodynamics brought together the speculations of Joule and of Carnot. The first law stated that energy could neither be created nor destroyed but only converted from one form to another. The second law can be stated in a number of ways. Thomson's own version was that 'it is impossible, by means of inanimate material agency, to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest of the surrounding objects'. But a simpler way of putting it is that heat will not spontaneously flow from a cold region to a hotter one, but only from a hotter to a colder region. It follows from this that it is impossible to convert heat completely into work or motive power.

Box 22.3 Religion and Science in Energy Physics

We saw in the previous chapter (see Box 21.2) that a group of Victorian scientists deliberately sought (chiefly as a strategy for professionalizing science) to break the age-old alliance between natural philosophy and theology. So, what did these scientists make of Joule's and Thomson's overt pronouncements about God in their physics? As you might expect, perceiving that they undermined their efforts to divorce religion and science, they could not let them pass unchallenged.

John Tyndall (1820–93), professor of physics at the Royal Institution went so far as to put his opposition to religion in science above his jingoistic pride in English achievement. Tyndall took the side of Julius R. Mayer (1814–78), a German physicist, who engaged in a priority dispute with Joule in 1848. Mayer had published on the equivalence of motive force and heat in 1842 but his ideas had had little impact. Joule, grateful to be brought to continental attention by Mayer's attack, conceded that Mayer might have had the idea first but that he, Joule, had been the first to prove it empirically. It wasn't until the 1860s, when Tyndall was embarked with others on the campaign to separate religion and science, that he started to praise Mayer's achievements over Joule's. Peter Guthrie Tait (1831–1901), as devout a religious believer as Joule and Thomson, entered the fray on Joule's side, dismissing Mayer as speculative and amateurish, compared with the painstaking experimentalist, Joule. It seems hard to believe that the sub-text of these exchanges did not include an element of rejecting on the one hand, and defending on the other, the reliance on notions of God in natural philosophy.

Certainly, Tait took issue with Tyndall on the relationship between science and religion after Tyndall delivered his notorious Belfast address in 1874, as President of the British Association for the Advancement of Science. Tait's *Unseen Universe* (1875), co-authored with fellow Scottish physicist Balfour Stewart (1828–87), was an attempt to use the new physics 'to show that the presumed incompatibility of Science and Religion does not exist'.

Furthermore, William Thomson himself had earlier been drawn into public dispute with John Tyndall and T. H. Huxley. In 1868 Huxley had presented an address to the Geological Society of London, in which he insisted upon the extremely prolonged age of the Earth. Thomson responded the following year in his 'Of Geological Dynamics'. Here he used his own considerable expertise to establish, on thermodynamic grounds, that the Earth could not have been habitable for more than 100 million years, and by 1897 had revised this down to 20–40 million years. Again, it is clear that Thomson's motivation was not just to correct what he saw as bad science, but to counter attempts to suggest that religion had no place in science.

The first law derives from Joule's work, but the second confirms the irreversibility that is implicit in Carnot's idea that a heat engine is driven by the fall of temperature during its cycle. It is as though the water driving a waterwheel fell, after turning the wheel, into an abyss. The water has been lost to mankind and can no longer be used to do useful work (such as driving another water wheel). In Carnot's cycle not all the heat is lost, but it is impossible to use absolutely all of it to generate motive force.

The first law, therefore, is about the interconvertibility of different forms of energy, but the second is about the inevitability of the waste of some of that energy in the process.

In 1854, Thomson declared to a meeting of the British Association for the Advancement of Science that Joule's experimental work 'led to the greatest reform that physical science had experienced since the days of Newton'.

Subsequently, Thomson and his collaborators, including W. J. Macquorn Rankine (1820–72), Peter Guthrie Tait (1831–1901), Fleeming Jenkin (1833–85), and James Clerk Maxwell (1831–79), began to redefine natural philosophy as the study of energy and its transformations, and to regard energy as a unifying principle capable of explaining all phenomena. The major published statement of this enterprise was the *Treatise on Natural Philosophy* which he co-wrote with Tait in 1867. Although the authors re-interpreted Newton's third law ('action and reaction are equal and opposite') in terms of conservation of energy, they saw themselves as going beyond him, and writing a new *Principia*.

FURTHER READING

Peter J. Bowler and Iwan Rhys Morus, *Making Modern Science: A Historical Survey* (Chicago: University of Chicago Press, 2005), chapter 4.

Robert Fox, 'Laplacian Physics', in R. C. Olby, G. N. Cantor, J. R. R. Christie and M. J. S. Hodge (eds), *Companion to the History of Science* (London: Routledge, 1990), pp. 278–94.

Peter Harman, *Energy, Force, and Matter: The Conceptual Development of Nineteenth-Century Physics* (Cambridge: Cambridge University Press, 1982).

Iwan Rhys Morus, *When Physics Became King* (Chicago: University of Chicago Press, 2005).

Crosbie Smith, 'Energy', in R. C. Olby, G. N. Cantor, J. R. R. Christie and M. J. S. Hodge (eds), *Companion to the History of Science* (London: Routledge, 1990), pp. 326–41.

Crosbie Smith, *The Science of Energy* (London: Athlone, 1998).

Crosbie Smith and M. Norton Wise, *Energy and Empire: A Biographical Study of Lord Kelvin* (Cambridge: Cambridge University Press, 1989).

M. Norton Wise, 'Work and Waste: Political Economy and Natural Philosophy in Nineteenth-Century Britain', *History of Science*, 27 (1989), pp. 263–301, and 391–449; and 28 (1990), pp. 221–61.

23

Newton Deposed: Einstein and Relativity Theory

If physicists in the late nineteenth century saw themselves as going further than Newton, they at least saw themselves as being on the same road. Physicists in the twentieth century, however, saw themselves as displacing Newton and producing an entirely new physics. Writing to a colleague in 1931, Max Planck (1858–1947), founder of quantum theory, wrote of how he and his colleagues at the beginning of the twentieth century felt that the old approaches were fundamentally inadequate. 'Classical physics was not sufficient', he wrote, 'that much was clear to me'. But, in referring to 'classical physics' in this way, Planck was using a newly formed designation for earlier physics.

The twentieth century saw the emergence of two completely new ways of understanding the world, relativity theory, and quantum theory. As the proponents of each of these theories diverged from more traditional physics, they emphasized the break they were making by referring to the earlier physics as 'classical'. The word had the right ring to it – it managed to convey the worthiness, even the importance, of the older physics, while also making it clear that its day had passed. Certainly, those who saw themselves as forging a new physics at the beginning of the twentieth century, came to see themselves as inaugurating another scientific revolution, but they did not make such a clean break with the past as they subsequently presented it. Indeed, it is clear that both strands of the new physics took their starting points from what had gone before, and that the protagonists of the new ways of thinking simply began by trying to solve problems that had emerged in traditional physics.

We will consider the two strands of this new physics separately in the last two chapters of this book. We begin with the theory of relativity, and we proceed by showing how it grew out of matters arising in what we now call 'classical' physics.

Theories of light

Although Descartes had supposed that light was a pressure wave transmitted through the jostling matter which filled his universe, he often resorted to explaining the behaviour of rays of light by analogy with tennis balls bouncing off walls, for example. Newton, whose universe was largely empty of matter, felt no compunction, therefore, about treating light as streams of particles. Thomas

Young (1773–1829), professor of natural philosophy at the Royal Institution, having studied the nature of sound waves, favoured the pressure wave theory, which even Newton had hinted at in his so-called 'aether' queries (a group of 'queries' speculating on an all-pervasive universal aether which he interpolated into the existing queries in the third edition of the *Opticks* in 1717). The intellectual implications of this new theory were noticed by Augustin Fresnel (1788–1827), one of the new generation of French thinkers, together with Fourier, who brought to an end the dominance of Newtonian physics as promoted by Laplace and Berthollet (see previous chapter). Drawing on Young's demonstration of the interference patterns created when light is shone through two parallel slits (a pattern of parallel light and dark lines are observed), and on phenomena attributed to what is now called the polarization of light, Fresnel persuasively argued that light was not only a wave but it must be a transverse wave (Figure 23.1).

Previous wave theories of light had assumed that light, like sound, consisted of variations of density of the medium through which the wave travelled. The compression of the air caused by striking a drum caused a shock wave of successively compressed air to travel spherically outward from the drum. This kind of wave is called a longitudinal wave. Fresnel's insistence that light was a transverse

Fig. 267.

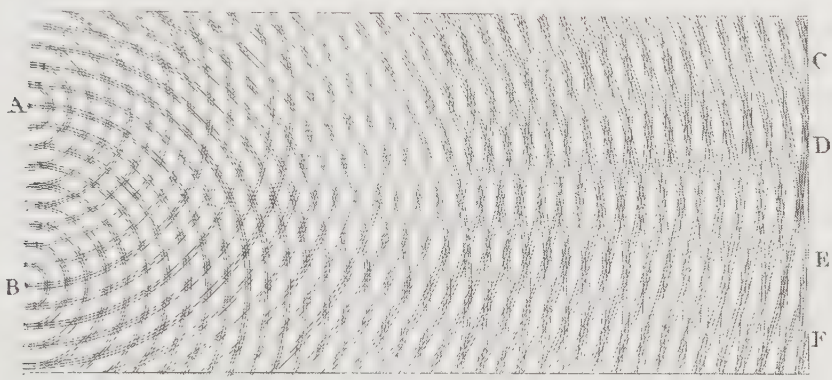

Figure 23.1 Illustration of the interference of two wave patterns from adjacent sources; from Thomas Young, *A Course of Lectures on Natural Philosophy and the Mechanical Arts*, 2 vols (London, 1807). Reproduced with permission from Edinburgh University Library, Special Collections Department (S.B.F., 53 You).

Thomas Young explained the parallel lines of light and dark thrown onto a screen when light was shone through two parallel slits in terms of what is now called 'interference'. Just as waves of water emanating from A and B (representing the parallel slits) sometimes reinforce one another (when the peak of a wave meets the peak of another, or trough meets trough), and sometimes cancel one another out (when the peak of one meets the trough of another), so light waves sometimes reinforce one another and remain bright, but in other places cancel one another out and darkness ensues. In the diagram, C, D, E, and F would supposedly be dark, while the places between would be light.

wave – that is, the kind of wave created by moving the end of a rope rapidly up and down, or side to side, so that the rope adopts a continuously moving sinuous shape – immediately created problems for physical models of how light worked. Young, who had also been led to notions of light as a transverse wave (because of the phenomenon of polarization), pointed out in 1823 that such a theory would require the aether to be 'not only highly elastic, but [also] absolutely solid!!!'

This problem about the nature of the aether was not such an obstacle to Continental theorists as it was to more mechanically minded British thinkers, who always required a workable physical model to be able to understand any physical phenomena. Even so, the transverse wave theory of light soon swept all before it and was accepted in Britain and America no less than on Continental Europe.

Theories of electricity and magnetism

Meanwhile, theories of electricity and magnetism were developing dramatically as a result of the accidental discovery in 1820, by the Danish natural philosopher Oersted, that a magnetic compass needle could be deflected from its north–south orientation by the flow of electric current in a nearby wire. In France Ampère discovered that a coil of wire acted like a magnet when an electric current passed through it, and concluded that magnetism was the result of the motion of electricity. Shortly after, William Sturgeon (1783–1850) began experimenting with so-called electromagnets (a horseshoe-shaped soft iron bar with a wire wound around it). In 1831, Faraday, at the Royal Institution, was able to show that a moving magnet produced an electric current in a surrounding coil of wire, and Sturgeon developed an electric motor capable of producing motive force. Joule, whom we met in the previous chapter, performed investigations of electromagnetic motors in the late 1830s before turning to investigate heat.

Faraday, seeking to make sense of these phenomena developed his notion of 'lines of force', and envisaged them as permeating space around magnets and sources of electricity, as could be made visible, for example, in the familiar experiments in which iron filings form filaments curving around from one pole to another of a bar magnet. In an early speculative note (1832) Faraday wrote that he was

> inclined to compare the diffusion of magnetic forces from a magnetic pole, to the vibrations upon the surface of disturbed water, or those of air in the phenomena of sound: i.e. I am inclined to think the vibratory theory will apply to these phenomena, as it does to sound, and most probably to light.

By the 1840s he was developing the idea (which was discussed at the end of Chapter 14) that the lines of force were more real than matter itself, which could simply be seen as focal points for the emanation of lines of force pervading the whole universe.

In 1846 he settled the idea of a link between magnetism (and therefore electricity) and light when he discovered that a magnet could rotate the plane of polarization of a polarized ray of light. A light beam is polarized when all of the waves of light are undulating in the same plane, let's say vertically up and down (normal, non-polarized light beams include light that is undulating in every plane, some up and down, some side to side, some on the diagonal plane between, and others on every other possible plane between). Faraday had now discovered that a beam of light, polarized so that all the light waves are orientated with vertical up and down motion, could be rotated by a magnet so that all the waves then undulated closer to the horizontal.

At a talk at the Royal Institution in the same year, Faraday effectively suggested that his image of *lines* of electromagnetic force was better able to explain phenomena of light than the image of an all-pervading fluid aether:

> The view which I am so bold as to put forth considers, therefore, radiation as a high species of vibration in the lines of force which are known to connect particles, and also masses of matter, together. It endeavours to dismiss the aether, but not the vibrations.

The advantage of Faraday's image was that, whereas a fluid aether suggested longitudinal pressure waves (like sound waves), vibrations in *lines* of force, could be envisaged as transverse waves, like so many skipping ropes emanating from a source of light.

Faraday was undeniably a brilliant and insightful thinker but, like most of us, he was hampered by a lack of facility in mathematics. Even his genius could only take him so far in a world where physics was becoming increasingly mathematical. His ideas on lines of force were not, however, unamenable to mathematical treatment. On the contrary, James Clerk Maxwell (1831–79), one of the greatest of Victorian mathematical physicists, provided a geometrical model of Faraday's ideas in an article 'On Faraday's Lines of Force', written in 1856. He extended this model in a four-part paper published in 1861–2, 'On Physical Lines of Force'.

Maxwell evidently felt the need to offer a model of Faraday's fields of force which was physical, indeed mechanical (see Box 22.2), and presented an elaborate scheme of vortices arranged in a honeycomb pattern, surrounded by spherical particles rotating in the opposite direction and acting as idle wheels to transmit the motions of the vortices to other parts of the field. The vortices themselves were inspired by earlier responses, notably by William Thomson, to the magneto-optical rotation discovered by Faraday (in which it was assumed that a rotation in the magnetic force field could cause a rotation – change of polarization – of transverse light waves). Although this complex model provided precise mechanical explanations for the different electromagnetic quantities discerned by the physicists, Maxwell himself admitted its awkwardness and expressed doubts as to this field of force as 'a mode of connexion existing in nature' (Figure 23.2).

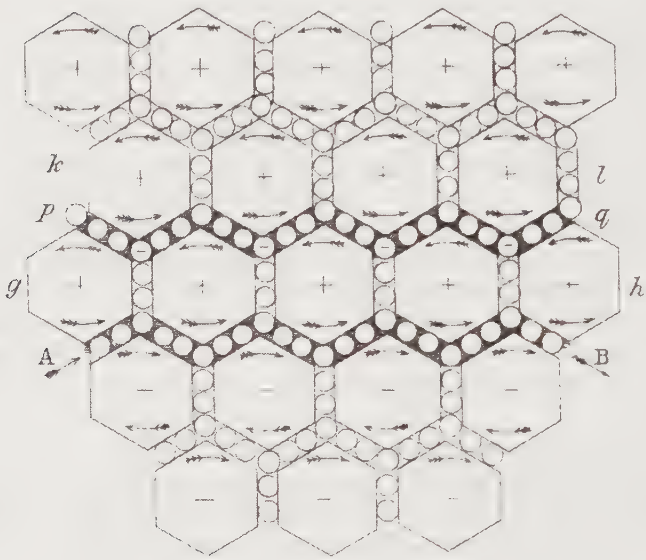

Figure 23.2 Maxwell's depiction of the electromagnetic aether from his article 'On Physical Lines of Force', in *The London, Edinburgh, and Dublin Philosophical Magazine*, April 1861. With thanks to the Trustees of the National Library of Scotland.

The rotating vortices required by Maxwell's theory are shown as hexagons. If we imagine these hexagons as cogwheels intermeshing with one another, then the rotation of one in an anti-clockwise direction would force a vortex intermeshed with it to rotate in a clockwise direction. Maxwell needed them all to rotate in the same direction, and so he introduced the 'idle wheels', as he called them, between the vortices – shown as the small circles (representing spheres) separating the hexagons. Maxwell on the one hand acknowledged the implausibility of this scheme, while on the other insisting that it served 'to bring out the actual mechanical connections between the known electromagnetic phenomena'.

The most important aspect of this work, however, was that when Maxwell calculated the speed of transmission of waves across such an electromagnetic medium the answer was the speed of light (which had just been measured with great accuracy by Léon Foucault, 1819–68). Accordingly, Maxwell jumped to a dramatic conclusion:

> We can scarcely avoid the inference that *light consists in the transverse undulations of the same medium which is the cause of electric and magnetic phenomena.*

By 1865, in his 'Dynamical Theory of the Electromagnetic Field', Maxwell dropped the mechanical model of the aether and concentrated instead on the mathematics of the electromagnetic field. His resulting sets of equations were reduced later to what are now known as the four Maxwell equations by Oliver

Heaviside (1850–1925) and Heinrich Hertz (1857–94). Hertz also established experimentally that, as Maxwell had predicted, there were other forms of electromagnetic waves, radio waves, with much longer wavelengths, but otherwise part of the same spectrum, a spectrum that was now recognized – given the links between electricity, magnetism, and light – to be an 'electromagnetic spectrum'.

It is not clear, however, that Maxwell intended to completely abandon the typical British mechanical approach in favour of a more abstract mathematical approach. In his *Treatise on Electricity and Magnetism* of 1873 he rejected attempts by continental mathematicians to dispose of the notion of a field and return to forces acting at a distance, by insisting on the reality of energy and that 'there must be a medium or substance in which the energy exists'. He evidently still wanted 'to construct a mental representation' of the aether. The precise status of the aether for Maxwell, and his contemporaries, is still being debated by historians and philosophers of science.

What is clear, however, is that Maxwell's equations were seen as worthy of standing alongside Newton's laws and his universal principle of gravitation, so that together they could be used to sum up the whole of physics. There was, however, a slight problem; what the philosopher of science Jonathan Powers called an 'almost invisible, but fatal, crack' in the edifice of classical physics.

Maxwell's equations included a constant, designated as c, which was then, because of its value (as we saw above), identified with the speed of light. But, what was the speed of light? Surely, the answer to that question depends on the circumstances – according to Newtonian mechanics the motion of the source of the light would need to be taken into account. If I am sitting at the front of a rocket ship flying at 150,000 kilometres a second, and I switch on a torch pointing in the direction of travel, then the light leaving the torch ought to have a speed of about 450,000 kilometres a second (the speed of light in air having been measured at very nearly 300,000 kilometres per second). The addition (and subtraction) of speeds in this way is a standard assumption in classical physics (we've already met it in the account of Galileo's theory of the tides – in Chapter 9). But Maxwell's equations make no provision for such concerns – according to them the speed of light is simply a constant. Perhaps Maxwell was wrong to identify this mathematical constant with something physical, such as the speed of light?

It seems that there is something precarious, after all, about combining Newton's laws and Maxwell's equations. It might look as though, taken together, they can be used to solve all problems in physics, but appearances, as we all know, can be deceptive.

Aetherial troubles

As if this wasn't bad enough, empirical efforts to establish the existence of the aether were also failing.

The most famous of these attempts were carried out over a number of years by Albert A. Michelson (1852–1931), who was later joined by Edward Morley (1838–1923). What Michelson and Morley did, culminating with great accuracy in 1887, was to measure the speed of light in different directions through space. Their apparatus was designed to split a beam of light into two components which were then sent out at right angles to one another, and reflected back. But the apparatus could be rotated so that the two directions in which the light travelled could be varied at will (although always maintaining the 90º separation between the two beams).

The intention was to detect any variation in the speed of light due to the motion of the Earth (which would have to be added to, or subtracted from, the speed of light). As the Earth moves through the aether, there should be an 'aether wind' – a current set up by the motion of the Earth through the stationary aether. The speed of a beam of light travelling into the headwind, and coming back with the wind behind it, should differ measurably from the speed of a beam travelling across the line of the wind.

It should; but as far as Michelson and Morley could tell, it did not. But nobody panicked. Physicists didn't give up.

By the end of the nineteenth century the electrodynamic theories of the Dutch physicist Hendrik A. Lorentz (1853–1928) seemed to answer all difficulties by introducing what came to be called the 'Lorentz transformations'. Lorentz assumed that matter was composed of electrically charged particles and that their motions somehow generated the all pervasive electromagnetic field around them. This field was then able, in turn, to affect the movements of the charged particles. Lorentz used these assumptions to suggest that the electromagnetic field of the aether caused bodies moving through it to contract in the direction of motion. The contraction was due to the field affecting the forces acting between the atomic particles of the body.

According to the mathematics, the contraction was precisely that required to thwart the empirical efforts to detect the difference in the speed of light expected as a result of the Earth's motion through the aether. Even Michelson and Morley's apparatus shrank in the direction of travel through the aether by just the amount needed to make measurement of the speed of light in that direction the same as the speed measured at a right angle to it.

Those of us who are not mathematicians might be wondering if there was some fudging taking place here: it all seems too convenient. But this coincidence is no more surprising than the coincidence that Kepler's geometrical archetype really did seem to define the positions and number of the planets visible to naked eye astronomy (see Chapter 8). Besides, there was no denying that the Michelson–Morley experiment always returned a null result, and yet if the Earth was perpetually moving through a stationary aether it ought not to have done.

Certainly, Lorentz believed the contraction was a real physical phenomenon, but he was less sure about another 'transformation' that he introduced to maintain the validity of Maxwell's equations. The absence of anticipated differences between systems in motion and systems at rest (where the speeds of light ought to be different, but which according to Maxwell's equations remain constant) could be explained, hypothetically, by assuming that observers in either situation made the same observations as each other (rather than slightly different observations due to their relative motions). This depended on what Lorentz called the 'local time' of the moving experimenter.

Perhaps Lorentz did not clearly see how this explanatory device could work physically (as opposed to mathematically), but the leading French physicist Henri Poincaré (1854–1912) suggested in 1900 that it was due to the fact that moving experimenters synchronized their clocks (with clocks that are stationary relative to the aether) using light signals that are assumed to be unaffected by their (the experimenters' and the clocks') motions (but really the light should be affected by their motions – hence the need for the Lorentz transformation).

It was shortly after this that Albert Einstein (1879–1955) stepped onto the stage.

Albert Einstein and the theory of relativity

Einstein has entered popular consciousness in the modern world like no other scientist before him, or since (so far). The highly individualistic look he developed in his old age has, ironically, come to represent the image of the scientist. Invoked by other cultural icons such as Marilyn Monroe (1926–62) and Bob Dylan (1941–), his face appears on the kind of posters which more usually depict pop singers or film stars and which can be found on sale on any high street. Einstein's posters are only distinguished from the rest by often including a supposedly amusing, or profound quotation (not usually available from popstars or film stars). Moreover, in his own lifetime he was very much a public figure, taking part in later life in anti-nuclear and civil rights movements; he was even offered the Presidency of Israel in 1952 (he had been prominent in efforts to establish the new state after the Second World War), though he turned it down. This isn't the place to discuss why he has become such a cultural icon, but it is hardly deniable that his science had much to do with it. If nothing else, Einstein's celebrity shows the huge impact that the latest ideas in science have had, and continue to have, on modern popular consciousness.

In view of Einstein's celebrity it is all too easy to see him as a unique genius who produced relativity theory out of nothing, and single-handedly deposed Newton, and undermined classical physics. This was certainly not the case, however. Einstein owed a great deal to the earlier work of Poincaré and Ernst Mach (1838–1916), and for a long time Einstein's new approach was routinely referred to as the Lorentz–Einstein theory. Even those who recognized the differences between Lorentz's and Einstein's views linked them together. Max Planck, for example, saw Einstein's theory of relativity as merely a more general form of Lorentz's principle. Indeed, Einstein himself once referred to 'the theory of

Lorentz and Einstein'. Furthermore, for many the theory only became acceptable after Hermann Minkowski (1864–1909) introduced an easier way of understanding the mathematics involved, by combining space and time into a four-dimensional non-Euclidean geometry, the space-time continuum (1906). As Max Planck declared in 1910:

> Under the pioneers of the new terrain we have first to mention Hendrik Antoon Lorentz, who discovered the concept of relative time and introduced it into electrodynamics... then Albert Einstein, who first had the boldness to proclaim the relativity of time as a universal postulate; and Hermann Minkowski, who succeeded in fashioning the theory of relativity into a consistent mathematical system.

With regard to Einstein's scientific success, his achievements seemed to stem from what he called 'my need to generalize', a need which perhaps had its origins in his reading of popular science books when he was a boy. According to two Einstein scholars, Jürgen Renn and Robert Schulmann, his preferred childhood reading consisted of introductory texts that presented the 'big picture', rather than dwelling on details. In an address that Einstein delivered in 1918 he spoke of the importance of providing a coherent worldview and insisted that 'the supreme task' of the physicist was 'to arrive at those universal elementary laws from which the cosmos can be built up'.

When Einstein graduated from the Zurich Polytechnic Institute in 1900 he had a solid grounding in the latest developments in physics, but he had also been profoundly affected by reading *The Science of Mechanics* (1883), a history of physics written by the Austrian physicist and philosopher, Ernst Mach. In some ways Mach can be seen as the culminating figure in the Continental tendency among physicists to avoid physical models and to rely on a more abstract mathematical, or positivist approach (accepting only what is directly accessible to the senses). Although many of his contemporaries in science felt he took things too far. Mach's rejection of the concept of atoms (on the grounds that their existence could never be empirically established), for example, was seen by many as incompatible with a scientific approach. Mach's *Science of Mechanics* also included a vigorous denunciation of Newton's concepts of absolute space and time, declaring them to be physically meaningless. It seems clear that Einstein took note of this.

When Einstein turned his attention to the 'Electrodynamics of Moving Bodies', which he published in the *Annalen der Physik* in 1905, his starting point was to show that in electrical induction what was important was only the relative motion of the magnet or the wire coil with regard to one another. Maxwell had offered different accounts depending upon whether the magnet or the coil is at rest, but Einstein insisted that either way, the case was the same. Furthermore, since all attempts have failed to measure the Earth's motion relative to an aether supposedly at absolute rest, absolute rest cannot be established and should be rejected.

Einstein's second point was simply to accept the clear implication of Maxwell's equations – that the speed of light is a constant. Where Lorentz, Poincaré and others had tried to adjust Maxwell's conclusions to make them fit in with Newtonian mechanics, Einstein simply assumed that the blame for the incompatibility lay with Newton, and that Maxwell was right: no matter how a system is moving, the speed of light is always a constant. If I am sitting at the front of a rocket ship travelling at 150 kilometres per second and I shine a torch in the same direction, the light from the torch will be travelling at very nearly 300 kilometres per second – just as it would if I shined my torch while standing on the Earth. Even with the powerful support provided by Maxwell's equations, this was a bold and radical move.

In order to work out just how Newtonian mechanics should be modified to make it fit in with these two presuppositions, Einstein turned to the work of Lorentz and Poincaré and, among other things, generalized Lorentz's notion of local time. Einstein developed his own transformations which were *mathematically* equivalent to the Lorentz transformations but the underlying physics was held to be dramatically different. Where Lorentz explained contraction in terms of changes in the attractive forces between the particles of a body, Einstein explained them in terms of changes in space itself. Again, this was a very bold move.

Although it was bold, it followed inexorably from Einstein's determination to uphold Maxwell's conception of the speed of light as a constant. If we imagine two experimenters passing by one another in transparent rocket-ships (so each can see what the other is doing), and we imagine that they each shine a ray of light from the back of their rocket-ships to the front, and then try to measure the speed of each other's light rays using a stop-watch and a measuring stick (or whatever high-tech equivalents you require to measure the speed of light – if your imagination is up to scratch a watch and a measuring stick ought to do), the result seems obvious. The measured speed of light will vary according to the relative movements of the passing rockets. If one overtakes the other at the speed of light, the ray of light in the slower ship should seem to be virtually at a stand-still to the overtaking experimenter, while the ray of light in the overtaking rocket will seem to be moving twice as fast as the light in the slower rocket when observed from that slower rocket. This is exactly like the case where I throw a ball in the direction of travel at 30 miles an hour from the front of a car moving at thirty miles an hour. An observer standing at the roadside would measure the speed of the ball at 60 miles an hour. Ordinarily, we add or subtract the relevant speeds, in accordance with the relative motions involved in the scenario. But if we wish to insist that the speed of light remains constant in all scenarios – as Einstein, following Maxwell, did – then we have to assume that something must happen to watches and to measuring sticks in these circumstances. The only way the speed of light can always be the same is if times and distances vary in such a way as to make it so.

By thinking about space and time and how we measure each of them, and by imaginatively considering how experimenters in different frames of reference, say passing rocket ships (passing with uniform motion) travelling near the speed of light, would measure what was going on in the different frames, Einstein developed a new physics. The relativization of space and time was to have profound and counter-intuitive consequences – indeed it has been said that the consequences of Einstein's theory were not difficult to understand, merely difficult to believe. The faster something moved the more it contracted along the direction of motion, and the slower time passed for it; only this way could the speed of light be maintained as a constant. But that wasn't all; a speeding body also increased in inertial mass. Einstein noticed that the increase in mass with speed was proportional to the energy of its motion, and this led him to consider that mass and energy were equivalent concepts. Furthermore, just as other kinds of energy could be converted into one another (see previous chapter), Einstein showed that inertial mass could be converted into energy, again using Maxwell's conversion factor, the speed of light: $E = mc^2$ (see Box 23.1).

Einstein's 'On the Electrodynamics of Moving Bodies' gave rise to what is now known as the theory of relativity, but it was a restricted theory, confined to

Box 23.1 $E = mc^2$ AND THE ATOMIC BOMB

After John Cockroft (1897–1967) and Ernest Walton (1903–95) split the nucleus of the atom in 1932 by bombarding lithium with accelerated protons (hydrogen nuclei), they also succeeded in confirming Einstein's famous equation $E = mc^2$. The lithium atom was split into two helium nuclei which had slightly less mass than the original atom and the proton that split it. The lost mass had been converted into energy in accordance with Einstein's equation. Although the lost mass was very small, because its value was multiplied by the speed of light squared, the amount of energy was quite significant (17 million electron volts). This is the basis of the atomic bomb – splitting the nucleus, or nuclear fission, could be used to make a bomb of immense destructive potential. But this leap was not made from Cockcroft and Walton's experiment. Leó Szilárd (1898–1964), Enrico Fermi (1901–54), and Frédéric Joliot-Curie (1900–58) were among the first to think of using radioactive materials (natural emitters of sub-atomic particles) to bombard themselves and thereby set up a so-called 'chain reaction' (if uranium-235 is split, for example, it emits neutrons which can cause other uranium-235 atoms to spilt, so producing neutrons which can split yet more uranium-235 atoms…). It seemed possible that this kind of self-replicating nuclear fission might be exploited to create a bomb.

On the eve of the outbreak of the Second World War, Leó Szilárd and others, fearing that the Nazis might succeed in developing such a bomb, drafted a warning letter to President Franklin D. Roosevelt (1882–1945) and persuaded Einstein (a celebrity even then) to put his name to it. The letter, dated 2 August 1939, opened like this:

> Some recent work by E. Fermi and L. Szilard, which has been communicated to me in manuscript, leads me to expect that the element uranium may be

systems in uniform motion (or at rest) – hence the designation special theory of relativity. Einstein, feeling the need to generalize, was restless to extend his ideas to cover accelerated motions and began to work on what he first called his 'generalized' theory of relativity in 1907.

Meanwhile, the world of professional physics was slow to catch on to Einstein's special theory. One of Einstein's former teachers, Hermann Minkowski, recognizing its importance, was able to make it more mathematically accessible by uniting space and time into four-dimensional space-time. Gradually numerous empirical tests of various dynamical, optical, and electromagnetical consequences of Einstein's special theory were carried out and the theory passed them all. Even so, for a long time British physicists, for example, continued, in the spirit of Lorentz, to make adjustments to Maxwellian theory and to refine their models of the aether. French physicists paid little attention until the 1920s, after the appearance of the general theory.

The general theory, encompassing accelerating as well as uniform systems, proved a tough nut to crack but it gained even greater significance for Einstein

> turned into a new and important source of energy in the immediate future. Certain aspects of the situation which has arisen seem to call for watchfulness and if necessary, quick action on the part of the Administration. I believe therefore that it is my duty to bring to your attention the following facts and recommendations.
>
> In the course of the last four months it has been made probable through the work of Joliot in France as well as Fermi and Szilard in America – that it may be possible to set up a nuclear chain reaction in a large mass of uranium, by which vast amounts of power and large quantities of new radium-like elements would be generated. Now it appears almost certain that this could be achieved in the immediate future.
>
> This new phenomenon would also lead to the construction of bombs, and it is conceivable – though much less certain – that extremely powerful bombs of this type may thus be constructed. A single bomb of this type, carried by boat and exploded in a port, might very well destroy the whole port together with some of the surrounding territory.

This eventually led to the establishment of the so-called Manhattan Project in 1942, which brought together many of the leading physicists of the day, under the direction of J. Robert Oppenheimer (1904–67), and which eventually succeeded in exploiting radioactive chain reactions to create a bomb with all the destructive force implied by Einstein's equation.

FURTHER READING

Jeff Hughes, *The Manhattan Project: Big Science and the Atom Bomb* (Cambridge: Icon Books, 2003).

after he hit upon what is known as his 'principle of equivalence' – which he described as 'the happiest thought of my life'. Einstein realized that a system, a frame of reference, affected by a gravitational field is equivalent to a frame of reference accelerating at the same rate as the acceleration produced by the gravitational field. We are all familiar with the fact that when a lift, or an elevator, begins to take us upwards through a tall building we feel as though we are being pulled downwards. The effect is due to the acceleration we undergo as the lift starts to move. However, we would experience exactly the same effect if a gravitational body, exerting the right amount of downward attraction on us, were to suddenly pass beneath the lift while it was stationary. The effects of acceleration, and of gravity are the same. If this is so, then it ought to be possible to consider gravity not as a mysterious force capable of operating at a distance, but as reducible to kinematics – that is to say, reducible to the physics of motion. What this meant was that a general theory of relativity would not just extend to accelerating bodies, but could also embrace gravitation and possibly lead to its unification with the other forces, such as electricity and magnetism, as once envisaged by Faraday (see the end of Chapter 14), and as noted as an important desideratum by Maxwell.

Using the principle of equivalence, Einstein soon predicted that time would run slower in an increased gravitational field (in the same way that it runs slower at higher speeds), and that gravitational objects would act as 'lenses', bending light rays which passed them by. Again, these predictions have been subjected to repeated empirical tests and completely confirmed.

Reminiscing in 1922, Einstein claimed to have realized in 1912 that if all accelerated (and gravitational) systems are equivalent, then Euclidean geometry cannot hold in all of them. Eventually, by 1915, with the help of the mathematician Marcel Grossman (1878–1936) who introduced him to the non-Euclidean geometry developed by Bernhard Riemann (1826–66), Einstein realized that gravitational attraction could be understood as the result of the distortion of space-time caused by a large mass. The Sun's mass, for example, deforms space, bending it in such a way that the inertial paths of the planets are bent around it in an ellipse. The planets are not acted upon by a force acting at a distance, they are merely following the curvature of space caused by the nearby Sun. Gravity was not really a separate force at all, but merely an aspect of the intrinsic nature of space-time.

Newton's reign as the greatest physicist who had ever lived, was well and truly over.

FURTHER READING

Peter J. Bowler and Iwan Rhys Morus, *Making Modern Science: A Historical Survey* (Chicago: University of Chicago Press, 2005), chapters 4 and 11.

Peter Dear, *The Intelligibility of Nature: How Science Makes Sense of the World* (Chicago: University of Chicago Press, 2006), chapter 5.

Peter Harman, *Energy, Force, and Matter. The Conceptual Development of Nineteenth-Century Physics* (Cambridge: Cambridge University Press, 1982).

Peter Harman, *Metaphysics and Natural Philosophy. The Problem of Substance in Classical Physics* (Brighton: Harvester Press, 1982).

Peter Harman, *The Natural Philosophy of James Clerk Maxwell* (Cambridge: Cambridge University Press, 1998).

Tetu Hirosige, 'The Ether Problem, The Mechanistic Worldview, and the Origins of the Theory of Relativity', *Historical Studies in the Physical Sciences*, 7 (1976), pp. 3–82.

Gerald Holton, *Thematic Origins of Scientific Thought from Kepler to Einstein* (Cambridge: Cambridge University Press, 1973).

Sean F. Johnston, *Holographic Visions. A History of New Science* (Oxford: Oxford University Press, 2006).

Helge Kragh, *Quantum Generations: A History of Physics in the Twentieth Century* (Princeton: Princeton University Press, 1999).

Iwan Rhys Morus, *When Physics Became King* (Chicago: University of Chicago Press, 2005).

Abraham Pais, *'Subtle is the Lord..': The Science and the Life of Albert Einstein* (Oxford: Oxford University Press, 1982).

Jonathan Powers, *Philosophy and the New Physics* (London: Methuen, 1982).

Crosbie Smith, 'Energy', in R. C. Olby, G. N. Cantor, J. R. R. Christie and M. J. S. Hodge (eds), *Companion to the History of Science* (London: Routledge, 1990), pp. 326–41.

Crosbie Smith, *The Science of Energy* (London: Athlone, 1998).

Loyd S. Swenson, *The Genesis of Relativity: Einstein in Context* (New York: Burt Franklin & Co., 1979).

24

Mathematics Instead of a World Picture: From Atomism to Quantum Theory

The history of physics, which we have been looking at in the last two chapters, owed more and more to mathematics. Many of the innovations were in fact only possible because of newly developed, or newly appropriated, mathematical techniques. Needless to say, these were all highly technical and defy any attempt to make them easily understandable to the non-mathematician (it would take many pages, and even then there would be a real danger that the attempt would result not just in simplification but in misdirection, and therefore a failure to properly understand the significance of the mathematics). Among others, these new techniques included quaternions and bi-quaternions, non-Euclidian geometries, and tensor calculus. The culmination of this trend can be seen in the development of quantum theory and its extensions, such as quantum electrodynamics. The mathematical approach to understanding nature became so dominant that it effectively supplanted earlier efforts to offer physical models as a way of understanding the world.

Max Born (1882–1970), Werner Heisenberg (1901–76), and Paul Dirac (1902–84), leading figures in the development of quantum theory, developed so-called matrix calculus and effectively sneered at any attempt to explain quantum theory in physical, visualizable, terms. Heisenberg was even scornful of the more physically accessible work of Erwin Schrödinger (1887–1961), another leading player in quantum physics and no mathematical slouch himself. Dirac, writing in 1930, insisted that:

> The only object of theoretical physics is to calculate the results that can be compared with experiment... it is quite unnecessary that any satisfactory description of the whole course of the phenomena should be given.

It seems true to say that quantum theory simply cannot be reduced to easily visualizable analogies without introducing misleading elements into the picture. Its success in dealing with experimental results testifies to the fact that quantum theory is highly successful and really does seem to capture some fundamental truths about the nature of the world. Nevertheless, for most of us, the fact that quantum theory cannot provide an explanation of what is going on in the microworld of atomic physics in terms of causes and effects makes it seem unscientific. This is perhaps what is most dizzying about quantum physics – the fact that it offers no proper physical explanation of how the microworld works. As Niels Bohr (1885–1962), chief architect of quantum theory, once said, 'if anybody says he can think about quantum theory without getting giddy it merely shows that he hasn't understood the first thing about it!'

Perhaps Richard Feynman (1918–88), one of the founders of quantum electrodynamics, had this in mind when he insisted that nobody (including himself) really understood it. If he had understood it he would, as Bohr suggested, have been too giddy to be able to make a serious contribution. The fact that he was awarded the Nobel Prize in 1965 indicates that he *did* make a serious contribution, so he couldn't have been giddy enough to have really understood it. Bohr is reputed to have delighted in paradoxes as a starting point for making progress, and there can be little doubt that quantum theory has been supremely successful and has led to great progress in physics and chemistry, but it has done so by leaving behind any attempt to explain the world in pictorial or any other physical terms.

Let us look, nevertheless , at how physical science came to this pass. As we'll see, in its beginnings, it was an attempt to offer a visualizable picture of the atom.

Bohr's picture of the atom

The familiar picture of the atom as a miniature solar system with the nucleus, like the Sun, at the centre perpetually circled by electrons first emerged as a result of the work of J. J. Thomson (1856–1940), Cavendish Professor of Physics at Cambridge. Extending earlier work which had demonstrated that 'rays' bearing a negative electric charge were emitted from the negative electrode when an electric current was discharged between two plates in a glass tube containing gas at low pressure, Thomson came to the conclusion that the rays consisted of streams of particles, or corpuscles, much smaller than hydrogen atoms, each of which was negatively charged. Now known as electrons, Thomson assumed that these were a constituent of the atoms of the electrode which had been stripped out and that, therefore, atoms were not homogeneous indivisible particles. Thomson's subsequent picture of the atom was one in which the atomic weight was made up by large numbers of electrons (1,836 in the hydrogen atom), which were kept together in the atom by being surrounded by a sphere of positive charge (often referred to as the 'plum pudding' model).

Ernest Rutherford (1871–1937), a former student at the Cavendish Laboratory who was then professor of physics at Manchester, surmised that one way of understanding the structure of the atom would be to try to split it. In 1909, using what he had called alpha rays, which are given off by radioactive substances and by the positive electrode in a gas discharge tube, and which were evidently very heavy and emitted with high energy, Rutherford bombarded thin sheets of metal foil. Rutherford found that most of the particles passed straight through, with only a few being deflected (or even more rarely, bounced straight back). Rutherford concluded that, rather than being like a stodgy plum pudding, the atom must consist mostly of empty space, but with something at its heart capable of bouncing a heavy alpha particle back. The result of these speculations was our familiar planetary system model of the atom.

The trouble was that this model could not possibly work. Orbiting electrons would emit their energy and very rapidly spiral into the nucleus. The picture was saved, however, by bringing in a new concept which had been recently developed in another area of physics.

In thermodynamics, attempts to understand the behaviour of a hypothetical 'black body' – that is to say a body which absorbs and emits all radiation with perfect efficiency – had proved elusive. The problem was that theoretical expectations bore little relation to experimental results using the closest thing to a 'black body' that could be physically achieved. Satisfactory explanations for the empirical results could be established for high frequency radiations, or for low frequency radiations, but not for both. In 1900 Max Planck, professor of physics at Berlin, cut the Gordian knot by assuming (in what he called an act of despair) that energy was not radiated continuously, as was always taken for granted, but only in discreet packets, or quanta.

Physically unjustifiable as this was – Planck regarded it merely as a mathematical trick – the important thing was that it seemed to provide the solution. At higher frequencies the black body can only radiate large packets of energy, at lower frequencies, however, it can radiate smaller packets. Planck's mathematics led him to a constant, which he called the quantum of action but is now known as Planck's constant, h, which, when multiplied by the frequency, determined the energy involved.

The implication of this, which was incompatible with classical mechanics, was that when electromagnetic radiation, such as heat or light, interacted with matter it did so only at particular levels of energy, not at others. Attitudes to the notion of a quantum of action began to change, however, when Einstein used it (in 1905) to offer an explanation of the photo-electric effect.

It had been noted that light and other parts of the electromagnetic spectrum, notably ultra-violet radiation, could cause the emission of electrons from metal and other surfaces. It had seemed strange that the effect did not depend upon the intensity of the light but its frequency. Einstein proposed that an electron could only be knocked out of a surface by light with the right energy level. What Einstein called 'light quanta' were required. Increasing the intensity of light simply meant firing more light quanta at the surface, but these would have no effect if the projectiles had insufficient power to knock an electron out of the metal surface, only projectiles with sufficient power could do it. These light quanta, these projectiles, looked suspiciously like particles and they later came to be called photons. Einstein seemed to be rejecting the wave-theory of light and returning to the Newtonian idea that light was particulate.

The problems with Rutherford's model of the atom could also be solved by recourse to quanta but it was not Rutherford who saw it, but a young postdoctoral student from Copenhagen in Denmark, who had gone to Manchester to study with him, Niels Bohr (1885–1962). In Bohr's variation on the Rutherford atom, announced in 1913, the orbiting electrons could not spiral into the nucleus as they continuously emitted radiation, because they could not *continuously* emit radiation. They could only emit (or absorb) energy in quanta, in packets of specific sizes, and this meant that they must otherwise remain in stable orbits.

Bohr's model could also account for the highly specific (indeed characteristic) spectra displayed by particular substances (at least it could for hydrogen, and therefore in principle, if not in practice, it could for more complex atoms). The emission spectrum of hydrogen shows characteristic lines which were now realized to correspond to the frequencies of the light emitted when hydrogen's single electron jumped from a higher orbit down to a lower one. Electrons could only jump from one orbit to another; they could not move continuously across the gap. Accordingly, as they jumped from one orbit to another they were also jumping from one energy level to another and had to emit (or absorb) the corresponding difference in energy, which in the case of emission appeared to an observer (or a recording apparatus) as quanta of light or photons.

Bohr's model was an instant success even though there were obvious problems with it. In particular, it was seen to be a mish-mash of incompatible ideas – classical physics on the one hand, and ideas taken from the new speculations about quanta on the other. In the sequel, however, classical physics was to be gradually forced out of the picture as the new quantum theory increasingly took over.

The new quantum theory

The ideas developed by Planck, Einstein and Bohr are often referred to as the 'Old Quantum Theory', and in a sense did not constitute a unified theory at all, merely a number of *ad hoc* ways of dealing with particular problems. The new quantum theory, developed around 1926, by Bohr, Heisenberg, Shrödinger, Dirac and others, was a distinct theory, however.

One of Bohr's early collaborators after he had moved back to Copenhagen was Werner Heisenberg. In 1925, Heisenberg, then working in Göttingen with the highly mathematical Max Born, and self-professedly enamoured of the positivism of Ernst Mach, decided to reject all attempts to form a physical picture of the atom. Seeking only to provide a workable mathematical account of what was empirically observable, Heisenberg arrived, by 1927, at his famous 'uncertainty principle', or as he preferred to call it, the 'indeterminacy principle'.

Heisenberg was able to show that there is an inbuilt indeterminacy in nature which made it impossible to precisely determine both the position and the momentum of an electron at the same time. The uncertainty was *mathematically defined* and therefore built in to Heisenberg's physics. Heisenberg himself tried to present a physical account of why these two aspects of the electron could not be simultaneously established but in so doing he presented a misleading picture.

Even now, it is often said in popularizations of science that the problem arises because in trying to determine the position of the electron the experimenter must shine a powerful light on it, in order to see it; but (at the level of the sub-microscopic) light waves will unavoidably have a major impact on the electron and will change the momentum of the electron at the moment that you fix its position (it is as if you tried to establish the position of a kitten, say, by directing a high-pressure jet of water at it).

As with all such analogies in quantum theory, the result is misleading. It seems to suggest that, given enough ingenuity, a scientific instrument maker might be able to build an apparatus which could measure the momentum of an electron but could only do it at a very precise location – any electron which arrived at that location would have its momentum measured, and its location at that moment would also be known. But Heisenberg's uncertainty principle is more robust than that. To avoid such clutching at visualizable straws, it is perhaps safer to say that Heisenberg's principle states that an electron does not at any given time have both a precise location and a specific momentum. Only one or the other can be mathematically, and therefore physically, defined. Following on from this, it has to be concluded that quantum mechanics cannot be deterministic. In order to predict the future behaviour of a particle, both its initial position and its initial momentum would have to be known. Since both cannot be known, deterministic calculations are impossible (see Box 24.1).

This is bound up, of course, with the famous wave–particle duality, which is another seemingly baffling aspect of quantum theory. If an electron is not a particle at all, but a wave, then it is difficult to define its precise position anyway. But let us take a closer look at this aspect of quantum theory.

Einstein's treatment of the photo-electric effect seemed to suggest that light was particulate and yet it had long been accepted that light was a wave formation in the electromagnetic aether. While the jury was still out, Louis de Broglie (1892–1987), a historian turned physicist, drew upon the equivalence of mass and energy expressed in Einstein's $E = mc^2$ equation, to suggest that electrons should also have the properties of both particles (matter) and waves (energy). This theory was enthusiastically endorsed by Einstein and empirically confirmed in 1925 when streams of electrons – supposedly particles – were shown to provide interference patterns in the same way that light waves do.

The Austrian physicist Erwin Schrödinger then considered the implications of this view of electrons for the atom. Using the phenomenon of what are called standing waves (i.e. waves that cannot travel freely like waves across the ocean but which are confined to a fixed space, such as the length of a violin string), Schrödinger developed equations which suggested that the fixed orbits of the Bohr atom were in fact the peaks of standing waves over which the charge of the electron was distributed.

Schrödinger believed that the electric charge of the electron really was diffused over the whole wave form, but Max Born suggested in 1926 that the wave merely represented the statistical probability of finding the electron in any given place.

Box 24.1 Uncertainty in Weimar Germany

There is a 'Forman Thesis' (see recommended readings below) which states that the rejection of traditional concepts of cause and effect and the concomitant development of the uncertainty principle and the emphasis on probabilistic physics was the result of a more generalized scepticism, and even despair, which profoundly affected German intellectuals after Germany's defeat in the First World War. The period of the Weimar Republic, 1919–33, saw Germans struggling with imposed reparations, runaway inflation, and a general cultural malaise. A book which argued that history was cyclical and that the West was entering a phase of decline, *The Decline of the West* (1918) by Oswald Spengler (1880–1936), a dilettante scholar based in Munich, became a best-seller because it suggested to Germans that the crisis was beyond their control and part of a much wider movement. Meanwhile the arts began to flourish, though often reflecting bitterness, cynicism, and despair. It was in this atmosphere, where old certainties no longer seemed to hold good, that Heisenberg, Born, Pascual Jordan (1902–80), and others, began to develop quantum probabilism and uncertainty.

It is hard to deny the likely effect of this atmosphere on German physicists, but it is also impossible to be sure of its real significance for the development of quantum theory overall. It is important to note, of course, that quantum mechanics was not developed solely by German thinkers. Moreover, it was enthusiastically taken up by non-German thinkers who presumably did not feel the Weimar angst. Bohr's Denmark remained neutral throughout the First World War, for example, and Paul Dirac's England had been victorious.

The fact is, as a result of the increasing professionalization of science, and its concomitant fragmentation into elite specialisms, the major factors affecting the thought and practice of innovators in science were developments within the sciences themselves, either of a technical kind (keeping up with which was increasingly demanding), or of an organizational or institutional kind. The politics of seeking jobs, securing funding, and ensuring one's contributions are widely recognized, had become, and remain, much more relevant to what scientists do than wider cultural concerns. Heisenberg and Born were not so much working in the atmosphere of the Weimar Republic as in that of the international Republic of Scientific Letters.

FURTHER READING

Joseph Ben-David, *The Scientist's Role in Society: A Comparative Study* (Chicago: University of Chicago Press, 1984).

Paul Forman, 'Weimar Culture, Causality and Quantum Theory, 1918–1927: Adaptation by German Physicists and Mathematicians to a Hostile Intellectual Environment', *Historical Studies in the Physical Sciences*, 3 (1971), pp. 1–115.

John Hendry, 'Weimar Culture and Quantum Causality', *History of Science*, 18 (1980), pp. 115–80.

Steven Shapin, *The Scientific Life: A Moral History of a Late Modern Vocation* (Chicago: University of Chicago Press, 2008).

The peak of the wave, supposedly corresponding to the position of the Bohr orbit, simply suggested that the probability of finding the electron there was greater than elsewhere. Born's probabilistic interpretation suggested once again, as in Heisenberg's work, that the electron could not be precisely located.

The result of this work was that so-called wave–particle duality was here to stay. Although De Broglie and Schrödinger both hoped they could provide a picture of a particle of matter as a tiny standing wave capable of moving through space, and so sometimes appearing like a wave and at other times appearing like a particle, the Heisenberg principle told against them.

Insofar as we can picture what is going on in the case of a single electron, we have to see it as an aggregate of probabilities. An experiment intended to pin the electron down causes what is called the 'collapse of the wave function', leaving one wave packet which defines that electron. For Heisenberg and Bohr, this effectively meant that the electron was forced into something like particulate existence by this very experimental process, by the act of observing it. If there was no such observation, all that would exist would be the array of probabilities.

Schrödinger famously objected to this in what is often referred to as a paradox, but is really an argument-form known as a *reductio ad absurdum* (if a set of starting assumptions leads to an absurd consequence, then there must have been something wrong with the starting assumptions). Schrödinger asked his readers to imagine a cat trapped in a box with a vial of poison gas. This vial would be automatically cracked open if a certain sub-atomic event takes place (for example, the emission of a sub-atomic particle from a radio-active substance), and would certainly kill the cat if it did open. If we suppose that the probability of the sub-atomic event taking place is 50/50, then the implication of the Heisenberg approach, according to Schrödinger, would be that the cat was both alive and dead all the time that it was in the box, and would only become definitely alive, or definitely dead, when we open the box and take a look in.

This is, of course, absurd when applied to a creature like a cat. A cat must be alive or dead, it cannot be both at once. But this does nothing to invalidate the uncertainty principle with regard to sub-atomic particles. After all, Bohr and others had already insisted that there was no continuity between the quantum world, in which Planck's constant plays a significant part, and our everyday world in which the influence of Planck's constant (an extremely miniscule number) is negligible.

One way of dismissing Schrödinger's *reductio ad absurdum* is to suggest that the vial and the cat, rather than being vicariously involved as part of the sub-atomic realm, constitute an experimental set-up to make an automatic observation of the sub-atomic event in question and record the result. We don't know the result until we open the box and look to see what has been recorded (a dead cat for a positive result, or a live cat for a negative one), but the cat has never been both alive and dead.

Bohr dealt with wave–particle duality simply by declaring these two aspects of sub-atomic reality to be complementary to one another. At the macroscopic level of cats and cabbages, waves and particles are distinct and incompatible with one another, but in the quantum world they are complementary aspects of being. We should not ask, therefore, whether an electron, or any other sub-atomic entity, is a wave or a particle. We should rather ask when and under what circumstances does it act like a wave, and when like a particle.

Bohr's principle of complementarity, announced in 1927, the same year as Heisenberg's uncertainty principle, became a mainstay of what became known as the Copenhagen Interpretation of the quantum theory which was generally accepted by physicists, and became the starting point for subsequent developments.

Einstein, however, could never bring himself to accept the uncertainty principle and the necessarily probabilistic nature of quantum physics. For Einstein it seemed to leave more questions than it answered. As he famously wrote to Max Born at the end of 1926:

> The theory says a lot, but does not really bring us any closer to the secret of the 'old one'. I, at any rate, am convinced that *He* does not throw dice.

Underlying Einstein's point was his conviction that fundamentally the world must be causal and deterministic. If quantum mechanics could only arrive at probabilities, and forever remain indeterminate, this merely revealed its incompleteness – these shortcomings of quantum mechanics were the result of our ignorance, not of the nature of the world (see also Box 24.2).

Together with Boris Podolsky (1896–1966) and Nathan Rosen (1909–95), Einstein wrote an article asking, 'Can Quantum Mechanical Description of Physical Reality Be Considered Complete?' (1935). In this the authors pointed out that, in principle, Heisenberg's indeterminacy principle could be worked around. When two particles are created together in a high-energy physics laboratory, conservation principles allow us to infer the momentum of one from the momentum of the other. In principle, therefore, we could directly measure the position of one of the particles and infer the momentum of that same particle by measuring the momentum of the other. Consequently, we would know both the position and the momentum of a particular particle.

Bohr responded to this by insisting that 'the extent to which an unambiguous meaning can be attributed to such an expression as 'physical reality'... must be founded on a direct appeal to experiments and measurements'. In other words, if we cannot establish both the position and the momentum of a particle by experiment and measurement, then we cannot claim these things (position and momentum) as undeniable aspects of physical reality. Bohr was putting his foot down. But not everyone found this satisfactory. One alternative position, inspired by the Einstein-Podolsky-Rosen paper, was that underlying quantum mechanics there must be a (so far) unobservable world of hidden variables.

One of the most interesting speculations along these lines was provided by David Bohm (1917–94) who claimed that his causal and deterministic theory could reproduce all the results of quantum mechanics. Unfortunately, his theory

Box 24.2 Something rotten in the state of Denmark?

Since the forging of the Copenhagen Interpretation, quantum theory has gone on from strength to strength and is widely recognized, certainly among physicists, as the most powerful research tradition in physics. The agreement between experiment and mathematical prediction of results has been recognized as far closer than in any other area of physics.

Relativity theory has also been almost entirely successful in matching mathematical prediction to experimental observations.

Furthermore, relativity theory and quantum theory have come together in highly fruitful ways. We've already noted, for example, that Louis de Broglie extended wave–particle duality from light to electrons (and this was subsequently extended to all sub-atomic particles, as they were successively discovered) by noting Einstein's equivalence between matter (particles) and energy (waves). Similarly, the highly successful branch of quantum theory known as quantum electrodynamics, also grew out of a partnership between quantum theory and relativity theory, brokered at first by Paul Dirac and subsequently by Richard Feynman and others.

In spite of this highly fruitful partnership, however, quantum theory and relativity theory don't really get along. And this is a major cause of concern for modern physicists. It is as though only one of them can be right, and yet if that is true how can they both be so successful?

Try as they might to overlook it, physicists are convinced that either something about quantum theory is rotten, or relativity theory is somehow misconceived.

Most of the betting men among them seem to have their money on quantum theory as the one that will win out, but others are still hoping for something to come along to show how they can both be right.

Putting it in simple terms, the problem is that relativity theory considers space-time to be fundamental, while for quantum theory neither space nor time are in any way fundamental.

Banesh Hoffmann (1906–86), a theoretical physicist who collaborated with Einstein, put the quantum case rather nicely. A single atom cannot meaningfully be

➠

had to assume that action at a distance was a reality. If Newtonianism (which took actions at a distance for granted) still ruled his ideas might have been taken more seriously, but in fact Bohm's claims have never been properly assessed.

If the point of science is to get to the truth about the way the world is, then we might wonder why Bohm's ideas have not been thoroughly examined and assessed. If he was right, then the whole course of modern physics would be different. But, of course, scientists are not driven by some abstract notion, such as 'The Truth'; like the rest of us, they are driven by careerist ambitions, financial security, and at least making a name for oneself if not actually seeking worldwide acclaim. This being so, no ambitious physicist, emerging from a long period of training in which they tried to master the latest ideas dominating the field, would ever think of stepping back to examine Bohm's attempt to develop a physics-of-said to have fluidity. Fluidity is defined in terms of the behaviour of lots of atoms acting together. Perhaps, he wrote, 'the fundamental particles of the universe individually lack the quality of existing in space and time'. Space and time, in other words, are only defined in quantum theory in terms of the behaviour of lots of atoms, or sub-atomic particles, acting together. 'Individual particles', he suggested, 'have no existence in space and time'.

According to Brian Greene (1963–), a theoretical physicist writing four decades later, the uncertainty principle tells against the relativistic assumption that empty space has zero gravitational field. According to the uncertainty principle all we can say is that *on average* the field is zero, but that the actual value can be higher or lower due to so-called 'quantum fluctuations'. Given that gravity is, according to relativity theory, a curvature of space-time, what this means is that in quantum terms space will be continually increasing and decreasing in curvature on a finely localized scale, at the tiniest dimensions it will be undulating so rapidly as to result in what the physicist John Wheeler (1911–2008) called 'quantum foam'. When dealing with extremely small distances of space (and time), therefore, the uncertainty principle is in direct conflict with the smooth geometrical picture of space-time which is fundamental to relativity theory.

There seems to be a fundamental incompatibility between quantum theory and relativity theory and this, for many physicists, indicates that there is a hidden flaw in our understanding of the world.

There seem to be only two ways to proceed from here: either to abandon the long cherished assumption that the physical world in its entirety can be explained by a single consistent, unified, rational and harmonious system; or to search for a unified theory capable of combining, or replacing, quantum and relativity theories. But this is a book about the history of science, not its future.

FURTHER READING

Brian Greene, *The Elegant Universe* (New York: Random House, 1999).
Banesh Hoffmann, *The Strange Story of the Quantum* (New York: Dover, 1959).

what-might-have-been. After all, his theory offered no new results, it merely offered a way of returning to a more classical kind of physics – a return to more traditional views of reality. But thanks to the work of Bohr, Heisenberg, Dirac and others – the objections of Einstein notwithstanding – cutting-edge physicists, as a professional group, had already moved far beyond mundane realities, and there was no turning back.

FURTHER READING

Peter J. Bowler and Iwan Rhys Morus, *Making Modern Science: A Historical Survey* (Chicago: University of Chicago Press, 2005), chapter 11.

David Cassidy, *Uncertainty: The Life and Science of Werner Heisenberg* (New York: Freeman, 1992).

P. C. W. Davies and J. R. Brown (eds), *The Ghost in the Atom: A Discussion of the Mysteries of Quantum Physics* (Cambridge: Cambridge University Press, 1986).

Peter Dear, *The Intelligibility of Nature: How Science Makes Sense of the World* (Chicago: University of Chicago Press, 2006), chapter 6.

Banesh Hoffmann, *The Strange Story of the Quantum* (New York: Dover, 1959).

Christa Jungnickel and Russell McCormach, *The Intellectual Mastery of Nature*, Vol. 2 (Chicago: University of Chicago Press, 1986).

Helge Kragh, *Quantum Generations: A History of Physics in the Twentieth Century* (Princeton: Princeton University Press, 1999).

Roger G. Newton, *From Clockwork to Crapshoot: A History of Physics* (Cambridge, MA: Harvard University Press, 2007).

Abraham Pais, *Niels Bohr's Times in Physics, Philosophy and Polity* (Oxford: Clarendon Press, 1991).

Bruce R. Wheaton, *The Tiger and the Shark: Empirical Roots of Wave-Particle Dualism* (Cambridge: Cambridge University Press, 1983).

Afterword

I have brought the story of science up to about the 1930s. It was at this time that Darwinian evolution and what came to be called Mendelian genetics came together to form what is usually known as neo-Darwinism. This was also the period which saw the general acceptance of the new quantum theory and of the theory of general relativity. Of course, the development of science did not stop there – far from it. As is generally acknowledged, the *pace* of scientific advance has increased throughout the twentieth century. For most lay persons, looking at science from the outside, this rapid advancement of science is seen in terms of new technological advances, and certainly such developments have a greater impact on our daily lives than scientific advances in the past were wont to do on those who lived in earlier times. But it is easy to see that this is not just a matter of scientific advances suddenly becoming more useful than they used to be.

The current alliance between science and technology, or between scientists and engineers, which ensures the exploitation of the latest scientific ideas, wherever possible, for pragmatic purposes (the fulfilment, at last, of the Baconian dream), has as much to do with developments in capitalism as it does with developments within science. Such are the economic realities that science, which is judged to be pragmatically useful, earns more funding than science that is not. Even so called 'blue sky' research in physics, which is supposed to have no obvious practical implications, is funded by governments (even to the extent of funding something as colossally expensive as the Large Hadron Collider near Geneva) because of the by-now inescapable hope that perhaps the most seemingly recondite scientific research may lead to something useful (and worth far more than its weight in Higgs bosons). Who would have believed, after all, that relativity theory could lead to nuclear weapons, or to GPS navigation systems?

The fact is, we have now entered the age of what commentators on science refer to as 'Big Science'. Science is now big in more ways than one. That being so, I make my excuses and leave off my story at this point. Even when considering nineteenth-century science, with its increasing professionalization, inflating the numbers of its personnel throughout the world, and its increasing specialization, resulting in a simultaneous diversification of science and a narrowing of the specialist experts' concerns, my task of providing a succinct account has become increasingly difficult. Big Science has now become so sprawling that it would certainly require a book as big as this one to do it even the scant justice that I have given to the two millennia of science leading up to the 1930s. I hope, however, that I have said enough to provide an overview of the foundations on which current Big Science has been built, and from which it continues to expand.

* * *

In a lecture delivered in 1959, 'Planck's Quantum theory and the Philosophical Problems of Atomic Physics', Werner Heisenberg pointed to one of the most profound implications raised by quantum physics:

> Here we find a consequence of the fact that natural science is not concerned with Nature itself, but with Nature as man describes and understands it. This does not mean that an element of subjectivity is introduced into natural science – no one claims that the processes and phenomena that take place in the world are dependent on our observation – but attention is brought to the fact that natural science stands between man and Nature and that we cannot dispense with the aid of perceptual concepts or other concepts inherent in the nature of man.

As Heisenberg seemed to note, the point is a general one, not just a point about quantum theory. As should be perfectly clear from the foregoing history, natural philosophy or natural science has *always* stood between man and Nature. The scientific thought of any age does not reveal the truth of the physical world to its contemporaries; it merely reveals how man describes and understands the world at that moment in history.

We like to think that we are making progress towards the real, ultimate, truth about the world itself, but no matter how close our mathematical predictions become to observed empirical results, we are still dependent upon our interpretations of nature, rather than on the reality of nature itself.

The same is also true (only more so – as any mathematical physicist would be quick to point out) of the *history* of science. Paul Valéry (1871–1945), poet, thinker, and correspondent of De Broglie and Einstein, famously said that 'History is the science of things which never happen twice'. In presenting this history to you I have not been presenting you with historical reality but with my interpretation of the evidence which seems to reveal the past to us. But what I have written stands between us and the historical reality. I have not been concerned with the past itself but with history as I have tried to understand it and describe it.

Authors of books on history are obliged by convention and by 'other concepts inherent in the nature of man' to impose a narrative structure on the past. Our assumptions about the inescapable chain of causes and effects leads us to try to explain the way Darwinism grew out of ways of thought that preceded it, or were already current when it was conceived. But, it is important to note, that the narrative imposed here is only one of many possible narratives that would have helped us to make sense of the evidence which reveals the past to us.

It is important, therefore, not to finish reading this book with the idea that you now know the history of science. At best, you only know its history as I have presented it here. If you are serious about wanting to understand the history of science, you must look into it yourself, by pursuing your own reading, and (ideally) your own research. Perhaps then, with your help, historians too will be able to feel, the way scientists do about their science, that gradually we are approaching closer to real historical truth.

FURTHER READING

Derek J.de Solla Price, *Little Science, Big Science*, New York: Columbia University Press,1963).

Peter Galison and Bruce Hevly (eds), *Big Science: The Growth of Large-Scale Research* (Stanford: Stanford University Press, 1994).

Index

CPSIA information can be obtained at www.ICGtesting.com
Printed in the USA
LVOW10*1724081214

417825LV00011B/156/P

9 780230 019423

palgrave macmillan

'A wide-ranging history of scientific thought, with a judicious selection of carefully explained topics. It is clear throughout that the work is based on extensive and deep knowledge of the field. . . an excellent text.'
– **David Philip Miller**, *Professor of History and Philosophy of Science, University of New South Wales, Australia*

'This is an excellent, well-balanced overview. Especially notable for its emphasis on the significance of magic in the development of modern science, it is also shrewd and perceptive on the relations between science and religion, and on science's broader cultural role.'
– **Michael Hunter**, *Professor of History, Birkbeck, University of London, UK*

From Plato to Einstein, via Copernicus and Darwin, this fascinating text shows how science came to be such a dominating and important aspect of modern culture. Presented as a lively and highly readable narrative, it provides a rich historical survey of the major developments in scientific thought, from the Ancient Greeks to the twentieth century. Along the way, John Henry revises many false assumptions that have been made about the history of science and offers new interpretations. Grounded firmly upon the latest historical scholarship, the book explains how new scientific theories have emerged and analyses their impact on contemporary thinking.

A Short History of Scientific Thought places the history of science within its social, cultural and historical contexts, covering key topics such as:

* the relationship between science and religion
* the history of science and magic
* experimental and mechanical philosophy
* Newton's legacy
* theories of evolution.

John Henry is Professor of the History of Science and Director of the Science Studies Unit at the University of Edinburgh, UK. He has published widely in the history of science from the sixteenth to the nineteenth centuries, including *The Scientific Revolution and the Origins of Modern Science*, now in its third edition.

Cover image: Diagram of Copernican Cosmology: a model of the solar system according to Copernican theory. Reproduced by permission of the Science Photo Library, London, UK

ISBN 978-0-230-01942-3
90000
9 780230 019423

www.palgrave.com