

1

Understanding the physical universe

AIMS

- to show how matter can be described in terms of a series of *models* (mental pictures of the structures and workings of systems) of increasing scale, starting with only a few basic building blocks
- to describe how, despite the great complexity of the material world, interactions between its building blocks can be reduced to no more than four distinct interactions
- to describe how natural phenomena can be studied methodically through observation, measurement, analysis, hypothesis and testing (the *scientific method*)

1.1 The programme of physics

Humans have always been curious about the environment in which they found themselves and, in particular, have sought explanations for the way in which the world around them behaves. All civilisations have probably engaged in science in this sense but sadly not all have left records of their endeavours. It would seem, however, that sophisticated scientific activity was carried out in ancient Babylonian and Egyptian civilisations and, certainly, many oriental civilisations had expert astronomers – every appearance of Halley’s comet over a time span of 1000 years was recorded by Chinese astronomers. Science as we know it today developed from the Renaissance in Europe which in turn owed much to the rediscovery of the work of the great Greek philosopher/scientists such as Aristotle, Pythagoras and Archimedes, work that had been further developed in the Islamic world between the seventh and sixteenth centuries.

Common to all scientific activity is the general observation that, in most respects, the physical world behaves in a regular and predictable manner. All other things being equal, an archer knows that if he fires successive arrows with the same strength and in the same direction they follow the same path to their target. Similar rules seem to govern the trajectories of stones, spears, discs and other projectiles. Regularities are also evident in phenomena involving light, heat, sound, electricity and magnetism (a magnetic compass would not be much use if its orientation changed randomly!). The primary objective of physics is to discover whether or not basic ‘rules’ exist and, if they do, to identify as exactly as possible what these ‘rules’ are. As we shall see, it turns out that most of the everyday behaviour of the physical universe can be explained satisfactorily in terms of rather few simple ‘rules’. These basic ‘rules’ have come to be called *laws of nature*, examples of which include the Galilean/Newtonian laws of motion (Sections 3.2, 3.3, 6.1), Newton’s law of gravitation (Section 5.1) and the laws of electromagnetism associated with the names of Ampère (Section 16.5), Faraday (Section 17.1), Coulomb (Section 15.5) and Maxwell (Section 18.1). In addition to these basic laws there are also ‘laws’ of a somewhat less fundamental nature which are used to describe the general behaviour of specific systems. Examples of the latter include Hooke’s law for helical springs (Section 3.5), Boyle’s (or Mariotte’s) law for the mechanical behaviour of gases (Section 10.10) and Ohm’s law for the conductivity of metals (Section 14.4).

The objective in studying physics, therefore, is to investigate all aspects of the material world in an attempt to discover the fundamental laws of nature and hence to understand and explain the full range of phenomena observed in the physical universe. This programme must include a satisfactory explanation of the structure of matter in all its forms (e.g. solids, liquids, gases), which in turn requires an understanding of the interactions between the basic building blocks from which all matter is constituted. How these interactions are responsible for the mechanical, thermal, magnetic and electrical properties of matter must also be explained. Such explanations, once discovered, can be applied to develop descriptions of phenomena ranging from the subatomic to the cosmic and to develop practical applications for the benefit of, and use by, society.

In the next three sections we review the language and images currently used by physicists to describe the structure of matter and the fundamental interactions of nature.

1.2 The building blocks of matter

Fundamental particles

Our present view of the nature of matter is very different from that which prevailed even fifty years ago. All matter is currently viewed as comprising various combinations of two classes of elementary particles – the basic building blocks – called, respectively, **quarks** and **leptons**. We give below an introductory account of the terminology and models used in the quark/lepton description of matter. The quark/lepton model will be discussed in more detail in Section 22.12.

Quarks and leptons occur in three distinct **generations** but only those in the first generation are involved in ordinary stable everyday matter. The first generation comprises two quarks, the up quark (symbol u) and the down quark (d), and two leptons, the electron (e) and the electron neutrino (ν_e). Matter comprising particles of the second and third generations is invariably unstable and is normally only formed when particles collide at very high speeds, such as those prevailing at the beginning of the Universe or in experiments with particle accelerators.

Leptons can exist as free isolated particles. Quarks, on the other hand, do not exist in isolation and are only observed grouped together, usually in threes, to form the wide range of different **particles** which form ordinary matter or which are produced in high-speed collisions.

In this section we describe how quarks and leptons, the basic building blocks of matter, combine to form larger building blocks which, in turn, combine to form even larger building blocks etc., as summarised in Table 1.1. Let us consider each stage in more detail, starting with combinations of quarks.

Table 1.1. Building blocks of matter

Building block	Scale/m
Quarks	$<10^{-20}$
Particles	$\sim 10^{-15}$
Nuclei	$\sim 10^{-14}$
Atoms	$\sim 10^{-10}$
Molecules	10^{-10} to 10^{-8}
Bulk matter	$> 10^{-9}$

Nuclei

The simplest combinations of first generation quarks which are observed are three-quark combinations called **nucleons**. As illustrated in Figure 1.1 two different types of nucleon are observed, namely the **proton** (p), which comprises two u quarks and one d quark, and the **neutron** (n), which comprises one u quark and two d quarks. The electric charge of the proton is $+e$ (e is called the fundamental electric charge), while that of the neutron is zero. While a proton is stable, a free neutron is not and decays radioactively to form a proton and two leptons. Further three quark combinations, involving quarks from other generations, will be considered when we come to discuss subnuclear particles in Section 22.11.

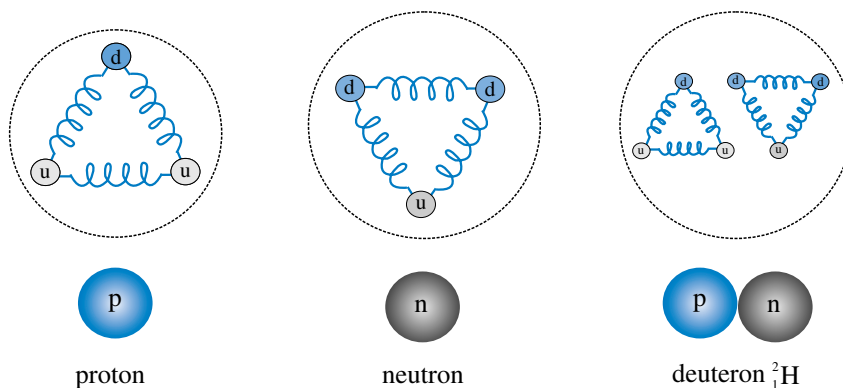


Figure 1.1. The quark and nucleon compositions of the proton (${}^1_1\text{p}$), neutron (${}^1_0\text{n}$) and deuteron (${}^2_1\text{H}$).

The next simplest combination, also illustrated in Figure 1.1, comprises six quarks ($uuuddd$), equivalent to one p and one n . This combination occurs in the **nucleus** of the deuterium atom (discussed below) and is called the deuteron. The electric charge of the deuteron,

like that of the proton, is $+e$. Two combinations of nine quarks, equivalent to pnn and ppn, are known; the first combination (pnn) is radioactive and the second (ppn) stable. When we consider atoms below we will identify these combinations as nuclei of tritium and helium atoms, respectively. Hundreds of stable particles (nuclei), comprising various combinations of u and d quarks (or, equivalently, protons and neutrons), are the basis of ordinary matter and will be discussed in Chapter 22. A great many other combinations can be created artificially, for example in nuclear reactors, and, while these are unstable, their lifetimes are often sufficiently long for them to be studied in detail and put to practical use (Chapter 22).

Atoms and molecules

All nuclei have an electric charge of $+Ze$, where Z is an integer; Z can be thought of as the number of protons in the nucleus. We will discover later (Chapter 15) that positive and negative charges are attracted to one another. Under normal conditions (by which is meant an environment which is not too hot and in which the matter density is not too low) the positively charged nuclei attract electrons to form electrically neutral systems called **atoms**. In atoms the electrons do not coalesce with the nuclei but, instead, behave as though they are moving around them in orbits with radii of the order of 10^{-10} m. This picture of an atom is something like that illustrated in Figure 1.2 – a very small nucleus of charge $+Ze$ surrounded by Z orbiting electrons, each of charge $-e$. The overall charge on the atom is thus zero – it is electrically neutral. The radius of an atom is 10 000 times greater than the radius of the nucleus (which is about 10^{-14} m). The electron is a very light particle, nearly 2000 times lighter than the proton, so nearly all the matter in an atom is concentrated in the nucleus. The nucleus and electrons are bound together in an atom by electrostatic attraction, a process which we will examine in detail when we study the structure of the atom in Chapter 20.

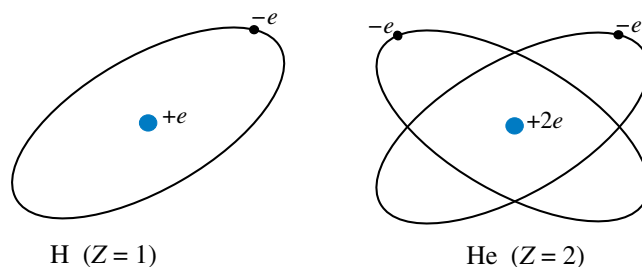


Figure 1.2. The electronic structure of the hydrogen and helium atoms.

As argued above, the electric neutrality of the atom requires that the nuclear charge $+Ze$ is balanced by the negative charge of Z electrons; Z therefore also gives the number of electrons in a neutral atom and is called the **atomic number**. The chemical properties of an atom are determined by the number of electrons it contains. An atom with $Z=1$, that is with a single proton in its nucleus and hence containing a single electron, is known as a hydrogen atom (Figure 1.2). The hydrogen nucleus can also contain one or two neutrons. Such atoms are called deuterium or tritium atoms, respectively, and are known as **isotopes** of hydrogen because they are chemically identical. Helium atoms have $Z=2$ (Figure 1.2); two different stable isotopes exist, ${}^3_2\text{He}$ (two p and one n) and ${}^4_2\text{He}$ (two p and two n). The chemical **elements**, listed in Appendix D (inside back cover), correspond to different values of Z ($Z=3$ for lithium, $Z=4$ for boron and so on). Note that the conventional notation used to specify an atomic nucleus (or **nuclide**) is ${}^A_Z\text{X}$ where X is the chemical symbol for the particular element, Z is the atomic number (the number of protons in the nucleus) and A (the number of nucleons – that is protons plus neutrons – in the nucleus) is called the **mass number**. Isotopes of an element therefore have the same Z but different values of A .

If an atom loses or gains an electron it will end up with a net positive or negative electric charge and is called an **ion**. The number of electrons lost or gained is conventionally denoted by a suffix to the notation for the atomic nucleus e.g. ${}^A_Z\text{X}^+$ (one electron lost), ${}^A_Z\text{X}^{2+}$ (two electrons lost) or ${}^A_Z\text{X}^-$ (one electron gained).

When atoms come sufficiently close together that their electron systems begin to overlap, they may form stable groupings of two or more atoms which are called **molecules**. Representations of some common molecules are illustrated in Figure 1.3. Molecular sizes vary from atomic dimensions ($\sim 10^{-10}$ m) to dimensions which are many hundreds of times larger in the case of biological molecules such as proteins and nucleic acids.

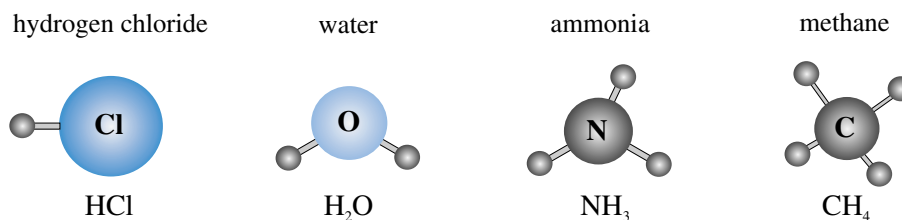


Figure 1.3. The atomic compositions of some common molecules – the smaller gray spheres represent hydrogen atoms.

The conventional notation for a molecule places the number of each type of atom in the molecule at the bottom right of the symbol for that atom. For example, a water molecule (a grouping of two atoms of hydrogen and one atom of oxygen) is denoted by the symbol H_2O (or ${}^1_1\text{H}_2 {}^{16}_8\text{O}$, if the isotopic species of each atom is also to be shown). We will consider the various processes by which atoms can bind together to form molecules in Section 21.1.

The description of matter which we have outlined in this section is summarised in Figure 1.4.

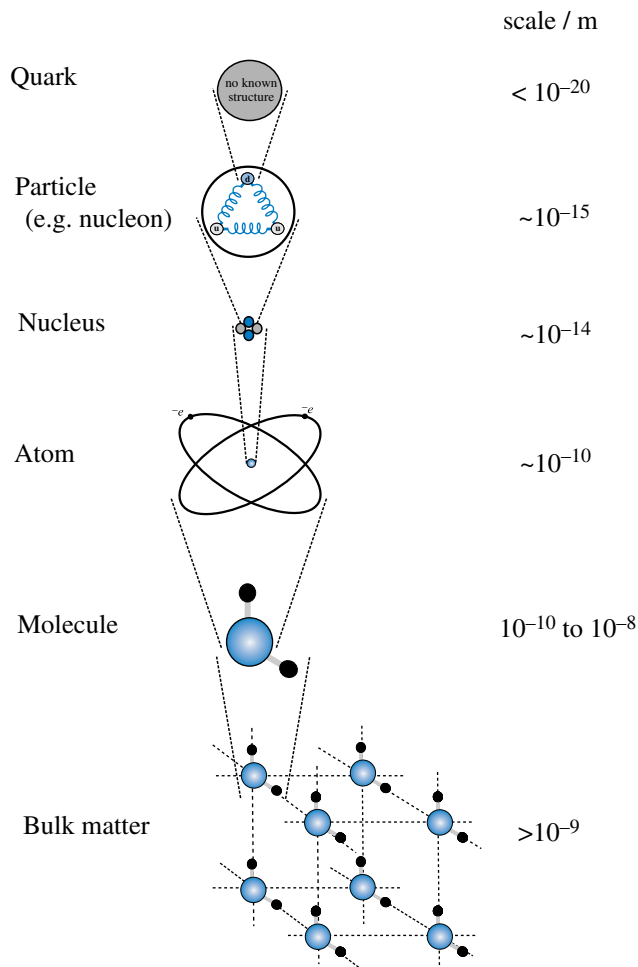


Figure 1.4. Models of the structure of matter – from the quark scale to the bulk matter scale.

1.3 Matter in bulk

When large numbers of atoms or molecules are bound closely together the atoms tend to arrange themselves in regular patterns, some examples of which are illustrated in Figure 1.5.

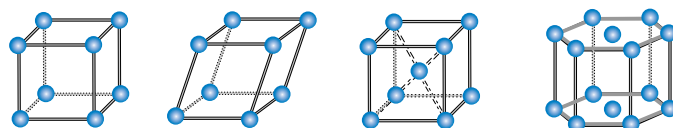


Figure 1.5. Some crystal lattice structures.

These patterns can extend over a very large number of atoms to form crystal lattices. Most **solids** are aggregates of crystals formed in this way and, if care is taken in their preparation, a solid may even be grown as one large single crystal.

Gases, on the other hand, comprise large numbers of molecules which are spaced so that the average distance between them is much greater than the molecular diameters. Molecules in gases move around rapidly and only interact with one another when they collide; otherwise they move in straight lines between collisions. The molecules in **liquids** are very close together but remain mobile and do not

form crystal lattices. Thus liquids fall somewhere between gases and solids. Many materials, glass for example, do not fall into these simple categories and have properties which are somewhere between those of solids and liquids.

Our everyday experience of solids, liquids and gases does not give any hint of their microscopic nature, that is, of their molecular, atomic or subatomic composition. Indeed, matter in bulk appears continuous – most materials seem to be uniform in their composition and properties at this level. Thus, if we are interested in answering questions such as ‘where is a stone going to land if I throw it from the top of a cliff?’ or ‘how much will the air in a balloon compress if I squeeze it?’, it hardly seems sensible to consider what happens to the atoms in the stone or to the quarks in the air! Questions like this are best addressed by employing **macroscopic models** (large-scale pictures) of the systems being investigated rather than the **microscopic models** which we have outlined in Section 1.2. Clearly a range of different models is available to us and the choice as to which one is best to use depends on the question being asked. The criterion which we must use here is that of *simplicity* – in attempting to explain any phenomenon only those concepts necessary for the explanation should be included in the theory. This principle, which is central to all scientific endeavour, is known as *Occam’s razor* after the medieval philosopher William of Occam (1285–1349), although the formulation in which it is normally stated (*entia non sunt multiplicanda praeter necessitatem* – entities are not to be multiplied unnecessarily) is attributed to John Ponce (1603–1661).

In this book we adhere to this principle as far as possible. We generally begin a discussion of a phenomenon from a macroscopic viewpoint. There will be many cases in which we are also able to discuss a phenomenon starting from a microscopic viewpoint (e.g. kinetic theory in Section 10.11). An important test of the microscopic approach will be whether its predictions agree with those of the macroscopic approach. We will find that when the two approaches agree we can be more confident that the microscopic approach is correct and, perhaps more importantly, we will gain some rewarding insights into the meaning of macroscopic concepts at a more basic level.

1.4 The fundamental interactions

We have seen that, despite the extraordinary complexity of the material world, all matter is made up from a relatively small number of basic building blocks. Equally remarkably we find that the way in which these building blocks interact with one another can be reduced to no more than four distinct interactions, namely

1. **The strong interaction:** This is the force between quarks which keeps them bound together within a particle or an atomic nucleus. It is responsible for the force between nucleons in a nucleus, as described in Chapter 22. The range over which the strong interaction operates is very small – it has negligible effect if the distance between particles is much greater than 10^{-15} m.
2. **The electromagnetic interaction:** This is the force which exists between all particles which have an electric charge, such as the force which keeps the electrons bound to the nucleus in an atom. The electromagnetic interaction is long range, extending in principle over infinite distances, but it is over 100 times weaker than the strong interaction within the range over which the strong interaction operates.
3. **The weak interaction:** Leptons are not affected by the strong interaction but interact with one another and with other particles via a much weaker force called the weak interaction, whose strength is only 10^{-14} times that of the strong interaction. While all particles interact weakly, the effect is noticeable only in the absence of the strong and electromagnetic interactions. The weak interaction is very short range ($\sim 10^{-18}$ m) and only plays a role at the nuclear and subnuclear level.
4. **The gravitational interaction:** By far the weakest of the fundamental interactions is the gravitational interaction, the interaction which, for example, gives a body weight at the surface of the Earth. Its strength is 10^{-38} times that of the strong interaction. All particles interact gravitationally and, like the electromagnetic interaction, the gravitational interaction operates over an infinite range.

There is a long tradition in physics of attempting to unify theories which were originally distinct. For example, for a long time magnetism and electricity were considered to be quite different phenomena but during the nineteenth century the two areas were united in a single theory of electromagnetism (Chapter 16). During the past fifty years the theories covering the fundamental interactions have been undergoing a similar unification process. In the 1960s, due principally to the work of Weinberg, Salam, Ward and Glashow, the electromagnetic and weak interactions came to be seen as different aspects of the same phenomenon (known as the **electroweak interaction**). Since then considerable progress has been made towards the unification of the electroweak interaction with the strong interaction (so-called **Grand Unification**) and this objective is still being pursued. Today, the goal of unifying gravity with the strong-electroweak interaction has become a ‘Holy Grail’ of physics but, to date, even the possibility of such a single theory of all four fundamental interactions, a *Theory of Everything*, remains in the realm of speculation.

1.5 Exploring the physical universe: the scientific method

Our aim in physics is to explore the physical universe, to observe, analyse and (hopefully) eventually understand the natural phenomena and processes which underlie the workings of the universe. In the process of achieving an understanding of natural phenomena we will often acquire an ability to predict their future course and hence an ability to apply our knowledge – to use it for practical purposes.

How then can we investigate natural phenomena? We outline below an approach known as the **scientific method**. It is a method which has proved its value over many centuries but it is important to note that there is nothing particularly remarkable about it – it has not been handed to us on ‘tablets of stone’. As we shall see, it is merely a series of practical steps that anyone who wishes to study a natural phenomenon methodically might well devise on his or her own initiative. We outline the steps below.

Observation

The first step is simply to observe the phenomenon – to watch it unfold. Careful systematic observation leads us inevitably to take notes on what we see – to **record** our observations. With records we can later remind ourselves, or others, of what we have observed. The process of recording what we see in a thorough and rigorous manner leads us quickly to make measurements. For example, if we are observing the motion of a moving object we could describe its motion in words by stating that ‘the object is first a long way from us, then not so far, then nearer and finally very near’. It is clear however that words alone soon become inadequate; they are not sufficiently precise and can be ambiguous. One person’s idea of ‘very near’ may not be the same as that of the next person. Measurement is therefore the next step in the scientific method.

Measurement

In making measurements we must decide which (physical) quantities associated with the phenomenon that we are observing can be measured most conveniently and accurately. Note that the process is already becoming a little arbitrary. One person’s idea of what can be measured conveniently may not be the same as that of the next person. As experience is built up, a consensus usually emerges on the best way to make a certain measurement. Sometimes, as we will see, technical developments can force a change in the consensus and hence even in the way in which physics is formulated. The development of physics has always been rooted strongly in empirical observation and hence in the process of measurement.

In making a measurement we inevitably have to choose a **unit** in which to make the measurement. In the case of the moving object we would naturally tend to measure its distance from us in metres because a unit of distance, the metre, has already been defined for us. Had it not been defined we would have had to invent some such unit. In choosing units for measurement it is also sensible to coordinate our choice with that of others, i.e. to choose agreed **measurement standards** and **systems of units**. This enables us to communicate our observations to colleagues on the other side of the world in such a way that they will know precisely what we mean.

Analysis and hypothesis

Having observed a phenomenon and having collected a set of measurements – our **experimental data** – the next step in the scientific method, in our attempt to understand the phenomenon, is to look for relationships between the quantities we have measured. For example in the case of a moving object we may have a set of measurements which gives the object’s position at certain times. In comparing the measurements of position with those of time can we see any pattern? Can we put forward any **hypothesis** (inspired guess) which describes and accounts for the relationship between the quantities? Can we go further and put forward a **model** of the situation, an idealised picture of what is happening, usually based on situations we already understand – i.e. on our experience?

At this stage the scientific method becomes arbitrary and personal. Different people from different backgrounds and with different experiences may see different patterns and may put forward different models. There is not necessarily any one correct interpretation. In time it may turn out that one approach is simpler and easier to follow than the others but it does not follow that this is the only correct approach. It is always wise to keep an open mind in studying natural phenomena – we are less likely to spot new patterns if we have already decided what we expect to see. We must always be on our guard against introducing prejudices when drawing on our experience.

A number of procedures may help us to identify patterns in our observations. As will be illustrated in Section 2.3 for the case of a moving object, we can assemble tables of data and can draw graphs of one measured quantity against another. We will see in Section 2.3 how analyses of tables and graphs often enable us to deduce relationships between observed quantities. Very general relationships are described as **laws** of physics. One of the things which makes physics such a rewarding subject to study is that not only are the fundamental laws few in number but they are also usually of relatively simple form. Because of the essential simplicity of the laws, the simplest and most natural way to express them is through the language of mathematics.

When we are successful in identifying relationships between observed quantities we are usually able to express them as mathematical equations, which, as we will see in Section 2.3, are usually the most concise and unambiguous way of expressing relationships.

The description of relationships between quantities as ‘laws’ of physics is perhaps unfortunate because these laws should not be regarded as incontrovertible edicts. They are merely well-established principles. Sometimes laws are found not to be as well established as was first believed. It is important therefore to **test** hypotheses and models regularly. This brings us to the final step in the scientific method.

Testing and prediction

It is now necessary to establish the range of applicability of any hypotheses and models which may have been proposed. We use these hypotheses and models, therefore, to *predict* results in situations in which measurements have not yet been made. We then make measurements in the new situations and see how well these measurements match predictions. Sometimes they do not match, although this

does not necessarily mean that our prior hypotheses and models were wrong. It means that they are limited in their applicability and that we have to extend the hypotheses and models to cover the new situations.

As we shall see, developments in physics in the twentieth century have shown that many apparently universal laws of classical physics do not apply at velocities which approach the speed of light or to particles on the microscopic (atomic and nuclear) scale. It has been necessary to develop new more comprehensive theories, namely the special theory of relativity (Chapter 9) and quantum mechanics (Chapter 13), to interpret and understand these situations.

As is apparent from the account of the scientific method given above, there is nothing particularly remarkable about the method. It has been described quite simply as 'organised common sense', a method which a person without a scientific background might well adopt when faced with the task of trying to understand a physical process. In physics we have the advantage of a wealth of techniques for observation and analysis that have been developed by the scientific community over a long period of time. This gives us a head start in seeking to understand new phenomena although we should always be aware of the possible limitations of established thinking.

In this book, therefore, we will not only describe the discoveries and the models which have been put forward by physicists, we will also, in the process, learn the skills and techniques which have been developed to analyse natural phenomena. We will then be able to apply these skills and techniques ourselves as we study the physical universe. The end product will be the ability to describe a whole range of apparently disconnected and complex phenomena in terms of an underlying simplicity of mathematically expressed structures. On many occasions we will see how advances in knowledge have led to new theories or models which replace a whole range of different models which were needed previously. This unifying process is one of the most satisfying aspects of physics. New understanding can actually simplify a situation, or a number of situations; we then feel instinctively that we are closer to the truth. The methods which we will uncover are powerful, intellectually satisfying and useful. We will not be able to reveal all the mysteries of the physical universe in this book but we will take some steps along the way and, perhaps more importantly, we will emerge equipped to explore further ourselves.

1.6 The role of physics: its scope and applications

In Sections 1.2 to 1.4 we saw how physics describes the basic components of matter and their mutual interactions. We also saw how physics endeavours to describe the physical world on all its scales – from that of the quark to that of the universe. In this sense, physics provides the basic conceptual and theoretical framework on which other natural sciences are founded and may therefore be regarded as the most fundamental and comprehensive of the natural sciences.

The techniques which have been developed to analyse the physical world can be used in almost any area of pure and applied research. Physics provides an excellent testing ground for the scientific method. Moreover, in seeking to unify understanding of the natural world, physics can play an important simplifying role in science, reducing complex situations to more understandable forms. In doing so, physics can also counteract the fragmentation into separate disciplines which tends to accompany the ever-expanding growth in scientific and technical knowledge.

Physics is at the basis of most present technology and is sure to be at the basis of much future technology, tackling problems as pressing and diverse as the development of new energy sources, of more powerful and less intrusive medical diagnostics and treatments and of more effective electronic devices. The growth of physics has spawned a multitude of technological advances which impact on almost all areas of science. Engineering practice must be revised regularly to take advantage of opportunities presented by the advance of physics.

In the previous section we noted that new, and more comprehensive, theories, namely the special theory of relativity and quantum mechanics, were developed in the last century to account for situations in which the laws of classical physics do not apply. The new theories have stimulated important new technologies, such as quantum engineering (the development of new microelectronic devices), laser technology and nuclear technology, technologies which could hardly have been dreamt of at the beginning of the twentieth century.

A sound knowledge of physics is needed by scientists and technologists if they are to be able to understand and adjust to the rapidly changing world in which they find themselves. Moreover this understanding should stimulate them to devise and initiate further changes themselves.

